

1 SEARCH FOR  $t\bar{t}Z' \rightarrow t\bar{t}t\bar{t}$  PRODUCTION IN THE MULTILEPTON FINAL STATE IN  
2  $pp$  COLLISIONS AT  $\sqrt{s} = 13$  TEV WITH THE ATLAS DETECTOR

3 By

4 Hieu Le

5 A DISSERTATION

6 Submitted to  
7 Michigan State University  
8 in partial fulfillment of the requirements  
9 for the degree of

10 Physics — Doctor of Philosophy

11 2025

**ABSTRACT**

13 Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut  
14 labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco  
15 laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in  
16 voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat  
17 non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

**ACKNOWLEDGMENTS**

- 19 Advisor: Reinhard Schwienhorst
- 20 Postdoc: Binbin Dong
- 21 Committee
- 22 MSU group
- 23 ATLAS analysis group
- 24 Friend: Daniel, Grayson, Bella, Eric, Jordan
- 25 Other friends: Jasper, Adam, Brittany
- 26 Parents
- 27 Spouse: Allen Sechrist
- 28 ATLAS in general & funding agencies

## PREFACE

30 This is my preface. remarks remarks remarks

# TABLE OF CONTENTS

31	<b>List of Tables</b>	<b>vii</b>
32	<b>List of Figures</b>	<b>viii</b>
33	<b>KEY TO ABBREVIATIONS</b>	<b>ix</b>
34	<b>Roadmap</b>	<b>1</b>
35	<b>Chapter 1. Introduction</b>	<b>2</b>
36	<b>Chapter 2. Theoretical Overview</b>	<b>3</b>
37	2.1 The Standard Model	3
38	2.1.1 Elementary particles	3
39	2.1.2 Mathematical formalism	6
40	2.1.3 Four-top quark production	12
41	2.2 Beyond the Standard Model	13
42	<b>Chapter 3. LHC &amp; ATLAS Experiment</b>	<b>14</b>
43	3.1 The Large Hadron Collider	14
44	3.1.1 Overview	14
45	3.1.2 LHC operations	14
46	3.2 The ATLAS detector	15
47	3.2.1 Inner detector	16
48	3.2.2 Calorimeter systems	17
49	3.2.3 Muon spectrometer	19
50	3.2.4 Forward detectors	21
51	3.2.5 Magnetic systems	21
52	3.2.6 Trigger & data acquisition	21
53	<b>Chapter 4. Data &amp; Simulated Samples</b>	<b>23</b>
54	4.1 Data samples	23
55	4.2 Monte Carlo samples	24
56	4.2.1 $t\bar{t}Z'$ signal samples	24
57	4.2.2 Background samples	24
58	<b>Chapter 5. Particle Reconstruction &amp; Identification</b>	<b>26</b>
59	5.1 Primary reconstruction	26
60	5.1.1 Topological clusters	26
61	5.1.2 Tracks	27
62	5.1.3 Vertices	28
63	5.2 Jets	29
64	5.2.1 Flavor tagging	30

65	5.3	Leptons . . . . .	33
66	5.3.1	Electrons . . . . .	33
67	5.3.2	Muons . . . . .	37
68	5.4	Missing transverse momentum . . . . .	39
69	5.5	Overlap removal . . . . .	41
70	5.6	Object definition . . . . .	41
71	<b>Chapter 6.</b>	<b>Analysis Strategy . . . . .</b>	<b>43</b>
72	6.1	Event selection . . . . .	43
73	6.2	Analysis regions . . . . .	45
74	6.2.1	Signal regions . . . . .	46
75	6.2.2	Control regions . . . . .	46
76	6.2.3	Validation regions . . . . .	48
77	6.3	Background estimation . . . . .	49
78	6.3.1	Template fitting for fake/non-prompt estimation . . . . .	51
79	6.3.2	Charge misidentification data-driven estimation . . . . .	52
80	6.3.3	$t\bar{t}W$ background data-driven estimation . . . . .	53
81	6.3.4	$t\bar{t}Z$ background validation . . . . .	57
82	<b>Chapter 7.</b>	<b>Systematic Uncertainties . . . . .</b>	<b>58</b>
83	7.1	Experimental uncertainties . . . . .	58
84	7.1.1	Leptons . . . . .	58
85	7.1.2	Jets . . . . .	59
86	7.1.3	Missing transverse energy . . . . .	61
87	7.2	Modeling uncertainties . . . . .	61
88	7.2.1	Signal and irreducible background uncertainties . . . . .	61
89	7.2.2	Reducible background uncertainties . . . . .	64
90	<b>Chapter 8.</b>	<b>Results . . . . .</b>	<b>67</b>
91	8.1	Statistical model . . . . .	67
92	8.1.1	Binned profile likelihood fit . . . . .	67
93	8.1.2	Signal significance . . . . .	67
94	8.1.3	Limit exclusion . . . . .	67
95	8.2	Fit results . . . . .	67
96	8.3	Limits . . . . .	68
97	<b>Chapter 9.</b>	<b>Summary . . . . .</b>	<b>69</b>
98	<b>References . . . . .</b>	<b>70</b>	

99 **List of Tables**

100	Table 4.1: Caption . . . . .	23
101	Table 4.2: Summary of all Monte-Carlo samples used in this analysis. . . . .	25
102	Table 5.1: <a href="#">[1]</a> . . . . .	41
103	Table 5.2: Caption . . . . .	42
104	Table 6.1: Caption . . . . .	46
105	Table 6.2: Caption . . . . .	47
106	Table 6.3: Caption . . . . .	50
107	Table 7.1: Summary of the experimental systematic uncertainties considered in this	
108	analysis. . . . .	62
109	Table 7.2: Caption . . . . .	65
110	Table 7.3: Caption . . . . .	66

# 111 List of Figures

<small>112</small>	Figure 2.1: <a href="#">[2]</a> . . . . .	4
<small>113</small>	Figure 5.1: . . . . .	27
<small>114</small>	Figure 5.2: <a href="#">[3]</a> <a href="#">[4]</a> <a href="#">[5]</a> . . . . .	32
<small>115</small>	Figure 5.3: <a href="#">[6]</a> . . . . .	34
<small>116</small>	Figure 5.4: <a href="#">[7]</a> . . . . .	37



# KEY TO ABBREVIATIONS

118	Physical & mathematical quantities
119	$\chi^2$ chi-squared
120	$\Delta R$ angular distance
121	$\eta$ pseudorapidity
122	$E_T$ transverse energy
123	$E_T^{\text{miss}}$ missing transverse momentum
124	$\hat{H}$ Higgs oblique parameter
125	$I$ weak isospin
126	$L$ instantaneous luminosity
127	$\mu$ signal strength
128	$p_T$ transverse momentum
129	Particles
130	$b$ bottom quark
131	$pp$ proton-proton
132	$t\bar{t}$ top/anti-top quark
133	$t\bar{t}t\bar{t}$ four-top-quark
134	$tW$ single-top
135	Acronyms
136	<b>1LOS</b> one lepton, or two leptons of opposite charges
137	<b>ATLAS</b> A Toroidal LHC ApparatuS
138	<b>BDT</b> boosted decision tree
139	<b>BSM</b> Beyond the Standard Model
140	<b>CERN</b> European Organization for Nuclear Research
141	<b>CMS</b> Compact Muon Solenoid
142	<b>CR</b> control region

143 **ECIDS** Electron Charge ID Selector  
144 **EM** electromagnetic  
145 **EW** electroweak  
146 **GNN** graph neural network  
147 **HLT** High-Level Trigger  
148 **ID** inner detector  
149 **JER** jet energy resolution  
150 **JES** jet energy scale  
151 **JVT** Jet Vertex Tagger  
152 **L1** Level 1  
153 **LH** likelihood  
154 **LLH** log-likelihood  
155 **LO** leading order  
156 **LAr** liquid argon  
157 **LHC** Large Hadron Collider  
158 **MET** missing transverse energy  
159 **NF** normalization factor  
160 **NLO** next-to-leading order  
161 **NNLO** next-to-next-to-leading order  
162 **NP** nuisance parameter  
163 **OP** operating point  
164 **PS** parton shower  
165 **PDF** parton distribution function  
166 **PCBT** pseudo-continuous  $b$ -tagging  
167 **QED** quantum electrodynamics  
168 **QCD** quantum chromodynamics  
169 **QFT** quantum field theory

170	<b>QmisID</b>	charge mis-identification
171	<b>SF</b>	scale factor
172	<b>SM</b>	Standard Model
173	<b>SR</b>	signal region
174	<b>SSML</b>	two leptons of the same charge, or more than two leptons (multilepton)
175	<b>TDAQ</b>	Trigger and Data Acquisition

# Roadmap

1. Finish adding bullets for all sections ..... 06/04

Remaining

- introduction

2. Fill in details ..... 06/13

- Add missing figures

- Add missing bib

3. Finalize analysis

4. String everything together

5. Miscellaneous/logistics (proofreading, review, ATLAS approval, etc.) .....

6. Submission to the graduate school ..... 07/01

7. Defense ..... 07/15

# Chapter 1. Introduction

1. background and context

2. problem to be solved in thesis

3. aim of analysis:  $Z'$  consequences of many BSM theories, searching for  $Z'$

4. hypothesis/research question: searching for  $Z'$  in  $t\bar{t}t\bar{t}$  SSML channel

5. methodology: data collection -  $\sqrt{s}$  analysis regions -  $\sqrt{s}$  binned likelihood fit

6. thesis structure:

- ch2: SM/BSM theoretical background

- ch3: LHC/ATLAS experiment

- ch4: samples used in the analysis

- ch5: ATLAS particle reconstruction and identification techniques, and object definitions for the analysis

- ch6: analysis strategy

- ch7: systematic uncertainties affecting the analysis

- ch8: final results

- ch9: summary

## Chapter 2. Theoretical Overview

### 2.1 The Standard Model

The Standard Model of physics (SM) is currently the most successful formalism to describe the physical world at a microscopic scale.

The SM provides descriptions for all currently known elementary particles and three out of four fundamental forces with the exception of gravity.

#### 2.1.1 Elementary particles

Elementary particles in the SM can be classified into two groups: bosons, consisting of particles following Bose-Einstein statistics with integer spin and fermions, consisting of particles following Fermi-Dirac statistics with half-integer spin

Fermions are the building blocks of composite particles and consequently all known matter, and can be further split into quarks & leptons.

Bosons act as force mediators for all fundamental forces described by the SM. Bosons have two types: a scalar boson with spin 0 and vector gauge bosons with spin 1.

For each elementary particle there also exists a corresponding antiparticle with identical mass and opposite charge (electric or color).

#### Fermions

Quarks and leptons each has six flavors, grouped into three generations of doublets.

The six quark flavors consist of up ( $u$ ), down ( $d$ ), charm ( $c$ ), strange ( $s$ ), bottom ( $b$ ) and top

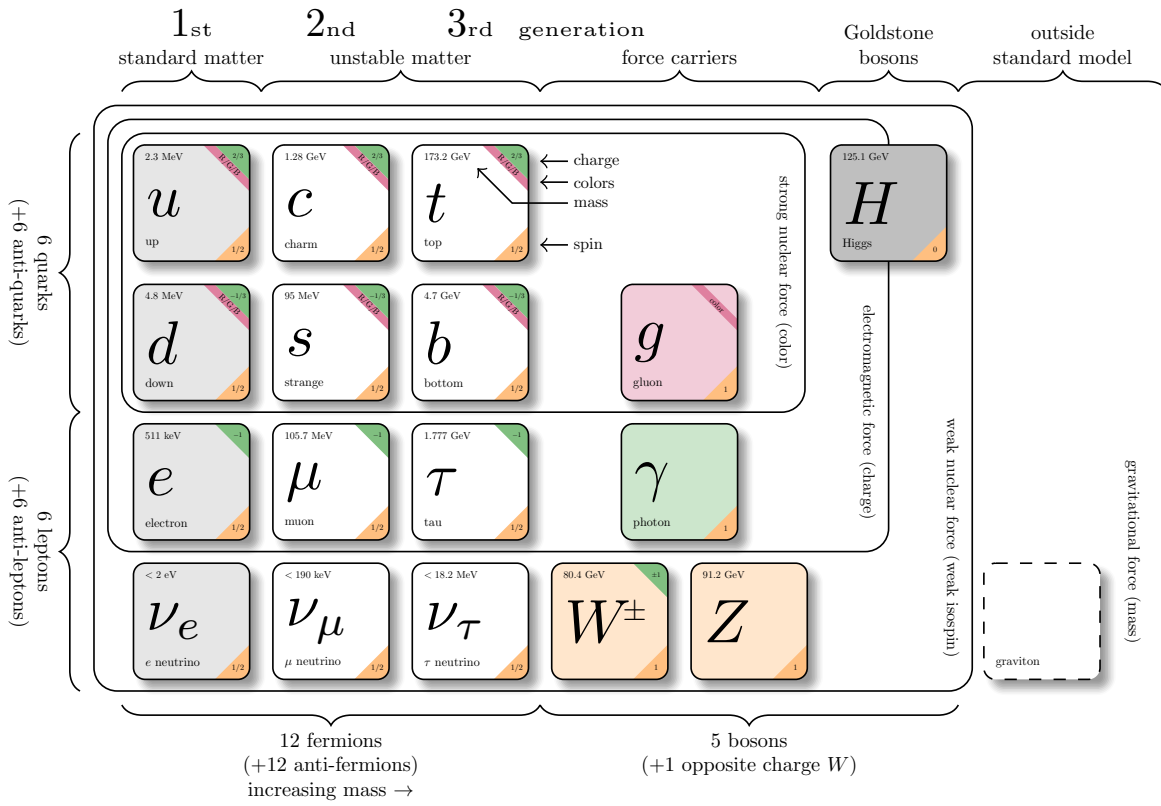


Figure 2.1: [2]

( $t$ ) quark flavors in increasing order of mass, forming three doublets ( $u, d$ ), ( $c, s$ ) and ( $t, b$ ).

Each doublet consists of one quark with electric charge of  $+2/3$  ( $u, s, t$ ), and one with charge of  $-1/3$  ( $d, c, b$ ).

Each quark also has a property known as color charge, with possible values of red ( $R$ ), green ( $G$ ), blue ( $B$ ) or antired ( $\bar{R}$ ), antigreen ( $\bar{G}$ ), and antiblue ( $\bar{B}$ ). Color charge follows color confinement rules, which allows only configurations of quarks with neutral color charge to exist in isolation. Neutral charge configurations can be formed from either a set of three colors ( $R, G, B$ ), a set of a color and its anticolor ( $q, \bar{q}$ ), or any combination of the two. Consequently, no isolated quark can exist in a vacuum and can only exist in bound states called hadrons.

Quarks are the only elementary particles in the SM that can interact with all four fundamental forces.

The three leptons doublets consist of electron ( $e$ ), muon ( $\mu$ ), tau ( $\tau$ ) and their respective neutrino flavors: electron neutrino ( $\nu_e$ ), muon neutrino ( $\nu_\mu$ ) and tau neutrino ( $\nu_\tau$ )

Charged leptons ( $e, \mu, \tau$ ) carry an electric charge of  $-1$ , while their antiparticles carry the opposite charge  $+1$  and their corresponding neutrino flavors carrying no charge (charge neutral).

Charged leptons interact with all fundamental forces except the strong force, while neutrinos only interact with the weak force and gravity.

## Bosons

The SM classify bosons into two types: one scalar boson with spin 0 known as the Higgs ( $H$ ) boson, and vector gauge bosons with spin 1 known as gluons ( $g$ ), photon ( $\gamma$ ),  $W^\pm$  and  $Z$  bosons.



247 The gluons and photon are massless, while the  $W^\pm$ ,  $Z$  and  $H$  are massive.  
 248 Each vector gauge boson serves as the mediator for a fundamental force described by the  
 249 SM.  
 250 Gluons are massless mediator particles for the strong interaction between quarks according  
 251 to quantum chromodynamics (QCD), and carry the color charge in a strong interaction.  
 252 Each gluon carries a non-neutral color charge out of eight linearly independent color states  
 253 in the gluon color octet.  
 254 Photon is the massless and charge-neutral mediator particle for the electromagnetic interac-  
 255 tion following quantum electrodynamics (QED).  
 256 The  $W^\pm$  and  $Z$  bosons are massive mediator particles for the weak interaction, with the  
 257  $W^\pm$  boson carrying an electric charge of  $\pm 1$  while the  $Z$  boson is charge neutral.  
 258 Other than the vector gauge boson, the only scalar boson in the SM is the Higgs boson which  
 259 is massive with electric charge of 0.  
 260 The Higgs boson does not mediate a fundamental force like vector bosons, but serve to  
 261 provide the rest mass for all massive elementary particles in the SM through the Higgs  
 262 mechanism.

### 263 **2.1.2 Mathematical formalism**

264 The SM can be described within the formalism of quantum field theory (QFT) with a  
 265 composite Lagrangian

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{QCD}} + \underbrace{(\mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{fermion}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}})}_{\mathcal{L}_{\text{EW}}} \quad (2.1)$$

266 where  $\mathcal{L}_{\text{QCD}}$  is the QCD term and  $\mathcal{L}_{\text{EW}}$  is the electroweak (EW) term of the Lagrangian.  
 267 The SM Lagrangian is gauge invariant under global Poincaré symmetry and local  $SU(3)_C \times$   
 268  $SU(2)_L \times U(1)_Y$  gauge symmetry, with the gauge term  $SU(3)_C$  corresponding to the strong  
 269 interaction and  $SU(2)_L \times U(1)_Y$  to the EW interaction.  
 270 Global Poincaré symmetry ensures that  $\mathcal{L}_{\text{SM}}$  satisfies translational symmetry, rotational  
 271 symmetry and Lorentz boost frame invariance. By Noether's theorem, these symmetries  
 272 give rise to conservation of momentum, angular momentum and energy.

## 273 **Quantum chromodynamics**

274 QCD is a non-Abelian gauge theory (Yang-Mills theory) describing the strong interaction  
 275 between quarks in the SM with the gauge group  $SU(3)_C$ , where  $C$  represents conservation  
 276 of color charge under  $SU(3)_C$  symmetry.  
 277 According to QFT, quarks can be treated as excitations of corresponding quark fields  $\psi$ .  
 278 Quark fields are invariant under  $SU(3)_C$  transformation

$$\psi \rightarrow e^{i\theta(x)T_a}\psi \quad (2.2)$$

279 where  $T_a$  are generators of  $SU(3)_C$ , represented as  $T_a = \lambda_a/2$  with  $\lambda_a$  being the eight Gell-  
 280 Mann matrices.

281 The free Dirac Lagrangian

$$\mathcal{L}_0 = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi \quad (2.3)$$

282 is invariant under global  $SU(3)$  symmetry, but not under local  $SU(3)_C$  symmetry. To  
 283 establish invariance under local  $SU(3)_C$  symmetry, the gauge covariant derivative  $D_\mu$  is

284 defined so that

$$D_\mu \psi = (\partial_\mu - ig_s G_\mu^a T_a) \psi, \quad (2.4)$$

285 where  $g_s = \sqrt{4\pi\alpha_s}$  is the QCD coupling constant,  $G_\mu^a(x)$  are the eight gluon fields that  
 286 transform under  $SU(3)_C$  as

$$G_\mu^a \rightarrow e^{iT_a \theta_a(x)} \left( G_\mu^a + \frac{i}{g_s} \partial_\mu \right) e^{-iT_a \theta_a(x)} = G_\mu^a - \frac{1}{g_s} \partial_\mu \theta_a(x) - f_{abc} \theta_b(x) G_\mu^c, \quad (2.5)$$

287 and  $T_a$  are the generators of  $SU(3)_C$  defined as  $T_a = \lambda_a/2$  with  $\lambda_a$  being the eight Gell-Mann  
 288 matrices.

289 Defining the gluon field strength tensor  $G_{\mu\nu}^a$  as

$$G_{\mu\nu}^a \equiv \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_s f^{abc} G_\mu^b G_\nu^c, \quad (2.6)$$

290 where  $f^{abc}$  are the structure constants of  $SU(3)_C$ , the gauge invariant QCD Lagrangian is

$$\mathcal{L}_{\text{QCD}} = \bar{\psi} (i\gamma^\mu D_\mu - m) \psi - \frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu}, \quad (2.7)$$

291 which can be expressed in the form of

$$\mathcal{L}_{\text{QCD}} = \underbrace{-\frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu}}_{\text{gluon kinematics \& self-interaction}} + \underbrace{\bar{\psi} (i\gamma^\mu \partial_\mu - m) \psi}_{\text{quark kinematics}} + \underbrace{\bar{\psi}^i (g_s \gamma^\mu (T_a)_{ij} G_\mu^a) \psi^j}_{\text{quark-gluon interaction}}. \quad (2.8)$$

292 with  $i, j$  being the color indices with integer values from 1 to 3. The noncommutativity of  
 293  $SU(3)_C$  gives rise to an additional term consisting of only gluon fields and gluon-gluon in-  
 294 teractions. Additionally, the Lagrangian also forces gluons to be massless to maintain gauge

295 invariance.

296

## 297 Electroweak theory

298 The electroweak interaction is the unified description of the weak interaction and electro-  
299 magnetism under the  $SU(2)_L \times U(1)_Y$  symmetry group, where  $L$  represents the left-handed  
300 chirality of the weak interaction and  $Y$  represents the weak hypercharge quantum number.  
301 The quantum number associated with the weak chirality is the weak isospin  $I$ . The EW  
302 quantum numbers are connected by the Gell-Mann-Nishijima relation

$$Q = I_3 + Y/2 \quad (2.9)$$

where  $Q$  is the electric charge and  $I_3$  is the third component of weak isospin  $I$ .

Fermions can have either left-handed or right-handed chirality, and can be divided into  
left-handed doublets and right-handed singlets

$$\psi_L = \begin{pmatrix} \nu_e \\ e_L \end{pmatrix}, \begin{pmatrix} \nu_\mu \\ \mu_L \end{pmatrix}, \begin{pmatrix} \nu_\tau \\ \tau_L \end{pmatrix}, \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \begin{pmatrix} c_L \\ s_L \end{pmatrix}, \begin{pmatrix} t_L \\ b_L \end{pmatrix} \quad (2.10)$$

$$\psi_R = e_R, \mu_R, \tau_R, u_R, d_R, c_R, s_R, t_R, b_R, \quad (2.11)$$

303 with the exception of neutrino which can only have left-handed chirality in the SM.

304 Both left-handed and right-handed fermion fields are invariant under  $U(1)_Y$  transformation

$$\psi \rightarrow e^{iY\theta(x)/2}\psi. \quad (2.12)$$

305 Similar to QCD, to establish invariance under local  $U(1)_Y$  symmetry, the  $U(1)_Y$  gauge  
 306 covariant derivative  $D_\mu$  is defined as

$$D_\mu \phi = \left( \partial_\mu - ig' \frac{Y}{2} B_\mu \right) \psi \quad (2.13)$$

307 where  $B_\mu(x)$  is a vector gauge field that transforms under  $U(1)_Y$  as

$$B_\mu \rightarrow B_\mu + \frac{1}{g'} \partial_\mu \theta(x) \quad (2.14)$$

and  $g'$  is the  $B_\mu$  coupling constant.

Right-handed fermion singlets are not affected by  $SU(2)_L$  transformation, so fermion fields transform under  $SU(2)_L$  as

$$\psi_L \rightarrow e^{iI_3 \vec{\theta}(x) \cdot \vec{\sigma}/2} \psi_L \quad (2.15)$$

$$\psi_R \rightarrow \psi_R. \quad (2.16)$$

308 where  $\vec{\sigma}/2$  are generators of  $SU(2)_L$  and  $\vec{\sigma}$  are Pauli matrices. In order to preserve local  
 309 symmetry, the gauge covariant derivative for  $SU(2)_L$  is defined as

$$D_\mu \psi_L = \left( \partial_\mu - ig \frac{\sigma_i}{2} W_\mu^i \right) \psi_L \quad (2.17)$$

310 where  $W_\mu^i(x)$  ( $i = 1, 2, 3$ ) are three boson gauge fields that transform under  $SU(2)_L$  as

$$W_\mu^i \rightarrow e^{i \frac{\sigma_i}{2} \theta_i(x)} \left( W_\mu^i + \frac{i}{g} \partial_\mu \theta_i(x) \right) e^{-i \frac{\sigma_i}{2} \theta_i(x)} = W_\mu^i + \frac{2}{g} \partial_\mu \theta_i(x) + \epsilon^{ijk} \theta_j(x) W_\mu^k, \quad (2.18)$$

with  $g$  as the gauge coupling constant for  $W_\mu^i$ , and  $\epsilon^{ijk}$  as the structure constant for  $SU(2)_L$ .

The gauge covariant derivative for  $SU(2)_L \times U(1)_Y$  can then be written as

$$D_\mu \psi_L = \left( \partial_\mu - ig' \frac{Y_L}{2} B_\mu - ig \frac{\sigma_i}{2} W_\mu^i \right) \psi_L \quad (2.19)$$

$$D_\mu \psi_R = \left( \partial_\mu - ig' \frac{Y_R}{2} B_\mu \right) \psi_R. \quad (2.20)$$

Similar to QCD, the kinetic term is added by defining field strengths for the four gauge fields

$$B_{\mu\nu} \equiv \partial_\mu B_\nu - \partial_\nu B_\mu \quad (2.21)$$

$$W_{\mu\nu}^i \equiv \partial_\mu W_\nu^i - \partial_\nu W_\mu^i - ge^{ijk} W_\mu^j W_\nu^k. \quad (2.22)$$

The local  $SU(2)_L \times U(1)_Y$  invariant EW Lagrangian can then be expressed as

$$\mathcal{L}_{\text{EW}} = i\bar{\psi}(\gamma^\mu D_\mu)\psi - \frac{1}{4}W_{\mu\nu}^i W_i^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \quad (2.23)$$

$$= \underbrace{i\bar{\psi}(\gamma^\mu \partial_\mu)\psi}_{\text{fermion kinematics}} - \underbrace{\bar{\psi} \left( \gamma^\mu g' \frac{Y}{2} B_\mu \right) \psi - \bar{\psi}_L \left( \gamma^\mu g \frac{\sigma_i}{2} W_\mu^i \right) \psi_L}_{\text{fermion-gauge boson interaction}} - \underbrace{\frac{1}{4}W_{\mu\nu}^i W_i^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}}_{\text{boson kinematics \& self-interaction}}. \quad (2.24)$$

The SM EW bosons can be extracted from  $\mathcal{L}_{\text{EW}}$  by reparameterizing the gauge fields  $B_\mu$  and  $W_\mu^i$  as

$$W^\pm \equiv \frac{1}{\sqrt{2}} \left( W_\mu^1 \mp iW_\mu^2 \right) \quad (2.25)$$

$$\begin{pmatrix} A \\ Z^0 \end{pmatrix} \equiv \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix} \quad (2.26)$$

311 where  $\theta_W \equiv \cos^{-1} \left( g / \sqrt{g^2 + g'^2} \right)$  is the weak mixing angle. The EW Lagrangian can then  
 312 be rewritten as

$$\begin{aligned}
 \mathcal{L}_{\text{EW}} = & \underbrace{e A_\mu \bar{\psi} (\gamma^\mu Q) \psi}_{\text{electromagnetism}} + \underbrace{\frac{e}{2 \sin \theta_W \cos \theta_W} \bar{\psi} \gamma^\mu (v_f - a_f \gamma_5) \psi Z_\mu}_{\text{neutral current interaction}} \\
 & + \underbrace{\frac{g}{2\sqrt{2}} \sum_{\psi_L} [\bar{f}_2 \gamma^\mu (1 - \gamma_5) f_1 W_\mu^+ + \bar{f}_1 \gamma^\mu (1 - \gamma_5) f_2 W_\mu^-]}_{\text{charged current interaction}}
 \end{aligned} \tag{2.27}$$

313 where  $a_f = I_3$ ,  $v_f = I_3(1 - 4|Q| \sin^2 \theta_W)$  and  $f_1, f_2$  are up and down type fermions of a  
 314 left-handed doublet.

## 315 **Higgs mechanism**

316 fermions & bosons still massless from previous section, resolved by introduction of the Higgs  
 317 mechanism

318 (show Higgs field, potential & Lagrangian)

319 (show minimum of Higgs potential aka VEV)

320 —————-[continue later]—————

### 321 **2.1.3 Four-top quark production**

322 - Top: heaviest particle, strong coupling to many BSM particles in BSM models.

323 - 4top: xsec relevant to and enhanced by many BSM models

324 - Predicted by SM and observed [observation paper]

325 - Predicted xsec and observed xsec

326 - (insert Feynman diagrams)

327 - Decay products & final state topologies

## 328 **2.2 Beyond the Standard Model**

### 329 **Top-philic vector resonance**

330 - (briefly introduce composite pseudo-nambu-Goldstone boson and motivation)

331 - hypothesis: top quark large mass results from high mixing between a "true" top quark and  
332 a colored, fermionic composite state

333 - composite vector resonance can be modeled as a top-philic  $Z'$  boson (without QCD color)  
334 or top-philic KK-gluon (with QCD color)

335 - color singlet vector boson ( $Z'$ ) model coupling strongly to top and weakly or not at all to  
336 others

337 - (show Lagrangian for interaction)

338 - two body decay  $Z'$  into  $t\bar{t}$  with  $m_{Z'}$  in TeV range  $\rightarrow$  top mass

339 - decay channels:  $t\bar{t}Z'$  s & t channels,  $tWZ'$ ,  $tjZ'$

340 - (show decay width at LO)

341 - (Feynman diagrams here)

### 342 **Higgs-top Yukawa coupling**

343 (show Lagrangian of Higgs-top Yukawa coupling)

344 (show dependence of  $t\bar{t}t\bar{t}$  xsec on Yukawa coupling at LO)



# Chapter 3. LHC & ATLAS Experiment

## 3.1 The Large Hadron Collider

theoretical predictions are tested with experimental data obtained from particle accelerators world's largest accelerator built by CERN situated on the border of Switzerland and France has been operating since xxxx lifetime divided into 3 runs, currently on Run 3 with planned upgrades on the horizon responsible for a number of discoveries aka Higgs, etc.

### 3.1.1 Overview

[Basic info: location, size, main working mechanism, main detectors, main physics done]  
- 27 km circumference, reusing LEP tunnels 175 m below ground level  
- 7-13-13.6 TeV center of mass energies for pp collisions  
- other than pp, also collides pPb, PbPb at 4 points with 4 main detectors: ATLAS, CMS (general purpose detectors), ALICE (heavy ion physics, ion collisions), LHCb (*b*-physics)

### 3.1.2 LHC operations

- focuses mainly on pp collisions for this thesis - beams split into bunches of  $1.1 \times 10^{11}$  protons with instantaneous luminosity of up to  $2 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$   
- beam energies ramp up in other accelerators before injection, full ramp up to 6.5 GeV about 20 minutes  
(insert full diagram of accelerator chain)  
Linac 4: hydrogen atoms, accelerated up to 160 MeV

365 PSB: H atoms stripped of electrons before injection, accelerated to 2 GeV  
366 PS: 26 GeV, SPS: 450 GeV  
367 LHC: injection in opposite directions, 6.5 TeV per beam  
368  
369 Run 1: 2010-2012, Run 2: 2015-2018, Run 3: 2022-2025, HL-LHC: 2029-?  
370 COM energies: 7 & 8 TeV, 13 TeV, 13.6 TeV, 13.6 & 14 TeV  
371 inbetween periods: long shutdowns (LS1, LS2, LS3)  
372 (add HL-LHC timeline graph)  
373 (insert LHC SM processes cross sections chart)

## 374 **Top quark production at the LHC**

375 history (CDF/D0)  
376 LHC as a top factory: show luminosity and cross section for top processes  
377 couples to Higgs as heaviest elementary particle  
378 Higgs produced mainly from ggH (90%) via top loop and from ttH  
379 (Feynman diagram of related processes)

## 380 **3.2 The ATLAS detector**

381 multipurpose particle detector with a symmetric cylindrical geometry and a solid angle  
382 coverage of almost  $4\pi$   
383 44m long, 25m diameter  
384 inner detector, solenoid/toroid magnet, EM & hadronic calorimeters, muon spectrometer  
385 (insert figure)

right-handed cylindrical system, z-axis follows beamline, azimuthal and polar (0 in the beam direction) angles measured with respect to beam axis.

pseudorapidity  $\eta = -\ln \tan(\theta/2)$ , approaches  $\pm \infty$  along and 0 orthogonal to the beamline

distance  $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$

transverse energy  $E_T = \sqrt{p_T^2 + m^2}$

transverse momentum  $p_T$  component of momentum orthogonal to the beam axis  $p_T =$

$\sqrt{p_x^2 + p_y^2}$

### 3.2.1 Inner detector

- measures tracks of charged particles with high momentum resolution ( $\sigma_{p_T}/p_T = 0.05\% \pm 1\%$ )

- covers particles with  $p_T > 0.5$  GeV,  $|\eta| < 2.5$

pixel detector -i semiconductor tracker -i transition radiation tracker, innermost to outermost

- pixel detector:

- innermost, 250  $\mu\text{m}$  silicon pixel layers

- detects charged particles from electron-hole pair production in silicon

- measures impact parameter resolution & vertex identification for reconstruction of short-lived particles

- spatial resolution of 10  $\mu\text{m}$  in the  $R - \phi$  plane and 115  $\mu\text{m}$  in the z-direction

- 80.4m readout channels

- sct:

- surrounds pixel detector, silicon microstrip layers with 80  $\mu\text{m}$  strip pitch
- particle tracks cross 8 strip layers
- measures particle momentum, impact parameters, vertex position
- spatial resolution of 17  $\mu\text{m}$  in the  $R - \phi$  plane and 580  $\mu\text{m}$  in the z-direction
- 6.3m readout channels.

• trt:

- outermost, layers of 4 mm diameter gaseous straw tubes with transition radiation material (70%  $Xe + 27\% CO_2 + 3\% O_2$ ) & 30  $\mu\text{m}$  gold-plated wire in the center
- tubes 144 cm length in barrel region ( $|\eta| < 1$ ), 37 cm in the endcap region ( $1 < |\eta| < 2$ ), arranged in wheels instead of parallel to beamline)
- gas mixture produces transition radiation when ionized for electron identification
- resolution/accuracy of 130  $\mu\text{m}$  for each straw tube in the  $R - \phi$  plane
- 351k readout channels

### 3.2.2 Calorimeter systems

surrounds the inner detector & solenoid magnet, covers  $|\eta| < 4.9$  and full  $\phi$  range. Alternates passive and active material layers. Incoming particles passing through calorimeter produce EM cascades or hadronic showers in passive layer. Energies deposited and convert to electric signals in active layers for readout.

EM calorimeter:

- innermost, lead-LAr detector (passive-

428 active) 448

429 • measures EM cascades (bremsstrahlung 449  
430 & pair production) produced by elec- 450  
431 trons/photons 451

432 • divided into barrel region ( $|\eta| < 1.475$ ) 452  
433 & endcap regions ( $1.375 < |\eta| < 3.2$ ) 453  
434 with transition region ( $1.372 < |\eta| <$  454  
435  $1.52$ ) containing extra cooling materi- 455  
436 als for inner detector 456

437 • end-cap divided into outer wheel 457  
438 ( $1.372 < |\eta| < 2.5$ ) & inner wheel 458  
439 ( $2.5 < |\eta| < 3.2$ ) 459

440 • higher granularity in ID ( $|\eta| < 2.5$ ) 460  
441 range for electrons/photons & precision 461  
442 physics, coarser elsewhere for jet recon- 462  
443 struction & MET measurements 463

444 hadronic calorimeter: 464

445 • outermost 465

446 • measures hadronic showers from inelas- 466  
447 tic QCD collisions 467

- thick enough to prevent most particles  
showers from reaching muon spectrom-  
eter
- split into tile calorimeter in barrel re-  
gion ( $|\eta| < 1.0$ ) & extended barrel re-  
gion ( $0.8 < |\eta| < 1.7$ ), LAr hadronic  
end-cap calorimeter (HEC) in end-cap  
regions ( $1.5 < |\eta| < 3.2$ ) & LAr forward  
calorimeters (FCal) in  $3.1 < |\eta| < 4.9$   
range.
  - tile calorimeters: steel-plastic  
scintillating tiles, readout via pho-  
tomultiplier tubes
  - hec: behind tile calorimeters, 2  
wheels per end-cap. copper plates-  
LAr. overlap with other calorime-  
ter systems to cover for gaps be-  
tween subsystems
  - fcal: 1 copper module & 2 tung-  
sten modules-LAr. copper op-  
timized for EM measurements,  
tungsten for hadronic.

### 3.2.3 Muon spectrometer

- ATLAS outermost layer. measures muon momenta & charge in range  $|\eta| < 2.7$
- momentum measured by deflection in track from toroid magnets producing magnetic field orthogonal to muon trajectory
  - large barrel toroids in  $|\eta| < 1.4$ , strength 0.5 T
  - 2 smaller end-cap toroids in  $1.6 < |\eta| < 2.7$ , strength 1 T
  - transition region  $1.4 < |\eta| < 1.6$ , deflection provided by a combination of barrel and end-cap magnets
- chambers installed in 3 cylindrical layers, around the beam axis in barrel region & in planes perpendicular to beam axis in the transition and end-cap regions
- split into high-precision tracking chambers (monitored drift tubes & cathode strip chambers) & trigger chambers (resistive plate chambers & thin gap chambers)
- trigger chambers provide fast muon multiplicity & approximate energy range information with L1 trigger logic
  - mdt:
    - \* range  $|\eta| < 2.7$ , innermost layer  $|\eta| < 2.0$
    - \* precision momentum measurement
    - \* layers of 30 mm drift tubes filled with 93% *Ar* & 7% *CO*<sub>2</sub>, with a 50  $\mu$ m gold-plated tungsten-rhenium wire at the center
    - \* muons pass through tube, ionizing gas and providing signals. Combining signals from tubes

496 forms track 517 – rpc:  
 497 \* maximumn drift time from wall 518 \* range  $|\eta| < 1.05$   
 498 to wire 700 ns 519 \* provide fast meas  
 499 \* resolution: 35  $\mu\text{m}$  per chamber, 520 – tgc:  
 500 80  $\mu\text{m}$  per tube 521 \* range  $1.05 < |\eta| < 2.7$   
 501 – csc:  
 502 \* forward region  $2.0 < |\eta| < 2.7$ ,  
 503 highest particle flux and density  
 504 region  
 505 \* multiwire proportional chambers  
 506 with higher granularity, filled  
 507 with 80% *Ar* & 20% *CO*<sub>2</sub>  
 508 \* shorter drift time than MDT,  
 509 plus other features making CSC  
 510 suitable for high particle den-  
 511 sities and consequently able to  
 512 handle background conditions  
 513 \* resolution: 40  $\mu\text{m}$  in bending  $\eta$ -  
 514 plane, 5 mm in nonbending  $\phi$ -  
 515 plane due to coarser cathode seg-  
 516 mentation, per CSC plane

### 3.2.4 Forward detectors

- LUCID (LUMinosity measurement using Cherenkov Integrating Detector):  $\pm 17$  m from interaction point, measures luminosity using  $pp$  scattering in the forward region
- ALFA (Absolute Luminosity for ATLAS):  $\pm 240$  m, measures  $pp$  scattering at small angles
- ZDC (Zero-Degree Calorimeter):  $\pm 140$  m, measures centrality in heavy-ion collisions

### 3.2.5 Magnetic systems

superconducting solenoid & toroid magnets cooled to 4.5 K with liquid helium

solenoid: 2.56 m diameter, 5.8 m length, 2 T strength axial magnetic field, encloses inner detector

toroid = barrel + endcap toroid x2

barrel toroid: 9.2/20.1 m inner/outer diameter, 25.3 m length, 0.5 T strength

endcap toroid: 1.65/10.7 m inner/outer diameter, 5 m length, 1 T strength

(show magnet system diagram)

### 3.2.6 Trigger & data acquisition

LHC produces large amount of data (40 MHz with 25 ns bunch crossing), necessitates a way

to filter out trash from interesting events

handles online processing, selecting and recording interesting events for further offline pro-

cessing and more in-depth analyses



- Level-1 (L1) trigger: online, fast hardware-based trigger, reduces to 100 kHz
  - L1 calorimeter triggers (L1Calo): selects high energy objects & MET
  - L1 muon triggers (L1Muon): selects using hit information from RPC & TGC
  - L1 topological trigger (L1Topo): select based on topological selection synthesized using information from L1Calo & L1Muon
  - Central Trigger Processor (CTP): uses L1Calo/Muon/Topo for final L1 trigger decision within  $2.5 \mu\text{s}$  latency. Also identify regions of interest in  $\eta$  and  $\phi$  to be processed directly by HLT
- L1 trigger information read out by Front-End (FE) detector electronics then sent to ReadOut Drivers (ROD) for preprocessing and subsequently to ReadOut System (ROS) to buffer
- High-Level Trigger (HLT): offline, software-based trigger, using dedicated algorithms and L1 output as input, reduces to 1 kHz
- Send to storage for analyses after HLT

overall trigger process reduces original collision data rate by a factor of about 10000 after HLT

(show TDAQ diagram)

# Chapter 4. Data & Simulated Samples

## 4.1 Data samples

LHC Run 2 data collected at  $\sqrt{s} = 13$  TeV between 2015-2018

luminosity  $140 \text{ fb}^{-1}$

(include uncertainty for Run 2 only)

Triggers used:

Table 4.1: Caption

Trigger	Data period			
	2015	2016	2017	2018
Single electron triggers				
HLT_e24_lhmedium_L1EM20VH	✓	-	-	-
HLT_e60_lhmedium	✓	-	-	-
HLT_e120_lhloose	✓	-	-	-
HLT_e26_lhtight_nod0_ivarloose	-	✓	✓	✓
HLT_e60_lhmedium_nod0	-	✓	✓	✓
HLT_e140_lhloose_nod0	-	✓	✓	✓
Di-electron triggers				
HLT_2e12_lhloose_L12EM10VH	✓	-	-	-
HLT_2e17_lhvloose_nod0	-	✓	-	-
HLT_2e24_lhvloose_nod0	-	-	✓	✓
HLT_2e17_lhvloose_nod0_L12EM15VHI	-	-	-	✓
Single muon trigger				
HLT_mu20_iloose_L1MU15	✓	-	-	-
HLT_mu40	✓	-	-	-
HLT_mu26_ivarmedium	-	✓	✓	✓
HLT_mu50	-	✓	✓	✓

## 4.2 Monte Carlo samples

Monte Carlo simulated samples are used to estimate signal acceptance before unblinding, profile the physics background for the analysis and to study object optimizations. Simulated samples for this analysis use are generated from ATLAS' generalized MC20a/d/e samples for Run 2, using full detector simulation (FS) based on Geant4.

### 4.2.1 $t\bar{t}Z'$ signal samples

Run 2  $t\bar{t}Z'$  sample

samples: 6 samples for each mass point from [1, 1.25, 1.5, 2, 2.5, 3] TeV

generator: MADGRAPH5\_AMC@NLO v.2.8.1p3.atlas9 at LO with NNPDF3.1LO pdf

event: PYTHIA8 [v.244p3.rangefix] using A14 tune & NNPDF2.3LO pdf

parameters:

- chirality  $\theta$ : does not affect the strong production mode for  $t\bar{t}Z'$ , therefore picking default value  $\pi/4$  to suppress loop-production of the  $Z'$  resonance
- top- $Z'$  coupling  $c_t = 1$

resonance width computed with MADGRAPH5\_AMC@NLO to be 4% of model configuration with these parameters

### 4.2.2 Background samples

Run 2 mc20 samples (2015-2018)

(explain most important backgrounds:  $t\bar{t}t\bar{t}$  &  $t\bar{t}t$ ,  $t\bar{t}V$ ,

Table 4.2: Summary of all Monte-Carlo samples used in this analysis.

Process	ME Generator	ME Order	ME PDF	PS	Tune	Sim.
<b>Signals</b>						
$t\bar{t}Z'/tjZ'/tWZ'$	MADGRAPH5_AMC@NLO	LO	NNPDF3.1LO	PYTHIA8	A14	FS
$t\bar{t}t\bar{t}$ and $t\bar{t}t$						
$t\bar{t}t\bar{t}$	MADGRAPH5_AMC@NLO	NLO	NNPDF3.0nlo	PYTHIA8	A14	AF3
	MADGRAPH5_AMC@NLO	NLO	MMHT2014LO	HERWIG7 H7-UE-MMHT		AF3
	SHERPA	NLO	NNPDF3.0 NNLO	HERWIG7 SHERPA		FS
$t\bar{t}t$	MADGRAPH5_AMC@NLO	LO	NNPDF2.31o	PYTHIA8	A14	AF3
$t\bar{t}V$						
$t\bar{t}H$	POWHEGBOX v2	NLO	NNPDF3.0nlo (mc20)	PYTHIA8	A14	FS
			PDF4LHC21 (mc23)			AF3
	POWHEGBOX v2	NLO	NNPDF3.0nlo	HERWIG7 H7.2-Default		FS
$t\bar{t}Z$	MADGRAPH5_AMC@NLO	NLO	NNPDF3.0 NNLO	PYTHIA8	A14	FS
$t\bar{t}W$	SHERPA	NLO	NNPDF3.0 NNLO	SHERPA	SHERPA	FS
	SHERPA	LO	NNPDF3.0 NNLO	SHERPA	SHERPA	FS
$t\bar{t}$ and Single-Top						
$t\bar{t}$	POWHEGBOX v2	NLO	NNPDF3.0nlo	PYTHIA8	A14	FS
$tW$	POWHEGBOX v2	NLO	NNPDF3.0nlo	PYTHIA8	A14	FS
$t(q)b$	POWHEGBOX v2	NLO	NNPDF3.0nlo (s)	PYTHIA8	A14	FS
			NNPDF3.0nlo 4f (t)			FS
$tWZ$	MADGRAPH5_AMC@NLO	NLO	NNPDF3.0nlo	PYTHIA8	A14	FS
$tZ$	MADGRAPH5_AMC@NLO	LO	NNPDF3.0nlo 4f	PYTHIA8	A14	FS
$t\bar{t}VV$						
$t\bar{t}WW$	MADGRAPH5_AMC@NLO	LO	NNPDF3.0nlo	PYTHIA8	A14	FS
	(mc20)					
	MADGRAPH (mc23)					FS
$t\bar{t}WZ$	MADGRAPH	LO	NNPDF3.0nlo	PYTHIA8	A14	AF3
$t\bar{t}HH$	MADGRAPH	LO	NNPDF3.0nlo	PYTHIA8	A14	AF3
$t\bar{t}WH$	MADGRAPH	LO	NNPDF3.0nlo	PYTHIA8	A14	AF3
$t\bar{t}ZZ$	MADGRAPH	LO	NNPDF3.0nlo	PYTHIA8	A14	AF3
$V(VV)+\text{jets}$ and $VH$						
$V+\text{jets}$	SHERPA	NLO	NNPDF3.0 NNLO	SHERPA	SHERPA	FS
$VV+\text{jets}$	SHERPA	NLO	NNPDF3.0 NNLO	SHERPA	SHERPA	FS
		LO ( $gg \rightarrow VV$ )				FS
$VVV+\text{jets}$	SHERPA	NLO	NNPDF3.0 NNLO	SHERPA	SHERPA	FS
$VH$	POWHEGBOX v2	NLO	NNPDF3.0 AZNLO	PYTHIA8	A14	FS

# Chapter 5. Particle Reconstruction & Identification

Reconstruction software reconstructs basic objects from signals collected from the event: interaction vertices, tracks, topological clusters of energy deposits. These quantities then used to reconstruct physics objects i.e. particles (electron, muon), jets, MET

## 5.1 Primary reconstruction

### 5.1.1 Topological clusters

[8][9]

Topological cluster (topo-cluster): Clusters of topologically connected cell signals in the calorimeter at the EM scale. This scale does not consider loss of signal from hadrons. Singular hits without hits from neighboring cells are considered noise.

Done in an effort to extract signal while minimizing electronic effects and physical fluctuations. Used to reconstruct hadronic objects and particles decaying hadronically i.e.  $\tau$  leptons. Signal hits with significance above a cell signal significance level  $\varsigma_{\text{cell}}^{\text{EM}}$  are seeded in as part of a proto-cluster. Neighboring cells satisfying a cluster growth threshold are collected into the cluster.

Two clusters are merged if a cell is matched to both

If a cluster has two or more local signal maxima satisfying  $E_{\text{cell}}^{\text{EM}} > 500$  MeV, the cluster is split accordingly.

The process continues iteratively until all cells with significant signal efficiency have been matched to a cluster.

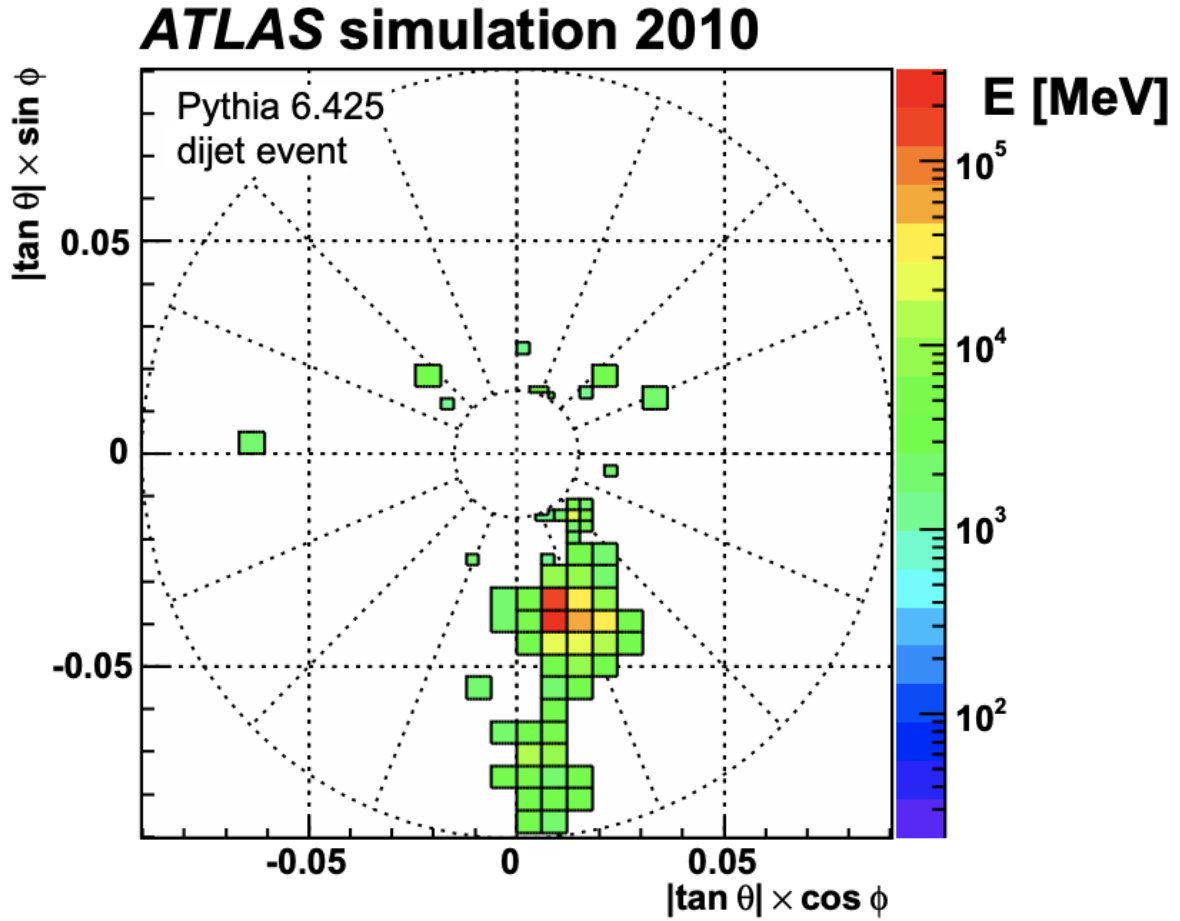


Figure 5.1

### 5.1.2 Tracks

[10]

Charged particles deposit energy in different layers of the inner detector and muon spectrometer

ID reco software: inside-out and outside-in algorithms

- Inside-out: [11]

Starts with seeded hits in the silicon detector in pixel & SCT

Loosely matched to an EM cluster to form a track candidate

Hits are added to form a track candidate using a pattern recognition algorithm based on a Kalman filter formalism [12]

Track candidates are then fitted with a  $\chi^2$  filter [13] and loosely matched to a fixed-sized EM cluster. Successfully matched track candidates are re-fitted with a Gaussian-sum filter (GSF) [14]

This is followed by a track scoring strategy to resolve fake tracks & hit ambiguity between different tracks [15]

Extend to TRT to form final tracks, filtered by threshold  $p_T > 400$  MeV.

- Outside-in: [16]

Reverse, starts with segments in TRT extending inward to silicon hits in pixel & SCT

Targeting secondary tracks (decays/interactions of primary particles) or long-lived particles

### 5.1.3 Vertices

Vertices: interaction or decay point

Primary vertex: pp interaction point

Important for reconstruction of the hard scattering pp interaction, resulting trajectories and

kinematic information of the event

- Vertex finding:

Uses the z-position of a track as input

Vertices require to have at least 2 tracks

Iterative  $\chi^2$  algorithm evaluate track-vertex compatibility, using the track as new seed for another vertex if large discrepancy

- Vertex fitting:

Adaptive multi-vertex fitter (AVF) algorithm assigns weights that depend on the track-vertex compatibility to each track to measure the probability of the track being an outlier vs inlier.

Vertex is then estimated by iteratively minimizing an objective function of these weights

## 5.2 Jets

- Quarks, gluons & other non-color-neutral hadrons cannot be observed individually due to QCD color confinement

- A non-color-neutral hadron will almost immediately undergo hadronization producing a cone of color-neutral hadrons also known as a jet

- Jet signals can be used to reconstruct and consequently indirectly observe the original quarks/gluons the jets originated from

- Jet reconstruction:



- PFlow: energy deposited in the calorimeter systems by charged particles is removed and replaced by particle objects created with the remaining energy in the calorimeter and tracks matched to the topo-clusters. (include PFlow graphics)
- anti- $k_t$  algorithms: sequential recombination jet algorithms
- pile-up jets: multiple interactions associated with one bunch crossing in addition to the hard scattering of interest and reconstructed as jets in the final states. Reconstructed pile-up jets can result from Pile-up jets are usually from soft interactions and can be distinguished with JVT algorithm using tracking information from the ID.
- JES/JER calibration: Jet reconstruction at EM scale does not accurately account for energy from QCD interactions and needs to be calibrated to jets reconstructed at particle level. This is done via a MC-based JES calibration sequence and additional JER calibration to match jet resolution in simulation to data using dijet events. For this analysis, jets are reconstructed using PFlow method with anti- $k_t$  algorithm, using radius parameter  $\Delta R = 0.4$ . JVT applied to reconstructed jets with  $p_T < 60$  GeV and  $|\eta| < 2.4$ .

### 5.2.1 Flavor tagging

- Classification of hadronic jets is an important task for many LHC analyses especially ones studying final states (Higgs decay/4top)
- Flavor tagging is namely interested in identifying jets containing  $b$ -hadrons,  $c$ -hadrons,  $uds$ -hadrons (light-jets), and hadronic decays from  $\tau$ .
- Of these, identifying  $b$ -jets is of particular interest due to their characteristically long

lifetime ( $\approx 1.5$  ps) from decay suppression by CKM factor, with a displaced secondary decay vertex and usually a tertiary vertex from  $c$ -hadron decays.

## Efficiency calibration

- [17]

- Performance of  $b$ -taggers are studied on MC simulated samples. However, the  $b$ -tagging efficiency predicted by simulation  $\varepsilon_b^{\text{sim}}$  is usually not the same as the efficiency measured in data  $\varepsilon_b^{\text{data}}$ .

- The correction for the rate of events after applying a  $b$ -tagging requirement is calibrated and applied jet by jet in the form of data-to-simulation scale factors  $\text{SF} = \varepsilon_b^{\text{data}} / \varepsilon_b^{\text{sim}}$ .

- Usage of  $b$ -tagger in this analysis is done via five operating points (OPs), corresponding to 60%, 70%, 77%, 85% and 90%  $b$ -jet tagging efficiency  $\varepsilon_b$  in simulated  $t\bar{t}$  events in order from loosest to tightest discriminant cut points. - OPs are defined by selection on the tagger output to provide a pre-defined level of  $\varepsilon_b$ , and act as a variable trade-off between  $b$ -tagging efficiency and  $c$ -/light-jet rejection i.e.  $b$ -jet purity

- A jet is considered  $b$ -tagged if it passes the efficiency criteria for a given OP. A pseudo-continuous  $b$ -tagging (PCBT) score is defined to summarize the OP criteria a jet passes into a variable. The PCBT score can take integer values between 1 and 6, where a score of 6 means a jet passes all four OP thresholds (passing 65% OP), a score of 2 for jets that pass only the 90% OP, and a score of 1 for jets that don't pass any OP. Additionally, PCBT defines a value of -1 for any jet that does not satisfy  $b$ -tagging criteria.

- For this analysis, jets containing  $b$ -hadrons are identified and tagged with the GN2v01 algorithm, described in subsection 5.2.1. A jet is considered  $b$ -tagged if it passes the 85%

WP; this gives the best sensitivity to the signal out of all five possible  $b$ -tagging WPs. The  $b$ -tagged jet is then assigned a PCBT score accordingly.

**bttag optimization table?**

## GN2 $b$ -tagging algorithm

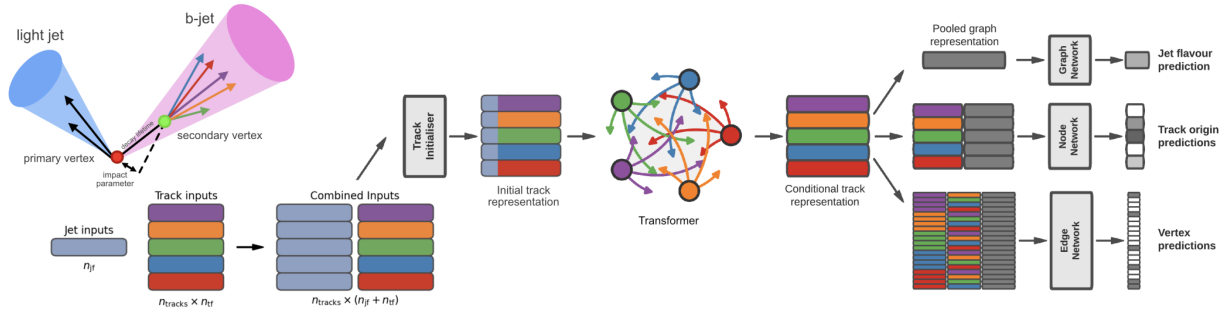


Figure 5.2: [3][4][5]

- GN2 transformer-based  $b$ -tagging algorithm, utilized for analysis of Run 2 and Run 3 data

- GN2 gives a factor of 1.5-4 improvement in experimental applications compared to the previous convolutional neural network-based standard  $b$ -tagging algorithm, DL1d, without dependence on the choice of MC event generator.

- Attention-based architecture, modified to incorporate domain knowledge and additional auxiliary physics objectives: grouping tracks originating from common vertices and prediction of the underlying process for each track

- MC simulated SM  $t\bar{t}$  and BSM  $Z'$  events from  $pp$  collisions were used as training and evaluation samples. In order to minimize bias, both  $b$ - and light-jet samples are re-sampled to match  $c$ -jet distributions.

- GN2 concatenates 2 jet and 19 track reconstruction variables of up to 40 tracks to form

the input feature vector, normalized to zero mean and unit variance.

- The output consists of a jet classification layer of size 4 consisting of  $p_b$ ,  $p_c$ ,  $p_u$  and  $p_\tau$  for the probability of each jet being a  $b$ -,  $c$ -, light- or  $\tau$ -jet respectively; a track-pairing output layer of size 2, and a track origin classification layer of 7 output categories.

## 5.3 Leptons

- Lepton reconstruction is concerned mainly with electron and muon construction, since tau decays quickly and can either be reconstructed using jets or light leptons. From here on out lepton will be used mostly to refer to electrons and muons

- Leptons can be classified into two categories: prompt leptons resulting from heavy particle decays, or non-prompt leptons resulting from detector or reconstruction effects, or from  $b$ - or  $c$ - hadron decays

- Reconstruction of leptons is therefore important to study the underlying physics and suppressing background

### 5.3.1 Electrons

- [6][7]

- Electrons lose energy interacting with the detector materials via bremsstrahlung. The bremsstrahlung photon can then produce an electron-positron pair which can itself leaves signals in the detector, creating a collimated object that can leave multiple tracks in the ID or EM showers in the calorimeter, all considered part of the same EM topo-cluster.

- Electron signal signature has three characteristic components: localized energy deposits in the calorimeter, multiple tracks in the ID and compatibility between the above tracks and

energy clusters in the  $\eta \times \phi$  plane. Electron reconstruction in ATLAS follows these steps accordingly - Seed-cluster reconstruction and track reconstruction are performed sequentially

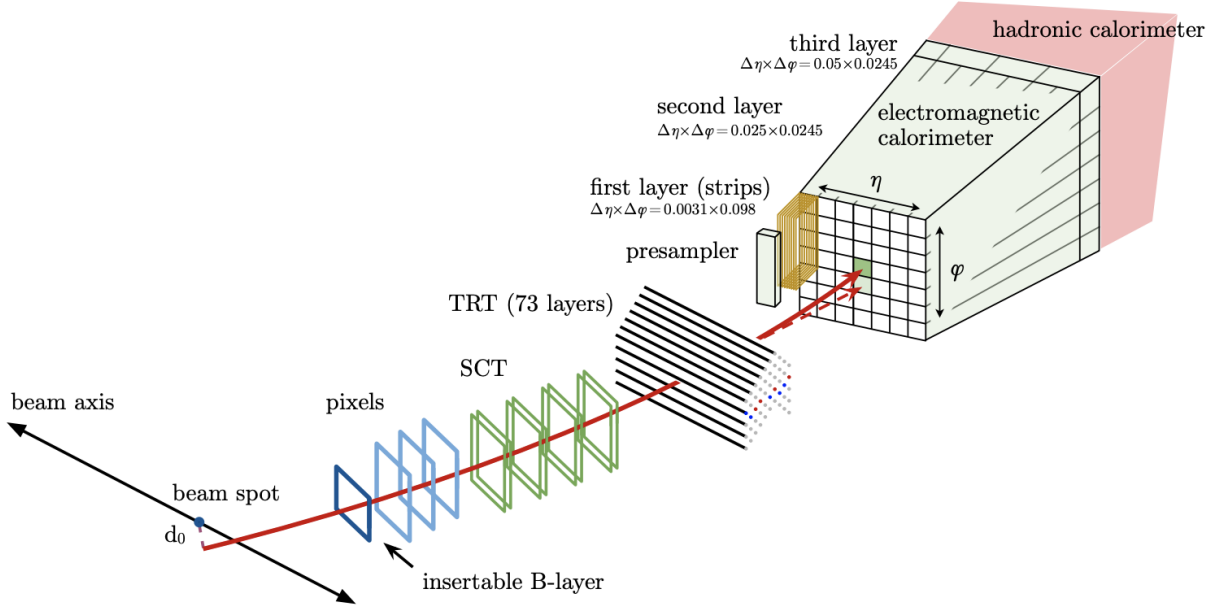


Figure 5.3: [6]

in accordance with the iterative clustering algorithm and track reconstruction method respectively, described in section 5.1

- The seed-cluster and track candidate associated with a conversion vertex are then matched to form an electron candidate.

- A reconstructed cluster is expanded from the seed-cluster in either  $\phi$  or  $\eta$  in the barrel or endcap region respectively

- The cluster energy is then calibrated to compute the original electron energy.

## Electron identification

- Additional likelihood-based identification selections using ID and EM calorimeter information are implemented to further improve the purity of the reconstructed electrons and

photons. These selections also help suppress background from hadronic jet deposits, photon conversions or electrons from heavy-flavor decays.

- Three operating points are defined for physics analyses: Loose, Medium and Tight, optimized for 9 bins in  $|\eta|$  and 12 bins in  $E_T$  with each corresponding to a fixed efficiency requirement for each bin. The target efficiencies for Loose, Medium and Tight start at 93%, 88% and 80% respectively for typical EW processes and increases with  $E_T$ . Similar to  $b$ -tagging OPs, the electron OPs represent a trade-off in signal efficiency and background rejection. The electron efficiency are estimated using tag-and-probe method [6] on samples of  $J/\Psi \rightarrow ee$  and  $Z \rightarrow ee$ [7].

## Electron isolation

- A characteristic distinction between prompt electrons and electrons from background processes is the relative lack of activity in both the ID and calorimeter within an area of  $\Delta\eta \times \Delta\phi$  surrounding the reconstruction candidate

- Electron isolation variables are needed to quantify the amount of activity around the electron candidate.

- Calorimeter-based isolation variables  $E_T^{\text{cone}XX}$  is calculated by first summing the energy of topological clusters with barycenters falling within a cone of radius  $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = XX/100$  around the direction of the electron candidate.

- The final isolation variable is then obtained by subtracting energy at the core of the cone belonging to the candidate electron from the sum, then applying corrections for energy leakage outside of the core and pile-up effects.

- Similar to calorimeter-based variables, track-based isolation variables  $p_T^{\text{varcone}XX}$  are calculated by summing all track  $p_T$  within a cone of variable radius  $\Delta R$  around the electron

candidate, minus the candidate's contribution. The cone radius is variable as a function of  $p_T$

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

with  $\Delta R_{\text{max}}$  being the maximum cone size, to account for the closer proximity of decay products to the electron in high-momentum heavy particle decays.

- Four isolation operating points are implemented to satisfy specific needs by physics analyses: Loose, Tight, HighPtCaloOnly and Gradient. The first three OPs are fixed in isolation variables, while the Gradient OP fixes the isolation efficiency to a  $p_T$  dependent function defined as  $\varepsilon = 0.1143 \times p_T + 92.14\%$  with  $p_T$  in GeV, using  $\Delta R = 0.2$  for calorimeter isolation and  $\Delta R_{\text{max}} = 0.2$  for track isolation[7].

## Electron charge misidentification

[6][18]

Electron charge is determined by the curvature of the associated track. Misidentification of charge can then occur via either an incorrect curvature measurement, or an incorrectly matched track.

The former is more likely for electrons with high  $p_T$  due to the small curvature in track trajectories at such scale, while the latter usually results from bremsstrahlung pair-production, creating additional secondary tracks in the vicinity.

Charge misidentification is a crucial irreducible background for analyses with charge selection criteria, and suppression of this background is assisted via a boosted decision tree discriminant known as the Electron Charge ID Selector (ECIDS) [6]. The addition of ECIDS removed 90% of electrons with incorrect charge while selecting 98% of electrons with correct

charge from electrons in  $Z \rightarrow ee$  events satisfying Medium/Tight identification and Tight isolation criteria.

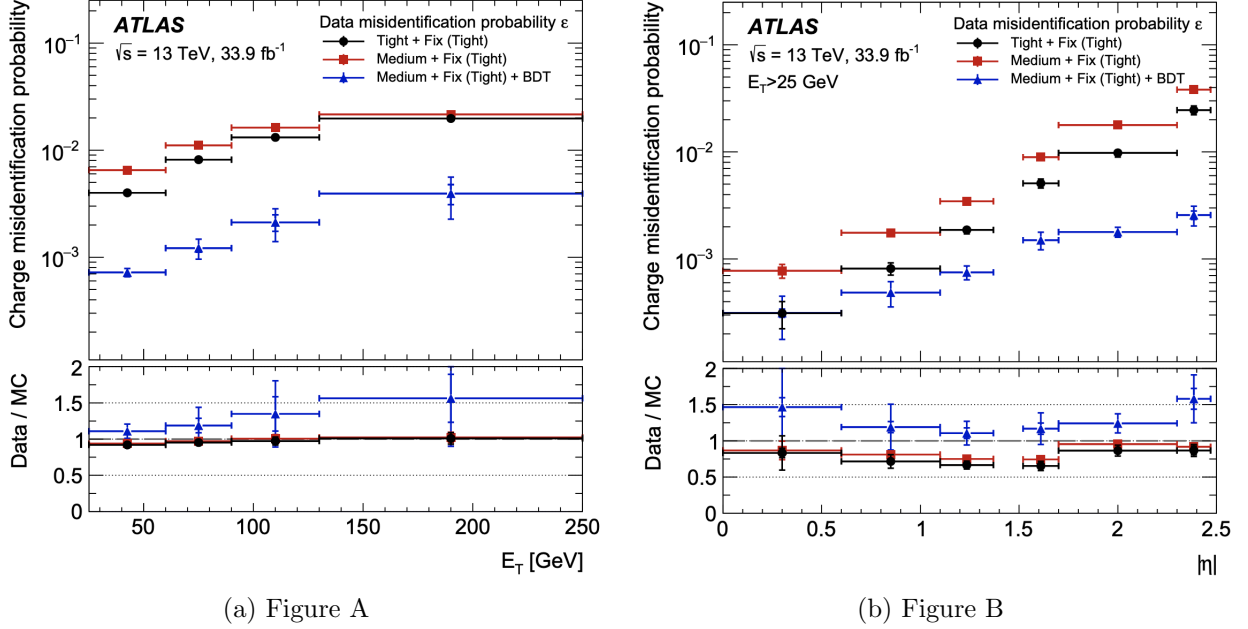


Figure 5.4: [7]

### 5.3.2 Muons

Signature: minimum-ionizing particle leaves tracks in the MS or characteristics energy deposits in the calorimeter

Muons can be reconstructed globally using information from the ID, MS and calorimeters.

Five reconstruction strategies, each corresponding to a muon type:

- Combined (CB): primary ATLAS muon reconstruction method. Muons first reconstructed using MS tracks then extrapolated to include ID tracks (outside-in strategy).

A global combined fit is then performed on both MS and ID tracks

- Inside-out combined (IO): Complementary to CB algorithm. Muon tracks are extrap-



olated from ID to MS, then fitted together with a combined track fit. Useful for muons without good MS information.

- MS extrapolated (ME): ME muons are defined as muons with a MS track that cannot be matched to an ID track using CB method. ME muons allow extension of muon reconstruction acceptance to regions not covered by the ID ( $2.5 < |\eta| < 2.7$ )
- Segment-tagged (ST): ST muons are ID tracks satisfying tight angular matching criteria to at least one reconstructed local segment in the MDT or CSC chambers when extrapolated to the MS. Used primarily when muons only crossed one layer of MS chambers.
- Calorimeter-tagged (CT): CT muons are ID tracks that when extrapolated through the calorimeter, can be matched to energy deposits consistent with those of a minimum-ionizing particle. Extends acceptance range to regions in the MS with sparse instrumentation ( $|\eta| < 0.1$ ), with a higher  $p_T$  threshold of 5 GeV compared to 2 GeV threshold used by other muon reconstruction algorithms due to large background contamination at the low  $p_T$  range of  $15 < p_T < 100$  GeV

## Muon identification

[\[19\]](#)[\[20\]](#)

Reconstructed muons are further filtered by identification criteria to select for high-quality prompt muons for physics analyses. Requirements include number of hits in the MS/ID, track fit properties and compatibility between measurements of the two systems.

Three standard WPs (Loose, Medium, Tight) are defined to better match the needs of different physics analyses concerning prompt muon ID efficiency,  $p_T$  resolution and non-prompt

muon rejection. Of the three, Medium WP is the default ID WP for ATLAS, by virtue of being optimized in efficiency and purity for a wide range of analyses while minimizing non-prompt rejection and systematic uncertainties[19].

## Muon isolation

Muons from heavy particle decays are often produced in an isolated manner compared to muons from semileptonic decays. Muon isolation is therefore an important tool for background rejection in physics analyses

Muon isolation strategies are similar to that of electron in Figure 5.3.1, with track-based and calorimeter-based isolation variables.

Seven isolation WPs are defined to satisfy analyses' needs.

## 5.4 Missing transverse momentum

[21]

Collisions at the LHC happen along the z-axis of the ATLAS coordination system between two particle beam of equal center-of-mass energy. By conservation of momentum, the sum of transverse momenta of outgoing particles should be zero. A discrepancy between measured momentum and zero would then suggest the presence of undetectable particles, which would consist of either SM neutrinos or some unknown BSM particles. This makes missing transverse momentum ( $E_T^{\text{miss}}$ ) an important observable to reconstruct.

Reconstructing  $E_T^{\text{miss}}$  utilizes information from fully reconstructed leptons, photons, jets and other matched track-vertex objects not associated with a prompt object (soft signals),

830 defined with respect to the  $x(y)$ -axis as

$$E_{x(y)}^{\text{miss}} = - \sum_{i \in \{\text{hard objects}\}} p_{x(y),i} - \sum_{j \in \{\text{soft signals}\}} p_{x(y),j}, \quad (5.1)$$

where  $p_{x(y)}$  is the  $x(y)$ -component of  $p_{\text{T}}$  for each particle. The following observables can then be defined:

$$\mathbf{E}_{\text{T}}^{\text{miss}} = (E_x^{\text{miss}}, E_y^{\text{miss}}), \quad (5.2)$$

$$E_{\text{T}}^{\text{miss}} = |\mathbf{E}_{\text{T}}^{\text{miss}}| = \sqrt{(E_x^{\text{miss}})^2 + (E_y^{\text{miss}})^2}, \quad (5.3)$$

$$\phi^{\text{miss}} = \tan^{-1}(E_y^{\text{miss}}/E_x^{\text{miss}}), \quad (5.4)$$

831 where  $E_{\text{T}}^{\text{miss}}$  represents the magnitude of the missing transverse energy vector  $\mathbf{E}_{\text{T}}^{\text{miss}}$ , and  
 832  $\phi^{\text{miss}}$  its direction in the transverse plane. Since physics analyses have differing requirements  
 833 for object selection, the vectorial sum  $\mathbf{E}_{\text{T}}^{\text{miss}}$  can be broken down into

$$\mathbf{E}_{\text{T}}^{\text{miss}} = - \underbrace{\sum_{\text{selected electrons}} \mathbf{p}_{\text{T}}^e - \sum_{\text{selected muons}} \mathbf{p}_{\text{T}}^\mu - \sum_{\text{accepted photons}} \mathbf{p}_{\text{T}}^\gamma - \sum_{\text{accepted } \tau\text{-leptons}} \mathbf{p}_{\text{T}}^\tau - \sum_{\text{accepted jets}} \mathbf{p}_{\text{T}}^{\text{jet}}}_{\text{hard term}} - \underbrace{\sum_{\text{unused tracks}} \mathbf{p}_{\text{T}}^{\text{track}}}_{\text{soft term}}. \quad (5.5)$$

834 Two WPs are defined for  $E_{\text{T}}^{\text{miss}}$ , Loose and Tight [22], with selections on jet  $p_{\text{T}}$  and JVT  
 835 criteria. The Tight WP is used in this analysis, and reduces pileup dependence of  $E_{\text{T}}^{\text{miss}}$  by  
 836 removing the phase space region with more pileup jets than hard-scatter jets, at the expense  
 837 of resolution at low pileup and scale of the reconstructed  $E_{\text{T}}^{\text{miss}}$ .

## 5.5 Overlap removal

Since the reconstruction processes for different objects are performed independently, it is possible for the same detector signals to be used to reconstruct multiple objects. An overlap removal strategy to resolve ambiguities; the overlap removal process for this analysis applies selections listed in Table 5.1 sequentially, from top to bottom.

Table 5.1: [1]

Remove	Keep	Matching criteria
Electron	Electron	Shared ID track, $p_{T,1}^e < p_{T,2}^e$
Muon	Electron	Shared ID track, CT muon
Electron	Muon	Shared ID track
Jet	Electron	$\Delta R < 0.2$
Electron	Jet	$\Delta R < 0.4$
Jet	Muon	$(\Delta R < 0.2 \text{ or ghost-associated}) \ \& \ N_{\text{track}} < 3$
Muon	Jet	$\Delta R < \min(0.4, 0.04 + 10\text{GeV}/p_T^\mu)$

## 5.6 Object definition

Table 5.2 shows the selections used in this analysis. Each selection comes with associated calibration scale factors to account for discrepancies between data and MC simulation, and are applied multiplicatively to the MC event weights.

Table 5.2: Caption

Selection	Electrons	Muons	Jets
$p_T$ [GeV]	$> 15$ $p_T(l_0) > 28$	$> 15$	$> 20$
$ \eta $	$1.52 \leq  \eta  < 2.47$ $< 1.37$	$< 2.5$	$< 2.5$
Identification	TightLH pass ECIDS ( $ee/e\mu$ )	Medium	NNJvt FixedEffPt ( $p_T < 60$ , $ \eta  < 2.4$ )
Isolation	Tight_VarRad	PflowTight_VarRad	
Track-vertex assoc.			
$ d_0^{\text{BL}}(\sigma) $	$< 5$	$< 3$	
$ \Delta z_0^{\text{BL}} \sin \theta $ [mm]	$< 0.5$	$< 0.5$	

## Chapter 6. Analysis Strategy

### 6.1 Event selection

Events for the analysis first are preselected following a list of criteria to optimize for event quality and background rejection.

The criteria are applied sequentially, from top to bottom

1. **Good Run List (GRL):** data events must be part of a predefined list of suitable runs and luminosity blocks.
2. **Calorimeter cleaning:** events containing signal hits indicating an error in the calorimeter are removed.
3. **Primary vertex:** events must have at least one reconstructed vertex matched to 2 or more associated tracks with  $p_T > 500$  MeV.
4. **Trigger:** events must be selected by at least one trigger documented in ??.
5. **Jet cleaning:** events must pass the LooseBad WP for jet cleaning using jets passing preselection criteria in section 5.6. This is done to remove events with significant number of calorimeter hits from non-prompt sources (e.g. instrumental effects, cosmic ray background, non-collision particles)
6. **Bad muon veto:** events are removed if they contain at least one muon before overlap removal with insufficient  $p_T$  resolution.
7. **Kinematic selection:** events must have exactly two Tight leptons with the same electric charge, or at least three Tight leptons of any charge. The leading lepton must

have  $p_T > 28$  GeV, and all leptons must satisfy  $p_T > 15$  GeV.

Events are separated into two channels based on the number of leptons: same-sign di-lepton (SS2L) for events with exactly two leptons of the same charge, or multilepton (ML) for events with three or more leptons. The channels are further separated into regions defined in section 6.2 to prepare for analysis.

Further selections are applied based on the lepton flavors present. In the SS2L channel, if both leptons are electrons, the invariant mass  $m_{ll}$  must satisfy  $m_{ll} < 81$  GeV and  $m_{ll} > 101$  GeV to suppress background involving  $Z$ -bosons. In the ML channel, the same criteria must be satisfied for every opposite-sign same-flavor pair of leptons in an event.

## Event categorization

Simulated events are categorized using truth information of leptons ( $e/\mu$ ) and their originating MC particle (mother-particle).

Each lepton can be classified as either prompt or non-prompt, with non-prompt leptons further categorized for background estimation purposes.

If an event contains only prompt leptons, the event is classified as its corresponding process.

If the event contains one non-prompt lepton, the event is classified as the corresponding type of the non-prompt lepton. If the event contains more than one non-prompt lepton, the event is classified as other.

- **Prompt:** if the lepton originates from  $W/Z/H$  boson decays, or from a mother-particle created by a final state photon.

- **Non-prompt:**

- **Charge-flip ( $e$  only)**: if the reconstructed charge of the lepton differs from that of the first mother-particle.
- **Material conversion ( $e$  only)**: if the lepton originated from a photon conversion and the mother-particle is an isolated prompt photon, non-isolated final state photon, or heavy boson.
- **$\gamma$ -conversion ( $e$  only)**: if the lepton originated from a photon conversion and the mother-particle is a background electron.
- **Heavy flavor decay**: if the lepton originated from a  $b$ - or  $c$ -hadron.
- **Fake**: if the lepton originated from a light- or  $s$ -hadron, or if the truth type of the lepton is hadron.
- **Other**: any lepton that does not belong to one of the above categories.

## 6.2 Analysis regions

Events are selected and categorized into analysis regions belonging to one of two types: control regions (CRs) enriched in background events, and signal regions (SRs) enriched in signal events. This allows for the examination and control of backgrounds and systematic uncertainties, as well as study of signal sensitivities.

The signal is then extracted from the SRs with a profile LH fit using all regions. The full selection criteria for each region are summarized in Table 6.3



### 6.2.1 Signal regions

[include blinding strategy]

- All events selected for SS2L and 3L signal regions must satisfy the following criteria:

- Contains 6 or more jets, with at least 2 jets  $b$ -tagged at the 85% WP
- Scalar sum of the transverse momenta of all leptons and jets  $H_T > 500$  GeV
- Dilepton invariant mass  $m_{\ell\ell}$  does not coincide with the  $Z$ -boson mass range of  $81 - 101$  GeV

- The SR is further granularized by the number of  $b$ -jets and leptons to further study and improve signal sensitivity

Table 6.1: Caption

SR	Selection criteria	
	$b$ -jets	leptons
2b2l	$N_b = 2$	$N_l = 2$
2b3l4l	$N_b = 2$	$N_l \geq 3$
3b2l	$N_b = 3$	$N_l = 2$
3b3l4l	$N_b = 3$	$N_l \geq 3$
4b	$N_b = 4$	

### 6.2.2 Control regions

Control regions are defined for each background to be enriched in the targeted background events, in order to maximize the targeted background's purity and minimize contamination from other sources within the region.

920 This helps to constrain and reduce correlation between background normalization factors.  
 921 Fit variables and selection criteria are determined via optimization studies on CRs to achieve  
 922 the largest discriminating power possible between the target background and other event  
 923 types.

#### 924 $t\bar{t}W$ background CRs

925 Two types of CRs are defined to estimate the flavor composition and normalization of  $t\bar{t}W$   
 926 +jets background: CR  $t\bar{t}W^\pm$ +jets to constrain flavor composition, and CR 1b( $\pm$ ) to con-  
 927 strain jet multiplicity spectrum.

928 These are further split into CR  $t\bar{t}W^\pm$  and CR 1b( $\pm$ ) due to the pronounced asymmetry in  
 929  $t\bar{t}W$  production from  $pp$  collisions, with  $t\bar{t}W^+$  being produced at approximately twice the  
 930 rate of  $t\bar{t}W^-$ . Selections on  $H_T$  and  $N_{\text{jets}}$  to ensure orthogonality to SR  
 931 Selections on total charge for each charged  $W^\pm$  boson

932

#### 933 Fake/non-prompt background CRs

934 Selection for fake/non-prompt CRs are determined using the DFCommonAddAmbiguity (DF-  
 935 CAA) variable for reconstructed leptons.

Table 6.2: Caption

DFCAA	Description
-1	No 2nd track found
0	2nd track found, no conversion found
1	Virtual photon conversion candidate
2	Material conversion candidate

936 Four CRs for three main types of fake/non-prompt backgrounds: virtual photon ( $\gamma^*$ )

conversion, photon conversion in detector material (Mat. Conv.) and heavy flavor decays (HF).

- Low  $m_\gamma^*$ : events with an  $e^+e^-$  pair produced from a virtual photon

Selects two same-sign leptons with at least one electron reconstructed as an internal conversion candidate and neither as with a material conversion candidate ( $\text{DFCAA}_{\ell_1(\ell_2)} = 1$  and  $\neq 2$ )

NF constrained using yield count only.
- Mat. Conv.: events with an electron originating from photon conversion within the detector material.

Selects two same-sign leptons with at least one electron reconstructed as a material conversion candidate ( $\text{DFCAA}_{\ell_1(\ell_2)} = 2$ ).

NF constrained using yield count only.
- HF  $e/\mu$ : events with a reconstructed non-prompt lepton from semi-leptonic decays of  $b$ - and  $c$ -hadrons (heavy flavor decays)

Selects three leptons with at least two electrons/muons, with no lepton reconstructed as a conversion candidate ( $\text{DFCAA} < 0$ ).

NFs constrained by fitting with  $p_T$  of the third leading lepton  $\ell_3$ .

### 6.2.3 Validation regions

In addition, validation regions are also defined to validate the normalization and modeling of  $t\bar{t}Z$  and  $t\bar{t}W$  background without being used in the fit.

- $t\bar{t}Z$ : Selects events with at least two  $b$ -tagged jets, at least four total jets and three leptons with at least one same-flavor opposite-sign lepton pair possessing invariant mass  $m_{\ell\ell}$  within the  $Z$ -boson mass window of  $81 - 101$  GeV
- $t\bar{t}W$ : Main charge asymmetric background leaning  $t\bar{t}W^+$ , validated using the difference in number of positively and negatively charged events  $N_+ - N_-$  instead of total number of events.  
Selects using CR  $t\bar{t}W$  and CR 1b criteria, with one VR not orthogonal to SR and one orthogonal VR with more limited statistics.

### 6.3 Background estimation

Background events in this analysis consist of SM processes that can result in a  $t\bar{t}t\bar{t}$  SSML final state.

Can be divided into two types: reducible and irreducible.

Reducible background consists of processes that do not result in SSML final state physically, but are reconstructed as such due to erroneous detector and reconstruction effects.

Three main types: charge misidentification (QmisID), fake leptons and non-prompt leptons.

Estimated using template fitting method to adjust MC predictions via floating normalization factors constrained in the CRs.

Irreducible background consists of SM processes that result in SSML final states physically, with all leptons being prompt.

Main irreducible background considered in this analysis:  $t\bar{t}t\bar{t}$ ,  $t\bar{t}W$ ,  $t\bar{t}Z$ , and  $t\bar{t}H$  with smaller contributions from  $VV$ ,  $VVV$ ,  $VH$  and rarer processes like  $t\bar{t}VV$ ,  $tWZ$ ,  $tZq$  and  $t\bar{t}t$ .

Table 6.3: Caption

Region	Channel	$N_{\text{jets}}$	$N_b$	Other selections	Fitted variable
CR Low $m_{\gamma^*}$	SS $e\ell$	[4, 6)	$\geq 1$	$\ell_1/\ell_2$ is from virtual photon decay $\ell_1 + \ell_2$ not from material conversion	event yield
CR Mat. Conv.	SS $e\ell$	[4, 6)	$\geq 1$	$\ell_1/\ell_2$ is from material conversion	event yield
CR HF $\mu$	$\ell\mu\mu$	$\geq 1$	1	$\ell_1 + \ell_2$ not conversion candidates $100 < H_T < 300$ GeV $E_T^{\text{miss}} > 35$ GeV total charge = $\pm 1$	$p_T(\ell_3)$
CR HF $e$	$e\ell\ell$	$\geq 1$	1	$\ell_1 + \ell_2$ not conversion candidates $100 < H_T < 275$ GeV $E_T^{\text{miss}} > 35$ GeV total charge = $\pm 1$	$p_T(\ell_3)$
CR $t\bar{t}W^+$	SS $\ell\mu$	$\geq 4$	$\geq 2$	$ \eta(e)  < 1.5$ for $N_b = 2$ : $H_T < 500$ GeV or $N_{\text{jets}} < 6$ for $N_b \geq 3$ : $H_T < 500$ GeV total charge $> 0$	$N_{\text{jets}}$
CR $t\bar{t}W^-$	SS $\ell\mu$	$\geq 4$	$\geq 2$	$ \eta(e)  < 1.5$ for $N_b = 2$ : $H_T < 500$ GeV or $N_{\text{jets}} < 6$ for $N_b \geq 3$ : $H_T < 500$ GeV total charge $< 0$	$N_{\text{jets}}$
CR 1b(+)	SS2L+3L	$\geq 4$	1	$\ell_1 + \ell_2$ not from material conversion $H_T > 500$ GeV total charge $> 0$	$N_{\text{jets}}$
CR 1b(-)	SS2L+3L	$\geq 4$	1	$\ell_1 + \ell_2$ not from material conversion $H_T > 500$ GeV total charge $< 0$	$N_{\text{jets}}$
VR $t\bar{t}Z$	3L $\ell^\pm\ell^\mp$	$\geq 4$	$\geq 2$	$m_{\ell\ell} \in [81, 101]$ GeV	$N_{\text{jets}}, m_{\ell\ell}$
VR $t\bar{t}W + 1b$	SS2L+3L			CR $t\bar{t}W^\pm$    CR 1b( $\pm$ )	$N_{\text{jets}}$
VR $t\bar{t}W + 1b + \text{SR}$	SS2L+3L			CR $t\bar{t}W^\pm$    CR 1b( $\pm$ )    SR	$N_{\text{jets}}$
SR	SS2L+3L	$\geq 6$	$\geq 2$	$H_T > 500$ GeV $m_{\ell\ell} \notin [81, 101]$ GeV	$H_T$

Most irreducible backgrounds are estimated using MC simulations normalized to their theoretical SM cross sections (template fitting), with the exception of  $t\bar{t}W$  background due to MC mismodeling of the process at high jet multiplicities.

The  $t\bar{t}W$  is instead given four dedicated CRs, and estimated using a data-driven method with a fitted function parameterized in  $N_{\text{jets}}$

All CRs and SR are included in the final LH-fit to data.

### 6.3.1 Template fitting for fake/non-prompt estimation

Template fit method is a semi-data-driven approach that estimates fake/non-prompt background distributions by fitting the MC kinematic profiles of background processes arising from fake/non-prompt leptons to data.

Each of the four main sources of fake/non-prompt leptons is assigned a free-floating normalization factor constrained by a CR enriched with the corresponding background. The NFs are determined simultaneously with the signal.

- $\text{NF}_{\text{HF } e(\mu)}$ : events with one reconstructed non-prompt electron (muon) from heavy flavor decays,
- $\text{NF}_{\text{Mat. Conv.}}$ : events with one reconstructed non-prompt electrons from photon conversion in the detector material
- $\text{NF}_{\text{Low } m_{\gamma^*}}$ : events with one reconstructed non-prompt electrons in an  $e^+e^-$  pair from virtual photon ( $\gamma^*$ ) conversion.

### 6.3.2 Charge misidentification data-driven estimation

The same-sign di-lepton channel in the analysis gives rise to a major background contamination in opposite-sign di-lepton events with one misidentified charge.

Charge misidentification occurs via incorrect track curvature measurements or trident electron contamination from bremsstrahlung, and therefore mainly concerns electrons due to muons' low bremsstrahlung rate and precise curvature information using the ID and MS.

The charge misidentification rates is significant at higher  $p_T$  and varies with  $|\eta|$  as a proxy for the amount of detector material the electron interacted with, and is consequently estimated in this analysis using a data-driven method with assistance from ECIDS.

The charge flip probability  $\epsilon$  is estimated using a sample of  $Z \rightarrow e^+e^-$  events with additional constraints on the invariant mass  $m_{ee}$  to be within 10 GeV of the  $Z$ -boson mass.

The  $Z$ -boson mass window is defined to be within  $4\sigma$  to include most events within the peak, and is determined by fitting the  $m_{ee}$  spectrum of the two leading electrons to a Breit-Wigner function, resulting in a range of  $[65.57, 113.49]$  for SS events and  $[71.81, 109.89]$  for OS events.

Background contamination near the peak is assumed to be uniform and subtracted using a sideband method.

Since the  $Z$ -boson decay products consist of a pair of opposite-sign electrons, all same-sign electron pairs are considered to be affected by charge misidentification.

Assuming the charge flip probabilities of electrons in an event are uncorrelated, the number of events with same-sign electrons  $N_{ij}^{\text{SS}}$  with the leading electron in the  $i^{\text{th}}$  2D bin in  $(p_T, |\eta|)$  and the sub-leading electron in the  $j^{\text{th}}$  bin can be estimated as

$$N_{ij}^{\text{SS}} = N_{ij}^{\text{tot}}(\epsilon_i(1 - \epsilon_j) + \epsilon_j(1 - \epsilon_i)), \quad (6.1)$$

where  $N_{ij}^{\text{tot}}$  is the total number of events in the  $i^{\text{th}}$  and  $j^{\text{th}}$  bin regardless of charge, and  $\epsilon_{i(j)}$  is the charge flip rate in the  $i^{\text{th}}(j^{\text{th}})$  bin.

Assuming  $N_{ij}^{\text{SS}}$  follows a Poisson distribution around the expectation value  $\bar{N}_{ij}^{\text{SS}}$ , the charge flip rate  $\epsilon$  can be estimated by minimizing a negative-LLH function parameterized in  $p_{\text{T}}$  and  $|\eta|$ ,

$$\begin{aligned}
-\ln(\mathcal{L}(\epsilon|N_{\text{SS}})) &= -\ln \prod_{ij} \frac{(N_{ij}^{\text{tot}})^{N_{ij}^{\text{SS}}} \cdot e^{-N_{ij}^{\text{tot}}}}{N_{ij}^{\text{SS}}!} \\
&= -\sum_{ij} \left[ N_{ij}^{\text{SS}} \ln(N_{ij}^{\text{tot}}(\epsilon_i(1-\epsilon_j) + \epsilon_j(1-\epsilon_i))) - N_{ij}^{\text{tot}}(\epsilon_i(1-\epsilon_j) + \epsilon_j(1-\epsilon_i)) \right].
\end{aligned} \tag{6.2}$$

$$\tag{6.3}$$

1019 The charge flip rate is then calculated separately for SR and CRs with different electron  
1020 definitions (CR Low  $m_{\gamma^*}$ , CR Mat. Conv., CR  $t\bar{t}W$ ) using events satisfying 2LSS kinematic  
1021 selections but with OS electrons, after applying region-specific lepton selections and ECIDS.  
1022 The following weight is applied to OS events to correct for misidentified SS events within  
1023 the region:

$$w = \frac{\epsilon_i + \epsilon_j - 2\epsilon_i\epsilon_j}{1 - \epsilon_i - \epsilon_j + 2\epsilon_i\epsilon_j}. \tag{6.4}$$

### 1024 **6.3.3 $t\bar{t}W$ background data-driven estimation**

1025 -  $t\bar{t}W$  represents a major source of irreducible background contamination in SM and BSM  
1026 analyses with  $t\bar{t}t\bar{t}$  final states.  
1027 - Measured cross section for  $t\bar{t}W$  background has been consistently higher than predicted  
1028 values as seen in previous analyses ( $t\bar{t}H/t\bar{t}W$  multilepton [23][24],  $t\bar{t}t\bar{t}$  analyses [25][26]) due  
1029 to mismodeling, especially at higher  $N_{\text{jets}}$



1030 (show postfit  $t\bar{t}W$  VR distribution)

1031 - Previously, this was handled by assigning large ad-hoc systematic uncertainties to  $t\bar{t}W$   
 1032 events with 7 or more jets. - A semi-data-driven method originally employed in the R-parity-  
 1033 violating-supersymmetry search [27] was used to mitigate this problem. - This method was  
 1034 shown to be effective in the SM  $t\bar{t}t\bar{t}$  observation analysis [26] by improving  $t\bar{t}W$  modeling  
 1035 especially in the showering step and switching  $t\bar{t}W$  systematic uncertainties from predomi-  
 1036 nantly modeling to statistical.

1037 - MC kinematic distributions for  $t\bar{t}W$  are applied with correction factors obtained from a  
 1038 fitted function parameterized in  $N_{\text{jets}}$ .

1039 - The function describes scaling patterns for QCD [28] can be represented by ratio of suc-  
 1040 cessive exclusive jet cross-sections

$$R_{(n+1)/n} = e^{-b} + \frac{\bar{n}}{n+1} = a_0 + \frac{a_1}{1+(j-4)}, \quad (6.5)$$

1041 where  $n$  is the number of jets in addition to the hard process,  $j$  is the inclusive number of  
 1042 jets, and  $\bar{n}$  is the expectation value for the Poisson distribution for exclusive jet cross-section  
 1043 at jet multiplicity  $n$ , described as  $P_n = \sigma_n/\sigma_{\text{tot}}$ .

1044 - Same-sign di-lepton  $t\bar{t}W$  events dominate the  $t\bar{t}W$  background and produce 4 jets in the  
 1045 matrix element at tree level for the hard process, so  $n$  is defined starting from 5 jets and  $j$   
 1046 is defined as inclusive number of jets with 4 or more jets, or  $j \equiv n + 4$ .

1047 - The two terms in the equation correspond respectively to staircase and Poisson scaling  
 1048 between successive multiplicity cross sections, defined as constant ratios  $e^{-b}$  and ratios be-  
 1049 tween Poisson probability for  $n+1$  and  $n$  jets. Staircase scaling is sensitive to events with  
 1050 high jet multiplicity, while Poisson scaling is sensitive to events with low jet multiplicity [28].

1051 - The scaling pattern can then be re-parameterized in  $a_0$  and  $a_1$  to obtain the  $t\bar{t}W$  yield at

1052  $j'$

$$\text{Yield}_{t\bar{t}W(j')} = \text{Yield}_{t\bar{t}W(j=4)} \times \prod_{j=4}^{j'-1} \left( a_0 + \frac{a_1}{1 + (j - 4)} \right) \quad (6.6)$$

1053 where  $j'$  is defined as  $j' \equiv j + 1$  with  $j \geq 4$  since the parameterization starts at the 4<sup>th</sup> jet.

1054 The  $t\bar{t}W$  yield at the 4-jet bin can be represented by a normalization factor applied to  $t\bar{t}W$

1055 MC simulation as  $\text{Yield}_{t\bar{t}W(j=4)} = \text{NF}_{t\bar{t}W(j=4)} \times \text{MC}_{j=4}$ .

1056 To account for the disparity in  $t\bar{t}W^+$  and  $t\bar{t}W^-$  cross-section, assuming the scaling is the

1057 same for both processes,  $\text{NF}_{t\bar{t}W(j=4)}$  can be further split into  $\text{NF}_{t\bar{t}W^+(j=4)}$  and  $\text{NF}_{t\bar{t}W^-(j=4)}$ .

1058 Both NFs are left free-floating to constrain  $t\bar{t}W$  yields at the 4-jet bin in CR 1b(+) and CR

1059 1b(-).

1060 The final  $N_{\text{jets}}$ -parameterized function can then be represented by  $\text{NF}_{t\bar{t}W(j')}$  as

$$\text{NF}_{t\bar{t}W(j')} = \left( \text{NF}_{t\bar{t}W^+(j=4)} + \text{NF}_{t\bar{t}W^-(j=4)} \right) \times \prod_{j=4}^{j'-1} \left( a_0 + \frac{a_1}{1 + (j - 4)} \right). \quad (6.7)$$

1061 This normalization is calculated and applied separately for each sub-sample of  $t\bar{t}W^+$  and

1062  $t\bar{t}W^-$  in an  $N_{\text{jets}}$  bin for  $4 \leq N_{\text{jets}} < 10$ .

1063 Due to small contributions in the CRs, events with  $N_{\text{jets}} < 4$  and  $N_{\text{jets}} \geq 10$  are not

1064 normalized with this scheme.

1065 Instead,  $N_{\text{jets}} < 4$   $t\bar{t}W$  events are fitted by propagating normalization in the 4-jet bin

1066 without additional shape correction. The correction factor for  $t\bar{t}W$  events with  $N_{\text{jets}} \geq$

1067 10 is obtained by summing up the overflow from  $N_{\text{jets}} = 10$  to  $N_{\text{jets}} = 12$ , described as

1068  $\sum_{j'=10}^{12} \prod_{j=4}^{j'-1} \left( a_0 + \frac{a_1}{1 + (j - 4)} \right)$ . Events with  $N_{\text{jets}} \geq 13$  are negligible and thus not included

1069 in the sum.

## Control region definitions

Four control regions CR  $t\bar{t}W^+$ , CR  $t\bar{t}W^-$ , CR 1b(+), CR 1b(-) are constructed to fit  $N_{t\bar{t}W^\pm(j=4)}$  and the scaling parameters  $a_0, a_1$  for the  $t\bar{t}W$  background, as well as validating the parameterization.

Events in CR  $t\bar{t}W^\pm$  are required to contain at least two  $b$ -tagged jets similar to the SR to determine the  $t\bar{t}W$  normalization within an SR-related phase space. Orthogonality with SR is satisfied by requiring  $H_T < 500$  GeV or  $N_{\text{jets}} < 6$  when  $N_b = 2$ , and  $H_T < 500$  GeV when  $N_b \geq 3$ .

The remaining CR 1b( $\pm$ ) require events to have  $H_T > 500$  GeV and at least four jets to encompass events with high  $N_{\text{jets}}$ , which can be used to determine the  $t\bar{t}W$  jet multiplicity spectrum for fitting  $a_{0,1}$ . The selection criteria also include exactly one  $b$ -tagged jet to maintain orthogonality with SR. Assuming the  $t\bar{t}W$  jet multiplicity distribution is similar across different  $N_b$ , a fitted  $N_{\text{jets}}$  distribution in CR 1b( $\pm$ ) can be used to describe the  $t\bar{t}W$  parameterization at higher  $N_{\text{jets}}$ . The full selection criteria for all four regions are shown in ??

Validating the  $t\bar{t}W$  parameterization in Equation 6.7 makes use of the unique charge asymmetry in  $t\bar{t}W$  production that's not present in other background or signal processes. The number of events with all negatively charged leptons is subtracted from that of events with all positively charged leptons, which cancels out charge symmetric events and leaves the  $t\bar{t}W$  background. Validation is done via a statistical-only (stat-only) fit to the  $t\bar{t}W$  MC prediction in CR 1b( $\pm$ ). **validation results**

1092 **6.3.4**  $t\bar{t}Z$  background validation

# Chapter 7. Systematic Uncertainties

(nuisance parameters)

- Heavy pruning, 5% on shape and normalization pruning (to fit timeline?)

## 7.1 Experimental uncertainties

Instrumental & minor:

- uncertainty on the integrated luminosity of the 2015-2018 Run 2 data set is 0.83%, obtained by the LUCID-2 detector for the primary luminosity measurements complemented by the ID and calorimeters

- Pile-up modeling in MC was calibrated to data through pile-up reweighting, resulting in a set of calibration SFs and associated uncertainties.

In general, calibrating MC simulations to match performance in data incurs uncertainties associated with the MC-to-data scale factors obtained from the calibration, which are in turn propagated to observables in the analysis.

### 7.1.1 Leptons

The trigger/reconstruction/ID/isolation efficiencies of electrons and muons (with separate systematic and statistical components for muon) differ between MC simulation and data, and require correction in the form of SFs with its associated uncertainties.

Similarly, electron and muon energy-momentum scale and resolution also incur uncertainties from MC-to-data correction, calculated by varying scale and resolution during simulations.

Muons have additional uncertainties for charge-dependent and charge-independent momen-

tum scale, and detector-specific (ID, MS, CB) track resolution.

The charge identification/ECIDS efficiency also gives rise to an additional uncertainty component.

## 7.1.2 Jets

Experimental uncertainties on jets are dominated by flavor tagging-related uncertainties, with subleading contributions from jet energy scale/resolution (JES/JER) and NNJvt calibration.

### Jet energy scale

JES and its associated uncertainties are determined using data from test-beam and LHC collisions and MC simulated samples, decomposed into uncorrelated components:

- Effective nuisance parameters (NPs): 15  $p_T$ -dependent uncertainty components in total measured in situ, grouped based on their origin (2 detector-related, 4 modeling-related, 3 mixed, 6 statistical-related)
- $\eta$  intercalibration: 6 total components (1 modeling-related, 4 non-closure and 1 statistical-related) associated with the correction of the forward jets' ( $0.8 \leq |\eta| < 4.5$ ) energy scale to that of the central jets ( $|\eta| < 0.8$ ).
- Flavor composition/response: 2 components for relative quark-gluon flavor compositions in background and signal samples, and 2 components for uncertainty in responses to gluon-initiated versus quark-initiated jets

- Pile-up subtraction: 4 components, two for uncertainty in  $\mu$  (`OffsetMu`) and  $N_{\text{PV}}$  (`OffsetNPV`) modeling, one for residual  $p_{\text{T}}$ -dependency (`PtTerm`) and one for topology dependence on the per-event  $p_{\text{T}}$  density modeling (`RhoTopology`)
- Punch-through effect treatment: two terms (AF3 fast simulation and full detector simulations) for GSC punch-through jet response correction between data and MC.
- Non-closure: one term to account for difference between AF3-simulated samples and full detector simulations.
- High- $p_{\text{T}}$  single-particle response: one term for response to high- $p_{\text{T}}$  jets from single-particle and test-beam measurements
- $b$ -jets response: one term for uncertainty in the response to  $b$ -jets

## Jet energy resolution

JER measured separately in data and MC simulations using in situ techniques as a function of  $p_{\text{T}}$  and  $\eta$  for a given jet. Associated uncertainties are defined as quadratic difference between data and MC simulations.

This analysis uses the full JER uncertainty set provided for Run 2 searches with 14 total components: 12 effective NPs and 2 for difference between data and MC simulation, separately for AF3 and FS.

## Jet vertex tagging

JVT associated uncertainty is obtained by varying the JVT efficiency correction SFs within their range of uncertainty. This uncertainty accounts for remaining contamination from pile-up jets after applying pile-up suppression and MC generator choice.

## Flavor tagging

SFs for  $b$ -jets tagging efficiencies and  $c$ -/light-jets mis-tagging rates are obtained as a function of  $p_T$  for  $b$ -/ $c$ -/light-jets and PCB scores. The covariance matrix of systematic and statistical uncertainties is diagonalized and reduced in dimensions using principle component analysis (PCA), resulting in a set of orthogonal NPs: 85 for  $b$ -jets, 56 for  $c$ -jets and 42 for light-jets.

### 7.1.3 Missing transverse energy

Uncertainties for  $E_T^{\text{miss}}$  arise from possible miscalibration of its soft-track component, and are estimated using data-MC comparison of the  $p_T$  scale and resolution between the hard and soft  $E_T^{\text{miss}}$  terms. These uncertainties are represented by three independent terms: one for scale uncertainty and two resolution uncertainties for the parallel and perpendicular components.

## 7.2 Modeling uncertainties

### 7.2.1 Signal and irreducible background uncertainties

- scale variations - 6-point variation method, varying  $\mu_R$  &  $\mu_F$  vs central values to cover missing higher-order QCD corrections (signal & all major irreducible background)

$(\mu_R, \mu_F) = (0.5, 0.5), (0.5, 1), (1, 0.5), (1, 2), (2, 1), (2, 2)$  - pdf uncertainty: flat 1% for  $t\bar{t}Z'$ ,  $t\bar{t}t\bar{t}$ ,  $t\bar{t}Z$ ,  $t\bar{t}H$ , envelope of differences between nominal vs. other pdf choices for  $t\bar{t}t$

$t\bar{t}Z'$  signal

- parton distribution function: 1%



Table 7.1: Summary of the experimental systematic uncertainties considered in this analysis.

Systematic uncertainty	Terms	Scale [%]
<b>Event</b>		
Luminosity	1	0.83
Pile-up reweighting	1	$\mathcal{O}(1) \sim \mathcal{O}(10)$
<b>Electrons</b>		
Trigger efficiency	1	$\mathcal{O}(10^{-2}) \sim \mathcal{O}(10^{-1})$
Reconstruction efficiency <sup>†</sup>	1	$\mathcal{O}(10^{-1}) \sim \mathcal{O}(1)$
Identification efficiency <sup>†</sup>	1	$\mathcal{O}(10^{-1}) \sim \mathcal{O}(1)$
Isolation efficiency <sup>†</sup>	1	$\mathcal{O}(10^{-1}) \sim \mathcal{O}(1)$
Energy scale	1	$\mathcal{O}(10^{-2}) \sim \mathcal{O}(10^{-1})$
Energy resolution	1	$\mathcal{O}(10^{-2}) \sim \mathcal{O}(10^{-1})$
Charge identification (ECIDS) efficiency <sup>†</sup>	1	$\mathcal{O}(10^{-1}) \sim \mathcal{O}(1)$
<b>Muons</b>		
Trigger efficiency (stat/sys)	2	$\mathcal{O}(10^{-1}) \sim \mathcal{O}(1)$
Track-to-vertex association efficiency (stat/sys)	2	$\mathcal{O}(10^{-2}) \sim \mathcal{O}(10^{-1})$
Reconstruction/identification efficiency (stat/sys)	2	$\mathcal{O}(10^{-1}) \sim \mathcal{O}(1)$
Low- $p_T$ (< 15 GeV) reconstruction/identification efficiency (stat/sys)	2	$\mathcal{O}(10^{-1}) \sim \mathcal{O}(1)$
Isolation efficiency (stat/sys)	2	$\mathcal{O}(10^{-1}) \sim \mathcal{O}(1)$
Charge-independent momentum scale	1	$\mathcal{O}(10^{-2}) \sim \mathcal{O}(10^{-1})$
Charge-dependent momentum scale	4	$\mathcal{O}(10^{-2}) \sim \mathcal{O}(10^{-1})$
Energy resolution (CB)	1	$\mathcal{O}(10^{-2}) \sim \mathcal{O}(10^{-1})$
Energy resolution (ID & MS)*	2	$\mathcal{O}(10^{-2}) \sim \mathcal{O}(10^{-1})$
<b>Jets</b>		
JES effective NP	15	$\mathcal{O}(10^{-2}) \sim \mathcal{O}(1)$
JES $\eta$ intercalibration	3	$\mathcal{O}(10^{-1}) \sim \mathcal{O}(1)$
JES flavor composition	2	$\mathcal{O}(10^{-1}) \sim \mathcal{O}(1)$
JES flavor response	1	$\mathcal{O}(10^{-1}) \sim \mathcal{O}(1)$
JES pile-up	4	$\mathcal{O}(10^{-1}) \sim \mathcal{O}(10)$
JES punch-through (FS/AF3*)	2	$< \mathcal{O}(10^{-2})$
JES non-closure	1	$\mathcal{O}(10^{-2}) \sim \mathcal{O}(10^{-1})$
JES high- $p_T$ single particle	1	$< \mathcal{O}(10^{-2})$
JES $b$ -jet response	1	$\mathcal{O}(10^{-1}) \sim \mathcal{O}(1)$
JER effective NP	12	$\mathcal{O}(10^{-1}) \sim \mathcal{O}(1)$
JER data/MC (FS/AF3*)	2	$\mathcal{O}(10^{-1}) \sim \mathcal{O}(1)$
JVT efficiency	1	$\mathcal{O}(10^{-1}) \sim \mathcal{O}(1)$
GN2v01 $b$ -tagging efficiency ( $b$ -jets)	85	$\mathcal{O}(10^{-2}) \sim \mathcal{O}(1)$
GN2v01 $b$ -tagging efficiency ( $c$ -jets)	56	$\mathcal{O}(10^{-2}) \sim \mathcal{O}(1)$
GN2v01 $b$ -tagging efficiency (light-jets)	42	$\mathcal{O}(10^{-2}) \sim \mathcal{O}(1)$
<b><math>E_T^{\text{miss}}</math>-Terms</b>		
Track-based soft term for transversal resolution	1	$\mathcal{O}(10^{-2}) \sim \mathcal{O}(10^{-1})$
Track-based soft term for longitudinal resolution	1	$\mathcal{O}(10^{-2}) \sim \mathcal{O}(10^{-1})$
Track-based soft term for longitudinal scale	1	$\mathcal{O}(10^{-2}) \sim \mathcal{O}(10^{-1})$

1174 **SM  $t\bar{t}t\bar{t}$  background**

- 1175 - cross section: 20% from NLO prediction in QCD+EW
- 1176 - generator uncertainty: madgraph5\_amc@nlo (nominal) vs sherpa 2.2.10
- 1177 - parton shower uncertainty: pythia8 (nominal) vs herwig7

1178 **SM  $t\bar{t}t$  background**

- 1179 - cross section: 30% from NLO prediction in QCD+EW
- 1180 - additional  $b$ -jets: 50% for  $t\bar{t}t$  events with 4+ truth  $b$ -jets

1181  **$t\bar{t}W$ ,  $t\bar{t}Z$ ,  $t\bar{t}H$  background**

- 1182 - cross section:  $t\bar{t}Z$  12%,  $t\bar{t}H$  10% (from CERN yellow report)
- 1183 no cross-section and pdf uncertainties for  $t\bar{t}W$  since normalizations and jet multiplicity spec-
- 1184 trum are estimated with data-driven method
- 1185 - parton shower uncertainty:  $t\bar{t}H$  powhegbox+pythia8 (nominal) vs powhegbox+herwig7
- 1186 - additional  $b$ -jets: events with additional HF jets can contaminate SR and are challenging
- 1187 to model w/ MC - 50% for events with an additional truth  $b$ -jet not from top-quark decay,
- 1188 additional 50% for 2 or more
- 1189 - generator uncertainty **table?**
- 1190
  - $t\bar{t}W$ - sherpa (nominal) vs madgraph5\_amc@nlo
- 1191
  - $t\bar{t}Z$ - madgraph5\_amc@nlo (nominal) vs sherpa 2.2.10
- 1192
  - $t\bar{t}H$ - powheg8/PhPy8 (nominal) vs powheg8/PhPy8 pthard

## Other backgrounds

- $t(\bar{t})X$ : cross section 30%
- $VV$ : cross section (STDM-2018-03) uncorrelated 20%/50%/60% for events with 3-  
/4/5+ jets; events with 1+ truth  $b$ -jets not from top decay 50%
- $t\bar{t}VV, VVV, VH$ : cross section 50%; additional  $b$ -jets same as  $VV$

## 7.2.2 Reducible background uncertainties

- Electron charge misidentification background:
- Material and internal (low  $\gamma^*$ ) conversion background: estimated based on data/MC differences in a region enriched with  $Z \rightarrow \ell^+ \ell^- \gamma$ ; 30% & 21% for material & internal conversion
- Heavy-flavor non-prompt lepton background: estimated based on data/MC differences in CR/SR distributions, ranging from 20-100%
- Light-flavor decays and other fake/non-prompt background: Conservative normalization uncertainty of 100% for light-flavor non-prompt lepton background (ATLAS-CONF-2019-045), 30% for normalization of all other fake backgrounds.
- +HF: contaminates SR phase space with large  $b$ -jet multiplicity, estimated from data/MC discrepancy, 30% for events with

Table 7.2: Caption

Systematic uncertainty	Terms	Scale [%]
<b><math>t\bar{t}Z'</math> modeling</b>		
Renormalization & factorization scale		
PDF		
<b>SM <math>t\bar{t}t\bar{t}</math> modeling</b>		
Cross-section		
Renormalization & factorization scale		
PDF		
Generator choice		
Parton shower model		
<b>SM <math>t\bar{t}t</math> modeling</b>		
Cross-section		
Renormalization & factorization scale		
PDF		
Additional $b$ -jets		
<b><math>t\bar{t}W</math> modeling</b>		
Renormalization & factorization scale		
Generator choice		
Additional $b$ -jets		
<b><math>t\bar{t}Z</math> modeling</b>		
Cross-section		
Renormalization & factorization scale		
PDF		
Generator choice		
Additional $b$ -jets		
<b><math>t\bar{t}H</math> modeling</b>		
Cross-section		
Renormalization & factorization scale		
PDF		
Generator choice		
Parton shower model		
Additional $b$ -jets		
<b>Other background modeling</b>		
Cross-section		
Additional $b$ -jets		

Table 7.3: Caption

Systematic uncertainty	Terms	Scale [%]
<b>Reducible SM background</b>		
$t\bar{t}/V/t+\text{jets}$	2	
Charge misidentification	1	
<b>Fake &amp; non-prompt background</b>		
Low $\gamma^*$	1	
Material conversion	1	
HF $e$	1	
HF $\mu$	1	
Light-flavor decays	1	100
Other fakes	1	30

# Chapter 8. Results

## 8.1 Statistical model

### 8.1.1 Binned profile likelihood fit

### 8.1.2 Signal significance

### 8.1.3 Limit exclusion

## 8.2 Fit results

### Fit setup

- Plain Asimov fit (**only mentioning briefly**): all regions included; simulated data used in the fit match exactly to MC prediction with nominal  $\mu_{t\bar{t}Z'}$  set to 0 and allowed to free-float.  
Purpose: to perform studies on optimizing fitted parameters and expected sensitivity; refining background estimation techniques; optimizing region definition and object definition
- Real SRs-blinded fit: similar to plain Asimov, but use observed data in CRs.  
Purpose: study the behavior of background estimation using real observed data in CRs on Asimov data in SRs and assessing the influence of statistical effects on fitted parameters and expected sensitivity
- Real SRs-unblinded/ $H_T$  fit: all regions included,

## 8.3 Limits

1229 **Chapter 9. Summary**



# References

- [1] ATLAS Collaboration, *Observation of electroweak production of two jets in association with an isolated photon and missing transverse momentum, and search for a Higgs boson decaying into invisible particles at 13 TeV with the ATLAS detector*, *Eur. Phys. J. C*, vol. 82, p. 105, 2022. DOI: [10.1140/epjc/s10052-021-09878-z](https://doi.org/10.1140/epjc/s10052-021-09878-z). arXiv: [2109.00925](https://arxiv.org/abs/2109.00925) [[hep-ex](#)] (cit. on pp. vii, 41).
- [2] C. Burgard and D. Galbraith. *Standard Model of Physics*. (), [Online]. Available: <https://texample.net/model-physics/> (visited on 06/02/2025) (cit. on pp. viii, 4).
- [3] ATLAS Collaboration, *Graph Neural Network Jet Flavour Tagging with the ATLAS Detector*, ATL-PHYS-PUB-2022-027, 2022. [Online]. Available: <https://cds.cern.ch/record/2811135> (cit. on pp. viii, 32).
- [4] A. Duperrin, *Flavour tagging with graph neural networks with the ATLAS detector*, CERN, Geneva, Tech. Rep., 2023, Presented at DIS2023, Michigan State University, USA. arXiv: [2306.04415](https://arxiv.org/abs/2306.04415) [[hep-ex](#)]. [Online]. Available: <https://cds.cern.ch/record/2860610> (cit. on pp. viii, 32).
- [5] S. Van Stroud, A. Duperrin, Y. Wang, *et al.*, *FTAG Run-3 Algorithms Performance*, CERN, Geneva, Tech. Rep., 2023. [Online]. Available: <https://cds.cern.ch/record/2872884> (cit. on pp. viii, 32).

- [6] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at  $\sqrt{s} = 13$  TeV*, *Eur. Phys. J. C*, vol. 79, p. 639, 2019. DOI: [10.1140/epjc/s10052-019-7140-6](https://doi.org/10.1140/epjc/s10052-019-7140-6). arXiv: [1902.04655](https://arxiv.org/abs/1902.04655) [[physics.ins-det](#)] (cit. on pp. viii, 33–36).
- [7] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data*, *JINST*, vol. 14, P12006, 2019. DOI: [10.1088/1748-0221/14/12/P12006](https://doi.org/10.1088/1748-0221/14/12/P12006). arXiv: [1908.00005](https://arxiv.org/abs/1908.00005) [[hep-ex](#)] (cit. on pp. viii, 33, 35–37).
- [8] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, *Eur. Phys. J. C*, vol. 77, p. 490, 2017. DOI: [10.1140/epjc/s10052-017-5004-5](https://doi.org/10.1140/epjc/s10052-017-5004-5). arXiv: [1603.02934](https://arxiv.org/abs/1603.02934) [[hep-ex](#)] (cit. on p. 26).
- [9] W. Lampl *et al.*, *Calorimeter clustering algorithms: description and performance*, ATLAS-PUB-2008-002, 2008. [Online]. Available: <https://cds.cern.ch/record/1099735> (cit. on p. 26).
- [10] ATLAS Collaboration, *Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2*, *Eur. Phys. J. C*, vol. 77, p. 673, 2017. DOI: [10.1140/epjc/s10052-017-5225-7](https://doi.org/10.1140/epjc/s10052-017-5225-7). arXiv: [1704.07983](https://arxiv.org/abs/1704.07983) [[hep-ex](#)] (cit. on p. 27).
- [11] T. Cornelissen, M. Elsing, S. Fleischmann, W. Liebig, E. Moyse, and A. Salzburger, *Concepts, design and implementation of the ATLAS New Tracking (NEWT)*, CERN, Geneva, Tech. Rep., 2007. [Online]. Available: <https://cds.cern.ch/record/1020106> (cit. on p. 28).
- [12] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers,*

*Detectors and Associated Equipment*, vol. 262, no. 2, pp. 444–450, 1987, ISSN: 0168-9002. DOI: [https://doi.org/10.1016/0168-9002\(87\)90887-4](https://doi.org/10.1016/0168-9002(87)90887-4). [Online]. Available: <https://www.sciencedirect.com/science/article/pii/0168900287908874> (cit. on p. 28).

[13] T. Cornelissen, M. Elsing, I. Gavrilenko, *et al.*, *The global  $\chi^2$  track fitter in ATLAS*, *Journal of Physics: Conference Series*, vol. 119, no. 3, p. 032 013, 2008. DOI: [10.1088/1742-6596/119/3/032013](https://doi.org/10.1088/1742-6596/119/3/032013). [Online]. Available: <https://dx.doi.org/10.1088/1742-6596/119/3/032013> (cit. on p. 28).

[14] ATLAS Collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, ATLAS-CONF-2012-047, 2012. [Online]. Available: <https://cds.cern.ch/record/1449796> (cit. on p. 28).

[15] D. Wicke, *A new algorithm for solving tracking ambiguities*, Oct. 1998 (cit. on p. 28).

[16] A. Salzburger and on behalf of the ATLAS Collaboration, *Optimisation of the ATLAS Track Reconstruction Software for Run-2*, *Journal of Physics: Conference Series*, vol. 664, no. 7, p. 072 042, 2015. DOI: [10.1088/1742-6596/664/7/072042](https://doi.org/10.1088/1742-6596/664/7/072042). [Online]. Available: <https://dx.doi.org/10.1088/1742-6596/664/7/072042> (cit. on p. 28).

[17] ATLAS Collaboration, *Measurements of  $b$ -jet tagging efficiency with the ATLAS detector using  $t\bar{t}$  events at  $\sqrt{s} = 13$  TeV*, *JHEP*, vol. 08, p. 089, 2018. DOI: [10.1007/JHEP08\(2018\)089](https://doi.org/10.1007/JHEP08(2018)089). arXiv: [1805.01845](https://arxiv.org/abs/1805.01845) [[hep-ex](#)] (cit. on p. 31).

[18] ATLAS Collaboration, *Electron Identification with a Convolutional Neural Network in the ATLAS Experiment*, ATL-PHYS-PUB-2023-001, 2023. [Online]. Available: <https://cds.cern.ch/record/2850666> (cit. on p. 36).

[19] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2  $pp$  collision data set at  $\sqrt{s} = 13$  TeV*, *Eur. Phys. J. C*, vol. 81,

p. 578, 2021. DOI: [10.1140/epjc/s10052-021-09233-2](https://doi.org/10.1140/epjc/s10052-021-09233-2). arXiv: [2012.00578](https://arxiv.org/abs/2012.00578) [[hep-ex](#)]  
(cit. on pp. 38, 39).

[20] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at  $\sqrt{s} = 13$  TeV*, *Eur. Phys. J. C*, vol. 76, p. 292, 2016. DOI: [10.1140/epjc/s10052-016-4120-y](https://doi.org/10.1140/epjc/s10052-016-4120-y). arXiv: [1603.05598](https://arxiv.org/abs/1603.05598) [[hep-ex](#)] (cit. on p. 38).

[21] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at  $\sqrt{s} = 13$  TeV*, *Eur. Phys. J. C*, vol. 78, p. 903, 2018. DOI: [10.1140/epjc/s10052-018-6288-9](https://doi.org/10.1140/epjc/s10052-018-6288-9). arXiv: [1802.08168](https://arxiv.org/abs/1802.08168) [[hep-ex](#)] (cit. on p. 39).

[22] ATLAS Collaboration,  *$E_T^{miss}$  performance in the ATLAS detector using 2015–2016 LHC  $pp$  collisions*, ATLAS-CONF-2018-023, 2018. [Online]. Available: <https://cds.cern.ch/record/2625233> (cit. on p. 40).

[23] ATLAS Collaboration, *Analysis of  $t\bar{t}H$  and  $t\bar{t}W$  production in multilepton final states with the ATLAS detector*, ATLAS-CONF-2019-045, 2019. [Online]. Available: <https://cds.cern.ch/record/2693930> (cit. on p. 53).

[24] ATLAS Collaboration, *Measurement of the total and differential cross-sections of  $t\bar{t}W$  production in  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, *JHEP*, vol. 05, p. 131, 2024. DOI: [10.1007/JHEP05\(2024\)131](https://doi.org/10.1007/JHEP05(2024)131). arXiv: [2401.05299](https://arxiv.org/abs/2401.05299) [[hep-ex](#)] (cit. on p. 53).

[25] ATLAS Collaboration, *Evidence for  $t\bar{t}t\bar{t}$  production in the multilepton final state in proton–proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, *Eur. Phys. J. C*, vol. 80, p. 1085, 2020. DOI: [10.1140/epjc/s10052-020-08509-3](https://doi.org/10.1140/epjc/s10052-020-08509-3). arXiv: [2007.14858](https://arxiv.org/abs/2007.14858) [[hep-ex](#)] (cit. on p. 53).

- 1317 [26] ATLAS Collaboration, *Observation of four-top-quark production in the multilepton*  
 1318 *final state with the ATLAS detector*, *Eur. Phys. J. C*, vol. 83, p. 496, 2023. DOI: [10.1140/epjc/s10052-023-11573-0](https://doi.org/10.1140/epjc/s10052-023-11573-0). arXiv: [2303.15061](https://arxiv.org/abs/2303.15061) [[hep-ex](#)] (cit. on pp. 53, 54).  
 1319
- 1320 [27] ATLAS Collaboration, *Search for R-parity-violating supersymmetry in a final state*  
 1321 *containing leptons and many jets with the ATLAS experiment using  $\sqrt{s} = 13$  TeV*  
 1322 *proton-proton collision data*, *Eur. Phys. J. C*, vol. 81, p. 1023, 2021. DOI: [10.1140/epjc/s10052-021-09761-x](https://doi.org/10.1140/epjc/s10052-021-09761-x). arXiv: [2106.09609](https://arxiv.org/abs/2106.09609) [[hep-ex](#)] (cit. on p. 54).  
 1323
- 1324 [28] E. Gerwick, T. Plehn, S. Schumann, and P. Schichtel, *Scaling Patterns for QCD Jets*,  
 1325 *JHEP*, vol. 10, p. 162, 2012. DOI: [10.1007/JHEP10\(2012\)162](https://doi.org/10.1007/JHEP10(2012)162). arXiv: [1208.3676](https://arxiv.org/abs/1208.3676)  
 1326 [[hep-ph](#)] (cit. on p. 54).