

Search for the supersymmetric partner to the top quark in the all-hadronic final state with the ATLAS detector

Matthew T. Anthony

A thesis presented for the degree of
Doctor of Philosophy



Student ID no.: 1601241341
University of Sheffield
Department of Physics and Astronomy
Hicks Building, Hounsfield Rd
Sheffield, Yorkshire
United Kingdom
4th October 2020

Search for the supersymmetric partner to the top quark in the all-hadronic final state with the ATLAS detector

Matthew T. Anthony

Abstract

The Large Hadron Collider and the experiments which rely on it constitute the largest research project for particle physics, both in terms of physical hardware and computing resources, to allow us to probe the frontiers of our understanding of nature, commonly referred to as the Standard Model (SM) of particle physics. A well motivated extension to the existing SM is that of supersymmetry (SUSY), posited to resolve some of the pressing limitations of the SM. SUSY predicts that for each given fermion spin state, there is an associated boson partner and vice-versa. This thesis will cover the decay where the scalar partner to the top quark ($\tilde{t}_1$) is pair produced and decays to a top quark and a stable supersymmetric particle called a neutralino ($\tilde{\chi}^0$). This search is undertaken using the ATLAS detector between LHC run periods 2015-2018 (LHC Run 2) corresponding to a centre-of-mass energy $\sqrt{s} = 13$ TeV and a total integrated luminosity of 139.0fb^{-1} , in particular considering events with no electrons or muons in the final state. Since the supersymmetric particles must be reconstructed from their final-state decay products, efforts were undertaken to constrain the SM backgrounds which produce the same final states, primarily the backgrounds arising from the $t\bar{t} + Z(\rightarrow \nu\nu)$ and $Z(\rightarrow \nu\nu) + \text{jets}$ processes. This thesis in addition also covers the workflows required from collision event (simulated using Monte Carlo or detected) to a final result.

In the absence of any significant excesses, exclusion limits at the 95% confidence level have been applied on the visible cross section for models with a branching ratio $BR(\tilde{t}_1 \rightarrow t + \tilde{\chi}_1^0) = 100\%$ up to 1.3 TeV in the case where $\Delta(m_{\tilde{t}}, m_{\tilde{\chi}^0}) > m_{\text{top}}$. Exclusion limits at the 95% confidence level have been applied up to $(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (600, 450)$ GeV for signal models where $\Delta(m_{\tilde{t}}, m_{\tilde{\chi}^0}) \sim m_{\text{top}}$, similarly up to $(m_{\tilde{t}} \sim 660 \text{ GeV})$ for models where $\Delta(m_{\tilde{t}}, m_{\tilde{\chi}^0}) < m_{\text{top}/W}$ through the use of a new search strategy targeting four body decays ($\tilde{t}_1 \rightarrow bqq'\tilde{\chi}^0$).

Acknowledgements

I would like to credit the inspiring words of [1] for providing me the motivation to write this thesis.

I would also like to thank my family and friends for their invaluable moral support in completing this endeavour.

Quote

“So long and thanks for all the fish”

Douglas Adams, 9 November 1984.

Grants

This project was funded by the European Research Council (ERC) Horizon 2020 grant at the University of Sheffield (UK).

Contents

Declaration	vii
List of Figures	x
List of Tables	xiv
1 Introduction	1
2 The Standard Model	3
2.1 Gauge field theory	5
2.1.1 Construction of the Standard Model with Gauge interactions	5
2.2 Higgs model	6
2.2.1 Abelian spontaneous symmetry breaking	6
2.2.2 Non-abelian case	7
2.2.3 Glashow-Weinberg-Salam (GWS) theory of weak interactions	7
2.2.4 Giving masses to fermions	10
2.2.5 Higgs boson mass in the GWS model	11
2.3 Renormalisation	11
3 Looking beyond the standard model of particle physics	15
3.1 Limitations of the Standard Model	15
3.1.1 Hierarchy problem	16
3.2 Dark matter	17
3.3 Physics beyond the standard model: Supersymmetry (SUSY)	19
3.3.1 Poincaré group of rotations, boosts and translations	20
3.3.2 Coleman-Mandula No-Go theorem and associated exceptions	22
3.3.3 SUSY model from first principles: Wess-Zumino Model	22
3.3.3.1 Interactions	25
3.3.3.2 SUSY invariant gauge interactions	25
3.4 Minimal Supersymmetric Standard Model	27
3.4.1 Mass hierarchy	28
3.4.2 R-Parity Violation	32
3.4.3 Simplified Models	32
3.5 The story so far in third generation searches	35

4	ATLAS detector and infrastructure as part of the LHC	39
4.1	LHC overview	39
4.2	The ATLAS detector	43
4.3	Hardware	44
4.3.1	The ATLAS tracking system	44
4.3.2	Magnetic fields	45
4.3.3	ATLAS Calorimetry	46
4.3.3.1	Electromagnetic Calorimeter (ECAL) for ATLAS	46
4.3.3.2	Hadronic Calorimeter (HCAL)	47
4.3.4	Muon Spectrometer	50
4.4	Trigger	51
4.4.1	Analysis level implementation of a trigger	52
4.4.2	Data Acquisition and Online Data Quality Monitoring	55
4.5	Simulation	56
4.5.1	Monte-Carlo matrix element generation and Parton Showering	59
4.5.2	ATLAS detector simulation & digitisation	59
4.5.3	CPU and wall-time enhancements: FastSimulation	60
4.5.4	Pileup simulation: FastChain Pileup	61
4.5.4.1	Author contribution to Fast Chain transform	62
4.6	Reconstruction	65
4.6.1	Jet reconstruction	65
4.6.1.1	Calorimeter Jets	66
4.6.1.2	Track jets	66
4.6.2	Heavy flavour jet (B-jet) reconstruction in calorimeter jets	67
4.6.3	Trackjet b-tagging	68
4.6.3.1	Large radius jets and jet substructure	69
4.6.4	Leptons (electrons, muons)	71
4.6.4.1	Electrons	71
4.6.4.2	Photon reconstruction	73
4.6.4.3	Muons	73
4.6.5	Jet vertex tagging (JVT)	75
4.6.6	E_T^{miss}	75
4.7	Distributed computing and analysis using the WLCG	77
5	Statistics	78
5.1	Probability, statistics and fitting	78
5.1.0.1	Neyman confidence belt construction	79
5.1.1	Profile Likelihood fitting	80
5.1.2	Hypothesis testing	82
5.1.3	Limit setting	84
5.1.3.1	CL_S exclusion limits	85
5.2	Statistical use case: ATLAS specific analysis design in context of SUSY searches	86
5.2.1	Analysis configuration	86
5.2.2	Likelihood and fits used	87
5.2.3	Systematic handling in HistFitter	88
5.2.4	Multi-bin fits	89
5.2.5	Fit types	89

5.3	Multi-variate analysis and machine learning	90
6	Moving beyond the 2015-16 analysis.	94
6.1	Analysis configuration	95
6.1.1	Key backgrounds	95
6.1.2	Key variables and discriminants	95
6.2	Signal Regions (2015-16)	97
6.2.1	Preselection	97
6.2.2	Signal Region SRA and SRB (2015-16)	98
6.2.3	Signal Region SRC (2015-16)	99
6.2.4	Other signal regions (SRD, SRE)	101
6.3	Control regions	101
6.3.1	$t\bar{t}$ control regions (2015-16)	101
6.3.2	W+jets and single top control regions.	101
6.3.3	$t\bar{t} + Z$ control region using $t\bar{t} + \gamma$	102
6.3.4	Z+jets control regions	103
6.3.5	Multi-jet estimation	104
6.4	Validation Regions	105
6.4.1	$Z + \text{jets}$ validation regions	106
6.4.2	$t\bar{t}$ and $W + \text{jets}$ validation regions	106
6.5	Results	107
7	Re-optimisation of the 2015-16 signal regions	111
7.1	Jet substructure and boosted object identification	111
7.2	Boosted Decision tree application to signal discrimination	114
7.3	Practical limitations of the boosted top tagging and jet substructure observables.	118
7.4	Tau veto studies	118
7.5	Hybrid B-tagging	125
8	Progress toward 139fb^{-1} result.	130
8.1	New variable studies using receiver operating characteristic (ROC) scans	130
8.2	Control region specific optimisations: $\text{CR}t\bar{t} + Z\text{-AB}$ ($t\bar{t} + Z$) design	136
8.2.1	Formulation 1: Di-leptonic $\text{CR}t\bar{t} + Z$	136
8.2.2	Formulation 2: Tripletonic control region for $t\bar{t} + Z$	138
8.2.3	Optimising the tri-lepton control region: final definition of $\text{CR}t\bar{t} + Z$	142
8.2.4	Theoretical uncertainties in cross sections for other backgrounds in $\text{CR}t\bar{t} + Z$	144
8.3	Control region specific: CRZ-AB ($Z + \text{jets}$) design	149
8.3.1	CRZ definition and construction	149
8.3.2	Final definition of CRZ-AB	152
9	2015-18 analysis: Summary	154
9.1	Object definitions	154
9.2	Signal regions	156
9.2.1	Signal Regions SRA and SRB	156
9.3	Signal region C (SRC)	158
9.4	Four body signal regions: SRD	159
9.5	Control regions	161

9.5.1	Control regions for $t\bar{t}$, W + jets and single top for signal regions SRA & SRB	161
9.5.2	Control region for the compressed signal region: CRTC	162
9.5.3	Multi-jet background estimation	163
9.5.4	Four body control regions	164
9.6	Validation regions	166
9.7	Systematics	169
9.7.1	$t\bar{t} + Z$ theory uncertainties: scale variations, alternate generator and parton shower uncertainties	171
10	Fit Results	179
10.1	Fit configuration	179
10.2	Fit results	181
10.3	Discussion	194
10.3.1	Context of results in terms of other ATLAS measurements	195
10.3.2	Reflection: results in light of CMS searches	196
11	Concluding remarks	199
11.1	Final summary	199
11.2	Future work	199
11.2.1	Analysis extensions: RECAST framework	200
Appendix A	ATLAS specific configurations	202
A.1	Simulated samples	202
Appendix B	Group theory	203
B.0.1	Groups and Symmetries: QED and QCD	204
B.0.1.1	Covariant derivative D_μ and propagators	205
B.0.1.2	Consequences of gauge invariance	207
Appendix C	The renormalisation group	208
C.1	Renormalisation group flow	208
Appendix D	Systematic tables for full Run 2 stop0L analysis	210
Bibliography		212

Declaration

Statement

I, the author, confirm that the Thesis is my own work. I am aware of the Universitys Guidance on the Use of Unfair Means (www.sheffield.ac.uk/ssid/unfair-means). This work has not been previously been presented for an award at this, or any other, university.

Some of the work presented in this thesis has been published in journals and conference proceedings - the relevant publication is listed below.

Publications

[2]: Search for a scalar partner of the top quark in the all-hadronic $t\bar{t}$ plus missing transverse momentum final state at $\sqrt{s} = 13$ TeV with the ATLAS detector, *ATLAS Collaboration*, Aug 2020, European Physical Journal C (Article No.: 737 (2020)), DOI link: <https://doi.org/10.1140/epjc/s10052-020-8102-8>.

Author contributions

The contributions listed here will be expanded upon throughout this thesis, but the author's principal work will be summarised here. Most of it constitutes work supporting [2], as well as material that was not used in the final analysis.

- FastChainPileup simulation: software development including implementation of WLCG based nightly testing/validation framework. This includes development on ATLAS/Athena software distribution.
- Undertook R&D studies into the usage of jet substructure as a potential improvement over the signal regions in [3].
- Evaluated the performance of the veto on τ -leptons as an improvement over the signal regions in [3]. This also includes pre-reconstruction analysis.
- Evaluated the use of Object based E_T^{miss} sig. and other variables in the design of the signal regions used in [2].
- Used WLCG distributed computing and HTCondor batch systems to process the simulated samples and data into a final dataset, for signal regions A,B,C,D and the control regions.
- Contribution to software development, testing and implementation using the ShefAnalysis framework [link](#) (CERN auth required) to process samples/data.
- Designed and optimised a control region for the $t\bar{t} + Z$ background in [2].
- Contributed to the design of the Z+jets control region CRZ-AB, and to the Z+jets validation regions with the studies listed in this thesis.
- Contributed to the implementation of the background-only and exclusion fits, both in back-end development (sample-processing) and front-end (fit evaluation, systematic configuration).
- Implemented a script on HTCondor batch (CondorDriver - [link](#)) to handle HistFitter backend sample processing for substantial workflow speed-up.

- Calculated theoretical systematics due to $t\bar{t} + Z$ renormalisation and factorisation scale uncertainties, and calculated uncertainty on cross section for $t\bar{t} + Z$. Including the implementation in the fit.
- Produced the post-fit summary plots listed in [2] from the background-only fit results of the fits corresponding to SRA, SRB, SRC and SRD.

List of Figures

2.1	Loop diagrams in renormalisation problems	12
3.1	1-loop renormalisation correction to the Higgs mass from a top quark loop. . . .	16
3.2	Scalar one-loop corrections to Higgs mass term	17
3.3	Rotational velocities for seven galaxies as a function of distance from their galactic centres	18
3.4	Rotational velocity curves of galaxies NGC 7541 and NGC 801	18
3.5	Bullet Cluster	19
3.6	Feynman diagrams corresponding to different interaction vertices of the SUSY Yang Mills theory	26
3.7	An example R-parity conserving interaction vertex between SUSY scalars (sfermions) and sfermions. Note this is in addition to all of the usual interaction diagrams in the SM.	28
3.8	Differential cross section observed as a function of the $t\bar{t}$ p_T by ATLAS.	29
3.9	Schematic SUSY mass spectra	32
3.10	Representative diagram of the two body stop decay mode.	34
3.11	Representative diagram of the three body stop decay mode.	34
3.12	Representative diagram of the four body stop decay mode.	35
3.13	CDF exclusion limit [26] in the $m_{\tilde{t}}, m_{\tilde{\chi}^0}$ plane, for simplified models with $m_{\tilde{\chi}^0} = 40$ GeV.	36
3.14	ATLAS run 1 exclusion limit in the $m_{\tilde{t}}, m_{\tilde{\chi}^0}$ plane, for simplified models with 100% branching ratio.	36
3.15	CMS run 1 exclusion limit in the $m_{\tilde{t}}, m_{\tilde{\chi}^0}$ plane, for simplified models with 100% branching ratio	37
3.16	$m_{\tilde{t}}, m_{\tilde{\chi}^0}$ exclusion limits from the ATLAS PMSSM scan	38
4.1	LHC with all of its sub-component accelerator complexes, and smaller experiments not covered in this thesis.	41
4.2	Detector operations plots from LHC run 2	42
4.3	ATLAS detector schematic	44
4.4	Inner Detector schematic	45
4.5	Magnetic field configuration for ATLAS	46
4.6	Schematic cut-away diagram of the ATLAS calorimeter system, presented in [30].	47

4.7	ECAL layers, including the readout electronics corresponding to a single trigger tower of cells, as per [30].	48
4.8	Tile calorimeter schematic	49
4.9	HEC schematic	49
4.10	FCAL schematic	50
4.11	Muon spectrometer layout	51
4.12	Configuration of the ATLAS trigger and data acquisition (DAQ) system.	52
4.13	Comparison of the E_T^{miss} distributions of events accepted by the trigger algorithms only using the calorimeter energy measurements for each of the different E_T^{miss} trigger algorithms.	54
4.14	Trigger efficiency curve for E_T^{miss}	54
4.15	DQM monitoring interface	55
4.16	ATLANTIS event display	56
4.17	Schematic summary of the ATLAS workflow from the underlying matrix element process to final end-user analysis.	58
4.18	ATLAS detector simulation pipeline	60
4.19	Visualisation of detector using GEANT4 and Fatras	61
4.20	Pileup and hard scatter event configuration for the fast simulation chain	62
4.21	Snapshot of the ART interface webpage	64
4.22	Schematic of a B-hadron decay	68
4.23	Boosted top schematic	69
5.1	Neyman belt construction	80
5.2	Feldman-Cousins limit construction	85
5.3	ATLAS CR/VR/SR schematic	87
5.4	Multi-bin fit schematic.	89
5.5	A schematic diagram of a single decision tree.	91
6.1	Schematic diagram of SRATT/SRBTT events	99
6.2	Schematic SRC diagram	100
6.3	$t\bar{t} + \gamma$ CR plot.	103
6.4	N-jets distribution in a QCD control region used to constrain the QCD scale factors.	105
6.5	2015-16 fit results: validation regions	107
6.6	2015-16 fit results: signal regions	108
6.7	Post-fit distributions of the signal regions in [3].	110
6.8	2015-16 exclusion limits	110
7.1	Comparison the Gradient and AdaBoost boosting algorithms for a Boosted Decision Tree (BDT) based on simple kinematics (configuration I from table 7.6) in the loose preselection region listed in 7.5.	116
7.2	Comparison using TMVA of the Gradient and AdaBoost boosting algorithms for a Boosted Decision Tree (BDT) based on simple kinematics and jet substructure observables (configuration II from table 7.6) in the loose preselection region listed in 7.5.	116
7.3	BDTG distributions	117
7.4	τ -veto efficiencies as a function of p_T for $t\bar{t}$.	122
7.6	$t\bar{t}$ composition study for SRAT0	123
7.7	$t\bar{t}$ composition study for SRBT0	123

7.8	$t\bar{t}$ composition study for SRC1	124
7.9	Tau veto alternate variable selection plots	125
7.10	Distributions of kinematic variables in the 0L preselection, comparing the influence of the hybrid b-tagger against the fixed MV2c10 selection.	127
7.11	Distributions of kinematic variables in the SRA-type signal region defined in table 7.13, comparing the influence of the hybrid b-tagger against the fixed MV2c10 selection.	127
7.12	Distributions of the generator level filtering of the Z + jets sample in the 0 lepton pre-selection. The $t\bar{t}$ sample is used as a reference.	128
8.1	Preselection distributions of variables used in the receiver operating curve scans	132
8.2	Receiver operating scans for SRA-type signal	133
8.3	ROC scan for SRB-type signal	134
8.4	ROC scans for SRC-type signal	135
8.5	Distributions for 8.5a $p_T^{\ell\ell}$ and 8.5b $m(\ell, \ell)$ in the region “CR-ttZ-2LOS-5j2b”.	137
8.6	Distributions for 8.6a $p_T^{\ell\ell}$ and 8.6b $m(\ell, \ell)$ in the region “CR-ttZ-2LOS-6j2b”.	137
8.7	Distributions for 8.7a $m(\ell, \ell)$ and 8.7b $p_T^{\ell\ell}$ in the region “CR-ttZ-orig”.	140
8.8	Distributions for 8.8a $m(\ell, \ell)$ and 8.8b $p_T^{\ell\ell}$ in the region “CR-ttZ-orig-pT4j”.	140
8.9	Distributions for 8.9a $m(\ell, \ell)$ and 8.9b $p_T^{\ell\ell}$ in the region “CR-ttZ-LEPMET”.	140
8.10	Distributions for 8.10a $m(\ell, \ell)$ and 8.10b $p_T^{\ell\ell}$ in the region “CR-ttZ-LEPMET250”.	141
8.11	Distributions for 8.11a $m(\ell, \ell)$ and 8.11b $p_T^{\ell\ell}$ in the region “CR-ttZ-test250”.	141
8.12	Distributions for 8.12a N_b and 8.12b N_{jets} in the region “CR-ttZ-LEPMET”.	141
8.13	Distributions for 8.13a N_b and 8.13b N_{jets} in the region “CR-ttZ-LEPMET250”.	142
8.14	$p_T^{\ell\ell}$ distribution of CR $t\bar{t}$ + Z -AB with the relaxed $p_T^{\ell\ell}$ selection	143
8.15	Uncertainty scan in $\mu_{t\bar{t}+Z}$ for differing bins in $p_T^{\ell\ell}$ starting from the base region “CR-ttZ-orig”, evaluating the impact of normalisation of the tZ background. The post-fit $\mu_{t\bar{t}+Z}$ and its uncertainty are presented.	145
8.16	Uncertainty scan in $\mu_{t\bar{t}+Z}$ for differing bins in $p_T^{\ell\ell}$ starting from the base region “CR-ttZ-orig”, evaluating the impact of normalisation of the tZ background. The uncertainty on the post-fit $\mu_{t\bar{t}+Z}$ is presented independent of the central value.	145
8.17	Uncertainty scan in $\mu_{t\bar{t}+Z}$ for differing bins in $p_T^{\ell\ell}$ starting from the base region “CR-ttZ-orig”, evaluating the impact of normalisation of the Fakes background. The post-fit $\mu_{t\bar{t}+Z}$ and its uncertainty are presented.	146
8.18	Uncertainty scan in $\mu_{t\bar{t}+Z}$ for differing bins in $p_T^{\ell\ell}$ starting from the base region “CR-ttZ-orig”, evaluating the impact of normalisation of the Fakes background. The uncertainty on the post-fit $\mu_{t\bar{t}+Z}$ is presented independent of the central value.	146
8.19	Uncertainty scan in $\mu_{t\bar{t}+Z}$ for differing bins in $p_T^{\ell\ell}$ starting from the base region “CR-ttZ-orig”, evaluating the impact of normalisation of the WZ background. The post-fit $\mu_{t\bar{t}+Z}$ and its uncertainty are presented.	147
8.20	Uncertainty scan in $\mu_{t\bar{t}+Z}$ for differing bins in $p_T^{\ell\ell}$ starting from the base region “CR-ttZ-orig”, evaluating the impact of normalisation of the WZ background. The uncertainty on the post-fit $\mu_{t\bar{t}+Z}$ is presented independent of the central value.	147
8.21	Uncertainty scan in $\mu_{t\bar{t}+Z}$ for differing bins in $p_T^{\ell\ell}$ starting from the base region “CR-ttZ-orig”, evaluating the impact of normalisation of the tWZ background. The post-fit $\mu_{t\bar{t}+Z}$ and its uncertainty are presented.	148

8.22	Uncertainty scan in $\mu_{t\bar{t}+Z}$ for differing bins in $p_T^{\ell\ell}$ starting from the base region “CR-ttZ-orig”, evaluating the impact of normalisation of the tWZ background. The uncertainty on the post-fit $\mu_{t\bar{t}+Z}$ is presented independent of the central value.	148
8.23	Zjets combinatoric schematic	150
9.1	Schematic of a four body signal event in SRD, noting the orientation of the ISR and the stop decay system.	160
9.2	Schema for the renormalisation and factorisation scale uncertainties for $t\bar{t} + Z$	171
9.3	$m_T^{b,\min'}$ (9.3a), $m_{T2}^{\chi_{2'}^2}$ (9.3b) $m(J1; R = 1.2)$ 9.3c and object based $E_T^{\text{miss}'}$ significance (9.3d) in the loose 2L preselection as defined in table 9.18.	175
10.1	Full Run2 control region summary plot	182
10.2	Post-fit distributions of the control regions CRTAB, CRSTAB, CRTC, CRTD1 and CRTD2	183
10.3	Post-fit distributions of the control regions CRWAB, CRWD0, CRWD1 and CRWD2	184
10.4	Post-fit distributions of the control regions CRZAB-T0, CRZAB-TT-TW and CRZD0	184
10.5	Post-fit plots of $CRt\bar{t} + Z$ -AB	185
10.6	Full Run2 validation region summary plot	185
10.7	Full Run2 signal region summary plot	186
10.8	Post fit signal regions SRA and SRB	188
10.9	Post-fit signal regions SRC and SRD	188
10.10	Fit parameter “pull” plot for the background-only fit for SRA and SRB	191
10.11	Correlation matrix for the fitted nuisance parameters in a background-only fit for signal regions A and B.	192
10.12	Full Run2 exclusion limits.	194
10.13	2020 ATLAS summary exclusion limit plot	195
10.14	CMS signal regions (pre-fit) covering the hadronic supersymmetry searches	197
10.15	CMS exclusion limit covering the hadronic stop decay search covered in this thesis, from [116]. In the CMS paper, these are denoted as T2tt models.	197

List of Tables

3.1	Multiplets associated to the Higgs and fermions	27
3.2	Gauge & gaugino super-multiplets	27
3.3	Sparticle content of the MSSM in addition to the standard model [6], assuming minimal mixing of the first and second generation quarks and leptons, in the context of gauge and mass eigenstates. This thesis particularly looks at the existence of the $\tilde{t}_1$ and $\tilde{\chi}_1^0$ states, which are linear combinations of the sfermion partners to the chiral states of the top quarks or the bosino partners to the B and neutral Higgsinos.	30
4.1	LHC design specification	42
4.2	ATLAS design specification	43
4.3	CPU timing of fast chain configurations	63
4.4	MV2c10 input summary	68
4.5	Egamma reconstruction observables summary	72
4.6	Muon isolation working points	75
6.1	Preselection criteria	97
6.2	2015-16 SRA and SRB definitions.	98
6.3	Selection criteria for the compressed region SRC from [3], in addition to the preselection.	100
6.4	$t\bar{t} + \gamma$ control region def	103
6.5	2015-16 $Z + \text{jets}$ control region definitions	104
6.6	2015-16 $Z + \text{jets}$ validation region	106
6.7	Fitted yields for signal region SRA and SRB (2015-16)	108
6.8	Fitted yields for 2015-16 signal region SRC	108
6.9	Discovery p-value and significances for each signal region given the background-only hypothesis	109
7.1	A reminder of the SRA-T0 region definition [3] including the pre-selection conditions.	112
7.2	Event yields in the 2015-16 Signal Region SRA-T0, with a luminosity of 36.1fb^{-1} . The significance is the expected binomial significance given a 30% systematic. Stated significances of "-1.00" occur where the number of signal events are less than 1.	113

7.3	Event yields for signal Region SRA-T0 (table 7.1) with the $R = 0.8$ jet selections removed, and the $R=1.2$ reclustered jet mass selections replaced with $R=1.0$ jet mass selections, also at 36.1fb^{-1} . Stated significances of "-1.00" occur where the number of signal events are less than 1.	113
7.4	Signal Region SRA-T0 with the $R = 0.8$ jet selections removed, and the $R = 1.2$ reclustered jet selections replaced with $R=1.0$ jet selections. In addition, the leading- p_T jet is tagged using the 80% signal efficiency working point. Stated significances of "-1.00" occur where the number of signal events is less than 1. . .	114
7.5	Definition of loose preselection region used for training	115
7.6	Table of kinematic and jet substructure variables considered in the BDT study, with the two configurations (configuration I and configuration II respectively). . .	115
7.7	BDT/BDTG hyperparameter sets for all configurations used in table 7.6.	117
7.8	Table of yields after applying the BDTG for the simple kinematics and the jet substructure configurations as listed in table 7.6, compared to the "cut and count" signal region SRATT, as defined in [3].	117
7.9	Comparison tables in 2015-16 region SRA-T0 for the tau veto using a veto on very loose taus, no veto or the default jet based veto.	119
7.10	Comparison tables in 2015-16 region SRB-TT for the tau veto using a veto on very loose taus, no veto or the default jet based veto.	120
7.11	Comparison tables in 2015-16 region SRC1 for the tau veto using a veto on very loose taus, no veto or the default jet based veto.	120
7.12	Preselection event yield before and after an application of a $M_T(\tau\text{-cand}, E_T^{\text{miss}})_{\min} > 80$ GeV selection	125
7.13	SRA-type region used for hybrid b-tagging studies.	126
7.14	Table outlining the yields for selections incorporating either hybrid b-tagged jets or b-jets tagged with the MV2c10 fixed selection for the 0L preselection or the SRA-type region.	128
8.1	Region definitions for a 2LOS control region for $t\bar{t} + Z$	137
8.2	Yields tables for the 2LOS $t\bar{t} + Z(\rightarrow \ell\ell)$ control regions listed in Table 8.1	138
8.3	Candidate control region definitions for the $t\bar{t} + Z$ background.	139
8.4	Statistical error only yields for the regions defined in table 8.3.	142
8.5	Final definition of CRttZ-AB	143
8.6	Pre-fit yields for the 3L region CRttZ-AB (Table 8.5).	143
8.7	CRZ with loosened sub-leading $R=1.2$ reclustered jet mass selections and $\Delta R(b_1, b_2)$ slicing schemes.	151
8.8	Yields in the signal regions, and impact on μ_Z of each of the slicing schemes listed in table 8.7 to the total uncertainty.	152
8.9	Final definition of Full Run2 CRZ-AB	153
9.1	Object definition summary	156
9.2	Signal Regions SRA and SRB definitions as per [2].	157
9.3	Selection criteria for SRC from [2]. The signal regions are separated into five categories based on ranges of R_{ISR}	159
9.4	Four body signal region (SRD) selections.	161
9.5	Table of definitions for the $t\bar{t}$, $W + \text{jets}$ and single top control regions from [2] .	162
9.6	Definition of the $t\bar{t}$ control region for the compressed signal regions, as listed in [2].	163
9.7	Control regions for $W + \text{jets}$ in the four body channels as per [2]	164

9.8	Control regions for $t\bar{t}$ in the four body channels, as per [2].	165
9.9	Control regions for $Z + \text{jets}$ in the four body channels.	165
9.10	Validation regions for the $Z + \text{jets}$ background as used in [2].	166
9.11	Full Run2 $t\bar{t}$ validation region definition (SRA/SRB)	167
9.12	Table of selections for validation region VRTC in [2].	167
9.13	Validation regions for the $t\bar{t}$ background corresponding to the four body signal regions SRD1 and SRD2, as used in [2].	168
9.14	Validation regions for the $Z + \text{jets}$ background corresponding to the four body signal regions SRD0, SRD1 and SRD2, as used in [2].	168
9.15	$2e/\mu$ emulated pre-selection.	173
9.16	$2e/\mu$ emulated version of SRATT (SRA-2L-Tight).	173
9.17	Yields for the different variation samples in $\text{CR}t\bar{t} + Z$ and the full 2L emulated version of SRA-TT, as defined in table 9.16.	174
9.18	Table of 2L Emulated SRATT selections.	174
9.19	Yields and transfer factors for the combination region in 9.18.	176
9.20	Direct calculation of the transfer factors from $\text{CR}t\bar{t} + Z$ to SRATT for the renormalisation and factorisation scale uncertainties.	176
9.21	Direct calculation of the transfer factors from $\text{CR}t\bar{t} + Z$ to SRAT0 for the renormalisation and factorisation scale uncertainties.	176
9.22	Direct calculation of the transfer factors from $\text{CR}t\bar{t} + Z$ to SRATW for the renormalisation and factorisation scale uncertainties.	176
9.23	Direct calculation of the transfer factors from $\text{CR}t\bar{t} + Z$ to SRBT0 for the renormalisation and factorisation scale uncertainties.	177
9.24	Direct calculation of the transfer factors from $\text{CR}t\bar{t} + Z$ to SRBTT for the renormalisation and factorisation scale uncertainties.	177
9.25	Direct calculation of the transfer factors from $\text{CR}t\bar{t} + Z$ to SRBTW for the renormalisation and factorisation scale uncertainties.	177
10.1	2015-18 fit configuration tables	181
10.2	Signal Region Yields (post-fit) for signal regions SRA-TT, SRA-TW, SRA-T0, SRB-TT, SRB-TW and SRB-T0. The highlighted backgrounds ($t\bar{t}$, $W + \text{jets}$, $Z + \text{jets}$, $t\bar{t} + Z$, single top) correspond to those normalised in control regions in the fit.	187
10.3	Signal Region Yields (post-fit) for signal regions SRC1, SRC2, SRC3, SRC4 and SRC5. The highlighted $t\bar{t}$ background is normalised in the fit for SRC.	187
10.4	Signal region yields for 2015-18 SRD	187
10.5	Discovery fit results for SRA, SRB, SRC and SRD	193
A.1	Simulated background and signal samples used in the 139.0fb^{-1} analysis in [2].	202
D.1	Systematic breakdown for signal regions SRA and SRB from [2]. Only contributions over 1% are included.	210
D.2	Systematic breakdown for signal regions SRC and SRD from [2]. Only contributions over 1% are included.	211

Introduction

Particle physics is the endeavour to understand the fundamental questions of the building blocks that form our universe at the most intrinsic level. Throughout the twentieth century, armed with the ground-breaking tools of quantum mechanics, atomic structure and later sub-atomic structure could be understood, modelled and accurately predicted, leading to what is now well known as the Standard Model (SM) of particle physics. The latest milestone in this quest for understanding was the discovery of the Higgs boson in 2012. This model has led to predictions of particles (observed as the Higgs Boson, $W^\pm$ bosons, Z^0 boson, top quark, bottom quark, charm quark, muon and τ leptons), through a variety of experiments at a variety of different energy scales. This allows us to understand the universe through the expression of two fundamental objects relating to their intrinsic quantum mechanical spin states: fermions and boson, and their respective composite particles (baryons, mesons, etc).

However, even the SM, an excellent baseline for our understanding of the universe as it is, is incomplete. Like any who wish to understand the universe, theorists and experimentalists pose the question: *is this the most fundamental way to describe reality?* Some aspects of the Standard Model remain inconsistent with the existence of dark matter and dark energy confirmed by astronomical observations, and the direct inconsistency of General Relativity with the Standard Model are large open questions which cannot be explained within the current framework. In addition, within the SM itself, there are inconsistencies such as the Hierarchy Problem, the existence of neutrino mass and the observation of neutrino flavour oscillation in neutrinos originating from our sun.

From a theoretical perspective, we discuss the attempts undertaken to resolve some of these problems, namely through the predictions of a Beyond Standard Model (BSM) principle known as *Supersymmetry* (SUSY), which predicts a bosonic partner for each fermion, and vice versa. This can resolve primarily the hierarchy problem and proposes a candidate particle state for dark matter, and SUSY often forms a principal component of popular advanced fundamental theories. This theory endeavour will be summarised in chapters two (2) and three (3.3), covering the key mechanisms of the standard model as we observe today, the Higgs mechanism and then looking in detail at supersymmetry, both from a fundamental perspective, and in terms of predictable results.

On the experimental side, substantial effort has been undertaken using the Large Hadron Collider (LHC) at *Centre Europeen de la Recherche Nucleaire* (English: *European Organisation for Nuclear Research*, CERN) in an attempt to either observe or exclude supersymmetric

phenomena. This endeavour commenced with the early data-taking runs with 7/8 TeV in the LHC run years 2010-2014, and was substantially improved with the vast dataset (corresponding to a total luminosity of 139 fb^{-1}) collected with a centre of mass energy of 13 TeV in LHC run years 2015-2018 (the majority of this dataset being collected in 2017-2018). The ATLAS (**A** Toroidal **L**H**C** **A**pparatu**S**) detector has been an integral part of this endeavour.

This thesis will outline in chapter four, the configuration and implementation of the LHC and ATLAS detector, the process for estimating Standard Model backgrounds through simulation, analysis workflows and a substantial contribution to the “future-proofing” of the experimental simulation. Since the most accurate description of any result should be considered in terms of the relevant statistics, chapter five will cover the main components that an ATLAS analysis will undertake in terms of likelihood fitting, construction of model-dependent and model-independent exclusion limits and computation of significance.

Chapters six through ten will outline a search for supersymmetric partners to the top quark in the final state without electrons/muons, which is based on the pair-produced decay of the supersymmetric partner to the top quark ($\tilde{t}_1$ to a top quark and a neutralino ($\tilde{\chi}_1^0$). Supersymmetric particles, with exception of the neutralino, are unstable and are assumed to have lifetimes that are too short to be detected directly. The neutralino itself interacts only weakly, and will also not be observed directly, merely inferred from a deficit of missing net momentum in the centre-of-mass frame of an event. Hence, the events containing the decay products of supersymmetric phenomena must be extracted from a set of Standard Model backgrounds which mimic the same identical final state, using event and object level selections. Dedicated modelling of these Standard Model backgrounds, with particular focus on the $t\bar{t} + Z$ and $Z + \text{jets}$ processes, is undertaken to better understand these processes and their prevalence within a region of interest which is sensitive to supersymmetric phenomena. With the standard model backgrounds well constrained, any statistically significant excess in data could be considered an indication of new phenomena. These chapters will start from the results obtained in 2015-16 (chapter 6), move to an analysis strategy targeting the full 139 fb^{-1} dataset (chapters 7, 8 9 and then discuss the results obtained (chapter 10).

The Standard Model

Introduction

For the purposes of brevity, this thesis will not derive the fundamentals of quantum field theory from first principles, and will assume the reader is familiar with the fundamental parts, such as the concepts of the Dirac and Klein Gordon equations, Lagrangian densities and the equations of motion therein, and the notion of the scattering matrix $\mathcal{S}$ with its relevant topics. In addition, relativistic electrodynamics will also be assumed prior knowledge, as well as knowledge of 4 vectors, momentum etc. This thesis will assume the convention of the $(+, -, -, -)$ signature of the Minkowski metric $g^{\mu\nu}$.

First, we must recall some facts from QFT, namely the concept of the action, Lagrangian and the equations of motion

QFT reminder

For scalar fields acting on Minkowski spacetime $\phi(x)$, we can define a Lagrangian as a function of those scalar fields and their spacetime derivatives. In classical mechanics, this can be written in terms of the canonical coordinates and their time derivatives as $L = K.E. - P.E.$, and analogously for QFT, the Lagrangian density $\mathcal{L}(\phi, \partial_\mu \phi)$, where the Lagrangian L is the space-integral $L = \int d^3\mathbf{x} \mathcal{L}(\phi, \partial_\mu \phi)$. The action is a functional defined between two reference events (a, b) , with $\mathcal{S} = \int_{t_b}^{t_a} dt L(\phi(x))$. The dynamics of the system is governed by the principle of least action[4], which is the set of fields ϕ such that the variational derivative:

$$\frac{\delta \mathcal{S}}{\delta \phi} = 0. \quad (2.1)$$

This also can be equivalently stated as the solutions of the *Euler-Lagrange* equation:

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right) - \frac{\partial \mathcal{L}}{\partial \phi} = 0. \quad (2.2)$$

The Lagrangian densities for free scalar fields (ϕ) and fermion spinors (ψ) with mass m can be specified as [4]:

$$\mathcal{L}_\phi = \frac{1}{2}(\partial^\mu \phi)(\partial_\mu \phi) - \frac{1}{2}m^2 \phi^2, \quad (2.3)$$

$$\mathcal{L}_\psi = i\bar{\psi}(\not{\partial} - m)\psi, \quad (2.4)$$

with the latter equation using the conventional Feynman slash notation $\not{\partial} = \gamma^\nu \partial_\nu$.

A reminder on symmetries

Since this thesis will discuss the concept of symmetries in quantum field theories (and later supersymmetries). These symmetries are transformations of the fields in the theory of interest, which leave the Lagrangian invariant (or the Lagrangian density invariant up to a surface term which would vanish under integration). For example, in scalar field theory (so a Klein Gordon Lagrangian with no mass), it can be shown that the action (spacetime integral of the Lagrange density) is invariant under spatial transformations (where ϵ^μ is some small quantity) [4].

$$x^\mu \rightarrow x'^\mu = x^\mu + \epsilon^\mu, \phi(x^\mu) \rightarrow \phi(x^\mu + \epsilon^\mu), \quad (2.5)$$

Taylor expanding the components and neglecting terms of $O(\epsilon^2)$, it can be found that

$$\phi(x'^\alpha) = \phi(x^\alpha) + \epsilon^\mu \partial_\mu \phi(x^\alpha) + O(\epsilon^2), \quad (2.6)$$

using this as an equivalent equation for the derivative term,

$$\partial \phi(x'^\alpha) = \partial_\mu \phi(x^\alpha) + \epsilon^\nu \partial_\nu \partial_\mu \phi(x^\alpha). \quad (2.7)$$

considering the difference between the Lagrangian under the spatial shift and the original, it can be found that:

$$\delta \mathcal{L} = \mathcal{L}(\phi(x'), \partial_\mu \phi(x')) - \mathcal{L}(\phi(x), \partial_\mu \phi(x)) = \epsilon^\nu \partial_\nu (\partial_\mu \phi(x)) \partial^\mu \phi(x), \quad (2.8)$$

and this remaining component vanishes under integration over 4-dimensional Minkowski spacetime (since both ϕ and $\partial_\mu \phi$ vanish according to their boundary conditions). Thus the action is invariant under spatial symmetries.

This last term can also be written in more general terms as:

$$\partial_\nu \left[\frac{\partial \mathcal{L}}{\partial (\partial_\nu \phi)} \partial_\mu \phi \right] \epsilon^\mu = \epsilon^\mu \partial_\mu \mathcal{L}. \quad (2.9)$$

Since the Lagrangian has been shown to be invariant of the spatial symmetries, if true for all values of ϵ^μ , this can be written as:

$$\partial^\nu T_{\nu\mu} = 0, \quad (2.10)$$

which is the energy momentum tensor

$$T_{\nu\mu} = \frac{\partial \mathcal{L}}{\partial (\partial_\nu \phi)} \partial_\mu \phi - g_{\mu\nu} \mathcal{L}. \quad (2.11)$$

This energy-momentum tensor contains the conserved energy density (T_{00}), and a momentum density $T_{0\mu}$ can be integrated over space to give:

$$P_\mu = \int d^3\mathbf{x} T_{0\mu}. \quad (2.12)$$

In general, this conserved current J^μ can be written for some infinitesimal symmetry transformation as (for a scalar Lagrangian invariant under a particular transformation $\phi \rightarrow \phi + \epsilon\delta\phi$):

$$\phi \rightarrow \phi + \epsilon\delta\phi, \mathcal{L} \rightarrow \mathcal{L}' = \mathcal{L} + \epsilon\partial_\mu V^\mu, \quad (2.13)$$

for some vanishing total divergence V^μ , and the conserved current can be written in general as:

$$J^\mu = \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right) \delta\phi - V^\mu. \quad (2.14)$$

2.1 Gauge field theory

The form of the Lagrangian density corresponding to EM fields alone in the presence of some source J^μ can be defined as:

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + J^\mu A_\mu, \quad (2.15)$$

which give rise to the familiar equations of motion, namely:

$$-\partial_\mu F^{\mu\nu} - J^\nu = 0, \quad (2.16)$$

which correspond to Maxwell's equations (albeit in a less readily recognisable form).

From electrodynamics, recall that the field A^μ can be defined by the scalar and vector potentials:

$$A^\mu = (\phi, \mathbf{A}), \quad (2.17)$$

and that Maxwell's equations are invariant under the gauge transformation of the field A^μ :

$$\begin{aligned} \phi &\rightarrow \phi - \frac{\partial \lambda}{\partial t}, \mathbf{A} \rightarrow \mathbf{A} + \nabla \lambda \\ A^\mu &\rightarrow A^\mu - \partial^\mu \lambda. \end{aligned} \quad (2.18)$$

This fact will reappear when we discuss U(1) symmetries, which produce exactly (up to a constant factor e) this exact transformation, hence the gauge transformation can be considered a consequence of the U(1) symmetry. For the purposes of brevity, this discussion will be directed to an appendix (B.0.1), which details a short summary of group theory and relevant components of Abelian/Non-Abelian Gauge Theories.

2.1.1 Construction of the Standard Model with Gauge interactions

Given the gauge symmetries of the standard model Lagrangian, namely the symmetry group combination $SU(3) \times SU(2)_L \times U(1)_Y$, the standard model of particle physics can be written in the following form:

$$\mathcal{L} = \mathcal{L}_{\text{kinetic}} + \mathcal{L}_{QCD} + \mathcal{L}_{SU(2)_L} + \mathcal{L}_{U(1)_Y} + \mathcal{L}_{\text{Other}}, \quad (2.19)$$

which combines all of the ingredients, namely the kinetic terms associated to the massive particles, and the dynamics of the respective gauge bosons. This last term will be the topic of the next section, since this is the term which induces spontaneous symmetry breaking (which will be covered in the next section).

2.2 Higgs model

This section will consider the concept of spontaneous symmetry breaking, which spontaneously breaks the $SU(3)_c \times SU(2)_L \times U(1)_Y$ symmetry into $SU(3)_C \times U(1)_{EM}$, giving rise to the well known massive charged and neutral currents ($W^\pm, Z^\pm$ as well as the photon from QED), with their respective masses (or massless status).

2.2.1 Abelian spontaneous symmetry breaking

We start (as per [4]) with a Lagrangian corresponding to a complex scalar field ϕ coupled to itself and an EM field:

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + (D_\mu\phi)^*(D_\mu\phi) - V(\phi), \quad (2.20)$$

with the functional form of this potential being:

$$V(\phi) = -\mu^2\phi^*\phi + \frac{\lambda}{2}(\phi^*\phi)^2. \quad (2.21)$$

A non-trivial vacuum expectation value (VEV) can occur in the case where $V(\phi) = 0, \phi \neq 0$, which occurs in the case that $\mu^2 > 0$. This VEV can be calculated as:

$$\langle 0 | \phi | 0 \rangle = \phi_0 = \left(\frac{\mu^2}{\lambda} \right)^{\frac{1}{2}}; \mu^2 > 0 \quad (2.22)$$

Now, the field and potential can be re-written about the vacuum state as (assuming the variation is small):

$$\phi(x) = \phi_0 + \frac{1}{\sqrt{2}}(\phi_1(x) + i\phi_2(x)), \quad V(\phi) = -\frac{1}{2\lambda}\mu^4 + \mu^2\phi_1^2 + O(\phi_i^3). \quad (2.23)$$

In the potential now there are only term dependent on ϕ_1 , thus the field ϕ_1 has a non-zero VEV (hence is massive with mass $\sqrt{2}\mu$), while the field ϕ_2 has a trivial VEV (hence is massless). This also gives rise to a massive photon term:

$$m^2 = 2e^2\phi_0^2. \quad (2.24)$$

In the unitary gauge (which we are allowed to do freely since all terms are locally U(1) invariant), we can require that the field ϕ is purely real, which then allows the Lagrangian to be re-written as:

$$\mathcal{L} = -\frac{1}{4}(F^{\mu\nu}F_{\mu\nu}) + (\partial_\mu\phi)(\partial^\mu\phi) + e^2\phi^2A_\mu A^\mu - V(\phi). \quad (2.25)$$

Running back through the above procedure instead results in a massive field A^μ , a massless physical field ϕ_1 and a coupling between them. This gives mass to the gauge boson, and is known as spontaneous symmetry breaking, and the general case is referred to as the *Higgs mechanism*.

2.2.2 Non-abelian case

The non-abelian case follows similarly to that of the abelian case, and the argument will be as follows [4]:

Let ϕ_i be a set of scalar fields defined in a Lagrangian under a symmetry group G , namely (:

$$\phi_i \rightarrow (\mathbb{I} + \alpha^a t^a)_{ij} \phi_j. \quad (2.26)$$

In order to make the ϕ take real values only, the matrices of the group t^a are Hermitian). Define

$$T_{ij}^a := -it_{ij}^a \quad (2.27)$$

which means the T^a are purely real and manifestly anti-symmetric. With this in hand, if we consider the group G to be a gauge symmetry, we can define the kinetic term as:

$$D_\mu \phi = (\partial_\mu - g A_\mu^a T^a) \phi, \quad (2.28)$$

$$\frac{1}{2} (D_\mu \phi_i)^* (D^\mu \phi_i) = \frac{1}{2} (\partial^\mu \phi_i) (\partial_\mu \phi_i) - g A_\mu^a (\partial_\mu \phi_i) + g^2 A_\mu^a A^{a\mu}. \quad (2.29)$$

Assuming the fields acquire some VEV $\langle 0 | \phi_i | 0 \rangle = (\phi_0)_i$, and expanding about these terms, a mass term can be found corresponding to:

$$\delta \mathcal{L} = \frac{1}{2} m_{ab}^2 A_\mu^a A^{b\mu}, \quad m_{ab}^2 = g^2 (T^a \phi_0)_i (T^b \phi_0)_i, \quad (2.30)$$

with m_{ab} being the “mass-matrix” of gauge bosons. Given the mass matrix, some entries will be 0 (which will correspond to massless degrees of freedom), which correspond to those generators that leave the vacuum invariant ($T^a \phi_0 = 0$). It can be shown (see [4] for a discussion of this) that if we consider $SU(2)$, all of the generators of $SU(2)$ do not leave the vacuum invariant and thus acquire mass.

2.2.3 Glashow-Weinberg-Salam (GWS) theory of weak interactions

Since in the last subsection, it was observed that the spontaneous symmetry breaking for both $U(1)$ and $SU(2)$ lead to massive vector bosons only. Since it is known that there is a massless vector boson (the photon) in the final theory (i.e. after spontaneous symmetry breaking). We follow the method of the GWS theory [4], namely we propose a symmetry $SU(2) \times U(1)$, which means a field transforms as:

$$\phi \rightarrow e^{i\alpha^a(x)t^a} e^{i\beta/2} \phi, \quad \text{or } \phi \rightarrow e^{i\alpha^a(x)t^a + i\frac{\beta}{2}\mathbb{I}}, \quad (2.31)$$

with the generator $t^a = \frac{\sigma^a}{2}$ (the Pauli matrices - see Appendix B). If the potential $V(\phi)$ is such that the VEV can be defined as (switching gauge to “unitary gauge” where appropriate):

$$\langle 0 | \phi | 0 \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \quad (2.32)$$

It can be found that a combination of gauge transformations leaves ϕ invariant, namely:

$$\alpha^1 = \alpha^2 = 0, \alpha^3 = \beta, \quad (2.33)$$

which leads to a specific case where there are three massive vector bosons (which will later correspond to $W^\pm$ and Z bosons as will be seen), and one single massless boson.

In order to obtain the mass terms for these bosons in our theory, we first define the covariant derivative under $SU(2) \times U(1)$ as:

$$D_\mu \phi = (\partial_\mu - igA_\mu^a t^a - \frac{i}{2}g'B_\mu)\phi, \quad (2.34)$$

for an $SU(2)$ associated vector boson A_μ^a and B_μ the $U(1)$ associated vector boson, and g, g' are coupling constants. Choosing only the terms in the mod-squared expansion of the above equation where the terms, we find:

$$\mathfrak{L}_{\text{Boson mass}} = \frac{1}{2} \begin{pmatrix} 0 & v \end{pmatrix} \left(gA_\mu^a t^a + \frac{1}{2}g'B_\mu \right) \left(gA^{b,\mu} t^b + \frac{1}{2}g'B^\mu \right) \begin{pmatrix} 0 \\ v \end{pmatrix}, \quad (2.35)$$

doing the calculations explicitly, we get:

$$\mathfrak{L}_{\text{Boson mass}} = \frac{1}{2} \begin{pmatrix} 0 & v \end{pmatrix} \left(gA^{b,\mu} A_\mu^a t^a t^b + \frac{1}{4}g'^2 B_\mu B^\mu + gg'A^{a,\mu} t^b B_\mu \right) \begin{pmatrix} 0 \\ v \end{pmatrix}. \quad (2.36)$$

We then relate the definition of the generators of $SU(2)$ with respect to the Pauli matrices σ^a : $t^a = \frac{\sigma^a}{2}$, and make use of a useful identity of the Pauli matrices: $\sigma^a \sigma^b = \delta^{ab} \mathbb{I}_2 + i\epsilon^{abc} \sigma^c$. The last term in the second identity is effectively irrelevant, as we are doubly contracting $t^a t^b$ with obviously symmetric terms $A^{a,\mu} A_\mu^b$. With a little algebra skipped for the purposes of brevity (including the explicit definition of the Pauli matrices used), it can be found that:

$$\mathfrak{L}_{\text{Boson mass}} = \frac{1}{2} \frac{v^2}{4} \left[g^2 A^{1,\mu} A_\mu^1 + g^2 A^{2,\mu} A_\mu^2 + (g'B^\mu - gA^{3,\mu})(g'B_\mu - gA_\mu^3) \right]. \quad (2.37)$$

From the GWS model, it is expected that three massive bosons are seen and a remaining massless boson, so we define the linear combinations of the A, B vector fields. This is effectively a linear transformation from the gauge eigenstates into the mass eigenstates ($W_\mu^\pm, Z_\mu^0, \Gamma_\mu$), and for the GWS model, the transformations are:

$$\begin{pmatrix} W_\mu^+ \\ W_\mu^- \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} \begin{pmatrix} A_\mu^1 \\ A_\mu^2 \end{pmatrix}, \quad m_W = \frac{gv}{2}, \quad (2.38)$$

$$\begin{pmatrix} Z_\mu^0 \\ \Gamma_\mu \end{pmatrix} = \frac{1}{\sqrt{g^2 + g'^2}} \begin{pmatrix} g & -g' \\ g' & g \end{pmatrix} \begin{pmatrix} A_\mu^3 \\ B_\mu \end{pmatrix}, \quad m_\Gamma = 0, \quad m_Z = \frac{v\sqrt{g^2 + g'^2}}{2}. \quad (2.39)$$

Typically in most literature, the field Γ_μ is denoted A_μ (denoting the EM field), but this is not used here mostly to avoid the case that the field in the mass eigenstate is confused with that in the gauge eigenstate. Note that in this case, there are three massive bosons and one massless eigenstate.

Defining now the covariant derivative in its mass eigenstates as (assume Y is a constant):

$$D_\mu = \partial_\mu - igA_\mu^a T^a - ig'Y\mathbb{I}_2 B_\mu \implies \quad (2.40)$$

$$\begin{aligned}
 D_\mu = & \partial_\mu - \frac{ig}{\sqrt{2}} \left(W_\mu^+ (T^1 + iT^2) + W_\mu^- (T^1 - iT^2) \right) \\
 & - \frac{i}{\sqrt{g^2 + g'^2}} Z_\mu (g^2 T^3 - g'^2 Y \mathbb{I}_2) \\
 & - i \frac{gg'}{\sqrt{g^2 + g'^2}} \Gamma_\mu (T^3 + Y \mathbb{I}_2).
 \end{aligned} \tag{2.41}$$

The usual electron charge can be found in terms of these coupling constants, and has the form:

$$e = \frac{gg'}{\sqrt{g^2 + g'^2}}, \tag{2.42}$$

and the charge operator can be defined as $\hat{Q} = T^3 + Y \mathbb{I}_2$ (a number of books drop the use of the identity operator, as it would be assumed by default), with the associated quantum number Q . Substituting $Q=-1$ gives the usual form of the EM field. In addition, following the logic of the above linear transformations from gauge eigenstates to mass eigenstates, these can be referred to as functions of the *Weinberg angle* θ , such that:

$$\begin{pmatrix} Z_\mu^0 \\ \Gamma_\mu \end{pmatrix} = \begin{pmatrix} \cos \theta_w & -\sin \theta_w \\ \sin \theta_w & \cos \theta_w \end{pmatrix} \begin{pmatrix} A_\mu^3 \\ B_\mu \end{pmatrix}, \quad \cos \theta_w = \frac{g}{\sqrt{g^2 + g'^2}}, \sin \theta_w = \frac{g'}{\sqrt{g^2 + g'^2}}. \tag{2.43}$$

From the definitions, it can now readily be observed that the constant g can be re-written as:

$$g = \frac{e}{\sin \theta_w}, \tag{2.44}$$

hence, the mass of the W and Z bosons are hence coupled:

$$m_W = m_Z \cos \theta_w. \tag{2.45}$$

Since it is known from nuclear physics that the W boson only couples to left handed fermions (i.e those with left handed chirality), while the Z boson and photon couple to all chiral states, first the kinetic term must be decomposed into its chiral components:

$$\mathfrak{L}_{f,\text{kin}} := i\bar{\psi}\not{\partial}\psi = i\bar{\psi}_L\not{\partial}\psi_L + \psi_R\not{\partial}\psi_R. \tag{2.46}$$

For the right handed fermions, there are singlet up or down type fermions (since there is only the U(1) symmetry), so $Y_R = Q_R$ (hypercharge is equivalent to the actual EM charge for right handed fermions), while for left handed fermions, they are coupled in doublets:

$$E_L = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \quad Q_L = \begin{pmatrix} u \\ d \end{pmatrix}_L, \tag{2.47}$$

with the eigenvalues of the T^3 operator being $\pm 1/2$, and the respective hypercharge assignments $Y=-1/2$ and $Y=+1/6$. A conventional dirac mass term is manifestly gauge symmetry violating, so will need to be introduced later when coupling to the Higgs (through the *Yukawa interaction*, discussed later). Thus the kinetic terms for the first fermion generation (e, ν_e, u, d) take the form:

$$\mathfrak{L}_{f,\text{kin}} = i\bar{E}_L\not{D}E_L + i\bar{E}_R\not{D}E_R + i\bar{Q}_L\not{D}Q_L + i\bar{u}_R\not{D}u_R + i\bar{d}_R\not{D}d_R. \tag{2.48}$$

Noting the coupling of each, the value of the hypercharge for each specific set, the following result can be obtained for the kinetic term [4]:

$$\begin{aligned}
 \mathcal{L} &= i\bar{E}_L \not{\partial} E_L + i\bar{E}_R \not{\partial} E_R + i\bar{Q}_L \not{\partial} Q_L + i\bar{u}_R \not{\partial} u_R + i\bar{d}_R \not{\partial} d_R \\
 &\quad + g(W_\mu^+ J_{W^+}^\mu + W_\mu^- J_{W^-}^\mu + Z_\mu^0 J_Z^\mu) + e\Gamma_\mu J_{EM}^\mu \\
 J_{W^+}^\mu &= \frac{1}{\sqrt{2}}(\nu_L \gamma^\mu e_L + \bar{u}_L \gamma^\mu d_L), \\
 J_{W^-}^\mu &= \frac{1}{\sqrt{2}}(\nu_L \gamma^\mu e_L + \bar{u}_L \gamma^\mu d_L), \\
 J_Z^\mu &= \frac{1}{\cos \theta_w} \left[\frac{1}{2} \bar{\nu}_L \gamma^\mu \nu_L + \bar{e}_L \gamma^\mu \left(-\frac{1}{2} + \sin^2 \theta_W \right) e_L + \sin^2 \theta_w \bar{e}_R \gamma^\mu e_R \right. \\
 &\quad + \bar{u}_L \gamma^\mu \left(\frac{1}{2} - \frac{2}{3} \sin^2 \theta_W \right) u_L - \frac{2}{3} \sin^2 \theta_W \bar{u}_R \gamma^\mu u_R \\
 &\quad \left. + \bar{d}_L \gamma^\mu \left(-\frac{1}{2} + \frac{1}{3} \sin^2 \theta_W \right) d_L + \frac{1}{3} \sin^2 \theta_W \bar{d}_R \gamma^\mu d_R \right], \\
 J_{EM}^\mu &= -\bar{e} \gamma^\mu e + \frac{2}{3} \bar{u} \gamma^\mu u - \frac{1}{3} \bar{d} \gamma^\mu d.
 \end{aligned} \tag{2.49}$$

This determines the coupling of the first generation of fermions to the charged and neutral electroweak bosons as well as to the photon, with the last coupling term being as expected from QED.

2.2.4 Giving masses to fermions

Given the assumptions of the previous sections, there is a coupling referred to as the *Yukawa interaction*, between the scalar field ϕ (the Higgs field) and the left/right handed doublets/singlets, namely (in the specific case of the electron):

$$\mathcal{L}_{e,\text{mass term}} = -\lambda_e \bar{E}_L \phi e_R + \text{h.c.} \tag{2.50}$$

This structure is a scalar because ϕ is a two component vector of two complex scalar fields, E_L is a two component spinor (a doublet state), while the right handed component e_R is a singlet state. The combination of all of these objects produces a scalar. The additional constant λ_e is referred to as the “Yukawa coupling constant”. Now, if we replace the ϕ by the VEV previously discussed, it can be found that the lepton mass term looks like:

$$\mathcal{L}_{e,\text{mass term}} = \frac{1}{2} \lambda_e v \bar{e}_L e_R + \text{h.c.} \tag{2.51}$$

which gives a mass to the electron. We can also find the same for the up and down quarks:

$$m_d = \frac{1}{\sqrt{2}} \lambda_d v, \quad m_u = \frac{1}{\sqrt{2}} \lambda_u v. \tag{2.52}$$

However, the basis which produces the diagonal mass terms is not necessarily the basis corresponding to the independent generations. Following the method of [4], define:

$$u_L^i = (u_L, c_L, t_L), \quad d_L^i = (d_L, s_L, b_L), \tag{2.53}$$

while the primed versions u'_L, d'_L denote the basis such that the Higgs couplings are diagonal. These are related via a unitary transformation, and normally the coupling term for the W boson is written as:

$$J_{W^+}^\mu = \frac{1}{\sqrt{2}} \bar{u}^i \gamma^\mu V_{ij} d_L^j, \quad (2.54)$$

with the matrix V being referred to as the Cabibbo-Kobayashi-Maskawa (CKM) matrix. The off-diagonal terms couple the differing generations, while the diagonal terms couple the same generation doublets together.

2.2.5 Higgs boson mass in the GWS model

Starting from the unitary gauge, we define the scalar field $\phi(x)$ to be:

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \quad (2.55)$$

and given a Lagrangian of type $\mathfrak{L} = (D_\mu \phi)^\dagger (D_\mu \phi) + \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2$, the minimum of the $V(\phi)$ is:

$$v = \left(\frac{\mu^2}{\lambda} \right), \quad (2.56)$$

using the solution for ϕ above (which corresponds to the “unitary gauge”) the potential can be written as:

$$V(\phi) = -\mu^2 h^2 - \lambda v h^3 - \frac{1}{4} \lambda h^4, \quad (2.57)$$

which looking at the quadratic term in $h(x)$ gives the mass term, namely:

$$m_h = \sqrt{2} \mu^2 \equiv \sqrt{\frac{\lambda}{2}} v. \quad (2.58)$$

This now couples the Higgs boson mass to the fundamental value of the VEV (v) (This is however, not the measured Higgs mass directly: we will cover what an experiment likely will see when we come to discussing renormalisation in 2.3).

2.3 Renormalisation

Since a perturbative approach is used in powers of the coupling constants, physically measurable observables such as the mass will differ from their “bare” masses. This is because these observables must be expressed with a perturbative expansion which is likely highly divergent in the energy scale Λ at high orders in the relevant theory’s coupling constant. In order to recover observable results from these divergences, we apply counter terms to cancel some divergences. This procedure turns a problem which is a function of some arbitrarily large energy scale Λ into one where we can measure the value of a given coupling constant at some reference scale q^2 (this is typically taken to be the order of the Zboson mass, but other scales such as ~ 1 TeV can be used), and then obtain the value of the observable at some other energy scale. This is the general idea of *renormalisation*.

Renormalisation arises because of the assumption that beyond some hypothetical maximum energy scale Λ_{UV} , we cannot claim to accurately understand the physics (hence consider this limit a “UV cutoff”), thus the physically measurable quantities become a function of both the theoretical parameter and the energy scale in which they arise. This procedure is primarily used to resolve divergent loop diagrams such as:

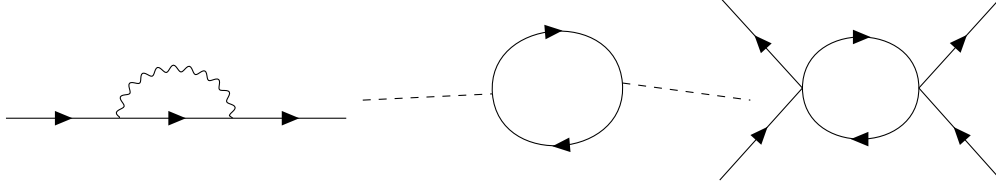


Figure 2.1: Some sample loop diagrams which may appear in calculations of observables.

The degree of divergence for a given diagram can be calculated as (in QED), as per the method and definitions in [4]:

- N_f : number of external fermion lines
- N_γ : number of external photon lines
- N_ϕ : number of external scalar lines.

A typical many loop diagram might have a functional form of the type (including the propagators in loop corrections etc):

$$I_D \sim \int \frac{d^4 k_1 d^4 k_2 \dots d^4 k_L}{(k_i - m) \dots (k_j)^2 (k_n)^2}, \quad (2.59)$$

integrating over the intermediate momenta $(k_1, \dots, k_L)$. Without actually doing this integral, it can be observed that there are two potential singularities in diagrams: cases where $|\mathbf{k}| \rightarrow 0$ (known as an “infrared/IR divergence - normally these are handled by small analytic extensions of the propagators to include $+i\epsilon$, under the understanding that $\epsilon \rightarrow 0$) or in the case of $|\mathbf{k}| \rightarrow \infty$ or $E_k \rightarrow \infty$ (known as an ultraviolet/UV divergence). Treating first the case of the UV divergences, the degree of divergence D can be defined by the rule of thumb [4, 5]:

$$\begin{aligned} D &:= N(\text{powers of } k \text{ in numerator}) - N(\text{powers of } k \text{ in denominator}) \\ &= 4 - N_\gamma \text{ the same applies also to } N_\phi \text{ and } N_f, \end{aligned} \quad (2.60)$$

such that in the UV limit, if $D < 0$, the integral is finite at this upper limit. For the cases where this is not the case, instead we apply a cut-off for the upper limit, Λ (where the end result is that we want to take $\Lambda \rightarrow \infty$). Then instead of defining the indefinite integral with an upper limit, we define it with the cut-off, thus making it finite.

For the integral I_D , we can instead define the finite integral as some function of the cut-off scale Λ and say a coupling constant λ , the energy scale q^2 (the origin of this shall be discussed when it comes to an example, but for now shall be treated as a free parameter), namely:

$$I_D = F(\Lambda, q^2, \lambda) \quad (2.61)$$

Given we do not know the value of λ , and since the integrals depend manifestly on this parameter, we need to consider that the end result should functionally be invariant to such scale choices (since it is effectively arbitrary). So instead, we parametrise a physical parameter λ_p at some scale q^2 , such that:

$$-i\lambda_p = F(\Lambda, q^2, \lambda). \quad (2.62)$$

Thus effectively the integrals can be re-written in some form of the physical parameter λ_p instead of the theoretical parameter λ . This is discussed best with an example, namely by [5] considering some scattering amplitude in ϕ^4 theory up to one loop order (which is a candidate model for meson scattering), such that:

$$\mathcal{M} = -i\lambda + iC\lambda^2 \left[\log\left(\frac{\Lambda^2}{s}\right) + \log\left(\frac{\Lambda^2}{t}\right) + \log\left(\frac{\Lambda^2}{u}\right) \right] + O(\lambda^3), \quad (2.63)$$

with s,t, u referred to as the ‘‘Mandelstam invariants’’ (for the purposes of this example, define them all to be equivalent to some scale μ^2) given a system of input particles p_1, p_2 and output particles p_3, p_4 . This result will be stated without proof, but this calculation can be found in [5][4], as it is somewhat involved and is beyond the scope of this thesis.

According to the ϕ^4 theory, the physical coupling constant is measured at some physical scale $s = t = u = q^2$ to be (up to one loop order):

$$-i\lambda_p = -i\lambda - iC\lambda^2 \left[\log\left(\frac{\Lambda^2}{q^2}\right) + \log\left(\frac{\Lambda^2}{q^2}\right) + \log\left(\frac{\Lambda^2}{q^2}\right) \right] + O(\lambda^3) \quad (2.64)$$

or hiding a few terms under re-definitions:

$$\begin{aligned} -i\lambda_p &= -i\lambda + iC\lambda^2 L_0 + O(\lambda^3) \\ \mathcal{M} &= -i\lambda + iC\lambda^2 L + O(\lambda^3) \end{aligned} \quad (2.65)$$

Now, instead of writing these as two separate equations, since we know what λ_p is, we substitute λ for λ_p :

$$-i\lambda = -i\lambda_p - iC\lambda^2 L_0 + O(\lambda^3) \quad (2.66)$$

which to the appropriate order of perturbation theory (as in, the difference is $O(\lambda^3)$), the λ^2 term can be replaced with the equivalent term but with λ_p^2 . This approximation arises by taylor expanding λ about λ_p and taking the first term only since all higher order terms are neglected, both of λ and λ_p). Thus the theory parameter λ can be written:

$$-i\lambda = -i\lambda_p - iC\lambda_p^2 L_0 + O(\lambda_p^3) \quad (2.67)$$

and similarly, the scattering amplitude looks like:

$$\mathcal{M} = -i\lambda_p - iC\lambda_p^2 L_0 + iC\lambda_p^2 L + O(\lambda_p^3) \quad (2.68)$$

now, by the laws of logarithms, the term $L - L_0$ has the form:

$$L - L_0 = \log\left(\frac{q^2}{\mu^2}\right). \quad (2.69)$$

From this, it can now be observed that we have now removed the dependence of the scattering amplitude on the cut-off scale Λ . This also allows us to physically measure a parameter at a given energy scale, and then again be able to predict its physically measurable value at some different energy scale, with this scale dependent evolution of physical parameters referred to as “renormalisation group flow”. In order to ascertain a good estimate of the physical parameter, customarily measurements are taken at either the Z-mass energy scale or another appropriate scale, then evolved to the energy scale of interest (e.g. to the 1 TeV scale).

This reformulation involving renormalised quantities rather than their bare parameters leads to an alternate form of the lagrangian, namely a renormalised lagrangian with counter terms corresponding to the corrections to the physical theory mentioned above. These counter terms in the lagrangian then cancel the respective divergences.

This particular type of re-normalisation is known as “Pauli-Villars” renormalisation [4, 5], and is not the only method of re-normalisation. Others such as using $D = 4 - 2\epsilon$ with $\lim_{\epsilon \rightarrow 0} \lambda$ also occur, and this is known as “Dimensional regularisation”, but these will not be discussed here. An interested reader is directed to [4, 5] for such additional techniques. In order for Pauli-Villars renormalisation to work, it is required that the one loop corrections are at most logarithmic, producing the required cancellations as seen above.

Since cancellation of higher order loop divergences in known physical parameters secures the stability of the various physical parameters under the perturbative approximations, the capability of a theory to be renormalised is usually used as a benchmark test of a given theory and is generally considered a “necessary” condition of a theory/ set of parameters, but is strictly not a “sufficient” condition.

Appendix C.1 briefly discusses the structure of some renormalisable theories, and examples of the renormalised couplings of QCD and QED are presented as examples of renormalisable theories.

Concluding remarks

This chapter was a brief review of some of the necessary concepts in quantum field theory that will be drawn upon later in this thesis, covering topics in Quantum Field Theory, Gauge Field Theory and a short discussion on Renormalisation. For a more complete summary, we direct the interested reader to [5] for a thorough overview of the material, and [4] for the more complete picture.

Looking beyond the standard model of particle physics

3.1 Limitations of the Standard Model

Given the results of the previous chapter, the full standard model of particle physics can be defined in the lagrangian:

$$\mathcal{L} = \mathcal{L}_{\text{Kin,fermion}} + \mathcal{L}_{QCDterms} + \mathcal{L}_{EW} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa} \quad (3.1)$$

or in the more “compact” notation (hiding the group element summations as an implicit trace where necessary) [4]:

$$\begin{aligned} \mathcal{L} = & -\frac{1}{4}(\text{Tr})(F^{\mu\nu a} F_{\mu\nu a}) + i\bar{\psi}\not{D}\psi + \text{h.c.} \\ & + \bar{\psi}_i y_{ij} \psi_j \phi + \text{h.c.} \\ & + (D_\mu \phi)^\dagger (D^\mu \phi) + V(\phi). \end{aligned} \quad (3.2)$$

This form of the standard model is compactified using a number of short-hands, the first being the definition of the covariant derivative to include each of the transformations under $SU(3)_C \times SU(2)_L \times U(1)_Y$, and the second being the implicit trace of field strength tensors corresponding to the relevant element of the gauge groups.

Line by line, this corresponds to:

1. Kinematic terms of the gauge bosons and the fermion sector, including the gauge couplings of the fermions.
2. Couplings of the fermions to the Higgs Boson through the Yukawa coupling.
3. Kinematic terms of the Higgs Boson and the gauge couplings of the Higgs to the boson sector.

However, this is by no means a complete understanding of nature, and there are several major limitations. The ones primarily relevant to this thesis are as follows:

1. The hierarchy problem: renormalisation of the Higgs mass leads to divergences which cannot be manifestly countered. This also relates to the concept of *naturalness*, and both of these will be discussed in section 3.1.1
2. Dark matter relic abundance: from cosmological observations, rotational velocity curves of galaxies do not match those expected using matter-only models. Dark matter models provide corrections to the rotational velocity curves such that these deviations are removed. This will be covered in section 3.2

and the other remaining items that are still open questions at time of writing this thesis are as follows:

- Neutrino oscillation and masses spectrum observed by solar neutrino observations. Mass ordering of the lepton neutrinos, and/or the existence of sterile neutrinos.
- In the CKM matrix V , the origin of the CP violating phase δ .
- Unification of the standard model with general relativity: An active area of research, limited at the experimental level by the differing energy scales involved between the electroweak sector, and gravity.

3.1.1 Hierarchy problem

The hierarchy problem can be formulated in terms of the renormalisation of the Higgs mass, which can also be re-phrased in terms of the differences in scales between the electroweak (EW) scale and the Planck scale. In terms of the Higgs mass formulation, following on from the discussion in renormalisation in section 2.3, the physical mass (not the theoretical mass) can be estimated as a function of the energy scale Λ , and that the mass is proportional to the value of the VEV.

Following the argument of [6], a one-loop renormalisation correction involving a top quark loop (Figure 3.1) takes the form [7]:

$$\delta m_{H, \text{top loop}}^2 = -2N_c |y_t| \int \frac{d^4 k}{(2\pi)^4} \frac{k^2 + m_t^2}{(k^2 - m_t^2)^2}. \quad (3.3)$$

where N_c is the number of colors, y_t is Yukawa coupling constant for the top-quark and m_t the top mass. The factor of two arises from the distinct chiral states t_L, t_R

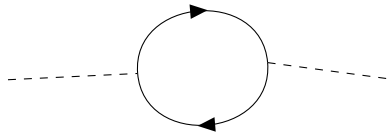


Figure 3.1: 1-loop renormalisation correction to the Higgs mass from a top quark loop.

When applying a cut-off of some scale Λ , it can be found that the mass correction takes the form [7]:

$$\delta m_{H, \text{top loop}}^2 = -\frac{N_c |y_t|}{8\pi^2} \left[\Lambda^2 - 3m_t^2 \ln \left(\frac{\Lambda^2 + m_t^2}{m_t^2} + O(\Lambda^3) \right) \right]. \quad (3.4)$$

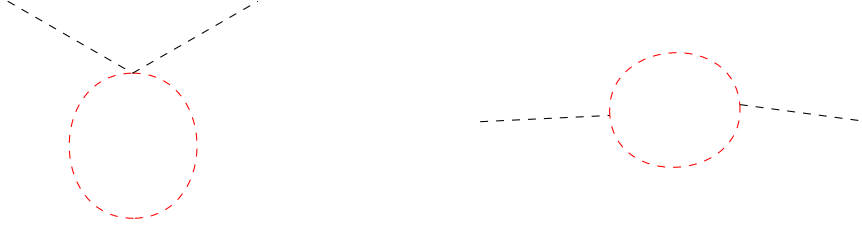


Figure 3.2: One-loop corrections to the higgs (dashed black lines) mass associated to the scalars $\tilde{t}_L, \tilde{t}_R$ (dashed red lines).. Left denotes the quartic coupling loop and right denotes the trilinear coupling

Now, suppose that for each of the N_C fermion states we have N scalar states ($\tilde{t}_L, \tilde{t}_R$, which couple to the higgs via either a quartic or trilinear coupling. Suppose these scalars couple to the Higgs with a constant λ in the case μ_L, μ_R and have masses $m_{\tilde{t}_L}, m_{\tilde{t}_R}$. The scalar contribution would take the forms [7]:

$$\delta m_{H, \text{quartic loop}}^2 = \frac{\lambda N}{16\pi^2} \left[2\Lambda^2 - m_{\tilde{t}_L}^2 \ln \left(\frac{\Lambda^2 + m_{\tilde{t}_L}^2}{m_{\tilde{t}_L}^2} \right) - m_{\tilde{t}_R}^2 \ln \left(\frac{\Lambda^2 + m_{\tilde{t}_R}^2}{m_{\tilde{t}_R}^2} \right) + \dots \right]. \quad (3.5)$$

$$\delta m_{H, \text{tri-linear loop}}^2 = -\frac{N}{16\pi^2} \left[\mu_L^2 \ln \left(\frac{\Lambda^2 + m_{\tilde{t}_L}^2}{m_{\tilde{t}_L}^2} \right) + \mu_R^2 \ln \left(\frac{\Lambda^2 + m_{\tilde{t}_R}^2}{m_{\tilde{t}_R}^2} \right) + \dots \right]. \quad (3.6)$$

Looking at this and the previous statements relating to the top loop coupling, a cancellation of the Λ^2 term in the fermion 1 loop expansion can be found if $N = N_C$, $\lambda = |y_t|^2$. In addition, if $m_t = m_{\tilde{t}_R} = m_{\tilde{t}_L}$, $\mu_L^2 = \mu_R^2 = 2\lambda m_t^2$ this leads to a cancellation of the logarithmic divergences.

However, within the SM, it is possible to reconcile this divergence without the scalar partner, largely by noting that these large cancellations could be offset by some higher order contribution not specified in the theory (since we truncate the perturbation series to one loop order), known as *fine tuning* corrections. This fine tuning, however would require large corrections, which are not considered to be “natural”, namely that free parameters of the standard model are generally invariant of these large fine tunings. Thus, *Supersymmetry* produces the required cancellations without further fine tuning to the standard model, whereby for each fermion chiral state, an equivalent scalar partner particle exists, and vice versa for the vector bosons and scalars of the standard model.

3.2 Dark matter

In this section, we move from the focus of the subatomic processes discussed in collider experiments to cosmological observations based on extra-solar observations, which cannot be otherwise covered by either Newtonian dynamics or General relativity.

A substantial motivator is that according to Newtonian dynamics, for a rotating galaxy (assuming a uniform density for simplicity), the velocity for a given object as a function of distance from galactic centre r takes the form $v \propto r^{-1/2}$. However, in practice, what is

observed is that the velocity is effectively flat as a function of radius (see figures 3.3,3.4), thus an additional counter term is required. This term is usually referred to as a “Dark matter halo”, since the mass density increases toward the exterior of the galactic spirals, forming a ring of dark matter toward the extremities of the galaxies.

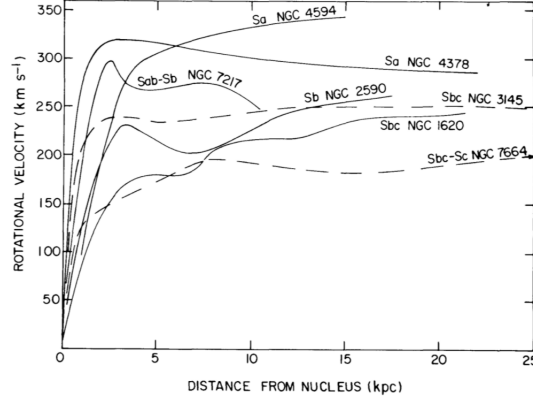


Figure 3.3: Rotational velocities for seven galaxies as a function of distance from their galactic centres as observed in [8].

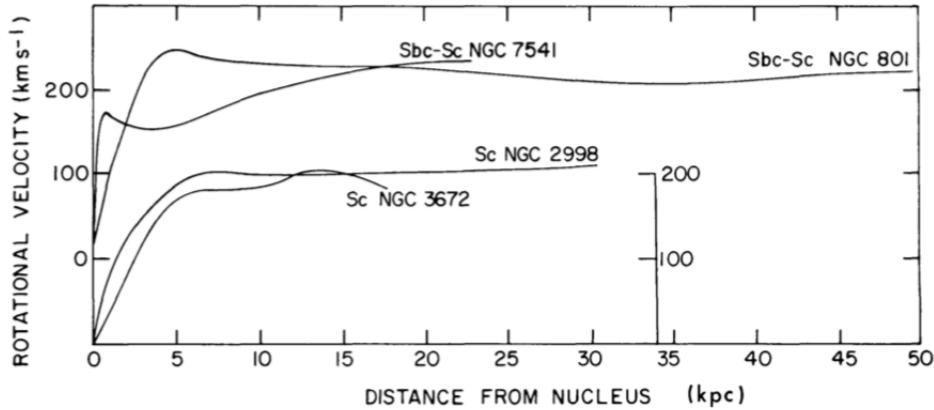


Figure 3.4: Rotational velocity curves of galaxies NGC 7541 and NGC 801 from [8].

However, these rotational velocity curves are not the only motivating evidence for dark matter. Another is the interaction of two colliding galaxies in the Bullet cluster (1E 0657-558 - figure 3.5), whereby through direct observation by the Chandra satellite and via weak lensing in observations of the Hubble Space Telescope. This particular cluster in a visual image looks like a shockwave between two colliding galaxies (hence the name), but notably from the gravitational lensing of light from the cluster, the majority of the mass of the galactic system is displaced from the visible region (as observed by the gravitational lensing).

A notable point is that these dark matter particles are most likely neutral (or only very slightly charged) given current experimental limits from Planck [10], and are most likely non-baryonic. Neutrino abundances are also considered too small, at least insofar as the current standard model left handed chiral neutrinos (or right handed chiral anti-neutrinos) are concerned. These dark matter particles are often referred to as “weakly interactive

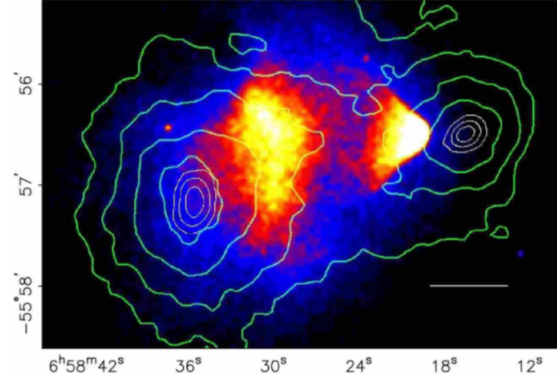


Figure 3.5: Chandra satellite X-ray image super-imposed with the weak gravitational lensing mass measurements, from [9]. The contours refer to contours on the mass density of each galaxy from the lensing measurements. The centers of mass observed for each galaxy are readily observable to be displaced from their visible (or at least most luminous) centres.

massive particles”, since principally, the dark matter states are not expected to be strongly coupled to the standard model (besides the weak interaction on high energy scales, and gravity on low energy scales).

The exact form of dark matter is widely varied, depending on the class of theory that the reader wishes to look at, but specific classes of Supersymmetry models (particularly those which preserve “R-parity” - more on this in the next section) that have a lightest stable particle can also be considered a dark matter candidate.

3.3 Physics beyond the standard model: Supersymmetry (SUSY)

We require an extension to Lie Algebras (see appendix for a brief discussion of these) to allow for symmetry generators which manifestly anti-commute with each other, as well as commute, known as a *graded lie algebra*[11, 12]. This forms a composite space of a set of operators which are purely bosonic (defined by commutation relations) and fermionic (defined by anti-commutation relations), or generators of the symmetry group defined by:

$$T^a T^b - (-1)^{y_a y_b} T^b T^a = i f^{abc} T^c \quad (3.7)$$

with f^{abc} the structure constants as per in the bosonic case, while the term y_a being an indicator function (where $y_a = 1$ if T^a is a fermionic generator, and $y_a = 0$ in the case that T^a is bosonic). This effectively means that the vector space spanned by the generators T^a can be decomposed into two components, fermionic and bosonic which are direct-summed together.

Since the quantities involved manifestly include spinors, henceforth additional notation will be defined, namely that in this section, roman indices (a,b,c) unless otherwise stated will correspond to spinor indices, namely two component spinors. In the Weyl basis, a four component spinor can be decomposed into two two-component spinors corresponding to

the left and right helicity states (when the fermion is massless). Thus, we define as four component fermion spinor $\Psi(x)$ as:

$$\Psi(x) = \begin{pmatrix} \epsilon^a \\ \chi^{\dagger\dot{a}} \end{pmatrix}; \quad a, \dot{a} = 1, 2. \quad (3.8)$$

The dotted and non-dotted indices are independent co-ordinates and the functions ϵ and χ denote the eigenvectors of the helicity operator (for massless fermions). A metric can also be defined on these spinors, lowering or raising the operators, by applying the transform:

$$g_{ab} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad g^{ab} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad (3.9)$$

and similarly for the dotted coordinates.

A *super-algebra* can be defined using two fermionic generators ($Q, Q^\dagger$) in addition to generators of the Lorentz group (in particular, the momentum operator P^μ), defined by[12]:

$$\begin{aligned} \{Q_a, Q_b\} &= 0, & \{Q_a^\dagger, Q_b^\dagger\} &= 0, \\ \{Q_a, Q_{\dot{a}}^\dagger\} &= 2\sigma_{a\dot{a}}^\mu P_\mu. \end{aligned} \quad (3.10)$$

In addition, an additional requirement of these generators is that they commute with the generators of translations P_μ , or:

$$[Q_a, P_\mu] = [Q_a^\dagger, P_\mu] = 0. \quad (3.11)$$

The effect of these Q operators defined have the effect of:

$$Q |\text{fermion}\rangle \propto |\text{boson}\rangle, \quad Q^\dagger |\text{boson}\rangle \propto |\text{fermion}\rangle, \quad (3.12)$$

such that they transform the fermion and boson states into their respective partner state.

3.3.1 Poincaré group of rotations, boosts and translations

In order to discuss the Poincaré group, we first discuss the Lorentz group, a strict subset of the Poincaré group (for all relevant cases, we assume proper orthochronous Lorentz transformations only), corresponding to Lorentz boosts, rotations and translations.

$$x^\mu \rightarrow x'^\mu = \Lambda_\nu^\mu x^\nu \quad (3.13)$$

can be defined by a form [12]:

$$\Lambda_\nu^\mu = [\exp(i\omega_{\rho\sigma} M^{\rho\sigma})]_\nu^\mu \quad (3.14)$$

the number $\omega_{\rho\sigma}$ is anti-symmetric under interchange of indices, with six generators spanning this group. Normally, these are subdivided into the usual angular rotation generators $J_i, i = 1, 2, 3$ and generators of Lorentz boosts $K_i, i = 1, 2, 3$, and this lie algebra is defined by the commutation relations:

$$\begin{aligned} [J_i, J_j] &= i\epsilon_{ijk} J_k \\ [J_i, K_j] &= i\epsilon_{ijl} K_l \\ [K_i, K_j] &= -i\epsilon_{ijk} J_k \end{aligned} \quad (3.15)$$

From these generators, the matrix $[M_{\mu\nu}]$ is defined as [12]:

$$M_{\rho\sigma} := \begin{pmatrix} 0 & K_1 & -K_2 & -K_3 \\ K_1 & 0 & J_3 & -J_2 \\ K_2 & -J_3 & 0 & J_1 \\ K_3 & J_1 & -J_2 & 0 \end{pmatrix} \quad (3.16)$$

or in a more readily useful form:

$$[M_{\rho\sigma}]_{\nu}^{\mu} = i(g_{\sigma\nu}\delta_{\rho}^{\mu} - g_{\rho\nu}\delta_{\sigma}^{\mu}).$$

The form of the matrix M and the parameter ω define the most general case of all Lorentz transformations and rotations, with a given parameter choice $\omega_{\rho\sigma}$ defining the explicit transformation type chosen. Note the Poincaré group is an extension on the Lorentz group to include space-time transformations, and an additional four elements are added to the group (corresponding to translations in the four spacetime dimensions).

From the definitions, the following identities can be obtained, and hold for the entire Poincaré group:

$$\begin{aligned} [M^{\mu\nu}, M^{\rho\sigma}] &= -i(g^{\mu\rho}M^{\nu\sigma} - g^{\mu\sigma}M^{\nu\rho} - g^{\nu\rho}M^{\mu\sigma} + g^{\nu\sigma}M^{\mu\rho}) \\ [M^{\mu\nu}, P^{\alpha}] &= i(g^{\nu\alpha}P^{\mu} - g^{\mu\alpha}P^{\nu}) \\ [M^{\mu\nu}, P^2] &= 0 \end{aligned} \quad (3.17)$$

with this last operator also referred to as a “Casimir operator”, and when applied to a mass state on-shell, it will have the eigenvalue m^2 . In addition, an additional Casimir operator W^2 [12] (the square of a “Pauli-Ljubanski polarization vector”) can be defined, with:

$$W_{\mu} := \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}P^{\nu}M^{\rho\sigma}, \quad [M_{\mu\nu}, W^2] = [P_{\mu}, W^2] = 0 \quad (3.18)$$

It can be shown with some algebra in addition (a full calculation can be found in [12][pp30]):

$$W^{\mu}P_{\mu} = 0 \quad (3.19)$$

in the case of a massive particle (i.e. such that $P^{\mu} = (m, 0, 0, 0)$), the eigenvalues of the second Casimir operator can be denoted [12] as $-m^2s(s+1)$ where s denotes the spin quantum number with the usual rules. At rest, the W ’s can be defined as:

$$W_i = \frac{1}{2}\epsilon_{i0jk}P^0S^{jk}; \quad W^{\mu}W_{\mu} = -|\mathbf{W}|^2 = -m^2|\mathbf{S}|^2 \quad (3.20)$$

and this means that the individual components of the $\mathbf{S}$ vector can be defined as:

$$S^i = \frac{1}{2}\epsilon^{ijk}S_{jk} \quad (3.21)$$

This manifestly couples spin to the Poincaré group transformations (through the W^2 invariant) and is such that the spin of a particle is held invariant under transformations of the Poincaré group.

3.3.2 Coleman-Mandula No-Go theorem and associated exceptions

A key theorem in the underpinning supersymmetry (or more a limitation that must be avoided) is the famous theorem by S. Coleman and J. Mandula in 1967 [13], referred to as the *No go theorem*. A short summary of the theorem [13], and discuss its consequences will be mentioned here, and a more formal statement for an interested reader can be found in [13] or [11].

The Coleman-Mandula No Go Theorem states that a symmetry group can only consist of two distinct sub-groups (which may themselves be empty) when defined as a Lie Algebra by commutation relations. These consist of a subset of the Poincaré group up to basis transformations, and a separate sub-group that corresponds to internal symmetries. These two sub-groups are independent, such that any internal symmetry group operator will commute with any elements of the Poincaré group. Since from the earlier discussion, the spin of a particle is an invariant of the Poincaré group, no internal symmetry can change the particle's spin.

However, it can be noted that it does not specify any such generators which anti-commute, and this particular extension [14][11][15] is known as the Haag, Łopuszański, and Sohnius extension. This allows a symmetry group which is defined by a Graded-Lie Algebra. In fact, the supersymmetry algebra defined in the previous subsection will cover both internal transformations (such as changing spin of a particle) and Poincaré transformations.

3.3.3 SUSY model from first principles: Wess-Zumino Model

The first class of supersymmetric model was proposed by Wess and Zumino [7, 16], in the case of one weyl basis fermion, and one scalar. Some pre-amble definitions need to also be defined, namely the objects σ^μ (a subset of the standard γ operators, corresponding to either the 2×2 identity matrix $\mathbb{I}_2$ or the Pauli Matrices) [7]:

$$\sigma_{a\dot{a}}^\mu = (\mathbb{I}_2, \sigma^i)_{a\dot{a}}, \quad \bar{\sigma}^{\mu\dot{a}a} = (\mathbb{I}_2, -\sigma^i)^{\dot{a}a} \quad (3.22)$$

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (3.23)$$

In the Weyl basis, gamma matrices take the form:

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix} \quad (3.24)$$

and this basis is chosen since in the case that $m = 0$, the spinor $\Psi(x)$ decomposes into two helicity eigenstates (which can be associated to the first two and last two components respectively).

With all of this pre-amble in hand, for a single Weyl spinor (so a single fermion helicity state ψ), and a single scalar ϕ , the simplest model that is supersymmetric will take the form (assume for the moment that both particles are massless) [7]:

$$\mathcal{L} := \mathcal{L}_{\text{Boson}} + \mathcal{L}_{\text{Fermion}} := \partial^\mu \phi (\partial_\mu \phi)^* + i\psi^\dagger \bar{\sigma}^\mu \partial_\mu \psi \quad (3.25)$$

In order to show this simple lagrangian is invariant under SUSY, for an infinitesimal transformation of the type:

$$\begin{aligned}\phi &\rightarrow \phi + \delta\phi \\ \psi &\rightarrow \psi + \delta\psi\end{aligned}\tag{3.26}$$

and since we expect the rotation of a scalar field into a fermion field, we can identify

$$\begin{aligned}\delta\phi &= \epsilon^a \psi_a \equiv \epsilon\psi \\ \delta\phi^* &= \epsilon^{\dagger\dot{a}} \psi_a^\dagger\end{aligned}\tag{3.27}$$

Note, due to the properties of spin summation the object $\epsilon\psi \equiv \epsilon^a \psi_a = \psi^a \epsilon_a \equiv \psi\epsilon$. Thus, the change in the lagrangian from the scalar term looks like

$$\delta\mathcal{L}_S = \epsilon(\partial^\mu\psi)(\partial_\mu\phi)^* + \epsilon^\dagger(\partial^\mu\psi^\dagger)(\partial_\mu\phi)\tag{3.28}$$

For the fermion term, two parts must be considered: the correct combination of spinor and Lorentz indices, given that the ϵ has a spinor index, and the correct form looks like:

$$\delta\psi_a = -i(\sigma^\nu\epsilon^\dagger)_a\partial_\nu\phi, \quad \delta\psi_a^\dagger = i(\epsilon\sigma^\nu)_{\dot{a}}(\partial_\nu\phi)^*\tag{3.29}$$

and the change in the fermion term will look like

$$\delta\mathcal{L}_f = -\epsilon\sigma^\nu\partial_\nu\phi^*\bar{\sigma}^\mu\partial_\mu\psi + \psi^\dagger\bar{\sigma}^\mu\sigma^\nu\epsilon^\dagger\partial_\mu\partial_\nu\phi\tag{3.30}$$

massaging this term using some properties of the Pauli matrices (which can be shown by direct calculation):

$$[\sigma^\mu\bar{\sigma}^\nu + \sigma^\nu\bar{\sigma}^\mu]_a^b = 2g^{\mu\nu}\delta_a^b, \quad [\bar{\sigma}^\mu\sigma^\nu + \bar{\sigma}^\nu\sigma^\mu]_{\dot{a}}^{\dot{b}} = 2g^{\mu\nu}\delta_{\dot{a}}^{\dot{b}}$$

the fermion piece of the lagrangian change in the SUSY transformation instead looks like:

$$\begin{aligned}\delta\mathcal{L}_f &= -\epsilon(\partial^\mu\psi)(\partial_\mu\phi)^* - \epsilon^\dagger(\partial^\mu\psi^\dagger)(\partial_\mu\phi) \\ &\quad + \partial_\mu(\epsilon\sigma^\mu\bar{\sigma}^\nu\psi(\partial_\nu\phi)^* - \epsilon\psi(\partial^\mu\phi)^* + \epsilon^\dagger\psi^\dagger\partial^\mu\phi)\end{aligned}\tag{3.31}$$

When the scalar and fermion terms are added together, the only remaining term is a total derivative, thus the action is invariant under the SUSY transformations. In addition, to prove that we have a complete group, we need to show that the group of SUSY transformations Q closes [7, 11], namely that the commutator of two infinitesimal transformations as above is itself a member of the group.

The closure proof [7] is complete, when all particles are on-shell, namely when the conditions $\partial^\mu\partial_\mu\phi = 0$ and $\bar{\sigma}^\mu\partial_\mu\psi = 0$ (since we did not specify mass terms) are satisfied. To correctly define the SUSY transformations in general for all fields ψ, ϕ , a non-physical *auxilliary term* must be added to the lagrangian, namely:

$$\mathcal{L}_{\text{aux}} = \mathcal{F}^*\mathcal{F}\tag{3.32}$$

This is a “non-physical” term since the equation of motion is trivial $\mathcal{F} = 0$. The transformation of this non-physical book-keeping term can be very carefully chosen by construction [7] to

remove any terms which break the closure of the SUSY algebra. The commutator of two infinitesimal SUSY transformations $\delta_{\epsilon_1}, \delta_{\epsilon_2}$ including the $\mathcal{F}$ term take the form:

$$\begin{aligned} [\delta_{\epsilon_1}, \delta_{\epsilon_2}]X &= -i(\epsilon_1 \sigma^\mu \epsilon_2^\dagger - \epsilon_2 \sigma^\mu \epsilon_1^\dagger) \partial_\mu X \\ X &= \phi, \phi^*, \psi, \psi^\dagger, \mathcal{F}, \mathcal{F}^* \text{ (a super-multiplet)} \end{aligned} \quad (3.33)$$

The commutator of two SUSY transformations can be shown by considering four parts:

1. The infinitesimal transformations of the fermion and scalar terms (including the auxiliary terms), which has the form:

$$\delta\psi_a = -i(\sigma^\nu \epsilon^\dagger)_a \partial_\nu \phi^\dagger + \epsilon_a \mathcal{F}, \quad \delta\phi = \epsilon^a \psi_a, \quad +\text{h.c} \quad (3.34)$$

2. The action of multiple transformations on an element of the super-multiplet $\delta_{\epsilon_1} \delta_{\epsilon_2} X = \delta_1(\delta_{\epsilon_2} X)$.
3. The infinitesimal transformation of the $\mathcal{F}$ term: $\delta\mathcal{F} = -i\epsilon^\dagger \bar{\sigma}^\mu \partial_\mu \psi$, $\delta\mathcal{F}^* = i\partial_\mu \psi^\dagger \bar{\sigma}^\mu \epsilon$.
4. Identities of products of spinors [7], primarily for spinors χ, ϵ, η : $\chi_a(\epsilon^b \eta_b) + \epsilon_a(\chi^b \eta_b) + (\epsilon^b \chi_b)\eta_a = 0$ (known as ‘Fierz identities’).

In the case of the variation of a Weyl spinor ψ_a , this commutator can be shown to hold:

$$\begin{aligned} \delta_{\epsilon_1} \delta_{\epsilon_2} \psi_a &= -i(\sigma^\nu \epsilon_2)_a \partial_\nu (\epsilon_1^b \psi_b) + \epsilon_{2a} \delta_{\epsilon_1} \mathcal{F} \\ [\delta_{\epsilon_1}, \delta_{\epsilon_2}] \psi_a &= -i \left(\epsilon_1^b \sigma_{bb}^\mu \epsilon_2^{\dagger b} - \epsilon_2^b \sigma_{bb}^\mu \epsilon_1^{\dagger b} \right) \partial_\mu \psi_a \\ &\quad + i \left(\epsilon_{1a} \epsilon_2^{\dagger b} \bar{\sigma}_{bb}^\mu (\partial_\mu \psi)^b - \epsilon_{2a} \epsilon_2^{\dagger b} \bar{\sigma}_{bb}^\mu (\partial_\mu \psi)^b \right) \\ &\quad + \epsilon_{1a} (-i\epsilon_{b2}^\dagger \bar{\sigma}^{\mu bb} \partial_\mu \psi_b) - \epsilon_{2a} (-i\epsilon_{1b}^\dagger \bar{\sigma}^{\mu bb} \partial_\mu \psi_b) \end{aligned} \quad (3.35)$$

Performing all of the cancellations yields the result $[\delta_{\epsilon_1}, \delta_{\epsilon_2}] \psi_a = -i(\epsilon_1 \sigma^\mu \epsilon_2^\dagger - \epsilon_2 \sigma^\mu \epsilon_1^\dagger) \partial_\mu \psi_a$.

The charges associated to the SUSY transformation are related to the supercurrent [7]:

$$\begin{aligned} J_a^\mu &= (\sigma^\nu \bar{\sigma}^\mu \psi)_a (\partial_\nu \phi)^*, \\ J_a^\dagger &= (\psi^\dagger \bar{\sigma}^\mu \sigma^\nu)_a \partial_\nu \phi, \\ Q_a &= \sqrt{2} \int d^3x J_a^0 Q_a^\dagger = \sqrt{2} \int d^3x J_a^{\dagger 0}, \end{aligned} \quad (3.36)$$

and the (anti-)commutation relations of these Q ’s are exactly that of the super-algebra defined earlier.

Hence, the simplest form of the *Wess-Zumino model* then can be described as [7, 11, 16]:

$$\mathcal{L} = (\partial^\mu \phi^j)^* (\partial_\mu \psi^j) + i\psi^j \bar{\sigma}^\mu \partial_\mu \psi_j + \mathcal{F}^{*j} \mathcal{F}_j \quad (3.37)$$

where j runs over the possible super-multiplets (the case discussed here is $j = 1$).

3.3.3.1 Interactions

The general form of terms which are invariant under SUSY transformations that can be used to model interactions can take the form [7, 11]:

$$\mathcal{L}_{int} = -\frac{1}{2}W^{jk}\psi_j\psi_k + W^j\mathcal{F}_j + h.c., \quad (3.38)$$

where W^j is a function of ϕ, ϕ^* , and W^{jk} is symmetric under index interchange. SUSY invariance requires that the W^{jk} are analytic in fields ϕ , and can be specified as:

$$W^{jk} = \frac{\partial^2}{\partial\phi_j\partial\phi_k}W; \quad W = E^j\phi_j + \frac{1}{2}M^{jk}\phi_j\phi_k + cy^{jkn}\phi_j\phi_k\phi_n, \quad (3.39)$$

with the matrices M and y are the mass and Yukawa matrices, and W is the “super-potential”. In addition, the term W^j can be defined as:

$$W^j = \frac{\partial W}{\partial\phi_j} \quad (3.40)$$

Thus the interaction term can be written as:

$$\mathcal{L} = \frac{\partial^2 W}{\partial\phi_j\partial\phi_k}\psi_j\psi_k + \frac{\partial W}{\partial\phi_j}\mathcal{F}_j + h.c. \quad (3.41)$$

3.3.3.2 SUSY invariant gauge interactions

Some SUSY invariant interaction terms can be described using the SUSY Yang Mills Lagrangian (this would, for instance cover SU(3)) [7]:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^t F^{t\mu\nu} + i\lambda^{\dagger t}\bar{\sigma}^\mu D_\mu\lambda^t + \frac{1}{2}\mathcal{D}^{*t}\mathcal{D}^t + \mathcal{L}_{\text{free}} \quad (3.42)$$

where the parameter t runs over the elements of the gauge group (i.e. SU(3)), $\mathcal{D}$ is the non-physical auxilliary field, and

$$\begin{aligned} D_\mu\lambda^t &= \partial_\mu\lambda^t - gf^{tuv}A_\mu^u\lambda^v \\ F_{\mu\nu}^t &= \partial_\mu A_\nu^t - \partial_\nu A_\mu^t - gf^{tuv}A_\mu^uA_\nu^v \end{aligned} \quad (3.43)$$

which is invariant under the SUSY transformation [7]:

$$\begin{aligned} \delta A_\mu^t &= -\frac{1}{\sqrt{2}}\left[\epsilon^\dagger\bar{\sigma}_\mu\lambda^t + \lambda^{\dagger t}\bar{\sigma}_\mu\epsilon\right], \\ \delta\lambda_a^t &= -\frac{i}{2\sqrt{2}}(\sigma^\mu\bar{\sigma}^\nu\epsilon)_a F_{\mu\nu}^t + \frac{\sqrt{1}}{\sqrt{2}}\epsilon_a\mathcal{D}^t, \\ \delta\lambda_a^{\dagger t} &= \frac{i}{2\sqrt{2}}\left(\epsilon^\dagger\bar{\sigma}^\nu\sigma^\mu\right)_{\dot{a}} F_{\mu\nu}^t + \frac{1}{\sqrt{2}}\epsilon_{\dot{a}}^{\dagger t}\mathcal{D}^t, \\ \delta\mathcal{D}^t &= \frac{-i}{\sqrt{2}}\left[\epsilon^\dagger\bar{\sigma}^\mu D_\mu\lambda^t - D_\mu\lambda^{\dagger t}\bar{\sigma}^\mu\epsilon\right], \end{aligned} \quad (3.44)$$

(t runs over group elements, a over spin indices.)

The proof of this invariance, the closure etc and associated super-charges will be assumed here, but can be found in [6, 7]. The super multiplet in this case is the vector field A , the fermion spinor λ , its hermitian conjugate $\lambda^\dagger$ and the auxiliary field $\mathcal{D}$. This couples each boson state to a single fermion state (the auxiliary term will have no physical meaning when considered “on-shell”).

Having looked at the free terms for both the fermions and gauge bosons, we can now look at the interaction terms. In a similar way to the standard model, the covariant derivatives for the fermion and their scalar superpartner can be defined:

$$\begin{aligned} D_\mu \phi_j &= \partial_\mu \phi_j + ig A_\mu^a G^a \phi_j, \\ (D_\mu \phi_j)^* &= (\partial_\mu \phi)^* - ig A_\mu^a G^a \phi_j^*, \\ D_\mu \psi_j &= \partial_\mu \psi_j + ig A_\mu^a G^a \psi_j, \end{aligned} \quad (3.45)$$

which correspond to the covariant derivatives coupling non-abelian gauge groups spanned by generators G^a to either the scalars or fermions respectively (assuming the same coupling to both).

The allowed supersymmetric interaction terms between the gauge field, its superpartners and the matter fermion and its superpartner (along side the auxiliary fields $\mathcal{D}, \mathcal{F}$) turn out to be renormalisable if [7]:

$$\mathcal{L}_{\text{SUSY}} = \mathcal{L}_{\text{Free-WZ}} + \mathcal{L}_{\text{Gauge sector}} - \sqrt{2}g \left[(\phi^* G^a \psi) \lambda^a + \lambda^{\dagger a} (\psi^\dagger G^a \phi) \right] + g(\phi^* G^a \phi) \mathcal{D}^a. \quad (3.46)$$

These interactions can be represented as Feynman diagrams, are presented in 3.6.

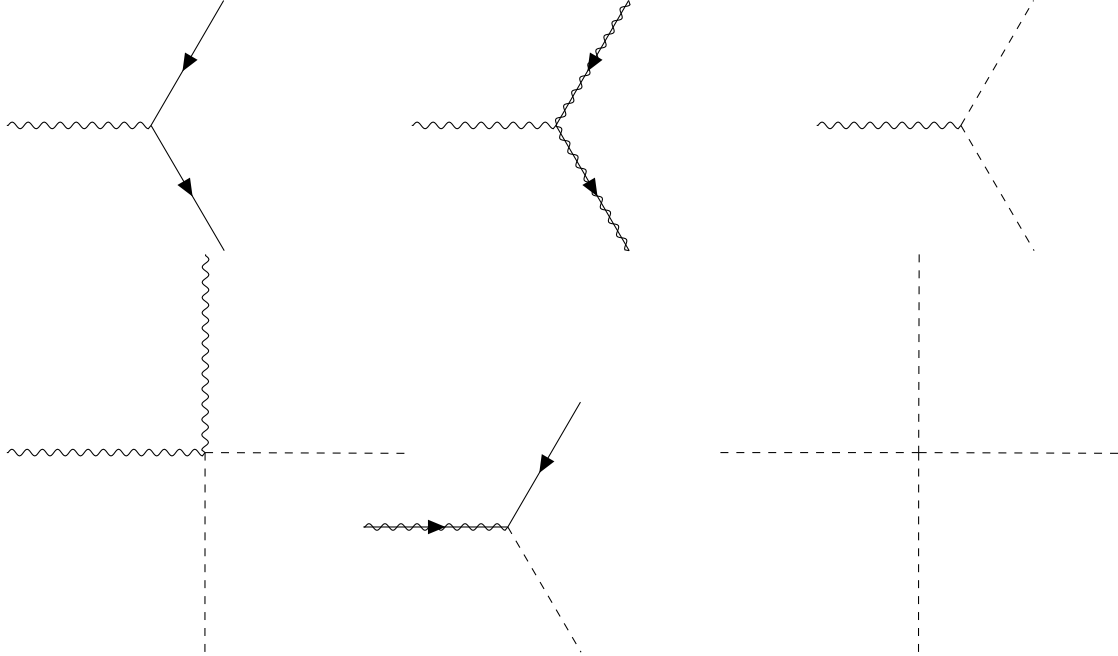


Figure 3.6: Feynman diagrams corresponding to different interaction vertices of the SUSY Yang Mills theory. Solid lines refer to the fermion, while dashed refer to the sfermion scalar. Wavy lines correspond to either the gauge boson (no-arrow) or the gaugino fermion (arrows).

3.4 Minimal Supersymmetric Standard Model

The set of multiplets for each particle/sparticle in the minimal supersymmetric standard model (MSSM) is [6, 11, 12] defined in ftables 3.1 and 3.2.

Description	Label	Spin-0 content	Spin 1/2 content
Quarks/squarks (3 generations)	Q	$(\tilde{u}_L, \tilde{d}_L)$	(u_L, d_L)
	$\bar{u}$	$\tilde{u}_R^*$	$u_R^\dagger$
	$\bar{d}$	$\tilde{d}_R^*$	$d_R^\dagger$
Leptons/sleptons (3 generations)	L	$(\tilde{\nu}, \tilde{e}_L)$	(ν, e_L)
	$\bar{e}$	$\tilde{e}_R^*$	$e_R^\dagger$
Higgs, Higgsino	H_u	(H_u^+, H_u^0)	$(\tilde{H}_u^+, \tilde{H}_u^0)$
	H_d	(H_d^-, H_d^0)	$(\tilde{H}_d^-, \tilde{H}_d^0)$

Table 3.1: Multiplets [6] for the left and right handed Weyl up-type fermions (u) or down-type fermions (d) and similarly for leptons. In addition, the partners to the Higgs boson are presented. The SUSY partners are normally labelled as per common convention marked with a tilde. The SM Higgs field is retrieved from the linear combination of the H_u^0 and H_d^0 states. The Higgs sector is extended to include charged and heavier than SM states (in what is known as a *two Higgs Doublet Model* (2HDM)).

Name	Spin-1/2 content	Spin 1 content
Gluino, gluon	$\tilde{g}$	g
Wino, W bosons	$\tilde{W}^\pm, \tilde{W}^0$	$W^\pm, W^0$
bino, B boson	$\tilde{B}^0$	B^0

Table 3.2: Gauge & gaugino super-multiplets prior to spontaneous symmetry breaking of the SM $SU(2)_L \times U(1)_Y$ symmetry. The supersymmetric partner particles are denoted with a tilde.

In the formulation of a supersymmetric standard model, it is possible to include terms which violate baryon and lepton number, which is normally preserved under the standard model. The consequences of this means that the proton can decay into an anti-lepton + meson via a SUSY particle exchange, however such a process has a lifetime [7] on the order of $\tau > 10^{32}$. Such a long decay time is many orders of magnitude longer than the lifetime of the universe!

In this thesis, we consider an additional assumed symmetry known as *R-parity*, with the eigenvalues on some eigenstate $|\psi\rangle$ is defined as:

$$\hat{R}|\psi\rangle = (-1)^{(3(B-L)+2s)}|\psi\rangle, \quad (3.47)$$

where B is the baryon number, L the lepton number and s is the particle spin. The notion of the R parity is such that:

$$\begin{aligned} \hat{R}|\text{Matter particle}\rangle &= +1 \times |\text{Matter particle}\rangle, \\ \hat{R}|\text{Super partner}\rangle &= -1 \times |\text{Super partner}\rangle. \end{aligned} \quad (3.48)$$

Any MSSM terms which preserve this symmetry are considered “R-parity conserving”, and as a consequence, any proton decay inducing terms are removed. Some ATLAS LHC

analyses look at R-parity violating models, such as [17], however this thesis assumes R-parity conserving models only. A consequence of this is such that in order to convert between standard model processes and supersymmetry, they must be pair produced (or be a process involving a final state SUSY particle). A possible interaction R-parity conserving interactions is presented in figure 3.7.

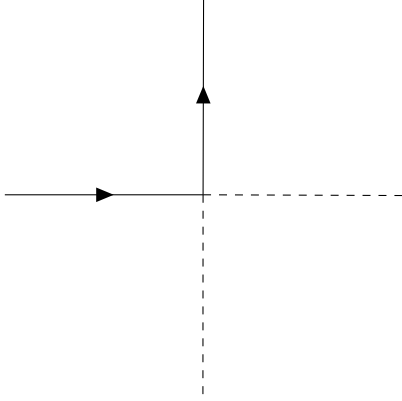


Figure 3.7: An example R-parity conserving interaction vertex between SUSY scalars (sfermions) and sfermions. Note this is in addition to all of the usual interaction diagrams in the SM.

3.4.1 Mass hierarchy

Stop sector

Since the $\hat{P}^2$ operator commutes with the SUSY generators as previously shown, the masses of the SUSY partners are required to be identical to their SM counterparts (and by extension the Yukawa coupling). With equal couplings and masses as their SM particles, such a theory should be seen as an enhancement of the standard model cross section in processes such as $t\bar{t}$. Since the $t\bar{t}$ cross section has been measured to a high degree of precision to be consistent with the standard model [18] (figure 3.8). These enhancements are not observed on the same mass scales as the standard model particles. Taking a lead from the GWS model, and the fact that a conventional fermion/scalar mass term is not manifestly gauge invariant nor invariant under the supersymmetry transformations, we can give masses to the super-partners through supersymmetry breaking of the extended Higgs/Higgsino sector. Thus the masses of the super-partners are now no longer mass degenerate with their associated standard model particles.

Following the discussion in [7], the superpotential corresponding to the Higgs sector in the minimal supersymmetric standard model is the following:

$$W_{\text{Higgs}} = \bar{u}Y_uQH_u - \bar{d}Y_dQH_d - \bar{e}Y_eLH_d + \mu H_uH_d. \quad (3.49)$$

where the Yukawa couplings denote all of those for each generation of quark and lepton. Applying the approximation that we only look at the third generation of quarks and leptons

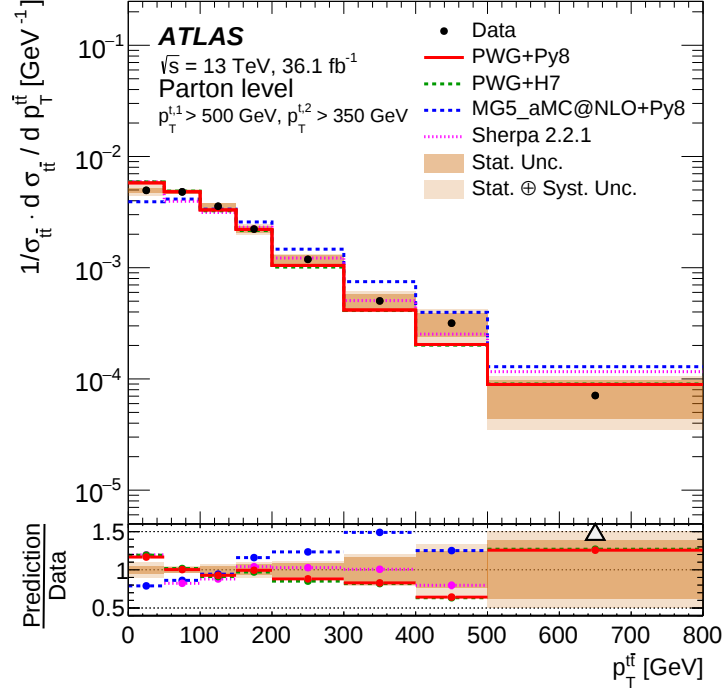


Figure 3.8: Differential cross section observed as a function of the $t\bar{t}$ p_T by ATLAS. Taken from [18].

(since these are in general much heavier than the lower generations). Thus the terms of the superpotential under this approximation look like:

$$Y_u \approx \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & y_t \end{pmatrix} \quad Y_d \approx \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & y_b \end{pmatrix} \quad Y_e \approx \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & y_\tau \end{pmatrix}$$

$$W_{\text{Higgs}} = y_t(\bar{t}tH_u^0 - \bar{t}bH_u^+) - y_b(\bar{b}tH_d^- - \bar{b}bH_d^0) - y_\tau(\bar{\tau}\nu_\tau H_d^- - \bar{\tau}\tau H_d^0) + \mu(H_u^+ H_d^- - H_d^0 H_d^0). \quad (3.50)$$

In addition the Higgs mass terms [7] are:

$$\begin{aligned} \mathcal{L}_{\mu, \text{quadratic terms}} &= -\mu(\tilde{H}_u^+ \tilde{H}_d^- - \tilde{H}_u^0 \tilde{H}_d^0) + h.c. \\ &\quad - |\mu|^2(|H_u^0|^2 + |H_u^+|^2 + |H_d^0|^2 + |H_d^-|^2). \end{aligned} \quad (3.51)$$

The first term can be seen as a mass term for the Higgsino, and the second can be seen as mass terms for the real Higgs doublets. Thus the MSSM soft SUSY breaking terms will have the form [6]:

$$\begin{aligned} \mathcal{L}_{\text{soft, MSSM}} &= -\frac{1}{2} \left(M_{\tilde{G}} \tilde{G}^\mu \tilde{G}_\mu + M_{\tilde{W}} \tilde{W}^\mu \tilde{W}_\mu + M_{\tilde{B}} \tilde{B}^\mu \tilde{B}_\mu \right) + h.c. \\ &\quad - \left(\tilde{u} A_u \tilde{Q} H_u - \tilde{d} A_d \tilde{Q} H_d - \tilde{e} A_e \tilde{L} H_d \right) + h.c. \\ &\quad - \tilde{Q}^* V_Q \tilde{Q} - \tilde{L}^* V_L \tilde{L} - \tilde{u}^* V_{\tilde{u}} \tilde{u} - \tilde{d}^* V_{\tilde{d}} \tilde{d} - \tilde{e}^* V_{\tilde{e}} \tilde{e} \\ &\quad - m_{H_u}^2 H_u^* H_u - m_{H_d}^2 H_d^* H_d - (b H_u H_d + h.c.). \end{aligned} \quad (3.52)$$

where the matrix V is often referred in several sources [6, 7] as the mass matrix $\mathbf{m}^2$, while the A matrices have dimension mass and correspond to the Yukawa matrices of the superpotential (suppressing the indices over gauge states). In the same vein as the GWS model, a particular choice of bases for $(Q, L, \bar{u}, \bar{e})$ leads to these being associated with the usual mass terms for fermions. The sparticle content of the MSSM is listed in table 3.3.

Label	Spin	P_R	Gauge Eigenstates	Mass Eigenstates
Higgs bosons	0	1	$H_u^0, H_d^0, H_u^+, H_d^-$	$(H^0, A^0, H^\pm) + \text{SM } h^0$
Squarks	0	-1	$\tilde{u}_L, \tilde{u}_R, \tilde{d}_L, \tilde{d}_R$ $\tilde{c}_L, \tilde{c}_R, \tilde{s}_L, \tilde{s}_R$ $\tilde{t}_L, \tilde{t}_R, \tilde{b}_L, \tilde{b}_R$	(same labels) (same labels) $\tilde{t}_1, \tilde{t}_2, \tilde{b}_1, \tilde{b}_2$
Squarks	0	-1	$\tilde{e}_L, \tilde{e}_R, \tilde{\nu}_e,$ $\tilde{\mu}_L, \tilde{\mu}_R, \tilde{\nu}_\mu,$ $\tilde{\tau}_L, \tilde{\tau}_R, \tilde{\nu}_\tau$	(same labels) (same labels) $\tilde{\tau}_1, \tilde{\tau}_2, \tilde{\nu}_\tau$
Neutralinos	1/2	-1	$\tilde{B}^0, \tilde{W}^0, \tilde{H}_u^0, \tilde{H}_d^0$	$\tilde{\chi}_1^0, \tilde{\chi}_2^0, \tilde{\chi}_3^0, \tilde{\chi}_4^0$
Chargions	1/2	-1	$\tilde{W}^\pm, \tilde{H}_u^\pm, \tilde{H}_d^\pm$	$\tilde{\chi}_1^\pm, \tilde{\chi}_2^\pm$
Gluino	1/2	-1	$\tilde{g}$	(same)

Table 3.3: Sparticle content of the MSSM in addition to the standard model [6], assuming minimal mixing of the first and second generation quarks and leptons, in the context of gauge and mass eigenstates. This thesis particularly looks at the existence of the $\tilde{t}_1$ and $\tilde{\chi}_1^0$ states, which are linear combinations of the sfermion partners to the chiral states of the top quarks or the bosino partners to the B and neutral Higgsinos.

Since at time of writing, none of the constituent particles of the MSSM have been readily observed, the form of the mixing matrices has not been explicitly determined (at least beyond a minimal set of parametrisations) [6]. For the mass and chiral states of the stop $\tilde{t}$ (scalar partner to the top quark) a matrix V of:

$$\begin{pmatrix} \tilde{t}_1 \\ \tilde{t}_2 \end{pmatrix} = V_{\tilde{t}} \begin{pmatrix} \tilde{t}_L \\ \tilde{t}_R \end{pmatrix}, \quad V_{\tilde{t}} = \begin{pmatrix} c_{\tilde{t}} & -s_{\tilde{t}}^* \\ s_{\tilde{t}} & c_{\tilde{t}}^* \end{pmatrix} \quad (3.53)$$

The only requirement on this matrix V that can be applied without knowing the explicit values of its entries is that the matrix must be unitary (and hence diagonalisable) into the separate eigenstates. The class of models that this thesis will cover are primarily those where $m(\tilde{t}_1) < m(\tilde{t}_2)$ (since we don't actually know the explicit values of the mass matrix entries, it is possible to re-arrange the labels such that this is always the case, neglecting the mass degenerate case). Similar logic can be used to determine the masses of the sbottom and staus, assuming that the mixing of the third generation with the other generation sparticles is negligible.

Neutralinos mass hierarchy

Since as per the stop sector, and the GWS, the mass eigenstates do not correspond to the mass eigenstates of the respective particles in the theory, a neutralino is formed, which as mentioned earlier is a mass basis combination of the $\Phi = (\tilde{B}, \tilde{W}^0, \tilde{H}_d^0, \tilde{H}_u^0)$ gauge eigenstates, into four mass eigenstates (neutralinos).

The mass term of the neutralinos in the lagrangian under soft SUSY breaking take the form [6]:

$$\mathcal{L}_{\tilde{\chi}-mass} = -\frac{1}{2}\Phi^T \mathbf{M}_{\tilde{\chi}} \Phi + \text{complex conjugate}, \quad (3.54)$$

$$\mathbf{M}_{\tilde{\chi}} := \begin{pmatrix} M_1 & 0 & -c_\beta s_W m_Z & s_\beta s_W m_Z \\ 0 & M_2 & c_\beta c_W m_Z & -s_\beta c_W m_Z \\ -c_\beta s_W m_Z & c_\beta c_W m_Z & 0 & -\mu \\ s_\beta s_W m_Z & -s_\beta c_W m_Z & -\mu & 0 \end{pmatrix}. \quad (3.55)$$

The shorthands $\{s, c\}_\theta = \{\sin \theta, \cos \theta\}$ are used, and the W denotes the Weinberg angle θ_W from the SM. The terms M_1 and M_2 come directly from the soft Lagrangian already specified earlier, and the term $-\mu$ comes from the Higgsino mass terms. This complex mass matrix in the basis of gauge eigenstates can be diagonalised to produce a diagonal hierarchy of particles $\tilde{\chi}^0$, with masses:

$$\mathcal{D}^{-1} \mathbf{M}_{\tilde{\chi}} \mathcal{D} = \begin{pmatrix} m_{\tilde{\chi}_1^0} & 0 & 0 & 0 \\ 0 & m_{\tilde{\chi}_2^0} & 0 & 0 \\ 0 & 0 & m_{\tilde{\chi}_3^0} & 0 \\ 0 & 0 & 0 & m_{\tilde{\chi}_4^0} \end{pmatrix} \quad (3.56)$$

A similar argument can be made for the two charged gauginos, and their respective chargino mass eigenstates, but will not be covered here. An interested reader is directed to [6]

Mass Hierarchy

The . Thus, the possible choice of spectra is not constrained, with exception a small number of observations (as per [6]):

1. The lightest neutralino (denoted as $\tilde{\chi}_1^0$ by mass ordering convention) is the LSP, unless a gravitino has a smaller mass, assuming that R-parity is conserved. R-parity requires that the lightest SUSY particle is stable. The neutralino can either be a perfect admixture, a “bino-like” or “wino-like”, which can have some impact on the mass of the lightest chargino states in addition.
2. Gluinos will be heavier than neutralinos and charginos.
3. Squarks of the first and second generations are nearly degenerate in both gauge and mass eigenstates (similar to the approximate degeneracy of u,d,s in the SM) and are expected to be heavier than sleptons, with mass splitting depending on the chiral states.
4. The lightest stop $\tilde{t}_1$ and lightest sbottom $\tilde{b}_1$ are the lightest, resulting from the mixing introduced via the mixing matrices. This last point is the principal motivator for the study in this thesis.

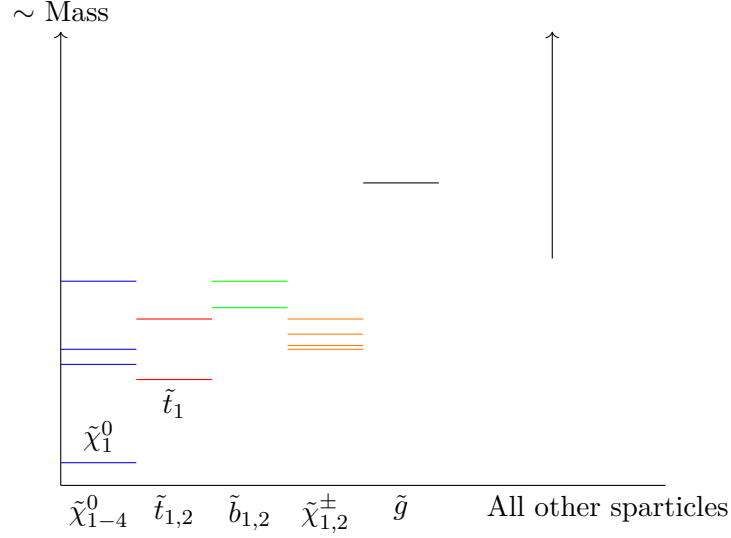


Figure 3.9: A possible spectrum of squarks and neutralinos given the mixing matrices and mass terms. This diagram is merely a schematic of a possible SUSY scenario of which the $\tilde{t}_1$ would be more massive than the LSP $\tilde{\chi}_1^0$, hence the magnitudes of mass splittings/orderings etc in this diagram may not actually reflect reality. The lightest stop and neutralino are highlighted for clarity.

3.4.2 R-Parity Violation

This thesis has principally assumed only decay chains which preserve R-parity, which forbids proton decay. However, this assumption is not manifestly required in all SUSY scenarios, and considerable experimental interest covers the additional R-parity violating (RPV) decays of the MSSM. Removing this restriction allows for the sparticles to decay directly into SM particles, although which processes are allowed to occur (either Baryon/Lepton number violation) may vary on the exact model of interest. In particular, this also implies that the lightest sparticle, $\tilde{\chi}_1^0$ is unstable, and can decay in various ways, as noted from [6]. Since this class of decay is beyond the scope of this thesis, additional details can be found in [6].

3.4.3 Simplified Models

Since the mass spectrum of the SUSY sector is not predicted beyond some minimal result, and the fact that only the LSP is expected to be stable, a heavier particle is expected to follow a decay chain (emission of standard model particles and/or SUSY pairs along the way). The mass spectrum is also dependant on some free parameter choice in the minimal supersymmetric standard model, which are inserted into the model by construction (and are not tightly constrained). In the minimal supersymmetric standard model, this set of parameters corresponds to 120 free parameters, which are all freely floating and not otherwise defined (unlike most of the standard model measurements, which could be constrained independently from either electrodynamics, nuclear physics or deep inelastic scattering, as well as collider experiments).

Since the > 120 free parameters are mostly free floating under minimal constraints, we apply a common restriction known as the phenomenological MSSM (pMSSM) [19], which is motivated by being CP conserving (i.e. SUSY introduces no new CP phases except those already present in the SM) and using “minimal flavour violation” (MFV). We also use the assumption from the previous subsections, namely the degeneracy of the first two generations sfermions. The remaining parameters are as follows:

1. Gaugino masses (3 parameters).
2. Higgsino mixing parameter μ .
3. Ratio of Higgs VEVs $\tan \beta$.
4. Mass of the pseudoscalar neutral Higgs “heavy Higgs”: m_A .
5. Masses of the sfermions (5 for the first two generations, 5 for the third generation).
6. Yukawa couplings for the third generation sfermions (3 since the yukawa coupling is not dependent on chirality).
7. Neutralino $\tilde{\chi}_1^0$ is the LSP, arising from the R-parity invariance

Without a defined mass spectrum in the pMSSM, the branching ratios for different become quite complex and mass spectrum dependant. This mass spectrum dependancy also translates to observable final states, since visible (and invisible) SM particle emission can occur with differing branching ratios. Without a strong constraint, it is impossible to determine the existence of any process, thus a simplified model is required.

The simplified models are in no way true reflections of reality (assuming the MSSM is valid), but reflect the case where a given mass splitting occurs and the initial state particle decays along its specified decay path with a fixed defined branching ratio (in the cases this thesis will cover, this is 100%). A simplified model reduces the number of free parameters in the pMSSM to two, namely in the cases of interest here, the mass of the LSP ($\tilde{\chi}_1^0$) and the mass splitting between it and the stop ($\Delta M(\tilde{t}_1, \tilde{\chi}_1^0)$). With these two parameters set as points on a grid, it is then possible to generate simulated events, and apply exclusion limits based on both the available data and the given signal model.

A strong caveat of this simplified model approach is that in all of the restrictions from the MSSM to the simplified are based on assumptions, of which the validity of those is not known. Thus given an exclusion limit on a particular simplified model does not imply an exclusion on a different class of simplified model - this is why exclusion limits are referred to as “model dependent exclusion limits”. The only “model independent” sections of the analysis are known as discovery fits, validating the existence of any excess of data over the SM background, although the regions used therein are designed using simplified models as a guide. In addition to the simplified model searches on ATLAS, such as those listed in this thesis, automated scans on this parameter space have been undertaken (see [20]).

Specific signal models: $\tilde{t} \rightarrow t + \tilde{\chi}^0$ decays

In this section the simplified model of pair-produced $\tilde{t}_1 \rightarrow t + \tilde{\chi}^0$ is considered, which is a restricted section of the PMSSM [2, 19], and the stop is assumed to decay in this mode with a 100% branching ratio. This particular decay mode is R-parity conserving (RPC), and the top decays either semi-leptonically or hadronically through the emission of a W boson. The search strategy used in this thesis covers the case where the search phase space primarily covers the case where the top pairs decay hadronically, and other analyses cover the different decay modes of the top pairs. The alternate top pair decay modes require somewhat distinct analysis strategies, which are discussed further in [21–25]). The initial decay (where $\Delta(m_{\tilde{t}}, m_{\tilde{\chi}^0}) > m_t$) is referred to as a two body decay, due to the number of final state particles in the initial decay (not any subsequent top quark decays).

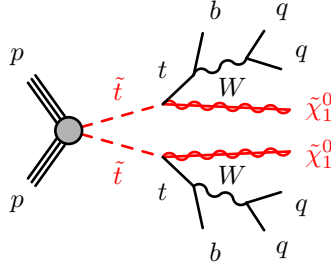


Figure 3.10: Representative diagram of the two body stop decay mode.

In addition to the two body decay mode, there is also a decay mode via 3-body decay, in the case where $\Delta(m_{\tilde{t}}, m_{\tilde{\chi}^0}) \leq m_t$, namely that there is only an on-shell b quark and an on-shell W produced, again assumed to decay hadronically as per the analysis strategy.

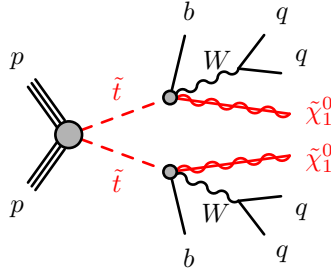


Figure 3.11: Representative diagram of the three body stop decay mode.

In the particular special case of small mass splittings between the stop and neutralino, particularly below the threshold of an on-shell W being produced from the decay, a four-body decay mode is postulated, namely that the stop directly decays into a b-quark, two quarks and the neutralino.

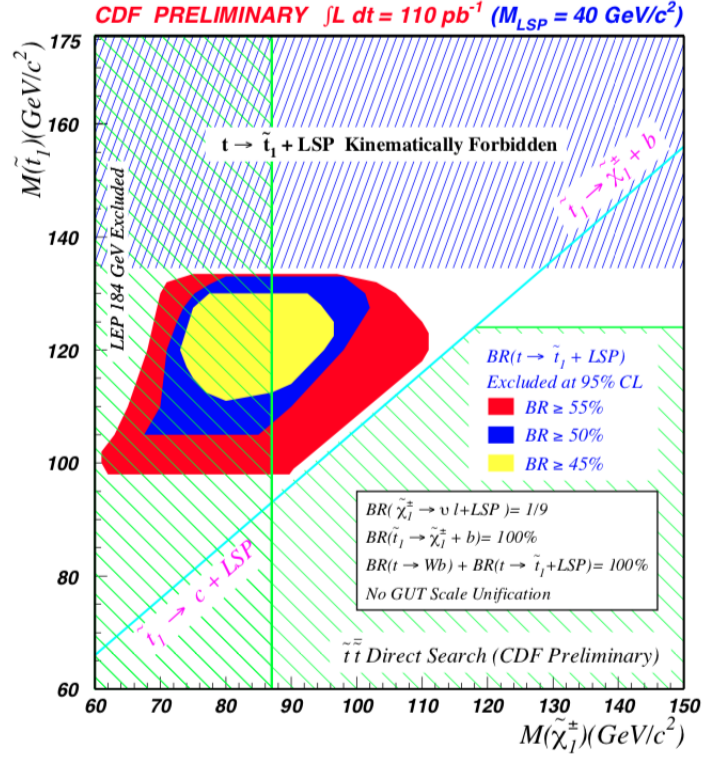


Figure 3.13: CDF exclusion limit [26] in the $m_{\tilde{t}_1}, m_{\tilde{\chi}_1^0}$ plane, for simplified models with $m_{\tilde{\chi}_1^0} = 40 \text{ GeV}$.

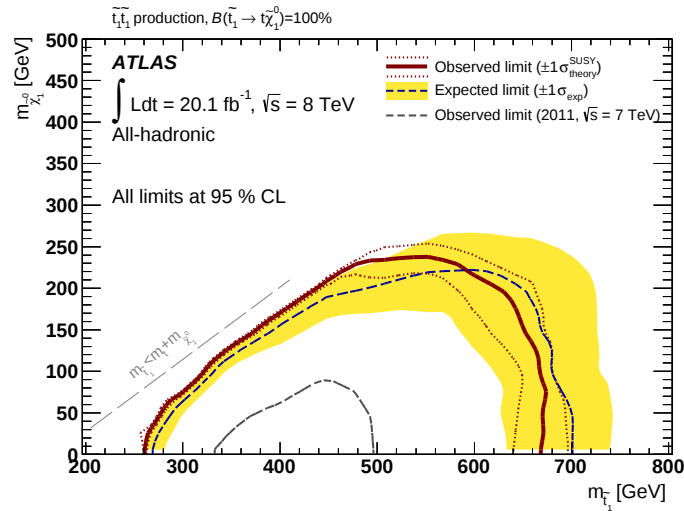


Figure 3.14: ATLAS run 1 exclusion limit [27] in the $m_{\tilde{t}_1}, m_{\tilde{\chi}_1^0}$ plane, for simplified models with 100% branching ratio.

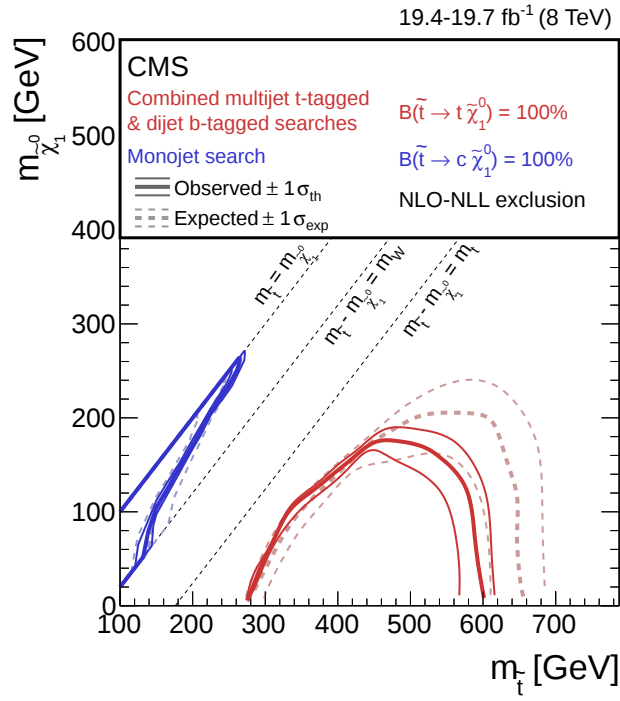


Figure 3.15: CMS [28] run 1 exclusion limit in the $m_{\tilde{t}}, m_{\tilde{\chi}_1^0}$ plane, for simplified models with 100% branching ratio.

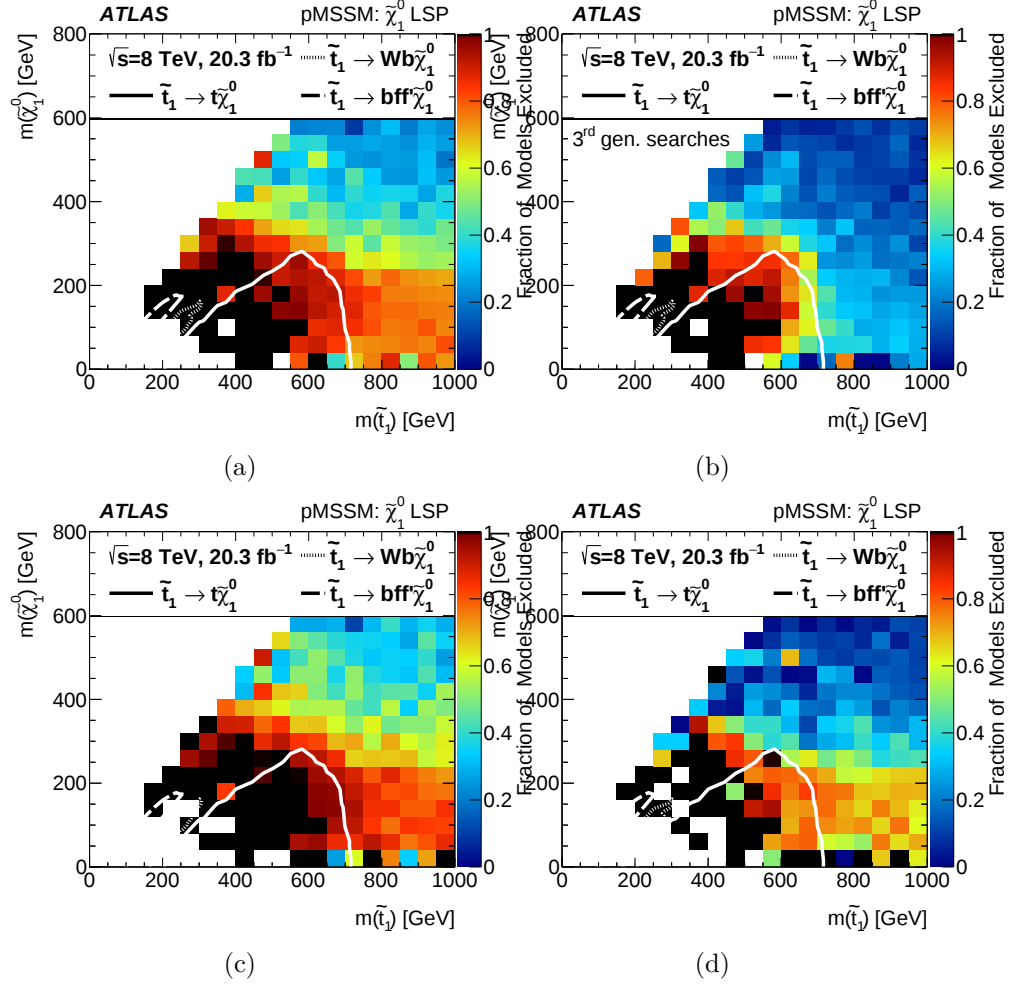


Figure 3.16: $m_{\tilde{t}}, m_{\tilde{\chi}_1^0}$ exclusion limits from the ATLAS PMSSM scan[20]. Figures 3.16a and 3.16b denote the exclusion limits for PMSSM models in either all searches or third-generation specific searches. Figures 3.16c and 3.16d denote exclusion limits on models with primarily left or right handed $\tilde{t}_1$. All of these plots were taken from [20].

ATLAS detector and infrastructure as part of the LHC

This chapter outlines the specification of the Large Hadron Collider (LHC) and the ATLAS detector, from which the results based in this thesis were obtained. The purpose of a proton-proton collider is to produce in a controlled manner, high energy particles arising from the collision interactions, and their associated decay products, which is at the time of writing, at the forefront of the high energy frontier. In addition to the design specification of the LHC, the operations pipeline will be discussed, both in the case of data collection from the LHC to simulation and the overall analysis workflow from input to final result.

4.1 LHC overview

This section covers the design of the LHC[\[29\]](#), which was installed from 2000-2010 in the same tunnel as had previously been constructed for the Large Electron Positron Collider (LEP)[\[29\]](#). The LHC is designed principally for proton-proton (p-p) and Heavy ion (Lead ions, Lead ion+proton, Xenon-Xenon) collisions, using oscillating and varying electromagnetic fields to both accelerate and direct the beams of protons/ions, and is a circular collider with several main components:

- Linear accelerators for initial beam injection up to an energy of 50 MeV.
- Proton Synchotron (PS) and boosters (628m circumference) to increase the proton beam energy to 1.4 GeV.
- Super proton Synchotron (SPS) - 6.9km circumference- as a further booster step, accelerating to a beam energy of 450 GeV.
- LHC main ring, circular ring approximately 27km long, consisting of two counter-propagating beams which are allowed to interact with each other only at specific points and under specific control conditions. This accelerates the proton beams up to the final beam energy of 6.5 TeV per beam.

A notable feature of the LHC is that the two counter-propagating beams are constrained within the same beam yoke.

Since there is a substantial demand to study different types of high energy processes using p-p or Pb-Pb collisions, the LHC main ring consists of four principal experiments, A Toroidal LHC ApparatuS (ATLAS) [30] and the Compact Muon Solenoid (CMS)[31] as general purpose particle detectors, LHCb (LHC beauty)[32] covering flavour physics and rare decays, and A Large Ion Collider Experiment (ALICE)[33] principally used for Heavy Ion Physics. These experiments are placed at the crossing points, where beams of protons or heavy ions can be brought into collision. The two general purpose detectors are principally designed to cross-validate each other, and the analyses, organisations and analysis techniques used are independently undertaken.

The full configuration of the LHC and its associated components can be found in figure 4.1.

The LHC for p-p collision operation was designed to operate at a collision centre of mass energy ($\sqrt{s}$) of 14TeV[29], and has steadily increase in beam collision energy, from 7/8 TeV in LHC Run 1 (2010-2014), to 13 TeV of LHC Run 2 (2015-2018). This outstrips the capabilities of any other (at time of writing) collider (with the next nearest being the now discontinued Tevatron experiment, at Fermilab, IL, USA with $\sqrt{s}$ of 1.96 TeV[34]). Since there are no alternative collider experiments which cover the same $\sqrt{s}$ range that the LHC obtains, the need for cross-validation of independently operated experiments is considerable.

In addition to the pp collision schedule, heavy ion collisions (principally Pb ions, but p-Pb and Xe ion collisions have been performed) have also been undertaken as part of the LHC run programme, with Run 2 centre of mass energy (per nucleon) being on the order of 5TeV. This thesis will, however, not pertain to heavy ion collisions.

LHC luminosity and pile-up

Approximately 10^{11} protons are accelerated in radio frequency electromagnetic field “buckets” referred to as bunches, with a spacing of approximately 25ns between bunches [29]. The proton beams are controlled by two sets of dipole magnets, consisting of 1232 superconducting 8.33T dipole superconductor magnets cooled to 1.9K, all within a single iron yoke and cryostat. The acceleration is undertaken at a single point along the ring using a radio frequency (RF) cavity, increasing from the injection energy of 450GeV to the final beam energy of 6.5TeV. The full process from injection to “Good for physics” collisions takes about 1hr (This includes detector cooling, calibration, supplying the relevant voltages to the detector components), and consists of a finite number of input bunches, in a fill which will normally last for around 6-8hrs.

At LHC interaction point 1 (IP1 : ATLAS) based on the CERN Meyrin site, the beams are allowed to intersect with an incidence of $285\mu\text{rad}$. The produced number of events for a process at the interaction point then relates to the luminosity and cross section using the formula:

$$N = L \times \sigma. \quad (4.1)$$

The machine instantaneous luminosity, hence can be estimated, assuming the beam profile

CERN's Accelerator Complex

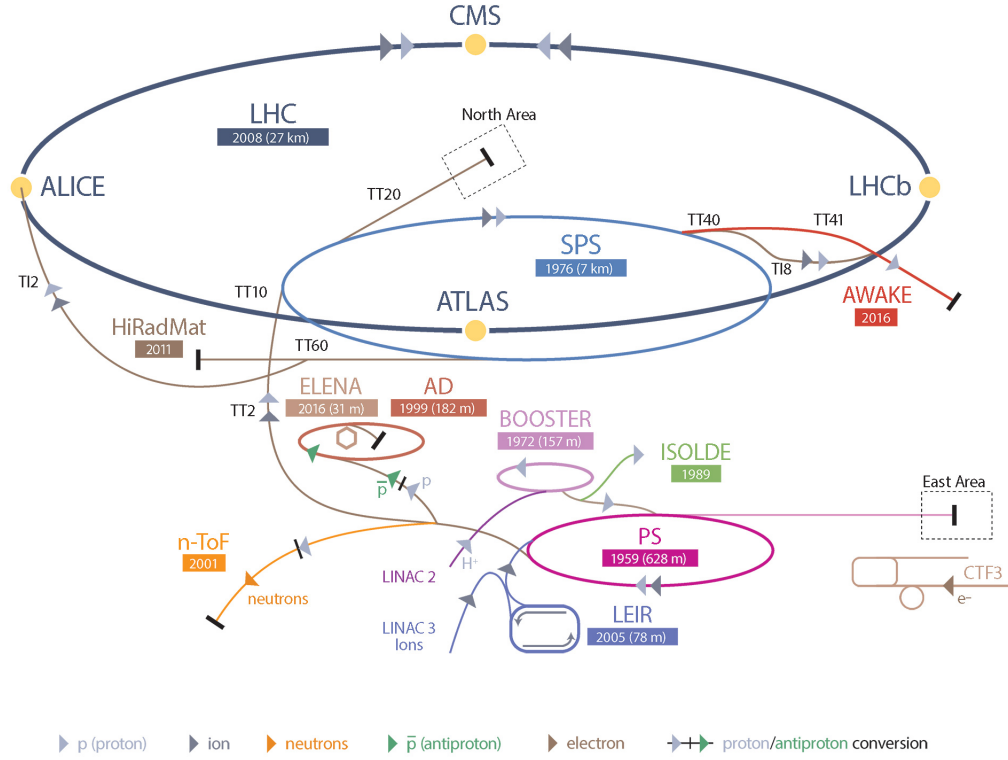


Figure 4.1: LHC with all of its sub-component accelerator complexes, and smaller experiments not covered in this thesis. Taken from [35]

has a Gaussian dispersion, and takes the form [30]:

$$L = \frac{N_b^2 n_b f_{rev} \gamma_{rel}}{4\pi \epsilon_n \beta^*} \mathcal{F} \quad (4.2)$$

with N_b defined as the number of protons per bunch, n_b the number of bunches, f_{rev} the revolution frequency, ϵ_n the transverse emittance (width of beam), and the value of the amplitude function (β) at the interaction point $\beta^* = \beta(IP1)$. The $\mathcal{F}$ term relates to the detector geometry and the crossing angle. Naturally, due to effects such as either beam losses or the reduced number of protons per bunch with subsequent collisions, the luminosity recorded on ATLAS and other LHC experiments will decrease as a function of time (and may vary even between beam injection periods for a variety of reasons). Thus it is preferable to consider the *integrated luminosity*, which is simply the time-integral of the instantaneous luminosity defined as above:

$$\mathcal{L} = \int_{start}^{end} L dt \quad (4.3)$$

This is normally recorded in time-intervals on the order of 1-2minutes where the luminosity is assumed to be constant (referred to as Lumi-blocks) at the LHC, and is compared between each of the relevant LHC experiments. Figure 4.2(a) refers to the integrated luminosity over the entirety of the LHC Run 2: 2015-18 (with a smaller subset being recorded and validated

for data quality). In this period, the ATLAS experiment recorded 139 fb^{-1} of “Good for physics” data within Run 2.

In addition, another metric of importance is the average number of collisions per bunch crossing or *pileup*, commonly denoted as $\langle\mu\rangle$ (or without the angle brackets). This was on average around 34 for the LHC data periods 2015-2018, but can vary depending on a variety of factors, such as beam conditions, detector conditions etc. A particular point that can impact pile-up is the coordination of bunches in the LHC into trains of 25ns space bunches, such as the *8b4e* scheme [36], which means in short a train of 12 RF bunches spaced by 25ns, of which 8 are considered filled with protons, and four are empty (to allow spacing between the bunch trains), however this scheme is not the only such available.

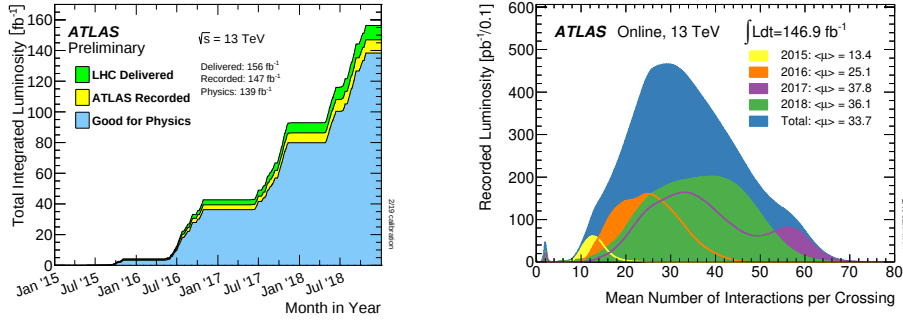


Figure 4.2: Detector operations plots from LHC Run 2 concerning luminosity (left) and pileup (right). [37]

The design specification of ATLAS beams [29] is summarised in Table 4.1.

		Injection (SPS)	Collision
Beam information			
Proton energy	[GeV]	450	7000
Relativistic γ		479.6	7461
Number of particles per bunch		1.15×10^{11}	
Number of bunches		2808	
Stored energy per beam	[MJ]	23.3	362
Peak luminosity conditions			
Avg. Bunch length (RMS)	[cm]	11.24	7.55
RMS beam size (IP1)	[μm]	375.2	16.7
Geometric luminosity reduction factor $\mathcal{F}$		-	0.836
Peak luminosity (IP1)	[$\text{cm}^{-2}\text{s}^{-1}$]	-	1.0×10^{34}
Peak luminosity/bunch crossing (IP1)	[$\text{cm}^{-2}\text{s}^{-1}$]	-	3.56×10^{30}

Table 4.1: Summary of the main LHC design parameters as per [29].

4.2 The ATLAS detector

This section will be based on material presented in [30], and summarise the key components of the ATLAS (**A** **T**oroidal **L**HC **A**pparatu**S**) detector, an experiment with full radial reconstruction capability, tracking performance up to an absolute pseudo-rapidity ($|\eta| < 2.5$) (calorimetry up to $|\eta| < 2.7$) and muon reconstruction for near-axis objects up to $|\eta| < 4.5$. A schematic diagram of the experiment is shown in figure 4.3. Key components of the detector include tracking, calorimetry, toroidal magnetic field and muon spectrometry. As per [30], the main aims of the ATLAS detector in design were:

- Use fast, radiation-hard electronics able to cope with high particle fluxes, and provide good granularity of reconstruction
- Full reconstruction performance across full range of polar angles, and maximal coverage in azimuthal angle.
- Efficient charged particle momentum resolution and reconstruction in the inner tracker, in addition to vertex detectors able to reconstruct τ -leptons and b-jets.
- Efficient electromagnetic and hadronic calorimetry for electron, photon and jet reconstruction.
- Efficient muon identification, charge evaluation and high muon momentum resolution over a large momentum range.
- Efficient triggering on physics objects of interest, producing datasets with sufficient background rejection.

These design goals can be best summarised in terms of the specification listed in table 4.2).^{*} The data recording rate from the proton-proton interaction point is expected to be of the

Component	Specified 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 Barrel (Bulk+Endcap)	$\sigma(E)/E = 50\%/\sqrt{E} \oplus 3\%$	± 3.2	± 3.2
Forward	$\sigma(E)/E = 100\%/\sqrt{E} \oplus 10\%$	$\in [3.1, 4.9]$	$\in [3.1, 4.9]$
Muon Spectrometer	$\sigma(p_T)/p_T = 10\%$ at $p_T = 1$ TeV	± 2.7	± 2.4

Table 4.2: ATLAS design specification and performance goals for the respective components, as listed in [30]. The resolution of the muon spectrometer is independent of the inner detector and calorimeters.

order of 1 GHz, requiring a trigger system to reduce the output data rate to approximately 1kHz, reducing the data by a factor of the order of 5×10^6 . This consists of a two tiered

^{*}ATLAS uses a modified version of the spherical polar coordinate system, where the azimuthal angle θ is replaced with the pseudo-rapidity $\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right)$, chosen such that $\Delta\eta$ for any two objects is Lorentz invariant under boosts along the Z axis (defined to be the beam axis).

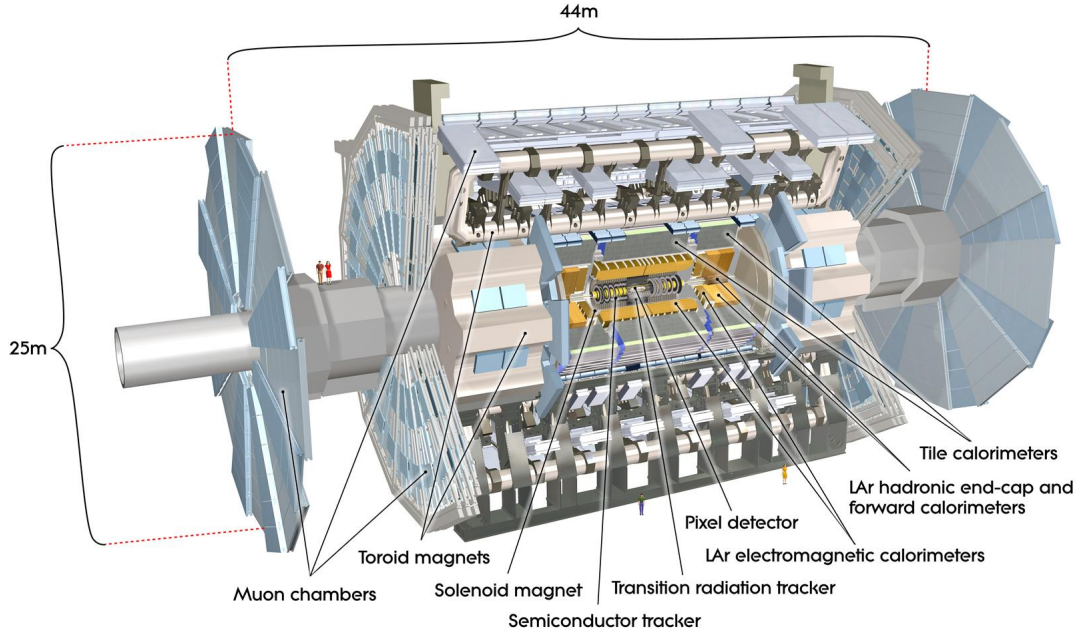


Figure 4.3: Schematic diagram of the ATLAS detector, listing the key components of the ATLAS detector and the overall dimensions.[30]

trigger system, which will be covered more in 4.4. Section 4.3 will cover the individual detector component hardware, while section 4.4 will cover the triggering and detector readout system. Any computing level processes (e.g. event level reconstruction) will be covered in section 4.6, since these primarily are offline (i.e. not directly covered as instantaneous detector readout).

4.3 Hardware

4.3.1 The ATLAS tracking system

The tracking system consists of several components in both the bulk barrel and the endcap, but no such system is present for the forward regions, and targets the reconstruction of charged particle tracks and where possible, secondary vertex finding. This tracking system consists of the semiconductor tracker (SCT), the transition radiation tracker (TRT), pixel detectors, and for the data periods 2015 onward (LHC run 2) the insertable B-layer (IBL)[38] was included to improve tracking performance. The entire inner detector is contained inside a 2T toroidal magnetic field. The precision tracking covers the range $|\eta| < 2.5$, with full coverage in ϕ . Each system requires different cooling systems, namely to mitigate the effect of radiation damage and avoid condensation on the interior, with the pixel detector and SCT embedded in dry N_2 [30] at around -7°C , while the TRT is maintained at room temperature in a bath of CO_2 .

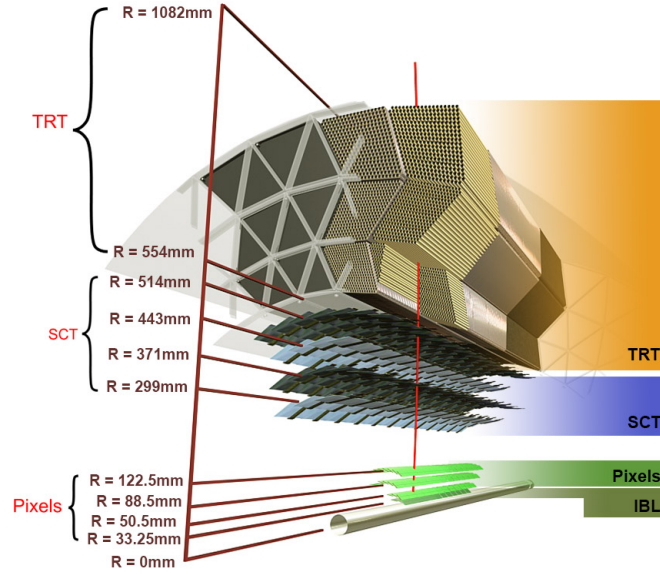


Figure 4.4: Schematic of the tracking system [39] as per the design specification [30]. The Insertable B-Layer (IBL) was inserted before the start of the 2015-2018 LHC data taking period as an extension beyond the original design specification.

The SCT uses silicon strips which are 6.4cm long and connected to readout channels. In the SCT these are laid out in stereo pairs of microstrip layers. In the pixel detector, for track and vertex finding (inferred by precision tracking). The TRT comprises gaseous straw tubes interleaved with polypropylene fibres (barrel) or foil (endcap), which are used to produce the transition radiation. A charged particle track with $p_T > 500$ MeV [30] through $|\eta| < 2.0$ will traverse through at least 36 straws, except in the region $|\eta| \in [0.8, 1.0]$, whereby due to the interface of the barrel and endcap TRT components, a track may pass through a minimum of 22 straws.

4.3.2 Magnetic fields

ATLAS consists of a system of four large superconducting magnets, consisting of a central solenoid, being found between the tracking and calorimetry components of the detector, and toroidal magnets consisting of one barrel toroid and two endcap toroids. This magnetic field system is important for using the bending of a charged particle using the Lorentz force to estimate the momentum of a charged particle track. The inner solenoid produces a uniform field of 2T covering the region of the tracking volume, while the toroidal magnet system is principally used to aid in muon reconstruction, and has a field of 0.5-1T for central/endcap regions. Principally the magnets consist of [30] NbTi conductor coils, cooled to 4.5K (main solenoid) and Nb/Ti/Cu conductor coils cooled to 4.6K in the case of the toroidal system. Figure 4.5 depicts the configuration.

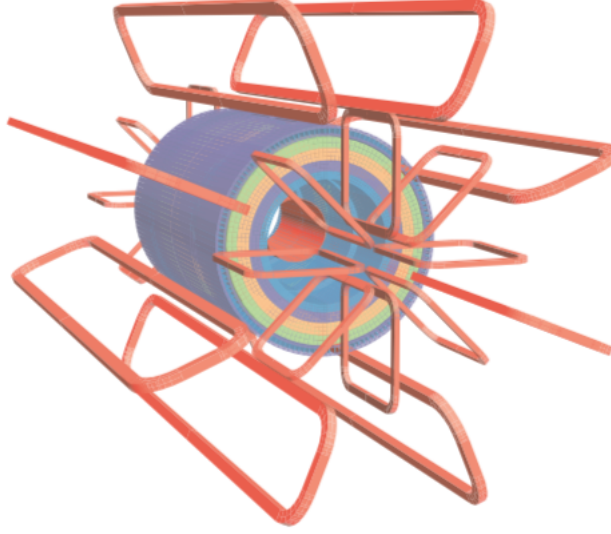


Figure 4.5: Magnet configuration for ATLAS[30], including the central solenoid inside the calorimeter volume. The central core denotes the central solenoid, while the large red magnets denote the toroidal muon system. The smaller toroidal systems related to the endcaps.

4.3.3 ATLAS Calorimetry

The calorimeter can be sub-divided into two constituent parts, the hadronic calorimeter (HCAL) and the electromagnetic calorimeter (ECAL), resolving both neutral and charged particle deposits respectively. In the barrel region, the ECAL is located inside a ring encapsulated by the hadronic calorimeter components. In the next two subsections, the configuration of these components will be discussed. Figure 4.6 depicts the calorimeter setup.

4.3.3.1 Electromagnetic Calorimeter (ECAL) for ATLAS

The ECAL [30, 40] is separated into two parts, as seen in fig 4.6, namely the barrel part ($|\eta| < 1.475$) and the end cap components ($|\eta| \in [1.375, 3.2]$). They consist of liquid argon calorimeters cooled to $59 \pm 1\text{K}$, consisting of both lead (an absorber required for shower production) and high-purity liquid argon (as the active medium), arranged in an accordion geometry, best visualised in 4.7. The calorimeter [40] consists of 3 layers of cells of size $\delta\eta \times \delta\phi = 0.025 \times 0.0245$ in layer 1, layer 2 $\delta\eta \times \delta\phi = 0.025 \times 0.0245$ and layer 3 $\delta\eta \times \delta\phi = 0.05 \times 0.0245$, with each layer covering a different range of η . The accordion geometry is chosen to minimise cracks between calorimeter cells, and keep the effective radiation length sufficiently small as to minimise muon interaction with the calorimeter. The entire configuration is about 47cm in thickness, surrounding the central barrel region, or in total corresponding to a radiation length of about $22 - 33X_0$ [40] depending on the angle of incidence (η).

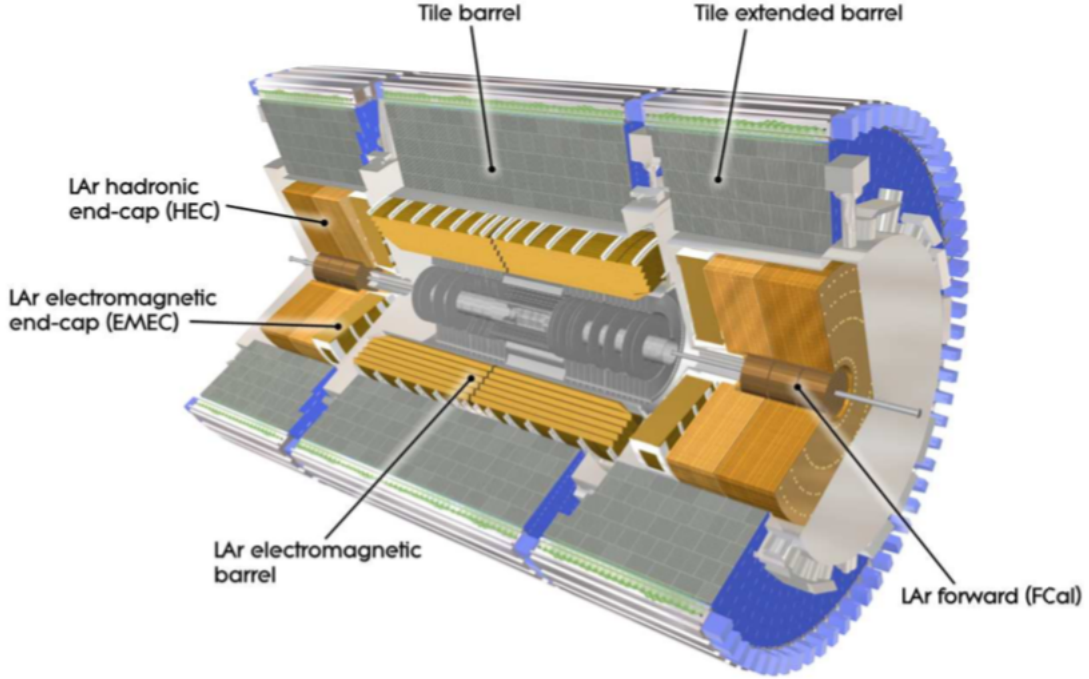


Figure 4.6: Schematic cut-away diagram of the ATLAS calorimeter system, presented in [30].

4.3.3.2 Hadronic Calorimeter (HCAL)

The HCAL consists of three parts, namely the liquid-argon forward calorimeter (FCAL), the hadronic endcap calorimeter (HEC) and the tile calorimeter (barrel). This calorimeter targets the deposition of energy originating from neutral hadrons that were not stopped in the electromagnetic calorimeter. In the tile calorimeter [41], the active absorber of neutral hadrons is steel, and tile shaped cells are embedded in the structure, which are filled with liquid argon. Photomultiplier tubes and readout electronics are attached to the ends with fibres connecting them to the scintillator chambers. This entire setup can be summarised in diagram 4.8. The hadronic endcap uses copper/liquid argon design in flat plates (see figure 4.9).

Forward Calorimeter (FCAL)

The forward calorimeter is not directly used in the reconstruction of jets or leptons due to kinematic selections on η for these objects. However, the forward calorimeter is particularly important in the measurement of the E_T^{miss} , with the inclusion of energy deposits close to the beamline.

The forward calorimeter consists of sets of three calorimeter modules (one EM calorimeter and two hadronic modules, using copper (EM calorimeter) or tungsten (Hadronic calorimeter)

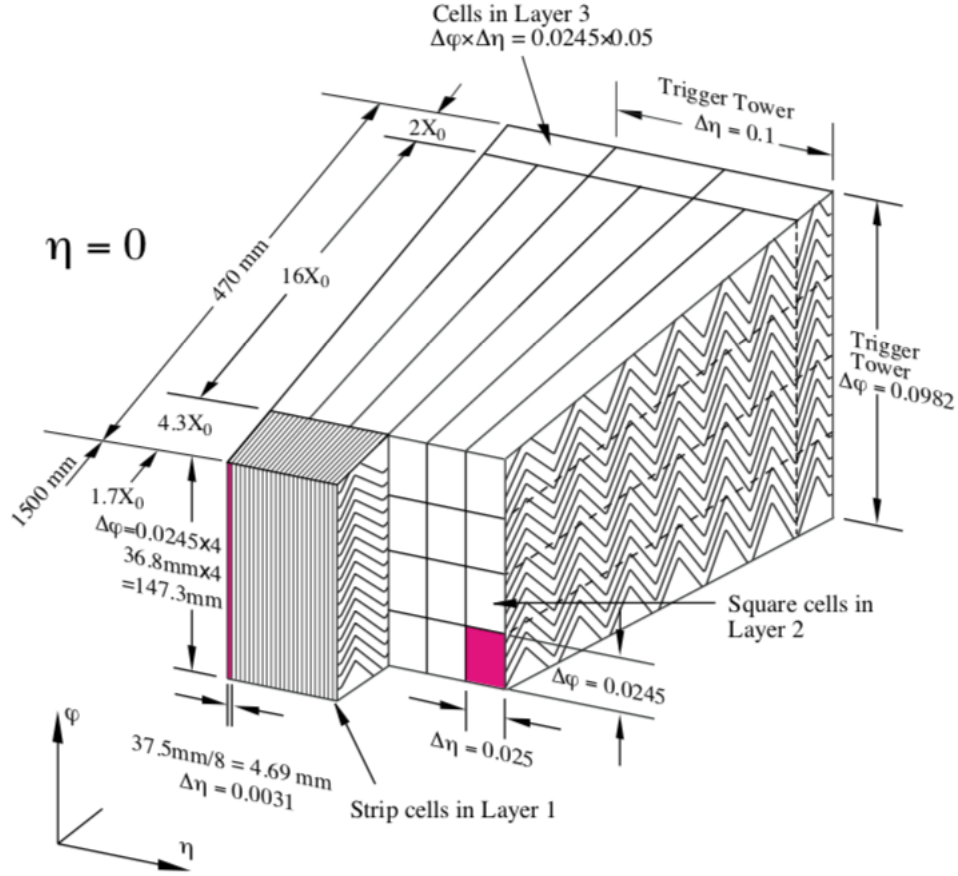


Figure 4.7: ECAL layers, including the readout electronics corresponding to a single trigger tower of cells, as per [30].

as the absorber. Each module consists of a metal matrix with regularly spaced electrodes parallel to the beam axis embedded in the LAr cryostat system. Figure 4.10 shows the location of the location of the FCAL modules in the calorimeter system.

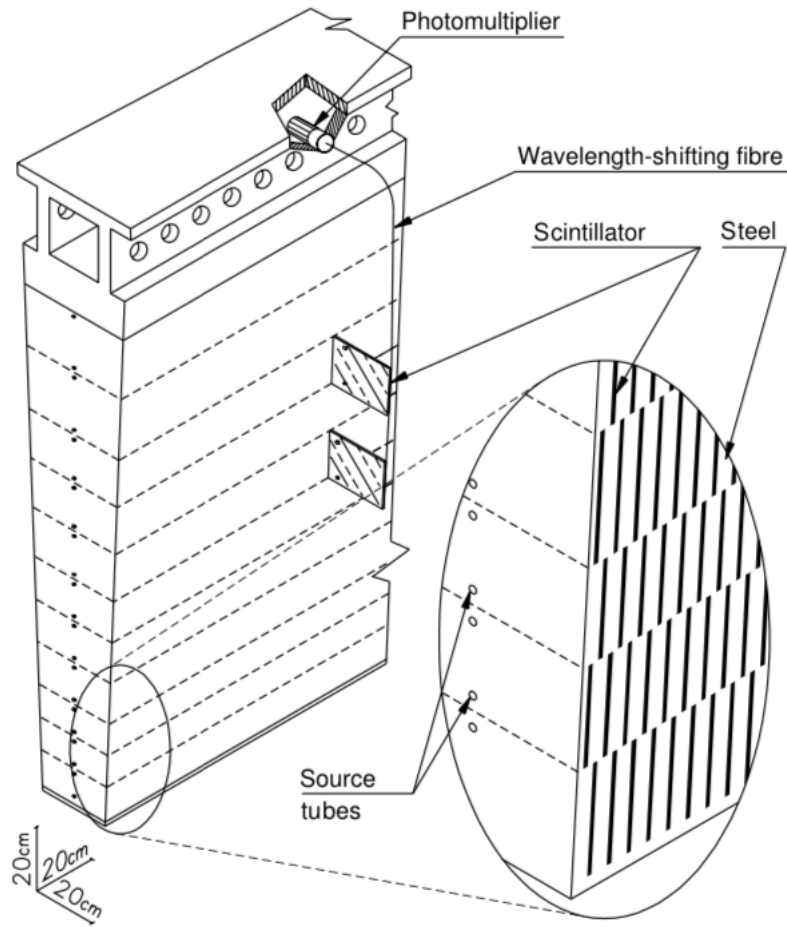


Figure 4.8: Schematic of the tile calorimeter layout as per [30].

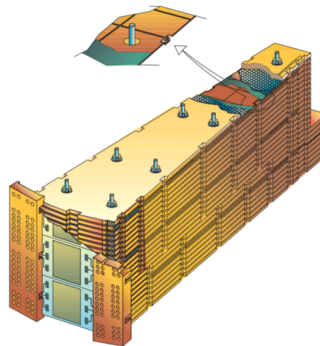


Figure 4.9: Schematic of the hadronic endcap calorimeter (HEC) layout as per [30].

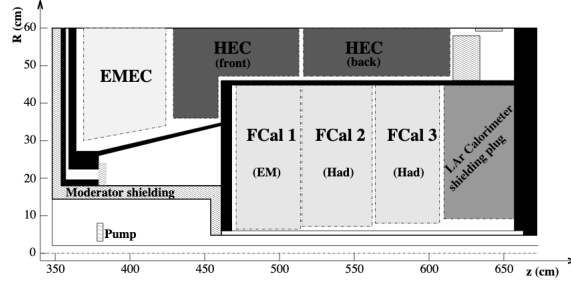


Figure 4.10: Schematic of the forward calorimeter (FCAL) layout as per [30].

4.3.4 Muon Spectrometer

The muon spectrometer forms the outermost component of the ATLAS detector, found outside the toroidal magnetic field, measuring charged particles that still persist after passing through the barrel or end-cap calorimeters in the range $|\eta| < 2.7$, and it is possible to explicitly trigger on events in this detector when $|\eta| < 2.4$. The specification requires that the muon spectrometer has at least a resolution of 10% for tracks with $p_T \sim 1$ TeV. This momentum is observed by measuring the sagitta of muon tracks (the deviation of the muon track from a straight path). In order to achieve this level of precision, the sagitta of a $p_T \sim 1$ TeV muon will be approximately $500\mu\text{m}$ with a resolution of $\leq 50\mu\text{m}$. The muon spectrometer is split into two components, the Monitored Drift Tube chambers (MDTs) (covering $|\eta| < 2.7$), which use 3-8 layers of gas drift tubes at a pressure of 3 bar, giving a resolution of $\approx 35\mu\text{m}$ per chamber, while the forward region $|\eta| \in [2, 2.7]$ consists of Cathode-Strip Chambers (CSC), which are multi-wire proportional chambers segmented into perpendicular strips, giving a resolution of $40\mu\text{m}$ in the bending plane, and 5mm in the direction parallel to the incident muon. From the position tracking in the magnetic field in the muon spectrometer, a high precision evaluation of the charge of the muon can also be identified (due to the direction of its bending in the magnetic field) as well as the momentum. The configuration is depicted in 4.11

For triggering and reconstruction in the third axis, there are two distinct systems covering $|\eta| < 1.05$ and $|\eta| \in [1.05, 2.7]$ (2.4 for triggering), namely the Resistive Plate Chambers (RPC) in the barrel region, and the Thin Gap Chambers (TGCs) in the end-cap respectively.

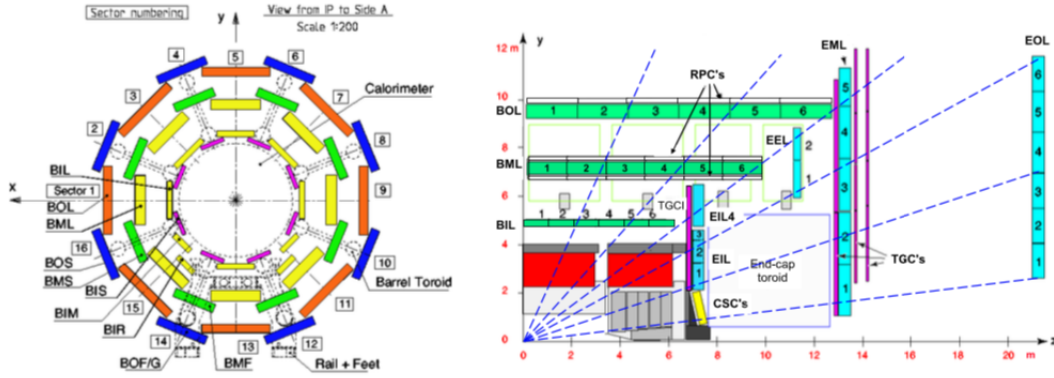


Figure 4.11: Muon spectrometer layout as per the ATLAS technical specification [30], in both a radial view and along the beam-pipe.

4.4 Trigger

Since the data production rate is substantially above the maximum possible write/ transmission speed on disk, specialist hardware and software is used to make decisions on the physics which may arise from hard scatter physics processes, instead of elastic p-p scattering. In addition, ATLAS searches for particularly rare processes, so a system to ensure these special cases are preserved is also important. The sorts of events that may be retained may have particular characteristic features like having a large E_T^{miss} , at least a single electron or muon or some other criteria. ATLAS uses a two level trigger system, based on both hardware and software based triggering systems [42]. The overall configuration is presented in figure 4.12.

The first level trigger system (L1) [43] is a hardware based system built on dedicated hardware, which is designed to operate at a rate of 40MHz. This first level trigger combines the basic calorimeter information and muon track segments, with an acceptance (output to next level) of 100kHz [43]. Dedicated event level reconstruction would not be feasible with such a high input rate, so this is normally processed from the data accepted by the L1 trigger.

A second level trigger known as the *High Level Trigger* (HLT) is used from events selected by the L1 trigger. With the lower input data rate, it becomes possible to perform global track reconstruction and receive the full set of ID tracks, calorimeter and muon track information, making decisions such as “does this event contain an electron/muon”, in addition to cases such as large E_T^{miss} , existence of b-jets in the event, etc. The HLT reduces the output rate to the order of 1 kHz, with the E_T^{miss} trigger [42] being configured to have an output rate of 50-110 Hz.

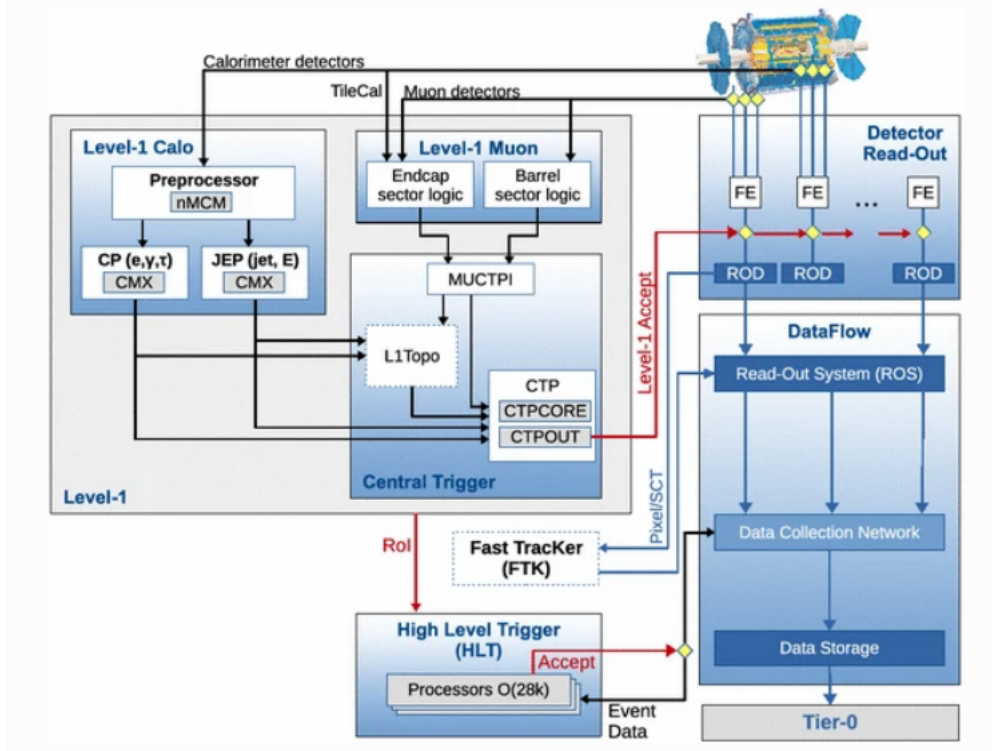


Figure 4.12: Configuration of the ATLAS trigger and data acquisition (DAQ) system, from the design specification in [42].

4.4.1 Analysis level implementation of a trigger

For each trigger, there are distributions of a given variable (of the particular interest in this thesis, the E_T^{miss} or lepton p_T), of which a trigger has a “turn-on” curve, from the minimum data threshold, to where the trigger is considered “maximally efficient”, such that events which truly satisfy the trigger threshold are not lost. For the E_T^{miss} trigger, this turn-on curve can be seen in figure 4.14, and the plateau of trigger efficiency (for offline E_T^{miss}) is around $E_T^{\text{miss}} \approx 250$ GeV. Another is the single electron or muon triggers used in this thesis, henceforth referred to as “single-lepton triggers”, which have a maximum efficiency for an electron/muon with $p_T > 27$ GeV. Selections are imposed in the analyses listed in this thesis such that the trigger is maximally efficient for a given event to be considered.

In some analyses (excluding those listed in this thesis), such as [44], in order to reduce the threshold of some observables, such as lepton p_T , a “pre-scaling” may be used. This pre-scaling could be that the trigger only accepts 10% of events per second, for example, that have passed its other threshold criteria. Since this pre-scaling is dependent on data rate, in this pre-scale condition may change over the course of an LHC operation period (known as a “run”). This exact configuration is performed as part of a run operation and is dependent on the observed instantaneous luminosity being reported during the run.

Since simulated datasets have no such triggering requirements (because they were produced independently of LHC operation), software emulation of such triggering must be applied. However for un-prescaled triggers on their respective efficiency plateau, in practice, this

should be a simple final weighting of the number of events to match the trigger efficiency (if any).

This analysis principally uses the E_T^{miss} trigger, which uses calorimeter tower information, and consists of several different trigger algorithms [42]:

- `xe` [42]: total missing momentum calculated from individual cells.
- `xe_tc_mht` [42]: E_T^{miss} reconstructed from jets.
- `xe_tc_lew`: topo-clusters are built from the calorimeter cells, and the E_T^{miss} is calculated from these topo-clusters.
- `xe_tc_pueta` [42]: pile-up suppression algorithm using the topo-cluster algorithm. Calculate the average topo-cluster energy and standard deviation in ten bins of $\eta \in [-5, 5]$. For each bin, remove topoclusters with an energy above 2σ (referred to as a ring), and calculate the average energy of the remaining clusters, and divide this by the solid angle of the ring to retrieve the pileup energy density. This density is then multiplied by the solid angle of the topocluster to produce a pileup correction which is subtracted from the energy of the topocluster. These pileup corrected topo-clusters are then used to calculate the E_T^{miss} .
- `xe_tc_pufit` [42]: partition the calorimeter into 112 towers of size $\eta \times \phi \approx 0.71 \times 0.79$. For each tower the p_x and p_y components of the topo-clusters centred on that tower are summed to obtain the transverse momentum of the k -th tower. The transverse energy sum of the tower $E_{T,k}$ is calculated from the clusters. If $E_{T,k} < 45$ GeV, the tower is below threshold and the energy is associated to pileup. The average pileup energy density is calculated as $\frac{\sum_k E_{T,k}}{\sum_k A_k}$, k denoting the list of towers below threshold, and A_k denoting the total area in (η, ϕ) coordinates of the towers. A fit estimates the E_T contribution from pileup in each tower above threshold using the average pileup energy density and requiring that there is no event-wide E_T^{miss} arising from pileup. These contributions are then subtracted, and the corrected E_T values are used to calculate the E_T^{miss} .

This trigger (or set of triggers with different turn-on thresholds) has different turn-on curves as a function of the calorimeter tower based E_T^{miss} , which are presented in 4.13, while the trigger efficiency for each algorithm is presented as a function of the offline E_T^{miss} (which includes full reconstruction including tracking and muon spectrometer information) in figure 4.14.

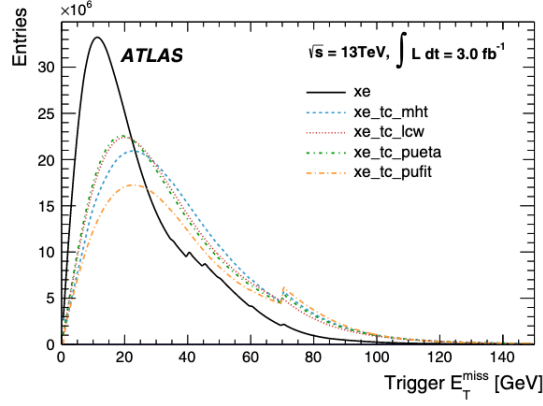


Figure 4.13: Comparison of the E_T^{miss} distributions of events accepted by the trigger algorithms only using the calorimeter energy measurements for each of the different E_T^{miss} trigger algorithms. Note this is smaller than the offline E_T^{miss} , which includes tracking and muon spectrometer information. The stepped behaviour relates to different trigger thresholds. Taken from [42].

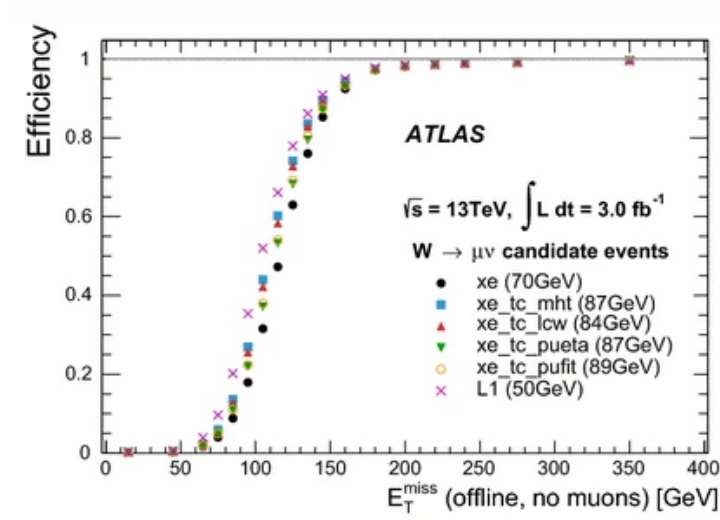


Figure 4.14: Trigger efficiency (Unprescaled trigger efficiency) as a function of the offline E_T^{miss} (so, in this case E_T^{miss} using considering all contributions except those attributed to muons), from [42], for events with a $W \rightarrow \mu\nu$ selection using the 2015 dataset. The key features can be noted as the minimum threshold of the trigger, and the efficiency plateau region. A similar curve for the electron/muon p_T can be found in the same reference.

4.4.2 Data Acquisition and Online Data Quality Monitoring

Events that pass the HLT are then sampled into different streams *Physics_main*, *Debug*, *Trigger* and *Express* streams. The last three are for monitoring of events both at the trigger level and event level for other irregularities, while the *Express* stream is a strict subset of the main dataset, targeting fast live reconstruction of observables, principally for online data quality monitoring by a shifter to debug event level problems. The *Express* stream is processed through the *Data quality monitoring framework (DQMF)* [45], to provide online debugging capability via histograms of key observables, such as the di-lepton mass (for Z peak reconstruction), J/Ψ reconstruction, E_T^{miss} distributions etc. Some preliminary statistical analysis is undertaken on the histograms, such as Kolgomorov-Smirnov tests against a reference histogram to highlight to the shifter any possible mis-modelling, but the primary work relies on the shifter being able to draw the comparisons themselves to debug unexpected trends. This online monitoring is useful in the mitigation of detector level problems, ensuring the maximum possible data available of sufficient quality for analysis. The full express stream is also produced within 24hrs for offline sign-off of the entire run by shifters, prior to full reconstruction and validation of the dataset on the full *Physics_main* stream. The analyses listed in this thesis only comprise of data saved in the *Physics_main* stream, and are strictly not present in the *debug* stream, and only includes runs approved by both online and offline shifters and validators.

The DQMF also monitors output from the other respective components of the ATLAS detector, particularly the status of trigger buffers in each respective component, as well as the monitoring of archiving of both the full dataset and the express debug stream. A typical set of outputs can be seen here, and can be useful for mitigating detector level problems (particularly relating to trigger buffering or monitoring of the number of SCT strips which are inactive due to radiation damage). Some sample screens used as part of the DQMF are presented in figure 4.15.

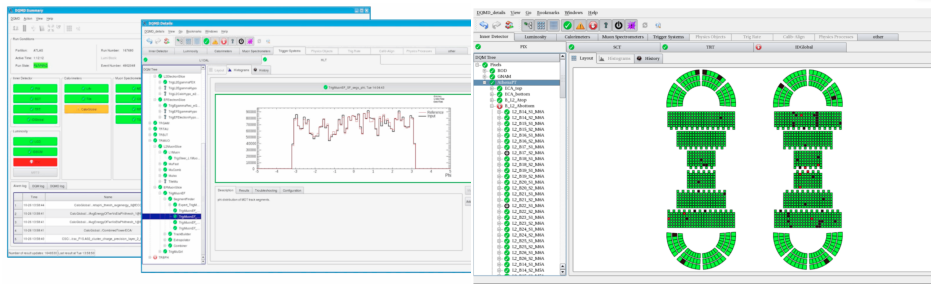


Figure 4.15: Sample screens of the DQM display (DQMD) tool used for in-run data quality monitoring and debugging, as presented in [45]. Left denotes the overview, while right denotes a detailed sub-detector readout.

The author undertook 104 hours of shifts in 2018 corresponding to the monitoring of the online DQMF during LHC p-p collision operation, which largely revolved around reports based on the outputs of the DQMF. This involved actioning online fixes relating to the trigger, calorimeter, muon spectrometer or inner detector based on the readouts of either the trigger monitoring systems or event level overviews. An event display could also be used for online debugging of specific events [46] during the run period, with an example presented in figure 4.16.

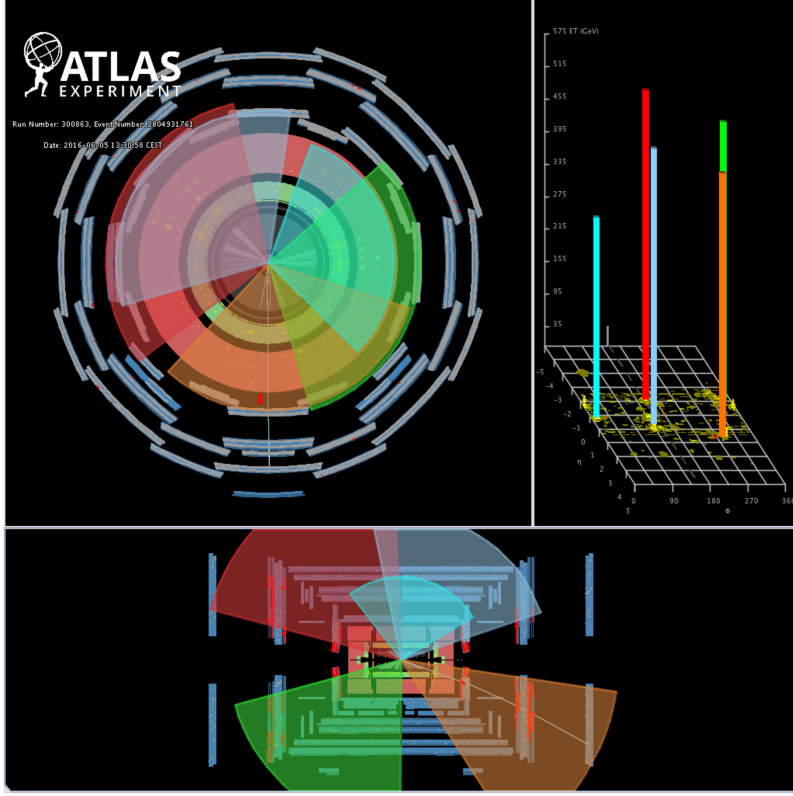


Figure 4.16: An event display [46] during the 2016 proton-proton collision period of a data event which was validated for use in physics analyses. This event [47] contains a $t\bar{t}$ pair with five large radius jets, and a sum of jet invariant masses $M_J^\Sigma = 0.62$ TeV.

4.5 Simulation

This section outlines the production of simulated samples from input particle physics process (Feynman diagrams/rules) to output events that emulate the readouts of the respective process in data. This simulation is largely undertaken on the WLCG, given the raw numbers of events required for some processes - SUSY signals are often relatively small given their cross sections, but backgrounds such as $t\bar{t}$ production can be several billion or more events (depending on which analysis needs the events). A typical ATLAS workflow could be summarised in the following steps (which will be outlined further in the later sections), with the last three steps being shared with data.

1. **Physics process generation:** Monte-carlo derived matrix element and parton showering, modelling the hard scatter process and the hadronisation. Outputs physics objects.
2. **Detector simulation:** Simulation of physics objects interacting with the material components of the detector and the associated infrastructure. Conversion of physics objects to experimentally observable objects (tracks etc).
3. **Pileup modelling:** simulating multiple proton interaction (pileup) through minimum bias samples overlaid with the dataset. This will be discussed further in 4.5.4

4. **Detector digitisation:** Simulation of the digitization (recording of track/cluster/muon spectrometer output). This is required to reflect the efficiencies and capabilities of the detector in different sectors, and produces the recorded tracks, calorimeter clusters, vertexing etc.
5. **Reconstruction:** Reconstruction of the event from the tracks, clusters etc recorded. This is the same as for recorded datasets. This step will be covered in section [4.6](#)
6. **End user analysis:** Application of calibration, reconstruction, identification, trigger efficiencies, other selections, variable definition, multi-variate analyses etc.

This workflow can also be depicted in figure [4.17](#).

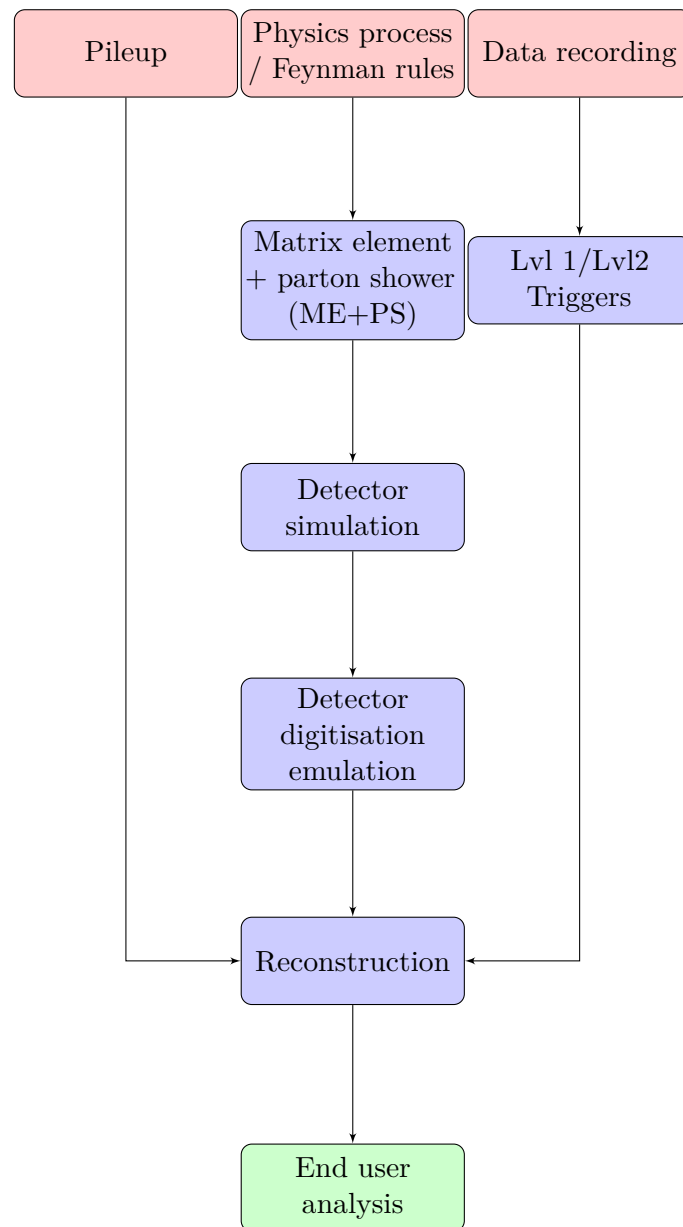


Figure 4.17: Schematic summary of the ATLAS workflow from the underlying matrix element process to final end-user analysis.

4.5.1 Monte-Carlo matrix element generation and Parton Showering

There are a wide variety of algorithms available to produce matrix elements (known as ME generators), depending on the physics process of interest, such as MADGRAPH5 [48] and its NLO version AMC@NLO, SHERPA 2.2.1/2.2.2 [49] or POWHEG [50]. These involve the conversion of feynman rules for a given process, generation of leading order (tree level) and next to leading order (one loop) diagrams (NLO), or higher order (NNLO). Given the NLO and LO (and where available NNLO) generated diagrams specified from the model, Monte Carlo methods are used to calculate the diagrams, producing a set of input events with given kinematic distributions which are used in the numerical calculation of the cross section and other kinematic observables. In addition, at this stage it is possible to vary several model parameters to estimate specific systematic uncertainties based on the theoretical model (MC generator) used.

Given the input events from the hard scatter process, the hadronisation of quarks must also be simulated (since quark colour confinement forbids the existence of isolated unbound quarks), using a variety of hadronisation models. In addition, the ISR and FSR must be considered. These hadronisation models include models such as Pythia 8 [51] or parton showering integrated into the matrix element generator (such as is the case for SHERPA). These models are generally referred to as a *Parton shower*(PS). The hadronisation model chosen, in short effectively takes an input quark state, and appends additional radiative corrections of soft gluons/ $q\bar{q}$ to form hadrons with given momentum, energy etc as well as kinematic angle for each particle produced, merging the output such that the energy scales are consistent between the final state particles and the input partons. The output from the parton showering (and the hard scatter) is referred to as *Truth*, since it reflects the *true* particle state as determined by the Monte Carlo generated events (although, these are inherently dependant on the choice of generator). This truth level output format in ATLAS is referred to as an EVTGEN or EVNT file (corresponding to the Event Generation), and can either be passed to the full detector simulation and treated as a *reconstructed* simulated sample, or converted into a truth-level dataset. Such truth level files can be used, for example, to estimate the systematic impact on the Monte Carlo simulation on changing the ME generator or PS configurations.

In addition, due to the difficulty in modelling the soft interactions (FSR QCD, multiple parton interactions etc), PYTHIA uses a particular set of tuning parameters (or *tune*) which form a group referred to as theA14 TUNE[51]. The parton density functions corresponding to the distributions of the partons within the proton must also be taken into account in the ME/PS generation, normally done by a tool called LHAPDF [52].

The specific details of algorithms used in this thesis will be summarised in an appendix, as the preferred choice of generator (and similarly parton shower) is related to modelling of key observables in data and varies between each respective process.

4.5.2 ATLAS detector simulation & digitisation

There are two principal algorithms used for detector simulation, a full simulation based on GEANT4 [53] and a fast simulation (some approximations based on simplified detector layouts, referred to as “geometries”). Since the full pipeline is somewhat involved, this is best summarised in figure 4.18.

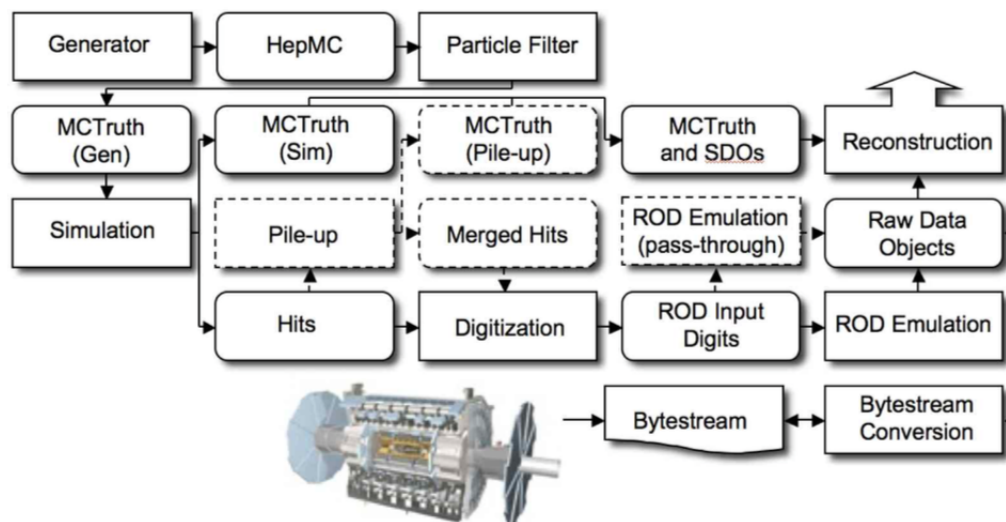


Figure 4.18: ATLAS detector simulation infrastructure subdivided into smaller component parts, as per [54]. This includes a complex set of individual objects: Generator level physics objects are referred to as “truth particles”, which are converted into “simulated data objects” (SDOs), as well as raw data objects (RDOs), the latter of which are either produced by real detector output or from simulation. Digitisation includes emulation of the read-out driver (ROD) electronics used in the detector simulation. Some samples use particle filters, i.e. Z+jets processes filtered to only be diagrams containing $Z \rightarrow e^+e^-$, or $t\bar{t}$ processes with $E_{\text{T}}^{\text{miss}}$ slicing. HepMC [55](or analogously “Les Houches Event”/LHE [56]) data format is a common generator level output format regardless of generator/parton shower type. The dashed boxes/lines correspond to the explicit processing pipelines used for pileup or optional components simulation chain.

GEANT4 simulates the full geometry of the detector, with the simulation of each respective calorimeter component calibrated using a test beam setup, as seen in [54]. The full scope of GEANT4 operation will not be covered here as it is beyond the scope of this thesis, but this can be found in [53].

4.5.3 CPU and wall-time enhancements: FastSimulation

The principal CPU runtime component of the simulation is spent simulating particles passing through the calorimeter volume, particularly in the case of low energy showers. In some cases this is tolerable (e.g. backgrounds for general ATLAS use by multiple teams), whereas for more “specialised” analyses requiring large numbers of signal events that are not readily shared across the entire ATLAS experiment (for example, the analyses listed in this thesis), this was not possible. An alternative is available to reduce runtime by a factor of three [54], namely ATLFAST-II (also referred to as *AFII*, or “fast-simulation”). AFII consists of two components [54]: fast ATLAS tracking simulation (Fatras) and fast calorimeter simulation (FastCaloSim) (normally only the latter is used, but when both are used, it is referred to as “AFII-F”). FastCaloSim [54] uses pre-simulated libraries of low energy showers, truth matched to the low energy electromagnetic particles, instead of the original fully simulated

particle. This simulation type has particular benefit for samples that are not strongly affected by modelling of calorimeter response or precision electromagnetic physics (this issue arises when looking at jet substructure). Fatras uses a simplified overall geometry of the detector volume compared to that used in the full GEANT4 simulation, of which can be seen in figure 4.19.

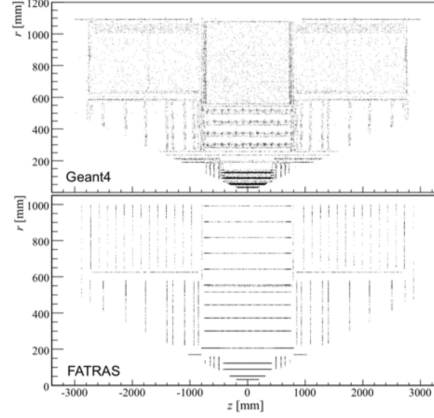


Figure 4.19: Visualisation [57] of the detector layout (geometry) used by the full GEANT4 simulation of tracks and the equivalent using Fatras. This image is derived from the photon conversion vertices.

In addition, due to the level of configuration of the ATLAS simulation infrastructure, it is also possible to simulate each respective component using either GEANT4 or the fast simulation respectively. Digitization can also be simulated using full/fast digitization.

4.5.4 Pileup simulation: FastChain Pileup

Pileup can arise from two sources [58]: *in time pileup*, where there are multiple proton interactions at the same instant in the bunch crossing, or *out of time pileup*, where a second p-p collision occurs within the response time of the detector, and the response associated to the new signal can be masked by the decaying signal of the previous event. Some detector components, such as the LAr calorimeter have response times (i.e. time to attenuate to a noise-only level) longer than the 25ns between bunch crossings, which lead to charges remaining in the calorimeter when a new signal is received (affecting the response of the calorimeter). This pileup distribution for LHC Run 2 was presented in figure 4.2. In time pile-up can either be modelled using minimum-bias data (i.e. special runs with no pileup) events, or simulated using $t\bar{t}$ pythia8 minimum bias events, and to reflect the average pileup conditions of ATLAS for a given hard scatter process, (e.g. for an in-time pileup of 30), a number of events randomly selected from the minimum bias sample are selected and overlaid with the hard scatter event, producing a file with an in-time pileup distribution (provisionally a Gaussian). This distribution is then re-weighted on the user analysis level to correct the original pileup distribution to reflect the pileup distribution of ATLAS for a given run or period. However, with the LHC high luminosity phase (where a pileup of order $O(>100)$ is expected) in the mid 2020s, simulation of min-bias events may become limited by excessive CPU time requirements. In addition, with increased luminosity, accessing and processing

increasingly large file-sets of minbias events becomes an overall limitation, thus *on the fly* pileup simulation has been explored. Such a method entails generating the minbias events required for the sample concurrent with the hard scatter event, rather than separately via overlay).

An alternative was proposed for the min-bias simulation production *on the fly*, namely the *Fast Simulation Chain*[57], exploiting the configurable nature of the ATLAS simulation infrastructure to perform fast-sim, fast digitisation for pileup, while retaining the useful Geant4 hard-scatter (HS) process pipeline, running the pileup (PU) simulation *on the fly* rather than separately. This would use AFII-F (Fatras+FastCaloSim) for the pileup, and several configurations have been considered. As this pipeline is still in development, a schematic overview of the simulation chain for both hard scatter and pileup events can be found in figure 4.20.

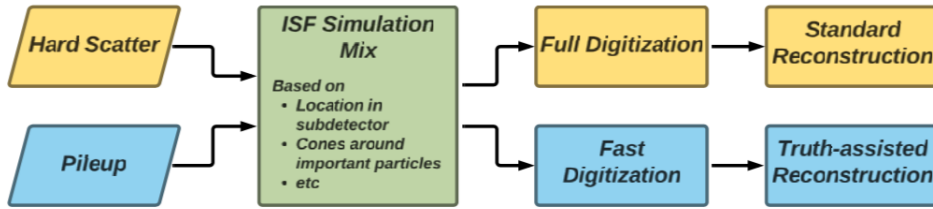


Figure 4.20: Pileup and hard scatter event configuration for the fast simulation chain, as listed in [57]. The simulation component for each can be varied, and can either be considered in sections around physics regions of interest.

In addition to the fast simulation using AFII/AFII-F, fast digitisation[57] may be used in addition, namely projecting a charged particle track through the Pixel and SCT detectors, smearing the readouts, and then using the outputs of the tracking to generate clusters (skipping the cluster finding step in the calorimeter). In the TRT, fast digitization emulates simulated hits from the generated track, smearing to account for the uncertainty on the position of the track from the simulation. Similarly, the generator level information (Truth) can be used to assign the correct detector hits associated (at generator level) to pileup interactions, skipping the cluster/track finding algorithms. However, this truth-assisted reconstruction approach is undergoing validation and is not available at time of writing of this thesis.

4.5.4.1 Author contribution to Fast Chain transform

The author in the configuration of the Fast/Full simulation and digitisation of pileup produced validation routines, primarily focussed on the CPU and memory consumption of the full and fast simulation, and worked on the configuration of some of the ancilliary objects used in the digitisation and reconstruction. Some profiling results produced as part of this study are listed in table 4.3, based on 50 event min-bias $t\bar{t}$ samples (this is chosen mostly for a reasonably low wall-time).

Module	Full HS, Full PU	Full HS, Fast PU	Fast HS, Fast PU
SCT	618	555	55
Pixel	792	699	62
TRT	970	812	176
LAr Calorimeter	1341	1506	1354
Overall CPU time	13664	15329	11261

Table 4.3: Number of CPU time intervals spent in each module for respective ATLAS modules given the configuration of either full/fast simulation and digitisation for hard scatter and pileup events as measured using the “pprof” profiling tool [59]. The configurations presented here should be considered preliminary at time of writing.

In addition, the author contributed to the WLCG based “**A**tlas **R**elease **T**ester” (ART)[60] for FastChainPileup package (which implements the tools listed in this section), which corresponds to nightly based testing of software releases and validation against reference benchmarks, sub-divided into functional components of the jobs (with one job per fast chain configuration, including special cases not listed here preserved for historical reasons). This is seen through a WLCG interface (see figure 4.21). A major consideration when implementing these tests is requiring a fixed CPU architecture: differences in floating point calculations such as $\sin \theta$ can be dependent on the underlying CPU operations, creating spurious warnings.

Package: FastChainPileup
Branch: 21.3/Athena/x86_64-slc6-gcc62-opt
Listed tests are for builds done from 11 apr 2020 to 11 apr 2020
Nightly tag is 2020-04-11T2148
Gitlab ID: 1554298

FastChainPileup									
21.3/Athena/x86_64-slc6-gcc62-opt									
Show output container	11 Apr 2020				Computing site	Duration, h:m:s	CPU time, s	MaxRSS, MB	CPU type
test_FastChain_fatras_mc16a_ttbar.sh	1 2	succeeded	T2148	1	1/2	0:35:35	1759	5065.2	s+Intel(R) Xeon(R) Gold 6150 CPU @ 2.70GHz 25344 KB
test_fastchain_g4ms_mc16a_ttbar.sh	1 2	succeeded	T2148	3	1/2	0:35:41	583	5531.4	s+Intel(R) Xeon(R) Gold 6150 CPU @ 2.70GHz 25344 KB
test_FastChain_mc16a_ttbar.sh	1 2 3	succeeded	T2148	2	1/2	0:38:26	2203	5539.9	s+Intel(R) Xeon(R) CPU E5-2690 v4 @ 2.60GHz 35840 KB
test_stdReco_fastSim_fullDigi.sh	1 2 3 4 5	succeeded	T2148	4	1/2	0:59:56	3354	3226.8	s+Intel(R) Xeon(R) CPU E5-2690 v4 @ 2.60GHz 35840 KB
test_stdReco_fullSim_fullDigi.sh	1 2 3 4 5	succeeded	T2148	5	1/2	0:59:36	3459	3310.5	s+Intel(R) Xeon(R) CPU E5-2690 v4 @ 2.60GHz 35840 KB
test_ttFC_fastSim_fastDigi.sh	1 2 3	succeeded	T2148	6	1/2	0:34:30	1940	3195.0	s+Intel(R) Xeon(R) Gold 6150 CPU @ 2.70GHz 25344 KB
test_ttFC_fastSim_fullDigi.sh	1 2 3	succeeded	T2148	7	1/2	0:31:25	1830	3227.3	s+Intel(R) Xeon(R) CPU E5-2660 v3 @ 2.60GHz 25600 KB
test_ttFC_fullIHS_fastPU_simDigi.sh	1 2 3	succeeded	T2148	8	1/2	0:58:33	3222	3313.4	s+Intel(R) Xeon(R) CPU E5-2690 v4 @ 2.60GHz 35840 KB
test_ttFC_fullSim_fullDigi.sh	1 2 3	succeeded	T2148	9	1/2	0:39:30	2410	3310.5	s+Intel(R) Xeon(R) Gold 6150 CPU @ 2.70GHz 25344 KB
test_ttFC_reco_newTracking_PseudoT_fastSim_fastDigi.sh	1 2 3 4 5	finished	T2148	13	1/2	1:09:45	3702	3156.1	s+Intel(R) Xeon(R) CPU E5-2660 v3 @ 2.60GHz 25600 KB
test_ttFC_reco_newTracking_PseudoT_fastSim_fullDigi.sh	1 2 3 4 5	succeeded	T2148	14	1/2	1:22:45	3376	3241.6	s+Intel(R) Xeon(R) CPU E5-2660 v3 @ 2.60GHz 25600 KB
test_ttFC_reco_newTracking_PseudoT_fullSim_fullDigi.sh	1 2 3 4 5	succeeded	T2148	15	1/2	0:56:31	3031	3309.2	s+Intel(R) Xeon(R) CPU E5-2690 v4 @ 2.60GHz 35840 KB
test_ttFC_reco_noSplit_noPseudoT_fastSim_fastDigi.sh	1 2 3 4 5	finished	T2148	19	1/2	1:03:39	3458	3162.9	s+Intel(R) Xeon(R) CPU E5-2660 v3 @ 2.60GHz 25600 KB
test_ttFC_reco_noSplit_noPseudoT_fastSim_fullDigi.sh	1 2 3 4 5	succeeded	T2148	20	1/2	1:04:49	3815	3225.4	s+Intel(R) Xeon(R) Gold 6150 CPU @ 2.70GHz 25344 KB
test_ttFC_reco_noSplit_noPseudoT_fullSim_fullDigi.sh	1 2 3 4 5	succeeded	T2148	21	1/2	1:14:38	4096	3311.6	s+Intel(R) Xeon(R) Gold 6150 CPU @ 2.70GHz 25344 KB

Figure 4.21: A snapshot of the ART interface webpage for FastChainPileup package including multiple configurations (as of 11/04/2020) and different sub-components on/off for the purposes of debugging/validation. The explicit choice of these implementations is considered preliminary, so will not be motivated further in this thesis. This can be accessed [here](#) - CERN auth required.

4.6 Reconstruction

Since useful physics objects are not manifestly observed directly, rather inferred via their interactions, the following section outlines the conversion of detector output: energy deposits in cells/clusters, hits in the tracking system, hits in the muon spectrometer, trigger decisions etc into useful physics objects (e.g. electrons, muons, τ -leptons, quark/gluon jets). As will be mentioned, this same reconstruction system will be used for simulated samples as well as for raw data, thus commonalities are shared between them, with the addition in reconstructed simulated samples, the original generator level “truth” information is available.

4.6.1 Jet reconstruction

As a result of colour confinement, a final state parton from a collision event undergoes a cascade of bremsstrahlung of collimated gluons and quark anti-quark pairs, known as hadronization. The final objects are colour-neutral hadrons, and groups of such approximately collimated hadrons are called jets, because of their cone-like signature. In simulation, this is emulated by the parton showering algorithm, which uses a given class of hadronisation models (such as PYTHIA 8 [51] or HERWIG [61]), which add additional gluons to the final state particles produced from the hard process (generator) depending on factors such as the QCD scale. These are configured to model the same effects that would be seen in data, and different parton showering algorithms have different modelling performance for a given hard scatter process (or even specific monte-carlo generator), such that there is no universal simulation of a parton shower.

In ATLAS [62], the algorithms used is principally the anti- k_T algorithm[63], using information either from tracks (see 4.6.1.2), calorimeter information (see 4.6.1.1) or a combination thereof (not used in this thesis, but can be found in [64]). This algorithm is referred to as a “sequential recombination algorithm”[63], and replaces cone-like algorithms implemented in earlier experiments (such as SIScone [63, 65]). This algorithm is as follows:

1. Identify a set of proto-jet (either a track, calorimeter cell/cluster) with momenta p_i^μ .
2. For each proto-jet (i, j) and for some constant R , define

$$d_i = p_{T,i}^{-2}, \quad \forall i \in \text{event} \quad (4.4)$$

Note, in several sources (e.g [63, 66]), this is normally referred to as the transverse energy of a sub-jet, but in ATLAS [62], the input sub-jets are treated as massless, hence the equivalence between transverse energy and magnitude of transverse momenta can be applied.

$$d_{ij} = \min(p_{T,i}^{-2}, p_{T,j}^{-2}) \times \frac{\sqrt{(\Delta\phi_{ij})^2 + (\Delta\eta_{ij})^2}}{R} \quad \forall (i, j) \in \text{event} \quad (4.5)$$

3. Find the smallest of the d_i and d_{ij} , labelling it $d_{\min}$.
4. If a d_{ij} is the minimum value, merge the proto-jets into a new proto-jet k , else if the minimum corresponds to d_i , then the proto-jet is removed and added to the list of jets.

5. Recursively iterate through this process until there are no further proto-jets.

Since this algorithm is independent of the type of object used in the reclustering (relying solely on position and momenta/energy resolution only). Thus, jets can be built either based on track-jets or calorimeter clusters.

4.6.1.1 Calorimeter Jets

ATLAS uses collections of individual calorimeter cell singles in a cluster referred to as a “Topo-cluster”, which are designed to follow spatial patterns generated by particle showers [67]. This is managed by the cell signal significance $\mathfrak{s}_{\text{cell}}$, namely

$$\mathfrak{s}_{\text{cell}} = \frac{E_{\text{cell}}^{\text{EM}}}{\sigma_{\text{noise,cell}}^{\text{EM}}}. \quad (4.6)$$

This significance threshold defines how a cluster is formed through a growing-volume [67] algorithm, determining the initial input seed cell (with high significance), and expanding the cluster to include additional neighbouring cells.

- Primary seed cell requirement: $|\mathfrak{s}| > 4$. Adjacent seed cells are merged where appropriate.
- $|\mathfrak{s}| > 2$: Cluster inclusion requirements of neighbouring cells/calorimeter layers.
- $|\mathfrak{s}| > 0$: Boundary / filter condition for a seed cell to not be included in a topo-cluster.

Noting the fact that the modulus of the cell significance is used, it is possible to reconstruct negative energy jets [67] from the calorimeter into these topo-clusters, which can either arise from electronic noise in the readout systems or from pile-up, but any jet which is determined to have $E_{\text{cluster}} < 0$ is neglected. However, including cells with negative energies into topo clusters which have net positive energy is also used to improve noise and pile-up suppression. These topo-clusters form the input to the anti- k_T algorithm with the distance parameter $R_0 := 0.4$, corresponding to a “cone” radius of 0.4, and are often referred to as “EMTopo” jets [62], since the energy inputs are calibrated to the “EM scale”, namely in detector design, prototyped with e^+e^- colliders and calibrated using electrons/photons during detector operation.

4.6.1.2 Track jets

In addition to jets produced in the calorimeters, it is possible to construct jets from the track information instead. This is particularly efficient in the lower p_T range of jets, since the reconstruction efficiencies of jets in the calorimeters is limited to about 17 GeV minimum (with maximal efficiency around 20 GeV). However, a major limitation of the track jets is that they only resolve charged particle jets (any neutral component is not recorded by the tracker). This is particularly used in the case of soft b-hadron identification [68], and involves tracks which are found in tracker acceptance ($|\eta| = 2.5$, $p_T > 500$ MeV). Such tracks also satisfy quality criteria including at least 7 SCT/Pixel hits, at most 1 hit shared with another

track and at most 1(2) missing hits in the pixel detector (SCT), alongside requirements on the impact parameter of the track $|z_0 \sin \theta| < 3\text{mm}$. Variable radius track jets are defined using the anti- k_T algorithm, but with the R parameter in the anti- k_T algorithm allowed to depend on the jet p_T .

$$R \rightarrow R_{\text{eff}} = \frac{\rho}{p_T} \quad (4.7)$$

The optimal set of parameters for this effective radius are [68] $\rho = 30 \text{ GeV}$, $R_{\text{min}} = 0.02$, $R_{\text{max}} = 0.4$, with the latter two parameters forming upper and lower bounds on the jet size. Effectively for low p_T (i.e. $< 30 \text{ GeV}$) such variable radius jets are considered of a fixed radius of $R = 0.4$ akin to their EMTopo equivalents, while a lower cut of $p_T > 5 \text{ GeV}$ is required as the lower bound limit of track-jet reconstruction.

4.6.2 Heavy flavour jet (B-jet) reconstruction in calorimeter jets

This section covers the identification of b-hadrons inside jets [69], allowing distinctions to be made for jets associated to b-quark final states, a key facet of reconstructing particular objects such as the Top and Higgs decay processes. Since the b-hadron has a finite lifetime of $\tau \approx 1.5\text{ps}$ ($\langle c\tau \rangle \sim 450\mu\text{m}$) [69], which relates to a decay length of $\langle L \rangle = \beta\gamma c\tau$ (for β, γ the relativistic factors).

Given this decay length being sufficiently large, a substantial component of b-jet identification is high-precision secondary vertex finding (since the b-hadron decay could be resolved at some distance away from the beamline interaction point, and tracks associated could be matched to said intermediate vertices). This sort of decay can be visualised in figure 4.22. With any algorithm that requires reconstruction of secondary vertices, two impact parameters must be considered: the actual decay length (commonly denoted as L_{xy}), and the transverse impact parameter d_0 is used, which is the projection of the distance between the primary and secondary vertices into the r, ϕ plane (where r is the radial distance from the collision point). B-tagging relies on three distinct “low level” classifiers, two of which correspond to detector impact parameter-based algorithms referred to as IP2D and I3PD respectively, while a secondary vertex finding algorithm, SV1, is used to find the b-hadron decay vertex in the tracker.

The IP2D algorithm [69] makes use of the signed impact parameter significance ($d_0 \text{ signif.} = d_0/\sigma(d_0)$) of tracks, while the IP3D algorithm considers the impact parameter significance in both the transverse and longitudinal directions (so, including $z_0 \sin \theta/\sigma(Z_0 \sin \theta)$). These algorithms are considered complementary to each other, with a 2D template fit applied to account for the correlations between the impact parameters, using a log-likelihood ratio (LLR) [69] of probability density functions (pdf) to classify the jet into either b or light-jet based on the impact parameter search. The SV1 algorithm is an iterative procedure [69] on candidate two-track vertices, rejecting those which are associated to K_S or Λ . photon conversions or other detector interaction related effects. These two track vertices are evaluated with a χ^2 test, undertaking “track to vertex association” with the track with the largest χ^2 removed at each step, with this vertex fit repeated until a jet has an invariant mass of tracks associated to the vertex of $< 6 \text{ GeV}$.

The remainder of the input algorithms are summarised in table 4.4. These low-level identifiers are combined in the high-level MVA classifier known as MV2c10[71](strictly this algorithm

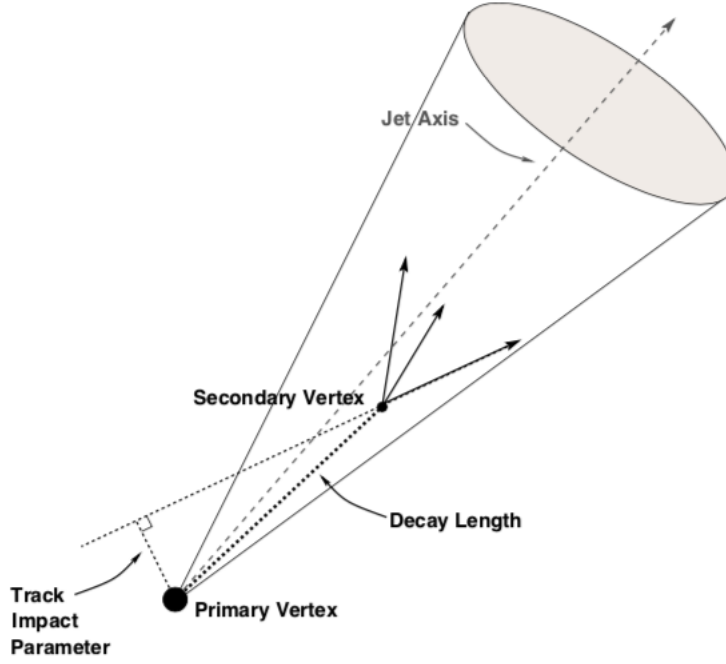


Figure 4.22: schematic[70] of a B-hadron decay inside a jet cone.

Input	Description
Kinematics	Jet p_T/η
IP2D/IP3D	Log Likelihood ratios between b-jet, light-jet and c-jet hypotheses
SV1	Secondary vertex classifier, using invariant masses, spatial distance between primary and secondary vertices and N_{tracks} for secondary vertex
JetFitter	Track mass from secondary vertices, displacement between tracks, jets, primary and secondary vertices. Number of displaced vertices
JetFitter c-tagging	Displacements of secondary vertex and mass/energy fractions.

Table 4.4: Summary of the low-level algorithms used in flavour tagging (b-jet identification) for ATLAS. A more detailed overview of each respective algorithm is discussed here in [69].

is a class, determining different fractions of b-jets and c-jets in the training). MV2 is the root algorithm using a BDT on the classifier outputs listed in 4.4, while the “c10” means that a sample including $\approx 7\%$ [72] charm quarks was used in the training). The BDT is trained on a hybrid sample of both $t\bar{t}$ and Z' , covering the full range of calorimeter jet p_T . The calibration principally uses $t\bar{t}$ samples up to a $p_T < 600$ GeV, but due to limited tail statistics, a Z' sample is used (since in addition to a lack of $t\bar{t}$ statistics in the high p_T tails, there is a lack of data events). This region is extrapolated from calibrations below 600 GeV)

4.6.3 Trackjet b-tagging

Track-jet b-tagging is motivated from the generator-level simulations of the b-hadron p_T in specific soft SUSY scenarios (corresponding to those of four body stop decay), which determines that the majority of b-jets would be produced with $p_T < 20$ GeV [68]. Since the calorimeter has maximal jet reconstruction efficiency for only > 20 GeV (and a minimum

threshold around 17 GeV), most of the b-hadrons in the event are lost in a given four body stop decay using calorimeter jets. Thus with the low- p_T performance of charged particle track-jets (maximally efficient for $p_T \in [5, 30]$ GeV), an approach based on track-jets is favoured.

The principal requirement for the jets that could be classified as b-jets are those which have at least two tracks associated of the jet, and a veto is applied if the track jets explicitly overlap (to avoid concentric structures of jet clusters). Reference [68] discusses a set of alternative possible vertexing algorithms, but these were deemed to be of comparable performance to direct application of the MV2c10 algorithm onto VR trackjets satisfying the above criteria. The MV2c10 algorithm on track jets is considered effective for track jets with $p_T > 5$ GeV, where below this track jet reconstruction is not efficient. The MV2c10 tagger is trained on $t\bar{t}$ with a nominal 60%, 70% or 77% efficiency working point, which corresponds to a fixed selection on the MV2 score.

4.6.3.1 Large radius jets and jet substructure

As a result of colour confinement, a final state parton from a collision event undergoes a cascade of bremsstrahlung of collimated gluons and quark anti-quark pairs, known as hadronization. The final objects are hadrons, and groups of such approximately collimated hadrons are called jets, because of their cone-like signature. A jet is then determined using reconstruction algorithms built from calorimeter clusters, predominantly the anti-kt algorithm [63]. ATLAS provides a set of calibrations $\Delta R = 0.4, 0.8, 1.2, 1.0$, where:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \quad (4.8)$$

However, determination of jets is fraught with challenges, such as cases where multiple jets overlap can be found inside a single reconstruction cone, which is common in the high p_T , or boosted decay scenarios. Thus an attempt must be made to understand the topology, or substructure within a single jet cone, to identify if a given jet is actually a single jet, or consisting of multiple jets overlapping in a single reconstruction cone, such as in Figure 4.23:

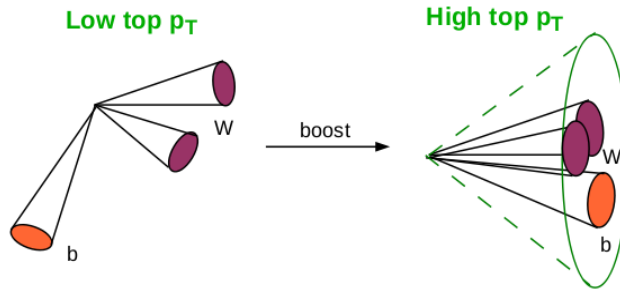


Figure 4.23: A top quark decay into jets and a b-jet in both low and high p_T regimes, with resolved jets in the low p_T regime, and all jets lying within a single jet cone in the boosted regime.[73].

The subject of jet topology is diverse, and many different variables have been suggested, although the most popular variables are [74–77]:

- N-subjettiness, defined by:

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min\{\Delta R_{1,k}, \Delta R_{2,k}, \dots, \Delta R_{N,k}\}, \quad d_0 = \sum_k p_{T,k} R_0, \quad (4.9)$$

for k running over the constituent particles in a jet, and R_0 the characteristic jet radius. This quantifies the extent to which a jet is actually consisting of N -subjets, and the ratio $\tau_{\{N+1,N\}} := \frac{\tau_{(N+1)}}{\tau_N}$ is a useful discriminant.

- Splitting scale, or k_T distance metric.

$$d_{ij} = \min(p_{T,i}^2, p_{T,j}^2) \Delta R_{ij}^2, \quad (4.10)$$

for i, j small radius sub-jets inside the jet. The sub-jets are reconstructed using the k_T method, using the metric above, while the large radius jets are constructed using the anti- k_T algorithm, which uses $p_{T,i}^2 \mapsto p_{T,i}^{-2}$ instead. Either the metric itself or its square root may be used in event discrimination.

- Sum of the masses of the small sub-jet constituents inside the large R jets.
- Energy Correlation Functions and their ratios (for a large R jet consisting of N constituents).

$$\text{ECF}(N, \beta) := \sum_{i_1 < i_2 < \dots < i_N \in \text{Jet}} \left(\prod_{k=1, k}^N p_{T, i_k} \right) \left(\prod_{l=1}^{N-1} \prod_{q=l+1}^N R_{i_l i_q} \right)^\beta, \quad (4.11)$$

where α denotes a free parameter, and R_{ij} denotes the (η, ϕ) distance between sub-jets i and j . To best visualise this formula, the cases when $N \in [1, 4]$ are shown:

$$\text{ECF}(1, \beta) = \sum_{\text{jet constituents}} p_{T, \text{constituent}}, \quad (4.12)$$

$$\text{ECF}(2, \beta) = \sum_{i < j \in J} (p_{T,i} p_{T,j} (R_{ij})^\beta), \quad (4.13)$$

$$\text{ECF}(3, \beta) = \sum_{i < j < k \in J} (p_{T,i} p_{T,j} p_{T,k} (R_{ij} R_{jk} R_{jk})^\beta), \quad (4.14)$$

$$\text{ECF}(4, \beta) = \sum_{i < j < k < l \in J} (p_{T,i} p_{T,j} p_{T,k} p_{T,l} (R_{ij} R_{ik} R_{il} R_{jk} R_{jl} R_{kl})^\beta). \quad (4.15)$$

If for a given jet, N is greater than the number of constituent jets within the jet, $\text{ECF}(N, \beta) = 0$. Much like the n -subjettiness, the ratios are also of interest, namely:

$$r_N^{(\beta)} := \frac{\text{ECF}(N+1, \beta)}{\text{ECF}(N, \beta)}, \quad (4.16)$$

and the double ratio [76]:

$$C_N^{(\beta)} := \frac{r_N^{(\beta)}}{r_{N-1}^{(\beta)}} \equiv \frac{\text{ECF}(N+1, \beta) \text{ECF}(N-1, \beta)}{(\text{ECF}(N, \beta))^2}. \quad (4.17)$$

Infrared and collinear (IRC) safety requirements must be considered, which requires that $\beta > 0$. These are the requirements that jet algorithms are stable under the cases where $p_T^{\text{const}} \rightarrow 0$ (Infrared safe), or in the case of two constituents (i, j) that the algorithm is stable under the limit $\Delta\theta_{ij} \rightarrow 0$ or $\Delta\phi_{ij} \rightarrow 0$ (collinear safe). A value [76] of $\beta = 1.0$ was determined to be optimal for boosted top identification from the $R=1.0$ anti- k_T jets.

These variables are chosen to discriminate between large- R jets which are anti- k_T jets with $R = 1.0$, according to whether they originate from either two or three prong processes, to associate with either those associated with the decay of a W boson jets (two prong), or top associated jets (three prong), using constituent jets reconstructed with anti- k_T and $R = 0.2$. The application of these objects will be discussed in section 4.6.3.1.

4.6.4 Leptons (electrons, muons)

This section will only cover the reconstruction of electrons and muons. τ -lepton reconstruction is not in general used in the analyses listed in this thesis (with the exception of the study in section 7.4).

4.6.4.1 Electrons

Electrons [78] are reconstructed from electromagnetic calorimeter cluster candidates (local energy deposits) and tracks, and the reconstruction of them consists of three parts: reconstruction of the electron from the calorimeter, identification of the electrons (and rejection of fake electron-like objects), and the isolation of said reconstructed and identified electrons. As electrons interact with the tracking system, they also produce showers of photons which can be observed in the calorimeter volume.

1. Electron reconstruction:

Electron reconstruction is defined in the precision tracking region $|\eta| < 2.47$, with full coverage in ϕ . Tracks are reconstructed from SCT/Pixel/IBL hits, and the full methodology of their identification can be found here [78].

For the purposes of reconstructing the electrons candidates from calorimetry, the EM calorimeter is separated into 200×256 calorimeter cells of size 0.025×0.025 (in the second calorimeter layer) (with a higher granularity in the first layer). Electron seed cluster (the first cluster used to reconstruct the electron) candidates are selected using a sliding window algorithm, namely searching for a cluster set of size 4×5 in (η, ϕ) cells such that the $\sum_{\text{cells}} E_T > 2.5$ GeV, iteratively finding such clusters until all elements of the calorimeter are evaluated. If there are two candidates within $\eta \times \phi = 5 \times 9$ cells, the seed cluster with the higher E_T is kept if $|\Delta E_T| > 10\%$, otherwise the cluster with the highest E_T central cell is kept only.

The reconstruction of electrons requires a matching between the track and the calorimeter seed cluster, and a full reconstructed cluster is used, namely a cluster of size 3×7 units in the barrel or 5×5 in the endcap (expanding the size of the original seed cluster). The final electron track is re-fitted to be consistent with the seed cluster

position information. The explicit electron charge can be reconstructed from the curvature of the tracks.

2. **Electron identification:** Electron identification is based on a likelihood based identification [78], namely $L(\mathbf{x}) = \prod_{i=1}^n P(x_i)$ for either the signal L_S or background L_B , with $P(x_i)$ denoting the pdf corresponding to signal or background for each variable listed in table 4.5. The backgrounds defined here are electrons originating from either photon conversions in the detector volume, non-prompt electrons or jets which fake electrons (combined into a single background). A discriminant d'_L is used for each background [78], namely:

$$d_L = \frac{L_S}{L_S + L_B}, \quad d'_L = -\frac{1}{15} \ln \left(\frac{1}{d_L} - 1 \right) \quad (4.18)$$

Different working points of the identification correspond to different values of d'_L , with the “Loose” operating point used in this thesis being a value chosen such that the efficiency of identification is $\approx 85\%$ for $E_T \approx 20$ GeV to $\approx 96\%$ at $E_T \approx 100$ GeV.

Type	Backgrounds vetoed
Hadronic leakages	jets, γ conv.
Ratio of third layer energy of calo to total	jets
Shower widths in 1st/2nd calo layer, energy ratios	all
Transverse impact parameters (Tracks)	γ conv. , non-prompt
TRT information	jets
Track/Cluster matching	jets, γ conv.

Table 4.5: Summary of the table of variables listed in [78]. The backgrounds in electron identification are light jets, photon conversions and non-prompt electrons from heavy flavour decays.

3. **Electron isolation:** This is undertaken [78] using topological clusters with signals corresponding to ≥ 4 times the expected cell noise as the seed cluster. Additional clusters are built in the neighbouring cells/layers with energy deposits twice the noise level. From an electron candidate, a cone of radius ΔR is drawn, and clusters whose barycenters lie within this cone are included in the raw isolation energy, alongside the electron energy. The effects of leakage (a shower depositing some energy outside of the bounds of the electron reconstruction window) and pileup are also considered. The isolation energy is defined as:

$$E_T^{\text{iso}} = E_T^{\text{cone}} - \sum_{\text{corrections}} E_T^i - E_T^{\text{e-cand}} \quad (4.19)$$

In addition to the calorimeter based isolation, track based isolation is used, using the cone of a radius ΔR around a given track to isolate a given electron [78], defined as

$$\Delta R = \min \left(\frac{10 \text{ GeV}}{p_T [\text{GeV}]}, R_{\text{max}} \right), \quad R_{\text{max}} := 0.2 \quad (4.20)$$

$$p_T^{\text{isol}} = p_T - \sum_{\text{Corrections}} p_T^i - p_T^e \quad (4.21)$$

This thesis covers the use of electrons isolated with the “FCLoose” working point, corresponding to a selection on $E_T^{\text{iso}}/p_T < 0.2$ ($\Delta R = 0.2$), $p_T^{\text{isol}}/p_T < 0.15$ ($R_{\text{max}} = 0.2$)

4. **Electron reconstruction efficiency measurements:** Electron reconstruction efficiencies can be estimated directly from “tag and probe” methods, namely using datasets that refer to either the $Z \rightarrow e^+e^-$ or $J/\psi \rightarrow e^+e^-$ resonances, using a strict requirement on one of the object (the “tag”), to acquire an unbiased estimate of the remaining electron (the “probe”) assuming a given value of the invariant mass between the two electrons. This efficiency can then be calculated as:

$$\epsilon_{\text{tot}} \equiv \epsilon_{\text{clusters}} \times \epsilon_{\text{reco}} \times \epsilon_{\text{ID}} \times \epsilon_{\text{isolation}} \times \epsilon_{\text{Trigger}} = \frac{N_{\text{Trigger}}}{N_{\text{all}}} \quad (4.22)$$

which incorporates the calculations of the cluster reconstruction, trigger, identification and isolation efficiencies. The $Z \rightarrow e^+e^-$ process is used to calculate [78] the electron reconstruction efficiencies in full, while the $J/\psi \rightarrow e^+e^-$ can only be used to extract the identification efficiency. The two tag and probe measurements allow reconstruction efficiencies to be estimated for $E_T > 4.5$ GeV, hence this forms the lower bound of electrons in the analyses listed in this thesis.

To correct for data/MC mis-modelling, two efficiencies are calculated, namely the data efficiency and that of simulated samples. The ratio of these is referred to as the “scale factor”, and is used to correct the efficiencies of simulated datasets to better reflect that observed in data.

4.6.4.2 Photon reconstruction

Photons were used in the analysis listed in [79](Chapter 6), principally in the definition of the $t\bar{t} + \gamma$ control region, and thus will be listed here. Photon reconstruction is analogous to that of electron reconstruction, using both the information from the tracker and calorimeter. A photon is not expected to have a direct track, but may induce a charged track by ionising particles along its flight path, or converting to an electron during its path. Principally, determining the track properties and vertices of possible photon conversions leads to the reconstruction of photons as separated from electrons. Also a major background are non-prompt photons which must be removed through selections on isolation variables. This thesis will not cover this particular topic in more detail, but the interested reader is directed to [80].

4.6.4.3 Muons

Muon reconstruction [81] is subdivided between two independent detector subsystems, namely via the inner detector (ID) in a manner similar to electrons as per the previous subsection, and via the muon spectrometer (MS). First a search for hit patterns in each muon chamber is undertaken, particularly searching for hits and hits aligned on a trajectory in the plane of bending of the muon track in the magnetic field. Muon track candidates are then reconstructed from the respective segment hits. For the purposes of some parts of the reconstruction, these tracks are extrapolated back to the interaction point (these are referred to as *ME tracks*). The ME tracks are used primarily in the region where there is no inner detector coverage ($|\eta| \in [2.5, 2.7]$). A matching criteria between the MS and ID tracks is applied, and afterward the entire track is re-fitted to be consistent between the two detector components, producing a *CB track* in both the ID and MS ([81] refers to this type of track as “combined”) . Calorimeter only tracks are referred to as *CT tracks*.

1. **Identification:** The key variables for identification of muons [81] are:

- q/p significance: Absolute value difference of the ratio of the measured charge and momenta of the muons as measured in the MS and ID (divided by their respective uncertainties).
- ρ' , the absolute value of the difference of p_T measured between the calorimeter muon and MS muon tracks divided by the p_T of the combined muon track.
- χ^2 from the fit of the muon tracks between the ID and the MS.

The analyses listed in this thesis use *medium* identified muons [81], which use CB and ME tracks. CB tracks are required to have at least 3 hits in at least two MDTs, except in the particular region $|\eta| < 0.1$ (this is where cabling and cooling services are connected to the ID and calorimeters, hence lower granularity), where only one MDT module is required to have hits, and at most one hole is allowed. ME tracks are used for tracks lying outside the tracker η -acceptance, requiring hits to at least three MDT/CSC layers. Additional restrictions on the q/p significance are applied in this region to veto hadrons which mimic muons.

2. **Reconstruction efficiency:** The efficiency of the medium working point is calculated using a tag and probe method in a similar manner to that undertaken for electrons, using $Z \rightarrow \mu^+ \mu^-$ or $J/\psi \rightarrow \mu^+ \mu^-$.

$$\epsilon(\text{Medium-}\mu) := \epsilon(\text{Medium-}\mu|\text{ID track}) \times \epsilon(\text{ID track}) = \epsilon(\text{Medium-}\mu|\text{CT}) \times \epsilon(\text{ID}|\text{MS}) \quad (4.23)$$

where the last term of the second definition defines the inner detector track reconstruction efficiency, given the muon spectrometer observation. As per electron reconstruction, a scale factor is extracted to correct simulation reconstruction efficiencies to match that of data.

3. **Isolation:** Muon isolation [81] (between muons and other particles) is undertaken by two variables, one based on tracking and another based on the calorimeter. Signal muons are often produced isolated from other particles, except from those embedded in jets from semi-leptonic decay (e.g. top-quark decay). The track based variable is $p_T^{\text{varcone30}}$, the p_T sum of tracks ($p_T > 1$ GeV) in within a p_T dependent cone of size $\Delta R = \min(10 \text{ GeV}/p_T, 0.3)$, excluding the muon track. From the calorimeter based isolation, $E_T^{\text{topocone20}}$ are used as the sum of the transverse energy of topological clusters (as mentioned in the previous section) of $\Delta R = 0.2$, correcting for pileup and subtracting the energy deposit of the muon itself. The analysis listed in thesis use either “FCLoose” or “GradientLoose” isolation working points, as listed in table 4.6.

The fixed cut isolation working point is readily described in table 4.6, while the gradient working points (and hence the loose) choose selection values of $p_T^{\text{varcone30}}/p_T$, $E_T^{\text{topocone20}}/p_T$ such that the efficiency of the isolation is 90% for 25 GeV and 99% efficient for 60 GeV. Equivalently, the loose isolation has efficiencies of 95% and 99% at the respective p_T benchmarks, as summarised in table 4.6. *FCLoose isolation* is used for the analysis in [2](9).

Working point	Selection
FCLoose	$p_T^{\text{varcone30}}/p_T < 0.15, E_T^{\text{topocone20}}/p_T < 0.30$
Gradient	$p_T^{\text{varcone30}}/p_T, E_T^{\text{topocone20}}/p_T \text{ s.t. } \geq 90(99)\% \text{ eff. at } 25(60) \text{ GeV}$
GradientLoose	$p_T^{\text{varcone30}}/p_T, E_T^{\text{topocone20}}/p_T \text{ s.t. } \geq 95(99)\% \text{ eff. at } 25(60) \text{ GeV}$

Table 4.6: Isolation working points used in the analyses listed in this thesis, as per [81].

4.6.5 Jet vertex tagging (JVT)

This subsection outlines the implementation of the jet vertex tagger [82], which targets the suppression of pileup from various sources (both in time and out of time pileup). This uses information on the primary vertex (denoted PV_0) and additional pileup vertices (the k -th vertex denoted as PV_k). JVT is a two dimensional likelihood discriminant, based on two variables:

$$\text{corrJVF} = \frac{\sum_k p_T^{\text{trk}_k}[PV_0]}{\sum_L p_T^{\text{trk}_L}[PV_0] + H}; \quad H = \frac{\sum_{n \geq 1} \sum_L p_T^{\text{trk}_L}[PV_n]}{C \times n_{\text{trk}}^{\text{PU}}}, \quad (4.24)$$

where C is a constant based on the slope of the average pileup track p_T , but the classifier is insensitive to the choice of C , while $n_{\text{trk}}^{\text{PU}}$ denotes the total number of pileup tracks. The second variable is calculated as the sum of primary vertex associated track momenta, divided by the momentum of the calibrated jet:

$$R_{pT} = \frac{\sum_k p_T^{\text{trk}_k}(PV_0)}{p_T^{\text{jet}}} \quad (4.25)$$

These are incorporated in a two-dimensional likelihood, as discussed further in [82] to produce a classifier output for a jet, determining if a jet is either originating from pileup or is a jet associated to the underlying hard scatter event.

4.6.6 E_T^{miss}

The missing transverse energy [83] is calculated using contributions from electrons, muons, jets and optionally photons, as well as tracks not otherwise associated to a reconstructed object, and can be defined as

$$E_T^{\text{miss}} = -\sum_i p_T^{e_i} - \sum_k p_T^{\mu_k} - \sum_l p_T^{\gamma_l} - \sum_{\text{jets}} p_T^j - \sum_n p_T^{\text{track}_n} \quad (4.26)$$

The last term of which is referred to as the track soft term (TST). The definition of “selected” is defined as the respective objects as defined elsewhere in a given analysis (identification criteria are assumed), plus additional overlap removal as will be discussed further in this section. This may be configured such that the “leptons are treated as invisible” in this thesis, such that the lepton contributions are removed, and/or the photon contribution being neglected as appropriate. In addition the total transverse energy ($\sum E_T$) is defined as:

$$\sum E_T = \sum_{\text{selected electrons}} E_T^e + \sum_{\text{selected muons}} E_T^\mu + \sum_{\text{selected-}\gamma} E_T^\gamma + \sum_{\text{jets}} E_T^j + \sum_{\text{unused tracks}} E_T^{\text{track}} \quad (4.27)$$

The last term is referred to as the soft term similarly to the E_T^{miss} calculation.

However, the principal difference between a simple sum over final states of the p_T or E_T one must consider that there are possible overlaps between electrons/photons, muons and jets:

1. **Electron/photon and jet overlap:** If an electron overlaps with a jet in an event (i.e. inside the $R = 0.4$ jet cone), define the ratio κ_E :

$$\kappa_E = \frac{E_{e/\gamma}^{\text{EM}}}{E_{\text{jet}}^{\text{EM}}} \quad (4.28)$$

and preserve the jet if $\kappa_E > 0.5$.

2. **Muon and jet overlap:** if a muon overlaps [83] with a jet in an event (i.e. inside the $R = 0.4$ jet cone) have a substantive contribution depending on the origin of the overlap: either energy loss along the path of the muon through the calorimeter, or pileup jets, while jets produced from FSR of the muon need to be preserved. This involves ghost associating the muon to the jet (inserting the muon into the jet cluster as an object with infinitesimally small transverse momentum), and if it is reconstructed inside the topo-clusters, then it is considered overlapping with the jet. JVT [82] is required to mitigate jets from pileup and mitigate the mis-tag rate of muons. In order to select the correct class of events to be preserved, the following requirements are applied on the tracks associated to the jets and muons, to mitigate pileup and these overlaps (note these criteria are only applied to jets which overlap with muons).

- $p_{T,\text{track}}^\mu / p_{T,\text{track}}^{\text{jet}} > 0.8$
- $p_T^{\text{jet}} / p_{T,\text{track}}^\mu < 2$
- Number of primary tracks associated to the jet < 5

Jets passing these criteria are preserved, while muons are retained regardless. In order to retain the jets associated to final state radiation, say from FSR photon emission of a muon in the calorimeter that is not otherwise identified. The jets preserved under this criterion have the following requirements:

- Less than 3 tracks associated to the primary vertex (few charged particles in the jet shower).
- 90% of the jet energy is deposited in the electromagnetic calorimeter.
- $p_T^{\text{jet}}(\text{Calo Pre-sampler}) > 2.5 \text{ GeV}$, used to determine that the jet originates from final state radiation, rather than some later non-prompt process.
- A jet width $w_{\text{jet}} < 0.1$, defined as:

$$w_{\text{jet}} = \frac{\sum_k \Delta R_k p_{T,k}}{\sum_i p_{T,k}} \quad (4.29)$$

This is used to separate the photon from similar electromagnetic deposits in the calorimeter.

- $p_{T,\text{track}}^{\text{jet}} / p_{T,\text{track}}^\mu > 0.8$: the sum of transverse momenta of all jet-associated tracks in the jet is similar to the muon track.

3. **Track soft term (TST):** The last term in the E_T^{miss} definition [83] denotes the track soft term, calculated using the tracks which are not associated to any other reconstructed physics object, while applying a signal overlap requirement on the tracks, removing:

- Tracks with $\Delta R(\text{Track}, e/\gamma \text{cluster}) < 0.05$, excluding tracks associated to electrons/photons.
- Tracks with $\Delta R(\text{track}, \tau) < 0.2$, excluding tracks associated with τ -leptons.
- Tracks associated to muons.
- Tracks which are ghost associated to jets (see previous item for this prescription) which are not identified as pileup jets by the JVT.

4.7 Distributed computing and analysis using the WLCG

After reconstruction, calibration and configuration etc of simulated samples/data, the majority of the user level analysis is simply a set of skimming selections, vetoing events which lie outside of the regions of interest (such as for example datasets which have $N_\ell = 0$, $E_T^{\text{miss}} > 250$ GeV in the analyses listed in this thesis) based on the object definitions. This is skimmed from AOD (Analysis Object Data) to derived AODs (known as DAODs), which are both pre-selected on kinematic variables, slimmed to remove extraneous objects and thinned to remove ancilliary variables not of benefit to end-user analysis (such as extensive vertexing information, some generator level information etc), mostly such that user datasets can be processed on the WLCG on the order of one day. Further sub-selections are also applied at the user analysis, further reducing the number of events and information to the minimum required for the purposes of region definition, allowing for an entire dataset to be stored on a local storage element.

Statistics

5.1 Probability, statistics and fitting

This section will be a brief summary of the requisite statistical methods required during the course of physics analysis, particularly in the concept of maximal likelihood estimation and fitting. This will briefly summarise a few preliminaries of probability, then focus on the construction of the likelihood function and confidence interval construction, from [84] among other literature.

Preliminaries

In this thesis, the axiomatic notion of probability is defined as a function $P : \Omega \rightarrow \mathbb{R}$, where Ω defines a sample space, and F defines an event space (or set of events) made of subsets of Ω , and the axioms assumed are:

$$P(E_i) \geq 0, \forall E_i \in F \quad (5.1)$$

$$P(\Omega) = 1$$

$$\forall (E_n, \dots, E_1) \in F^n \subseteq \Omega : E_i \cup E_j = \emptyset,$$

In addition to the fundamental axioms of probability, the notion of conditional probability can be defined for two sets $\{A, B\} \in \Omega$ namely $P(A \text{ "given" } B)$:

$$P(A|B) := \frac{P(A \cap B)}{P(B)}. \quad (5.2)$$

Probability density functions are distributions $f : \mathbb{R}^n \rightarrow \mathbb{R}$ that are integrable over their domain of definition, such that $\int_{-\infty}^{\infty} f(\mathbf{x}) d\mathbf{x} := 1$. The notion of probability from such functions is defined (in the 1-D sense, readily extended to multiple dimensions) as:

$$P(x_a \leq x \leq x_b) = \int_{x_a}^{x_b} f(x) dx \quad (5.3)$$

for $[x_a, x_b] \in \mathbb{R}, x_a \leq x_b$. The cumulative distribution function (CDF) is also used, namely:

$$\text{CDF}(\mathbf{x}) = \int_{-\infty}^{\mathbf{x}} f(\mathbf{y}) d\mathbf{y}. \quad (5.4)$$

Bayes' theorem also posits the relation between conditional probabilities:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}, \quad (5.5)$$

with $P(A)$ being referred to as the prior probability while $P(A|B)$ is referred to as the “posterior” probability. This posterior probability has a useful notion, in that if the set A corresponds to real data events (say event yield over background), given a hypothesis A , we can infer the extent to which hypothesis A is true or false given the data yield over background.

In addition, in a real experiment, we have multiple random variables [84] $x_i, \dots, x_n$ and some nuisance parameters $\theta_1, \dots, \theta_m$ (e.g. detector systematics, normalisation parameters), thus an overall PDF would be referred to as a likelihood function:

$$L = f(x_1, \dots, x_n; \theta_1, \dots, \theta_m). \quad (5.6)$$

And equivalently for N independent measurements, this is extended to

$$L = \prod_{i=1}^N f(x_1^i, \dots, x_n^i; \theta_1^i, \dots, \theta_m^i). \quad (5.7)$$

Using Bayes' theorem, the posterior $P(\boldsymbol{\theta}|\mathbf{x})$ can be determined for an unknown set of parameters $\boldsymbol{\theta}$:

$$P(\boldsymbol{\theta}|\mathbf{x}) = \frac{L(\mathbf{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\int L(\mathbf{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}}. \quad (5.8)$$

the function $\pi(\boldsymbol{\theta})$ is referred to as the prior. There is freedom in the choice of prior ([84] covers several distinct priors), but for smooth priors, if L is peaked around the true values for $\boldsymbol{\theta}$, then the posterior is only weakly dependent on the choice of prior [85]. The prior that will be assumed throughout will be the uniform prior (such that the PDF is flat in $\mathbf{x}$).

Bayesian inference means that the most probable value of $\boldsymbol{\theta}$ can be taken as the central value for the parameter $\boldsymbol{\theta}$. A one dimensional central interval $[a, b], \hat{\theta} \in [a, b]$ can be determined to have a given probability:

$$\int_a^b P(\theta|x)d\theta = 1 - \alpha. \quad (5.9)$$

This value of α can be corresponding to either 0.317, 0.1, 0.05 depending on the choice of probability interval, and depending on the type of interval (one-sided or symmetric). Symmetric intervals would be of the form $b = \hat{\theta} + \delta, a = \hat{\theta} - \delta$, while asymmetric intervals would be either $[-\infty, a]$ or $[b, \infty]$ respectively.

5.1.0.1 Neyman confidence belt construction

Confidence intervals can also be extended into the concept of a confidence belt following the Neyman construction [85, 86]. Since this is best visualised in a simple example, consider the case of one nuisance parameter and one parameter of interest, with the PDF $P(x|\theta)$, and prior $P(\theta)$. If the form of prior is known (such as is the case in Bayesian interpretations), a region R in the $(x - \theta)$ plane can be determined such that

$$P(R) = \iint_R P(x|\theta)P(\theta)dx d\theta. \quad (5.10)$$

This can be visualised in the schematic in figure 5.1.

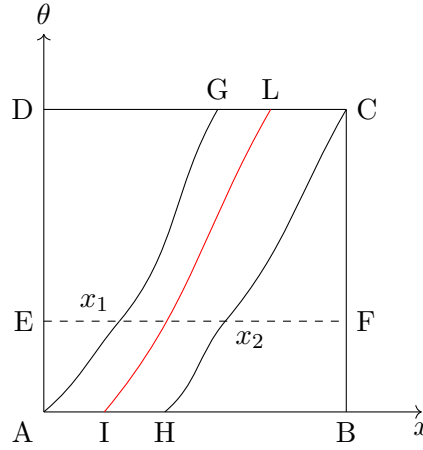


Figure 5.1: Neyman confidence belt construction in the simple model of one parameter of interest x and one nuisance parameter θ , which are allowed to vary between 0 and some values (B, D) . Curve $\bar{I}\bar{L}$ denotes the fitted value of $\hat{x}$ given a particular value θ or vice versa and is purely representative. The region R is defined as the region bounded by $\bar{A}\bar{G}$, $\bar{G}\bar{C}$, $\bar{C}\bar{H}$ and $\bar{H}\bar{A}$, which forms the confidence belt.

However, for some fixed value of θ (e.g. line $\bar{E}\bar{F}$), we can determine the x_1, x_2 , and some confidence interval α such that:

$$\int_{x_1}^{x_2} P(x|\theta)dx = \alpha \quad (5.11)$$

And, using the full posterior probability and the previous result:

$$P(R) = \iint_R P(\theta)P(x|\theta)dx d\theta = \int_A^D P(\theta) \left(\int_{x_1(\theta)}^{x_2(\theta)} dx \right) d\theta = \int_A^D \alpha P(\theta) d\theta = \alpha \quad (5.12)$$

Hence the belt definition is in essence independent of the prior $P(\theta)$. Alternatively, given a fixed value of x , the confidence interval can be determined in θ . It is then possible to construct the confidence belt through the iteration of several such fixed values of θ , to construct the curves $\bar{A}\bar{G}$ and $\bar{H}\bar{C}$. This construction allows us to construct a belt R corresponding to a confidence level of $1 - \alpha$, and can be extended to higher dimensions [85].

5.1.1 Profile Likelihood fitting

Maximal likelihood estimates

Following the definition from [85], the maximum likelihood estimate (MLE) $\hat{\theta}$ is the parameter θ which maximises the likelihood function $L(x; \theta)$ given some fixed x , under the assumption that the prior $P(\theta)$ is uniform. Notably, this MLE is also transitive: any MLE of a function of the parameter θ (say $F(\hat{\theta})$) is equivalent to the function as applied to the MLE of θ ($F(\hat{\theta})$). In the case of large N , for a specific subset of assumptions on regularity (we assume in all cases relevant here that these hold) [85], the MLE also can be found.

Relation to ATLAS implementation

Given a sample of N measurements $\mathbf{x} = (x_1, \dots, x_N)$ and nuisance parameters $(\theta_1, \dots, \theta_m)$, the likelihood is defined in 5.7, but in most cases the sample size itself is a random variable, which is referred to as the extended likelihood [84]

$$L(\mathbf{x}) = P(N; \boldsymbol{\theta}) \prod_{i=1}^N f(\mathbf{x}^i; \boldsymbol{\theta}). \quad (5.13)$$

This $P(N; \boldsymbol{\theta})$ is a poisson distribution, with the mean $\lambda(\boldsymbol{\theta})$

$$P(N; \boldsymbol{\theta}) = \frac{\lambda(\boldsymbol{\theta})^N e^{-\lambda(\boldsymbol{\theta})}}{N!}$$

In an ATLAS analysis use case, we consider the signal (s) and background event yields (b) as parameters (as well as the signal strength μ). We define a measurement of a particular variable in event e as x_e , and the relative mixture of signal and background (e.g. distributions of signal and background in a given variable, denoted by g_s and g_b) must be considered. All of these ingredients result in the following likelihood [87]:

$$L(\mathbf{x}; \mu, s, b) = \frac{(\mu s + b)^N e^{-(\mu s + b)}}{N!} \left[\prod_{k=1}^N \frac{\mu s g_s(x_k) + b g_b(x_k)}{(\mu s + b)} \right] \quad (5.14)$$

This is referred to as a “marked Poisson model”. Assuming the data is fixed (as are the signal and background yields in this case), the only free parameter is the signal strength μ . Thus we define the likelihood as a function of μ only. A useful re-formulation of the likelihood is the log likelihood (or its negative):

$$-\ln L(\mu) = -(\mu s + b) + \ln n! - \sum_{e=1}^N \ln[\mu s g_s(x_e) + b g_b(x_b)] \quad (5.15)$$

The MLE of μ can be found, assuming a uniform prior (hence $L(\boldsymbol{\mu}; \mathbf{x}, s, b)$ is identical to $L(\mathbf{x}; \mu)$):

$$-\left. \frac{\partial}{\partial \mu} \ln L(\mathbf{x}; \mu, s, b) \right|_{\mu=\hat{\mu}} = -\frac{1}{L(\mathbf{x}; \hat{\mu}, s, b)} \left. \frac{\partial L(\mathbf{x}; \mu, s, b)}{\partial \mu} \right|_{\mu=\hat{\mu}} = 0 \quad (5.16)$$

If $L(\hat{\mu})$ is never 0 (this can never be the case by definition of the likelihood), we can use the derivative of the log-likelihood to produce the maximum likelihood or best-fit estimate $\hat{\mu}$.

In the case of the physical likelihood (Equation 5.15), this log-likelihood can be expressed as

$$-\ln L(\mathbf{x}; \mu, s, b) = \mu s + b - \sum_{e=1}^N \ln(\mu s g_s(x_e) + b g_b(x_e)) + \ln N! \quad (5.17)$$

with the last term being a constant, hence irrelevant for the minimisation. With the free parameters in the model being μ the normalisation and the signal/background yields (s, b). In the case of a known resonance, μ would be set to 1 (as we expect the signal to exist), and the parameter s would be the “parameter of interest”. All of the other parameters in the model that are not directly related to the final measurement are considered nuisance parameters. These distributions can then be re-written to incorporate the impact of systematic variations, but the concept still holds (with total derivatives being replaced with the vector gradient).

5.1.2 Hypothesis testing

Hypothesis tests are statistical tests used to evaluate given hypotheses. In all of the practical cases discussed in this thesis, at most two hypotheses are considered:

- H_0 , the null hypothesis (e.g. data is consistent with SM background only)
- H_1 , the alternative hypothesis (e.g. data is consistent with SM background and SUSY signal)

In order to make decisions or evaluate the validity of these hypotheses, given the data in hand, a test statistic is chosen. There are some specific quantities of interest in the use of any such statistical test:

1. α : Significance level, or the probability to reject H_0 given H_0 is assumed to be true (known as *type I error*).
2. β : probability to reject H_1 given that H_1 is assumed to be true (known as a *type II error*). A metric $(1 - \beta)$ is known as the “power” of the test.
3. p -value is the probability, assuming H_0 is true, of getting a value of the test statistic at least as extreme as that observed.

Following the result of the Neyman Pearson lemma [84], the most powerful statistical test is based on the likelihood ratio:

$$\lambda(x) = \frac{L(x|H_1)}{L(x|H_0)} > k_\alpha \quad (5.18)$$

For some constant k_α which depends on the significance α . However, in most practical cases this exact function cannot be analytically defined. Thus, the LHC uses a modified version of the likelihood called the “profile-likelihood” test statistic. Assume the only parameters used are the signal strength (μ), observations are denoted $\mathbf{x}$, and nuisance parameters are denoted $\boldsymbol{\theta}$. The nuisance parameters are allowed to be functions of μ :

$$\lambda(\mu) = \frac{L(\mathbf{x}; \mu, \hat{\boldsymbol{\theta}}(\mu))}{L(\mathbf{x}; \hat{\mu}, \hat{\boldsymbol{\theta}}(\mu))}. \quad (5.19)$$

The numerator of this equation denotes the likelihood given:

$$\nabla_{\boldsymbol{\theta}} L(\mathbf{x}; \mu, \boldsymbol{\theta}(\mu)) \big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}^{\mu \text{ fixed}} = \mathbf{0}; L(\mathbf{x}; \mu, \hat{\boldsymbol{\theta}}(\mu)) = \sup L \quad (5.20)$$

while the denominator is the likelihood given that both the μ and $\boldsymbol{\theta}$ are fit simultaneously. This choice of statistic arises from Wilks’ theorem [84, 88] such that in the asymptotic limit (which will be assumed for all analyses in this thesis), $-2 \ln \lambda(\mu)$ is equivalent to a χ^2 distribution. This is guaranteed in the condition that H_1 and H_0 are nested, i.e that the nuisance parameters associated with H_0 are a strict subset of those associated with H_1 , which is assumed to be true in all cases here.

In the case of discovery of new physics, a search is undertaken assuming that the background-only hypothesis or null hypothesis is true, while ignoring the case where the background over-estimates the data.

$$q_0 = \begin{cases} -2 \ln \lambda(0), & \hat{\mu} \geq 0 \\ 0, & \hat{\mu} < 0 \end{cases} \quad (5.21)$$

In the case of an upper limit, the test statistic q_μ is:

$$q_\mu = \begin{cases} -2 \ln \lambda(\mu), & \hat{\mu} \leq \mu \\ 0, & \hat{\mu} > \mu \end{cases} \quad (5.22)$$

while in general, the test statistic $q_\mu = -2 \ln(\mu)$ is used without any specific requirements, with the explicit type of fit specifying the conditions where the test statistic may be 0 as above.

In the case of BSM searches, the parameter of interest is the signal strength (reflecting the existence of a signal above the SM background), while all other parameters (background normalisation etc as well as any systematic uncertainties), are considered nuisance parameters. To best outline this, the functional form of the profile likelihood is as follows: define a simulated signal yield in an SR as “s”, the total background yield as “b”, μ_s as the signal strength and a systematic θ , which may depend on μ_s . The initially-constructed likelihood function would then be of the form $L(\mathbf{x}; \mu_s, s, b, \theta)$, and as the parameter of interest is the signal strength, all other parameters are profiled, which means

$$\lambda(\mu_s) = \frac{L(\mathbf{x}; \mu_s, \hat{\hat{s}}, \hat{\hat{b}}, \hat{\hat{\theta}}(\mu_s))}{L(\mathbf{x}; \hat{\mu}_s, \hat{\hat{s}}, \hat{\hat{b}}, \hat{\hat{\theta}}(\mu_s))} \quad (5.23)$$

Then, principally in the case of discovery, the discovery type test statistic (upward fluctuations above background only) is used to evaluate the significance of the result given the null hypothesis only (and neglecting under-fluctuations of data). The p -value is then [85, 89]:

$$\int_p^\infty g(q_0) dq_0 = p_0 \quad (5.24)$$

and customarily, relating to a Gaussian distribution, the significance level Z can be determined from the unit Gaussian CDF Φ :

$$Z = \Phi^{-1}(1 - p_0) \quad (5.25)$$

If the test statistic follows a Gaussian distribution (either exactly, or approximately in large N limit), this simplifies somewhat to [89]

$$Z = \sqrt{q_0} \quad (5.26)$$

However, a form of this can also be derived in the case that the test-statistic follows the Poisson distribution, namely:

$$Z = \sqrt{2(s + b) \ln \left(1 + \frac{s}{b} \right) - 2s} \quad (5.27)$$

which in specific limiting cases can be simplified further into a form:

$$Z \approx \frac{s}{\sqrt{b}} \quad (5.28)$$

This last equation is often used as a simple significance estimator in some ATLAS use cases.

5.1.3 Limit setting

In a search for new signal [84], a counting experiment can be undertaken, comparing a set of data events with the total signal yield s (with a signal strength μ , and background yield b), let H_0 be the hypothesis “ $s = 0$ ” be the expected scenario, and H_1 the alternative hypothesis: “ $s + b$ ” be the expected case. For the purposes of this example, the background uncertainties are neglected. The likelihood would be in this case (as before, and following the logic discussed in [84]):

$$L(n; s, b) = \frac{(s + b)^n}{n!} e^{-(s+b)} \quad (5.29)$$

In the case of a bayesian upper limit interval $[0, s^{\text{up}}]$ for the signal, given some confidence level α , the following relation can be defined

$$1 - \alpha = \int_0^{s^{\text{up}}} P(s|n) ds = \frac{\int_0^{s^{\text{up}}} L(n; b, s) \pi(s) ds}{\int_0^\infty L(n; b, s) \pi(s) ds} \quad (5.30)$$

In the choice of uniform prior $\pi(s) = 1$, this resolves to:

$$\alpha = e^{-s^{\text{up}}} \frac{\sum_{m=0}^n \frac{(s^{\text{up}}+b)^m}{m!}}{\sum_{m=0}^n \frac{b^m}{m!}} \quad (5.31)$$

This gives an upper limit using the bayesian statistics.

As an equivalent using the frequentist upper limits, the likelihood function can be defined for some fixed value n (say $n = 0, b = 0$), such that:

$$L(n = 0; s) = \text{Po}(0; s) = e^{-s} \quad (5.32)$$

Inverting the Neyman limit construction for $n=0$ to produce an upper limit allows the determination:

$$P(n \leq 0; s^{\text{up}}) = P(n = 0; s^{\text{up}}) = \alpha \implies s^{\text{up}} = -\ln \alpha \quad (5.33)$$

This can similarly be performed for a given value of n and b , inverting the Neyman belt and determining the threshold value. These numerically coincide with the values determined by the Bayesian upper limit.

In statistics, there are some cases of interest where a single composite test may be used for determining both an upper limit or a central value of a given parameter, depending on the data used. This cannot be readily undertaken using the Neyman limit construction, due to a problem known as flip-flopping (no smooth transition between the a central value estimation and an upper limit). This can be best seen in [90] figure 4. This is fixed using a “Feldman-Cousins” confidence interval, namely using an interval on a given parameter μ with confidence level α [90]:

$$R_\mu = \frac{L(\mathbf{x}; \mu)}{L(\mathbf{x}; \hat{\mu})} > k_\alpha \quad (5.34)$$

where k_α denotes a constant dependent on the confidence level as per the Neyman construction. A possible such set of limits can be seen in figure 5.2. However, these intervals generally require numerical computation, even in the simplest cases.

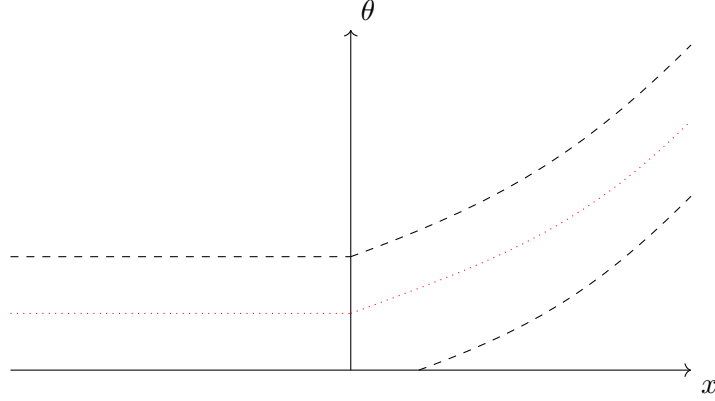


Figure 5.2: Composite upper limits and limit interval determined by Feldman Cousins approach (dashed lines). A line showing the best fit value of θ as a function of x is also plotted (red dotted line).

5.1.3.1 CL_S exclusion limits

In the case of no significant excess, a modified frequentist approach is proposed, based on Higgs searches, namely the use of CL_S [91]. Following on from the definition in [84, 91], the p-values corresponding to the null and alternate hypotheses (H_0, H_1) are calculated, namely:

$$p_{s+b} = P(\lambda(\mathbf{x}|H_1) \geq \lambda^{\text{obs}}) \quad (5.35)$$

$$p_b = P(\lambda(\mathbf{x}|H_0) \leq \lambda^{\text{obs}}) \quad (5.36)$$

Instead of the frequentist upper limit $p_{s+b} \leq \alpha$, the CL_s method uses a quantity:

$$CL_s := \frac{p_{s+b}}{1 - p_b} \quad (5.37)$$

Since often toy models are used, this can be approximated by using the number of pseudo experiments satisfying the same condition.

$$CL_s = \frac{N(\lambda_{s+b} \geq \lambda^{\text{obs}})}{N(\lambda_b \geq \lambda^{\text{obs}})} \quad (5.38)$$

and in the specific case of a SUSY search, it can be written in terms of a counting experiment, using the event count as the test statistic:

$$CL_s = \frac{P(n \leq n^{\text{obs}} | \mu = 1)}{P(n \leq n^{\text{obs}})} \quad (5.39)$$

The “ CL_s ” limit is then defined for some “confidence interval” α as:

$$CL_s \leq \alpha \quad (5.40)$$

instead of the fully frequentist or bayesian upper limits.

5.2 Statistical use case: ATLAS specific analysis design in context of SUSY searches

The analyses listed in this thesis use a package called HISTFITTER [92], which is a wrapper configuration on top of HISTFACTORY [87], ROOSTATS[93] and ROOFIT[94]. In this section, the usage of this package will be outlined in terms of analysis design, then the likelihood used will be outlined including systematic treatment.

5.2.1 Analysis configuration

ATLAS uses the concept of three types of regions in the configuration of analyses and fits:

1. **Signal region (SR):** search region where some given signal model is expected to be present (such as supersymmetry). These are optimised using simulated signals and simulated background, without looking at data.
2. **Control region (CR):** regions where no signal is expected by design, in order to correct the normalisation of standard model backgrounds that arise in the signal region to match the dataset. These need to be strictly non-overlapping with the signal regions (this is often referred to as “orthogonal” to the signal region).

Control regions should have similar kinematic profiles (at least in the given background of interest to normalise) to the signal region, avoiding extrapolation over variables that are not well modelled in simulation. In order to get an accurate estimate of the normalisations of the given sample, the purity of the sample must be maximised (which can also reduce correlations with normalisations of other backgrounds).

3. **Validation region (VR):** regions with negligible signal contamination that share a similar kinematic profile with the signal region. These are used to check that the normalisations of the standard model backgrounds are accurate given a region similar to the signal region. These are also required to be non-overlapping with either the control or signal regions

Figure 5.3 outlines a typical configuration of control, signal and validation regions.

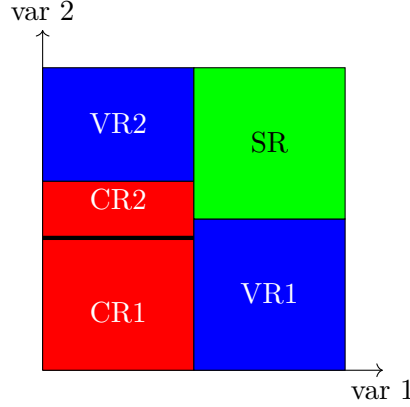


Figure 5.3: Schematic diagram of the control, validation and signal regions configured in two arbitrary variables. This particular case covers an analysis with one signal region, two control region and two sideband validation regions corresponding to those backgrounds. An additional note must be made that these validation regions, control regions etc need not be strictly adjacent in any phase space.

5.2.2 Likelihood and fits used

This is configured as a start point using a binned likelihood function, which is an extension of the likelihood function described earlier. This incorporates the fact that some of the control regions may be binned histograms rather than single histograms (which occurs several times in this thesis). This binning criterion means we can define bins in a given control or signal region, such that the total signal yield s and total background yield b can be defined in the marked poisson model as [87]:

$$L(\mathbf{x}; \mu_s, s, b) = \frac{(\mu_s s + b)^N e^{-(\mu_s s + b)}}{N!} \left[\prod_{k=1}^N \frac{\mu_s s g_s(x_k) + b g_b(x_k)}{(\mu_s s + b)} \right] \quad (5.41)$$

where $g_{\{s,b\}}(x_k)$ denotes the histogram shape for either signal/background, thus an individual bin content for either signal (ν^{sig}) or background (ν^{bkg}) can be related to the shape via the bin width Δ_{bin_k} :

$$g_s(x_k) = \frac{\nu_{\text{bin}_k}^{\text{sig}}}{s \Delta_{\text{bin}_k}}, \quad g_b(x_k) = \frac{\nu_{\text{bin}_k}^{\text{bkg}}}{b \Delta_{\text{bin}_k}}. \quad (5.42)$$

Thus, the likelihood could be reformulated as a binned likelihood [87], where $\mathbf{n}$ outlines the number of data events in a given histogram bin (a function of $\mathbf{x}$):

$$L(\mathbf{n}; \mu) = \text{Po}(n^{\text{total}} | \mu_s s + b) \left[\prod_{c \in \text{bins}} \frac{\mu_s \nu_c^{\text{sig}} + \nu_c^{\text{bkg}}}{\mu_s s + b} \right] \propto \prod_{c \in \text{bins}} \text{Po}(n_c; \mu_s \nu_c^{\text{sig}} + \nu_c^{\text{bkg}}). \quad (5.43)$$

Using the binned likelihood defined as above, and incorporating an additional systematic term $C(\boldsymbol{\theta})$ [92]

$$L(\mathbf{n}; \mu_s, b, \boldsymbol{\theta}) = P(n_{SR} | \nu(\mu_s, b, \boldsymbol{\theta})) \times \prod_{i \in CR} P(n_i | \nu(\mu_s, b, \boldsymbol{\theta})) \times C(\boldsymbol{\theta}) \quad (5.44)$$

This last term is a constraint term based on the systematics and nuisance parameters only, and is typically a multi-dimensional Gaussian. The function $\nu(\mu_s, b, \theta)$ encodes the information from the binning, the yields in each bin, and the nuisance parameters. These nuisance parameters can be divided into “useful parameters” such as the normalisation parameters (typically also labelled μ_{bkg}) of the simulated SM background, and “true nuisance parameters” (systematics, statistical uncertainties in the signal and control regions).

5.2.3 Systematic handling in HistFitter

This subsection outlines the systematic handling in HistFitter, of the nuisance parameters corresponding to systematic uncertainties. Histograms corresponding to the up/down variations of systematics are used, alongside the nominal variation. However, it must be noted that these up/down variations refer to the original systematic configuration, not the effect, such that the upper and lower limits may be above the nominal. These variations may also be correlated *a priori*. In order to account for all of the effects of systematics, a parametrisation $\eta(\theta)$ is chosen, where θ denotes the nuisance parameters corresponding to systematics, with the parametrisation mapping the up and down variations to parameters varying between +1 and -1 respectively. HistFactory has two main options for this parametrisation, one is listed primarily to demonstrate a simple parametrisation (piecewise linear), while the second more complex implementation forms the primary one used in the ATLAS searches discussed in this thesis.

Definition 1 *Piecewise linear interpolation of systematics*

Let [87] $\eta(\alpha)$ be the parametrised normalisation uncertainty as a function of the set of nuisance parameters for each systematic α , and for a given systematic p , the up variation is defined as η_p^+ , the nominal is defined as η_p^0 , and the down variation as η_p^- , and the corresponding nuisance parameter as α_p . The parametrisation requires that for a given variation $\eta(\alpha_p = 1) = \eta_p^+$, $\eta(\alpha_p = -1) = \eta_p^-$ and $\eta(\alpha_p = 0) = \eta_p^0$.

Piecewise-linear interpolation is defined by the relations:

$$\eta(\alpha) := 1 + \sum_{p \in \text{Systs}} F(\alpha_p; 1, \eta_p^+, \eta_p^-) \quad (5.45)$$

$$F(\alpha_p, I^0, I^+, I^-) := \begin{cases} \alpha_p(I^+ - I^0) & \text{if } \alpha \geq 0 \\ \alpha_p(I^0 - I^-) & \text{otherwise} \end{cases} \quad (5.46)$$

Definition 2 *Piecewise exponential/polynomial interpolation of systematics*

Use all of the definitions of systematics, η as per definition 1. Boundary conditions are applied to $\eta(\alpha = \pm 1)$ as before, but there is an additional requirement of $\frac{\partial \eta}{\partial \alpha_p}|_{\alpha=\pm 1}$ and $\frac{\partial^2 \eta}{\partial \alpha_p^2}|_{\alpha=\pm 1}$ being defined such that the function is both differentiable and continuous up to the second derivative. The parametrised uncertainty is defined as [87]:

$$\eta(\alpha) := \prod_{p \in \text{Systs}} H(\alpha_p; 1, \eta_p^+, \eta_p^-, 1) \quad (5.47)$$

$$H(\alpha_p; 1, \eta_p^+, \eta_p^-, 1) = \begin{cases} (\eta_p^+)^{\alpha_p} & \alpha \geq 1 \\ 1 + \sum_{i=1}^6 c_i (\alpha_p)^i & |\alpha| < 1 \\ (\eta_p^-)^{-\alpha} & \alpha < -1 \end{cases} \quad (5.48)$$

This equation forms the definition of the HistFitter “OverallSys” configuration of systematics used by ATLAS. The free parameters c_i are chosen such that the boundary conditions on the derivatives match.

5.2.4 Multi-bin fits

Multi-bin fits make use of the binned likelihood to extract a single parameter for a distribution, whereas in a one-bin fit, only the total normalisation is known. If a sufficiently adequate binning choice is chosen, then this may cover defects in the shapes of the kinematic distributions in the simulation/data or better reflect a sample over a given set of kinematic categorisations. This can be best motivated visually, namely by figure 5.4.

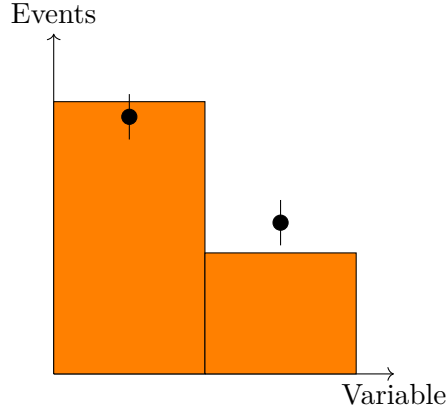


Figure 5.4: A schematic diagram of a binned control region. The background corresponds to the orange histogram, with data points overlaid. The total normalisation μ_{bkg} of this region considered alone (if a single bin) may exacerbate differences in shape between data and simulation, whereas a multi-bin fit should better reflect this shape discrepancy between bins.

5.2.5 Fit types

This section outlines the key configuration differences used in HistFitter for different stages of analysis, namely the fit definition/validation, derivation of exclusion limits where appropriate and hypothesis testing based on the background-only hypothesis. The three fit types considered are [92]:

1. Background only fit: performed primarily for validation of the backgrounds in the signal and validation regions. The constraints on the parameters in the PDF are only applied to the control regions. Given these parameters, an estimate of the background

contribution in the validation and signal regions can be inferred (including some unconstrained systematics which may apply to the signal/validation regions specifically). This fit type is entirely signal model independent, with no direct inclusion of the signal region except for an output of the fit.

2. Model-dependent signal fit: Assuming no significant excess is observed in the background-only fit, a parameter of interest is included in the likelihood, namely μ_s for a given signal model (say a particular SUSY signal point corresponding to some specific mass parameter). CL_s limits on this particular parameter can be deduced from this fit type. Note unlike the background only fit, additional constraints can occur on nuisance parameters since the signal regions are included in the fit (alongside systematics which are attached to the signal regions such as signal theory systematics).
3. Model-independent signal fit: This fit type includes SR and CRs, but only one SR at a time per fit. A dummy signal with a yield of 1 event is included in the given signal region, and the normalisation parameter μ_s is inserted into the fit as a free parameter. This dummy signal model allows for upper limits to be established on any other signal model, given the actual observed dataset and background yields. This fit type also covers the hypothesis testing of the background-only hypothesis using the test statistic for discovery mentioned earlier, allowing for the extraction of p -values and significances of any excesses, extracting confidence levels associated to the background-only hypothesis.

5.3 Multi-variate analysis and machine learning

In order to discriminate between topologies, there is the possibility to use physical techniques, such as new variable definitions and discriminants with cut and count, or through the application of machine learning. Machine Learning (ML) works through self-improving (or "learning") algorithms, as a class of multi-variate analysis strategy, where input variables are specified, and the output is a classification (for example, saying if a jet is associated with a top quark or not) mapped onto a scale $[-1,1]$ or $[0,1]$ [95].

ML has a variety of different approaches available, such as the Neural Network (NN), which is a series of nodes (or neurons) connected together, with some back propagation for learning, or Boosted Decision Trees (BDT).

Definitions

Definitions used throughout the sub-section:

- β , a free parameter chosen by the user.
- $\mathbf{x}$: Input ntuple, with $\mathbf{x}_i$ denoting the i^{th} data point in an ntuple. Each point is a vector with entries corresponding to the relevant event variables.
- $y(\mathbf{x}_i)$ denotes if the data point in the training set truly is a signal (+1) or a background (-1).

- $T_i(\mathbf{x}_j)$ is the classification for the i^{th} tree for the j^{th} data point. Each individual tree can only have the values -1 or 1 respectively. A tree classifier is depicted in figure 5.5.
- $T(\mathbf{x}_j) \in [-1, 1]$ denotes the final output classification for the data point j after boosting.
- w_i denotes the weighting applied to each tree in the boosting. $\sum_i w_i = 1$.
- $\text{Indicator}(y = x) := 1$ if $y = x$, vanishes otherwise.

Decision Tree Implementation

Decision trees are classifiers based on various pass or fail conditions, with the head being the input or trunk, and the output branches known as leaves. A diagrammatic representation can be shown in figure 5.5. However, an individual tree may not correctly classify the objects

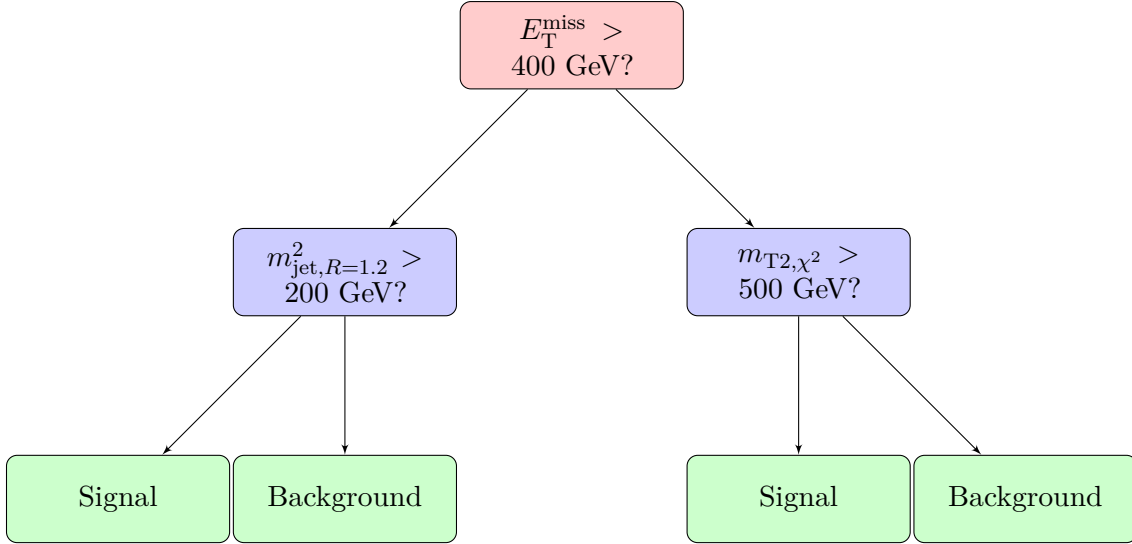


Figure 5.5: A schematic diagram of a single decision tree.

in the first pass, or lack sufficient stability in converging to the correct values, thus we use a random forest of many trees, with weights applied (boostings) to the misclassified results. This is the definition of the BDT method, a weighted sum over classifiers $T(\mathbf{x}_j)$:

$$T(\mathbf{x}_j) = \sum_i w_i T_i(\mathbf{x}_j). \quad (5.49)$$

To retrieve the correct weightings, the BDT must first be trained on a training sample of known signal and background proportions, then evaluated on a separate sample, known as a test sample. This will be done by optimising a chosen set of signal points and all of the background samples, initially trained, by splitting the samples in half and training on half the signal and half the background (randomly selected with no replacement) and the other half being allocated as the test set. The weights are then constructed on the training set (the trees are said to have been "grown"), and the testing set is for end user classification performance. Events in this set are passed through the decision trees that have already

been constructed. Once satisfied with the classification performance on the test sample, the classifier is then applied to the test set or to data not included in either set.

Boost weighting is non-universal, and several boost methods have been implemented, namely Adaptive Boosting (AdaBoost)[96], Gradient Boosting[97] and Bagged Boosting[98]. These were implemented using the ROOT Multivariate Analysis package: TMVA[95].

AdaBoost

The error of the m^{th} tree is defined [95][96] as:

$$\text{err}_m := \frac{\sum_{i=1}^N (w_i \times \text{Indicator}(y_i \neq T(x_i)))}{\sum_{i=1}^N w_i}, \quad (5.50)$$

with the weight α defined as:

$$\alpha_m := \beta \times \ln \left(\frac{1 - \text{err}_m}{\text{err}_m} \right). \quad (5.51)$$

The weights must then be changed and renormalised:

$$w_i \rightarrow w_i \times \exp[\alpha_m \times \text{Indicator}(y_i \neq T(x_i))], \quad \Rightarrow \quad w_i \rightarrow \frac{w_i}{\sum_{i=1}^N w_i}. \quad (5.52)$$

Thus, the final score for an event is

$$T(\mathbf{x}_j) = \sum_{m=1}^M \alpha_m T_m(\mathbf{x}_j) \quad (5.53)$$

Gradient Boost

Gradient boost[95][97], uses a set of weak learners $T(\mathbf{x}, \alpha_m)$, such as our decision trees, with α_m, β_m denoting free parameters of the trees, thus the weighted sum of trees is parametrised with a set of free parameters P such that:

$$T(\mathbf{x}_j, P) = \sum_{m=0}^M \beta_m T_m(\mathbf{x}_j, \alpha_m); \quad P \in \{\beta_m, \alpha_m\}_0^M \quad (5.54)$$

The algorithm then requires the choice of such parameters in P that the binomial log-loss function:

$$L(T(\mathbf{x}_j), y) = \ln \left(1 + e^{-2T(\mathbf{x}_j)y(\mathbf{x}_j)} \right) \quad (5.55)$$

is minimised, using the method of steepest descent (hence the name "Gradient Boost"). This is calculated by first calculating the initial gradient of the loss function, then producing a regression tree with leaves with values adjusted to be the mean gradient in each region defined by the tree, and then iterated over. However, the method is applicable to any reasonable (differentiable) loss function, such as the squared error

$$L(T(\mathbf{x}_j), y(\mathbf{x}_j)) = (T(\mathbf{x}_j) - y(\mathbf{x}_j))^2. \quad (5.56)$$

Bagged boost [\[95\]](#)

Despite the name, it is not precisely a method of boosting (weighting), instead it is a resampling method, namely sampling events in the training set with replacement. This is intended to stabilise the response of a classifier, and as a sampling method it is likely able to be combined with other boosting methods.

Moving beyond the 2015-16 analysis.

As discussed in 3.4.3, SUSY particles and their associated decay chains can result in a variety of final states. In this case we consider the principal model of RPC pair produced $\tilde{t}_1 \rightarrow t + \tilde{\chi}_1^0$ as the primary model of interest. Some variations may also be discussed equivalently, such as the three or four body decay modes $\tilde{t}_1 \rightarrow bW/bjj' + \tilde{\chi}_1^0$. Following this, the top pairs themselves are unstable and will decay into a variety of distinct final states: differing numbers of jets/b-jets, final state electrons/muons and E_T^{miss} . Since each class of decay has a slightly different background profile and event kinematics, typically ATLAS sub-divides the searches for direct stop production into teams by the number of leptons in the final state. These analyses cover the stops decaying into semi-leptonic top pairs [99][22], stops decaying into leptonic top pairs [23], and longer decay chains where the stop decays to a neutralino via the emission of a Z/H [24] as discussed in 3.4.3.

The author contributed to the search for the stop decays in the all-hadronic channel for simplified models with a branching ratio $BR(\tilde{t} \rightarrow t\tilde{\chi}^0)_1 = 100\%$. This channel has the primary benefit that the principal E_T^{miss} source are the neutralinos themselves, but with the limitation that jet and E_T^{miss} observables are substantially less precise than those used in the identification of electrons/muons, particularly where it concerns the identification of a top quark. This chapter forms a summary of the analysis listed in [3] covering ATLAS data periods of years 2015-2016 with an integrated luminosity of 36.1fb^{-1} with a centre-of-mass energy of 13 TeV. The author primarily contributed through the extraction of QCD scale factors to this publication.

Since it is often customary for ATLAS searches to expand upon their previous iterations as a starting point where possible, the motivations for the signal regions and control regions will be discussed here, since these same themes will again return in chapter 9. Chapter 9 covers the same class of signal models for a luminosity of 139fb^{-1} covering the ATLAS data periods of 2015-2018. In addition, chapters 7 and 8 will cover the development undertaken by the author between the two analyses.

6.1 Analysis configuration

6.1.1 Key backgrounds

A major limitation in SUSY searches is that we cannot directly detect SUSY particles in any way, merely inferring their existence from excesses over standard model backgrounds. In such case, there are multiple standard model processes which can overlap with that of the given signal model (at least insofar as the final state can be concerned). These processes can be separated into two classes, namely *reducible* backgrounds, and *irreducible* backgrounds. Reducible backgrounds by their definition are in some way distinct from the SUSY signal, such that there is a variable that may discriminate between that background and the given signal. Irreducible backgrounds, on the other hand. In a fictional perfect analysis, reducible backgrounds would be removed entirely, while a SUSY signal would be resolved as an excess over the irreducible background. The processes that produce final states equivalent to our signal are as follows:

- $Z \rightarrow \nu\nu + \text{jets/b-jets}$. The jets/b-jets either arise from the hard scatter event or gluons radiated from the event (either in the initial/final states).
- Top pair production ($t\bar{t}$), principally the case where $t\bar{t}$ decays semi-leptonically, either with an electron or muon out of acceptance or a hadronically decaying τ . The case of all-hadronic $t\bar{t}$ is principally covered by the multi-jet background, and is generally small.
- $t\bar{t} + W/Z$: The principal dominating process of this background is $t\bar{t} + Z(\rightarrow \nu\nu)$, which is irreducible compared to the SUSY signal. For completeness, $t\bar{t} + W(\rightarrow \ell\nu)$ also contributes, but the overall impact of this background is much smaller than the $t\bar{t} + Z$.
- $W \rightarrow \ell\nu + \text{jets/b-jets}$ (where ℓ is out of acceptance or is a hadronic τ). This background is reducible, and is generally sub-dominant except in signal region SRB.
- Single top processes (e.g. tW, tZq etc). This reducible background is generally small except for SRA/SRB.
- Multi-jet (often referred to in ATLAS searches as “QCD”), with the E_T^{miss} arising from jet mismeasurement. This is principally only relevant in the compressed region (SRC).
- Di-boson backgrounds ($ZZ + \text{jets}, WW + \text{jets}$). Very small background contribution, principally from $ZZ \rightarrow \nu\nu + \text{jets}$ and $WW \rightarrow \ell\nu + \text{jets}$ (with a lepton out of acceptance), but generally almost negligible.

6.1.2 Key variables and discriminants

Since most of the backgrounds above are reducible, kinematic selections have been applied on variables which exploit the differences in kinematics between the signal and background. The explicit values are usually selected based on a binned distribution of the given variable of interest, optimising for maximal expected signal significance. This section discusses the variables used in [3], which were also retained for [2], and is as follows:

Core components

The E_T^{miss} , number of jets, b-jets, jet- p_T and number of electrons/muons are the principal variables required for triggering and selecting the correct event kinematics.

m_{T2, χ^2}

The stransverse mass m_{T2} [3, 100, 101] is calculated from top candidates based on $R = 0.4$ jets. Candidates are selected through the minimising of the penalty function $\chi^2 = (m_{\text{cand}} - m_{\text{true}})^2 / m_{\text{true}}$, where m_{true} is set to either the m_W or m_t respectively (in this case 80.4 GeV and 173.2 GeV) depending on the identification. The reconstruction is performed first by identifying W candidates using the minimisation then combined with additional b-jets to form two top quark candidates.

Re-clustered jet masses

The calibrated and configured $R=0.4$ anti- k_T jets are used as inputs to reclustering using the k_t algorithm into jets with either $R=0.8$ or $R=1.2$ respectively. These jet radii are chosen to cover the reconstruction of boosted W or boosted top quarks respectively, and as will be discussed later, the masses of these is considered to be comparable to the actual masses of the W/top.

$$\min[\Delta\phi(\mathbf{j}_{1-4}, \mathbf{E}_T^{\text{miss}})], \min[\Delta\phi(\mathbf{j}_{1-2}, \mathbf{E}_T^{\text{miss}})] \text{ and } \mathbf{E}_T^{\text{miss, track}}$$

These selections are applied to reduce the multi-jet background by vetoing events where the E_T^{miss} principally arises from jet mis-measurement. These are based on the minimum $\Delta\phi$ between the relevant numbers of p_T ordered jets and the $\mathbf{p}_T^{\text{miss}}$ vector (with $\min[\Delta\phi(j_{1-4}, E_T^{\text{miss}})]$ selections a strict subset of the $\min[\Delta\phi(j_{1-2}, E_T^{\text{miss}})]$ selections).

$m_T^{b, \min}$ and $m_T^{b, \max}$

Transverse mass calculated from the E_T^{miss} and the b-jet which is either closest (furthest) in $\Delta\phi$ to the $\mathbf{p}_T^{\text{miss}}$.

$$m_T^{b, \min} = \min_{\Delta\phi(b, E_T^{\text{miss}})} \sqrt{2p_T^b E_T^{\text{miss}} [1 - \cos \Delta\phi(b, \mathbf{p}_T^{\text{miss}})]} \quad (6.1)$$

$$m_T^{b, \max} = \max_{\Delta\phi(b, E_T^{\text{miss}})} \sqrt{2p_T^b E_T^{\text{miss}} [1 - \cos \Delta\phi(b, \mathbf{p}_T^{\text{miss}})]} \quad (6.2)$$

The former of which is used to mitigate the $t\bar{t}$ background, since the upper limit of the transverse mass is expected to be smaller than the top mass. This applies particularly in the case of the semi-leptonic $t\bar{t}$ with a soft lepton, E_T^{miss} and b-jets on one branch. This means that the majority of the $t\bar{t}$ events should have $m_T^{b, \min} < 200$ GeV. $m_T^{b, \max}$ is principally used to aid in overall signal/background discriminations.

$\Delta R(b_1, b_2)$

The ΔR between the leading and the sub-leading b-jets. This helps with additional discrimination of signal against backgrounds where b-jets originate from gluon splitting (which, as will be discussed later, covers specific Z+jets processes, and to some extent W+jets). This specifically covers the ΔR between the leading and sub-leading b-jets. A study into the composition of the $Z + \text{jets}$ background was undertaken in section 8.3.

Recursive jigsaw reconstruction

This reconstruction technique [102] is used to separate a SUSY signal event recoiling against an ISR system into two components: ISR and the system associated to the SUSY decays (referred to as the *sparticle hemisphere*). This SUSY decay sector is then subdivided into two components: visible and invisible objects associated to the sparticle decay.

Since ISR objects can be defined using this method, angles between this system and the E_T^{miss} can be constructed as well as the p_T of the ISR to ensure that the system is recoiling against the ISR and is approximately anti-collinear. In addition, additional selections can be undertaken based on observables associated only to the sparticle hemisphere, such as the number of jets and b-tagged jets. In addition, the transverse mass of the visible sparticle system and the E_T^{miss} (so in essence, all components that do not include the ISR) is used (this is denoted as m_s), alongside the invariant mass of all the visible objects associated to the sparticle hemisphere m_V .

6.2 Signal Regions (2015-16)

6.2.1 Preselection

A preselection is selected for all signal regions, defining primarily the trigger and some multi-jet/hadronic $t\bar{t}$ veto selections. This is outlined in table 6.1.

Variable/Region	Presel-0L-2015/16
Trigger	E_T^{miss}
E_T^{miss}	$> 250 \text{ GeV}$
N_j	≥ 4
Jet- p_T	$> 80 \text{ GeV}, > 80 \text{ GeV}, > 40 \text{ GeV}, > 40 \text{ GeV}$
N_b	≥ 1
$\min[\Delta\phi(j_{1-2}, E_T^{\text{miss}})]$	> 0.4
$E_T^{\text{miss,track}}$	$> 30 \text{ GeV}$
$\Delta\phi(E_T^{\text{miss}}, E_T^{\text{miss,track}})$	$\pi/3$

Table 6.1: Table of preselections for every signal region in [3].

6.2.2 Signal Region SRA and SRB (2015-16)

Signal regions SRA and SRB [3] target two body $\tilde{t} \rightarrow t + \tilde{\chi}^0$ signal models with $\Delta(m_{\tilde{t}}, m_{\tilde{\chi}^0}) \gg m_t$ or $\Delta(m_{\tilde{t}}, m_{\tilde{\chi}^0}) > m_t$ respectively, corresponding to the “boosted” and “intermediate” selections respectively. The benchmark signal used for the optimisation of SRA is $(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (1000, 1)$ GeV and likewise $(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (600, 300)$ GeV for SRB.

A requirement on both signal regions is set to have at least 2 b-jets, at least four jets with $p_T > 20$ GeV, $\min[\Delta\phi(j_{1-3}, E_T^{\text{miss}})] > 0.4$.

Since the jets associated to the $t\bar{t}$ system may overlap in (η, ϕ) , re-clustering is used with radius of either $R=0.8$ or $R=1.2$ as discussed in section 4.6.1. The mass cuts on either the leading or sub-leading (in p_T) of such reconstructed jets are used to reconstruct either a top or a hadronic W system. A top is identified if the $R=1.2$ reclustered jet mass is > 120 GeV, a W if $m \in [60, 120]$ GeV, and is not classified otherwise. Top categories are applied to the signal regions, namely that there is a “TT” region which denotes two tops, “TW” and “T0” similarly categorised in order of the leading/sub-leading $R=1.2$ reclustered jets. In addition, a selection on the leading $R=0.8$ reclustered jet of $M(J_1; R=0.8) > 60$ GeV is required such that there is a boosted jet that is at least consistent with the W mass in the event.

The $t\bar{t}$ background is substantially mitigated through a threshold selection on the $m_T^{b,\min} > 200$ GeV. Another important variable in addition to the reclustered mass categories is the m_{T2,χ^2} . To offset the effect of an intermediate state $W^- \rightarrow \tau^- \bar{\nu}$ (or its charge conjugate), a tau veto is used, namely if any non b-tagged jet in the event has at most four tracks, and the aforementioned jet has $\Delta\phi(E_T^{\text{miss}}, \text{jet}) < \pi/5$, the event is vetoed. In SRB, there is the inclusion of $m_T^{b,\max}$ and $\Delta R(b_1, b_2)$. Table 6.2 outlines the selections, and figure 6.1 outlines the kinematics of a signal event. In addition to the case presented there, additional diagrams exist which include the case where top decay which decays leptonically, with the soft electron/muon not reconstructed or via a decay with a hadronically decaying τ .

Signal Region	Variable/Category	TT	TW	T0
	Preselection $m_{\text{jet}, R=1.2}^1$ $m_T^{b,\min}$ N_b τ -veto $\min[\Delta\phi(j_{1-3}, E_T^{\text{miss}})]$ $m_{\text{jet}, R=1.2}^2$	0L Preselection (Table 6.1) > 120 GeV > 200 GeV ≥ 2 $\checkmark$ > 0.4		
		> 120 GeV	$[60, 120]$ GeV	< 60 GeV
SRA	m_{T2,χ^2} E_T^{miss} $\Delta R(b_1, b_2)$ $m_{\text{jet}, R=0.8}^1$	> 400 GeV > 400 GeV > 1	> 500 GeV > 500 GeV -	> 500 GeV > 550 GeV
SRB	$m_T^{b,\max}$ $\Delta R(b_1, b_2)$	> 200 GeV > 1.2		

Table 6.2: Table of definitions for signal regions SRA and SRB from [3].

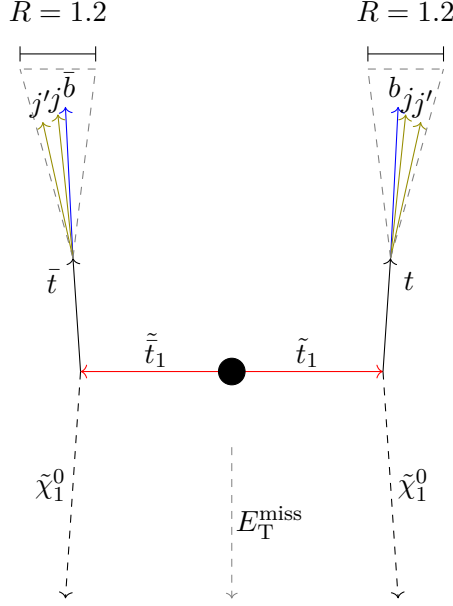


Figure 6.1: Schematic diagram of a signal event in signal region SRA/SRB (TT category). The other top categories cover the cases where one top branch is not fully reconstructed inside an $R=1.2$ reclustered jet. The TW category covers one full reconstructed top, and one reclustered jet which at least contains the W-associated jets. T0 contains only a single fully reconstructed top.

6.2.3 Signal Region SRC (2015-16)

Signal region SRC [3] targets the compressed signal models, namely $\Delta(m_{\tilde{t}}, m_{\tilde{\chi}^0}) \simeq m_t$. By default, the compressed SUSY signal will be indistinguishable from the $t\bar{t}$ background with a small E_T^{miss} , so in order to resolve such a system and to provide a sufficiently large E_T^{miss} such that the trigger is fired, the signal is assumed to be recoiled against a boosted ISR jet (or set of jets), such that the top p_T and E_T^{miss} can be resolved.

This region uses the “recursive jigsaw reconstruction” [3, 102] to separate the jets and E_T^{miss} in a given signal event into three categories, the ISR system, the visible decay products of the SUSY particles (in this case the hadronically decaying top and its daughters) and the invisible decay products of the SUSY particles (the E_T^{miss}). The diagram demonstrates the groupings of objects in the event.

A key discriminant of signal against background is the R_{ISR} defined as [2]:

$$R_{\text{ISR}} = \frac{E_T^{\text{miss}}}{p_T^{\text{ISR}}} \sim \frac{m_{\tilde{\chi}^0}}{m_{\tilde{t}}}. \quad (6.3)$$

This last approximate equality can be considered from the kinematics of an event where the ISR jet system and the $t\bar{t}$ component are anti collinear, boosting the neutralinos to produce a large E_T^{miss} . In the rest frame of the stop decay, the top and neutralinos will be approximately at rest, which would be indistinguishable from the $t\bar{t}$ background. Hence,

when recoiled against ISR [103], it can be observed that:

$$\mathbf{p}_T(\chi) \simeq \frac{m_{\tilde{\chi}^0}}{m_{\tilde{t}}} \mathbf{p}_{T\tilde{t}}. \quad (6.4)$$

To cover the full class of signal points along the compressed diagonal, a multi-bin fit is performed on five bins of the R_{ISR} distribution to maximise exclusion for each signal. Selections on $E_T^{\text{miss,track}}$ and $\min[\Delta\phi(j_{1-2}, E_T^{\text{miss}})]$ are used to mitigate multi-jet backgrounds. Table 6.3 outlines the selections used, while figure 6.2 depicts the event topology.

Variable/SR	SRC1	SRC2	SRC3	SRC4	SRC5
$N_{b\text{-jet}}$			≥ 1		
$N_{b\text{-jet}}^S$			≥ 1		
N_{jet}^S			≥ 5		
$p_{T,b}^{1,S}$			$> 40 \text{ GeV}$		
m_S			$> 300 \text{ GeV}$		
$\Delta\phi_{\text{ISR}, E_T^{\text{miss}}}$			> 3.0		
p_T^{ISR}			$> 400 \text{ GeV}$		
$p_T^{4,S}$			$> 50 \text{ GeV}$		
R_{ISR}	$\in [0.3, 0.4]$	$\in [0.4, 0.5]$	$\in [0.5, 0.6]$	$\in [0.6, 0.7]$	$\in [0.7, 0.8]$

Table 6.3: Selection criteria for the compressed region SRC from [3], in addition to the preselection.

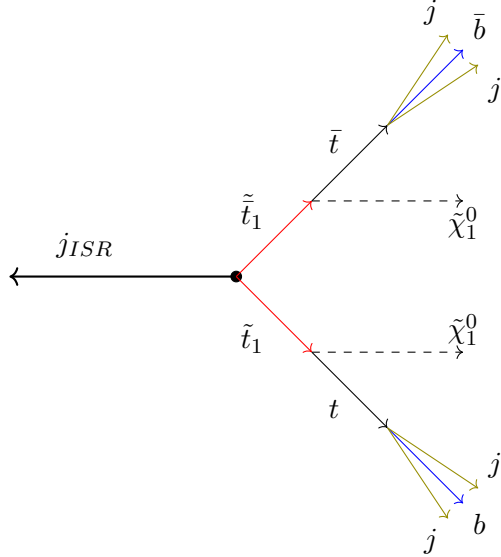


Figure 6.2: Schematic of a signal event that would be found in SRC: a pair produced stop system recoiled against an ISR jet. Each of the jets associated to the top decay will be spatially separate from each other. In addition to the pictured case, this signal region also covers the case where one of the top decays semi-leptonically (with the lepton mis-identified/out of acceptance), or including decay with hadronic τ s.

6.2.4 Other signal regions (SRD, SRE)

In this publication, additional SUSY signal models were covered which relate to alternative SUSY scenarios which give approximately the same final state, principally models that take the form $\tilde{t}_1 \rightarrow b + \tilde{\chi}_1^+ \rightarrow b + W^+ \tilde{\chi}_1^0$ (known as “b+chargino” models) and finally to the LSP via an emission of a Z/h. A different model with a similar final state is instead the pair-produced gluino model ($\tilde{g} \rightarrow t\tilde{t}_1 \rightarrow t\tilde{\chi}_1^0 + \text{soft}$). SRD covers the first case, while SRE covers the latter. However, the author did not contribute to these signal regions and so will be considered out of the scope of this thesis.

6.3 Control regions

6.3.1 $t\bar{t}$ control regions (2015-16)

The $t\bar{t}$ process which principally contributes to the signal region is the semi-leptonic $t\bar{t}$, where one of the W’s in the top decay chain produces either a lepton which is not reconstructed (either due to a detector crack or being produced outside of kinematic acceptance) or a hadronically decaying τ . Since the signal region requires $E_T^{\text{miss}} > 250$ GeV at minimum, the all-hadronic $t\bar{t}$ process is substantially suppressed, since the principal source of E_T^{miss} would originate from jet mis-reconstruction. This process is particularly dominant in the case of low E_T^{miss} , and is notably indistinguishable from compressed SUSY signals in the case of soft E_T^{miss} (so the target space of signal region SRC before an ISR boost is applied).

In order to accurately model the semi-leptonic $t\bar{t}$ background, a prescription referred to as “treating the lepton as a jet” is used. The $t\bar{t}$ background has one control region per signal region used, with the principal differing variables from the signal region (with exception of SRC, which is one for all SRC bins). All of them require a single electron/muon, inclusion therein to the selection for jet observables, the selection on $m_T(\ell, E_T^{\text{miss}})$ required primarily to ensure orthogonality with the signal region of the ATLAS one lepton stop-search and a selection on the $\Delta R(b_{1,2}, \ell)_{\text{min}}$ to ensure that the lepton is associated to a top decay in the event.

For control regions associated to SRA, selections on the E_T^{miss} are relaxed to 250/300/350 GeV for categories (TT, TW, T0), and the selections on m_{T2, χ^2} are removed, but otherwise are identical to the signal region. For control regions associated to SRB, the E_T^{miss} was loosened to > 250 GeV, but otherwise the signal selections were preserved.

6.3.2 W+jets and single top control regions.

The $W + \text{jets}$ background is a comparatively small background except in the cases of SRA/B, where a leptonically decaying W (either into an electron/muon out of acceptance, or a hadronically decaying τ), with jets arising from ISR/FSR virtual gluon emissions, producing jets/b-jets. For large sufficiently large virtual gluon p_T , these jets can then be found to be reconstructed within the $R=1.2$ jet mass cones (particularly in SRA-T0 and SRB-T0).

Similarly, single top processes normally consist of processes such as tZq and tWb , where $Z \rightarrow \nu\nu$, and $W \rightarrow \ell\nu$, with the lepton either an out of acceptance electron/muon or a

hadronically decaying τ , and the remainder producing sufficient b-jets and hadronic jets. Since there is only a single top in the event (any additional b-jets or jets which mimic a top will arise from ISR/FSR gluon emission), this process is particularly prevalent in signal regions SRAT0/SRBT0.

These control regions are implemented for all signal regions (primarily affecting SRA and SRB), although the impacts of them are comparatively low in SRC, since SRC is almost solely dominated by $t\bar{t}$.

The W-jets and single top control regions follow the “treating a lepton as a jet” prescription and baseline requirements as per the $t\bar{t}$ control regions, but the b-jets in the event and the lepton are spatially separated ($\Delta R(b_{1,2}, \ell)_{\min} > 2.0$) (which forms the orthogonality with the $t\bar{t}$ control regions). For the $W + \text{jets}$ control regions, these are defined by the requirement of a single b-jet, and $m_{\text{jet}, R=1.2}^1 < 120$ GeV. Control region CRST is defined by a large separation between b-jets, $m_{\text{jet}, R=1.2}^1 < 120$ GeV and a selection on $m_{\text{T}}^{b, \min}$ to reduce $t\bar{t}$ contamination.

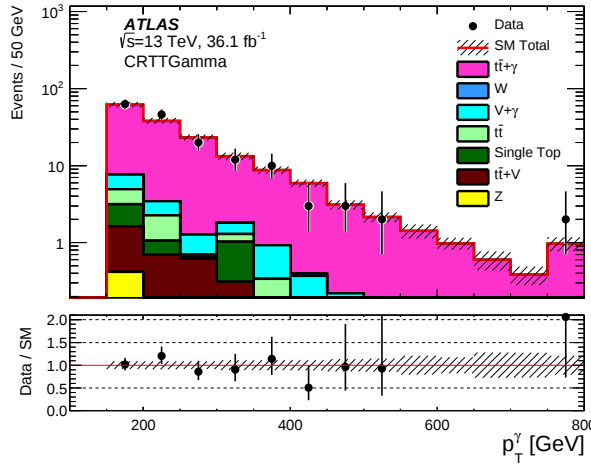
6.3.3 $t\bar{t} + Z$ control region using $t\bar{t} + \gamma$

$t\bar{t} + Z(\rightarrow \nu\nu)$ is a kinematically irreducible background, such that the final state topology is indistinguishable from the signal topology (even having nearly identical Feynman diagrams). Thus, this background is significant as any SUSY signal must be evaluated as an excess above this benchmark background. Hence, a high quality estimate is vitally important, and the exact process of interest cannot be constrained in a control region directly. To circumvent this problem, *nearest equivalent* models are used, such as $t\bar{t} + Z(\rightarrow \ell\ell)$, $\ell = e, \mu$, which would share approximately the same kinematics (with exception of FSR photons from leptons, that would not occur with neutrinos). Since then the lepton p_{T} could be measured (and hence the combined pair as $p_{\text{T}}(Z)/p_{\text{T}}^{\ell\ell}$), it could be used to model the $E_{\text{T}}^{\text{miss}}$ produced by $t\bar{t} + Z(\rightarrow \nu\nu)$ production in the signal region.

However, as will be noted by the author later in section 8.2, there are several challenges which must be overcome, namely low statistics in the high $p_{\text{T}}^{\ell\ell}$ tails and poor purity of $t\bar{t} + Z$ as compared to $Z + \text{jets}$ and di-leptonic $t\bar{t}$. The result in [3] was limited by statistics in the high $p_{\text{T}}^{\ell\ell}$ region overall, so instead a single isolated high p_{T} photon as the candidate model for the Z, since a photon could be well resolved. This issue would only be rectified later through the author’s work in section 8.2 with the benefit of a larger dataset.

The original definition follows the similar methodology to [21], requiring a single signal lepton to ensure purity of $t\bar{t} + \gamma$ against $\gamma + \text{jets}$ processes, and to avoid overlap with the signal region. Kinematically, the $t\bar{t} + Z$ and $t\bar{t} + \gamma$ will have a near identical set of Feynman diagrams (with the exception of some kinematically suppressed final state Z emission from neutrinos). This allowed for a normalisation to be extracted based on $t\bar{t} + \gamma$ that could be applied to the signal region, with some loosening of the photon p_{T} to allow for sufficient statistics in the control region. Table 6.4 outlines the selections for this region.

Variable/CR	CR $t\bar{t} + \gamma$
Trigger	Single electron/muon
N_ℓ	$== 1$
$p_{T(\ell)}$	$> 28 \text{ GeV}$
N_γ	$== 1$
p_T^γ	$> 150 \text{ GeV}$
N_{jets}	≥ 4
Jet p_T	$> (80, 80, 40, 40) \text{ GeV}$
$N_{b\text{-jet}}$	≥ 2

 Table 6.4: Table of selections for the $t\bar{t} + \gamma$ control region for $t\bar{t} + Z$ as per [3].

 Figure 6.3: Distribution in the $t\bar{t} + \gamma$ control region for the $t\bar{t} + Z$ process, applying the maximal likelihood estimates of the normalisation of the contributing backgrounds. This plot can be found in [3].

6.3.4 Z +jets control regions

As will be discussed further in section 8.3, the b-jets in the $Z \rightarrow \nu\nu + b$ -jets process can enter the signal region in two different ways: as the original Z-radiating jets or from ISR/FSR gluon decays to jets or b-jets. This is a particularly dominant process in SRA, SRB, since the boosted jets in SRA/SRB may mean the additional jets from gluon splitting also lie within an $R=1.2$ jet cone. As will be elaborated later, the $\Delta R(b_1, b_2)$ is a significant discriminant between the two sources of b-jets in the event, since the b-jets associated to the decay of a boosted gluon are likely to be close together.

The production of $Z \rightarrow \nu\nu$ is normally constrained in the signal region using the process $Z \rightarrow \ell\ell + \text{jets}$ ($\ell = e, \mu$), since like for $t\bar{t} + Z$ this is a near-equivalent model of $Z \rightarrow \nu\nu$. Since we are looking at a process whereby the Z is the source of the electrons/muons (and we wish to model E_T^{miss} in the signal region), the E_T^{miss} is replaced with $E_T^{\text{miss}'}$, whereby the E_T^{miss} is re-calculated, treating the electrons/muons as invisible. To allow for the fact that some of the E_T^{miss} in the signal region may arise from jet mismeasurement (although is not the principal E_T^{miss} source), an upper selection on real $E_T^{\text{miss}} < 50 \text{ GeV}$ is applied. Using $E_T^{\text{miss}'}$

instead of E_T^{miss} , analogous selections to the signal region are hence applied (all variables using the $E_T^{\text{miss}'}$ are themselves primed).

Selection	CRZAB-TT-TW	CRZAB-T0
Trigger	Single ℓ (e, μ)	
N_ℓ	$= 2$	
p_T^ℓ	$> (28, 20) \text{ GeV}$	
N_j	≥ 4	
Jet p_T	$> (80, 80, 40, 40) \text{ GeV}$	
E_T^{miss}	$< 50 \text{ GeV}$	
$E_T^{\text{miss}'} (e\mu \text{ invisible})$	$> 100 \text{ GeV}$	
N_b	≥ 2	
$m_{\text{jet}, R=1.2}^1$	$> 120 \text{ GeV}$	
$m_{\text{jet}, R=1.2}^2$	$> 60 \text{ GeV}$	$< 60 \text{ GeV}$

Table 6.5: 2015-16 $Z + \text{jets}$ control region definitions from [3], corresponding to SRA/SRB. Mass categories analogous to the signal region were used for SRA and SRB associated control regions, although due to low statistics, the TT and TW mass categories were merged.

6.3.5 Multi-jet estimation

In all regions, especially signal region SRC, the multi-jet background is non-negligible. This background is estimated using the data-driven jet smearing method [3, 104], using data collected with single-jet triggers instead of E_T^{miss} . This procedure is outlined as follows from [104]:

1. Select low- E_T^{miss} seed events in data with well measured jets (passing quality cuts)
2. Construct a response function in simulated di-jet MC comparing generator-level jet energy to reconstructed jet energy. This response function is constrained empirically in [104], based on matching the ratio of the generator level jet p_T with the reconstructed jet p_T if they are within some ΔR threshold.
3. Convolve or smear the jets in the seed events with the response function to generate pseudo-data events. This first set of pseudo-data is used to constrain the response function to match experimental data by matching the pseudo-data to the experimental data, and modifying the response function accordingly.
4. Apply additional gaussian smearing and tail corrections to the jet response.
5. Convolve or smear the jets in the seed events with the constrained response function determined in 3. to obtain the pseudo-data sample, recalculating E_T^{miss} with the smeared jet four-momenta.

This has been also applied for all signal regions SRA-C. Since this method is a data-driven background estimation, no normalisation control region has been used, and systematics relating to the multi-jet estimate are taken to be 30% as a conservative estimate of the

configuration of the seed event, response function and tail corrections. This estimate is largely an order of magnitude estimate of the uncertainties in the seed events (such as number of input data events used for the response function etc). However, as this background is effectively negligible in signal regions SRA and SRB for these regions a more detailed analysis is not required since the uncertainty is sub-dominant. For SRC, the 30% systematic uncertainties on the seed selections etc are not the dominant uncertainties (primarily this is associated with the estimation of $t\bar{t}$), so a modest reduction in this uncertainty would not have any noticeable changes to the fit configuration.

The author contributed to this estimate through providing correction scale factors of the data-driven method in a sideband control region, such that the estimate matches the given dataset and object definitions, trigger used etc.

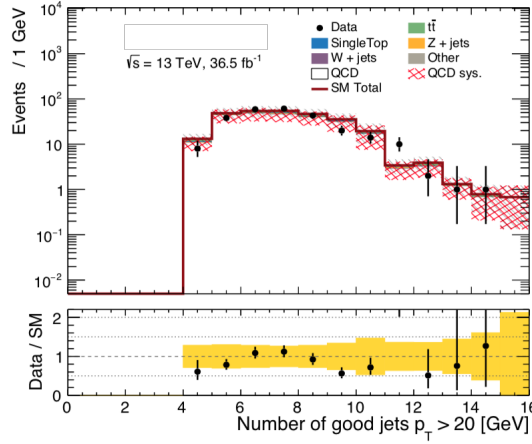


Figure 6.4: N-jets distribution in a QCD control region used to constrain the QCD scale factors.

6.4 Validation Regions

This section outlines the validation regions corresponding to $Z + \text{jets}$, $t\bar{t}$ and $W + \text{jets}$ backgrounds in [3]. The underlying principle for these validation regions is that they contain 0 electrons/muons (with exception for the validation region for the $W + \text{jets}$ background, which requires one), and generally chosen by inversion of a limited set of selections from the signal region, with the exact selections considered principally for purity of the validation region (which ideally is the principal background process in the validation region). Generally these validation regions are designed after the definition of signal regions, such that the process generally shares the same kinematic distribution of a given sample as the signal region (normally by inversion of a single selection).

The requirements for sample purity, magnitude of statistical uncertainties etc are, however, substantially weaker than those for the control region, since these validation regions are only required to validate the constrained fits on the normalisation parameters and systematics. In essence, this means that the leading background in the validation region should be

the background of interest (and as pure as is reasonably achievable while minimising the extrapolation to the signal region).

6.4.1 $Z + \text{jets}$ validation regions

As the $Z + \text{jets}$ background is primarily important in signal regions SRA and SRB, one validation region is constructed: VRZAB (for SRA/SRB). This validation region is presented in Table 6.6.

Variable/VR	VRZAB
Preselection	0L preselection (Table 6.1)
$p_T^{j_1}, p_T^{j_2}$	$> 80, > 80$ GeV
N_{jets}	≥ 4
$N_{\text{b-jet}}$	≥ 2
τ -veto	$\checkmark$
$m_T^{b,\min}$	> 200 GeV
$m_{\text{jet}, R=1.2}^1$	< 120 GeV
$\Delta R(b_1, b_2)$	< 1.0

Table 6.6: Table of selections for the $Z + \text{jets}$ validation region used in [3]

6.4.2 $t\bar{t}$ and $W + \text{jets}$ validation regions

$t\bar{t}$

Validation regions for the $t\bar{t}$ background have the same 0-lepton preselection as per the signal regions, except for the inversion of the $m_T^{b,\min} > 200$ selection in the signal region. There are six $t\bar{t}$ validation regions for SRA and SRB, and one each for signal regions SRC, SRD and SRE. These are as follows:

- Validation region VRTA shares the requirements as per SRA, except $m_T^{b,\min} \in [100, 200]$ GeV and E_T^{miss} is lowered by 100 GeV, binned by top categories as per the signal region SRA.
- VRTB shares all of the requirements with the SRB bins, except with $m_T^{b,\min} \in [100, 200]$ for VRTB-TT, $m_T^{b,\min} \in [140, 200]$ for VRTB-TW or $m_T^{b,\min} \in [160, 200]$ for VRTB-T0.
- Validation region VRTC uses the same requirements as signal region SRC, except for looser requirements $m_S > 100$ GeV, $p_T^{4,S} > 40$ GeV, $N_{\text{jet}}^S > 4$, the $\Delta\phi_{\text{ISR}, E_T^{\text{miss}}} < 3.0$, $m_V/m_S < 0.6$ and no binning in R_{ISR} .

$W + \text{jets}$

The $W + \text{jets}$ background is validated in a single validation region for all signal regions. This region is based on the one-lepton CRW, except that $\Delta R(b_{0,1}, \ell)_{\min} > 1.8$, $m_T^{b,\min} > 150$ GeV and $m_{\text{jet}, R=1.2}^1 < 70$ GeV.

Single top and other backgrounds

Since the single-top background is comparatively small compared to other backgrounds in the fit across all regions, no validation region is required since the overall effect of the single top normalisation on the overall background yield would be comparatively small (reducing the uncertainty on said background normalisation in the signal region is the principal motivation for constraining this background). All other backgrounds are not additionally validated in the fit, since they are generally either estimated by data-driven methods or are estimated purely by simulation.

6.5 Results

In [3], a simultaneous profile likelihood fit was used, constraining the background estimates for $Z + \text{jets}$, $t\bar{t}$, $W + \text{jets}$, single top and $t\bar{t} + Z$ using their respective control regions. The contributions of multi-jets and di-boson processes are fixed at the values expected either from the data-driven method or simulation respectively. This fit is undertaken using HistFitter [92] to construct the likelihood, incorporating systematics as nuisance parameters in the fit. In figure 6.5, the validation region distributions are plotted post-fit with the fitted normalisations and systematics used, and 6.6 outlines the distributions post-fit for the signal regions. The yields tables are outlined in tables 6.7 and 6.8, while the results of the discovery fit can be found in 6.9.

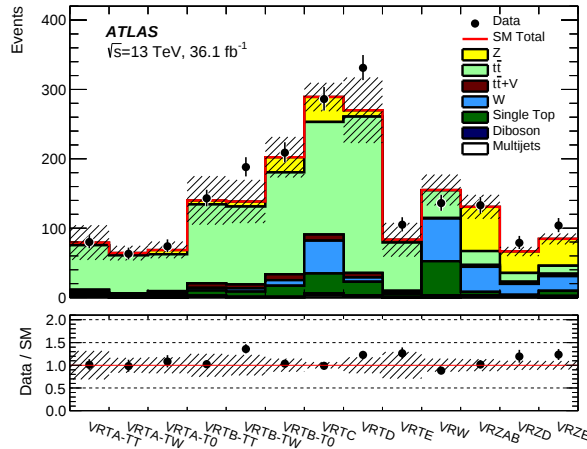


Figure 6.5: Background-only fit results (post-fit) of the validation regions from [3]. All detector systematic uncertainties and statistical uncertainties are included.

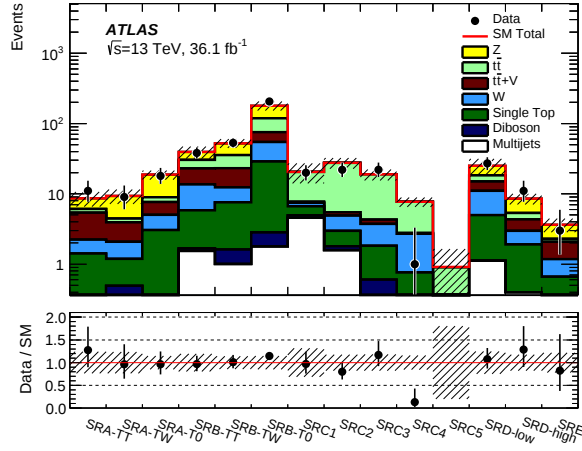


Figure 6.6: Background-only fit results (post-fit) of the signal regions from [3]. All detector systematic uncertainties and statistical uncertainties are included.

	SRA-TT	SRA-TW	SRA-T0	SRB-TT	SRB-TW	SRB-T0
Observed	11	9	18	38	53	206
Total SM	8.6 ± 2.1	9.3 ± 2.2	18.7 ± 2.7	39.3 ± 7.6	52.4 ± 7.4	179 ± 26
$t\bar{t}$	$0.71^{+0.91}_{-0.71}$	$0.51^{+0.55}_{-0.51}$	1.31 ± 0.64	7.3 ± 4.3	12.4 ± 5.9	43 ± 22
$W + \text{jets}$	0.82 ± 0.15	0.89 ± 0.56	2.00 ± 0.83	7.8 ± 2.8	4.8 ± 1.3	25.8 ± 8.8
$Z + \text{jets}$	2.5 ± 1.3	4.9 ± 1.9	9.8 ± 1.6	9.0 ± 2.8	16.8 ± 4.1	60.7 ± 9.6
$t\bar{t} + W/Z$	3.16 ± 0.66	1.84 ± 0.39	2.60 ± 0.53	9.3 ± 1.7	10.8 ± 1.6	20.5 ± 3.2
Single top	1.20 ± 0.81	0.70 ± 0.42	2.9 ± 1.5	4.2 ± 2.2	5.9 ± 2.8	16.8 ± 4.1
Diboson	-	0.35 ± 0.26	-	0.13 ± 0.07	0.60 ± 0.43	1.04 ± 0.73
Multi-jet (QCD)	0.21 ± 0.10	0.14 ± 0.09	0.12 ± 0.07	1.54 ± 0.64	1.01 ± 0.88	1.8 ± 1.5

Table 6.7: Fitted yields for signal region SRA and SRB as observed in [3].

	SRC1	SRC2	SRC3	SRC4	SRC5
Observed	20	22	22	1	0
Total SM	20.6 ± 6.5	27.6 ± 4.9	18.9 ± 3.4	7.7 ± 1.2	0.91 ± 0.73
$t\bar{t}$	12.9 ± 5.9	22.1 ± 4.3	14.6 ± 3.2	4.91 ± 0.97	$0.63^{+0.70}_{-0.63}$
$W + \text{jets}$	0.80 ± 0.37	1.93 ± 0.49	1.91 ± 0.62	1.93 ± 0.46	0.21 ± 0.12
$Z + \text{jets}$	-	-	-	-	-
$t\bar{t} + W/Z$	0.29 ± 0.16	0.59 ± 0.38	0.56 ± 0.31	0.08 ± 0.08	0.06 ± 0.02
Single top	1.7 ± 1.3	$1.2^{+1.4}_{-1.2}$	1.22 ± 0.69	0.72 ± 0.37	-
Diboson	0.39 ± 0.33	$0.21^{+0.23}_{-0.21}$	0.28 ± 0.18	-	-
Multi-jet (QCD)	4.6 ± 2.4	1.58 ± 0.77	0.32 ± 0.17	0.04 ± 0.02	-

Table 6.8: Fitted yields for signal region SRC as observed in [3].

The leading systematic contributions are for each signal region:

- SRA/SRB: $t\bar{t}$ theory uncertainties, single-top theory uncertainties, JER, Pileup, $Z + \text{jets}$ normalisation, $t\bar{t} + W/Z$ normalisation and single top normalisation.
- SRC: $t\bar{t}$ theory uncertainty, multi-jet estimates, $t\bar{t}$ normalisation, JER and JES, $E_{\text{T}}^{\text{miss}}$ soft term.

For an interested reader, the systematic contributions to each signal regions can be found in [3].

Signal Region	P-value	Significance
SRA-TT	0.23	0.74
SRA-TW	0.5	0.00
SRA-T0	0.5	0.00
SRB-TT	0.5	0.00
SRB-TW	0.5	0.00
SRB-T0	0.13	1.15
SRC1	0.5	0.00
SRC2	0.5	0.00
SRC3	0.27	0.62
SRC4	0.5	0.00
SRC5	0.23	0.74

Table 6.9: Discovery p-value and significances for each signal region given the background-only hypothesis for the regions in [3]. By construction in HistFitter, the significance of any under-fluctuation is neglected, defining the p-value to be 0.5 and the significance to be 0.00 in such cases.

As can be noted from figure 6.6 and table 6.9, no significant excess was observed in any signal region, thus exclusion limits were applied on the $\Delta(m_{\tilde{t}}, m_{\tilde{\chi}^0})$ plane for the direct decay and the results are listed in full in [3], and summarised in figures 6.8.

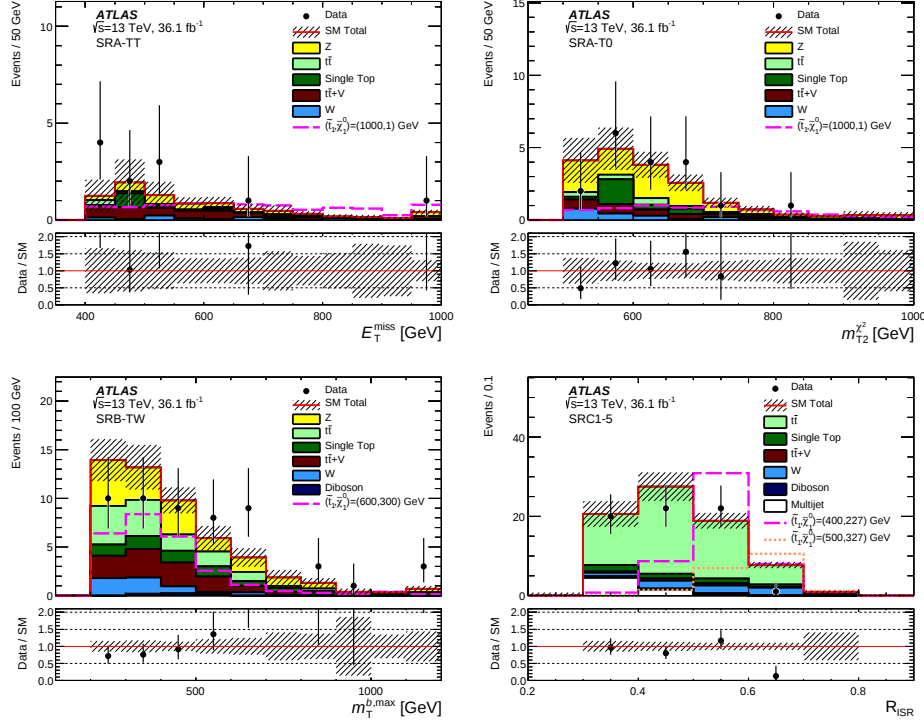
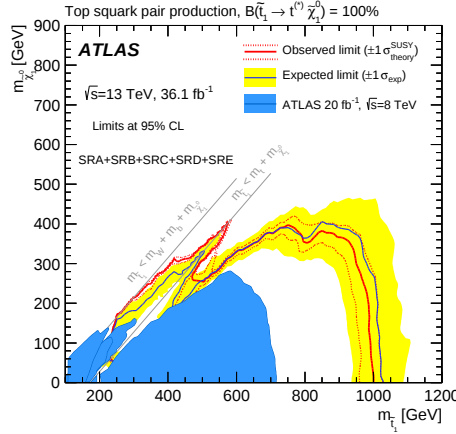


Figure 6.7: Post-fit distributions of the signal regions in [3].

Figure 6.8: Exclusion limits for signals in the simplified scenario where $\tilde{t}_1 \rightarrow t + \tilde{\chi}^0$ with 100% branching ratio. Previous results obtained in LHC run 1 [27] are plotted for comparison.

As concluded in [3], no statistically significant excess is observed in any signal region. Exclusion limits are applied in the case of the direct stop decay models, namely simplified models with stop masses up to 1TeV, and up to $(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (500, 350)$ GeV along the compressed diagonal.

This analysis forms the benchmark by which the later ATLAS search in chapter 9 will be compared against later.

Re-optimisation of the 2015-16 signal regions

Between the LHC data-taking periods of 2017-2018, another 103 fb^{-1} of data was collected. This enlarged dataset allowed us to radically re-think how we undertake searches for supersymmetry. This encompassed new techniques and ideas from within the HEP community, such as machine learning and a hybrid b-jet identification. In addition, larger datasets benefit the efforts of reconstructed object calibration substantially, giving significantly improved reconstruction performance and allowing us to better utilise the detector in more difficult regions of phase space (such as low p_T electrons or muons). With these improvements in hand, it is possible to better probe some of the underlying assumptions in the background composition to see if new variables, objects and discriminants can be applied that may have been previously limited by low statistics. This chapter and the next cover the author's extension work beyond the analysis of the last chapter.

This chapter specifically covers the studies undertaken by the author in the development of [2], however these were not included in the publication due to limitations that will be elaborated further in this chapter. Chapter 8 will reference the studies which formed a core component of [2] and form the motivation for some of the decisions made therein.

To avoid any confusion with other definitions, the significances used in this chapter specifically will be the expected significance that is extracted from a binomial distribution for some given signal yield s , some background yield b , and some fixed uncertainty size ($\sigma_b = 20\%$). For those readers familiar with RooStats[93], this pertains to the *Roostats::BinomialExpZ(s,b,20%)* function.

7.1 Jet substructure and boosted object identification

As the top quark produced from the stop decay is inherently unstable, it will not be explicitly observed, thus it must be reconstructed from its final decay products, in this analysis can be assumed to be a b jet and two jets from the emission of a W boson. In some sectors of the analysis it is desired to identify W associated jets (jet daughters of a W , which will have a net charge of ± 1) as well, namely in the 2015-16 signal region SRA-TW and control region CRW. In resolved scenarios, where all constituents are isolated from each other, it can be

possible to reconstruct the top quark objects, however in more boosted (high p_T) scenarios, multiple constituent jet cones can overlap and may not be individually reconstructed.

Several strategies from outside ATLAS SUSY searches have been investigated since the discovery of the top quark in 1995 at DØ [105] and CDF [106], mostly from analyses targeting $t\bar{t}$ resonances in boosted regimes.

Two methods are investigated for top tagging: using a boosted decision tree for several input variable selections to produce a classification using a BDT or kinematic fitting[77], based on comparing the signal mass spectrum to that of a Gaussian set at either the Top/W mass with width of 15(30) GeV, and the final cut is tweaked until the signal efficiency is 50(80)%. This was implemented using a smooth top tagger (smooth in jet p_T).

In this analysis, the author considered the improvement in signal efficiency for signal region SRA-T0, and applied a strategy, whereby we replace the reclustered $R=1.2$ and $R=0.8$ jets built from $R=0.4$ jets with clustered $R=1.0$ jets. We then investigated the signal significance and applied the tagger to the leading- p_T jet with a signal efficiency of 80%, as the 50% signal efficiency working point was deemed to be too restrictive. The signal significance is calculated from a binomial distribution comparing the expected values against the background, with an assumed 20% systematic error. The signal region selections used in this study are outlined in table 7.1. The results of this strategy are discussed below in tables 7.2, 7.3 and 7.4. **FIX THE TABLES!**

Variable	Selection
E_T^{miss}	$> 550 \text{ GeV} + E_T^{\text{miss}} \text{ trigger [3]}$
$ \Delta\phi(E_T^{\text{miss}}, E_T^{\text{miss,track}}) $	$< \frac{\pi}{3}$
$E_T^{\text{miss,track}}$	$> 30 \text{ GeV}$
$\tau\text{-veto}$	$\checkmark$
$M_{T2}^{\chi^2}$	$> 500 \text{ GeV}$
$m_{1,R=1.2}^{\text{jet}}$	$> 120 \text{ GeV}$
$m_{2,R=1.2}^{\text{jet}}$	$< 60 \text{ GeV}$
$m_{2,R=0.8}^{\text{jet}}$	$> 60 \text{ GeV}$
$ \Delta\phi(\text{jet}^{0,1,2,3}, E_T^{\text{miss}}) $	> 0.4
$m_T^{b,\text{min}}$	$> 200 \text{ GeV}$
No. b-jets	≥ 2
No. Leptons	0
No. Jets	> 4
Jet p_T (in order of p_T)	$> (80, 80, 40, 40) \text{ GeV}$

Table 7.1: A reminder of the SRA-T0 region definition [3] including the pre-selection conditions.

SRAT0 (default)	
$t\bar{t}$	1.39 ± 0.31
$t\bar{t} + W/Z$	1.96 ± 0.20
Single top	2.24 ± 0.78
Diboson	0.77 ± 0.31
$Z + \text{jets}$	8.64 ± 0.64
$W + \text{jets}$	1.65 ± 0.49
Total (SM)	16.65 ± 1.22
Signals	
$(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (1000, 1)$	6.42 ± 0.33 ($Z_N = 0.99$)
$(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (600, 300)$	0.58 ± 0.15 ($Z_N = -1.0$)
$(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (900, 1)$	11.76 ± 0.66 ($Z_N = 1.82$)
$(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (700, 400)$	0.34 ± 0.06 ($Z_N = -1.00$)
$(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (800, 1)$	18.53 ± 0.36 ($Z_N = 2.76$)
$(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (550, 250)$	0.30 ± 0.30 ($Z_N = -1.00$)

Table 7.2: Event yields in the 2015-16 Signal Region SRA-T0, with a luminosity of 36.1fb^{-1} . The significance is the expected binomial significance given a 30% systematic. Stated significances of "-1.00" occur where the number of signal events are less than 1.

JSSAT0STn0kt8	
$t\bar{t}$	1.43 ± 0.30
$t\bar{t} + W/Z$	1.99 ± 0.21
Single top	2.00 ± 0.76
Diboson	0.91 ± 0.33
$Z + \text{jets}$	8.54 ± 0.68
$W + \text{jets}$	1.56 ± 0.36
Total (SM)	16.43 ± 1.19
Signals	
$(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (1000, 1)$	6.55 ± 0.33 ($Z_N = 1.02$)
$(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (600, 300)$	0.73 ± 0.16 ($Z_N = -1.00$)
$(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (900, 1)$	11.88 ± 0.64 ($Z_N = 1.85$)
$(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (700, 400)$	0.46 ± 0.07 ($Z_N = -1.00$)
$(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (800, 1)$	19.45 ± 0.37 ($Z_N = 2.91$)
$(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (550, 250)$	1.80 ± 1.80 ($Z_N = 0.19$)

Table 7.3: Event yields for signal Region SRA-T0 (table 7.1) with the $R = 0.8$ jet selections removed, and the $R=1.2$ reclustered jet mass selections replaced with $R=1.0$ jet mass selections, also at 36.1fb^{-1} . Stated significances of "-1.00" occur where the number of signal events are less than 1.

JSSAT0STn0kt8	
$t\bar{t}$	0.93 ± 0.21
$t\bar{t} + W/Z$	1.47 ± 0.18
Single top	0.84 ± 0.15
Diboson	0.55 ± 0.28
$Z + \text{jets}$	4.49 ± 0.50
$W + \text{jets}$	0.85 ± 0.25
Total (SM)	9.14 ± 0.70
Signals	
$(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (1000, 1)$	5.83 ± 0.32 ($Z_N = 1.33$)
$(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (600, 300)$	0.49 ± 0.14 ($Z_N = -1.00$)
$(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (900, 1)$	10.17 ± 0.60 ($Z_N = 2.27$)
$(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (700, 400)$	0.25 ± 0.06 ($Z_N = -1.00$)
$(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (800, 1)$	15.99 ± 0.33 ($Z_N = 3.38$)
$(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (550, 250)$	1.80 ± 1.80 ($Z_N = 0.34$)

Table 7.4: Signal Region SRA-T0 with the $R = 0.8$ jet selections removed, and the $R = 1.2$ reclustered jet selections replaced with $R=1.0$ jet selections. In addition, the leading- p_T jet is tagged using the 80% signal efficiency working point. Stated significances of "-1.00" occur where the number of signal events is less than 1.

From these tables, looking at the signals near the SRA target mass points $(m_t, m_{\tilde{\chi}^0}) = (900 \text{ GeV}, 1 \text{ GeV})$, there is some improvement in the significance in region SRA-T0 by applying the clustering approach followed by tagging.

7.2 Boosted Decision tree application to signal discrimination

In addition to considering the already available identification techniques for top quarks, an analysis-specific multi-variate optimisation was undertaken. This particularly targeted the boosted signals in the 2015/16 SRA-TT/SRATW/SRAT0 signal regions [3], and used a TMVA Boosted Decision tree for binary signal and background classification. The primary aims of the Boosted Decision Tree was to both optimise the signal significance as compared to the direct observable based selections as listed in [3], and to determine if there were any underpinning correlations between variable selections that could not be observed directly (say due to a relation to some third variable not explicitly listed). Since this study was undertaken as a comparison to the regions defined in [3], 36.1fb^{-1} of MC were compared against the 2015-2016 run periods (integrated luminosity of 36.1fb^{-1}), and loose selections were applied, primarily as event cleaning, trigger requirements and basic topology selections.

Variable/Region	SRA-Loose
E_T^{miss}	$> 400 \text{ GeV} + \text{Trigger}$
N_ℓ	$=0$
$N_j(R = 0.4)$	≥ 4
$E_T^{\text{miss,track}}$	$> 30 \text{ GeV}$
$\min[\Delta\phi(j_{1-3}, E_T^{\text{miss}})]$	> 0.4
$ \Delta\phi(E_T^{\text{miss}}, E_T^{\text{miss,track}}) $	$\Pi/3$
$m_T^{b,\min}$	$> 200 \text{ GeV}$

Table 7.5: Definition of loose preselection region used for training

Variable/Configuration	Configuration I: Simple Kinematics	Configuration II: Jet substructure
E_T^{miss}		✓
m_{T2,χ^2}		✓
$\Delta R(b_1, b_2)$		✓
$N_j(R = 0.4)$	✓	✓
$m_{\text{jet}, R=1.2}^1$ (Reclus.)	✓	-
$m_{\text{jet}, R=1.2}^2$ (Reclus.)	✓	-
$m_{\text{jet}, R=0.8}^1$ (Reclus.)	✓	-
$m_{\text{jet}, R=0.8}^2$ (Reclus.)	✓	-
$\tau_{32}(J_1; R = 1.0)$	-	✓
$\tau_{32}(J_2; R = 1.0)$	-	✓
$\sqrt{d_{23}}(J_1; R = 1.0)$	✓	-
$\sqrt{d_{23}}(J_2; R = 1.0)$	✓	-
$m(J_1; R = 1.0)$	✓	-
$m(J_2; R = 1.0)$	✓	-

Table 7.6: Table of kinematic and jet substructure variables considered in the BDT study, with the two configurations (configuration I and configuration II respectively).

BDT Training

In order to successfully train the BDT, a fraction of the samples used as signal and background must be set aside for training, and the remainder can be used for validation/testing. Since statistics even in the loose preselection region are somewhat limited, instead of a single BDT trained on a kinematic subset of the samples, two BDTs were considered, with different training and test sets respectively. The samples were split into two, based on the event number of the sample, namely separating odd and even numbered events, and a BDT was trained using the odd numbered events as the training set and even numbered events as the test set, and vice versa. The output of these two BDTs on the respective event numbers in their respective test sets were henceforth treated as a single distribution.

BDT configuration checks

Since the boosted decision trees have alternate methods of undertaking the boosting, two of the main algorithms were trialled: AdaBoost [95, 96] and Gradient Boost (GradBoost)[95, 97] and compared in their performance. Table 7.7 outlines the settings used by the BDT/BDTG, while figures 7.1, 7.2 outline the comparisons in a receiver operating characteristic plot. A trial selection on the BDTG in both configurations listed in 7.6 is applied to signal in table 7.8.

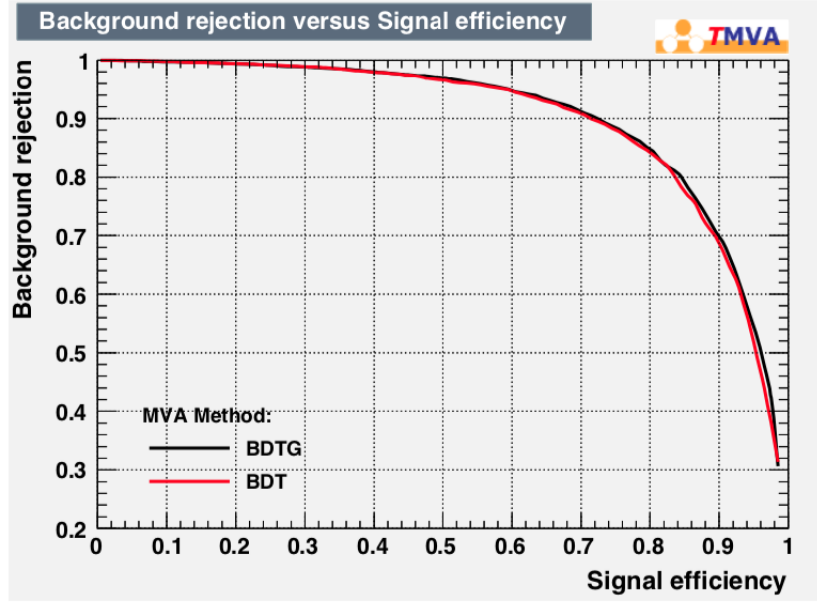


Figure 7.1: Comparison the Gradient and AdaBoost boosting algorithms for a Boosted Decision Tree (BDT) based on simple kinematics (configuration I from table 7.6) in the loose preselection region listed in 7.5.

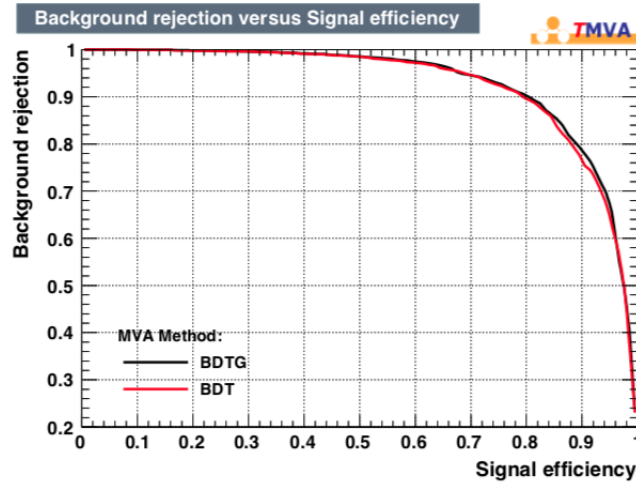


Figure 7.2: Comparison using TMVA of the Gradient and AdaBoost boosting algorithms for a Boosted Decision Tree (BDT) based on simple kinematics and jet substructure observables (configuration II from table 7.6) in the loose preselection region listed in 7.5.

Hyper-parameter	Setting (BDTG)	Setting (BDT)
N_{trees}	1000	850
Tree depth	3 levels	3 levels
Minimum leaf number of events	2.5%	-
Ada-Boost β	-	0.5
Node splitting index	Gini index	

Table 7.7: BDT/BDTG hyperparameter sets for all configurations used in table 7.6.

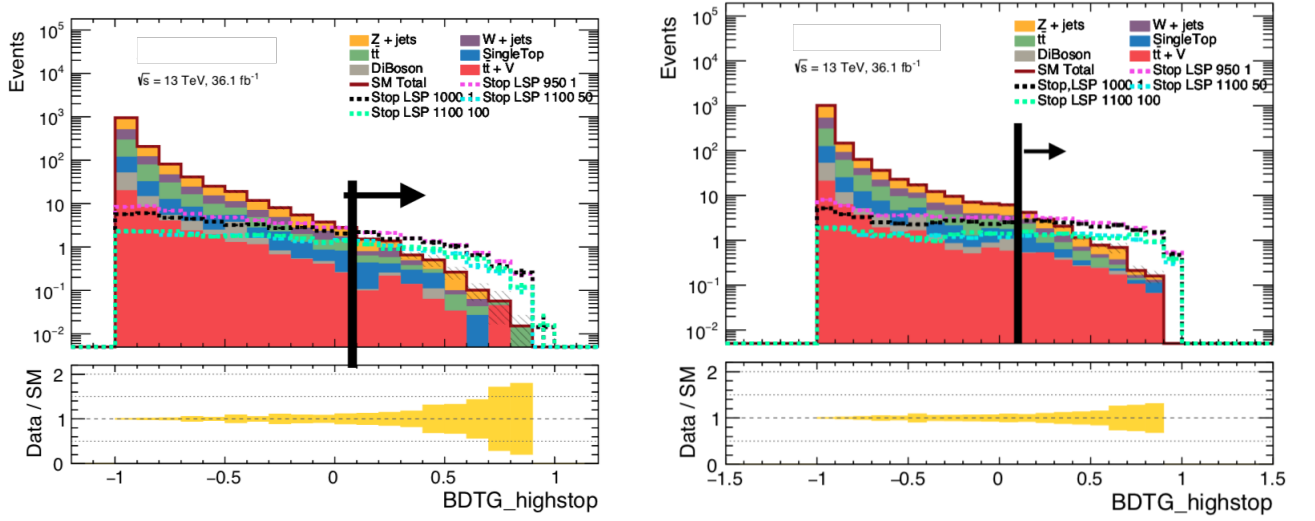


Figure 7.3: Distributions of the BDTG output score in both configuration I (left) and configuration II (right) as per table 7.6 to motivate the choice of selection. Selections were chosen to maximise expected significance (as arrows on the plots).

Sample/Region	SRA-TT	Presel+BDTG(I)>0.1	Presel+BDTG(II)>0.1
Total bkg	6.04 ± 0.42	4.51 ± 0.36	12.18 ± 0.57
$\Delta(m_{\tilde{t}}, m_{\tilde{\chi}_0}) = (1000, 1)$ GeV	$7.84 \pm 0.37, (Z_n = 2.22)$	$8.94 \pm 0.40, (Z_n = 2.87)$	$18.15 \pm 0.55, (Z_n = 3.24)$
$\Delta(m_{\tilde{t}}, m_{\tilde{\chi}_0}) = (950, 1)$ GeV	$9.33 \pm 0.30, (Z_n = 2.60)$	$9.82 \pm 0.32, (Z_n = 3.11)$	$20.84 \pm 0.47, (Z_n = 3.65)$
$\Delta(m_{\tilde{t}}, m_{\tilde{\chi}_0}) = (1100, 50)$ GeV	$4.30 \pm 0.22, (Z_n = 1.24)$	$5.67 \pm 0.28, (Z_n = 1.90)$	$10.24 \pm 0.37, (Z_n = 1.94)$
$\Delta(m_{\tilde{t}}, m_{\tilde{\chi}_0}) = (1100, 100)$ GeV	$4.37 \pm 0.23, (Z_n = 1.26)$	$5.55 \pm 0.28, (Z_n = 1.86)$	$10.38 \pm 0.38, (Z_n = 1.96)$

Table 7.8: Table of yields after applying the BDTG for the simple kinematics and the jet substructure configurations as listed in table 7.6, compared to the “cut and count” signal region SRA-TT, as defined in [3].

Discussion

For the boosted (large Δm) signals used in the study, a gradient boosted BDT based on jet substructure observables outperforms both the fixed selections for SRA-TT and the BDT trained only using simple kinematic observables. Thus even basic jet substructure information within large R jets can be used to gain additional discrimination against background. However, this must all be limited by caveats relating to statistical uncertainties in both the training and testing sets, and since training is applied based on kinematic distributions of given

signals, how stable these classifiers are to changes in the signal model parameters. The assumed uncertainty of 20% may not be sufficient to cover these variations, and the effect of these were not known at time of writing.

7.3 Practical limitations of the boosted top tagging and jet substructure observables.

Although there is some improvement for the SRA mass points, a major limitation is that the clustering information inside the $R=1.0$ anti-kt jet in ATLFASII is not consistent with fully simulated datasets at time of writing (although the jet itself is calibrated for its energy, p_T , mass etc). Since the clustering information inside the jet is particularly important for observables such as the N-subjettiness, this will lead to a noticeable differences in these distributions between full simulation and ATLFASII. Due to CPU and walltime limitations, the entire space of signal points required to test different model parameter points on the stop-neutralino mass plane must be produced using ATLFASII detector simulation instead of the full Geant4 detector simulation. Hence a comparison of ATLFASII signal against FullSim background cannot be considered correctly calibrated, and cannot be considered as part of the analysis strategy. The update required to potentially mitigate this incorrect calibration of the clusters inside $R=1.0$ jets is not expected until an updated version of FastCaloSim (FastCaloSimV2) is available, which is at time of writing not available and will not form the part of any publications in this thesis.

7.4 Tau veto studies

Overview

In a given $t\bar{t}$ event in the signal region, it is possible for the top decay to produce a $t \rightarrow bW \rightarrow b + \tau^+ + \nu_\tau$ (or its charge conjugate process), such that a τ can arise in the final state. The case of particular interest is when the τ decays hadronically. To mitigate the impact of the $t\bar{t}$ background, such events must be vetoed. In the 36.1fb^{-1} analysis [3] this was performed using a selection on an overall event, namely based on the conditions that:

1. If any jet in the event is not b-tagged, is in acceptance and has $N_{\text{tracks}} < 4$, define the jet as a τ -candidate
2. If the $\Delta\phi(\tau_{\text{cand}}, E_T^{\text{miss}}) < \frac{\pi}{5}$, for any τ -candidate in the event, veto the event

This veto is also used to exclude incorrectly reconstructed jets which occur due to areas of low reconstruction performance in the calorimeter, which can also produce jets with a low number of tracks close to the E_T^{miss} . A limitation of this approach is that it is reliant on the number of tracks, and may not be considered stable under the inclusion of additional pileup tracks in the event. However, ATLAS has a number of alternative measures for directly identifying hadronically decaying τ 's, namely [107], which uses a multivariate classifier based on calorimeter and tracking information.

This identifier is primarily optimised for usage in searches where the τ final state is desired, with several choices of working point, “tight”, “medium”, “loose” and “very-loose” being related to the efficiency of identification, the τ fake rate and background rejection.

Note that the tighter working points improve the background rejection (hence lower fake rate) at the expense of signal efficiency. For a veto, a high signal efficiency of the identification is important (since a higher signal efficiency is equivalent to a stricter veto). A high fake rate means that a large number of events which do not actually contain a τ are vetoed in addition, which will affect the significance, but this may be a tolerable effect if the significance is sufficiently increased.

Performance comparison against existing tau veto approach

In this study, the original definitions of the signal region as per [3] were used, and the τ -veto applied therein was compared against either:

- No τ veto.
- $N_\tau = 0$ for a given working identification working point

These were checked in the context of the 36.1 fb^{-1} results, looking at the impact on the signal significance for benchmark signals in each signal region with each τ veto combination.

Sample (Region=SRAT0)	No τ -veto	$N_\tau=0$	Jet based veto (Default)
$t\bar{t}$	2.52 ± 0.94	1.78 ± 0.84	0.82 ± 0.63
$t\bar{t} + Z$	2.12 ± 0.20	1.75 ± 0.19	1.94 ± 0.19
Single top	5.75 ± 0.74	3.06 ± 0.52	2.25 ± 0.42
Multi-jet (QCD)	0.03 ± 0.03	-	0.03 ± 0.03
$Z + \text{jets}$	10.87 ± 0.76	9.73 ± 0.72	10.11 ± 0.73
$W + \text{jets}$	3.69 ± 0.56	2.59 ± 0.50	2.13 ± 0.4
Other	1.96 ± 0.44	1.34 ± 0.37	1.22 ± 0.28
Total SM	26.93 ± 1.60	20.26 ± 1.38	18.50 ± 1.18
Signals			
$(m_{\bar{t}}, m_{\bar{\chi}^0}) = (900, 100) \text{ GeV}$	$12.87 \pm 0.56 (Z_n = 1.42)$	$10.75 \pm 0.51 (Z_n = 1.46)$	$11.83 \pm 0.53 (Z_n = 1.71)$
$(m_{\bar{t}}, m_{\bar{\chi}^0}) = (800, 1) \text{ GeV}$	$21.44 \pm 0.48 (Z_n = 2.31)$	$18.36 \pm 0.45 (Z_n = 2.43)$	$19.92 \pm 0.46 (Z_n = 2.76)$
$(m_{\bar{t}}, m_{\bar{\chi}^0}) = (1100, 1) \text{ GeV}$	$4.46 \pm 0.23 (Z_n = 0.44)$	$3.79 \pm 0.21 (Z_n = 0.46)$	$3.04 \pm 0.22 (Z_n = 0.54)$

Table 7.9: Comparison tables in 2015-16 region SRA-T0 for the tau veto using a veto on very loose taus, no veto or the default jet based veto.

Sample (Region=SRBTT)	No τ -veto	$N_\tau=0$	Jet based veto (Default)
$t\bar{t}$	17.09 ± 2.30	12.39 ± 1.97	9.41 ± 1.70
$t\bar{t} + Z$	6.93 ± 0.43	5.80 ± 0.39	5.94 ± 0.40
Single top	13.19 ± 1.33	6.27 ± 1.03	5.36 ± 0.98
Multi-jet (QCD)	0.34 ± 0.14	0.27 ± 0.12	0.24 ± 0.12
$Z + \text{jets}$	9.99 ± 0.71	9.01 ± 0.69	8.94 ± 0.68
$W + \text{jets}$	8.12 ± 0.90	5.26 ± 0.74	3.85 ± 0.53
Other	3.72 ± 0.61	2.98 ± 0.60	2.33 ± 0.50
Total SM	59.36 ± 2.98	41.98 ± 2.55	36.07 ± 2.24
Signals			
$(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (700, 400)$ GeV	11.72 ± 0.56 ($Z_n = 0.66$)	9.75 ± 0.51 ($Z_n = 0.74$)	10.43 ± 0.52 ($Z_n = 0.91$)
$(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (600, 400)$ GeV	2.68 ± 0.74 ($Z_n = 0.05$)	2.19 ± 0.3 $7(Z_n = 0.07)$	2.15 ± 0.64 ($Z_n = 0.09$)

Table 7.10: Comparison tables in 2015-16 region SRB-TT for the tau veto using a veto on very loose taus, no veto or the default jet based veto.

Sample (Region=SRC1)	No τ -veto (Default)	$N_\tau=0$	Jet based veto
$t\bar{t}$	6.64 ± 1.39	3.35 ± 0.96	2.81 ± 0.89
$t\bar{t} + Z$	0.13 ± 0.07	0.10 ± 0.05	0.12 ± 0.05
Single top	0.78 ± 0.29	0.30 ± 0.10	0.25 ± 0.08
Multi-jet (QCD)	0.01 ± 0.01	0.01 ± 0.01	0.01 ± 0.01
$Z + \text{jets}$	0.47 ± 0.13	0.43 ± 0.13	0.40 ± 0.13
$W + \text{jets}$	0.93 ± 0.24	0.56 ± 0.18	0.56 ± 0.20
Other	0.26 ± 0.04	0.13 ± 0.07	0.20 ± 0.08
Total SM	9.22 ± 1.45	4.88 ± 1.00	4.33 ± 0.92
Signals			
$(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (500, 327)$ GeV	$10.17 \pm 0.63(Z_n = 2.26)$	$10.17 \pm 0.63(Z_n = 2.54)$	$6.66 \pm 0.51(Z_n = 0.48)$

Table 7.11: Comparison tables in 2015-16 region SRC1 for the tau veto using a veto on very loose taus, no veto or the default jet based veto.

Inclusion of the very-loose working point has some comparable performance to using the stop0L jet based τ -veto, but not comprehensive across all signal region bins, likely limited by statistical fluctuation. This is principally because in addition to the removal of τ -like jets, poorly reconstructed jets (which may be close to the E_T^{miss} , and consist of few tracks) may be vetoed in addition. This additional effect allows for reduction in other backgrounds besides those with τ s expected in the final state.

The conclusion for SRC1 is that a veto of very-loose BDT identified τ 's may have some improvement in significance. However, the assumed 20% systematic may be too small when calculating the significance, since this may be an under-estimate (thus minimising the effect when corrected). A jet based-veto will also not be effective here (for the $(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (500, 327)$ GeV signal), since the mass splitting and the ISR jet boost will mean that the majority of jets except the ISR will be relatively close to the E_T^{miss} (and relatively soft, with few tracks).

This effect would later be confirmed with generator level or “truth-level” studies, as will be seen in the next subsections.

Tau truth matching and truth composition

Truth matching, namely matching a detector reconstructed τ jet object to the output from the original record of the MC generator can be performed in the case of τ : a truth matched τ is defined as a reconstructed τ -jet with $p_T > 10$ GeV and in η acceptance, with $\Delta R(\text{reconstructed } \tau - \text{jet}, \text{truth-}\tau) < 0.2$. If there is any ambiguity in truth matching, the truth τ most consistent in p_T is chosen.

Efficiency of the tau veto

Since specific working points of the tau identification can be used to produce an event veto with comparable performance to that using jets, the efficiencies and fake rates can be investigated when looking at MC, since we have in addition to the reconstructed detector output, the true particle identity of the given physics object from the MC generator and parton shower. In this section, we denote two relevant definitions:

- Reconstructed hadronic τ -jets: Reconstructed using the ATLAS hadronic τ -lepton reconstruction procedure and identified using the MVA based taggers or the topological condition $N_{\text{trk}}(\text{obj}) < 4$, $\Delta(\text{obj}, E_T^{\text{miss}}) < \Pi/5$.
- Reconstructed (as per the ATLAS hadronic τ reconstruction procedure, but not necessarily identified) τ -jet objects truth matched to a tau as per the previous subsection. Note identification at this stage is not necessary.

For each bin in p_T of the reconstructed τ identified with the respective identification procedure, we calculate the following:

$$\text{Fake rate} = \frac{N(\text{reconstructed } \tau | \text{ Not truth matched to } \tau)}{N(\tau \text{ not matched to a truth } \tau)} \quad (7.1)$$

$$\text{Efficiency} = \frac{N(\text{reconstructed } \tau | \text{ Truth matched})}{N(\tau \text{ Truth matched})} \quad (7.2)$$

to see which working point was maximally efficient for a given bin in τp_T .

These efficiencies and fake rates are then presented for the $t\bar{t}$ background for a loose preselection region ($0 < e/\mu$, $E_T^{\text{miss}} > 250$ GeV, $N_j \geq 4$ ($N_b \geq 2$), $p_T^{j_2} > 80$ GeV, $p_T^{j_4} > 40$ GeV) in figures 7.4, 7.5b.

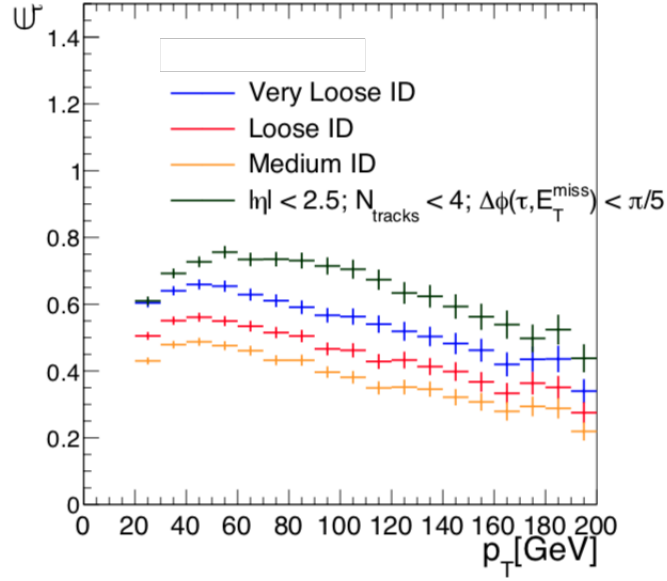
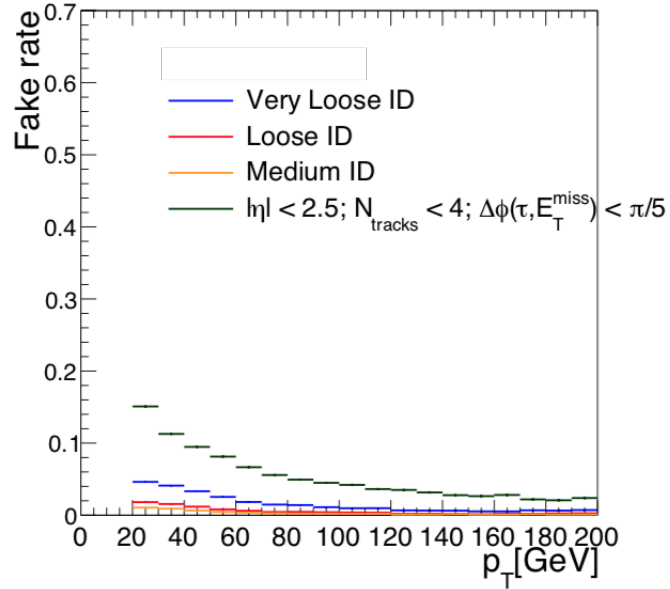


Figure 7.4: comparison of the efficiencies of the respective MVA based τ ID working points (including applying the topological selection to hadronic τ -jets) as a function of p_T^τ for the $t\bar{t}$ background.



(a)

(b) 7.5a: Comparison of the fake rates of the respective MVA based τ ID working points (including applying the jet-based veto to the hadronic τ -jets) as a function of p_T^τ for the $t\bar{t}$ background.

Impact on the tau veto on truth level composition

An important potential source of τ 's is the $t\bar{t}$ background, thus using truth level output the composition can be determined. The semi-leptonic $t\bar{t}$ background also can contribute to the all-hadronic signal region if the lepton falls out of acceptance, and the impact of this background source was observed. Since $t\bar{t}$ has an impact in the boosted, intermediate and compressed regions (particularly in the last case as the leading background), this study was undertaken with the truth leptons required to be originating from a Top/W emission (this origin information is saved during the simulation). The impact of each τ veto method on reconstructed objects was considered by the impact observed on the truth level $t\bar{t}$ background composition. The background was subdivided into different truth classifications based on number of leptons/ τ -s originating from the top/W in the event, while all cases where no truth lepton or τ associated with the top or W was present were treated as unclassified.

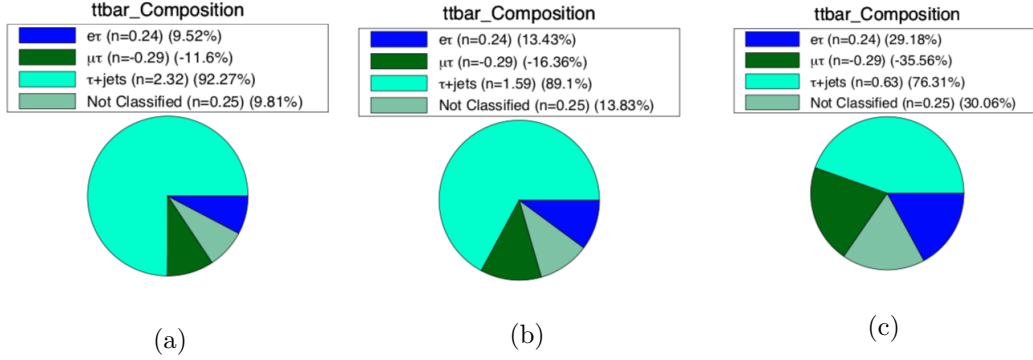


Figure 7.6: Truth level composition of the $t\bar{t}$ background in signal region SRAT0, when considering no τ -veto (7.6a), using the BDT based veto (7.6b) and using the jet based τ -veto (7.6c).

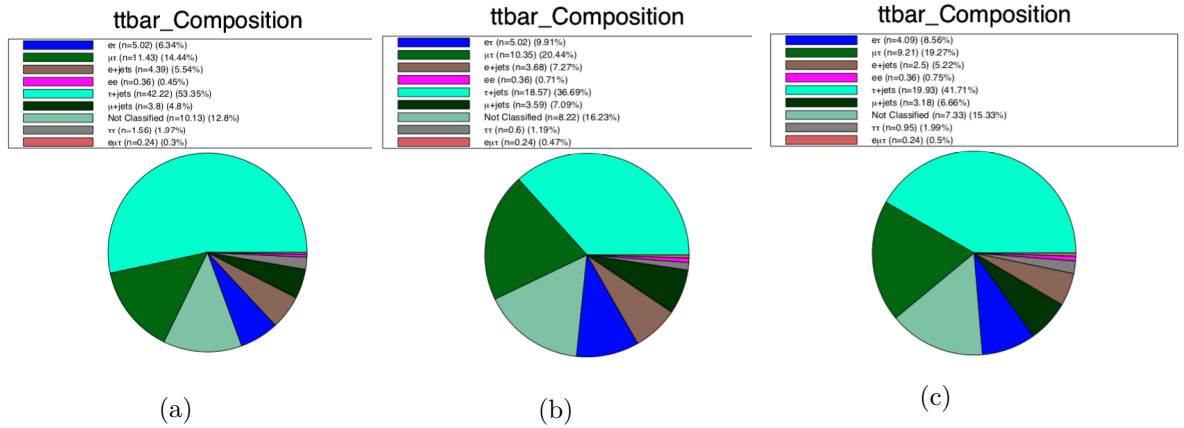


Figure 7.7: Truth level composition of the $t\bar{t}$ background in signal region SRBT0, when considering no τ -veto (7.7a), using the BDT based veto (7.7b) and using the jet based τ -veto (7.7c).

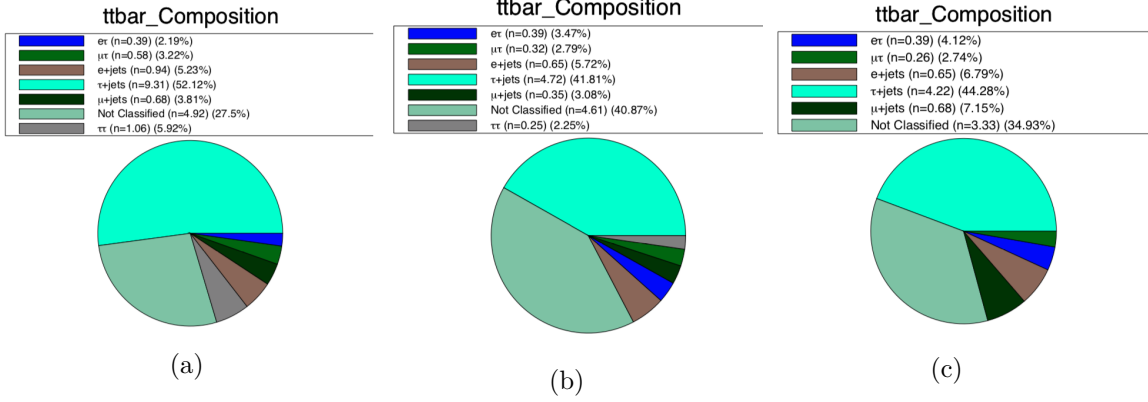


Figure 7.8: Truth level composition of the $t\bar{t}$ background in signal region SRC1, when considering no τ -veto (7.8a), using the BDT based veto (7.8b) and using the jet based τ -veto (7.8b).

Conclusion

The τ veto based on BDT based identification of hadronically decaying τ leptons has comparable or worse performance for all working points compared to the veto based on jets, because in addition to a veto on τ , the angular selection on jets also vetoes events containing mis-measured jets close to the E_T^{miss} , which can arise from say a jet or its constituents passing through a poorly reconstructed region in the detector.

Alternate variables: $M_T(\tau, E_T^{\text{miss}})$

In the definition of a jet-based τ veto, a tau candidate was defined as having $n\text{Tracks} \leq 4$, and then as a global veto, if any τ candidate had $\Delta\phi(\tau, E_T^{\text{miss}}) < \frac{\pi}{5}$, the event would be vetoed. In a semi-leptonic $t\bar{t}$ event with a τ , there will be a leptonically decaying W into a $\tau + \bar{\nu}_\tau$, a source of real E_T^{miss} (and assumed to be the dominant source given the large pre-selection requirement), so the transverse mass $M_T(\tau, E_T^{\text{miss}})$ can be used to reconstruct the W system, such that if a system is compatible with a $W \rightarrow \tau\nu$, then it will be vetoed. To measure the performance of these variables on a one dimensional plot, since an event can contain more than one τ candidate event (since $N_{\text{jets}} \geq 4$), the minimum $M_T(\tau, E_T^{\text{miss}})$ value was considered in the case of multiple candidate events, with a similar minimisation for the $\Delta\phi(\tau, E_T^{\text{miss}})$. This was checked at a loose preselection level, namely: $E_T^{\text{miss}} > 250$, $\min[\Delta\phi(j_{1-4}, E_T^{\text{miss}})]4 > 0.4$, $N_\ell = 0$, $N_b \geq 1$, $N_{\text{jets}} \geq 4$, and for an investigation of particular selection, an initial selection of $M_T(\tau - \text{cand}, E_T^{\text{miss}})_{\text{min}} < 80$ was trialled.

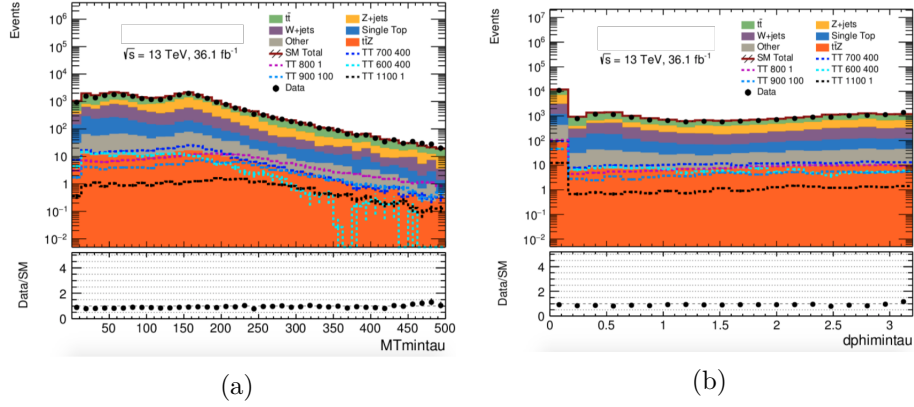


Figure 7.9: $M_T(\tau - \text{cand}, E_T^{\text{miss}})_{\min}$ 7.9a and $\Delta\phi(\tau - \text{cand}, E_T^{\text{miss}})_{\min}$ 7.9b in the preselection $E_T^{\text{miss}} > 250, \min[\Delta\phi(j_{1-4}, E_T^{\text{miss}})]_4 > 0.4, N_\ell = 0, N_b \geq 1, N_{\text{jets}} \geq 4$. Note that the $\Delta\phi(\tau - \text{cand}, E_T^{\text{miss}})_{\min}$ was defined by construction to be 0 if there were no such τ candidates in the event .

Preselection	
$t\bar{t}$	17710.12 ± 72.50
$t\bar{t} + Z$	336.63 ± 2.87
Single top	3077.71 ± 25.34
$Z + \text{jets}$	12056.26 ± 163.35
$W + \text{jets}$	7831.80 ± 59.55
Other	1229.72 ± 13.38
Total SM	42242.24 ± 191.42
Signals	
$(m_{\tilde{t}}, m_{\tilde{\chi}_0}) = (900, 100) \text{ GeV}$	$172 \pm 2.10 (Z_n = -0.11)$
$(m_{\tilde{t}}, m_{\tilde{\chi}_0}) = (800, 1) \text{ GeV}$	$350.62 \pm 1.98 (Z_n = -0.09)$
$(m_{\tilde{t}}, m_{\tilde{\chi}_0}) = (700, 400) \text{ GeV}$	$431.83 \pm 3.39 (Z_n = -0.11)$
$(m_{\tilde{t}}, m_{\tilde{\chi}_0}) = (600, 400) \text{ GeV}$	$237.85 \pm 5.96 (Z_n = -0.10)$
$(m_{\tilde{t}}, m_{\tilde{\chi}_0}) = (1100, 1) \text{ GeV}$	$44.53 \pm 0.74 (Z_n = -0.13)$

(a)

Preselection + $\min M_T(\tau - \text{cand}, E_T^{\text{miss}}) > 80 \text{ GeV}$	
$t\bar{t}$	11764.49 ± 59.09
$t\bar{t} + Z$	269.52 ± 2.52
Single top	1909.82 ± 20.24
$Z + \text{jets}$	10449.35 ± 150.16
$W + \text{jets}$	5277.11 ± 49.79
Other	924.95 ± 11.26
Total SM	30595.24 ± 170.47
Signals	
$(m_{\tilde{t}}, m_{\tilde{\chi}_0}) = (900, 100) \text{ GeV}$	$150.10 \pm 1.96 (Z_n = -0.11)$
$(m_{\tilde{t}}, m_{\tilde{\chi}_0}) = (800, 1) \text{ GeV}$	$310.73 \pm 1.87 (Z_n = -0.09)$
$(m_{\tilde{t}}, m_{\tilde{\chi}_0}) = (700, 400) \text{ GeV}$	$342.71 \pm 3.01 (Z_n = -0.11)$
$(m_{\tilde{t}}, m_{\tilde{\chi}_0}) = (600, 400) \text{ GeV}$	$163.75 \pm 4.96 (Z_n = -0.10)$
$(m_{\tilde{t}}, m_{\tilde{\chi}_0}) = (1100, 1) \text{ GeV}$	$39.10 \pm 0.69 (Z_n = -0.13)$

(b)

Table 7.12: Preselection event yield before and after an application of a $M_T(\tau - \text{cand}, E_T^{\text{miss}})_{\min} > 80 \text{ GeV}$ selection .

It is possible to observe that there is no substantial gain in signal significance at preselection level from applying a tau-veto based on the $M_T(\tau, E_T^{\text{miss}})_{\min}$ variable over that based on an angular separation between the E_T^{miss} and τ . The additional performance from a selection of $\Delta\phi(\tau, E_T^{\text{miss}})_{\min}$ arises from the additional vetoing of mismodelled jets in addition to the vetoing of events due to taus.

7.5 Hybrid B-tagging

Since the MV2c10 algorithm [69] forms the heavy flavour tagging component of the signal regions in [3], alternatives to the configuration of a fixed cut 77% efficient working point can be considered. Since the observables in the MV2c10 input have p_T dependence, a selection on the MV2c10 classifier may vary in efficiency as a function of p_T (see [69] for a plot of the efficiency as a function of p_T). Qualitatively, the efficiency falls off for large p_T above some threshold, thus high p_T b-jets are less likely to be identified. To mitigate

this behaviour a p_T -dependent selection on the MV2c10 score was trialled, combining the low p_T performance of the fixed cut working point and another tagger for high p_T that could maintain identification efficiency at high p_T . This hybrid approach for b-jets with $p_T^b < \lesssim 200$ GeV would be identical to a fixed selection on the MV2c10 output as per default (with an expected efficiency of approximately 77%), but for b-jets with $p_T^b > \sim 200$ GeV, the selection of the MV2c10 output would vary as a function of p_T such that the signal efficiency would remain at 77% regardless of p_T . The following selections were chosen (the definition of the b-jet depends on whether the hybrid or fixed selection was used). The following study was undertaken using simulated samples only, comparing the MV2c10 fixed selection against hybrid for $t\bar{t}, Z + \text{jets}$ and a signal with $(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (800, 1)$ GeV (only a single signal is used since this study is focussing on the key distributions which may arise from the choice of generator rather than particular model kinematics).

- A loose preselection: $E_T^{\text{miss}} > 200$ GeV, ≥ 2 b-jets, ≥ 4 jets, $\min[\Delta\phi(j_{1-4}, E_T^{\text{miss}})] > 0.4$, $N_\ell = 0$.
- SRA without top categories (table 7.13):

Variable	Selection (SRA-type signal region)
Trigger	E_T^{miss} trigger
E_T^{miss}	> 400 GeV
N_ℓ	$= 0$
p_T^{j2}	> 80 GeV
p_T^{j4}	> 40 GeV
$\min[\Delta\phi(j_{1-4}, E_T^{\text{miss}})]3$	> 0.4
$E_T^{\text{miss,track}}$	> 30
$ \Delta\phi(E_T^{\text{miss}}, E_T^{\text{miss,track}}) $	$< \pi/3$
τ -veto	$\checkmark$
$N_{b\text{-jet}}$	≥ 2
$m_{T, \text{b,min}}^1$	> 200
$m_{\text{jet}, R=0.8}^1$	> 60 GeV
$m_{\text{jet}, R=1.2}^1$	> 120 GeV
m_{T2, χ^2}	> 400 GeV

Table 7.13: SRA-type region used for hybrid b-tagging studies.

The distributions of several variables with these selections applied and the b-tagging schemes used are presented in figures 7.10, 7.11.

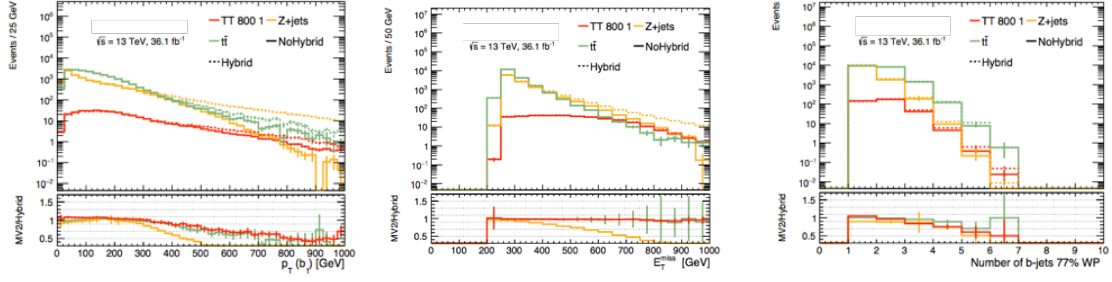


Figure 7.10: Distributions of kinematic variables in the 0L preselection, comparing the influence of the hybrid b-tagger against the fixed MV2c10 selection. Note that variables which do not explicitly depend on b-jets are impacted by the inclusion of a $N_{b\text{-jet}} \geq 2$ selection in the preselection.

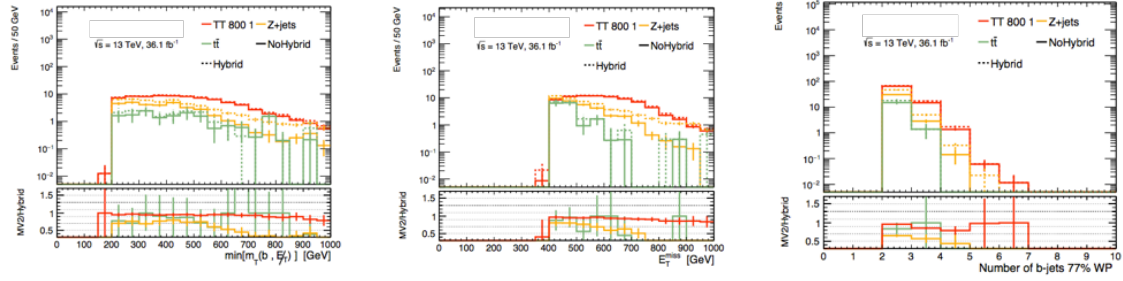


Figure 7.11: Distributions of kinematic variables in the SRA-type signal region defined in table 7.13, comparing the influence of the hybrid b-tagger against the fixed MV2c10 selection. Note that variables which do not explicitly depend on b-jets are impacted by the inclusion of a $N_{b\text{-jet}} \geq 2$ selection in the preselection.

Since the $Z + \text{jets}$ background seems to be substantially increased with respect to the fixed cut (unlike for the signal and $t\bar{t}$ background, which are generally in agreement), a study was undertaken into the predominant source of such a large number of high p_T b-jets. Since ATLAS provides simulated samples corresponding to different generator level filtering (chosen to increase statistics for a given process), the generator level filter can be used to identify what class of diagrams will be impacted by the hybrid b-tagging configuration. The generator level filters available are the b & c quark veto (CVetoBVeto), filtered only to include b-jets (BFilt), and veto the b-jets while including the c-jets (CFilterBVeto), and the distributions of these can be plotted. This generator level filtering can also be used to ascertain the mis-tag rate (since any b-jets originating from samples where b-jets are vetoed would correspond to fake b-jets). Figure 7.12 demonstrates the results.

Conclusion

For jet substructure observables, the major limitation of the study is the mismodelling of the jet constituents in AFII fast simulation (which would require the not currently available FastCaloSimV2 toolset). Since simulating signals in full simulation at time of writing during a period of high Grid CPU utilisation was not viable.

Studies into the veto of τ leptons was also promising, however a jet based veto out-performed object identification vetoes because of the supplementary effect of vetoing events containing mis-measured jets, in addition to the removal of τ s. In the compressed case, any gains made from applying a τ veto may be substantially offset by a loss of statistics (particularly in very low statistics regions such as the 2015-16 SRC5) and a large theoretical uncertainty on the $t\bar{t}$ background. Hybrid b-tagging was investigated and found to have a large fake-rate in the high p_T tails. Although this increases statistics overall, the loss in significance using this classifier over the standard MV2c10 fixed cut was not tolerable.

All of the studies presented in this chapter were not used in publication for [2] because of the limitations presented in this chapter. However, when aided with the relevant calibrations, the methods discussed here can be considered as a promising baseline for future searches, particularly those oriented toward machine-learning based searches, or more complex analysis strategies as compared to the “cut and count” analysis.

Progress toward 139fb^{-1} result.

This chapter outlines the work undertaken by the author in preparation/optimisation for the publication [2], and the studies presented here form an integral part of the design of the publication and the respective regions. The author first discusses the implementation of new variables in the pre-selection to mitigate the multi-jet background, then focusses on the design of control regions for the $t\bar{t} + Z$ (re-interpreting a signal region for measurement as a control region for a search) and $Z + \text{jets}$ backgrounds. In addition, taking the lead from the observation of the $t\bar{t} + Z$ process, additional background contributions are considered to the $t\bar{t} + Z$ control regions, namely discussing the need/lack thereof for ancilliary control regions.

Following this chapter, given all of the preliminary work undertaken here, chapter 9 will discuss the full publication of [2], and chapter 10 will outline the final results.

8.1 New variable studies using receiver operating characteristic (ROC) scans

Overview

In this subsection, the discrimination performance of different observables was considered, optimising the acceptance of signal and background rejection as a function of different values of the observable of interest, namely in reduction of the multi-jet background. The performance of each observable was measured using a Receiver Operating Characteristic (ROC) curves, which are particularly useful in classification problems in Machine Learning, which were produced by observing the signal acceptance of a benchmark signal against a given background, with iterative selections on the variables of interest. The particular background of interest to reject was the multi-jet (QCD) background, which is particularly prevalent in signal region SRC (SRA, SRB generally remove this through a selection on $\min[\Delta\phi(j_{1-4}, E_T^{\text{miss}})]$), and a relatively subdominant effect in the other signal regions.

The variables used in this study are:

- Object based E_T^{miss} significance [108]:

$$\text{Object based } E_T^{\text{miss}} \text{ sig.} = \frac{|E_T^{\text{miss}}|}{\sqrt{\sigma_L^2(1 - \rho_{LT}^2)}} \quad (8.1)$$

where the σ_L^2 is the longitudinal resolution of all objects in the event as a function of p_T of each object, and ρ_{LT} denotes the correlation between longitudinal and transverse resolutions. The more general [108] (although less tractable) form of this significance based on the likelihood $\mathcal{L}$ of two dimensional parameter p_T^{inv} :

$$\text{Object based } E_T^{\text{miss}} \text{ sig.} = \sqrt{2 \ln \left(\frac{\max_{p_T^{\text{inv}} \neq 0} \mathcal{L}(E_T^{\text{miss}} | p_T^{\text{inv}})}{\max_{p_T^{\text{inv}} = 0} \mathcal{L}(E_T^{\text{miss}} | p_T^{\text{inv}})} \right)} \quad (8.2)$$

- $\min[\Delta\phi(j_{1-2}, E_T^{\text{miss}})], \min[\Delta\phi(j_{1-3}, E_T^{\text{miss}})], \min[\Delta\phi(j_{1-4}, E_T^{\text{miss}})]:$ variables already used for mitigation of the multi-jet background.
- $\text{metsigHT} \left(\frac{E_T^{\text{miss}}}{\sqrt{H_T}} \right):$ This considers the significance of the E_T^{miss} given a resolution of $\sqrt{H_T}$, and is useful when the largest resolution effects originate from jet measurements.
- $\text{metsigET} \left(\frac{E_T^{\text{miss}}}{\sqrt{\sum E_T}} \right):$ This considers the significance of the E_T^{miss} given it has a resolution $\sqrt{\sum E_T}$ to cover overall resolution effects. $\sum E_T$ differs from the H_T through the inclusion of the tracks and soft terms.

ROC scans

This study was performed using a loose preselection: ($p_T^{j_4} > 50$, $E_T^{\text{miss}} > 250$, $N_b \geq 2$, $N_{\text{jets}} \geq 4$, $N_\ell = 0$) comparing a target signal against a given background composition, checking the signal acceptance and background rejection. The backgrounds considered were:

- Di-jet - Simulated estimate of the QCD background instead of the data-driven background estimate used in [3].
- JZ6 sample: Specific generator level slicing of the Di-jet background. This is the largest sub-sample of the di-jet background. Given the generator level slicing of this sample, it is the principle sub-sample which would be contributing to the di-jet background in SRC.
- $t\bar{t}$, Z + jets and W + jets.

For reference, the distributions of the variables used in the study are plotted in figure 8.1. The receiver operating characteristic curves (ROC) of each of the variables selected are presented in figures 8.2-8.4, comparing the performance of a selection on a particular variable between a signal which is a benchmark signal for signal regions SRA, SRB or SRC and a given background. This study also follows the same reasoning as a similar study in [108].

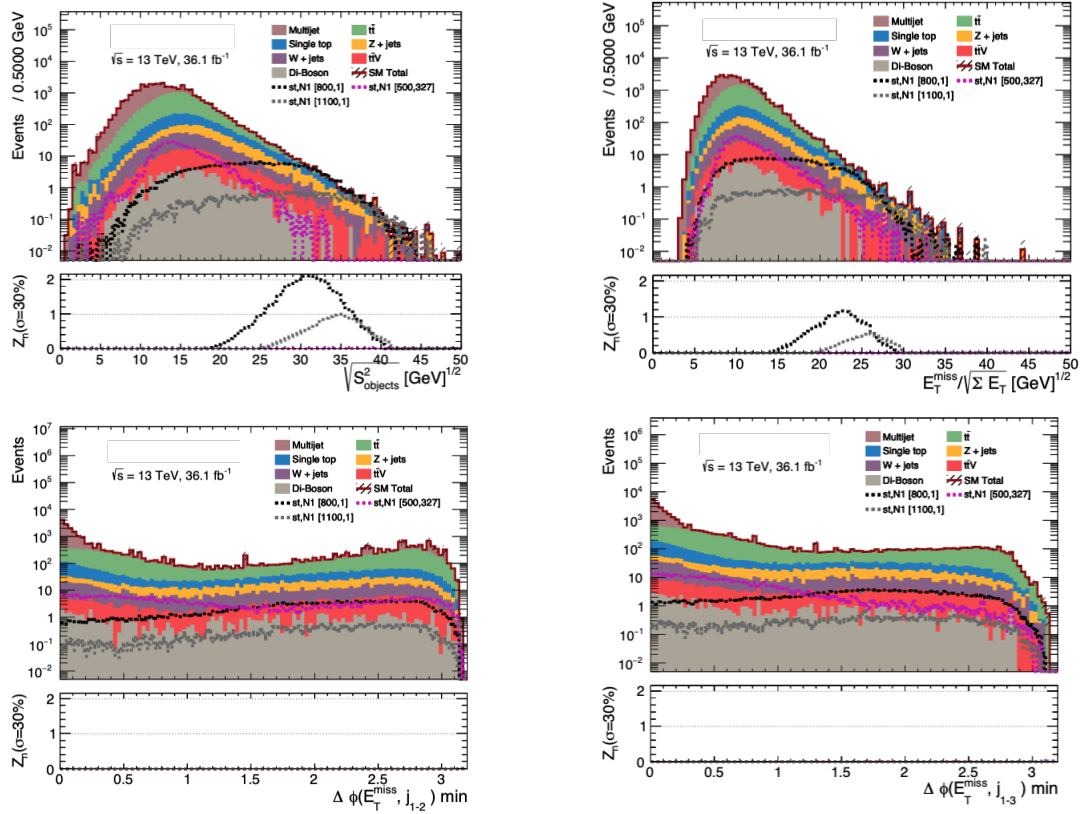


Figure 8.1: Preselection distributions of several variables used in the receiver operating curve scans. The range and bin intervals chosen in these plots match those for the ROC scans.

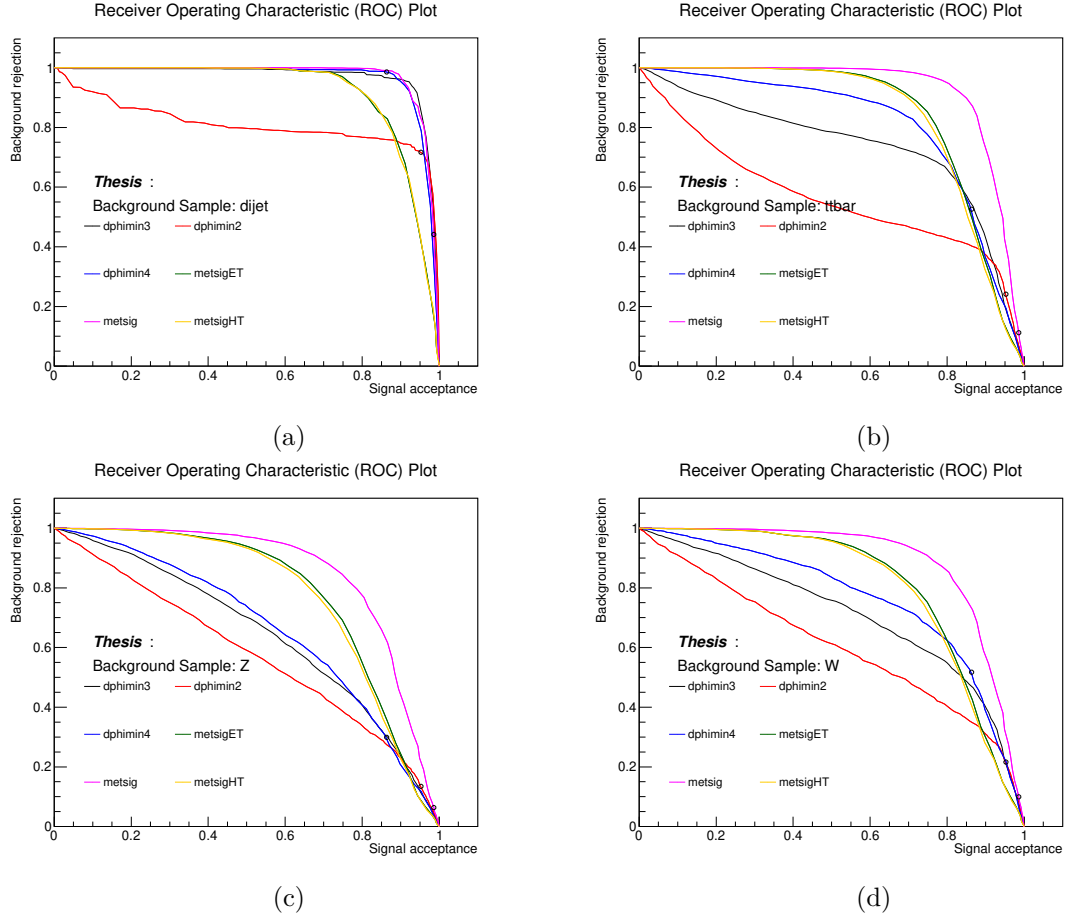


Figure 8.2: Receiver operating scans comparing di-jet (8.2a), $t\bar{t}$ (8.2b), $Z + \text{jets}$ (8.2c) and $W + \text{jets}$ (8.2d) backgrounds against the $m_{\tilde{t}}, m_{\tilde{\chi}^0} = (1100, 1)$ signal point, which corresponds to an SRA benchmark signal. A fixed selection of $\min[\Delta\phi(j_{1-4}, E_T^{\text{miss}})] > 0.4$, $\min[\Delta\phi(j_{1-2}, E_T^{\text{miss}})] > 0.4$ or Object based $E_T^{\text{miss}} \text{ sig.} > 11$ was plotted to show the acceptance and rejection of such fixed cuts.

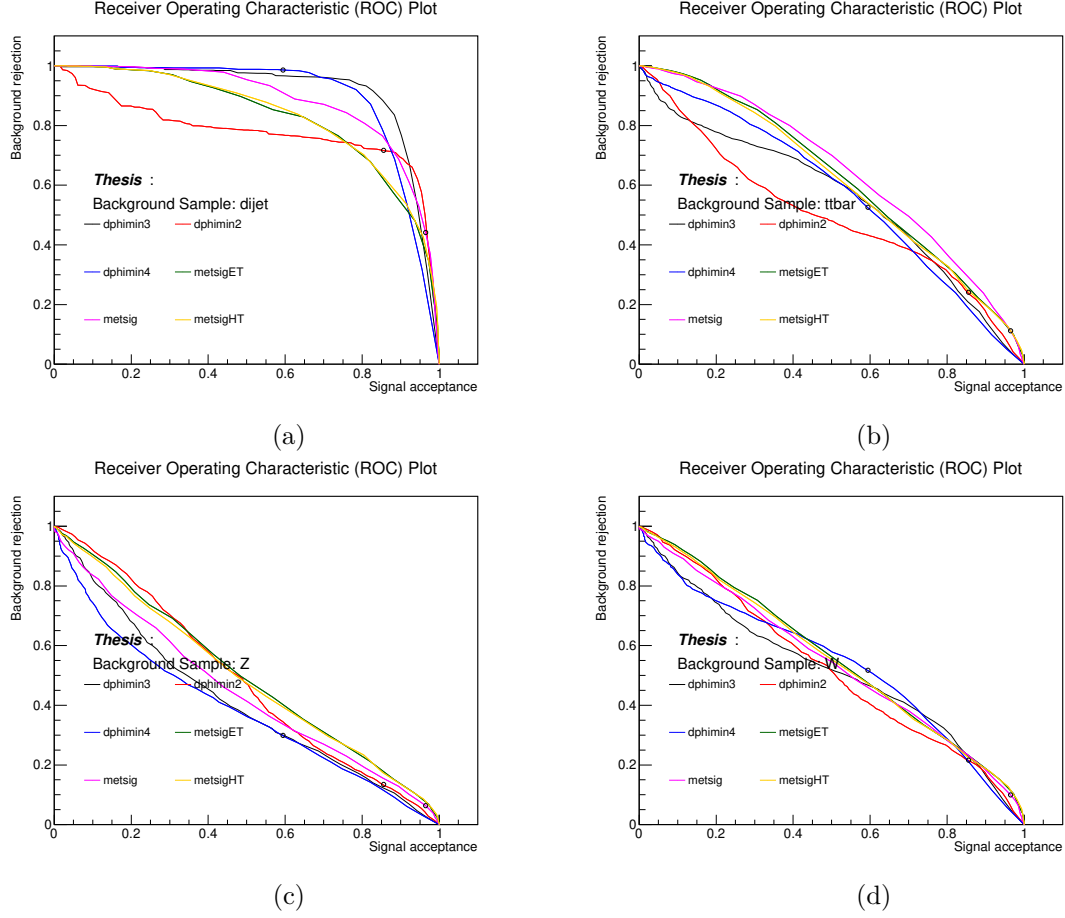


Figure 8.3: Receiver operating scans comparing di-jet (8.3a), $t\bar{t}$ (8.3b), $Z + \text{jets}$ (8.3c) and $W + \text{jets}$ (8.3d) backgrounds against an SRB benchmark signal point $(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (600, 400)$ GeV. A fixed selection of $\min[\Delta\phi(j_{1-4}, E_T^{\text{miss}})] > 0.4$ or Object based $E_T^{\text{miss}} \text{ sig.} > 11$ were additionally plotted (circled) to show the acceptance and rejection of such fixed cuts.

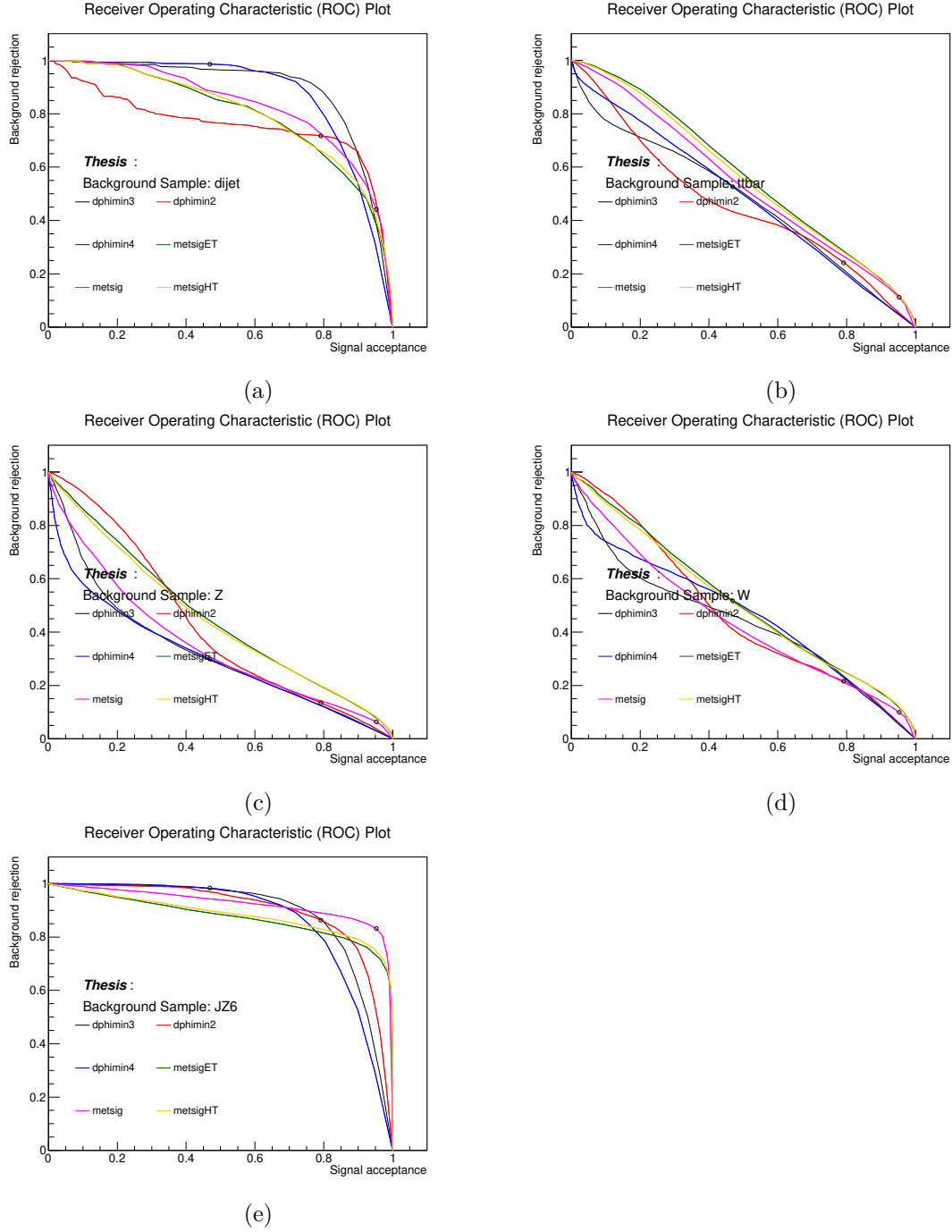


Figure 8.4: Receiver operating scans comparing di-jet (8.4a), JZ6 (8.4e), $t\bar{t}$ (8.4b), $Z + \text{jets}$ (8.4c) and $W + \text{jets}$ (8.3c) backgrounds against an SRC benchmark signal point $(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (500, 327)$ GeV. A fixed selection of $\min[\Delta\phi(j_{1-4}, E_T^{\text{miss}})] > 0.4, \min[\Delta\phi(j_{1-2}, E_T^{\text{miss}})] > 0.4$ and Object based $E_T^{\text{miss}} \text{ sig.} > 11$ were plotted to show the acceptance and rejection of such fixed cuts.

Overall, this discussion demonstrates that the Object based E_T^{miss} sig. is a performant variable both in discrimination of the QCD background against signal, as well as other background processes, particularly for signal regions SRA and SRB type signals. The ROC plots suggest that the Object based E_T^{miss} sig. should be used to supplement the $\min[\Delta\phi(j_{1-4}, E_T^{\text{miss}})]$ selections.

8.2 Control region specific optimisations: $CRt\bar{t} + Z\text{-AB } (t\bar{t} + Z)$ design

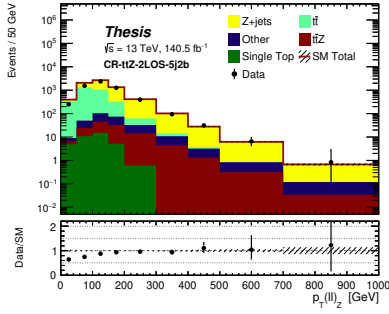
Considering the final state of two top quarks and missing energy, it is readily observed that $t\bar{t} + Z$, where $Z \rightarrow \nu\nu$, should be considered an irreducible background to a search for SUSY involving $t\bar{t}$ pair and missing energy. Thus the normalisation and modelling of this irreducible background is of particular importance. As a first attempt to understand the process, it is possible to consider the process $t\bar{t} + Z$, where $Z \rightarrow \ell\ell$ ($\ell = e, \mu$), as a model for $t\bar{t} + Z(\rightarrow \nu\nu)$. However, in [3], due to low statistics available in the high $p_T(Z)$ tails for aMC@NLO +PYTHIA 8 in the 36.1fb^{-1} regime, an estimate of the normalisation was undertaken using $t\bar{t} + \gamma$ in a single lepton channel with a single photon. Thus the visible photon could be used as a replacement for the Z boson.

In light of increased luminosity, and the measurement [109] of the $t\bar{t} + Z(\rightarrow \ell\ell)$, ($\ell = e, \mu$) cross-section in the tri-lepton channel, $t\bar{t} + Z \rightarrow \ell\ell$ was again considered. This section will discuss the formulation of such a control region, re-interpreting the signal region used for a measurement as a candidate control region for $t\bar{t} + Z$. Throughout this section, we will require that for a Z-associated pair of leptons is present in the event, namely that a pair of electrons/muons is classed as “opposite sign, same flavour”(OSSF) and that for this candidate pair $|m(\ell, \ell) - m(Z)| < 10$ GeV, choosing the minimum $\Delta M(\ell\ell, Z)$ combination in the case of ambiguities. $p_T^{\ell\ell}$ denotes the p_T of this candidate electron/muon pair.

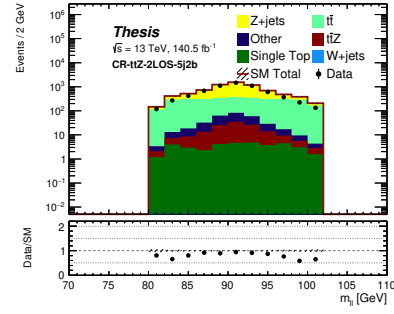
8.2.1 Formulation 1: Di-leptonic $CRt\bar{t} + Z$

This final state is desirable as it would be expected to have identical kinematics for $t\bar{t} + Z(\rightarrow \ell\ell)$ to that of $t\bar{t} + Z(\rightarrow \nu\nu)$. Strictly, the two processes differ in that di-leptonic $t\bar{t} + Z$, diagrams which can include FSR photon radiation from the leptons, but this effect is not substantial. A preselection was initially applied, following the prescription of the di-leptonic OSSF signal region of [109], in considering the cases of $= 5$ jet, ≥ 2 b-jet (2LOS-5j2b) and ≥ 6 jet, ≥ 2 b-jet (2LOS-6j2b). In addition, these preselections were further restricted by applying a selection of $p_T^{\ell\ell} > 200$ GeV, to minimise the extrapolation over the $p_T(Z)$ to the signal region. Table 8.1 outlines the initial selections, figures 8.5, 8.6 outline the $m(\ell, \ell)$ and $p_T^{\ell\ell}$ distributions and table 8.2 denote the event level yields.

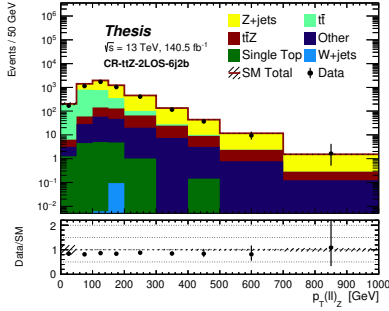
Selection	CR-ttZ 2LOS (5j2b)	CR-ttZ 2LOS(5j2b, $p_T(\ell)_Z > 200$)	CR-ttZ 2LOS (6j2b)	CR-ttZ 2LOS (6j2b, $p_T(\ell)_Z > 200$)
Trigger	Single e, μ			
Signal Leptons	==2			
Z like OSSF pair?	Yes			
N_b	≥ 2			
Lepton p_T 's	$>(30,20)$ GeV			
N_{Jets}	==5		≥ 6	
$p_T(\ell)_Z$	-	>200 GeV	-	>200 GeV

 Table 8.1: Region definitions for a 2LOS control region for $t\bar{t} + Z$


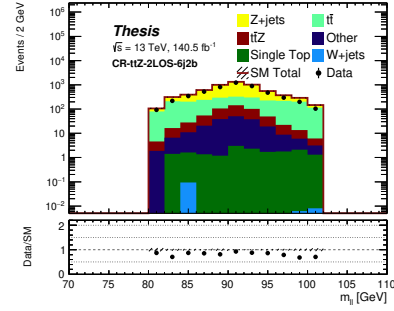
(a)



(b)

 Figure 8.5: Distributions for 8.5a $p_T^{\ell\ell}$ and 8.5b $m(\ell, \ell)$ in the region “CR-ttZ-2LOS-5j2b”.


(a)



(b)

 Figure 8.6: Distributions for 8.6a $p_T^{\ell\ell}$ and 8.6b $m(\ell, \ell)$ in the region “CR-ttZ-2LOS-6j2b”.

CR- $t\bar{t}Z$ -2LOS-5j2b	
Z+jets	4382.70 ± 31.82
$t\bar{t}$	2843.66 ± 28.30
$t\bar{t} + Z$	113.63 ± 0.78
Single Top	37.13 ± 3.06
W+jets	0.00 ± 0.00
Other	183.73 ± 3.41
SM	7560.85 ± 42.84
Data	6476.00 ± 80.47
CR- $t\bar{t}Z$ -2LOS-6j2b	
Z+jets	3711.97 ± 35.55
$t\bar{t}$	1864.28 ± 22.65
$t\bar{t} + Z$	342.76 ± 1.46
Single Top	17.79 ± 2.10
W+jets	0.11 ± 0.10
Other	186.60 ± 3.45
SM	6123.51 ± 42.38
Data	5181.00 ± 71.98

CR- $t\bar{t}Z$ -2LOS-5j2b-pTl1200	
Z+jets	943.51 ± 9.27
$t\bar{t}$	68.84 ± 4.36
$t\bar{t} + Z$	40.83 ± 0.43
Single Top	1.23 ± 0.56
W+jets	0.00 ± 0.00
Other	54.90 ± 2.03
SM	1109.31 ± 10.47
Data	1074.00 ± 32.77

CR- $t\bar{t}Z$ -2LOS-6j2b-pTl1200	
Z+jets	1043.45 ± 8.30
$t\bar{t}$	83.02 ± 4.81
$t\bar{t} + Z$	143.22 ± 0.93
Single Top	2.34 ± 0.78
W+jets	0.00 ± 0.00
Other	63.49 ± 2.16
SM	1335.52 ± 9.91
Data	1159.0 ± 34.04

 Table 8.2: Yields tables for the 2LOS $t\bar{t} + Z(\rightarrow \ell\ell)$ control regions listed in Table 8.1

From these results, it can be observed that the $Z + \text{jets}$ and di-leptonic $t\bar{t}$ backgrounds dominate in this region, instead of $t\bar{t} + Z(\rightarrow \ell\ell)$, thus the purity is low. This is mitigated in [109] using a neural network based on high-level jet observables, but this approach would not be preferable for a control region, namely when considering the extrapolation from the control region to the signal region.

8.2.2 Formulation 2: Trileptonic control region for $t\bar{t} + Z$

Following on from [109], the case of trileptonic $t\bar{t} + Z(\rightarrow \ell\ell)$ was considered, inspired by the inclusive “3L-2b-4j” signal region. In this region, three signal leptons were required, of which one OSSF pair consistent with the Z mass exists and at least 4 jets, of which at least 2 are b-tagged. This selection has a good purity by default, substantially reducing the contamination of $t\bar{t}$ and $Z + \text{jets}$ compared to the 2LOS case, however unlike the di-leptonic case there is some difference in event kinematics (and hence cross section) between the semi-leptonic $t\bar{t} + Z(\rightarrow \ell\ell)$ and all-hadronic $t\bar{t} + Z(\rightarrow \nu\nu)$, unlike the comparison of hadronic $t\bar{t} + Z(\rightarrow \ell\ell)$ and $t\bar{t} + Z(\rightarrow \nu\nu)$. This also incurs a different background profile containing backgrounds which are sub-dominant in the signal regions (WZ, tWZ, tZ, and other processes with a mis-identified lepton), but these alternate backgrounds will have a small contribution to the final definition of the control region (as covered later in section 8.2.4).

In order use a tri-leptonic $t\bar{t} + Z(\rightarrow \ell\ell)$ region to model all-hadronic $t\bar{t} + Z(\rightarrow \nu\nu)$, the $t\bar{t}$ component and the $p_T(Z \rightarrow \ell\ell)$. To model the Zp_T , we define the Z as the Z -associated pair of electrons/muons. The remaining electron/muon and the E_T^{miss} is then associated to the semi-leptonic $t\bar{t}$ component. In order to use the semi-leptonic $t\bar{t}$ component to model the hadronic $t\bar{t}$ component, we treat the objects not associated to the Z boson (including the E_T^{miss} , and the electron/muon not associated to the Z boson) as a jet. Given this, we apply jet p_T selections on this extended jet collection which are equivalent to those in the signal region, using the following algorithm:

1. Add the $\mathbf{p}_T^{\text{miss}}$ and the non Z -associated lepton to the collection of jets in the event.
2. Sort the extended jet collection in descending order in p_T .
3. Require the p_T of the sub-leading “jet” to be > 80 GeV and the p_T of the fourth “jet” to be > 40 GeV

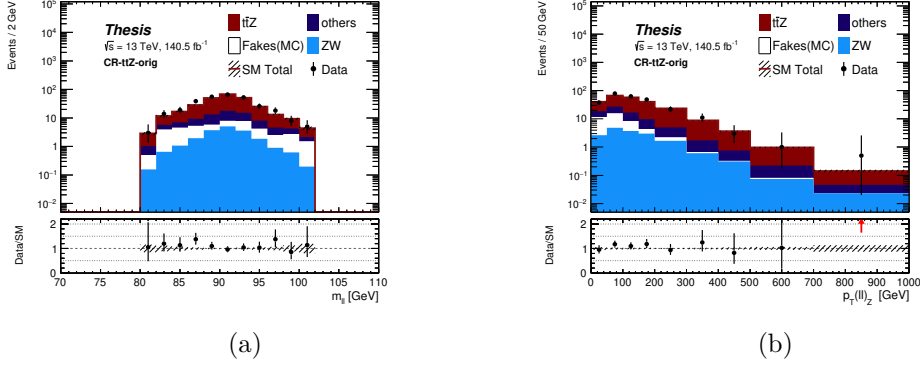
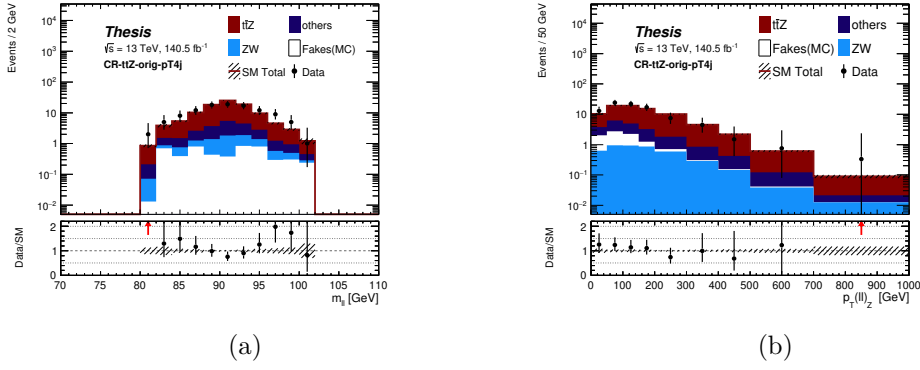
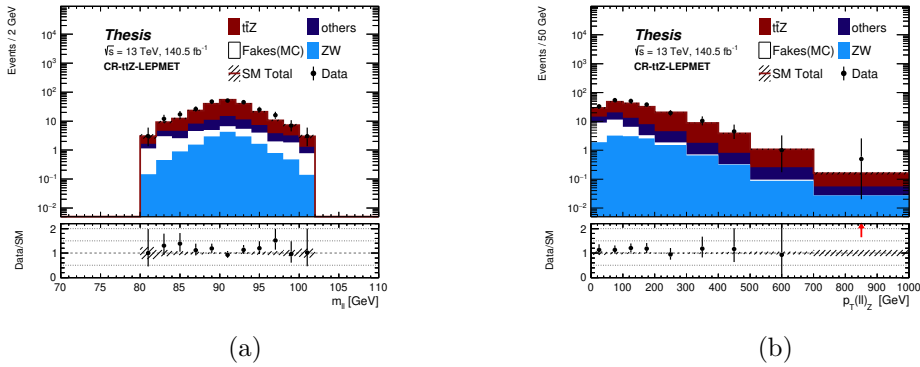
Since the E_T^{miss} and non Z -associated lepton are being treated as jets, a reduction in the number of jets from the default preselection of 4jets could be applied, lowering to two, but calorimeter effects lead to substantial mismodelling in N_{jets} for $N_{\text{jets}}=2,3$ bins. Thus the default requirement translates to an effective requirement of “ $N_{\text{jets}} \geq 6$ ”. To match the signal region E_T^{miss} selection, a selection of $p_T(Z) > 250$ GeV was trialed in addition.

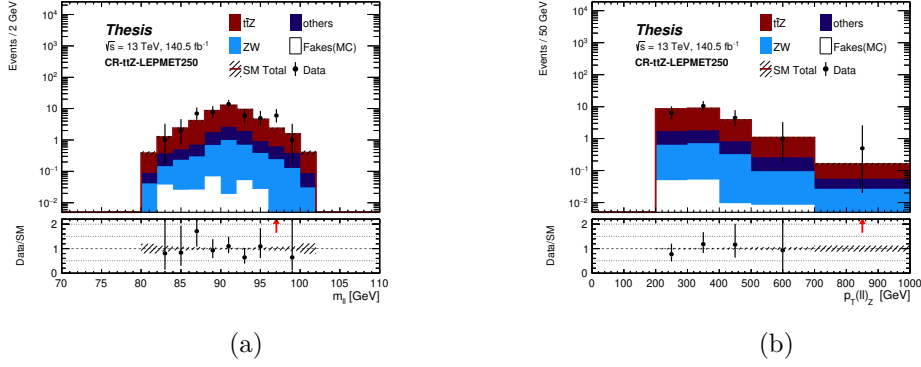
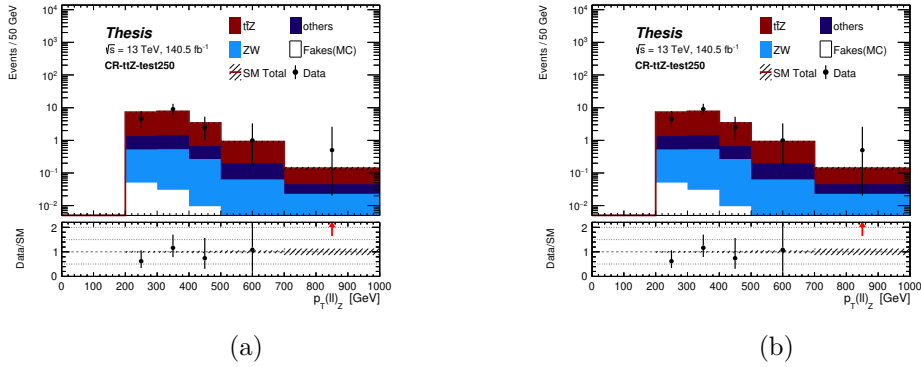
Table 8.3 outlines the different regions considered as candidates. These candidates were chosen to evaluate the yields and modelling of the $t\bar{t} + Z$ background in regions as close as possible to the signal region. Note that for triggering, some overlap handling is required such that the correct calibrations are applied in the case that an event fires the muon trigger and the electron trigger simultaneously.

Selection	CR-ttZ-orig	CR-ttZ-orig-pT4j	CR-ttZ-LEPMET	CR-ttZ-LEPMET250	CR-ttZ-test250
Trigger	Single e/μ				
N_ℓ (Signal)	$=3$				
p_T^1	> 27 GeV				
p_T^3	> 20 GeV				
Z-like OSSF pair	Yes				
$m(\ell, \ell)$	$m_{ll} \in [81.2, 101.2]$				
N_b	≥ 2				
N_{jets}	≥ 4		≥ 2		≥ 4
p_T^{j2}	-	> 80 GeV	> 80 GeV (incl. E_T^{miss} , non Z lep)		
p_T^{j4}	-	> 40 GeV	> 40 GeV (incl. E_T^{miss} , non Z lep)		
$p_T^{\ell\ell}$	-	-	-	> 250 GeV	

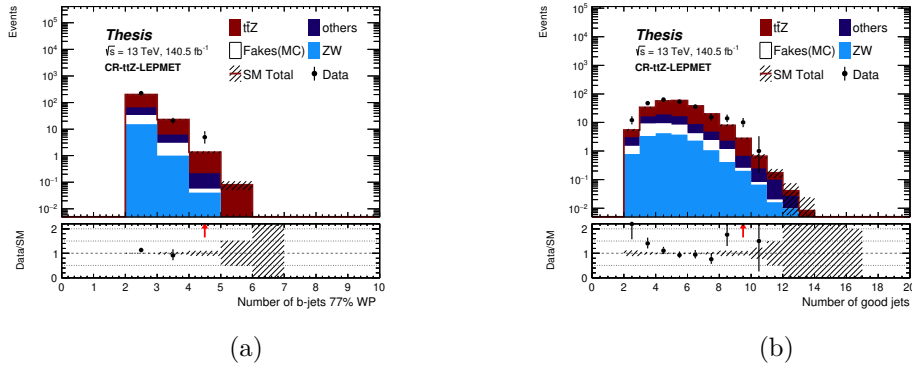
Table 8.3: Candidate control region definitions for the $t\bar{t} + Z$ background.

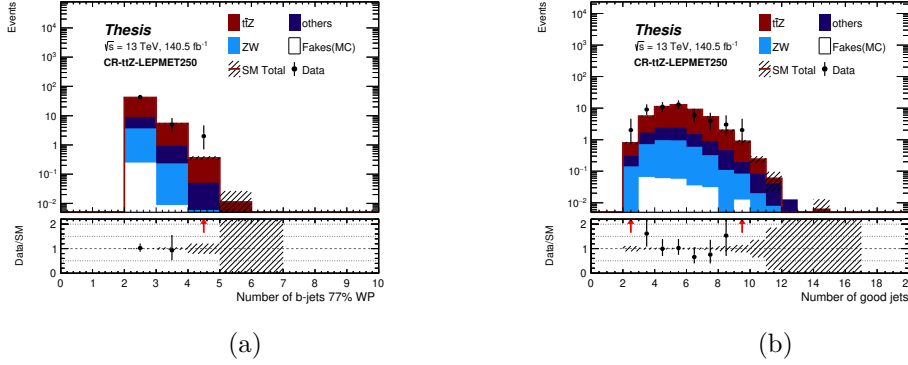
The following study was undertaken with the following Monte Carlo samples, corresponding to $t\bar{t} + Z$, WZ , and fake lepton backgrounds, which are subdivided into $t\bar{t}, Z + \text{jets}, W + \text{jets}$, single top (not associated to tZ), and others, which are defined as: $t\bar{t} + W, t\bar{t} + H, tZ, tWZ$ & ZZ . The study was performed using MC samples normalised to 140.5fb^{-1} and compared against data originating from periods 2015-2018 inclusive. This study considers the statistical errors of data and each sample only. Distributions for each region defined in 8.3 are presented in figures 8.7-8.11.


 Figure 8.7: Distributions for 8.7a $m(\ell, \ell)$ and 8.7b $p_T^{\ell\ell}$ in the region “CR-ttZ-orig”.

 Figure 8.8: Distributions for 8.8a $m(\ell, \ell)$ and 8.8b $p_T^{\ell\ell}$ in the region “CR-ttZ-orig-pT4j”.

 Figure 8.9: Distributions for 8.9a $m(\ell, \ell)$ and 8.9b $p_T^{\ell\ell}$ in the region “CR-ttZ-LEPMET”.


 Figure 8.10: Distributions for 8.10a $m(\ell, \ell)$ and 8.10b $p_T^{\ell\ell}$ in the region “CR-ttZ-LEPMET250”.

 Figure 8.11: Distributions for 8.11a $m(\ell, \ell)$ and 8.11b $p_T^{\ell\ell}$ in the region “CR-ttZ-test250”.

The number of jets distributions have also been checked in particular for “CR-ttZ-LEPMET” and “CR-ttZ-LEPMET250” in figures 8.12 and 8.13.


 Figure 8.12: Distributions for 8.12a N_b and 8.12b N_{jets} in the region “CR-ttZ-LEPMET”.


 Figure 8.13: Distributions for 8.13a N_b and 8.13b N_{jets} in the region “CR- $t\bar{t}$ Z-LEPMET250”.

CR- $t\bar{t}$ Z-orig		CR- $t\bar{t}$ Z-orig-pT4j		CR- $t\bar{t}$ Z-LEPMET	
$t\bar{t} + Z$	192.39 ± 1.01	$t\bar{t} + Z$	75.73 ± 0.69	$t\bar{t} + Z$	151.10 ± 0.90
WZ	19.80 ± 0.37	WZ	5.66 ± 0.18	WZ	16.29 ± 0.31
Fakes (MC)	27.27 ± 2.46	Fakes (MC)	5.19 ± 1.10	Fakes (MC)	20.56 ± 2.49
Others	41.43 ± 0.60	Others	14.18 ± 0.37	Others	35.40 ± 0.27
Total SM	280.89 ± 2.75	Total SM	100.76 ± 1.36	Total SM	223.37 ± 2.49
Data	306.0 ± 17.49	Data	108.00 ± 10.39	Data	252.00 ± 15.87

CR- $t\bar{t}$ Z-LEPMET250		CR- $t\bar{t}$ Z-test250	
$t\bar{t} + Z$	37.54 ± 0.46	$t\bar{t} + Z$	33.17 ± 0.44
WZ	3.67 ± 0.13	WZ	2.87 ± 0.11
Fakes (MC)	0.26 ± 0.08	Fakes (MC)	0.19 ± 0.07
Others	6.12 ± 0.27	Others	14.18 ± 0.37
Total SM	47.59 ± 0.56	Total SM	41.23 ± 0.53
Data	50.00 ± 7.07	Data	39.0 ± 6.24

Table 8.4: Statistical error only yields for the regions defined in table 8.3.

In particular, figure 8.13 motivates the choice of $N_{jets} \geq 4$, which produce mismodelling which arises from calorimeter effects in events with a low number of jets. A selection with at least 4 jets was chosen to avoid this mismodelling effects, and the initial candidate control region selected was “CR- $t\bar{t}$ Z-test-250”. The purity of this region was determined as 80.4%, and the normalisation factor was 0.95, which is a substantial improvement in purity over a di-leptonic $CR_{t\bar{t}} + Z$.

8.2.3 Optimising the tri-lepton control region: final definition of $CR_{t\bar{t}} + Z$

In this subsection, an optimisation of the $CR_{t\bar{t}} + Z$ was implemented to increase the overall statistics of $t\bar{t} + Z$ to reduce the statistical uncertainty on the normalisation $\mu_{t\bar{t}+Z}$. Starting with the candidate region $CR_{t\bar{t}} + Z$ -test-250 as defined in 8.3, a modest relaxing of the $p_T^{\ell\ell}$ selection to > 200 GeV was applied. The modelling in the low $p_T^{\ell\ell}$ bins was checked to ensure that the extrapolation was well motivated in the control region. This region was chosen as the control region for the $t\bar{t} + Z$ background in [2], and hence is defined in full in table 8.5.

Variable/CR	CRttZ-AB (CRttZ-pT200)
Trigger	Single e/μ
N_ℓ	$= 3$
Trigger	Single lepton
p_T^ℓ	$> 27, 20, 20$ GeV
Contains Z-like OSSF pair?	✓
$m(\ell, \ell)$	$\in [81.2, 101.2]$
N_b	≥ 2
N_j	≥ 4
p_T^{j2}	> 80 GeV (incl. E_T^{miss} , non-Z lepton)
p_T^{j4}	> 40 GeV (incl. E_T^{miss} , non-Z lepton)
$p_T^{\ell\ell}$	> 200 GeV

Table 8.5: Final definition of CRttZ-AB used in [2]. Note [2] defines the $m(\ell, \ell)$ interval is specified as $[81, 101]$ GeV instead (but is the same region).

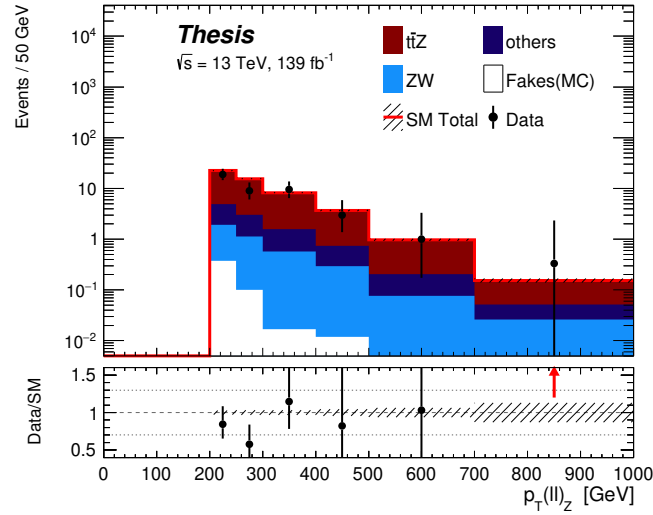


Figure 8.14: $p_T^{\ell\ell}$ distribution of $CRt\bar{t} + Z$ -AB with the relaxed $p_T^{\ell\ell}$ selection

CRttZ-AB	
$t\bar{t} + Z$	53.14 ± 0.58
ZW	4.71 ± 0.16
Fakes(MC)	0.55 ± 0.17
others	8.39 ± 0.32
SM	66.79 ± 0.70
Data	59.00 ± 7.68

Table 8.6: Pre-fit yields for the 3L region CRttZ-AB (Table 8.5).

From 8.14 it can be noted that the extrapolation of this variable is stable in the lower $p_T^{\ell\ell}$ bin, and the number of $t\bar{t} + Z$ events increases by approximately 60% as a result of the inclusion of the additional $p_T^{\ell\ell}$ bin. Purity is observed to be approximately 80%. Henceforth, this region is defined as CRttZ-AB.

8.2.4 Theoretical uncertainties in cross sections for other backgrounds in $CRt\bar{t} + Z$

Since the control region topology is a tri-lepton channel, there are significant background sources that in the signal region are sub-dominant or vanishing backgrounds, namely the tZ , tWZ , WZ and fake lepton backgrounds. For the signal region of the ATLAS $t\bar{t} + Z$ observation [110], these backgrounds are naturally constrained by control regions or (in the case of backgrounds including fake-leptons, a data-driven estimate). If this approach was followed also, a control region for a control region would be required (so a control region for WZ in tri-lepton $t\bar{t} + Z$, that has essentially no impact on the signal region directly).

In the case of the WZ background, this would have to be estimated in a region with only a small number of b-jets (and would imply extrapolation to $CRt\bar{t} + Z$, incorporating uncertainties therein to heavy flavour jet extrapolation). We use an estimate of $\pm 30\%$ for the magnitude of these uncertainties. If a sub-control region did not exist for WZ , we would have to use the estimate from simulation only, which would incur an effective systematic uncertainty on the cross section of $\pm 100\%$ on the estimation of the normalisation of this background. This study evaluates the impact of estimating these backgrounds from either simulation or with a dedicated sub-control region on the fit parameter which will be influenced by these decisions: the normalisation of the $t\bar{t} + Z$ background ($\mu_{t\bar{t}+Z}$).

In order to perform this check, a maximal likelihood background-only fit was undertaken with 1 parameter of interest (namely the normalisation $\mu_{t\bar{t}+Z}$ and one control region (with a dummy blinded signal region used for validation). In order to perform the fit, only the MC statistics were included, and to evaluate the impact of estimation of the given sample (tZ , tWZ , Fakes), either an uncertainty of $\pm 30\%$ was applied on the cross section of the given sample (this would reflect a typical uncertainty due to estimation of a background via a control region), or a $\pm 100\%$ was applied on the cross section of the given sample (uncertainty due to estimation via MC).

A scan over $p_T^{\ell\ell}$ bins in intervals of 50 GeV, was chosen, including only the statistical uncertainty in the fit according to the following algorithm:

1. Start with the baseline selection for the region (in this case, region “CR-ttZ-orig” from table 8.3).
2. Perform the background only fit with the dummy 30% cross section systematic and MC statistical uncertainties.
3. Extract the $\mu_{t\bar{t}+Z}$ and its uncertainty post-fit.
4. Add 50 GeV to the $p_T^{\ell\ell}$ selection and iterate again until end of interval (350 GeV).
5. Repeat procedure with the 100% cross section systematic on the given sample.

The results (central value and uncertainty of $\mu_{t\bar{t}+Z}$ as a function of $p_T^{\ell\ell}$ selection) are documented in figures 8.15-8.22. Each histogram bin relates to a fit on a region defined by the base region definition with the additional selection on the $p_T^{\ell\ell}$ applied.

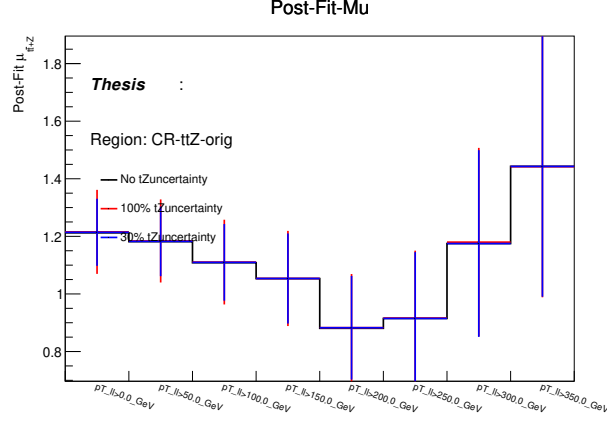


Figure 8.15: Uncertainty scan in $\mu_{t\bar{t}+Z}$ for differing bins in $p_T^{\ell\ell}$ starting from the base region “CR-ttZ-orig”, evaluating the impact of normalisation of the tZ background. The post-fit $\mu_{t\bar{t}+Z}$ and its uncertainty are presented.

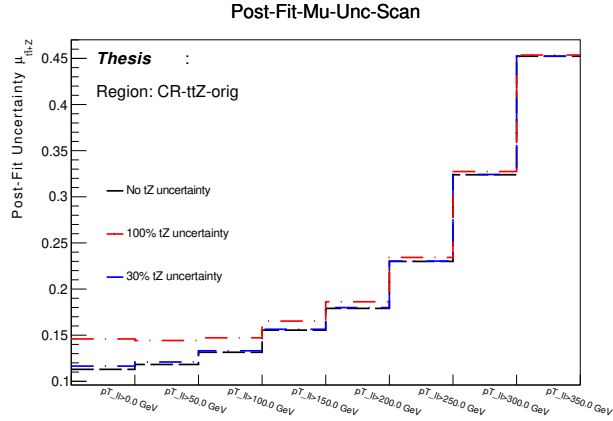


Figure 8.16: Uncertainty scan in $\mu_{t\bar{t}+Z}$ for differing bins in $p_T^{\ell\ell}$ starting from the base region “CR-ttZ-orig”, evaluating the impact of normalisation of the tZ background. The uncertainty on the post-fit $\mu_{t\bar{t}+Z}$ is presented independent of the central value.

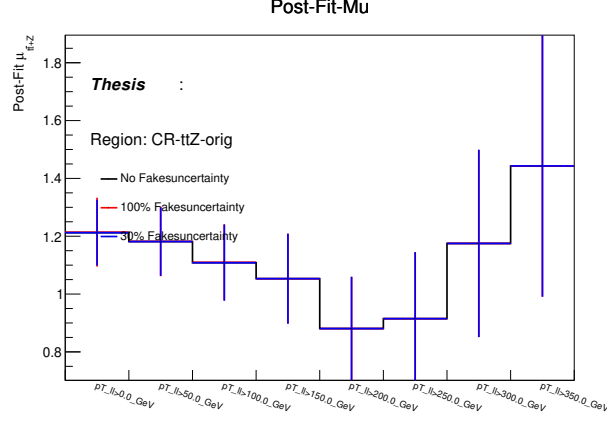


Figure 8.17: Uncertainty scan in $\mu_{t\bar{t}+Z}$ for differing bins in $p_T^{\ell\ell}$ starting from the base region “CR-ttZ-orig”, evaluating the impact of normalisation of the Fakes background. The post-fit $\mu_{t\bar{t}+Z}$ and its uncertainty are presented.

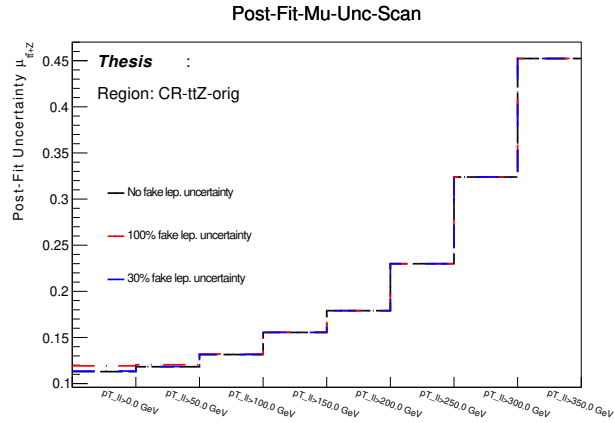


Figure 8.18: Uncertainty scan in $\mu_{t\bar{t}+Z}$ for differing bins in $p_T^{\ell\ell}$ starting from the base region “CR-ttZ-orig”, evaluating the impact of normalisation of the Fakes background. The uncertainty on the post-fit $\mu_{t\bar{t}+Z}$ is presented independent of the central value.

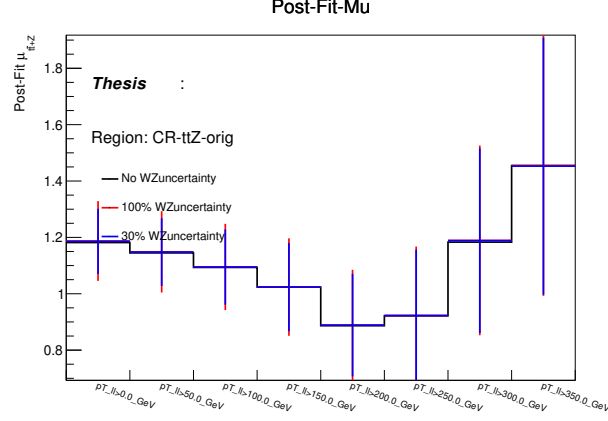


Figure 8.19: Uncertainty scan in $\mu_{t\bar{t}+Z}$ for differing bins in $p_T^{\ell\ell}$ starting from the base region “CR-ttZ-orig”, evaluating the impact of normalisation of the WZ background. The post-fit $\mu_{t\bar{t}+Z}$ and its uncertainty are presented.

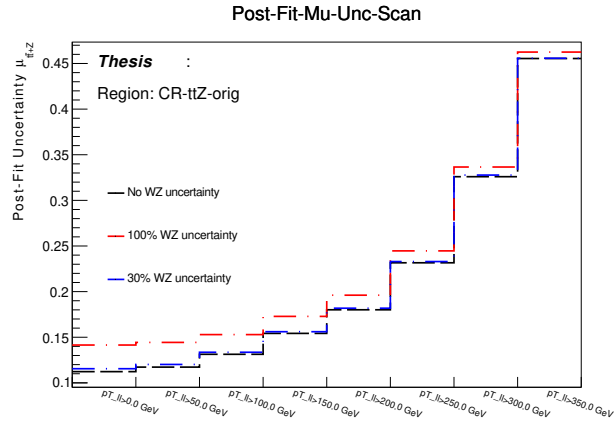


Figure 8.20: Uncertainty scan in $\mu_{t\bar{t}+Z}$ for differing bins in $p_T^{\ell\ell}$ starting from the base region “CR-ttZ-orig”, evaluating the impact of normalisation of the WZ background. The uncertainty on the post-fit $\mu_{t\bar{t}+Z}$ is presented independent of the central value.

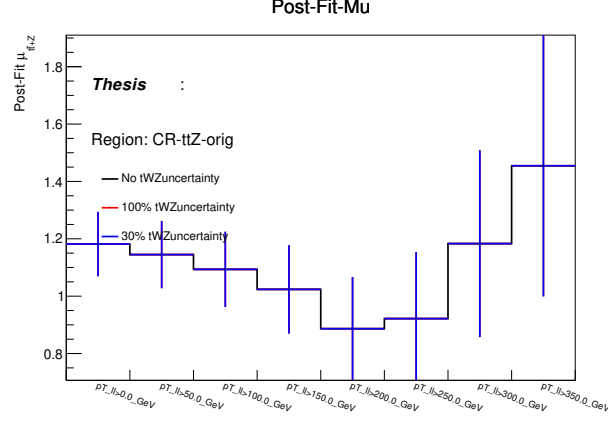


Figure 8.21: Uncertainty scan in $\mu_{t\bar{t}+Z}$ for differing bins in $p_T^{\ell\ell}$ starting from the base region “CR-ttZ-orig”, evaluating the impact of normalisation of the tWZ background.. The post-fit $\mu_{t\bar{t}+Z}$ and its uncertainty are presented.

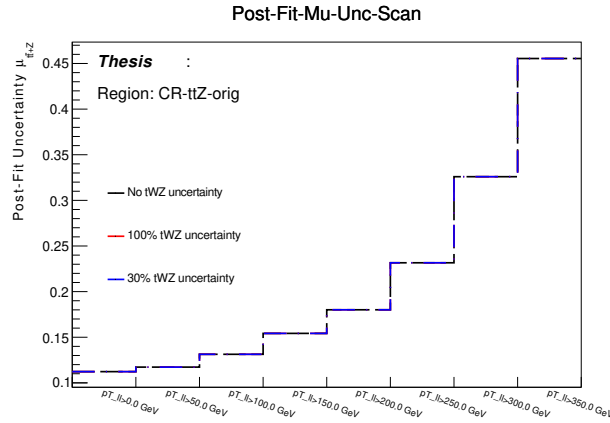


Figure 8.22: Uncertainty scan in $\mu_{t\bar{t}+Z}$ for differing bins in $p_T^{\ell\ell}$ starting from the base region “CR-ttZ-orig”, evaluating the impact of normalisation of the tWZ background. The uncertainty on the post-fit $\mu_{t\bar{t}+Z}$ is presented independent of the central value.

From figures 8.16-8.21, it can be observed that the impact on the uncertainty of the fitted $\mu_{t\bar{t}+Z}$ in each bin in $p_T^{\ell\ell}$ for each sample is small compared to the impact of the dominant uncertainty, the poissonian statistical uncertainty in data events, either in the case of the 30% or 100% cross section uncertainties. Thus estimation of the uncertainties from Monte-Carlo without further normalisation would not have any substantive impact on the estimate of the $t\bar{t} + Z$ normalisation in either the control region or signal region, compared to the statistical uncertainty in data events which is already accounted for in the uncertainty of $\mu_{t\bar{t}+Z}$. To conclude, it can be noted that the fake lepton, WZ and tWZ can be estimated from MC directly, rather than through dedicated control regions or alternate approaches.

8.3 Control region specific: CRZ-AB ($Z + \text{jets}$) design

The $Z + \text{jets}$ process is an important background in signal regions A and B, in particular with the decay $Z \rightarrow \nu\nu$. The jets/b-jets in the event can arise from gluon splitting or from diagrams which involve the radiation of a Z .

A good candidate for these processes is a direct replacement of neutrinos to visible leptons in the final state, thus using $Z \rightarrow \ell\ell + \text{jets}$ in a di-leptonic final state. To correctly model the E_T^{miss} in the signal region, in a di-lepton channel, we add the leptons to the E_T^{miss} or “treat leptons as invisible” (discussed in chapter 6), which is referred to henceforth as $E_T^{\text{miss}'}$ or “fake E_T^{miss} ”. The current region definition is targeting an opposite sign same flavour di-leptonic final state to model the Z . However, to maximise the constraints on systematics in the final fit, instead of a single bin control region, a multi-bin fit was chosen (multiple distinct regions with a single normalisation parameter). In order to maximise the efficacy of a multi-bin fit, a variable that is not well modelled in the control region is chosen as the variable to be binned, namely the $R=1.2$ reclustered jet mass, so as to constrain some of the mis-modelling on this variable in the signal region.

8.3.1 CRZ definition and construction

In the lowest order $Z + \text{jets}$ diagrams, there are two distinct event topologies, primarily relating to the source of b-jets in a $Z \rightarrow \nu\nu + \text{jets}$ event, where the b-jets can either arise from an ISR/FSR $g \rightarrow b\bar{b}$, or as a production from the scattering event. b-jets arising from $g \rightarrow b\bar{b}$ decays are expected to be correlated with each other due to the momentum carried by the decaying gluon, thus gluon splitting events can be characterised by an event with two b-jets in close proximity. These two topologies are labelled “gluon-splitting type” and “scattering type” diagrams respectively, and this is best shown by an example (figure 8.23), although these categories may overlap. Consider a typical di-leptonic $Z + \text{jets}$ event that would arise in the control region, which has 4 jets, of which precisely 2 are b-tagged:

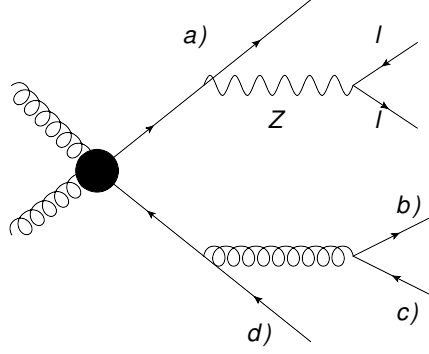


Figure 8.23: Combinatoric diagram of a $Z + \text{jets}$ event with 4 jets, of which 2 are b-tagged. “scattering type” denotes the case where quarks “a)” and “d)” are b-tagged, while “gluon splitting type” denotes the case where quarks “b)” and “c)” are b-tagged .

By this observation, it can be noted that the “gluon splitting type” diagrams should have a contribution only in events with small $\Delta R(b_1, b_2)$ due to the momentum of the parent gluon, while the “scattering type” diagrams should have an approximately uniform contribution across this same variable, thus a selection in this variable can be applied to pick the given topology required.

In the signal regions, there are several different regimes considered due to differing cuts on $\Delta R(b_1, b_2)$:

1. SRATT: $\Delta R(b_1, b_2) > 1.0$.
2. SRATW & SRAT0: Inclusive on $\Delta R(b_1, b_2)$.
3. SRB(TT/TW/T0) : $\Delta R(b_1, b_2) > 1.4$.

To allow for additional statistics in each $\Delta R(b_1, b_2)$ bin, a loosened selection on the sub-leading $R=1.2$ reclustered jet mass was applied, merging into two bins of $m_{\text{jet}, R=1.2}^2 > 60 \text{ GeV}$ or $m_{\text{jet}, R=1.2}^2 < 60 \text{ GeV}$. The full set of selections is denoted below:

Selection Scheme	Selection	CRZA-TT-TW	CRZA-T0	CRZB-TT-TW	CRZB-T0
Preselection	Trigger N_l OSSF? $p_T(l_1)$ $p_T(l_2)$ N_j N_b $p_T(j_2)$ $p_T(j_4)$ (Real) E_T^{miss} (Fake) E_T^{miss} $M(l\bar{l})$ $\min[m_T(b, E_T^{\text{miss}})]$ Object Based (Fake) E_T^{miss} Significance $m(J_1; R = 1.2)$	1L Trigger 2 signal, 2 baseline Yes $> 27 \text{ GeV}$ $> 20 \text{ GeV}$ ≥ 4 ≥ 2 $> 80 \text{ GeV}$ $> 40 \text{ GeV}$ $< 70 \text{ GeV}$ $> 200 \text{ GeV}$ $\in [81, 101] \text{ GeV}$ $> 100 \text{ GeV}$ > 10 $> 120 \text{ GeV}$			
Loose scheme	$m(J_2; R = 1.2)$	$> 60 \text{ GeV}$	$< 60 \text{ GeV}$	$> 60 \text{ GeV}$	$< 60 \text{ GeV}$
$\Delta R(b, b)$ slicing	$\Delta R(b_1, b_2)$ $m(J_2; R = 1.2)$	< 1.4 -		> 1.4 -	

Table 8.7: CRZ with loosened sub-leading $R=1.2$ reclustered jet mass selections and $\Delta R(b_1, b_2)$ slicing schemes.

Fit comparisons

In order to evaluate the efficacy of a given control region configuration, a maximal likelihood fit must be applied on the given control regions and applied to the signal region, with the parameter to be compared being the normalisation of the $Z + \text{jets}$ background $\mu_{Z + \text{jets}}$ and its uncertainties between the “CRZ-loose” scheme and the “CRZ-loose- $\Delta R(b, b)$ slicing scheme”. In this study, the control region CRZ bins are used from table 8.7 in a multi-bin fit (a single normalisation parameter for all Zjets CRs), and a single $t\bar{t}Z$ control region. Two fit parameters were considered, namely the $\mu_{t\bar{t}+Z}$ and $\mu_{Z\text{jets}}$, and using a set of systematics reflecting the dominant contributions to the fit:

1. Jet energy resolution and scale uncertainties
2. E_T^{miss} soft term uncertainties
3. Theory uncertainties for $Z + \text{jets}$: Merging, factorisation and renormalisation scale uncertainties
4. Theory uncertainties for $t\bar{t} + Z$: Factorisation/renormalisation scale
5. b-tagging extrapolation uncertainties

Signal Region	Slicing scheme	Fitted signal region yield (Total bkg)	Uncertainty due to μ_Z	Stat. uncert	Total syst.
SRAT0	Preselection only	14.92	± 0.73	± 3.86	± 2.29
	Loose-slicing	14.85	± 1.06	± 3.85	± 2.29
	$\Delta R(b_1, b_2)$ slicing	14.81	± 0.71	± 3.85	± 2.42
SRBTT	Preselection only	45.97	± 1.24	± 6.78	± 3.71
	Loose-slicing	45.82	± 1.81	± 6.77	± 4.72
	$\Delta R(b_1, b_2)$ slicing	44.86	± 1.20	± 6.70	± 3.53

Table 8.8: Yields in the signal regions, and impact on μ_Z of each of the slicing schemes listed in table 8.7 to the total uncertainty.

Comparing the uncertainties in SRAT0, it can be observed that the Z+jets background is generally well normalised in the control region CRZ post-fit. There are only modest changes between the $\Delta R(b, b)$ slicing regime and the “loose” strategy.

In the signal region SRAT0, the $\Delta R(b, b)$ slicing scheme results in a larger total uncertainty (increasing by $\approx 1\%$), while the uncertainty due to μ_Z decreases by $\approx 0.1\%$.

In all of the cases for SRAT0, the μ_Z is not the dominant systematic, mostly substantially smaller than the MC statistical uncertainty and somewhat smaller than some JET_JER variations, so the effect of the different schemes is negligible on μ_Z , and their impact on other systematics is also small. SRAT0 is also generally statistically limited with a large MC statistical uncertainty because of a low expected number of events in this region. SRBTT also shows a $\approx 2\%$ increase in total uncertainty in the top categories scheme compared to the loose scheme, and $\approx 1\%$ increase in the uncertainty of μ_Z , while the $\Delta R(b, b)$ scheme reduces the uncertainty overall by $\approx 0.3\%$, while having only a $\approx 0.02\%$ impact on μ_Z . The uncertainty on μ_Z is not the leading systematic in SRBTT, and the predominant uncertainties are that of the $t\bar{t} + Z$ normalisation, ttbar ISR and JES uncertainties.

Conclusion

Overall, it can be observed that the differences between the slicing schemes are almost negligible ($O(\approx 1\%)$), particularly in the signal region uncertainties where the μ_Z is not necessarily the leading uncertainty contribution. This is somewhat mitigated by the absence of MC16e $Z\nu\nu$ samples, although MC16a/d are re-scaled to cover this absence. SRA is dominated by MC statistical uncertainties, while SRB is dominated by uncertainties not associated with Z+jets. As a result, the “Loose” binning strategy is the preferred scheme despite the differences in event topology in $Z + \text{jets}$ events, thus no specific $\Delta R(b_1, b_2)$ binning is required.

8.3.2 Final definition of CRZ-AB

The final definition of CRZ was optimised based on the signal region selections, final region statistics and the study in section 8.3.1. To minimise extrapolation of the Z+jets background between signal and control region, the control regions were binned in $m_{\text{jet}, R=1.2}^2$ as per the signal region, although due to limitations of control region statistics, the TT and TW bins were merged. The primed variables refer to those defined using $E_{\text{T}}^{\text{miss}'}$ instead of $E_{\text{T}}^{\text{miss}}$ in their respective calculations, where $E_{\text{T}}^{\text{miss}'}$ is $E_{\text{T}}^{\text{miss}}$ calculated where the leptons are treated as invisible.

The selections are as follows [2]:

Variable/CR	CRZAB-T0	CRZAB-TTTW
Trigger	Single lepton	
N_ℓ	== 2 OSSF	
$m(\ell, \ell)$	[81, 101] GeV	
E_T^{miss}	< 50 GeV	
p_T^ℓ	> 27, > 20 GeV	
$E_T^{\text{miss}'}$	> 200 GeV	
N_j	≥ 4	
p_T^{j2}	> 80 GeV	
p_T^{j4}	> 40 GeV	
N_b	≥ 2	
$m_{\text{jet}, R=1.2}^1$	> 80 GeV	
$m_T^{b, \text{min}'}$	> 150 GeV	
Object based $E_T^{\text{miss}'}$ sig.	> 10 GeV	
$m_{\text{jet}, R=1.2}^2$	> 60 GeV	< 60 GeV

Table 8.9: Definition of control region CRZAB-TTTW and CRZAB-T0 as defined in [2].

Concluding remarks

To conclude, this chapter has discussed the implementation of a di-leptonic control region for the $Z + \text{jets}$ background (as a re-design from the 2015/16 implementation), and studied the effects of using binned control regions and multi-bin fits over different variables and their effects on the overall uncertainties in background estimation, deciding on a strategy which is equivalent to the top categories in the 2015/16 signal region.

In addition, a wholly new control region for the $t\bar{t} + Z$ background has been designed, re-interpreting the signal regions used for the observation of $t\bar{t} + Z$ as a control region for a supersymmetry search. This replaces the control region that was previously defined with a different kinematic process ($t\bar{t} + \gamma$). This region is comparable to those used by the ATLAS one-lepton stop search [22, 99], however differs somewhat in the object definitions used and selections used. Sub-control regions have also been discussed for the tri-lepton $\text{CR}t\bar{t} + Z$, and were observed not to be necessary since the uncertainty on $\mu_{t\bar{t}+Z}$ was principally driven by statistical uncertainties in data in the control region, rather than any uncertainties on ancilliary backgrounds, at least in the regions where $p_T^{\ell\ell} > 150$ GeV.

2015-18 analysis: Summary

This chapter outlines a detailed summary of the analysis published in [2], the search for the scalar partner of the top quark in the all-hadronic $t\bar{t}$ plus E_T^{miss} final state at $\sqrt{s} = 13$ TeV with the ATLAS detector. The principal signal models are the same as in the 2015-16 analysis, namely RPC $\tilde{t}_1 \rightarrow t + \tilde{\chi}_1^0$ alongside three and four body decay modes $\tilde{t}_1 \rightarrow bW/bjj' + \tilde{\chi}_1^0$, in a final state with large E_T^{miss} , zero electrons / muons, and ≥ 4 jets (of which at least two are b-tagged). This will discuss the signal, control and validation regions which are used, and make use of the author's contributed control region designs in the previous chapter, along with a discussion of the systematics.

Since ATLAS analyses often build on the ideas discussed in previous papers, some of the regions will appear similar to those in 2015/16 in their general aims, however the variables used will differ, based on optimisation by the author and others in the collaborative effort. This particularly applies to the signal regions SRA, SRB and SRC, while the four body signal regions (SRD) discussed later are newly conceived for this publication. Thus, some of the discussion of the major motivations of the signal regions for SRA, SRB and SRC will primarily be discussed in chapter 6, and the differences will be highlighted.

The author's principal contributions have been summarised in the previous chapters to the analysis design. However, in addition to the material provided in [2], a more detailed explanation of the estimation of several theory uncertainties (primarily related to the estimation of the $t\bar{t} + Z$ background) undertaken by the author has been discussed, alongside the general philosophy of uncertainty estimation in ATLAS searches.

9.1 Object definitions

Object definitions are usually agreed upon by a collective group made of many analysis teams with a similar set of final states. This agreement is always a balance of considerations such that signal regions (or worse, control regions) do not overlap with a signal region for another team or to gain additional statistics. A particular point of contention is the identification of electrons or muons and their p_T : a decrease in the electron/muon minimum p_T may increase the number of events entering the signal region for a study involving leptons, however the converse case will occur in the all-hadronic region. In addition, a decrease in the minimum

p_T threshold may impact the reconstruction efficiency and fake rates (decreasing the number of events in the signal region further).

There are also two major distinctions in the definitions of electrons/muons: baseline and control (also referred to as “signal” leptons in ATLAS nomenclature). Baseline e/μ are where leptons are identified with a loose identification criteria selection on p_T etc, and form the baseline definition of the signal region (and orthogonality with other analyses). Control/signal e/μ are those which have undergone isolation and overlap removal, namely removal of objects which are overlapping with other objects such as jets, and have a tighter requirement than the “baseline” selection, since in all cases they are a strict subset of the baseline. This strict subset will have tighter selections on the lepton p_T etc, dependent on the choice of trigger used. Regions using a E_T^{miss} trigger will have a requirement of $p_T^\ell > 20$ GeV, while those using single e/μ triggers will require the same, but with $p_T^{\ell_1} > 28$ GeV to account for the maximal trigger efficiency regions for the trigger (this electron/muon is required to be the one to fire the trigger). In all cases used in this thesis, the combination that both the number of baseline and signal e/μ required must be the same.

Since identification of b-hadron decay in jets is important, another principal component is using the MV2c10 algorithm introduced in 4.6.2, configured to be a selection on a fixed value to give an identification efficiency of approximately 77%.

Given the set of jets that we will have in our final state, re-clustering of jets was used, namely the use of well-calibrated jets as an input to the k_t -algorithm (an important note, the metric used is $O(k_t^2)$ rather than $O(k_t^{-2})$ for the *anti* - k_T algorithm). These are used with a radius of $R=1.2$ and $R=0.8$, to cover the reconstruction of the boosted top and W with a p_T over 200 GeV.

For the full details of the object definitions, table 9.1, which forms a summary of the full definitions presented in [2].

Object	Final configuration
Primary vertex	Vertex with ≥ 2 tracks w/ $p_T > 500$ MeV
Calorimeter Jets	anti- k_T ($R=0.4$) $p_T > 20$ GeV, $ \eta < 2.8$
JVT applied to jets with:	Jet $p_T < 120$ GeV, $ \eta < 2.4$
Calorimeter b-jets	Algorithm=MV2c10 (77% eff. WP), $ \eta < 2.5$, $p_T > 20$ GeV
Track b-jets	Variable radius track jets ($R_0 = 0.4$) Tracks: $p_T > 500$ MeV, $ \eta < 2.5$ Tracks: ≥ 7 SCT/Pixel hits Tracks: ≤ 1 shared hit, ≤ 2 missed hits Trackjet: $p_T > 5$ GeV Trackjet: Overlap removal of leading non-b calo jet Trackjet: $ Z_0 \sin \theta < 3mm$ Alg.=MV2c10 (70% eff. WP)
Leptons/Photons	
Electron (Baseline)	$ \eta < 2.47$, $p_T > 4.5$ GeV ID=Loose WP $ z_0 \sin \theta < 0.5mm$ Overlap removal
Muons (Baseline)	$ \eta < 2.7$, $p_T > 4$ GeV $ d_0 > 0.2mm$, $ z_0 > 1mm$ Overlap removal. Cosmic muon veto
Tau veto	Non b-jet, $N_{trk}(trk - p_T > 500 \text{ MeV}) \leq 4$, $\Delta\phi(jet, E_T^{\text{miss}}) < \frac{\pi}{5}$
Tight electron (Control)	$p_T > 20$ GeV, ID="Tight" $ d_0 < 5\sigma d_0$ Isolation=Fixed Cut(FCLoose)
Medium Muon (Control)	$p_T > 20$ GeV, ID="medium" Isolation="Fixed Cut (FCLoose)" $ d_0 < 3\sigma d_0$
E_T^{miss}	Tracks included: $p_T > 500$ MeV

Table 9.1: A summary table of object definitions found in [2].

9.2 Signal regions

In this section, the signal regions for the analysis published in [2] are outlined. These will be summarised briefly.

9.2.1 Signal Regions SRA and SRB

Signal regions SRA and SRB are optimised to target the two-body signal models where $\tilde{t} \rightarrow t + \tilde{\chi}^0$ in the case where $\Delta(m_{\tilde{t}}, m_{\tilde{\chi}^0}) > m_{\text{top}}$. SRA targets the class of models where $\Delta(m_{\tilde{t}}, m_{\tilde{\chi}^0}) \gg m_{\text{top}}$, with the predominant optimisation being designed to maximise exclusion of the $(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (1300, 1)$ GeV signal model. Since in such signal models due to the large mass splitting, the p_T of the top system and/or the E_T^{miss} is expected to be large, this region is referred to as the “boosted” region. SRB targets intermediate signal models, namely those with $\Delta(m_{\tilde{t}}, m_{\tilde{\chi}^0}) > m_{\text{top}}$, in particular optimised for the exclusion of the $(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (700, 400)$ GeV.

Variable/SR	SRA-TT	SRA-TW	SRA-T0	SRB-TT	SRB-TW	SRB-T0
$E_{\text{T}}^{\text{miss}}$	$> 250 \text{ GeV}$					
Trigger	$E_{\text{T}}^{\text{miss}}$					
N_j	≥ 4					
Jet- p_{T}	$> 80 \text{ GeV}, > 80 \text{ GeV}, > 40 \text{ GeV}, > 40 \text{ GeV}$					
$m_{\text{jet}, R=1.2}^1$	$> 120 \text{ GeV}$					
$m_{\text{jet}, R=1.2}^2$	$> 120 \text{ GeV}$	$60 - 120 \text{ GeV}$	$< 60 \text{ GeV}$	$> 120 \text{ GeV}$	$60 - 120 \text{ GeV}$	$< 60 \text{ GeV}$
$ \Delta\phi(\text{jet}^{1-4}, E_{\text{T}}^{\text{miss}}) $	> 0.4					
$N_{\text{b-jet}}$	≥ 2					
$m_{\text{T}}^{b, \text{min}}$	$> 200 \text{ GeV}$					
τ -veto	$\checkmark$					
$m_{\text{jet}, R=0.8}^1$	$> 60 \text{ GeV}$			–		
$j_1^{R=1.2}(b)$	$\checkmark$			–		
$j_2^{R=1.2}(b)$	$\checkmark$	–		–		
$\Delta R(b, b)$	> 1.0	> 1.4				
$m_{\text{T}}^{b, \text{max}}$	–			$> 200 \text{ GeV}$		
Object based $E_{\text{T}}^{\text{miss}}$ sig.	> 25			> 14		
$m_{\text{T}2, \chi^2}$	$> 450 \text{ GeV}$			$< 450 \text{ GeV}$		

Table 9.2: Signal Regions SRA and SRB definitions as per [2].

The primary objects of interest are the jets, reconstructed with the anti- k_t algorithm [2, 111] having a radius of $R = 0.4$, the flavour of those jets (in particular whether they are b-jets), the E_T^{miss} , and since the top decay system may be boosted (and hence each of the daughter components may result in coincident jets), reclustered jets.

In each region, at least two $R=1.2$ reclustered jets are required, with a top quark reconstruction in the event being based on the mass of the given reclustered jet, namely being identified as a top in the case that $M(\text{jet}; R = 1.2) > 120 \text{ GeV}$, a W in the case of $M(\text{jet}; R = 1.2) \in [60, 120]$ and unclassified otherwise. The signal regions SRA and SRB are subdivided into three categories based on these reconstructions, with the requirement that the leading (in p_T) $R=1.2$ reclustered jet must be consistent with the top mass, and the second reclustered jet is classified as consistent with a top (SRX-TT), a W (SRX-TW) or unclassified (SRX-T0). Since the reclustering incorporates the originally calibrated anti- k_t $R=0.4$ jets with their associated information, it is possible to ascertain the number of b-jets inside the reclustered jet. Additional selections are applied on $\Delta\phi(j_{1 \rightarrow 4}, E_T^{\text{miss}})$ and Object based E_T^{miss} sig. to reduce the multi-jet background (this variable is discussed in section 8.1) and $m_T^{b, \text{min}}/m_T^{b, \text{max}}$ to reduce the $t\bar{t}$ background [2].

The m_{T2} as previously defined in the 2015-16 analysis provides the orthogonalisation between SRA and SRB, namely that SRA requires $m_{T2} > 450 \text{ GeV}$ and SRB $m_{T2} < 450 \text{ GeV}$. The result of these optimisations is the signal region table for SRA and SRB, as presented in [2], in table 9.2.

These signal regions are combined for the purposes of extracting limits on cross sections, but an additional orthogonal region (“SRA-TT-Disc”) was added for the purpose of signal discovery. This region targets the lower stop mass signals, however this region has negligible improvement on expected exclusions of those processes. “SRA-TT-Disc” is defined by the selections of “SRA-TT”, with the Object based E_T^{miss} sig. > 25 replaced with Object based E_T^{miss} sig. > 11 .

9.3 Signal region C (SRC)

This region targets signals where $\Delta(m_{\tilde{t}}, m_{\tilde{\chi}^0}) \sim m_t$, which are referred to as “compressed” 2 body signals, similarly to the SRC from the 2015-16 analysis. The small mass-splitting of the stop and LSP leads to a small E_T^{miss} , which makes it effectively indistinguishable from the semi-leptonic $t\bar{t}$ background, and may have insufficient E_T^{miss} to be triggered on. In order to resolve the signal, the system considered is assumed to recoil from an ISR jet, with a $p_T > 400$ GeV. This region only considers anti- k_T $R=0.4$ jets, their b-tagging information and the E_T^{miss} , since the top system is expected to be resolved compared to the boosted case in the previous subsection. As per the 2015-16 analysis listed in chapter 6, the R_{ISR} is used as the key discriminant of interest, defined identically between the two publications.

To cover the full class of signal points along the compressed diagonal, a multi-bin fit is performed on five bins of the R_{ISR} distribution to maximise exclusion for each signal. Selections on $E_T^{\text{miss, track}}$ and $\min[\Delta\phi(j_{1-2}, E_T^{\text{miss}})]$ are used to mitigate multi-jet backgrounds. The full list of selections [2] as applied to SRC are outlined in Table 9.3. However, the main differing points between the 2015-16 and 2019 iterations of SRC are:

- Re-optimisation of the SRC selections, given the additional statistics.
- Re-optimisation of the R_{ISR} bins to maximise expected exclusion.
- Additional definition of a discovery region in addition to the R_{ISR} binned signal regions, as mentioned below, using the object based met significance.

A discovery region is also considered, which is optimised for discovery of compressed signals rather than their expected exclusion in the diagonal case where $m_{\tilde{t}}$ is large. This region is a subset of the full SRC bins, namely R_{ISR} bins SRC3-SRC5, with an additional constraint on Object based E_T^{miss} sig. > 11 , which is not used in the application of exclusion limits on cross sections.

Variable/SR	SRC1	SRC2	SRC3	SRC4	SRC5
N_ℓ	== 0				
E_T^{miss}	> 250 GeV				
Trigger	E_T^{miss}				
N_j	≥ 4				
Jet p_T	> (80, 80, 40, 40) GeV				
$N_{b\text{-jet}}$	≥ 2				
$E_T^{\text{miss,track}}$	> 30 GeV				
$ \Delta\phi(E_T^{\text{miss}}, E_T^{\text{miss,track}}) $	< $\pi/3$				
$ \Delta\phi(\text{jet}^{1,2}, E_T^{\text{miss}}) $	> 0.4				
N_{jet}^S	≥ 4				
p_T^{ISR}	> 400 GeV				
$p_{T,b}^{1,S}$	> 50 GeV				
$p_T^{4,S}$	> 50 GeV				
m_S	> 400 GeV				
$\Delta\phi_{\text{ISR}, E_T^{\text{miss}}}$	> 3.0				
R_{ISR}	0.30 – 0.40	0.40 – 0.50	0.50 – 0.60	0.60 – 0.70	≥ 0.70

Table 9.3: Selection criteria for SRC from [2]. The signal regions are separated into five categories based on ranges of R_{ISR} .

9.4 Four body signal regions: SRD

This signal region differs substantially from the signal region SRD presented in [3], and concerns the four body decay modes of the stop ($\tilde{t} \rightarrow b + ff + \tilde{\chi}_1^0$), and is considered a new set of signal regions, with associated control and validation regions.

These signal regions target models where the $\Delta(m_{\tilde{t}}, m_{\tilde{\chi}_1^0}) < m_{\text{top}}$, primarily with mass splittings substantially less than the top mass. Instead of a top+neutralino decay, the stop decay incorporates a four body vertex via an off-shell top and off-shell W, resulting in a $\tilde{\chi}_1^0$, b-jet and two light jets). In terms of signal models, this channel is optimised for searches with mass splittings $\Delta(m_{\tilde{t}}, m_{\tilde{\chi}_1^0}) = (20, 50, 80)$ GeV.

The kinematics of this particular event depend on the mass splitting, but since it is again compressed requires a recoil against ISR to be resolvable from the $t\bar{t}$ background. The key object of interest that differs between the sets of signal models considered is the b-jet, which for the most compressed signals ($\Delta M = 20$ GeV) would produce b-jets with $p_T < 20$ GeV. Since due to calorimeter limitations, b-jets cannot be reconstructed using calorimeter clusters (“EMTopo” jets) for $p_T < 20$ GeV, jets reconstructed from tracking information are used since their reconstruction performance is adequate in the case of $p_T < 20$ GeV, until $p_T > 5$ GeV. Vertexing and b-tagging of these jets can be performed for $p_T > 5$ GeV. For larger mass splitting models ($\Delta M = 50, 80$ GeV), the b-jet p_T would be > 20 GeV so calorimeter reconstructed b-jets are used.

The ISR in this case is denoted as the leading light (i.e. not b-tagged) calorimeter jet, and a large angular separation between the ISR and the E_T^{miss} is expected.

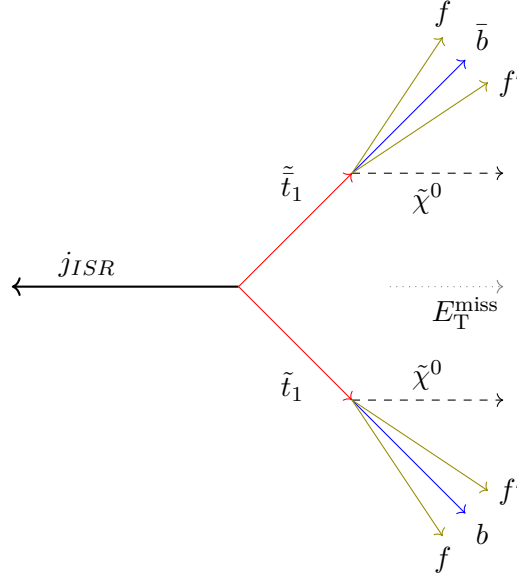


Figure 9.1: Schematic of a four body signal event in SRD, noting the orientation of the ISR and the stop decay system.

Signal region SRD0 targets signal models with the tightest mass splittings, SRD1 targets signal models with $\Delta M = 50$ GeV, and SRD2 targets signal models with $\Delta M = 80$ GeV, separated by the number of calorimeter b-jets. Each also makes use of the angular distributions of the light jets, track jets, and track (and calorimeter)-b-jets. Table 9.4 outlines the selections applied [2] to each signal region.

Variable/SR	SRD0	SRD1	SRD2
$N_{b\text{-jet}}$	exactly 0	exactly 1	≥ 2
$p_{T(\text{non } b\text{-jet}_1)}$	$> 250 \text{ GeV}$		
$ \Delta\phi(\text{non} - b_1, E_T^{\text{miss}}) $	> 2.4		
$E_T^{\text{miss,track}}$	$> 30 \text{ GeV}$		
$ \Delta\phi(E_T^{\text{miss}}, E_T^{\text{miss,track}}) $	$< \pi/3$		
$N_{b\text{-jet}}^{\text{VR}}$	≥ 1	–	
$ \Delta\phi(\text{jet}^{1-4}, E_T^{\text{miss}}) $	> 0.4	–	
$ \eta^{b_1, \text{VR}} $	< 1.2	–	
$\max \Delta\phi(\text{non} - b_1, b^{\text{VR}}) $	> 2.2	–	
$ \Delta\phi(b_1^{\text{VR}}, b_2^{\text{VR}}) $	< 2.5	–	
$p_T^{b_1, \text{VR}}$	$< 50 \text{ GeV}$	$> 10 \text{ GeV}$	–
$p_T^{1, \text{VR}}$	–	$< 40 \text{ GeV}$	–
$\min \Delta\phi(\text{jet}^{\text{VR } 1-4}, \text{non} - b) $	–	> 1.2	–
$ \eta^{b_1} $	–	< 1.6	–
$ \Delta\phi(\text{non} - b, b_1) $	–	> 2.2	
$ \eta^{b_2} $	–	< 1.2	
$p_T^{b_1}$	–	$< 175 \text{ GeV}$	
$ \Delta\phi(\text{non} - b, b_2) $	–	> 1.6	
$E_T^{\text{miss}}/\sqrt{H_T}$	$> 26\sqrt{\text{GeV}}$	$> 22\sqrt{\text{GeV}}$	

Table 9.4: Signal region selections for the four body signal regions (SRD) as per [2]. Track-jet observables are labelled “track” or “VR” (denoting the “Variable radius track-jets”), while the leading non b-tagged calorimeter jet is denoted as the ISR (labelled “non-b”). Any other jets not labelled as such are calorimeter based anti- k_t $R = 0.4$ jets.

9.5 Control regions

Control regions and their definitions for the $Z + \text{jets}$ and $t\bar{t} + Z$ for signal regions A and B have been defined in sections 8.3.2 and 8.2.3. The remainder of the control regions corresponding to the other contributing backgrounds and signal regions will be discussed in this section.

9.5.1 Control regions for $t\bar{t}$, $W + \text{jets}$ and single top for signal regions SRA & SRB

In this section, the control regions for $t\bar{t}$, single-top and $W + \text{jets}$ backgrounds will be covered [2]. These regions are all one-lepton regions, and due to the processes considered (backgrounds contributing to the signal region due to a lepton which falls out of acceptance),

the lepton is treated as a jet in the calculation of variables such as the E_T^{miss} and the lepton resolution is added to the jet resolution in the definition of the Object based E_T^{miss} sig.. This follows on from the same prescription used in [3] (chapter 6). A notable definition was required for orthogonality with the signal region for the ATLAS one lepton stop search [112], in particular for CRTAB, namely $m_T(\ell, E_T^{\text{miss}}) < 120$ GeV, while also mitigating possible signal contamination. The definitions are as follows:

The proportions of the single top and $W + \text{jets}$ background in the signal region are substantially than $t\bar{t}$, $Z + \text{jets}$ or $t\bar{t} + W/Z$ processes, hence the impact of the normalisations on these backgrounds upon the final fit results are considerably smaller. Their principle motivation for inclusion in [2] was to reduce systematic uncertainties on the normalisation of the given process. Since the single top backgrounds are relatively small for all regions except a sub-leading contribution in signal region SRB, only a single control region is configured for signal regions SRA & SRB. The $W + \text{jets}$ background has a small contribution to both SRA, SRB and SRD signal regions, and a relevant set of control regions are defined in tables 9.5 and 9.7.

Some additional variables defined here to either require orthogonality with the signal regions of other analyses (principally this is the transverse mass of the lepton $m_T(\ell, E_T^{\text{miss}})$) are the separation between $\Delta R(b, \ell)$ required such that the electron/muon cannot be associated to the b-hadron decay.

Variable/CR	CRTAB	CRWAB	CRSTAB
Trigger $E_{\text{T}}^{\text{miss}}$ N_{ℓ}	$E_{\text{T}}^{\text{miss}}$ $> 250 \text{ GeV}$ $==1$		
p_{T}^{ℓ}	$< 20 \text{ GeV}$		$> 20 \text{ GeV}$
$m_{\text{T}}(\ell, E_{\text{T}}^{\text{miss}})$	$< 120 \text{ GeV}$	$< 100 \text{ GeV}$	
N_{jets} $p_{\text{T}}^{j_1}$ $p_{\text{T}}^{j_3}$	≥ 4 (incl. Lepton) $> 80 \text{ GeV}$ (incl. Lepton) $> 40 \text{ GeV}$ (incl. Lepton)		
N_b	≥ 2	$== 1$	≥ 2
$\min[\Delta\phi(j_{1-4}, E_{\text{T}}^{\text{miss}})]$ $m_{\text{jet}, R=1.2}^1$	> 0.4 (incl. Lepton) $> 120 \text{ GeV}$		
$m_{\text{T}}^{b, \text{min}}$	$< 200 \text{ GeV}$	$< 60 \text{ GeV}$	$> 120 \text{ GeV}$
$\Delta R(b_1, b_2)$	> 1.4	-	> 1.4
$m_{l, b, \text{min}}$ τ -veto	- -		$> 100 \text{ GeV}$ $> 100 \text{ GeV}$
Object based $E_{\text{T}}^{\text{miss}}$ sig.	> 14		
$\Delta R(b, l)$	-	> 2.0	-

Table 9.5: Table of definitions for the $t\bar{t}$, $W + \text{jets}$ and single top control regions from [2]

9.5.2 Control region for the compressed signal region: CRTC

Due to the ISR requirement of the signal region, the dominant backgrounds for SRC are multi-jet (which will be covered in 9.5.3 and $t\bar{t}$, and are not otherwise covered by the control regions defined in 9.5.1 due to the requirements of recursive jigsaw reconstruction variables instead of re-clustered $R=1.2$ jets. The primary source of $t\bar{t}$ in the signal region is through

a decay where either a hadronically decaying tau (likely reconstructed as a jet), E_T^{miss} and jets are produced, or a lepton which falls out of acceptance, E_T^{miss} and jets. A control region which requires a lepton to be in acceptance can be used to constrain this $t\bar{t}$ background. The selections are outlined in Table 9.6 [2]. In this section, the lepton is also treated as a jet for the purposes of all jet and Object based E_T^{miss} sig. calculations.

Variable/CR	CRTC
Trigger	E_T^{miss}
E_T^{miss}	$> 250 \text{ GeV}$
ℓ	$== 1$ (signal, baseline)
p_T^ℓ	$> 20 \text{ GeV}$
$m_T(\ell, E_T^{\text{miss}})$	$< 100 \text{ GeV}$
N_{jets}	≥ 4
Jet p_T	$> (80, 80, 40, 40) \text{ GeV}$
N_b	≥ 2
Object based E_T^{miss} sig.	> 5
$\min[\Delta\phi(j_{1-2}, E_T^{\text{miss}})]$	> 0.2
N_{jet}^S	≥ 4
$N_{b\text{-jet}}^S$	≥ 2
p_T^{ISR}	$> 400 \text{ GeV}$
$p_{T,b}^{1,S}$	$> 40 \text{ GeV}$
m_S	$> 400 \text{ GeV}$
$\Delta\phi_{\text{ISR}, E_T^{\text{miss}}}$	> 3.0
m_V/m_S	< 0.75
$\Delta R(b_{1,2}, \ell)_{\text{min}}$	< 2.0
$E_T^{\text{miss,track}}$	$> 30 \text{ GeV}$
$ \Delta\phi(E_T^{\text{miss}}, E_T^{\text{miss,track}}) $	$< \pi/3$

Table 9.6: Definition of the $t\bar{t}$ control region for the compressed signal regions, as listed in [2].

9.5.3 Multi-jet background estimation

As noted in [2] and the result of the study of section 8.1, as a selection on $\min[\Delta\phi(j_{1-4}, E_T^{\text{miss}})] > 0.4$ and Object based E_T^{miss} sig. > 11 (minimum) is applied in SRA and SRB, the multi-jet background is negligible. In SRC, the multi-jet background is non-negligible, and is estimated using the data-driven Jet smearing method [2, 104], using data using single-jet triggers instead of E_T^{miss} . This method is the same as used in [3] (see csection 6.3.5, but is updated in terms of object definitions and calibrations.

This has been also applied in the signal regions SRA, SRB and SRD, but is confirmed to be effectively negligible in these cases (and in all later tables will be merged into the “other” background). This is principally through the use of the $\min[\Delta\phi(j_{1-4}, E_T^{\text{miss}})]$ and $E_T^{\text{miss,track}}$ selections. Since this method is a data-driven background estimation, no normalisation control region has been used, and systematics relating to the multi-jet estimate are taken to be 30% as a conservative estimate of the configuration of the seed event, response function and tail corrections.

9.5.4 Four body control regions

The important backgrounds in the four body channels are $t\bar{t}$, $Z + \text{jets}$ and $W + \text{jets}$, with all other background components being sub-dominant ($t\bar{t}$ is also sub-dominant in SRD0). Since the requirement on number of calorimeter b-jets substantially alters the composition of the respective backgrounds, differing control and validation regions must be implemented for the respective signal regions. For the purposes of brevity, the selections will be outlined in tables 9.7, 9.8 and 9.9, but this is discussed further in [2].

Variable/CR	CRWD0	CRWD1	CRWD2
E_T^{miss}	$> 250 \text{ GeV}$		
Trigger	E_T^{miss}		
N_ℓ	$=1$		
p_T^ℓ	$> 20 \text{ GeV}$		
$p_T(\text{non b-jet}_1)$	$> 250 \text{ GeV}$		
$E_T^{\text{miss,track}}$	$> 30 \text{ GeV}$		
$ \Delta\phi(E_T^{\text{miss}}, E_T^{\text{miss,track}}) $	$< \pi/3$		
$ \Delta\phi(\text{non} - b_1, E_T^{\text{miss}}) $	> 2.4		
$m_T(\ell, E_T^{\text{miss}})$	$< 120 \text{ GeV}$		$< 100 \text{ GeV}$
$N_{b\text{-jet}}^{\text{VR}}$	≥ 1		-
$N_{b\text{-jet}}$	$=0$	$=1$	≥ 2
$\Delta R(b_{1,2}, \ell)_{\text{min}}(\text{calo.})$	-	> 1.8	> 2.2
$\Delta R(b_{1,2}, \ell)_{\text{min}}$	> 1.6	-	
$\min[\Delta\phi(j_{1-4}, E_T^{\text{miss}})]$	> 0.4	-	
$ \eta^{b_1, \text{VR}} $	< 1.2	-	
$\max \Delta\phi(\text{non} - b_1, b^{\text{VR}}) $	> 2.2	-	
$ \Delta\Phi(b_1^{\text{VR}}, b_2^{\text{VR}}) $	< 2.5	-	
$p_T^{b_1, \text{VR}}$	$< 50 \text{ GeV}$	$> 10 \text{ GeV}$	-
$p_T^{1, \text{VR}}$	-	$< 40 \text{ GeV}$	-
$\min \Delta\phi(\text{jet}^{\text{VR } 1-4}, \text{non} - b) $	-	> 1.2	-
$ \eta^{b_1} $	-	< 1.6	-
$p_T^{b_1}$	-		$< 175 \text{ GeV}$
$ \eta^{b_2} $	-		< 1.0
$\Delta R(b_1, b_2)$	-		< 1.0
$E_T^{\text{miss}}/\sqrt{H_T}$	$> 14\sqrt{\text{GeV}}$	$> 8\sqrt{\text{GeV}}$	$> 12\sqrt{\text{GeV}}$

Table 9.7: Control regions for $W + \text{jets}$ in the four body channels as per [2]

Selection	CRTD1	CRTD2
E_T^{miss}	$> 250 \text{ GeV}$	
Trigger	E_T^{miss}	
N_ℓ	$== 1$	
p_T^ℓ	$> 20 \text{ GeV}$	
$m_T(\ell, E_T^{\text{miss}})$	$< 120 \text{ GeV}$	
$E_T^{\text{miss,track}}$	$> 30 \text{ GeV}$	
$ \Delta\phi(E_T^{\text{miss}}, E_T^{\text{miss,track}}) $	$< \pi/3$	
$p_{T(\text{non b-jet}_1)}$	$> 250 \text{ GeV}$	
$ \Delta\phi(\text{non} - b_1, E_T^{\text{miss}}) $	> 2.4	
$ \Delta\phi(\text{non} - b_1, b_1) $	> 2.2	
$N_j \text{ (Calo.)}$	≥ 3	-
N_b	$== 1$	≥ 2
$N_{b\text{-jet}}^{\text{VR}}$	≥ 1	-
$p_{T_1}^{b_1, \text{VR}}$	$> 10 \text{ GeV}$	-
$p_{T_1}^{1, \text{VR}}$	$< 40 \text{ GeV}$	-
$\max \Delta\phi(\text{non} - b_1, b^{\text{VR}}) $	> 1.2	-
$ \eta^{b_1} $	< 1.6	-
$E_T^{\text{miss}}/\sqrt{H_T}$	$> 8\sqrt{\text{GeV}}$	$> 14\sqrt{\text{GeV}}$
$ \eta^{b_2} $	-	< 1.2
$ \Delta\phi(\text{non} - b_1, b_2) $	-	> 1.6
$p_T^{b_1}$	-	$< 175 \text{ GeV}$

 Table 9.8: Control regions for $t\bar{t}$ in the four body channels, as per [2].

Variable/CR	CRZD0	CRZD1	CRZD2
Trigger	Single lepton		
N_ℓ	$== 2$, OSSF		
$m(\ell, \ell)$	$[81, 101] \text{ GeV}$		
E_T^{miss}	$< 70 \text{ GeV}$		
p_T^ℓ	$> 30, > 27 \text{ GeV}$		
$N_{b\text{-jet}}$	$== 0$	$== 1$	$== 2$
$p_{T(\text{non b-jet}_1)}$	$> 250 \text{ GeV}$	$> 200 \text{ GeV}$	$> 250 \text{ GeV}$
$ \Delta\phi(\text{non} - b_1, E_T^{\text{miss}}) $	> 2.4		
$N_{b\text{-jet}}^{\text{VR}}$	≥ 1		-
$\min[\Delta\phi(j_{1-4}, E_T^{\text{miss}})]$	> 0.4	-	
$ \eta^{b_1, \text{VR}} $	< 1.2	-	
$\max \Delta\phi(\text{non} - b_1, b^{\text{VR}}) $	> 2.2	-	
$ \Delta\phi(b_1^{\text{VR}}, b_2^{\text{VR}}) $	< 2.5	-	
$p_{T_1}^{b_1, \text{VR}}$	$< 50 \text{ GeV}$	$> 10 \text{ GeV}$	-
$p_{T_1}^{1, \text{VR}}$	-	$< 40 \text{ GeV}$	-
$\min \Delta\phi(\text{jet}^{\text{VR } 1-4}, \text{non} - b) $	-	> 1.2	-
$ \eta^{b_1} $	-	< 1.6	-
$ \Delta\phi(\text{non} - b_1, b_1) $	-	> 1.8	> 2.2
$ \eta^{b_2} $	-	< 1.2	
$p_T^{b_1}$	-	$< 175 \text{ GeV}$	
$ \Delta\phi(\text{non} - b_1, b_2) $	-	> 1.6	
$E_T^{\text{miss}}/\sqrt{H_T}$	$> 12\sqrt{\text{GeV}}$	$> 8\sqrt{\text{GeV}}$	

 Table 9.9: Control regions for $Z + \text{jets}$ in the four body channels.

9.6 Validation regions

Validation regions for signal regions A & B

In this subsection, the validation regions as defined in [2] will be outlined, in particular for SRA and SRB. Each of these validation regions is a region containing no electrons or muons, and validation regions were defined for the following backgrounds:

- $t\bar{t}$: Single validation region for SRA & SRB.
- $Z + \text{jets}$: 3 validation regions, one for SRA (VRZA), and two for VRZB, separated into two categories (VRZB-TT-TW, VRZB-T0).

For the remaining backgrounds ($W + \text{jets}$, single-top and $t\bar{t} + Z$), there are no validation regions provided. Concerning the $W + \text{jets}$ and single-top backgrounds, no satisfactory validation region could be found that was sufficiently close to the SR that had sufficient purity of either $W + \text{jets}$ or single-top, even in a combined validation region. For $t\bar{t} + Z$, a candidate validation region was determined to have sufficient signal contamination ($> 15\%$) that it cannot be used as a validation region (this result lead to the implementation for SRA-TT-Disc), and a di-lepton implementation would face the same problems as a di-lepton $CRt\bar{t} + Z$: poor purity compared to large $t\bar{t}$ and $Z + \text{jets}$ contamination.

The tables below outline the principal selections for the $Z + \text{jets}$ and $t\bar{t}$ validation regions. The variables used for orthogonality with the signal regions are the b-jet content of the leading $R=1.2$ reclustered jet in the case of VRZA, and the $\Delta R(b_1, b_2)$ in the case of VRZB. To minimise extrapolation between SR and VR, where sufficient statistics are available (VRZB), the validation regions are subdivided into two bins. This is not viable for VRZA due to the limited validation region statistics.

Selection	VRZA	VRZB-TT-TW	VRZB-T0
N_ℓ	== 0		
E_T^{miss}	$> 250 \text{ GeV}$		
Trigger	E_T^{miss}		
N_{jets}	≥ 4		
Jet p_T	$> (80, 80, 40, 40) \text{ GeV}$		
$N_{b\text{-jet}}$	≥ 2		
$m_{\text{jet}, R=1.2}^1$	$> 120 \text{ GeV}$		
$m_T^{b, \text{min}}$	$\geq 200 \text{ GeV}$		
$\tau\text{-veto}$	$\checkmark$		
b-tagged jet in $(J_1; R = 1.2)$?	No	-	
Object based E_T^{miss} sig.	[14, 26]	[15, 17]	
m_{T2, χ^2}	$> 450 \text{ GeV}$	$< 450 \text{ GeV}$	
$\Delta R(b_1, b_2)$	-	< 1.4	
$m_{\text{jet}, R=1.2}^2$	-	$> 60 \text{ GeV}$	$< 60 \text{ GeV}$

Table 9.10: Validation regions for the $Z + \text{jets}$ background as used in [2].

For the validation region of $t\bar{t}$ (table 9.11, the principal orthogonalisation between signal region and validation region is the $m_T^{b, \text{min}}$, since this selection is explicitly used in the signal region to reduce the impact of this background when $m_T^{b, \text{min}} > 200 \text{ GeV}$.

Selection	VRTAB
N_ℓ	$== 0$
E_T^{miss}	$> 250 \text{ GeV}$
$E_T^{\text{miss}} \text{ trigger}$	$\checkmark$
N_{jets}	≥ 4
$N_{\text{b-jet}}$	≥ 2
$\min[\Delta\phi(j_{1-4}, E_T^{\text{miss}})]$	≥ 0.4
p_T^{j2}	$> 80 \text{ GeV}$
p_T^{j4}	$> 40 \text{ GeV}$
$\Delta R(b_1, b_2)$	> 1.4
$\tau\text{-veto}$	$\checkmark$
$m_{\text{jet}, R=1.2}^1$	$> 120 \text{ GeV}$
Object based E_T^{miss} sig.	> 14
$m_T^{b, \min}$	$[150, 200] \text{ GeV}$

Table 9.11: Definition of the validation region for $t\bar{t}$ for signal regions SRA and SRB, as used in [2].

Validation region for the compressed signal regions

In the case of the compressed signal regions SRC1-5, there is only a single validation region for the $t\bar{t}$ background, VRTC (table 9.12).

Variable	VRTC
Trigger	E_T^{miss}
E_T^{miss}	$> 250 \text{ GeV}$
control ℓ	$== 0$
additional baseline ℓ	0
N_{jets}	$\geq 4 \text{ same}$
p_T^{j2}	$> 80 \text{ GeV}$
p_T^{j4}	$> 40 \text{ GeV}$
$\min \Delta\phi(\text{jet}^{1-2}, E_T^{\text{miss}}) $	> 0.2
$N_{\text{b-jet}}$	≥ 2
$p_{T,b}^{1,S}$	$> 40 \text{ GeV}$
m_S	$> 400 \text{ GeV}$
$\Delta\phi_{\text{ISR}, E_T^{\text{miss}}}$	$\in [2.5, 3.0]$
p_T^{ISR}	$> 400 \text{ GeV}$
$p_{T,b}^{4,S}$	$> 50 \text{ GeV}$
m_V/m_S	< 0.6

Table 9.12: Table of selections for validation region VRTC in [2].

Four body validation regions

In the context of the track-jet based four body signal regions, there are five validation regions implemented corresponding to two validation regions for the $t\bar{t}$ background for signal regions SRD1 and SRD2 (but not SRD0, where $t\bar{t}$ is a sub-dominant background) and three validation regions for the $Z + \text{jets}$ background. Since the $W + \text{jets}$ background is comparatively small in the signal region, no validation region was provided for this background, since the impact on the fit itself of the normalisation of $W + \text{jets}$ is comparatively small.

Selection/VR	VRTD1	VRTD2
Trigger	E_T^{miss}	
E_T^{miss}	$> 250 \text{ GeV}$	
N_ℓ	$= 0$	
N_{jets}	≥ 4	
Jet p_T	$> (80, 80, 40, 40)$	
$E_T^{\text{miss,track}}$	$> 30 \text{ GeV}$	
$ \Delta\phi(E_T^{\text{miss}}, E_T^{\text{miss,track}}) $	$< \pi/3$	
N_b	$= 2$	≥ 2
$p_{T(\text{non b-jet}_1)}$	$> 250 \text{ GeV}$	
$ \eta^{b_1} $	< 1.6	-
$ \eta^{b_2} $	-	< 1.2
$p_T^{b_1}$	$< 200 \text{ GeV}$	$< 175 \text{ GeV}$
$ \Delta\phi(\text{non} - b_1, b_1) $	< 0.8	< 2.0
$ \Delta\phi(\text{non} - b_1, b_2) $	-	> 2.0
$N_{b\text{-jet}}^{\text{VR}}$	≥ 1	-
$p_{T_1}^{\text{VR}}$	$< 60 \text{ GeV}$	-
$E_T^{\text{miss}}/\sqrt{H_T}$	$> 18 \text{ GeV}$	$> 21 \text{ GeV}$
$\Delta R(b_1, b_2)$	> 2.0	-

Table 9.13: Validation regions for the $t\bar{t}$ background corresponding to the four body signal regions SRD1 and SRD2, as used in [2].

Selection/VR		VRZD0	VRZD1	VRZD2
Trigger		$E_{\text{T}}^{\text{miss}}$		
$E_{\text{T}}^{\text{miss}}$		$> 250 \text{ GeV}$		
N_{ℓ}		$== 0$		
N_j		≥ 4		
Jet p_{T}		$> (80, 80, 40, 40)$		
$E_{\text{T}}^{\text{miss,track}}$		$> 30 \text{ GeV}$		
$ \Delta\phi(E_{\text{T}}^{\text{miss}}, E_{\text{T}}^{\text{miss,track}}) $		$< \pi/3$		
N_b		$== 0$	$== 1$	≥ 2
$p_{\text{T}}(\text{non b-jet}_1)$		$> 250 \text{ GeV}$		
$ \eta^{b_1} $		-	< 1.6	-
$ \eta^{b_2} $		-		< 1.2
$p_{\text{T}}^{b_1}$		-		$< 175 \text{ GeV}$
$ \Delta\phi(\text{non} - b_1, b_1) $		-	< 2.2	
$ \Delta\phi(\text{non} - b_1, b_2) $		-		< 2.0
$N_{\text{b-jet}}^{\text{VR}}$		≥ 1		-
$p_{\text{T}}^{b_1, \text{VR}}$		$< 50 \text{ GeV}$	$> 10 \text{ GeV}$	-
$p_{\text{T}}^{1, \text{VR}}$		-	$< 40 \text{ GeV}$	-
$ \eta^{b_1, \text{VR}} $		< 1.2	-	
max	$ \Delta\phi(\text{non} - b_1, b^{\text{VR}}) $	< 2.0	-	
	$ \Delta\phi(b_1^{\text{VR}}, b_2^{\text{VR}}) $	< 2.5	-	
$E_{\text{T}}^{\text{miss}}/\sqrt{H_{\text{T}}}$		$> 25\sqrt{\text{GeV}}$	$> 22\sqrt{\text{GeV}}$	$> 22\sqrt{\text{GeV}}$
min $[\Delta\phi(j_{1-4}, E_{\text{T}}^{\text{miss}})]$		> 0.4	-	

Table 9.14: Validation regions for the $Z + \text{jets}$ background corresponding to the four body signal regions SRD0, SRD1 and SRD2, as used in [2].

9.7 Systematics

Since we have applied object definitions and calibrations, each of these have respective systematics associated with their variations. These normally involve the variation of some reconstruction parameter, corrections or statistical uncertainties in the calibrations and identification. In the analysis from [2], the experimental systematics that contribute are as follows (a full breakdown of the systematics in the fit is presented in appendix D:

- Electrons: Resolution and energy scale calibrations, identification and isolation efficiency uncertainties, and trigger efficiencies in the case of regions where the electron trigger is fired (control regions).
- Muons: Muon Sagitta estimation uncertainties, identification uncertainties, isolation uncertainties, reconstruction uncertainties and trigger efficiencies in the case of regions where the muon trigger is fired (control regions).
- Tri-lepton specific uncertainties (One lepton trigger overlap). This is a special case in 3-electron/muon regions only, because the object firing the relevant trigger is ambiguous: in all other control regions, only one type of lepton is allowed in the event. This ambiguity is best demonstrated by example: a $t\bar{t} + Z$ event has two OSSF electrons (from the Z) and one muon (from a top decay), a valid event in the control region. It can be that one of the Z-associated leptons can fire the electron trigger (hence one must apply the uncertainties and calibration for that trigger efficiency) or the muon can fire the muon trigger (hence again uncertainties and calibration for that trigger), or in some cases both can occur. A software based tool is used to resolve this ambiguity using the event information, and correctly apply the trigger efficiency uncertainties.
- Jets: Jet Energy Scale (JES) calibrations, Jet Energy Resolution (JER), Jet Vertex Tagger (JVT) efficiencies/corrections.
- Flavour tagging uncertainties associated to MV2c10 b-tagging: b/c/light efficiency uncertainties, efficiency extrapolation uncertainties.
- E_T^{miss} : Uncertainties in measuring the track soft term, uncertainty in the resolution of the E_T^{miss} and its respective constituents.

These uncertainties (or subsets therein) may be manifestly correlated, and some studies into de-correlated set of systematics were performed.

Overview of theoretical systematics

Unlike the experimental systematics, which largely arise from variation of given parameters in reconstruction, calibration etc, the theoretical systematics (which arise from the validity of a given MC generated sample to reflect a given process) are somewhat more complex to obtain. The wide variety of input parameters to each MC generator and sample (of which for many, may not result in a substantive change in a given sample yield or distribution) may vary depending on the specific implementations of MC event generator and parton shower. The respective model assumptions will mean that there is no consistent approach

that correctly covers the full uncertainty on the theoretical cross section of a given sample considering all possible alternate parametrisations, let alone be consistent with any other generator or parton shower algorithm. Since each specific implementation is distinct, a good estimate of such theoretical cross section uncertainties is undertaken through variation of known key parameters or techniques. The following list of systematics underpins the theory systematics used in [2], which reflect a set of input assumptions into the given MC generator and parton shower algorithm.

- The renormalisation and factorisation scales in $a\text{MC@NLO} + \text{PYTHIA8}$. ($t\bar{t} + Z$)
- A14 variations in $a\text{MC@NLO} + \text{PYTHIA8}$ (Parton shower variations) ($t\bar{t} + Z$). These correspond to different variations in the Pythia configurations (referred to as a “tune”), and roughly approximate to an up/down variation of the parton shower uncertainties. [51]
- Variations in the scales associated to ISR and FSR in the configuration settings of $\text{POWHEG} + \text{PYTHIA}$ ($t\bar{t}$).
- Resummation, factorisation and renormalisation scales in Sherpa ($Z + \text{jets}$, $W + \text{jets}$).
- Heavy flavour composition uncertainties due to extrapolations in $N_{\text{b-jet}}$ between CRWAB and SRA or SRB ($W + \text{jets}$).
- Single top theory uncertainties: two different schemes used to separate the Wt diagrams from $t\bar{t}$ interference terms at NLO. These are referred to as “diagram subtraction” (DS) and “diagram reduction” (DR), which produce different NLO corrections. A full explanation of this is out of the scope of this thesis, but further elaboration for the interested reader can be found in [113].
- Parton shower variations, renormalisation and factorisation scales for signals.

To compare the large set of additional parameters and modelling assumptions that vary between generator and parton shower combinations (and even specific versions thereof), systematics are also either compared or included with systematics associated to producing the same process with an alternate MC event generator and/or parton shower algorithm, where sample availability permits. Some of these uncertainties may be calculated from the samples themselves, but others require the processing of additional samples, and where reconstructed samples are unavailable at time of publication, generator level comparisons are undertaken instead.

There is no coherent prescription that can be applicable for every MC sample in any generator/parton shower combination, the combination of these uncertainties could only be considered a conservative upper limit estimate compared to a full parameter variation calculation (or the true uncertainty on the cross section due to the specific generator and parton shower combination). Since this analysis is primarily dominated by statistical uncertainty in the signal region, often followed by statistical uncertainties due to the control regions, this prescription is sufficient to produce a conservative upper limit in the time available. However, this holds only in the assumption that a signal region does not have an expected signal significance of $\sim 5\sigma$ or is not a measurement, in which a more detailed theoretical prescription may need to be estimated.

The following subsections will outline the author's work into calculating some of the theory uncertainties, primarily for the $t\bar{t} + Z$ background. The remainder are briefly summarised in [2].

9.7.1 $t\bar{t} + Z$ theory uncertainties: scale variations, alternate generator and parton shower uncertainties

In order to estimate the effect of different generators or parton showers, a study in the signal region comparing $t\bar{t} + Z (\rightarrow \nu\nu)$ samples produced with alternate parton shower and generator combinations was required. The default (as mentioned in Appendix A.1) nominal sample for $t\bar{t} + Z$ (both $t\bar{t} + Z(\rightarrow \nu\nu)$ and $t\bar{t} + Z(Z \rightarrow \ell\ell)$) is aMC@NLO generator interfaced with PYTHIA (NLO cross section), with the underlying event configured using the A14 tuning scheme.

As an estimate for the uncertainties on the cross section for $t\bar{t} + Z (\rightarrow \nu\nu)$, NLO Sherpa 2.2.1 $t\bar{t} + Z (\rightarrow \nu\nu)$ sample would normally be sufficient to estimate a systematic due to an alternate MC generator and parton shower. To validate uncertainties due to discrepancies in the parton shower choice, a sample combining the aMC@NLO generator with the Herwig7 was considered. In order to consider uncertainties due to the variations in the renormalisation (μ_R) and factorisation scales (μ_F) in the generator, differing weights were applied to the sample to produce each respective variation. To consider the maximal effect of these scale variations, a set of alternate combinations of μ_F, μ_R , as per figure 9.2 was chosen.

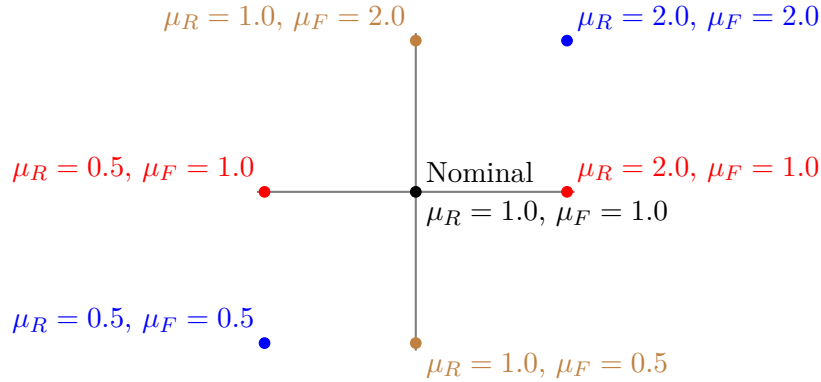


Figure 9.2: Schema for the different scale variations of μ_F and μ_R considered, namely by multiplying the nominal factor by either two or one half. Colour coding denotes the different uncertainties: variations due to μ_R , holding μ_F constant (red), the converse case (brown) and the uncertainty due to varying μ_R and μ_F simultaneously (blue).

Samples

In this study, the following samples were used for comparisons:

- aMC@NLO + Pythia 8 ($t\bar{t} + Z$ Nominal) - FullSim : $t\bar{t} + ee, t\bar{t} + \mu\mu$.
- aMC@NLO + Herwig7 - AFII (alternate parton shower sample): $t\bar{t} + ee, t\bar{t} + \mu\mu$.

- Sherpa NLO Multi-Leg (i.e. +1 parton) - FullSim: $t\bar{t} + (ee/\mu\mu)$ - alternate generator and alternate parton shower sample.
- aMC@NLO + Pythia8 ($t\bar{t} + Z$ μ_R, μ_F scale variations): $t\bar{t} + ee$, $t\bar{t} + \mu\mu$, $t\bar{t} + Z\nu\nu$.

Note, the $t\bar{t} + \tau\tau$ samples are not used, as the contributions to the 3L SR and 2L CR are very small $O(\ell^+\ell^-1\%)$ for the nominal sample, and the equivalent is not available in all of the respective generator and parton shower combinations.

At time of writing, the samples for $t\bar{t} + Z(\rightarrow \nu\nu)$ either with alternate generator or parton shower were not available. It is noted that although the $t\bar{t} + Z(\rightarrow \nu\nu)$ samples were unavailable, the di-leptonic $t\bar{t} + Z(\rightarrow \ell\ell)$ samples were available for both the alternate parton shower and Sherpa comparison. Since a direct reproduction of the signal region was unavailable, an emulation of the signal region was proposed in a di-lepton regime to make a comparison between the different generators and parton shower combinations.

The sherpa multi-leg sample was not generally a preferred sample either, because of the requirements of an additional parton at generator level (which will affect distributions such as the N_{jets}), but a suitable replacement sample was not available. In addition, there are substantive differences between the default $t\bar{t} + Z$ sample and Sherpa 2.2.1 also including an alternative form of parton shower (Sherpa is a full generator including both components internally), leading to some ambiguity in the comparisons. Hence, the principal estimator of the matrix element generator uncertainties will be taken using the generator scale variations, with this alternate sample considered as a cross-check.

Thus, a study is required for the following:

1. Extract the uncertainties due to μ_R, μ_F scale variations
2. Check that the μ_R, μ_F scale variations are comparable in scale to a possible variation one could extract using the alternate matrix element generator. If so, then use the scale variations as an estimate of the uncertainty due to change of matrix element generator.
3. Extract the uncertainty due to variation of parton shower algorithm, given a fixed input matrix element generator.
4. Extract the A14 variation uncertainties: this can be extracted by a generator level comparison and emulation of the control and signal regions and the relevant transfer factor. The choice of generator level extraction was primarily motivated solely by sample availability. The author did not contribute to this estimation and is not explicitly presented in [2], so will not elaborate on this further.
5. Compare the parton shower variation uncertainties to the variations due to alternation of Pythia tunes (A14 variations).

The first of these items can be directly calculated from the control and signal region definitions, since it is included in the sample as an event level weighting. The remainder must be extracted from emulated signal regions, due to the absence of the relevant samples with no generator level electrons/muons. The remainder of the section outlines the procedure undertaken by the author to construct the emulated regions, and extract the relevant uncertainties (therelative uncertainty in the transfer factor is applied in the fit).

Selecting a region to compare the variations

In this study, a full emulation of the signal region was performed by a requirement of 2 signal leptons in the region of interest, and calculating the E_T^{miss} (prime), treating the leptons as invisible particles - an identical treatment to that for control region CRZ. With the altered calculation of the E_T^{miss} , all variables dependent on the E_T^{miss} were recomputed to include the leptons treated as invisible in this di-lepton selection. A requirement that the leptons were oppositee sign and same flavour was applied in this selection. As a first pass, the full signal region SRA-TT selection was emulated and applied, since $t\bar{t} + Z$ is particularly prevalent in this signal region, replacing the previous E_T^{miss} trigger selection with a requirement on the single lepton trigger, and a trigger threshold p_T selection on $p_T(\ell 1) > 27$ GeV was required. The tables below outline the emulated preselection and the emulated signal region, while control region CRttZ-AB (table 8.5) is considered as the reference control region:

Selection	Threshold
Trigger	Single e/μ
N_L	$= 2$
p_T^ℓ	$> (27, 20)$ GeV
OSSF?	Yes
$E_T^{\text{miss}'}$	> 250 GeV
N_j	≥ 4
N_b	≥ 2
$p_T^{\text{jet},2}$	80 GeV
$p_T^{\text{jet},4}$	40 GeV
$\Delta\phi(\mathbf{p}_T^{\text{miss}'}, j^{1-4})$	> 0.4
Object based $E_T^{\text{miss}'}$ signif.	> 5
τ -veto?	Yes
$m_T(b, E_T^{\text{miss}'})$ min	> 50 GeV

Table 9.15: $2e/\mu$ emulated pre-selection.

Selection	Threshold
Preselection	2L emulated pre-selection
$m_T(b, E_T^{\text{miss}'})$ min	> 200 GeV
$m(J1; R = 1.2)$	> 120 GeV
$m(J1; R = 0.8)$	> 60 GeV
$m_{T2}^{\chi^{2'}}$	> 450 GeV
$E_T^{\text{miss}'}$ signif.	$> 25\sqrt{\text{GeV}}$
$(J1; R = 1.2)$ b-tag content?	Yes
$(J2; R = 1.2)$ b-tag content?	Yes
$m(J2; R = 1.2)$	> 120 GeV
$\Delta R(b_1, b_2)$	> 1.0

Table 9.16: $2e/\mu$ emulated version of SRATT (SRA-2L-Tight).

9.7.1. $t\bar{t} + Z$ theory uncertainties: scale variations, alternate generator and parton shower uncertainties

Region	Sample	Yield
CRttZ	aMc+Py8 (Nominal)	52.61 ± 0.58
	aMC+Herwig7 (altPS)	43.09 ± 0.34
	Sherpa Multi-leg	54.04 ± 0.42
	Sherpa NLO (incl)	-
SRA-2L-Tight	aMc+Py8 (Nominal)	0.142 ± 0.031
	aMC+Herwig7 (altPS)	0.105 ± 0.019
	Sherpa Multi-leg	0.157 ± 0.022
	Sherpa NLO (incl)	-

Table 9.17: Yields for the different variation samples in CR $t\bar{t} + Z$ and the full 2L emulated version of SRA-TT, as defined in table 9.16.

From table 9.17, it can be seen that a full emulation of SRA-TT is limited by statistics in the emulated region (which is also prevalent in the real 0 lepton SR), so a loose signal region emulation must be considered instead.

As a result, the key variables required in the signal region must be checked for stability over extrapolation, particularly those relating to jets (which are mostly fixed between CR and SR, with exception of the R=1.2 fat-jet masses), and $E_T^{\text{miss}'}$. The the $E_T^{\text{miss}'}$ dependent variables ($m_T^{b, \text{min}'}$, Object based $E_T^{\text{miss}'}$ sig. , m'_{T2, χ^2}) have been checked at preselection level, and the distributions are outlined in figures 9.3a-9.3d.

Selection	Threshold			
	SRA-2L-MTbmin	SRA-2L-metsig	SRA-2L-comb	
$E_{\text{T}}^{\text{miss}'}$	> 250 GeV			Preselection
N_j	≥ 4			
N_b	≥ 2			
p_T^{j2}	> 80 GeV			
p_T^{j4}	> 40 GeV			
N_l	$== 2$			
OSSF?	Yes			
$m(J1; R = 1.2)$	> 120 GeV			Top Categories
$m(J1; R = 1.2)$	> 120 GeV			
$m_T(b, E_{\text{T}}^{\text{miss}'}) \text{ min}$	> 200 Gev	-	> 100 GeV	
Object-based $E_{\text{T}}^{\text{miss}'}$ signif.	-	> 20	> 10	

Table 9.18: Table of 2L Emulated SRATT selections.

9.7.1. $t\bar{t} + Z$ theory uncertainties: scale variations, alternate generator and parton shower uncertainties

sample	CR-ttZ Yield	SRA-2L-comb Yield	Transfer factor	Ratio to nom. TF
Nominal ttll	52.61 ± 0.58	7.38 ± 0.23	0.1403 ± 0.0047	-
aMC+Herwig	43.09 ± 0.34	6.16 ± 0.13	0.1429 ± 0.0031	1.0190(+1.90%)
SherpaMultiLeg	54.04 ± 0.42	7.40 ± 0.15	0.1369 ± 0.0030	0.9761(-2.39%)
Scale Variations				
$\mu_R = 2.0, \mu_F = 2.0$	46.90 ± 0.45	6.74 ± 0.18	0.1436 ± 0.0041	1.0239(+2.39%)
$\mu_R = 2.0, \mu_F = 1.0$	48.84 ± 0.48	7.01 ± 0.20	0.1435 ± 0.0042	1.0230(+2.30%)
$\mu_R = 1.0, \mu_F = 2.0$	51.08 ± 0.53	7.19 ± 0.21	0.1408 ± 0.0044	1.0039(+0.39%)
$\mu_R = 0.5, \mu_F = 0.5$	56.78 ± 0.78	7.62 ± 0.32	0.1341 ± 0.0059	0.9561(-4.39%)
$\mu_R = 0.5, \mu_F = 1.0$	55.96 ± 0.71	7.61 ± 0.29	0.1359 ± 0.0054	0.9691(-3.09%)
$\mu_R = 1.0, \mu_F = 0.5$	54.37 ± 0.64	7.58 ± 0.26	0.1394 ± 0.0050	0.9936(-0.64%)

Table 9.19: Yields and transfer factors for the region SRA-2L-comb in 9.18. The key points to compare are: the uncertainty in the transfer factor (last column) for the altGen variation and the μ_R/μ_F scale variations. The alternate parton shower transfer factor uncertainty is also shown as included in the fit.

Since the scale variations (μ_R, μ_F) are derived from internal weights applied to the samples directly, we can calculate these directly using the samples available. Tables

Variation	Yield CRttZ-pT200 (tt+ll)	Yield SRATT (ttZvv)	TransferFactor	Ratio to nom. TF
Nominal	52.6126 ± 0.5758	1.2863 ± 0.1840	0.0244 ± 0.00	-
$\mu_R = 2.0, \mu_F = 2.0$	52.7332 ± 0.5017	1.3128 ± 0.1632	0.0249 ± 0.0031	1.0182(+1.82%)
$\mu_R = 2.0, \mu_F = 1.0$	53.3393 ± 0.5267	1.3480 ± 0.1745	0.0253 ± 0.0033	1.0337(+3.37%)
$\mu_R = 1.0, \mu_F = 2.0$	52.2040 ± 0.5469	1.2588 ± 0.1713	0.0241 ± 0.0033	0.9863(-1.37%)
$\mu_R = 0.5, \mu_F = 0.5$	51.7262 ± 0.7135	1.2221 ± 0.2244	0.0236 ± 0.0044	0.9664(-3.36%)
$\mu_R = 0.5, \mu_F = 1.0$	51.7407 ± 0.6585	1.2109 ± 0.2028	0.0234 ± 0.0039	0.9572(-4.28%)
$\mu_R = 1.0, \mu_F = 0.5$	52.9743 ± 0.6191	1.3126 ± 0.2014	0.0248 ± 0.0038	1.0135(+1.35%)

Table 9.20: Direct calculation of the transfer factors from CR $t\bar{t} + Z$ to SRATT for the renormalisation and factorisation scale uncertainties.

Variation	Yield CRttZ-pT200 (tt+ll)	Yield SRAT0 (ttZvv)	TransferFactor	Ratio to nom. TF
Nominal	52.6126 ± 0.5758	1.7471 ± 0.2016	0.0332 ± 0.00	-
$\mu_R = 2.0, \mu_F = 2.0$	52.7332 ± 0.5017	1.7141 ± 0.1769	0.0325 ± 0.0034	0.9789(-2.11%)
$\mu_R = 2.0, \mu_F = 1.0$	53.3393 ± 0.5267	1.7602 ± 0.1883	0.0330 ± 0.0035	0.9938(-0.62%)
$\mu_R = 1.0, \mu_F = 2.0$	52.2040 ± 0.5469	1.7030 ± 0.1883	0.0326 ± 0.0036	0.9824(-1.76%)
$\mu_R = 0.5, \mu_F = 0.5$	51.7262 ± 0.7135	1.7829 ± 0.2460	0.0345 ± 0.0048	1.0380(+3.80%)
$\mu_R = 0.5, \mu_F = 1.0$	51.7407 ± 0.6585	1.7404 ± 0.2236	0.0336 ± 0.0043	1.0129(+1.29%)
$\mu_R = 1.0, \mu_F = 0.5$	52.9743 ± 0.6191	1.7934 ± 0.2192	0.0339 ± 0.0042	1.0195(+1.95%)

Table 9.21: Direct calculation of the transfer factors from CR $t\bar{t} + Z$ to SRAT0 for the renormalisation and factorisation scale uncertainties.

Variation	Yield CRttZ-pT200 (tt+ll)	Yield SRATW (ttZvv)	TransferFactor	Ratio to nom. TF
Nominal	52.6126 ± 0.5758	0.8938 ± 0.1557	0.0170 ± 0.00	-
$\mu_R = 2.0, \mu_F = 2.0$	52.7332 ± 0.5017	0.9047 ± 0.1359	0.0172 ± 0.0026	1.0099(+0.99%)
$\mu_R = 2.0, \mu_F = 1.0$	53.3393 ± 0.5267	0.8955 ± 0.1430	0.0168 ± 0.0027	0.9883(-1.17%)
$\mu_R = 1.0, \mu_F = 2.0$	52.2040 ± 0.5469	0.9087 ± 0.1468	0.0174 ± 0.0028	1.0246(+2.46%)
$\mu_R = 0.5, \mu_F = 0.5$	51.7262 ± 0.7135	0.8613 ± 0.1905	0.0167 ± 0.0037	0.9802(-1.98%)
$\mu_R = 0.5, \mu_F = 1.0$	51.7407 ± 0.6585	0.8970 ± 0.1750	0.0173 ± 0.0034	1.0205(+2.05%)
$\mu_R = 1.0, \mu_F = 0.5$	52.9743 ± 0.6191	0.8696 ± 0.1673	0.0164 ± 0.0032	0.9663(-3.37%)

Table 9.22: Direct calculation of the transfer factors from CR $t\bar{t} + Z$ to SRATW for the renormalisation and factorisation scale uncertainties.

9.7.1. $t\bar{t} + Z$ theory uncertainties: scale variations, alternate generator and parton shower uncertainties

Variation	Yield CRttZ-pT200 (tt+ll)	Yield SRBT0 (ttZvv)	TransferFactor	Ratio to nom. TF
Nominal	52.6126 $\pm$ 0.5758	26.9200 $\pm$ 0.7873	0.5117 $\pm$ 0.02	-
$\mu_R = 2.0, \mu_F = 2.0$	52.7332 $\pm$ 0.5017	26.4136 $\pm$ 0.6909	0.5009 $\pm$ 0.0139	0.9789(-2.11%)
$\mu_R = 2.0, \mu_F = 1.0$	53.3393 $\pm$ 0.5267	26.7695 $\pm$ 0.7222	0.5019 $\pm$ 0.0144	0.9809(-1.91%)
$\mu_R = 1.0, \mu_F = 2.0$	52.2040 $\pm$ 0.5469	26.5880 $\pm$ 0.7501	0.5093 $\pm$ 0.0153	0.9954(-0.46%)
$\mu_R = 0.5, \mu_F = 0.5$	51.7262 $\pm$ 0.7135	27.4808 $\pm$ 0.9643	0.5313 $\pm$ 0.0200	1.0383(+3.83%)
$\mu_R = 0.5, \mu_F = 1.0$	51.7407 $\pm$ 0.6585	27.2120 $\pm$ 0.8929	0.5259 $\pm$ 0.0185	1.0279(+2.79%)
$\mu_R = 1.0, \mu_F = 0.5$	52.9743 $\pm$ 0.6191	27.2385 $\pm$ 0.8425	0.5142 $\pm$ 0.0170	1.0049(+0.49%)

Table 9.23: Direct calculation of the transfer factors from CR $t\bar{t} + Z$ to SRBT0 for the renormalisation and factorisation scale uncertainties.

Variation	Yield CRttZ-pT200 (tt+ll)	Yield SRBTT (ttZvv)	TransferFactor	Ratio to nom. TF
Nominal	52.6126 $\pm$ 0.5758	11.7330 $\pm$ 0.5755	0.2230 $\pm$ 0.01	-
$\mu_R = 2.0, \mu_F = 2.0$	52.7332 $\pm$ 0.5017	11.9410 $\pm$ 0.5025	0.2264 $\pm$ 0.0098	1.0154(+1.54%)
$\mu_R = 2.0, \mu_F = 1.0$	53.3393 $\pm$ 0.5267	12.0440 $\pm$ 0.5300	0.2258 $\pm$ 0.0102	1.0125(+1.25%)
$\mu_R = 1.0, \mu_F = 2.0$	52.2040 $\pm$ 0.5469	11.6903 $\pm$ 0.5429	0.2239 $\pm$ 0.0107	1.0042(+0.42%)
$\mu_R = 0.5, \mu_F = 0.5$	51.7262 $\pm$ 0.7135	11.2340 $\pm$ 0.7077	0.2172 $\pm$ 0.0140	0.9739(-2.61%)
$\mu_R = 0.5, \mu_F = 1.0$	51.7407 $\pm$ 0.6585	11.3489 $\pm$ 0.6487	0.2193 $\pm$ 0.0128	0.9836(-1.64%)
$\mu_R = 1.0, \mu_F = 0.5$	52.9743 $\pm$ 0.6191	11.7382 $\pm$ 0.6214	0.2216 $\pm$ 0.0120	0.9936(-0.64%)

Table 9.24: Direct calculation of the transfer factors from CR $t\bar{t} + Z$ to SRBTT for the renormalisation and factorisation scale uncertainties. .

Variation	Yield CRttZ-pT200 (tt+ll)	Yield SRBTW (ttZvv)	TransferFactor	Ratio to nom. TF
Nominal	52.6126 $\pm$ 0.5758	14.6962 $\pm$ 0.6184	0.2793 $\pm$ 0.01	-
$\mu_R = 2.0, \mu_F = 2.0$	52.7332 $\pm$ 0.5017	14.6979 $\pm$ 0.5408	0.2787 $\pm$ 0.0106	0.9978(-0.22%)
$\mu_R = 2.0, \mu_F = 1.0$	53.3393 $\pm$ 0.5267	14.8157 $\pm$ 0.5679	0.2778 $\pm$ 0.0110	0.9944(-0.56%)
$\mu_R = 1.0, \mu_F = 2.0$	52.2040 $\pm$ 0.5469	14.6295 $\pm$ 0.5861	0.2802 $\pm$ 0.0116	1.0033(0.33%)
$\mu_R = 0.5, \mu_F = 0.5$	51.7262 $\pm$ 0.7135	14.5174 $\pm$ 0.7591	0.2807 $\pm$ 0.0152	1.0048(0.48%)
$\mu_R = 0.5, \mu_F = 1.0$	51.7407 $\pm$ 0.6585	14.5789 $\pm$ 0.6991	0.2818 $\pm$ 0.0140	1.0087(0.87%)
$\mu_R = 1.0, \mu_F = 0.5$	52.9743 $\pm$ 0.6191	14.7346 $\pm$ 0.6648	0.2781 $\pm$ 0.0130	0.9958(-0.42%)

Table 9.25: Direct calculation of the transfer factors from CR $t\bar{t} + Z$ to SRBTW for the renormalisation and factorisation scale uncertainties.

Summary

To incorporate this into an estimate of the theoretical variations of the $t\bar{t} + Z$ background, we subdivide the theoretical uncertainties into four nuisance parameters: the systematic due to the renormalisation scale, while holding the factorisation scale fixed, vice versa, the systematic due to the effect of both and the uncertainty due to the parton showering algorithm.

For the scale variation uncertainties or generator level uncertainties, these are taken to be from the respective variations of renormalisation and factorisation scales from the MC generator, and this was cross checked against multi-leg Sherpa and found to be consistent in the 2L emulated regime, with three systematics used as per the diagram in figure 9.2.

For the parton shower level uncertainties, the uncertainty estimate is taken from the maximised-symmetrised envelope of the uncertainties in the transfer factor due to $t\bar{t} + Z$ radiation uncertainties (calculated for [2]), which are determined by alternate “tunes” of PYTHIA 8 (this is the A14 variations), and the uncertainty in the transfer factor from changing the parton shower algorithm to HERWIG 7 calculated in the SRA-2L-comb region defined in 9.18. In general, these uncertainties were compatible across all of the signal

regions for SRA and SRB. The strategy chosen in [2] was to use the “combination” 2L emulated region (where Object based $E_T^{\text{miss}'} \text{ sig.} > 10$ and $m_T^{b,\text{min}'} > 100$) and the respective top categories for each signal region (based on the sub-leading $R=1.2$ jet constraint as per signal regions SRA & SRB) to retrieve the alternate parton shower uncertainties.

Concluding remarks

This chapter has outlined the configuration of control, signal and validation regions for the analysis covering the run periods 2015-18 [2]. The re-optimisation of signal regions SRA, SRB and SRC has been considered, alongside a definition of the new four-body signal regions. A more detailed discussion of the systematics used is undertaken, outlining the author’s contribution.

The signal region design of ATLAS contrasts with that of CMS, where ATLAS signal region designs principally rely on a smaller subset of physically motivated “discovery” regions, binned using topologically relevant selections to improve expected exclusion performance. CMS, on the other hand often uses primarily high granularity signal regions (corresponding to individual bins in a given variable, or pair of variables), performing multi-bin fits to undertake the exclusion. This multi-bin fit on a large number of signal regions leads to stronger exclusion limits (in the cases of no significant excesses), while the ATLAS strategy primarily favours discovery of supersymmetry (since excesses in large number of signal regions are more likely due to statistical variation as per the look-elsewhere effect). This comment will be again discussed further when fit results are evaluated in the next chapter.

Fit Results

This chapter outlines the maximal likelihood fits to the normalisation of different backgrounds for signal, control and validation regions specified in the previous chapter, and the extraction of hypothesis test results to determine the existence/non-existence of SUSY signals in ATLAS searches, as discussed in Chapter 5. The author contributed substantially to the fits, both in front-end validation/publication and back-end implementation. A major novel component was the use of HTCondor batch systems in preparation of samples for fits, which substantially reduced the wall-time from the order of days to produce a fit, to the order of 8hrs from input file to final fit result. In addition, the implementation of the theory systematics calculated in the previous chapter was undertaken by the author.

This chapter then is sub-divided into three components, first the configuration of the fits, then the results and finally a discussion in light of results of other ATLAS searches and those from CMS, and will include material presented in [2] as well as supplementary material from the author relating to signal regions SRA and SRB.

10.1 Fit configuration

This section uses the profile-likelihood fitting methodology using HistFitter as discussed in section 5.1, and uses the configuration of control, signal and validation regions as mentioned in the last chapter.

Since the respective signal regions have different dominant and sub-dominant background profiles and different event topology, the analysis is configured with 3 distinct setups, namely a fit for all SRA and SRB combined, a fit for all SRC regions and a separate fit for all SRD regions. There are 3 types of fit performed, relating to the type of hypothesis test involved (or for validation):

1. Background only fit: control regions are fitted with their maximal likelihood estimated normalisations, nuisance parameters related to the control regions are also constrained, while the signal regions are treated as a validation region. This is mostly used for validation of the fit setup, and the outputs of this fit can be plotted as variable distributions for either signal, control or validation regions.

2. Model dependent exclusion fit (*exclusion fit*): For each group of signal regions, and a given signal, the signal model is inserted and the $p(\text{data}|\text{signal} + \text{background})$ is observed, to obtain the observed exclusion of the given class of signal models. For SRA and SRB, all of the bins are combined in multi-bin fit to maximise exclusion, and the same multi-bin fit strategy is used for SRD regions. SRC uses a shape fit on R_{ISR} over signal regions SRC1-5 to maximise the exclusion reach.
3. Model independent discovery fit: For each single signal region, nuisance parameters are fully constrained (both for control region and signal region), and a signal model independent hypothesis test is performed to consider the $p(\text{background-only}|\text{data})$. Additional regions are added here to this fit known as discovery regions, that have negligible impact on the expected exclusion limit (assuming no excess), but were designed to maximise the discovery significance up to the the $3 - \sigma$ level.

In total, 3 background only fits are performed, 3 sets of model-dependent exclusion fits per signal model - a single point point on the $(m_{\tilde{t}}, m_{\tilde{\chi}^0})$ plane- are performed stitching together the results of each fit, using the best CL_s value of the 3 fits to determine if a given signal is excluded. In addition, 16 model independent discovery fits are performed, corresponding to 3 SRA regions, 3 SRB regions, an SRA discovery region, 5 SRC regions, an SRC discovery region, and 3 SRD regions. Since the fits are subdivided for SRAB, SRC and SRD then stitched together as appropriate, the table outlines the allocation of control, signal and validation regions to each of the fits. Note that where a sample has more than one control region for a given fit category, a multi-bin fit on all such control regions in that category is applied (this is reflected in a single normalisation parameter for all control regions for that sample in that fit category). Each fit category is distinct and the exclusion fit limits are merged to form a final exclusion limit, but otherwise the fits are independently configured.

The table outlining the configuration of control, validation and signal regions is outlined in table [10.1](#).

Sample	Fit parameter	Norm. Region (SRAB Fit)	Norm. Region (SRC fit)	Norm. Region (SRD fit)
$t\bar{t}$	$\mu_{t\bar{t}}$	CRTAB	CRTC	CRTD1, CRDT2
Z+jets	μ_Z	CRZAB-T0, CRZAB-TT-TW	-	CRDZ0, CRDZ1, CRDZ1
W+jets	μ_W	CRWAB	-	CRDW0, CRDW1, CRDW2
Single top	μ_{ST}	CRSTAB	-	-
$t\bar{t} + Z$	$\mu_{t\bar{t}+Z}$	CR $t\bar{t} + Z$	-	-

Sample	Vaidation Region (SRAB fit)	Validation Region (SRC fit)	Validation Region (SRD fit)
$t\bar{t}$	VRTAB	VRTC	VRTD1, VRTD2
Z+jets	VRZA, VRZB-TT-TW, VRZB-T0	-	VRZD0, VRZD1, VRZD2

Fit type	Signal Regions (SRAB fit)	Signal Regions (SRC Fit) (SRC fit)	Signal regions (SRD fit)
Exclusion	SRA(TT/TW/T0), SRB(TT/TW/T0)	SRC[1-5]	SRD(0-2)
Discovery	SRA(TT/TW/T0), SRB(TT/TW/T0), SRATT-Disc	SRC[1-5], SRC-Disc	SRD(0-2)

Table 10.1: Tables outlining the control, validation and signal regions used in the SRA&B combined, SRC and SRD fits, mapping the control and validation regions to the respective signal regions. The exclusion fit regions are also implemented with the background-only fit.

10.2 Fit results

This section forms a summary of the results presented in [2], sub-divided into the background only fit, the discovery fit and the exclusion fit respectively.

Background only fit results

All results presented here are either directly from [2] or were produced as the contribution to the aforementioned publication. The yields tables are presented (Tables 10.2, 10.3, 10.4), which correspond to the background-fit results. These tables correspond to the standard model backgrounds post-fit in the signal region against data, allowing for observation of excesses or deficits of data in specific signal regions.

The blinded versions of the background only fits (where a dummy dataset is used in the signal region, equivalent to the sum of total standard model background) is used as a strong check on the stability of the fit. Fit stability means that normalisation parameters are not largely deviating from unity (except where this effect was explicitly known), and that the fitted values of other nuisance parameters are not substantially changed from their pre-fit values. Some diagnostic plots produced by the author to validate the fit stability will be discussed later.

The author produced the summary plots presented later and undertook the background only fits for the signal regions SRA and SRB, and in addition contributed to the production workflow (both in terms of the fit implementation, and in plotting) required to produce post-fit distributions for the control, signal and regions presented in this section.

Summary plots and tables

The following plots were also presented in [2] as the pre/post-fit summary plots for the background only fit. The summary plots are produced by the author, while the post-fit distributions in each control and signal region are from [2]. The summary plots for the signal regions are directly equivalent to the background-only yields tables for their respective regions, with the addition of a significance calculation equivalent to that used in the discovery fit. However, the values of significance calculated will vary slightly from the discovery fit, as the signal region itself is considered a “validation region”, so additional constraints from the signal region of the normalisations in the likelihood are not included, but for the cases used here, the difference is effectively negligible. A similar plot is produced for the validation regions in addition.

For the control region, a distinct plot is made, which reflects two different components of the fit: the top pad denoting the pre-fit yields of both the standard model background and data in the control region (reflecting the pre-fit normalisation), and the bottom pad denoting the fitted value of the normalisation parameter. Where the bins are merged between control regions for the bottom pad, this denotes the multi-bin fit for the given control region. The normalisation in the bottom pad corresponds to the fitted normalisation parameter $\hat{\mu}_{\text{bkg}}$ of the corresponding background (also in the same colour). Figure 10.1 outlines the results obtained.

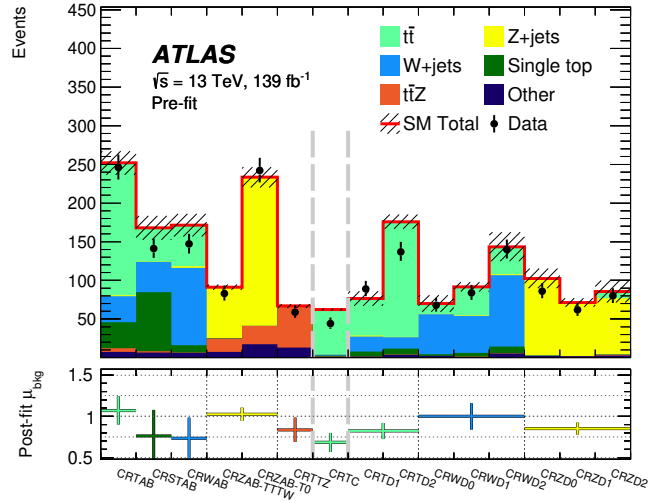


Figure 10.1: Summary plot of all the control regions used in the SRAB, SRC and SRD background-only fits respectively. The plot denotes the pre-fit normalisations and distributions used in the fit. The multi-bin fit implementation of control regions CRZAB, CRTD, CRDW and CRDZ can be seen as a multi-bin fit of the normalisations of the respective backgrounds.

In addition to the total normalisation, kinematic distributions of variables in the control regions are usually helpful to validate the efficacy of a fit. This is because some shape features in the control/validation regions may be incorrectly modelled when fitted (since the fit is on the total normalisation, not a particular distribution).

The following set of plots outlines a plot from each control region of a variable that would be obtained in a signal region, to check that the normalisation and/or relevant extrapolations are well understood (since in some cases, there are extrapolations over distributions between control and signal regions (figures 10.2-10.5). The unblinded signal region distributions are presented, and figures 10.6 and 10.7 then outline the one-bin summary plots for the validation and signal regions. Note the significance here may vary slightly from those calculated by discovery hypothesis tests, but this is due to the signal regions themselves being treated as validation regions to the fit, and hence systematics associated to these regions are not constrained

Tables 10.2-10.4 describe the final results of the background-only fit in terms of event-counts. Figure 10.9 then outlines the kinematic distributions of variables in the signal regions from the background-only fit.

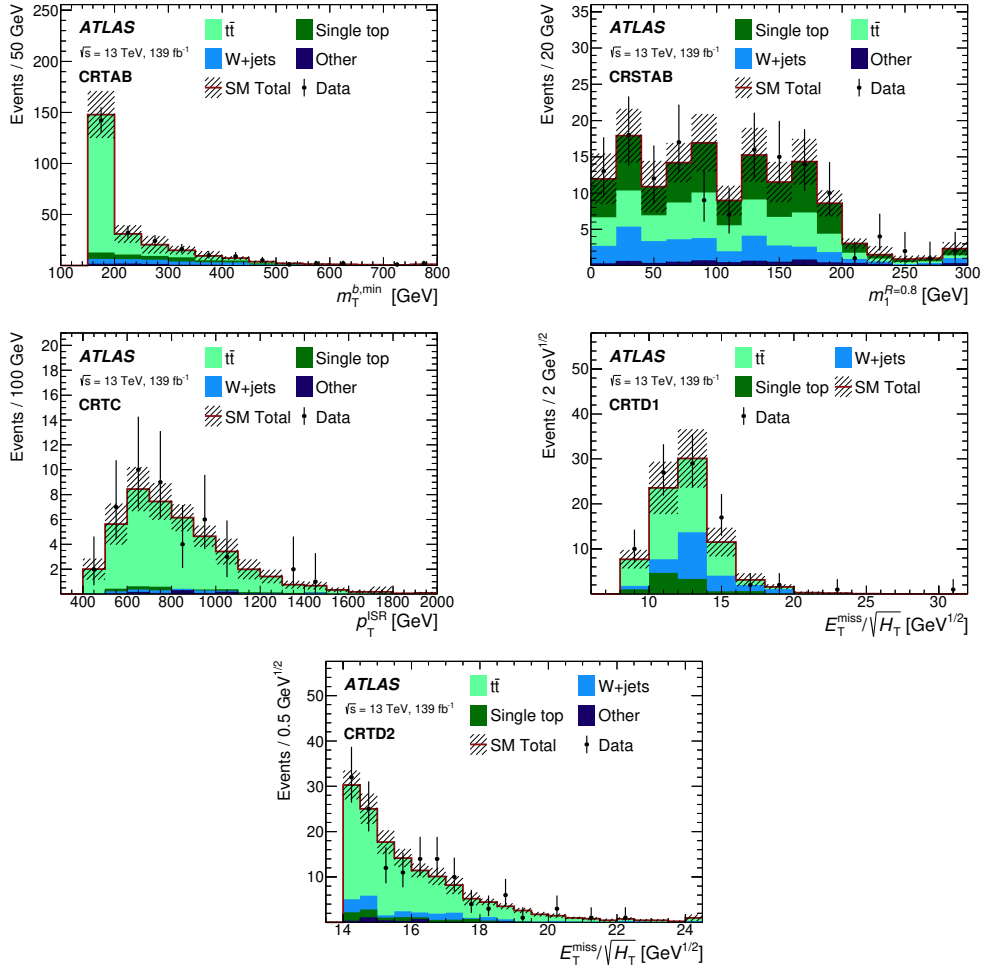


Figure 10.2: Post-fit distributions of the control regions CRTAB, CRSTAB, CRTC, CRTD1 and CRTD2 from [2].

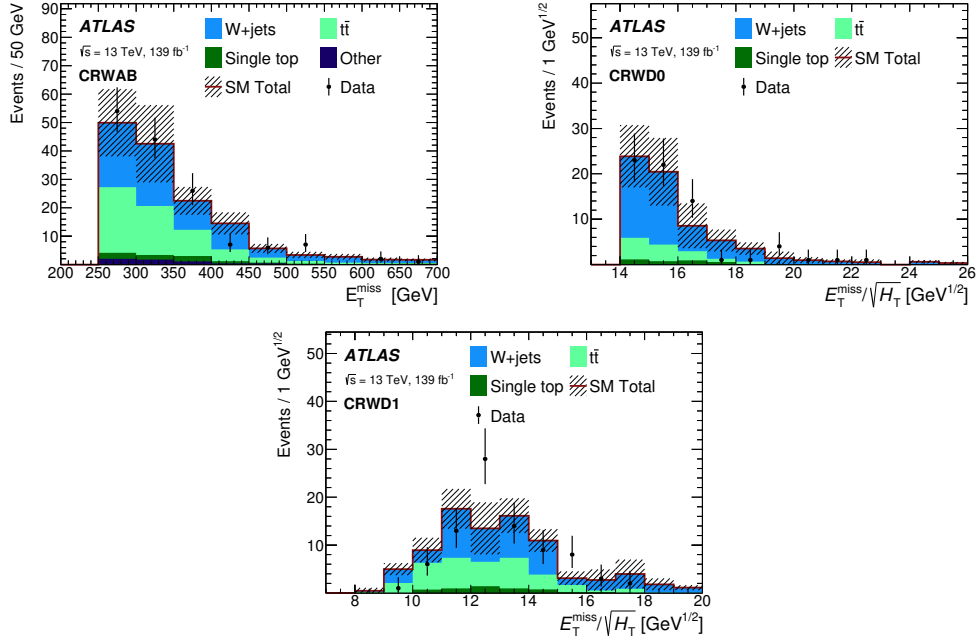


Figure 10.3: Post-fit distributions of the control regions CRWAB, CRWD0, CRWD1 and CRWD2 from [2].

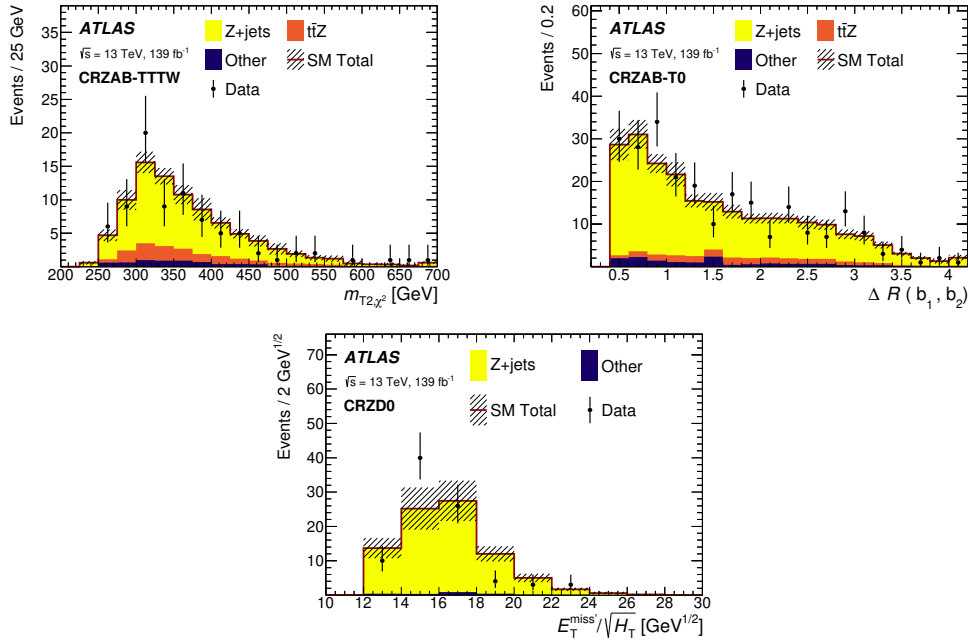


Figure 10.4: Post-fit distributions of the control regions CRZAB-T0, CRZAB-TT-TW and CRZD0 from [2].

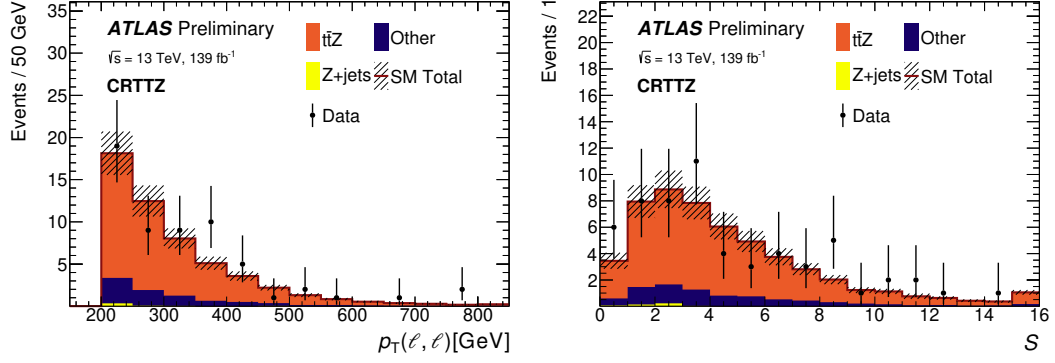


Figure 10.5: Post-fit distributions of the $p_T^{\ell\ell}$ and Object based E_T^{miss} sig. (S) in control region $\text{CR}t\bar{t} + Z\text{-AB}$ from [2].

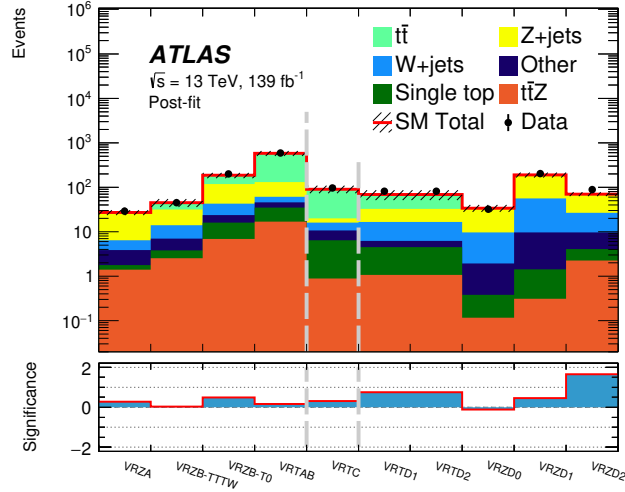


Figure 10.6: Summary plot of all the validation regions used in the SRAB, SRC and SRD background only fits. The significance is calculated using the formula derived in [114].

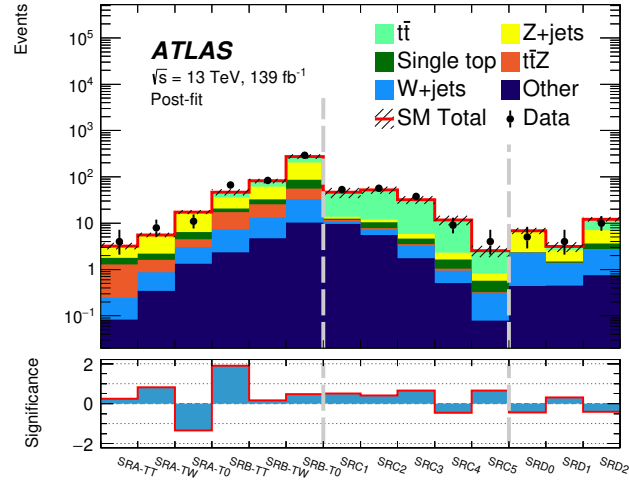


Figure 10.7: Summary plot of all the signal regions used in the SRAB, SRC and SRD background-only fits. The significance is calculated using the formula derived in [114], and includes the significance for under-fluctuations, which are dropped in the case of the model-independent discovery fit.

10.2. Fit results

	SRA-TT	SRA-TW	SRA-T0	SRB-TT	SRB-TW	SRB-T0
Observed	4	8	11	67	84	292
Fitted background events						
Total SM	3.22 ± 0.50	5.58 ± 0.66	17.3 ± 1.7	46.7 ± 6.7	81.7 ± 7.5	275 ± 24
$Z + \text{jets}$	1.35 ± 0.28	3.22 ± 0.42	10.5 ± 1.3	15.6 ± 3.3	28.7 ± 3.4	117 ± 14
Single top	0.50 ± 0.31	0.59 ± 0.29	1.95 ± 0.75	3.5 ± 1.2	7.0 ± 3.0	31 ± 15
$t\bar{t}$	0.08 ± 0.07	0.16 ± 0.10	0.36 ± 0.36	10.5 ± 5.3	20.5 ± 6.4	72 ± 19
$t\bar{t} + Z$	1.05 ± 0.29	0.74 ± 0.17	1.50 ± 0.34	9.9 ± 1.9	12.5 ± 2.5	22.7 ± 4.3
$W + \text{jets}$	0.16 ± 0.07	0.53 ± 0.20	1.71 ± 0.62	4.9 ± 1.9	8.5 ± 3.6	22.8 ± 8.7
Multijets	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
Other	0.08 ± 0.02	0.34 ± 0.24	1.32 ± 0.22	2.23 ± 0.35	4.4 ± 2.3	9.9 ± 1.3

Table 10.2: Signal Region Yields (post-fit) for signal regions SRA-TT, SRA-TW, SRA-T0, SRB-TT, SRB-TW and SRB-T0. The highlighted backgrounds ($t\bar{t}$, $W + \text{jets}$, $Z + \text{jets}$, $t\bar{t} + Z$, single top) correspond to those normalised in control regions in the fit.

	SRC1	SRC2	SRC3	SRC4	SRC5
Observed	53	57	38	9	4
Fitted background events					
Total SM	46 ± 12	52.2 ± 9.4	32.4 ± 6.5	11.8 ± 3.1	2.54 ± 0.69
$Z + \text{jets}$	1.17 ± 0.51	1.43 ± 0.77	1.17 ± 0.51	0.67 ± 0.32	0.24 ± 0.12
Single top	0.90 ± 0.34	2.69 ± 0.56	1.12 ± 0.73	0.61 ± 0.33	0.25 ± 0.05
$t\bar{t}$	32 ± 11	40.3 ± 9.2	26.6 ± 6.2	9.5 ± 2.7	1.73 ± 0.64
$t\bar{t} + Z$	0.74 ± 0.32	0.50 ± 0.31	0.24 ± 0.12	$0.09^{+0.11}_{-0.09}$	$0.01^{+0.03}_{-0.01}$
$W + \text{jets}$	1.28 ± 0.59	1.73 ± 0.62	1.51 ± 0.67	$0.44^{+0.62}_{-0.44}$	0.23 ± 0.09
Multijets	8.8 ± 2.9	4.2 ± 1.5	1.03 ± 0.43	0.15 ± 0.05	$0.03^{+0.03}_{-0.03}$

Table 10.3: Signal Region Yields (post-fit) for signal regions SRC1, SRC2, SRC3, SRC4 and SRC5. The highlighted $t\bar{t}$ background is normalised in the fit for SRC.

	SRD0	SRD1	SRD2
Observed	5	4	10
Fitted background events			
Total SM	6.9 ± 1.3	3.13 ± 0.98	12.2 ± 1.5
$Z + \text{jets}$	4.21 ± 0.82	1.07 ± 0.25	3.51 ± 0.57
Single top	$0.02^{+0.03}_{-0.02}$	$0.10^{+0.16}_{-0.10}$	0.84 ± 0.31
$t\bar{t}$	0.36 ± 0.29	0.56 ± 0.31	5.10 ± 0.97
$t\bar{t} + Z$	$0.02^{+0.04}_{-0.02}$	$0.01^{+0.01}_{-0.01}$	< 0.01
$W + \text{jets}$	1.87 ± 0.57	0.93 ± 0.60	2.02 ± 0.66
Multijets	< 0.01	< 0.01	0.24 ± 0.08
Other	0.44 ± 0.13	0.45 ± 0.21	0.51 ± 0.18

Table 10.4: Signal Region Yields (post-fit) for signal regions SRD0, SRD1 and SRD2. The highlighted backgrounds ($t\bar{t}$, $W + \text{jets}$, $Z + \text{jets}$) correspond to those normalised in control regions in the fit. Specifically for SRD0, no $t\bar{t}$ control region is configured.

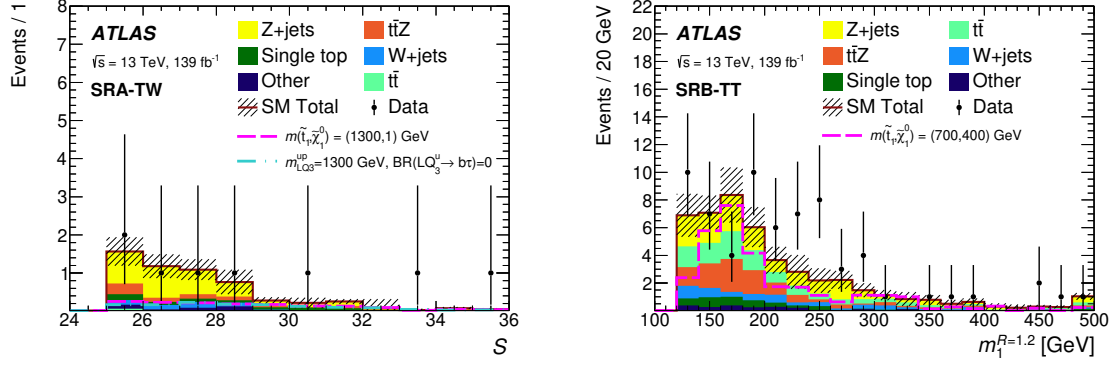


Figure 10.8: Post-fit distributions of the SRA and SRB signal regions presented in [2].

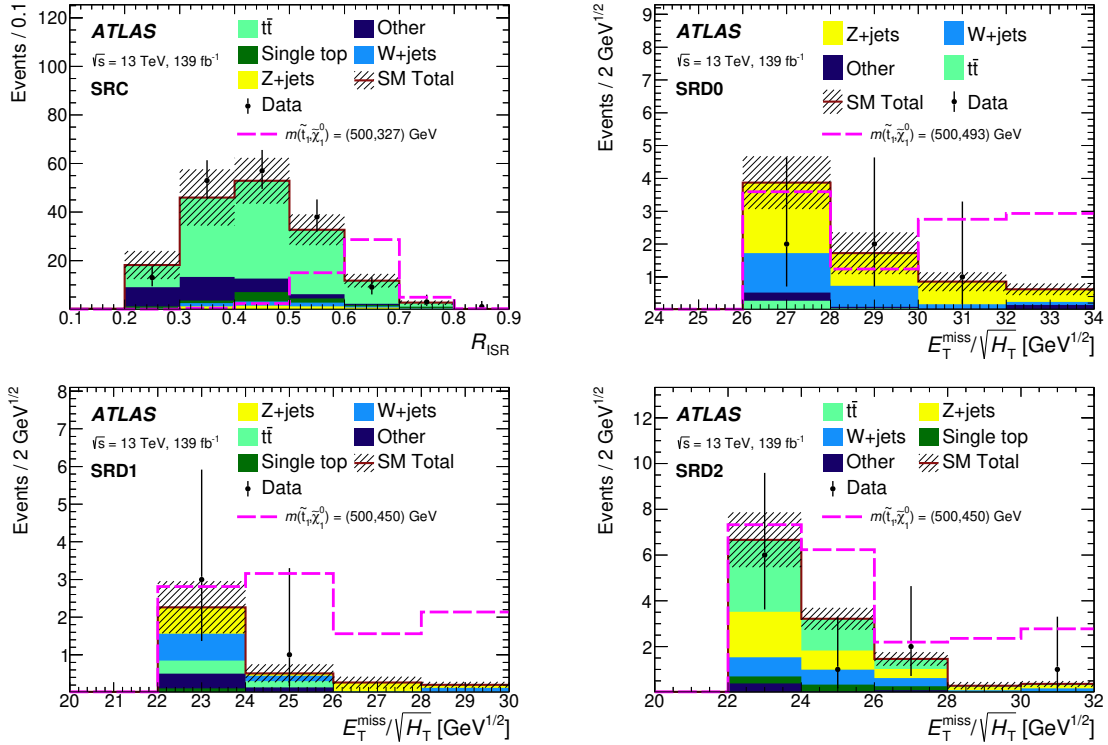


Figure 10.9: Post-fit distributions of the SRC and SRD signal regions presented in [2].

As can be observed in the tables and in the summary plots, no statistically significant excesses are observed in the background-only fit in any signal region, or across any binned distribution therein. Also, key variable distributions in the control regions are well modelled by the post-fit normalisations.

Additional diagnostic plots

The following plots were not presented as part of [2] but were produced concurrently with the fits that produced the summary plots, and were all produced by the author. The first outlines the impact of the fit (often referred to as *pulls*) on the nuisance parameters. The plot of these pulls presents [92] any deviations between the fitted estimate of a nuisance parameter $\hat{\alpha}$ and the pre-fit nuisance parameter estimate α . These pulls affect systematic uncertainties which are constrained in the fit due to normalisation (primarily those normalised in a control region as well as in the signal region), and can reflect either how the central value and up or down variations of the nuisance parameter are constrained by the fit. In order to demonstrate the stability of a possible fit, the pulls are required to be at most small deviations from their pre-fit values, and the uncertainty of the central value is expected to be consistent with 1. If a parameter has an uncertainty larger than 1 after fitting, it is said to be “under-constrained”, while an uncertainty in the pull of less than 1 is said to be constrained (the uncertainty associated to a given NP is reduced in the fit). This plot is particularly useful in determining misconfigured systematics in the fit and determining the stability of constraints on fit parameters. The pull-plot for the background-only fit signal regions SRA & SRB is presented in figure 10.10.

This plot involves several different parameters in a single plot which correspond to different components of the fit. In order to first understand the plot, these can be decomposed into several components (which relate to those discussed in 5.2):

1. μ_{sig} : the MLE of the parameter of interest is presented. However this is not used in the background-only fit (since it is by definition 0). An equivalent plot can be made with this parameter of interest for an exclusion fit.
2. μ_{bkg} : post-fit values (MLE) of the normalisation parameters of the standard model backgrounds and their uncertainties.
3. $\gamma_{\text{region-name}}$: statistical uncertainties in the given control region bin.
4. $\alpha_{\text{syst-name}}$: the MLE of a given parametrised systematic (with a central value plotted).

These systematics fall into several groups, corresponding to experimental or theoretical systematics already discussed in section 9.7, but these can be further sub-divided into several nuisance parameter groups:

- $\alpha_{\text{EL_<X>}}$: Uncertainties associated with the electron identification, isolation, reconstruction, charge identification, and electron trigger efficiencies in regions where the electron triggers are used

- $\alpha_{\text{FT}_{<X>}}$: Uncertainties associated with heavy flavour tagging (b-jets), either through the estimation of the b/c fractions in MV2c10 or any extrapolation over b-jet p_T .
- $\alpha_{\text{JET}_{<X>}}$: Uncertainties associated with calibration, jet energy scale estimation ($\alpha_{\text{Grouped_NP_X}}$), jet energy resolution($<X>_{\text{JER}}$), JVT efficiency.
- $\alpha_{\text{MET}_{<X>}}$: Uncertainties in the soft track resolution and energy scale measurements
- $\alpha_{\text{MUON}_{<X>}}$: Uncertainties associated with Muon reconstruction/bad muon rejection, identification and trigger efficiencies (where the muon trigger is used) in tracking, vertexing and momentum calculations.
- α_{PRW} : Uncertainties associated with correcting the pile-up profile of simulated samples with actual datasets.
- The remaining uncertainties are theory uncertainties as discussed in more detail in 9.7.
- Any uncertainties that are explicitly applied only in the signal regions ($t\bar{t} + Z$ theory variations) are not seen here, because the background-only fit treats the signal regions as validation regions, so no constraining is applied to these systematics.

Figure 10.10 indicates that the deviations on the nuisance parameters due to the fit is quite small, and that only a small amount of constraining is observed. Since the effects are small and the normalisations are approximately consistent with unity, the fits are stable.

In the configuration of the fit, some nuisance parameters may end up being correlated, and as such the correlations between the fitted values of different nuisance parameters can be plotted. This correlation occurs because the maximal likelihood estimate of a given parameter may be intrinsically correlated to another (for example where a background is shared between multiple control regions). It can be noted that for signal regions A and B, there is some coupling of the $t\bar{t}$, $W + \text{jets}$ and single top systematics, because of the substantial contamination of $t\bar{t}$ in the control regions for $W + \text{jets}$ and single top, and likewise for the other two samples. This can be readily seen from the control region summary plot in figure 10.1. Correlation plots are used in the background-only and exclusion fits both when blinded (whereby the data is a dummy dataset equivalent to the total background contribution) and un-blinded as a check of fit stability and to minimise spurious correlations between uncertainties. These spurious correlations typically may imply mis-configurations in the fit, and correlation is ideally desired to be negligible where possible. A plot outlining the correlations of uncertainties for the signal regions SRA and SRB fit is found in figure 10.11.

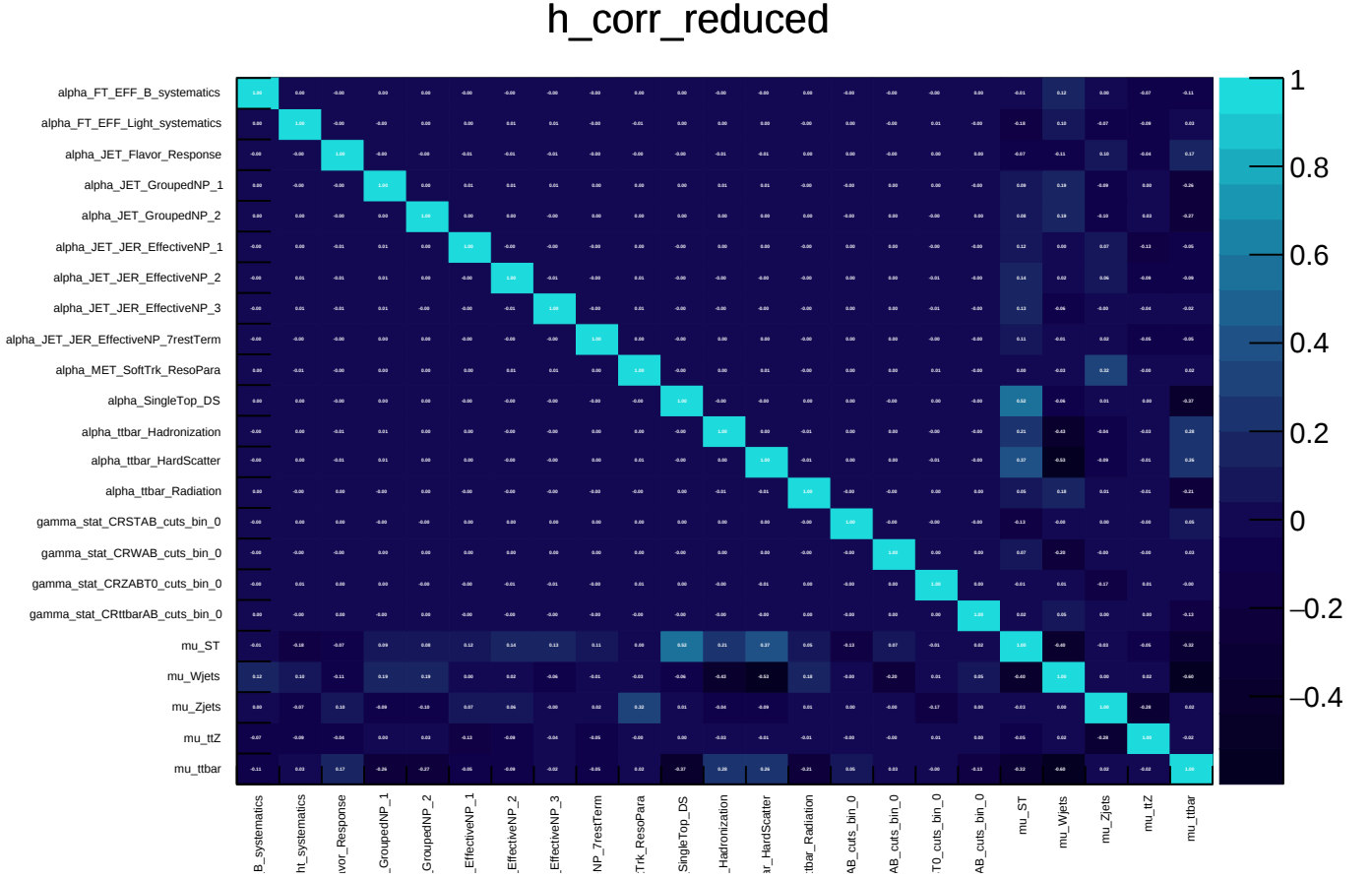


Figure 10.11: Correlation matrix for the fitted nuisance parameters in a background-only fit for signal regions A and B. This is a summary of the largest systematic correlations, dropping the correlations with coefficients below some threshold value for the purposes of clarity. The definitions of the systematics is the same as that for figure 10.10.

As seen in figure 10.11, the largest correlation is between the single-top normalisation and the theory uncertainties associated to single top or $t\bar{t}$. This is expected since the single top diagram subtraction/removal uncertainty corresponds to a comparison with an alternate sample, so an impact in the normalisation of the nominal sample will directly impact the magnitude of the uncertainty calculated between the nominal sample (diagram subtraction scheme) and the alternate diagram reduction scheme. Single-top normalisations have some correlation with the systematics for $t\bar{t}$ due to the substantial contamination of $t\bar{t}$ in the control region. However overall, there are no substantial correlations in any of the other systematics, and with this and the pull plot, it can be noted that this fit is stable (no substantial changes induced from the fits).

Discovery fit results

The following fit results outline the p -values and significances for each discovery fit in each respective signal region as defined in 10.1. The p -values are converted to significances through the following formula [92]

Region	Observed p -value	Observed significance (σ)
SRA-TT	0.34	0.40
SRA-TW	0.18	0.92
SRA-T0	0.50	0.00
SRA-TT-Disc	0.50	0.00
SRB-TT	0.03	1.87
SRB-TW	0.42	0.19
SRB-T0	0.30	0.53
SRC-1	0.49	0.01
SRC-2	0.22	0.77
SRC-3	0.23	0.75
SRC-4	0.5	0.00
SRC-5	0.22	0.78
SRC-Disc	0.49	0.02
SRD0	0.50	0.00
SRD1	0.34	0.40
SRD2	0.50	0.00

Table 10.5: Observed p -values and significances for each signal region in [2] calculated using HistFitter “UpperLimitTable” [92], from [2]. For any regions with under fluctuations of data against background, the p -value is defined by construction to be 0.5, and the significance is set to 0.

Table 10.5 notes that there is an excess of 1.87σ in SRB-TT only, but this is not statistically significant. An under-fluctuation of data against the SM background is also observed in signal region SRA-T0, but this is not statistically significant either. Thus it can be concluded that there are no statistically significant excesses or under-fluctuations in data in any signal region bin, including the discovery regions. Thus, the signal regions can be used to provide observed exclusion limits upon classes of simplified signal models.

Exclusion fit results

This section outlines the exclusion fit results for each signal point on the $(m_{\tilde{t}}, m_{\tilde{\chi}^0})$ grid, plotted as an exclusion limit based on expected significance (the expected exclusion) and that based on observed significance (the observed exclusion limit). The exclusion limit contours corresponding to the respective signal region groupings respectively have been merged to form a single limit. Since the exclusion limit is formed from the CL_s values of signal points in a grid, interpolation is used in HistFitter to construct the contour.

In order to interpret this exclusion limit, two limits are produced: an expected CL_s limit (assuming the total SM background is the actual event yield), and an observed CL_s limit. Typically, the exclusion limit is applied at a 95% confidence level.

The uncertainty on the expected limit includes all experimental systematics for both the standard model background normalisation and the experimental systematics associated to the signals. The uncertainty on the observed limit is due the variation on the cross section and other theoretical variations of the respective signal points in the signal grid. Figure 10.12 outlines the results of the exclusion fits over the $(m_{\tilde{t}}, m_{\tilde{\chi}^0})$ plane.

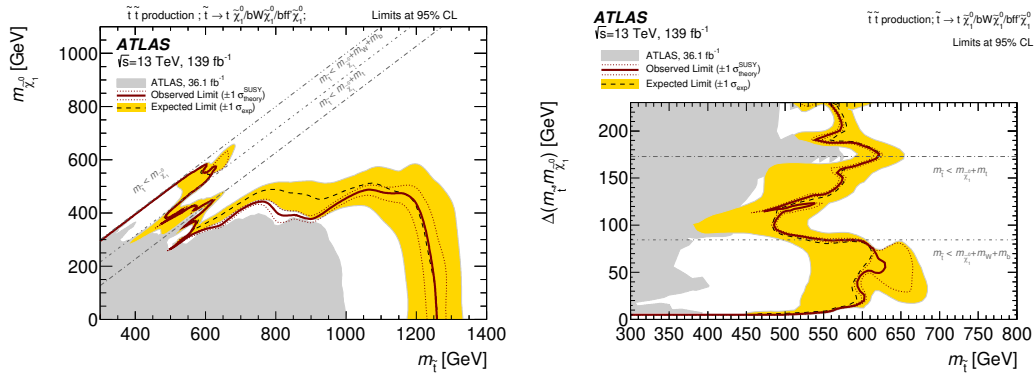


Figure 10.12: Exclusion limits (both expected and observed) in both the $(m_{\tilde{t}}, m_{\tilde{\chi}^0})$ plane and the equivalent $m_{\tilde{t}}, \Delta(m_{\tilde{t}}, m_{\tilde{\chi}^0})$ plane presented in [2].

10.3 Discussion

As observed in the previous sections, no significant excesses have been observed in any signal region (the only excess observed is in signal region SRA-T0, and has a significance of 1.87σ). Exclusion limits are applied up to around $m_{\tilde{t}} = 1.3$ TeV, and $(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (600, 400)$ GeV along the compressed diagonal. Four body signal model exclusion limits have been obtained up to $(m_{\tilde{t}}, m_{\tilde{\chi}^0}) \approx (610, 600)$ GeV.

Brief discussion of notable features in the exclusion limit

Some notable features must be discussed when constructing the limit, namely that each separate fit (SRAB, SRC, SRD) is used to constrain all signal models, using the best expected

or observed (i.e. lowest CL_s in the given region) to produce the final plot. However, this can procedure can produce discontinuities where two fits overlap in excluded signal models due to interpolation between excluded points, which can be seen at $\Delta(m_{\tilde{t}}, m_{\tilde{\chi}^0}) - 125$ GeV. We expect that this feature is merely an artificial effect of the fit structure and would be resolved with additional signal points in this region or a fit strategy which encompassed signals in this particular region of interest.

The second feature of interest is the feature at $\Delta(m_{\tilde{t}}, m_{\tilde{\chi}^0})$, which arises from the best expected exclusion limits from the SRD fit corresponding to signal regions SRD0, SRD1, SRD2. The effect of the large expected exclusion and uncertainty bands is driven by the multi-bin fits of signal regions SRD0, SRD1 and SRD2 to produce the exclusion limit, which for a given signal point have quite substantially different numbers of signal events for the same signal model. A future analysis could rectify this issue with additional signal points in this region, but this is observed to be an artefact of the multi-bin fit configuration that a future analysis may resolve with increased luminosity.

10.3.1 Context of results in terms of other ATLAS measurements

Since the analysis considered here is corresponding to a single final state (jets, E_T^{miss} and no e/μ), a complete picture of possible models excluded must include all of the possible final states. Normally this is presented in a combined exclusion curve across the entire signal grid (assuming no analysis shows a statistically significant excess). The relevant publications corresponding to the latest 139fb^{-1} searches are [2, 22, 99], which at time of writing were the latest results in stop searches.

The current status of the ATLAS stop searches (as of writing) can be summarised in a single plot combining all of the latest search results, as outlined in figure 10.13.

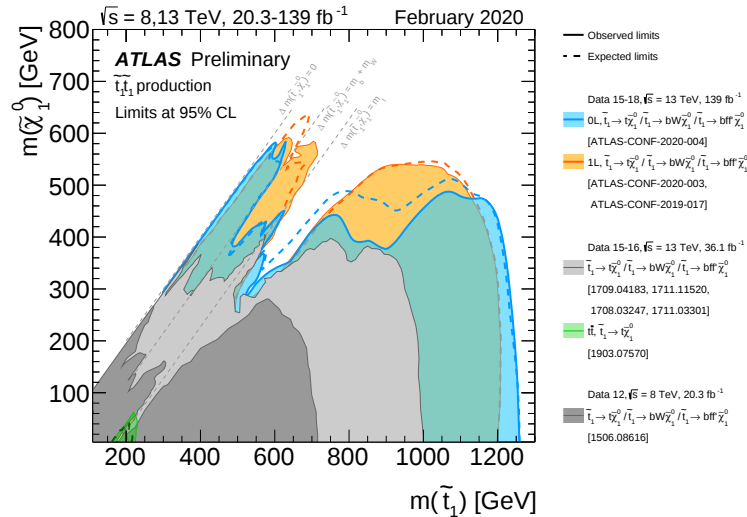


Figure 10.13: Exclusion limits (both expected and observed) in the $(m_{\tilde{t}}, m_{\tilde{\chi}^0})$ plane in the context of all similar results by ATLAS as of May 2020. [115]

A note of caution must be made in the comparisons of these combination plots, namely in that although the signal models and signal grids tested in these analyses are the same across

teams, the object definitions and optimisations used therein will differ, as well as the signal region design itself. This will lead to overlapping exclusion limits across different search channels, which are conditional on the assumptions undertaken in a given analysis channel.

10.3.2 Reflection: results in light of CMS searches

Since any result produced by ATLAS should naturally be contrasted against a result observed by CMS, the author will outline the equivalent signal model search result from the CMS experiment. CMS uses an alternative method to maximise the expected exclusion: a large number of binned signal regions and multi-bin fits.

The full details of their background estimation techniques can be found in [116], but will be somewhat analogous in the aims and objectives as the ATLAS analyses presented in this thesis. Post-fit distributions in a given variable are then one dimensional projections of these binned fits in the variable of interest (usually the variable which was selected for the binning of the signal region).

As a summary, the CMS signal regions require as inputs:

- $N_{\text{jets}} \geq 2$, for jets with $|\eta| < 2.4$.
- $H_T > 300$ GeV, using jets again with $|\eta| < 2.4$.
- $H_T^{\text{miss}} > 300$ GeV: This is the negative magnitude of the vector sum of jet p_T for jets with $|\eta| < 5$.
- $H_T^{\text{miss}} < H_T$ to remove jet mismeasurement backgrounds.
- No isolated e/μ with $p_T > 10$ GeV.
- No isolated track with $m_T < 100$ GeV, and $p_T > 10(5)$ GeV (the lower p_T selection is used if the track is associated to an e/μ).
- No isolated photons
- Angular separation between the direction of the H_T^{miss} and any other jet reconstructed with $|\eta| < 2.4$. This is analogous to use of $\min[\Delta\phi(j_{1-4}, E_T^{\text{miss}})]$ selections in ATLAS for multi-jet mitigation.

They are then binned in a four dimensional set of regions corresponding to either: N_{jets} , $N_{\text{b-jet}}$, H_T or H_T^{miss} , corresponding to 174 signal regions.

The advantage of this multi-bin approach is it is somewhat general in its approach to excluding a variety of model classes, as [116] covers a number of different interpretations using the less specific-model dependent final state. However, the tradeoff is that some excesses in the dataset will be solely due to statistical fluctuation. The 174 signal regions are plotted pre-fit in figure 10.14, and for the purposes of brevity, the underlying yields tables, significances and upper limits for the model-independent fits can be found in [116]. As no significant excesses were observed, the exclusion limits from the CMS scans of the signal model corresponding to $\tilde{t}_1 \rightarrow t + \tilde{\chi}_1^0$ can be noted in figure. This follows the same prescription

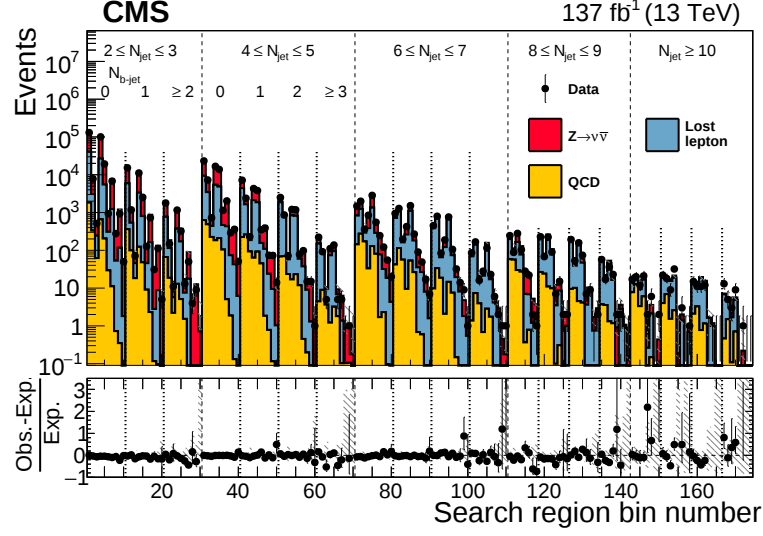


Figure 10.14: CMS signal regions (pre-fit) covering the hadronic supersymmetry searches, from [116]. These 174 bins cover the four dimensional binning scheme mentioned above.

as the ATLAS analysis in terms of construction of the observed and expected exclusion limits, however in addition the CL_s upper limit on the cross section is plotted. ATLAS does this equivalently, but normally this is provided in auxiliary material corresponding to each individual signal point used to produce the exclusion limit (see here).

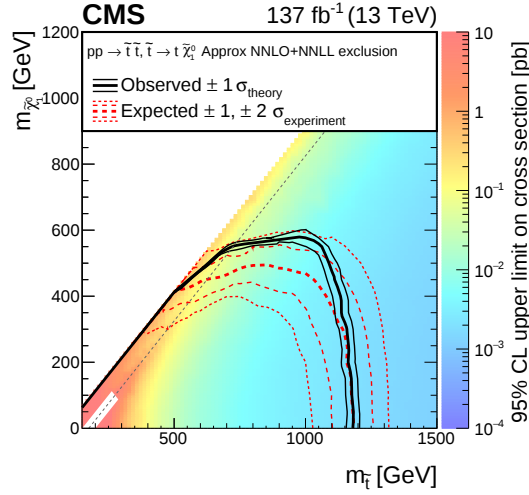


Figure 10.15: CMS exclusion limit covering the hadronic stop decay search covered in this thesis, from [116]. In the CMS paper, these are denoted as T2tt models.

The CMS exclusion limit extends in the boosted region to ~ 1.2 TeV, which is comparable to the ATLAS limit, but in the bulk region the exclusion reach is maximal for the signal point $(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (1000, 600)$ GeV, which is not excluded in the equivalent ATLAS analysis. However, the CMS analysis does not include four body decay models in the exclusion limit. Overall, it can be summarised that both results are comparable given the substantial

differences in detector design and analysis design, with neither collaboration presenting any statistically significant excess in the search for supersymmetry.

Summary

In this chapter, the fit configuration and results for [2] were presented. In addition a validation of the stability of the fit for signal regions SRA and SRB were presented. Since the results show no significant excess in any signal region (the largest excess in SRB-TT having a significance of 1.87σ), exclusion limits are applied up to 1.3 TeV in the boosted plane, up to $(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (600, 450)$ GeV in the compressed region, and $(m_{\tilde{t}}, m_{\tilde{\chi}^0}) = (700, 600)$ GeV in the four body region. These results are also comparable with similar results from CMS, although some regions of phase space are excluded by CMS that are not excluded by ATLAS, and vice-versa.

Concluding remarks

11.1 Final summary

This thesis has covered the search for the supersymmetric partner of the top quark in the all-hadronic final state, discussing both the analyses published using an integrated luminosity of 36.1fb^{-1} (covering LHC run years 2015-2016), and the new analysis for 139fb^{-1} (covering LHC run years 2015-2018). The largest excesses had been observed in the 2015-18 signal region SRB-TT with a significance of 1.9σ , and the 2015-16 signal region SRB-T0 with a significance of 1.2σ , neither of which are statistically significant.

The 95% confidence level exclusion limits for simplified stop models $\tilde{t}_1 \rightarrow t + \tilde{\chi}_1^0$ were extended by 260 GeV for signal models with large Δm , extended by another 50 – 100 GeV along the $\Delta(m_{\tilde{t}}, m_{\tilde{\chi}_1^0}) = m_t$ diagonal. Exclusion limits on four body signal models have been applied up to (610, 600) GeV with the implementation of the new signal model strategy for four body regions using low p_T track-jets and soft b -tagging.

At the time of writing of this thesis, supersymmetry has not been observed in any final state (either those presented here, or elsewhere), providing strong constraints on cross sections of the classes of simplified models considered by both the ATLAS and CMS collaborations. However, such limits must be considered carefully given the underpinning assumptions made in the construction of these simplified models (branching fractions, R-parity, coupling strength, mass parameters and spectrum). Any future work such as scans over parameter spaces of the PMSSM would make good use of these simplified model searches to apply more generally-applicable exclusion limits. These re-interpretation/extension studies are greatly aided by automated analysis pipelines (see section 11.2.1 for a brief discussion of this).

11.2 Future work

LHC data taking run 3, expected to commence in 2021, is expected to use principally the same instantaneous luminosity and pile-up conditions as per the data taking years 2017-2018[117], with an expected total luminosity of around 300fb^{-1} . Since there will be no new data obtained until likely late 2021, analyses will be poised to investigate more complex analysis strategies such as MVA driven analysis and approaches taking inspiration from

the either phenomenology or algorithms which push the frontier of event reconstruction. A notable improvement will be the use of particle flow (currently already used by the CMS collaboration)[64][118], which will match and merge track and calorimeter information in jets or other objects, improving the reconstruction efficiencies and resolution for low p_T objects (particularly jets with $p_T > 15$ GeV) as compared to the calorimeter topology based reconstruction as mentioned in this thesis. This will enable more studies into regions containing low p_T jets.

Improvements will also arrive from the choice of simulated samples available, covering improved NLO and NNLO corrections to Standard Model processes, modelling of various kinematic variables at generator level, and in addition improved detector simulations through improvements such as FastCaloSimV2.

In LHC run years 2027 onward, the LHC is expected operate in a state known as “High Luminosity LHC” mode, which greatly exceeds the original design luminosity, as compared to the 2018 run conditions which had instantaneous luminosities twice that of the design specification. This will induce a number of new features, such as a greatly increased pile-up (of around $\langle\mu\rangle > 100$) and a vastly increased dataset, still at a centre of mass energy of 13 TeV.

To avoid overloading the capacity of the WLCG CPUs in the endeavour, simulation components such as the FastChain handling of fully simulated hard scatter events combined with fast simulated pileup. This will ideally replace the current strategy of overlay pileup, removing the need for huge libraries of minimum bias ($\langle\mu\rangle = 1$) events overlaid with simulated events at random. In addition, further detail such as the integrated simulation framework will allow for more CPU utilisation on impactful physics processes in such conditions. However, at time of writing, this tool is not yet in production.

11.2.1 Analysis extensions: RECAST framework

Only a limited set of signal models can be at any time tested by the relevant analysis team (with the relevant assumptions therein). However, another analyst may wish to use the analysis results to apply exclusion limits to alternate classes of models that were not initially processed throughout the course of publication. This procedure is known as “re-casting” an analysis, and is generally where a published analysis is re-run with an alternate set of signal models. In the past, this can be somewhat a problem as researchers finish projects without sufficient future proofing of their work.

ATLAS uses a framework called “RECAST” [119] to produce an automated pipeline with a given input signal. This processes the input signal sample from the DAOD data format through to final exclusion fit results, with a fully automated single command pipeline. In order to ensure stability of the respective components and “future-proof” the analyses as published, each independent step is run on pre-compiled Docker [120] containers. The author contributed to the implementation of the RECAST pipeline for [2]. This workflow also has been demonstrated to work efficiently with Continuous Integration (CI) on web servers such as gitlab (which can fully automate the entire process from start to finish). The workflow for this analysis can be found in [this repository](#) (CERN auth. required).

Analysis combination

ATLAS subdivides the searches for the supersymmetric partner to the top quark into separate teams and publications depending on the targeted signatures (particularly under the number of leptons in the final state). Each of these differing analysis teams may choose different configurations of selections, object definitions etc for their respective publications. This means that exclusion limits combining the results of multiple analyses should be considered carefully, since they vary some of the underlying assumptions. During the analysis preparation stages, the object definitions are somewhat harmonised, such that at minimum no analysis channel contains a control/validation region which overlaps with a signal region for any other channel. However, some object definitions will also vary between channels, due to differing object optimisations in the respective channels, and the effects of these must be taken into account when considering the combination. Such ambiguities can arise in definitions of objects such as the MV2c10 b-jet efficiency working point selection, or the definitions of reconstructed electrons/muons.

However, despite the requirement for orthogonality between control and signal regions, there may be some overlap in either control regions or signal regions respectively of differing analysis channels, which must be taken into account/re-binned. This is intended to avoid overlaps when performing the exclusion fits (to minimise overconstraints) covering all analysis channels, to produce a final exclusion limit based on the combined fits of all channels. These combined limits are beneficial both for the simplified model searches, MSSM scans and even other signal models. The extension surveys involving merging the analysis strategies of the respective analysis channels, to see if a more stringent model-dependent exclusion limit can be obtained. This combination, however, is often a trade-off since different analysis strategies may rely on different physics object definitions and definitions of the control regions across analyses may overlap in a non-trivial manner. A combination approach however is normally considered only after the conclusion of single channel simplified model searches have been obtained, because the individual channel approach is designed for the observation of signal excesses, while the combination approach is principally designed for exclusion.

ATLAS specific configurations

Luminosity and configurations

The dataset for chapter 6 was collected between LHC run periods 2015-2016 was measured as $36.1 \pm \text{fb}^{-1}$. The dataset for chapter 9 was collected between LHC run periods 2015-2018 with a total integrated luminosity of $139.0 \pm 2.4 \text{fb}^{-1}$ [2][36] measured by Van-der-Meers luminosity scan.

A.1 Simulated samples

The parton distribution function set used for signal samples is the NNPDF2.3LO for SUSY signals, with the A14 tune of the underlying event and parton shower. The matching of the matrix element to the parton shower was undertaken using the CKKW-L [121] prescription. All signal cross sections are calculated to approximate NNLO in α_{QCD} , with the matching scale set to one quarter of the top squark mass. The resummation of soft gluon emission in the parton shower was included at next-to-next-to-leading-logarithm accuracy (approximate NNLO+NNLL). The top squark mixing was set to be maximal (which leads to the lightest possible $\tilde{t}_1$). Similar prescriptions were applied to the backgrounds, as is discussed in A.1.

Sample	Matrix element Generator	Parton Shower shower	Optimisation scheme (UE)	Cross section calculation
$W + \text{jets}/Z + \text{jets}$	Sherpa 2.2.1		Default	NNLO
$t\bar{t} + Z/t\bar{t} + W$	aMC@NLO 2.3.3	Pythia 8.210	A14 tune	NLO
$t\bar{t}$	Powheg-Box-v2	Pythia 8.330	A14 tune	NNLO+NNLL
Single top	Powheg-Box-v2	Pythia 8.230	A14 tune	NNLO+NNLL
Diboson	Sherpa 2.2.1-2.2.2		Default	NLO
$t\bar{t} + H$	aMC@NLO 2.2.3	Pythia 8.230	A14 tune	NLO
$tWZ, t\bar{t}Z$	aMC@NLO 2.3.3	Pythia 8.21/8.230	A14 tune	NLO
Signals	aMC@NLO 2.6.2	Pythia 8.230 (PS) EvtGen 1.6.0 (b/c-hadron decay)	A14 tune	NNLO+NNLL

Table A.1: Simulated background and signal samples used in the 139.0fb^{-1} analysis in [2].

Group theory

This section outlines the fundamentals of groups and symmetries, in relation to QED and QCD, with some relevant definitions. The fundamental structure of a group relies on containing a set of elements, and an operation (in the cases relevant here, it will be matrix multiplication $\times$), and has the four axiomatic requirements for a group G [122]:

1. Closure: for two elements $(a, b) \in G$, the operation $a \times b \in G$
2. Associativity: $\forall(a, b, c) \in G, a \times (b \times c) = (a \times b) \times c$.
3. An identity element exists (in all relevant cases this is the identity matrix $\mathbb{I}$) such that $a \times \mathbb{I} = \mathbb{I} \times a = a \forall a \in G$.
4. Existence of inverse: $\forall a \in G, \exists a^{-1} \in G : a \times a^{-1} = \mathbb{I}$.

Matrix multiplication guarantees almost all of the above properties by default (except the existence of the inverse), but all the matrices used will be considered invertible. The class of groups we will consider will be some sort of symmetry group, transforming:

$$T(x) |x\rangle = |y_0\rangle, \quad (\text{B.1})$$

Considering the transformation $T(x)$ as an infinitesimal transformation from $x \rightarrow x + \delta x$:

$$T(\delta x) = \mathbb{I} + iG\delta x, \quad (\text{B.2})$$

for some object G (which will be determined later, and will take the form of a real scalar), Taylor expanding around x to first order:

$$T(x + \delta x) = T(x) + dx \frac{dT(x)}{dx}, \quad (\text{B.3})$$

and by the group operation:

$$T(x + \delta x) = T(x)T(\delta x). \quad (\text{B.4})$$

Using the three identities, we can find that

$$T(x) = \exp(iGx). \quad (\text{B.5})$$

This group corresponds to the set of complex numbers $U(1)$, depending on the value of $G \in \mathbb{R}$. A group G can be defined by a mathematical construct known as a “Lie algebra”, such that for some elements $T_i \in G$ (T can be a matrix), the group can be defined by a basis of generators [5, 122]:

$$\begin{aligned} T_i \forall i \in \mathbb{N} &\in (0, \dim(G)) \\ \{e^{iT_1x}, e^{iT_2x}, e^{iT_3x}, \dots\} &\in G. \end{aligned} \quad (\text{B.6})$$

These generators T_i are then defined by the relation:

$$[T_i, T_j] = f_{ijk}T_k; \quad [T_i, T_j] = T_iT_j - T_jT_i. \quad (\text{B.7})$$

The group axiom requires that the f_{ijk} is a number, and these are referred to as the “structure constants”. The explicit value of these structure constants depends on the group in question, but if $\forall \{i, j, k\}; f_{ijk} = 0$ (namely the group generators commute), the group is referred to as *abelian*, or likewise *non-abelian* otherwise.

B.0.1 Groups and Symmetries: QED and QCD

Three important groups that will be discussed are $U(1)$, $SU(2)$ and $SU(3)$, which correspond to “unitary group in 1-dimension” (i.e. $z \in \mathbb{C}; |z| = 1$) and “special unitary groups in two or three dimensions” (i.e. $UU^\dagger = \mathbb{I}_{2,3} \cap \det(U) = 1$). The number of elements in each group corresponds to $N_{\text{dim}}^2 - 1$, so for $SU(2)$ this corresponds to three elements, and $SU(3)$ has 8 elements. The minimum dimension (as in number of matrix entries) set of elements that can define the group is referred to as the “fundamental representation”, but this basis is non-unique for $SU(n)$ under unitary transformations, so a choice of matrices is available.

For $SU(2)$, a readily available representation is the Pauli Matrices [4] (up to a normalisation):

$$\sigma_1 = \frac{i}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \frac{i}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \frac{i}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (\text{B.8})$$

which satisfy all of the requirements of $SU(2)$ (as does their matrix exponential), and using their commutation relations, the structure constants are equivalent to the components of the totally antisymmetric levi-civita tensor ϵ_{ijk} .

For $SU(3)$, a different representation must be used, namely the Gell-Mann matrices [4]:

$$\lambda_1 = \frac{1}{2} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_2 = \frac{1}{2} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_3 = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_4 = \frac{1}{2} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad (\text{B.9})$$

$$\lambda_5 = \frac{1}{2} \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \lambda_6 = \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \lambda_7 = \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda_8 = \frac{1}{2\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

From these Gell-Mann matrices, it can be observed that the structure constants are totally-antisymmetric under interchange of indices, and have the following values (these will be stated without proof, but these can be proven from the matrix multiplication):

$$f_{123} = 1, \quad f_{147} = -f_{156} = f_{246} = f_{257} = f_{345} = -f_{367} = \frac{1}{2}, \quad f_{458} = f_{678} = \frac{\sqrt{3}}{2}. \quad (\text{B.10})$$

B.0.1.1 Covariant derivative D_μ and propagators

Following the methodology of [4], we define the directional derivative in some unit direction n^μ as:

$$n^\mu \partial_\mu \psi = \lim_{\epsilon \rightarrow 0} \frac{\psi(x^\alpha + \epsilon n^\alpha) - \psi(x^\alpha)}{\epsilon}, \quad (\text{B.11})$$

and we define a quantity that transforms the field from one point to another ($U(y, x)$), namely

$$U(y, x) \rightarrow e^{i\alpha(y)} U(y, x) e^{-i\alpha(x)}; U(y, y) = 1, U(y, x) = e^{if(x, y)} (\text{U is a pure phase}). \quad (\text{B.12})$$

The directional covariant derivative D_μ can be defined with the above transformation:

$$n^\mu D_\mu \psi = \lim_{\epsilon \rightarrow 0} \frac{\psi(x + \epsilon n) - U(x + \epsilon n, x) \psi(x)}{\epsilon}, \quad (\text{B.13})$$

where $U(x + \epsilon n)$ is defined (by its infinitesimal expansion):

$$U(x + \epsilon n) = 1 - i\epsilon n^\mu A_\mu(x) + O(\epsilon^2). \quad (\text{B.14})$$

The exact form of $A_\mu(x)$ here is arbitrary (in general referred to as a “connection”), and the choice of the constants is also arbitrary (but chosen conveniently to match those in QED). Calculating the limit then leaves the definition:

$$n^\mu D_\mu \psi(x) = n^\mu \partial_\mu \psi(x) + n^\mu (ie A_\mu \psi(x)); D_\mu \psi(x) = \partial_\mu \psi(x) + ie A_\mu \psi(x), \quad (\text{B.15})$$

This field transforms under the U(1) local transformation as:

$$A_\mu(x) \rightarrow A_\mu - \frac{1}{e} \partial_\mu \alpha(x). \quad (\text{B.16})$$

Using the transformation of both $A_\mu(x)$ and $\psi(x)$ under the local U(1) transformation, it can be noted that:

$$D_\mu \psi(x) = e^{i\alpha(x)} D_\mu \psi(x), \quad (\text{B.17})$$

namely the covariant derivative ∂_μ transforms under the U(1) symmetry in the same way as for the field ψ .

In addition, it can be shown that the field strength tensor $F^{\mu\nu}$ is also invariant under the local U(1) transformation. This is because the field strength tensor can be defined as:

$$[D_\mu, D_\nu] \psi(x) := ie F_{\mu\nu} \psi(x) = ie (\partial_\mu A_\nu - \partial_\nu A_\mu) \psi(x). \quad (\text{B.18})$$

A similar logic holds for non-abelian gauge fields, instead defining the $U(y, x)$ as a matrix (2x2 for SU(2), 3x3 for SU(3) dependent on the choice of representation), such that it transforms under:

$$U(y, x) \rightarrow V(y) U(y, x) V^\dagger(x). \quad (\text{B.19})$$

This $U(y, x)$ can be determined to be:

$$U(x + \epsilon n, x) = \mathbb{I} - i\epsilon n^\mu A_\mu^i T^a + O(\epsilon^2), \quad (\text{B.20})$$

where g is an arbitrary coupling constant, and the roman indices correspond to the indices over group elements, while greek indices correspond to Minkowski spacetime. Thus the covariant derivative is:

$$D_\mu \psi(x) = \left(\partial_\mu - ig A_\mu^a T^a \right) \psi(x). \quad (\text{B.21})$$

These fields $A_\mu^a T^a$ transform in the following way:

$$A_\mu^a T^a \rightarrow A^a T^a + \frac{1}{g} \partial_\mu \alpha^a T^a + i \alpha^a A_\mu^b [T^a, T^b], \quad (\text{B.22})$$

and the field strength tensor is defined:

$$[D_\mu, D_\nu] \psi(x) := -ig F_{\mu\nu}^a T^a \psi(x), \quad (\text{B.23})$$

$$F_{\mu\nu}^a T^a = \partial_\mu A_\nu^a T^a - \partial_\nu A_\mu^a T^a + ig A_\mu^b A_\nu^c [T^b, T^c] T^a, \quad (\text{B.24})$$

where the last term can be re-written as $gf^{abc} A_\mu^b A_\nu^c T^a$. This last term is the main discerning difference between the abelian and non-abelian symmetry groups, since the field strength tensor now manifestly includes the structure constants f^{abc} . Evidently, this is no longer gauge invariant [4], but a term can be made with this that is gauge invariant, such as the kinematic lagrangian (trace summed over the number of dimensions in the group):

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu}^a F^{a,\mu\nu}, \quad (\text{B.25})$$

which is gauge invariant.

In addition, for an intermediate boson propagators, they have the propagator form (assuming a particular gauge choice) [4]:

$$\langle \psi_{l\alpha}(x) \bar{\psi}_{j\beta}(y) \rangle = \int \frac{d^4 k}{(2\pi)^4} \left(\frac{i}{\not{k} - m + i\epsilon_{\alpha\beta}} \right) \delta_{lj} e^{-ik^\mu(x_\mu - y_\mu)}, \quad (\text{B.26})$$

$$\langle A_\mu^l(x) A_\nu^j(y) \rangle = \int \frac{d^4 k}{(2\pi)^4} \left(\frac{-ig_{\mu\nu}}{k^\rho k_\rho - m^2 + i\epsilon} \right) \delta^{lj} e^{-ik_\sigma(x^\sigma - y^\sigma)}. \quad (\text{B.27})$$

In this case, the Roman indices j, l denote elements of the symmetry group either for the fermions or for the gauge bosons, and greek indices α, β denote Dirac spinor indices and μ, ν the usual Lorentz four vector indices. To point out the specific gauge fixing of this simpler propagator, it can be shown from functional methods in [4] section 16.2 that the full propagator form is:

$$\langle A_\mu^a(x) A_\nu^b(y) \rangle = \int \frac{d^4 k}{(2\pi)^4} \frac{-i}{k^2 - m^2 + i\epsilon} \left(g_{\mu\nu} - (1 - \xi) \frac{k_\mu k_\nu}{k^\rho k_\rho} \right) \delta^{ab} e^{-ik_\sigma(x^\sigma - y^\sigma)}, \quad (\text{B.28})$$

with the simplification known as the *Feynman-t'Hooft Gauge* ($\xi = 1$). In the case of the Abelian symmetries such as $U(1)$, this is the only remaining factor, but in the case of non-abelian symmetries, non-physical “ghost” terms appear. These are consequences only of the choice of gauge chosen, and are non-physical because they must be anti-commuting fields which transform as per scalars under Lorentz transformations. Effectively this corresponds

to a gauge fixing lagrangian term referred to as the Fadeev-Popov Lagrangian, namely for some ghost field $\mathcal{C}$ [4]:

$$\mathcal{L}_{\text{FP-ghost}} := \bar{\mathcal{C}}^a \left(-\partial^\mu D_\mu^{ac} \right) \mathcal{C}^c \equiv \bar{\mathcal{C}}^a \left(-\square \delta^{ac} - g \partial^\mu f^{abc} A_\mu^b \right) \mathcal{C}^c, \quad (\text{B.29})$$

(a, b, c) denoting the symmetry group indices. This lagrangian term relates to the feynman propagator for the ghost:

$$\langle \mathcal{C}^a \bar{\mathcal{C}}^b \rangle = \int \frac{d^4 k}{(2\pi)^4} \frac{i}{k^2} \delta^{ab} e^{-ik^\rho (x_\rho - y_\rho)}. \quad (\text{B.30})$$

B.0.1.2 Consequences of gauge invariance

For U(1) gauge symmetries, mass terms of type $m A^\mu A_\mu$ are not gauge invariant, so the field must be massless (and no other terms exist, hence no self interactions). However, in non-abelian theories, it is possible to have self interactions due to the non-commuting structure of the relevant Gauge group.

In addition, if we include a non-abelian gauge theory (e.g. SU(3) for QCD), the standard mass terms of both the Dirac and Klein Gordon equations. This arises because:

$$\begin{aligned} e^A \times e^B &= e^C, \\ C &= A + B + \frac{1}{2}[A, B] + \frac{1}{12}[A, [A, B]] - \frac{1}{12}[B, [A, B]] \\ &\quad + \text{Higher order commutators of } [A, B] \\ e^{-iA^a} e^{iA^b} &\neq \mathbb{I} \text{ if } [A^a, A^b] \neq 0. \end{aligned} \quad (\text{B.31})$$

Thus the gauge invariance of mass terms is not manifestly preserved, and as such a Gauge invariant theory cannot include bare mass terms. But, this does not stipulate such terms cannot exist in a particular gauge!). These terms will re-appear when the Higgs model is discussed in section 2.2.

The renormalisation group

C.1 Renormalisation group flow

In the specific case of massless ϕ^4 theory, [4][Ch12] propose that Green's functions corresponding to the intermediate propagator terms. The n-point Green's functions can be calculated from the direct diagrammatic terms (the reader is directed to either [4] or [5]). However, the physical propagators (hence the n-point greens functions) must depend in some way on an energy scale Q^2 , as well as a coupling constant λ .

This evolution is governed by the ‘‘Callan-Symanzik’’ equation, which takes the form:

$$\left[Q^2 \frac{\partial}{\partial Q^2} + \beta(\lambda) \frac{\partial}{\partial \lambda} + n\gamma(\lambda) \right] G^{(n)}(\mathbf{x}; Q^2, \lambda) = 0 \quad (\text{C.1})$$

for some universal functions $\beta(\lambda)$ and $\gamma(\lambda)$. The solution $\beta(\lambda)$ can be found to satisfy the equation:

$$\frac{d}{d \log(p/Q^2)} \lambda_p(p) = \beta(\lambda_p(p)); \lambda_p(p = Q^2) = \lambda \quad (\text{C.2})$$

The solutions of this equation differ depending on the theory of interest (which affects the functional form of β).

For QED, the coupling constant is governed by the running equation [123]:

$$\alpha_s(\mu^2) = \frac{\alpha_{EM}}{1 - 2 \frac{\alpha_{EM}}{3\pi} \ln \left(\frac{\mu^2}{\Lambda} \right)} \quad (\text{C.3})$$

and for QCD this takes the form [124] for some reference energy scale μ^2 :

$$\alpha_s(\mu^2) = \frac{1}{\beta_0 \ln(\mu^2/\Lambda^2)}. \quad (\text{C.4})$$

with the relevant constants then chosen to match the measured value of the coupling constant at some reference scale (e.g. Z mass for QCD, electron mass for QED).

To summarise these behaviours of the couplings, as a function of scale Λ , the coupling strength for QCD decreases with increasing energy (this corresponds to the *asymptotic*

freedom at large energy scales), while the coupling strength of the EM interaction increases as a function of μ^2 (due to the increase in sea charges formed from vacuum).

For a more complete discussion of renormalisation group flow, we direct the interested reader to either [\[4\]](#) or [\[5\]](#).

Systematic tables for full Run 2 stop0L analysis

These tables outline the main systematic contributions to the signal regions for [2].

	SRA-TT	SRA-TW	SRA-T0	SRB-TT	SRB-TW	SRB-T0
Total syst. unc.	15	12	10	14	9	9
$t\bar{t}$ theory	2	2	1	11	6	4
Single-top theory	7	5	4	1	<1	1
$t\bar{t}Z$ theory	3	<1	<1	<1	<1	<1
Z theory	<1	<1	1	<1	<1	<1
$\mu_{t\bar{t}}$	<1	<1	<1	4	4	4
$\mu_{t\bar{t}+Z}$	6	2	2	4	3	1
μ_Z	3	5	5	3	3	3
μ_W	2	3	3	4	4	3
$\mu_{\text{single top}}$	6	4	5	3	4	5
JER	7	3	2	6	2	3
JES	4	4	2	2	<1	<1
b -tagging	5	3	3	2	1	2
$E_{\text{T}}^{\text{miss}}$ soft term	2	1	1	<1	<1	<1
MC statistics	7	7	5	3	3	2

Table D.1: Systematic breakdown for signal regions SRA and SRB from [2]. Only contributions over 1% are included.

	SRC1	SRC2	SRC3	SRC4	SRC5	SRD0	SRD1	SRD2
Total syst. unc.	25	18	20	27	27	18	31	12
$t\bar{t}$ theory	20	11	12	16	21	4	9	5
Single-top theory	<1	<1	<1	<1	<1	<1	4	2
Z theory	<1	<1	1	2	4	7	3	2
W theory	<1	<1	1	2	3	<1	<1	<1
$\mu_{t\bar{t}}$	12	13	14	14	11	<1	2	5
μ_Z	<1	<1	<1	<1	<1	5	3	2
μ_W	<1	<1	<1	<1	<1	4	5	3
JER	5	<1	8	15	7	8	18	4
JES	<1	1	<1	4	6	1	4	2
b -tagging	2	2	2	2	2	3	5	7
Track-jet flavour	<1	<1	<1	<1	<1	4	7	<1
Track-jet flavour (low p_T)	<1	<1	<1	<1	<1	7	4	1
E_T^{miss} soft term	<1	<1	<1	<1	3	<1	<1	<1
Pile-up	<1	<1	<1	1	<1	2	12	<1
MC statistics	3	2	3	4	6	11	17	5

Table D.2: Systematic breakdown for signal regions SRC and SRD from [2]. Only contributions over 1% are included.

Bibliography

- [1] D. Upper, *The unsuccessful self-treatment of a case of "writer's block"*, J Appl Behav Anal **7** (1974) 497 (cit. on p. ii).
- [2] ATLAS Collaboration, *Search for a scalar partner of the top quark in the all-hadronic $t\bar{t}$ plus missing transverse momentum final state at $\sqrt{s}=13$ TeV with the ATLAS detector*, Eur. Phys. Journal C. **80** (2020), arXiv: [2004.14060 \[hep-ex\]](#) (cit. on pp. vii–ix, 34, 74, 95, 99, 111, 128–130, 142, 143, 153–172, 177–179, 181–185, 188, 189, 191, 193–195, 198, 200, 202, 210, 211).
- [3] ATLAS Collaboration, *Search for a scalar partner of the top quark in the jets plus missing transverse momentum final state at $\sqrt{s}=13$ TeV with the ATLAS detector*, JHEP **12** (2017) 085, arXiv: [1709.04183 \[hep-ex\]](#) (cit. on pp. viii, 94–100, 102–110, 112, 114, 117–119, 125, 131, 136, 159, 162, 163).
- [4] M. E. Peskin, *An introduction to quantum field theory*, eng, Addison-Wesley, 1995, ISBN: 0201503972 (cit. on pp. 3, 4, 6, 7, 10, 12–15, 204–209).
- [5] A. Zee, *Quantum field theory in a nutshell*, eng, 2nd ed., Princeton University Press, 2010, ISBN: 9780691140346 (cit. on pp. 12–14, 204, 208, 209).
- [6] S. P. Martin, *A Supersymmetry primer*, (1997), [Adv. Ser. Direct. High Energy Phys.18,1(1998)], preprint version 7 updated in 2016, chapters 1-4, arXiv: [hep-ph/9709356 \[hep-ph\]](#) (cit. on pp. 16, 26, 27, 29–32).
- [7] J. Terning, *Modern supersymmetry: Dynamics and duality*, Oxford, UK: Clarendon (2006) 324 p, chapters 1-3, 2006 (cit. on pp. 16, 17, 22–30).
- [8] V. C. Rubin, J. Ford W. Kent and N. Thonnard, *Extended rotation curves of high-luminosity spiral galaxies. IV. Systematic dynamical properties, Sa through Sc*, Astrophys. J. **225** (1978) L107 (cit. on p. 18).
- [9] D. Clowe et al., *A Direct Empirical Proof of the Existence of Dark Matter*, The Astrophysical Journal **648** (2006) L109, ISSN: 1538-4357, URL: <http://dx.doi.org/10.1086/508162> (cit. on p. 19).
- [10] A. D. Dolgov, S. L. Dubovsky, G. I. Rubtsov and I. I. Tkachev, *Constraints on millicharged particles from Planck data*, Physical Review D **88** (2013), ISSN: 1550-2368, URL: <http://dx.doi.org/10.1103/PhysRevD.88.117701> (cit. on p. 18).

-
- [11] S. Weinberg, *The quantum theory of fields. Vol. 3: Supersymmetry*, Cambridge University Press, 2013, ISBN: 978-0-521-67055-5, 978-1-139-63263-8, 978-0-521-67055-5 (cit. on pp. 19, 22–25, 27).
- [12] H. Müller-Kirsten and A. Wiedemann, *Introduction to Supersymmetry*, World Scientific lecture notes in physics, World Scientific, 2010, ISBN: 9789814293426, URL: <https://books.google.co.uk/books?id=65DkngEACAAJ> (cit. on pp. 19–21, 27).
- [13] S. Coleman and J. Mandula, *All Possible Symmetries of the S Matrix*, *Phys. Rev.* **159** (5 1967) 1251, URL: <https://link.aps.org/doi/10.1103/PhysRev.159.1251> (cit. on p. 22).
- [14] R. Haag, J. T. Lopuszanski and M. Sohnius, *All possible generators of supersymmetries of the S-matrix*, *Nuclear Physics B* **88** (1975) 257, ISSN: 0550-3213, URL: <http://www.sciencedirect.com/science/article/pii/0550321375902795> (cit. on p. 22).
- [15] Y. Golfand and E. Likhtman, *Extension of the Algebra of Poincare Group Generators and Violation of p Invariance*, *JETP Lett.* **13** (1971) 323 (cit. on p. 22).
- [16] J. Wess and B. Zumino, *Supergauge transformations in four dimensions*, *Nuclear Physics B* **70** (1974) 39, ISSN: 0550-3213, URL: <http://www.sciencedirect.com/science/article/pii/0550321374903551> (cit. on pp. 22, 24).
- [17] ATLAS Collaboration, *A search for pair-produced resonances in four-jet final states at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *The European Physical Journal C* **78** (2018), ISSN: 1434-6052, URL: <http://dx.doi.org/10.1140/epjc/s10052-018-5693-4> (cit. on p. 28).
- [18] ATLAS Collaboration, *Measurements of $t\bar{t}$ differential cross-sections of highly boosted top quarks decaying to all-hadronic final states in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, *Physical Review D* **98** (2018), ISSN: 2470-0029, URL: <http://dx.doi.org/10.1103/PhysRevD.98.012003> (cit. on pp. 28, 29).
- [19] C. F. Berger, J. S. Gainer, J. L. Hewett and T. G. Rizzo, *Supersymmetry without prejudice*, *Journal of High Energy Physics* **2009** (2009) 023, ISSN: 1029-8479, URL: <http://dx.doi.org/10.1088/1126-6708/2009/02/023> (cit. on pp. 33, 34).
- [20] ATLAS Collaboration, *Summary of the ATLAS experiments sensitivity to supersymmetry after LHC Run 1 interpreted in the phenomenological MSSM*, *Journal of High Energy Physics* **2015** (2015), ISSN: 1029-8479, URL: [http://dx.doi.org/10.1007/JHEP10\(2015\)134](http://dx.doi.org/10.1007/JHEP10(2015)134) (cit. on pp. 33, 35, 38).
- [21] ATLAS Collaboration, *Search for top-squark pair production in final states with one lepton, jets, and missing transverse momentum using 36 fb1 of $\sqrt{s} = 13$ TeV pp collision data with the ATLAS detector*, *Journal of High Energy Physics* **2018** (2018), ISSN: 1029-8479, URL: [http://dx.doi.org/10.1007/JHEP06\(2018\)108](http://dx.doi.org/10.1007/JHEP06(2018)108) (cit. on pp. 34, 102).
- [22] ATLAS Collaboration, *Search for direct top squark pair production in the 3-body decay mode with a final state containing one lepton, jets, and missing transverse momentum in $\sqrt{s} = 13$ TeV pp collision data with the ATLAS detector*, tech. rep. ATLAS-CONF-2019-017, CERN, 2019, URL: <https://cds.cern.ch/record/2676594> (cit. on pp. 34, 94, 153, 195).
-

-
- [23] ATLAS Collaboration, *Search for direct top squark pair production in final states with two leptons in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*, (2017), URL: <https://cds.cern.ch/record/2265806> (cit. on pp. 34, 94).
- [24] ATLAS Collaboration, *Search for direct top squark pair production in events with a Higgs or Z boson, and missing transverse momentum in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*, *Journal of High Energy Physics* **2017** (2017), ISSN: 1029-8479, URL: [http://dx.doi.org/10.1007/JHEP08\(2017\)006](http://dx.doi.org/10.1007/JHEP08(2017)006) (cit. on pp. 34, 94).
- [25] ATLAS Collaboration, *Search for new phenomena in final states with an energetic jet and large missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, *Phys. Rev.* **D94** (2016) 032005, arXiv: [1604.07773 \[hep-ex\]](https://arxiv.org/abs/1604.07773) (cit. on p. 34).
- [26] C. Pagliarone, *SUSY Searches at Tevatron Collider*, 2000, arXiv: [hep-ex/0011016 \[hep-ex\]](https://arxiv.org/abs/hep-ex/0011016) (cit. on pp. 35, 36).
- [27] ATLAS Collaboration, *ATLAS Run 1 searches for direct pair production of third-generation squarks at the Large Hadron Collider*, *The European Physical Journal C* **75** (2015), ISSN: 1434-6052, URL: <http://dx.doi.org/10.1140/epjc/s10052-015-3726-9> (cit. on pp. 35, 36, 110).
- [28] V. Khachatryan et al., *Searches for third-generation squark production in fully hadronic final states in proton-proton collisions at $\sqrt{s} = 8$ TeV (CMS)*, *Journal of High Energy Physics* **2015** (2015), ISSN: 1029-8479, URL: [http://dx.doi.org/10.1007/JHEP06\(2015\)116](http://dx.doi.org/10.1007/JHEP06(2015)116) (cit. on pp. 35, 37).
- [29] L. Evans and P. Bryant, *LHC Machine*, *Journal of Instrumentation* **3** (2008) S08001, URL: <https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08001> (cit. on pp. 39, 40, 42).
- [30] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *Journal of Instrumentation* **3** (2008) S08003, URL: <https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08003> (cit. on pp. 40, 41, 43–51).
- [31] CMS Collaboration, *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*, Technical Design Report CMS, CERN, 2006, URL: <https://cds.cern.ch/record/922757> (cit. on p. 40).
- [32] LHCb Collaboration, *The LHCb Detector at the LHC*, *Journal of Instrumentation* **3** (2008) S08005, URL: <https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08005> (cit. on p. 40).
- [33] P. Giubellino, *The ALICE detector at LHC*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **344** (1994) 27, ISSN: 0168-9002, URL: <http://www.sciencedirect.com/science/article/pii/0168900294906475> (cit. on p. 40).
- [34] S. Holmes, R. S. Moore and V. Shiltsev, *Overview of the Tevatron collider complex: goals, operations and performance*, *Journal of Instrumentation* **6** (2011) T08001, URL: <https://doi.org/10.1088%2F1748-0221%2F6%2F08%2Ft08001> (cit. on p. 40).
- [35] C. Lefèvre, ‘The CERN accelerator complex. Complexe des accélérateurs du CERN’, 2008, URL: <https://cds.cern.ch/record/1260465> (cit. on p. 41).
-

-
- [36] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, tech. rep. ATLAS-CONF-2019-021, CERN, 2019, URL: <http://cds.cern.ch/record/2677054> (cit. on pp. 42, 202).
- [37] ATLAS Collaboration, *ATLAS Luminosity Public Results Run2*, 2020, URL: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2> (cit. on p. 42).
- [38] M. Capeans et al., *ATLAS Insertable B-Layer Technical Design Report*, tech. rep. CERN-LHCC-2010-013. ATLAS-TDR-19, 2010, URL: <https://cds.cern.ch/record/1291633> (cit. on p. 44).
- [39] ATLAS Collaboration, *The upgraded Pixel detector and the commissioning of the Inner Detector tracking of the ATLAS experiment for Run-2 at the Large Hadron Collider*, tech. rep. ATL-PHYS-PROC-2016-104, 15 pages, EPS-HEP 2015 Proceedings: CERN, 2016, URL: <http://cds.cern.ch/record/2209070> (cit. on p. 45).
- [40] H. Zhang and ATLAS Liquid Argon Calorimeter Group, *The ATLAS Liquid Argon Calorimeter: Overview and Performance*, *Journal of Physics: Conference Series* **293** (2011) 012044, URL: <https://doi.org/10.1088%2F1742-6596%2F293%2F1%2F012044> (cit. on p. 46).
- [41] L. Heelan, *Performance of the ATLAS Tile Calorimeter*, *Journal of Physics: Conference Series* **623** (2015) 012014, URL: <https://doi.org/10.1088%2F1742-6596%2F623%2F1%2F012014> (cit. on p. 47).
- [42] ATLAS Collaboration, *Performance of the ATLAS trigger system in 2015*, *The European Physical Journal C* **77** (2017), ISSN: 1434-6052, URL: <http://dx.doi.org/10.1140/epjc/s10052-017-4852-3> (cit. on pp. 51–54).
- [43] ATLAS Collaboration, *The ATLAS Data Acquisition and High Level Trigger system*, *Journal of Instrumentation* **11** (2016) P06008, URL: <https://doi.org/10.1088%2F1748-0221%2F11%2F06%2Fp06008> (cit. on p. 51).
- [44] ATLAS Collaboration et al., *Search for Low-Mass Dijet Resonances Using Trigger-Level Jets with the ATLAS Detector in pp Collisions at $\sqrt{s} = 13$ TeV*, *Physical Review Letters* **121** (2018), ISSN: 1079-7114, URL: <http://dx.doi.org/10.1103/PhysRevLett.121.081801> (cit. on p. 52).
- [45] A. Corso-Radu et al., *Data Quality Monitoring Framework for the ATLAS experiment: Performance achieved with colliding beams at the LHC*, *Journal of Physics: Conference Series* **331** (2011) 022027, URL: <https://doi.org/10.1088%5C%2F1742-6596%5C%2F331%5C%2F2%5C%2F022027> (cit. on p. 55).
- [46] R. Bianchi and et al., *Event visualization in ATLAS*, tech. rep. ATL-SOFT-PROC-2017-046. 7, CERN, 2017, URL: <http://cds.cern.ch/record/2248396> (cit. on pp. 55, 56).
- [47] ATLAS Collaboration, *Search for massive supersymmetric particles in multi-jet final states produced in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, tech. rep. ATLAS-CONF-2016-057, CERN, 2016, URL: <https://cds.cern.ch/record/2206149> (cit. on p. 56).
-

-
- [48] J. Alwall et al., *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), ISSN: 1029-8479, URL: [http://dx.doi.org/10.1007/JHEP07\(2014\)079](http://dx.doi.org/10.1007/JHEP07(2014)079) (cit. on p. 59).
- [49] T. Gleisberg et al., *Event generation with SHERPA 1.1*, *Journal of High Energy Physics* **2009** (2009) 007, ISSN: 1029-8479, URL: <http://dx.doi.org/10.1088/1126-6708/2009/02/007> (cit. on p. 59).
- [50] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *Journal of High Energy Physics* **2010** (2010), ISSN: 1029-8479, URL: [http://dx.doi.org/10.1007/JHEP06\(2010\)043](http://dx.doi.org/10.1007/JHEP06(2010)043) (cit. on p. 59).
- [51] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, *Computer Physics Communications* **191** (2015) 159, ISSN: 0010-4655, URL: <http://dx.doi.org/10.1016/j.cpc.2015.01.024> (cit. on pp. 59, 65, 170).
- [52] A. Buckley et al., *LHAPDF6: parton density access in the LHC precision era*, *The European Physical Journal C* **75** (2015), ISSN: 1434-6052, URL: <http://dx.doi.org/10.1140/epjc/s10052-015-3318-8> (cit. on p. 59).
- [53] GEANT4 Collaboration et al., *Geant4 a simulation toolkit*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** (2003) 250, ISSN: 0168-9002, URL: <http://www.sciencedirect.com/science/article/pii/S0168900203013688> (cit. on pp. 59, 60).
- [54] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, *Eur. Phys. J.* **C70** (2010) 823, arXiv: 1005.4568 [physics.ins-det] (cit. on p. 60).
- [55] M. Dobbs and J. B. Hansen, *The HepMC C++ Monte Carlo event record for High Energy Physics*, *Comput. Phys. Commun.* **134** (2001) 41 (cit. on p. 60).
- [56] J. Alwall et al., *A standard format for Les Houches Event Files*, *Computer Physics Communications* **176** (2007) 300, ISSN: 0010-4655, URL: <http://dx.doi.org/10.1016/j.cpc.2006.11.010> (cit. on p. 60).
- [57] A. Basalaev and Z. Marshall, *The Fast Simulation Chain for ATLAS*, *Journal of Physics: Conference Series* **898** (2017) 042016, URL: <https://doi.org/10.1088/1742-6596/898/2/042016> (cit. on pp. 61, 62).
- [58] Z. A. Basalaev, *Simulation of Pile-up in the ATLAS Experiment*, *Journal of Physics: Conference Series* **513** (2014) 022024, URL: <https://doi.org/10.1088/1742-6596/513/2/022024> (cit. on p. 61).
- [59] *pprof (Profiling tool)*, 2017, URL: <https://github.com/google/pprof> (cit. on p. 63).
- [60] T. Cuhadar Donszelmann, W. Lampl and G. Stewart, *ART - ATLAS Release Tester using the Grid*, tech. rep. ATL-SOFT-PROC-2020-012, CERN, 2020, URL: <http://cds.cern.ch/record/2709656> (cit. on p. 63).
- [61] S. Plätzer and S. Gieseke, *Dipole showers and automated NLO matching in Herwig++*, *The European Physical Journal C* **72** (2012), ISSN: 1434-6052, URL: <http://dx.doi.org/10.1140/epjc/s10052-012-2187-7> (cit. on p. 65).
-

-
- [62] A. Hrynevich, *ATLAS jet and missing energy reconstruction, calibration and performance in LHC Run-2*, *Journal of Instrumentation* **12** (2017) C06038, URL: <https://doi.org/10.1088/2F1748-0221/2F12/2F06/2Fc06038> (cit. on pp. 65, 66).
- [63] M. Cacciari, G. P. Salam and G. Soyez, *The Anti- $k(t)$ jet clustering algorithm*, *JHEP* **04** (2008) 063, arXiv: 0802.1189 [hep-ph] (cit. on pp. 65, 69).
- [64] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, *The European Physical Journal C* **77** (2017), ISSN: 1434-6052, URL: <http://dx.doi.org/10.1140/epjc/s10052-017-5031-2> (cit. on pp. 65, 200).
- [65] G. P. Salam and G. Soyez, *A practical seedless infrared-safe cone jet algorithm*, *Journal of High Energy Physics* **2007** (2007) 086, ISSN: 1029-8479, URL: <http://dx.doi.org/10.1088/1126-6708/2007/05/086> (cit. on p. 65).
- [66] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, *Physical Review D* **48** (1993) 3160, ISSN: 0556-2821, URL: <http://dx.doi.org/10.1103/PhysRevD.48.3160> (cit. on p. 65).
- [67] G. Aad et al., *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, *The European Physical Journal C* **77** (2017), ISSN: 1434-6052, URL: <http://dx.doi.org/10.1140/epjc/s10052-017-5004-5> (cit. on p. 66).
- [68] ATLAS Collaboration, *Soft b -hadron tagging for compressed SUSY scenarios*, tech. rep. ATLAS-CONF-2019-027, CERN, 2019, URL: <https://cds.cern.ch/record/2682131> (cit. on pp. 66–69).
- [69] G. Aad et al., *ATLAS b -jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV*, *The European Physical Journal C* **79** (2019), ISSN: 1434-6052, URL: <http://dx.doi.org/10.1140/epjc/s10052-019-7450-8> (cit. on pp. 67, 68, 125, 128).
- [70] ATLAS Collaboration, *Performance of the ATLAS Secondary Vertex b -tagging Algorithm in 900 GeV Collision Data*, tech. rep. ATLAS-CONF-2010-004, CERN, 2010, URL: <https://cds.cern.ch/record/1273194> (cit. on p. 68).
- [71] A. Calandri, *Flavour tagging algorithms and performance at the ATLAS experiment*, (2016), URL: <https://cds.cern.ch/record/2209195> (cit. on p. 67).
- [72] *Optimisation of the ATLAS b -tagging performance for the 2016 LHC Run*, tech. rep. ATL-PHYS-PUB-2016-012, CERN, 2016, URL: <https://cds.cern.ch/record/2160731> (cit. on p. 68).
- [73] <http://www.quantumdiaries.org/wp-content/uploads/2012/08/boost.png>, Boosted top quark reconstruction image, accessed 7/6/2017 (cit. on p. 69).
- [74] J. Shelton, ‘*Jet Substructure*’, *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: Searching for New Physics at Small and Large Scales (TASI 2012): Boulder, Colorado, June 4-29, 2012*, 2013 303, arXiv: 1302.0260 [hep-ph], URL: <https://inspirehep.net/record/1217434/files/arXiv:1302.0260.pdf> (cit. on p. 69).
- [75] J. Thaler and K. Van Tilburg, *Identifying Boosted Objects with N -subjettiness*, *JHEP* **03** (2011) 015, arXiv: 1011.2268 [hep-ph] (cit. on p. 69).
- [76] A. J. Larkoski, G. P. Salam and J. Thaler, *Energy correlation functions for jet substructure*, *Journal of High Energy Physics* **2013** (2013), ISSN: 1029-8479, URL: [http://dx.doi.org/10.1007/JHEP06\(2013\)108](http://dx.doi.org/10.1007/JHEP06(2013)108) (cit. on pp. 69–71).
-

-
- [77] ATLAS Collaboration, *Performance of top-quark and W-boson tagging with ATLAS in Run 2 of the LHC*, *The European Physical Journal C* **79** (2019), ISSN: 1434-6052, URL: <http://dx.doi.org/10.1140/epjc/s10052-019-6847-8> (cit. on pp. 69, 112).
- [78] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC protonproton collision data at $\sqrt{s} = 13$ TeV*, *The European Physical Journal C* **79** (2019), ISSN: 1434-6052, URL: <http://dx.doi.org/10.1140/epjc/s10052-019-7140-6> (cit. on pp. 71–73).
- [79] ATLAS Collaboration, *Search for a scalar partner of the top quark in the jets plus missing transverse momentum final state at $\sqrt{s}=13$ TeV with the ATLAS detector*, *JHEP* **12** (2017) 085, arXiv: [1709.04183 \[hep-ex\]](https://arxiv.org/abs/1709.04183) (cit. on p. 73).
- [80] ATLAS Collaboration, *Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run 2 data collected in 2015 and 2016*, *Eur. Phys. J. C* **79** (2018) 205. 55 p, 55 pages in total, author list starting page 39, 18 figures, 2 tables, published in EPJC. All figures including auxiliary figures are available at <http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/PERF-2017-02>, URL: <http://cds.cern.ch/record/2643391> (cit. on p. 73).
- [81] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in protonproton collision data at $\sqrt{s} = 13$ TeV*, *The European Physical Journal C* **76** (2016), ISSN: 1434-6052, URL: <http://dx.doi.org/10.1140/epjc/s10052-016-4120-y> (cit. on pp. 73–75).
- [82] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, tech. rep. ATLAS-CONF-2014-018, CERN, 2014, URL: <https://cds.cern.ch/record/1700870> (cit. on pp. 75, 76).
- [83] ATLAS Collaboration et al., *Performance of missing transverse momentum reconstruction with the ATLAS detector using protonproton collisions at $\sqrt{s} = 13$ TeV*, *The European Physical Journal C* **78** (2018), ISSN: 1434-6052, URL: <http://dx.doi.org/10.1140/epjc/s10052-018-6288-9> (cit. on pp. 75–77).
- [84] L. Lista, *Practical Statistics for Particle Physicists*, 2016, arXiv: [1609.04150 \[physics.data-an\]](https://arxiv.org/abs/1609.04150) (cit. on pp. 78, 79, 81, 82, 84, 85).
- [85] R. von Mises, *Mathematical Theory of Probability and Statistics*, Academic press, 1964 (cit. on pp. 79, 80, 83).
- [86] N. J. and J. Harold, *Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability*, *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* (1937) (cit. on p. 79).
- [87] 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. CERN-OPEN-2012-016, New York U., 2012, URL: <https://cds.cern.ch/record/1456844> (cit. on pp. 81, 86–88).
- [88] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, *Ann. Math. Statist.* **9** (1938) 60, URL: <https://doi.org/10.1214/aoms/1177732360> (cit. on p. 82).
- [89] O. Behnke, K. Kröninger, G. Schott and T. Schörner-Sadenius, *Data analysis in high energy physics: a practical guide to statistical methods*, Wiley-VCH, 2013 (cit. on p. 83).
-

-
- [90] G. J. Feldman and R. D. Cousins, *Unified approach to the classical statistical analysis of small signals*, *Physical Review D* **57** (1998) 3873, ISSN: 1089-4918, URL: <http://dx.doi.org/10.1103/PhysRevD.57.3873> (cit. on p. 84).
- [91] A. L. Read, *Presentation of search results: the CLs technique*, *Journal of Physics G: Nuclear and Particle Physics* **28** (2002) 2693, URL: <https://doi.org/10.1088/2F0954-3899/2F28/2F10/2F313> (cit. on p. 85).
- [92] M. Baak et al., *HistFitter software framework for statistical data analysis*, *The European Physical Journal C* **75** (2015), ISSN: 1434-6052, URL: <http://dx.doi.org/10.1140/epjc/s10052-015-3327-7> (cit. on pp. 86, 87, 89, 107, 189, 193).
- [93] L. Moneta et al., *The RooStats Project*, 2010, arXiv: [1009.1003](https://arxiv.org/abs/1009.1003) [physics.data-an] (cit. on pp. 86, 111).
- [94] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, eConf **C0303241** (2003) MOLT007, ed. by L. Lyons and M. Karagoz, arXiv: [physics/0306116](https://arxiv.org/abs/hep-ph/0306116) (cit. on p. 86).
- [95] , *TMVA package*, <http://tmva.sourceforge.net/>, Accessed: 05-06-2017, User guide also referenced., 2011 (cit. on pp. 90, 92, 93, 115).
- [96] B. Kégl, *The return of AdaBoost.MH: multi-class Hamming trees*, 2013, arXiv: [1312.6086](https://arxiv.org/abs/1312.6086) [cs.LG] (cit. on pp. 92, 115).
- [97] J. H. Friedman, *Greedy Function Approximation: A Gradient Boosting Machine*, *The Annals of Statistics* **29** (2001) 1189, ISSN: 00905364, URL: <http://www.jstor.org/stable/2699986> (cit. on pp. 92, 115).
- [98] B. Ghojogh and M. Crowley, *The Theory Behind Overfitting, Cross Validation, Regularization, Bagging, and Boosting: Tutorial*, 2019, arXiv: [1905.12787](https://arxiv.org/abs/1905.12787) [stat.ML] (cit. on p. 92).
- [99] ATLAS Collaboration, *Search for new phenomena with top quark pairs in final states with one lepton, jets, and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, tech. rep. ATLAS-CONF-2020-003, CERN, 2020, URL: <http://cds.cern.ch/record/2711489> (cit. on pp. 94, 153, 195).
- [100] C. G. Lester and D. J. Summers, *Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders*, *Phys. Lett.* **B463** (1999) 99, arXiv: [hep-ph/9906349](https://arxiv.org/abs/hep-ph/9906349) [hep-ph] (cit. on p. 96).
- [101] A. Barr, C. Lester and P. Stephens, *A variable for measuring masses at hadron colliders when missing energy is expected; m_{T2} : the truth behind the glamour*, *Journal of Physics G: Nuclear and Particle Physics* **29** (2003) 2343, ISSN: 1361-6471, URL: <http://dx.doi.org/10.1088/0954-3899/29/10/304> (cit. on p. 96).
- [102] P. Jackson and C. Rogan, *Recursive Jigsaw Reconstruction: HEP event analysis in the presence of kinematic and combinatoric ambiguities*, (2017), arXiv preprint: [1705.10733](https://arxiv.org/abs/1705.10733) (cit. on pp. 97, 99).
- [103] H. An and L.-T. Wang, *Opening Up the Compressed Region of Top Squark Searches at 13 TeV LHC*, *Physical Review Letters* **115** (2015), ISSN: 1079-7114, URL: <http://dx.doi.org/10.1103/PhysRevLett.115.181602> (cit. on p. 100).
-

-
- [104] ATLAS Collaboration, *Search for squarks and gluinos with the ATLAS detector in final states with jets and missing transverse momentum using 4.7fb^{-1} of $\sqrt{s} = 7\text{ TeV}$ proton-proton collision data*, *Physical Review D* **87** (2013), ISSN: 1550-2368, URL: <http://dx.doi.org/10.1103/PhysRevD.87.012008> (cit. on pp. 104, 163).
- [105] D0 collaboration, *Observation of the top quark*, *Phys. Rev. Lett.* **74** (1995) 2632, arXiv: [hep-ex/9503003](https://arxiv.org/abs/hep-ex/9503003) [[hep-ex](#)] (cit. on p. 112).
- [106] CDF Collaboration, *Observation of top quark production in $\bar{p}p$ collisions*, *Phys. Rev. Lett.* **74** (1995) 2626, arXiv: [hep-ex/9503002](https://arxiv.org/abs/hep-ex/9503002) [[hep-ex](#)] (cit. on p. 112).
- [107] ATLAS Collaboration, *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*, tech. rep. ATL-PHYS-PUB-2015-045, CERN, 2015, URL: <http://cds.cern.ch/record/2064383> (cit. on p. 118).
- [108] ATLAS Collaboration, *Object-based missing transverse momentum significance in the ATLAS detector*, tech. rep. ATLAS-CONF-2018-038, CERN, 2018, URL: <https://cds.cern.ch/record/2630948> (cit. on p. 131).
- [109] ATLAS Collaboration, *Measurement of the $t\bar{t} + Z$ and $t\bar{t} + W$ cross sections in proton-proton collisions at $\sqrt{s} = 13\text{ TeV}$ with the ATLAS detector*, *Physical Review D* **99** (2019), ISSN: 2470-0029, URL: <http://dx.doi.org/10.1103/PhysRevD.99.072009> (cit. on pp. 136, 138).
- [110] ATLAS Collaboration, *Measurement of the $t\bar{t}Z$ and $t\bar{t}W$ cross sections in proton-proton collisions at $\sqrt{s} = 13\text{ TeV}$ with the ATLAS detector*, *Physical Review D* **99** (2019), ISSN: 2470-0029, URL: <http://dx.doi.org/10.1103/PhysRevD.99.072009> (cit. on p. 144).
- [111] M. Cacciari, G. P. Salam and G. Soyez, *The Anti- $k(t)$ jet clustering algorithm*, *JHEP* **04** (2008) 063, arXiv: [0802.1189](https://arxiv.org/abs/0802.1189) [[hep-ph](#)] (cit. on p. 157).
- [112] A. Collaboration, *Search for top-squark pair production in final states with one lepton, jets, and missing transverse momentum using 36fb^{-1} of $\sqrt{s} = 13\text{ TeV}$ pp collision data with the ATLAS detector*, *Journal of High Energy Physics* **2018** (2018), ISSN: 1029-8479, URL: [http://dx.doi.org/10.1007/JHEP06\(2018\)108](http://dx.doi.org/10.1007/JHEP06(2018)108) (cit. on p. 162).
- [113] C. D. White, S. Frixione, E. Laenen and F. Maltoni, *Isolating Wt production at the LHC*, *Journal of High Energy Physics* **2009** (2009) 074, ISSN: 1029-8479, URL: <http://dx.doi.org/10.1088/1126-6708/2009/11/074> (cit. on p. 170).
- [114] G. Choudalakis and D. Casadei, *Plotting the differences between data and expectation*, *The European Physical Journal Plus* **127** (2012), ISSN: 2190-5444, URL: <http://dx.doi.org/10.1140/epjp/i2012-12025-y> (cit. on pp. 185, 186).
- [115] *SUSY May 2020 Summary Plot Update*, tech. rep. ATL-PHYS-PUB-2020-013, CERN, 2020, URL: <http://cds.cern.ch/record/2718947> (cit. on p. 195).
- [116] A. M. Sirunyan et al., *Search for supersymmetry in proton-proton collisions at 13 TeV in final states with jets and missing transverse momentum*, *Journal of High Energy Physics* **2019** (2019), ISSN: 1029-8479, URL: [http://dx.doi.org/10.1007/JHEP10\(2019\)244](http://dx.doi.org/10.1007/JHEP10(2019)244) (cit. on pp. 196, 197).
-

- [117] F. Pastore, *ATLAS Run-2 status and performance*, [Nuclear and Particle Physics Proceedings](#) **270-272** (2016) 3, 18th Montpellier International Conference on Quantum Chromodynamics (QCD 15), ISSN: 2405-6014, URL: <http://www.sciencedirect.com/science/article/pii/S2405601416000031> (cit. on p. 199).
- [118] A. Sirunyan et al., *Particle-flow reconstruction and global event description with the CMS detector*, [Journal of Instrumentation](#) **12** (2017) P10003, ISSN: 1748-0221, URL: <http://dx.doi.org/10.1088/1748-0221/12/10/P10003> (cit. on p. 200).
- [119] ATLAS Collaboration, *RECAST framework reinterpretation of an ATLAS Dark Matter Search constraining a model of a dark Higgs boson decaying to two b-quarks*, tech. rep. ATL-PHYS-PUB-2019-032, CERN, 2019, URL: <https://cds.cern.ch/record/2686290> (cit. on p. 200).
- [120] D. Merkel, *Docker: Lightweight Linux Containers for Consistent Development and Deployment*, [Linux J.](#) **2014** (2014), ISSN: 1075-3583 (cit. on p. 200).
- [121] S. Höche, F. Krauss, S. Schumann and F. Siegert, *QCD matrix elements and truncated showers*, [JHEP](#) **05** (2009) 053, arXiv: 0903.1219 [[hep-ph](#)] (cit. on p. 202).
- [122] ‘Introduction to Group Theory’, *Mathematical Physics*, John Wiley & Sons, Ltd, 2007, chap. 15 597, ISBN: 9783527618132, eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/9783527618132.ch15>, URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/9783527618132.ch15> (cit. on pp. 203, 204).
- [123] G.-j. Ni, G.-h. Yang, R.-t. Fu and H. Wang, *Running Coupling Constants of Fermions with Masses in Quantum Electro Dynamics and Quantum Chromo Dynamics*, 1999, arXiv: [hep-ph/9906364](#) [[hep-ph](#)] (cit. on p. 208).
- [124] G. Proserpi, M. Raciti and C. Simolo, *On the running coupling constant in QCD*, [Progress in Particle and Nuclear Physics](#) **58** (2007) 387, ISSN: 0146-6410, URL: <http://dx.doi.org/10.1016/j.pnpnp.2006.09.001> (cit. on p. 208).

Colophon

This thesis is based on a template developed by Matthew Townson and Andrew Reeves. It was typeset with L^AT_EX 2_ε. It was created using the *memoir* package, maintained by Lars Madsen, with the *madsen* chapter style. The font used is Latin Modern, derived from fonts designed by Donald E. Kunith.