Estimating the Sensitivity to New Resonances Decaying to Boosted Quark Pairs and Produced in Association with a Photon in the ATLAS Experiment

Research Thesis

Submitted in partial fulfilment of the requirements for the degree of

Master of Science in Physics

Roy Schimmel Brener

Submitted to the Senate of
Technion – Israel Institute of Technology
Kislev 5781, Haifa, December 2020

This research was conducted under the supervision of
Dr. Enrique Kajomovitz at the Department of Physics

The generous financial help of Technion is gratefully acknowledged

Contents

List of Figures

List of Tables

List of Abbreviations

SM	Standard Model
BSM	Beyond the Standard Model
HEP	High Energy Physics
NP	New Physics
LO	Leading Order
NLO	Next-to-Leading Order
DoF	Degrees of Freedom
CoM	Centre of Mass
QFT	Quantum Field Theory
QED	Quantum Electrodynamics
QCD	Quantum Chromodynamics
SSB	Spontaneous Symmetry Breaking
PDF	Parton Distribution Function
LFV	Lepton Flavour Violation
DM	Dark Matter
CKM	Cabibbo–Kobayashi–Maskawa
GWS	Glashow–Weinberg–Salam
GIM	Glashow–Iliopoulos–Maiani
PMNS	Pontecorvo–Maki–Nakagawa–Sakata
DGLAP	Dokshizer–Gribov–Lipatov–Altarelli–Parisi
CERN	The European Organization for Nuclear Research
LHC	Large Hadron Collider

ATLAS	A Toroidal Large Hadron Collider ApparatuS
IP	Interaction Point
MS	Muon Spectrometer
ID	Inner Detector
ECAL	Electromagnetic Calorimeter
HCAL	Hadronic Calorimeter
LCW	Local-Cluster Weighting
JES	Jet Energy Scale
HLT	High-Level Trigger
ISR	Initial-State Radiation
FSR	Final-State Radiation
MC	Monte Carlo
ML	Machine Learning
ANN	Adversarial Neural Network
ROC	Receiver Operating Characteristic
MVA	Multivariate Analysis
CL	Confidence Level

Abstract

Particle physics' most solidified theory is the Standard Model, which includes three of the four known interactions in the Universe, namely the electromagnetic, weak and strong. Although the Standard Model's predictions are repeatedly tested to remarkable accuracy, there are eminent reasons to believe it is not a complete theory of Nature. These consist of the model's inability to incorporate gravity in its description of the fundamental forces, predict non-zero neutrino masses and account for the existence of Dark Matter. Numerous extensions known as beyond the Standard Model physics are proposed and explored by Large Hadron Collider experiments where bounds are set upon them. The study described in this thesis relates to mediators between Standard Model and Dark Matter particles and appraises novel techniques by which such resonances may be investigated with the ATLAS detector.

The benchmark model proposes a leptophobic vector boson, Z', charged under a new $U(1)'$ gauge symmetry which couples to Standard Model quarks and to Dark Matter Dirac fermions. Its interactions with Dark Matter are infinitesimally rare and thus searches involving such mediators primarily focus on their decays to Standard Model particles, identifiable by Large Hadron Collider detectors. This new type of hardonically-decaying particles remains illusive to current searches due to an extensive multijet background, saturating the trigger scheme. An alternative method is adopted in this thesis.

The method uses events wherein an initial-state radiation photon is the trigger object whilst the recoiling resonance is reconstructed from the hadronic final state as a single large-radius jet. These jets are reconstructed with the Hadronic Calorimeter and their energy measurement is of an approximately 3% accuracy. Due to the recoil of the resonance against the photon, the resonance is highly Lorentz-boosted. Hadronically-decaying bosons with high Lorentz boosts can be tagged using substructure observables, targeting their two-pronged characteristic.

Using Monte Carlo simulations, selection criteria are optimised. The usefulness of a machine learning tool based on jet substructure observables for tagging two-pronged resonances is tested. Following the optimised selection, Monte Carlo simulations and a control study are used to assess the sensitivity of the ATLAS experiment to the Z' in this channel. Within the control study, $36.0\,\text{fb}^{-1}$ data are compared with jet+γ and $q\bar{q}+\gamma \to W/Z \to q\bar{q}$ simulations. Thereby, it is estimated that this scheme bears the potential to discover processes similar to a Standard Model vector boson production in association with a photon and a diquark final state at cross sections as small as 16% and 46% of the W and Z production cross sections, respectively. Finally, using Monte Carlo simulations of processes where $q\bar{q} + \gamma \to Z' \to q\bar{q}$ for a mass range $m_{Z'} \in [100, 220]\,\text{GeV}$, the sensitivity to the new resonance is estimated and provided in terms of cross section and coupling confidence level limits.

Chapter 1

Theoretical Background

1.1 The Standard Model

Formalised by means of Quantum Field Theory (QFT) [1–3] and developed in the second half of the twentieth century, the Standard Model (SM) [4–6] constitutes the best theory of particle physics at hand today. Over the last 40 years, it has been tested repeatedly and rigorously, hitherto maintaining its precise description of fundamental particles. Notwithstanding, reasons to believe the SM is incomplete have accumulated. At worst, it is believed the SM is a low-energy effective theory of particle physics and studies extending its limits are coined Beyond the Standard Model (BSM) physics. In any event, the framework of particle physics is the SM: extensions are based upon the same language of describing bosons and fermions, particles of integer spin and half-integer spin, respectively, as quantum fields. Moreover, codifying three fundamental forces, electromagnetic, weak and strong, into a non-Abelian gauge symmetry of three Lie groups,

$$SU(3) \times SU(2) \times U(1), \tag{1.1}$$

has been proven as an insightful mathematical characterisation of matter and its interactions at elementary scales. In this scheme, fermions are organised in three generations, each containing a lepton pair and a quark pair. Unlike quarks, leptons do not interact through the strong force but are, however, subject to the electroweak interaction. The electromagnetic and strong forces are carried by massless bosons, photons and (eight) gluons, respectively, whilst the weak force carriers are massive particles, the $W^{\pm}$ and Z^0 bosons.

Each particle's representation of the Poincaré group, the non-Abelian Lie group of spacetime symmetries, rotations, translations and boosts, can be related to its mass and spin. Associating bosons with vector quantum fields (notwithstanding the unique case of the Higgs boson, a scalar field) and fermions with spinor quantum fields, each SM force is related to gauge symmetries of the SM Lagrangian. Relating to Eq. (1.1), the electroweak interaction results from the non-Abelian Lie group $SU(2) \times U(1)$ and the strong interaction from $SU(3)$. In the SM, electromagnetism is unified with the weak interaction as a Yang-Mills field with an $SU(2) \times U(1)$ gauge symmetry, otherwise known as electroweak symmetry. The generators of $SU(2)$ and $U(1)$ are associated with isospin, T, and hypercharge, Y, respectively. These relate to unphysical massless isospin W fields and a hypercharge B field. Through a process of spontaneous symmetry breaking (SSB) mediated by a spin-zero scalar particle, the Higgs boson, the $SU(2) \times U(1)$ symmetry is broken. In this process, fermions and weak vector bosons $W^{\pm}$ and Z^0 gain their mass. The SM Lagrangian is

$$\mathcal{L} = -\frac{1}{4}(F^a_{\mu\nu})^2 + \bar{\psi}\mathrm{i}\not{D}\psi + y_{ij}\bar{\psi}_i\psi_j\phi + |D_\mu\phi|^2 + \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2, \tag{1.2}$$

where $F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + \mathrm{g}f^{abc}A^b_\mu A^c_\nu$ is the gauge field strength tensor, $D_\mu = \partial_\mu - \mathrm{ig}A^a_\mu t^a_r$ the gauge covariant derivative, ψ the spinor fields, ϕ the scalar fields, A_μ the vector (gauge) fields, t^a_r the group representation matrices, f^{abc} the strucure constants and i, j generation indices. The Yukawa couplings y_{ij} and parameters λ and μ describe the Higgs field-fermion interaction and the Higgs field self-interaction, respectively. An immediate conclusion of a profound physical importance is to observe that for non-Abelian symmetries where $f^{abc} \neq 0$, $SU(2)$ and $SU(3)$, there exist terms where respective vector fields self-couple. For this reason, bosons of the weak and strong forces interact with themselves. Conversely, in electromagnetism where $f^{abc} = 0$, photons do not self-interact.

The Lagrangian in Eq. (1.2) is given before electroweak symmetry breaking. The fields there are outlined and thereafter, the phenomenon of SSB, exaplained. One may associate between fields in Eq. (1.2) to representations of Lie groups given in Eq. (1.1), as succinctly given in Table 1.1. There are three sets, each of right-handed up-type and down-type quarks. The left-handed up-type and down-type quark pair constitute an $SU(2)$ doublet. Likewise, there are three generations of left-handed charged leptons which, together with their corresponding neutrinos, are grouped as $SU(2)$ doublets. Another $SU(2)$ doublet, ϕ, consists of two complex scalar fields and bears key significance within SSB. The right-handed fermions, quarks and leptons alike, constitute $SU(2)$ singlets. Apart from a $U(1)$ gauge field, the SM Lagrangian before SSB contains $W^i \,|\, i \in \{1, 2, 3\}$ and B weak boson fields for $SU(2)\times U(1)$ symmetry and eight gluons, one for each generator of the $SU(3)$ Lie algebra, associated with the Gell-Mann matrices. For each dimension of the fundamental representation of $SU(3)$ there exists three quarks. Equivalently with electrical charge associated with $U(1)$ symmetry, partons (quarks and gluons) carry colour charge, stemming from $SU(3)$ gauge. This charge is consistent with three options of the symmetry and labelled red, green or blue, such that a quark triplet is effectively colourless whereas the strong force bosons, gluons, carry one colour charge and one anti-colour charge each. The fields and their charges and representations under SM gauge symmetries before SSB are indeed also given in Table 1.1.

Through a process involving SSB named the Higgs mechanism [7, 8], the symmetry given in Eq. (1.1) is disjoint and thereby SM fermions and gauge bosons gain mass. This occurs when the Higgs potential, $V(\phi) = -\mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2$, possesses a non-zero minimum for μ^2, $\lambda > 0$. The field value for such a minimum, the ground state of the Higgs field, is degenerate and often written as $\frac{1}{\sqrt{2}}(0 \quad v) \mid v = \sqrt{\frac{\mu^2}{\lambda}}$, which can be associated to any other (degenerate) ground state via a global $SU(2)$ transformation. For $\phi = (\phi^+ \; \phi^0)$ as in Table 1.1, the expectation value of ϕ^0 in the ground state coined the vacuum expectation value (VEV), is given as $\langle\phi^0\rangle = \frac{1}{\sqrt{2}}v$. The Higgs field may therefore be considered a fluctuation about this minimum and

Field	Particle content	Spin	$U(1)$	$SU(2)$	$SU(3)$
Q_i	$(u_L\ d_L)$	$\frac{1}{2}$	$\frac{1}{6}$	$\mathbf{2}$	$\mathbf{3}$
$u_{R,i}$	u_R	$\frac{1}{2}$	$\frac{2}{3}$	$\mathbf{1}$	$\bar{\mathbf{3}}$
$d_{R,i}$	d_R	$\frac{1}{2}$	$-\frac{1}{3}$	$\mathbf{1}$	$\bar{\mathbf{3}}$
ℓ_i	$(\ell_L\ \nu_{\ell,L})$	$\frac{1}{2}$	$\frac{1}{2}$	$\mathbf{2}$	$\mathbf{1}$
$\ell_{R,i}$	ℓ_R	$\frac{1}{2}$	-1	$\mathbf{1}$	$\mathbf{1}$
ϕ	$(\phi^+\ \phi^0)$	0	$\frac{1}{2}$	$\mathbf{2}$	$\mathbf{1}$
B	B	1	0	$\mathbf{1}$	$\mathbf{1}$
W	$(W^1\ W^2\ W^3)$	1	0	$\mathbf{3}$	$\mathbf{1}$
g	g	1	0	$\mathbf{1}$	$\mathbf{8}$

Table 1.1: Field and particle content of the SM prior to electroweak symmetry breaking. Values in the $U(1)$ column indicate electrical charge, originating from the Abelian symmetry, unlike entries under the $SU(2)$ and $SU(3)$ columns which are the respective representations of the entries given in the Field column. *E.g.* a $u_{R,i}$ field, a right-handed up-type quark, is a $SU(2)$ singlet in the fundamental representation and a $SU(3)$ triplet in the anti-fundamental representation of electrical charge 2/3.

rewritten as $\phi = \frac{1}{\sqrt{2}}(v + h(x))$ given a unitary gauge choice where $h(x) \in \mathbb{R}\,\forall\, x$. Here it is worthwhile noting that there are no explicit mass terms in Eq. (1.2). This is due to the fact they are forbidden by symmetry: a term relating left-handed and right-handed fermions, $m\bar{\psi}_L\psi_R$, would break the $SU(2)$ symmetry. However, having gone through SSB, the expanded Higgs field may be inserted into Eq. (1.2). Expanding the post-SSB SM Lagrangian, mass terms for fermions and electroweak gauge bosons are yielded. Respectively,

$$\begin{aligned}\mathcal{L}_{(\psi\bar{\psi}+W/Z)\text{-}v} =& -\frac{yv}{\sqrt{2}}\psi\bar{\psi} \\ &+ \frac{1}{2}\frac{v^2}{4}\Big\{g_2^2\big[(W_\mu^1)^2 + (W_\mu^2)^2\big] + \big(-g_2W_\mu^3 + \frac{1}{2}g_1B_\mu\big)^2\Big\},\end{aligned} \tag{1.3}$$

from which fermion masses are resolved as $m_f = yv/\sqrt{2}$ and the three eigenstates and corresponding masses for the gauge bosons as

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}}(W_\mu^1 \pm iW_\mu^2) && |\, m_W = \mathrm{g}v/2, \\ Z_\mu &= \cos(\theta_W)W_\mu^3 - \sin(\theta_W)B_\mu && |\, m_Z = m_W/\cos(\theta_W), \\ A_\mu &= \sin(\theta_W)W_\mu^3 + \cos(\theta_W)B_\mu, \end{aligned} \tag{1.4}$$

where A_μ is a single eigenstate for a massless boson field (photon) and $\theta_W = \cos^{-1}(g_2/\sqrt{{g_1}^2 + {g_2}^2})$ is the Weinberg angle. Three of the four degrees of freedom (DoF) of ϕ are absorbed by the massive gauge bosons whereas the fourth DoF is $h(x)$, the Higgs boson, which couples to all massive particles. The SM fermions and

	Leptons				Quarks			
Generation	Particle		Charge	Mass [GeV]	Particle		Charge	Mass [GeV]
First	electron neutrino	ν_e	0	$< 2 \times 10^{-9}$	up	u	$+\frac{2}{3}$	0.002
	electron	e^-	-1	5×10^{-4}	down	d	$-\frac{1}{3}$	0.005
Second	muon neutrino	ν_μ	0	$< 1.9 \times 10^{-4}$	charm	c	$+\frac{2}{3}$	1.275
	muon	μ^-	-1	0.1057	strange	s	$-\frac{1}{3}$	0.096
Third	tau neutrino	ν_τ	0	$< 18.2 \times 10^{-3}$	top	t	$+\frac{2}{3}$	173.2
	tau	τ^-	-1	1.777	bottom	b	$-\frac{1}{3}$	4.660

Table 1.2: Schematic overview of SM fermions [9]. Anti-fermions (not shown) possess identical properties up to conjugated charge.

Particle		Charge	Spin	Mass [GeV]
Photon	γ	0	1	0
Weak	$W^\pm$	± 1	1	80.4
Weak	Z^0	0	1	91.2
Gluon	$g\,(\times 8)$	0	1	0
Higgs	H	0	0	125

Table 1.3: Schematic overview of SM bosons [9].

bosons and their chief properties are summarised in Table 1.2 and Table 1.3, respectively. There is an important caveat to Table 1.2 with regard to neutrino masses. By the SM, neutrinos are massless particles. Insertion of the Higgs doublet into the SM Lagrangian simultaneously gives mass to charged leptons and yields neutral leptons as massless. A hypothetical renormalisable effective Yukawa term added to the Lagrangian which would have yielded massive neutrinos, would also break the global baryon and lepton flavour symmetry, $U(1)_B \times U(1)_\tau \times U(1)_\mu \times U(1)_\tau$, otherwise known as lepton flavour violation (LFV) [11]. However, since the SM had been formulated, observations suggestive of phenomena not predicted by the SM have called for BSM extensions to be proposed. Key amongst these were quantum mechanical processes coined neutrino oscillations, indicative of non-zero neutrino masses. These proposals have been put forward to explain experimental measurements of neutrino oscillations; since the late 1990s limits on neutrino masses have been tightening through collaborations primarily focusing on oscillations of solar and atmospheric neutrinos.

The quark fields in Table 1.1 may be described as mass eigenstates, leading to the SM quarks summarised in Table 1.2. Such parameterisation lead [10] to a 3×3 unitary matrix which describes transitions between quark types through weak interactions, namely the Cabibbo–Kobayashi–Maskawa (CKM) matrix, V_{CKM}. Its diagonal elements are $\mathcal{O}(10^1)$–$\mathcal{O}(10^2)$ greater than its off-diagonal elements, thereby allowing relatively low rates for cross-generation quark transitions. The post-electroweak symmetry breaking SM Lagrangian contains V_{CKM} elements in its charged-current (CC) terms,

$$\mathcal{L} \supset -\frac{g}{2\sqrt{2}}[\overline{u}_i\gamma^\mu(1-\gamma^5)V_{ij}d_jW_\mu + \bar{d}_j\gamma^\mu(1-\gamma^5)V_{ij}^*u_iW_\mu^\dagger], \tag{1.5}$$

where $\gamma^\mu \in \{\gamma^0, \gamma^1, \gamma^2, \gamma^3\}$ are Clifford algebra representations, $\gamma^5 = \mathrm{i}\gamma^0\gamma^1\gamma^2\gamma^3$ and $V_{ij} \in V_{\mathrm{CKM}}$. This coupling between quarks and $W^\pm$ bosons closely relates to symmetry conservation in the SM: unlike other interactions, this one uniquely allows for a violation referred to as charge-parity (CP) symmetry breaking. Occurence of this phenomenon at rate higher than permitted by the SM may be related to universal matter-antimatter imbalance. V_{CKM} contains four independent real parameters, values of which have been attained to higher precision over time and compiled into a triangle known as the Unitarity Triangle. A similar behaviour is manifested by neutrino oscillation, mentioned above. Neutrino flavour eigenstates are described as superpositions of their mass eigenstates [11]. The amplitude of a neutrino being in one mass eigenstate or the other is given by the 3×3 unitary Pontecorvo–Maki–Nakagawa–Sakata (PMNS) matrix.

The effective coupling of SM forces are energy-scale dependent. Treated with perturbation theory, the relation between the two is governed by the Callan–Symanzik equations [12, 13] and encoded in a β-function,

$$\beta(g) = \frac{\partial g}{\partial \log E}, \tag{1.6}$$

where $\beta(g) \neq 0$. Uniquely for the theory of the strong interaction, quantum chromodynamics (QCD), $\beta(g)_{\mathrm{QCD}} \propto -g^3 < 0$, which means that the strong force coupling is stronger at lower energies (confinement) and weaker at higher energies (asymptotic freedom). Aside from the top quark which decays before QCD confinement ($\tau_t = 3 \times 10^{-25}\,\mathrm{s}$), the other five quarks' lifetime is sufficiently long for them to be confined by the strong force into colour-neutral composite particles, hadrons. Baryons are hadrons made of an odd number of quarks — typically three — and mesons are hadrons made of an even number of quarks — typically two. Hadron masses are usually several orders of magnitude greater than the combined masses of their constituent quarks thanks to the binding energy of the strong force. In turn, this results in frequent exchange of gluons between hadron-bounded (valence) quarks.

The nature of the strong coupling calls for a more precise scheme to describe quarks and their confinement to hadrons. Let a be a parton coming from a hadron A. It carries a fraction $x_A \in [0, 1]$ of A's momentum, such that the probability density to find it is $f_{a/A}(x_A)$, known as the Parton Distribution Function (PDF) [14]. A crucial adjustment is that PDFs also depend upon the energy scale $|Q|$ at which A is probed and hence written generally as $f(x, Q^2)$. By a notion known as Bjorken scaling, at high $|Q|$ interacting hadrons behave as collections of point-like constituents, implied by asymptotic freedom. Therefore, at any given moment, hadrons are "composed" of their valence quarks inasmuch as their PDF predicts it, depending on the momentum fraction carried by the constituents and energy scale to which they are probed.

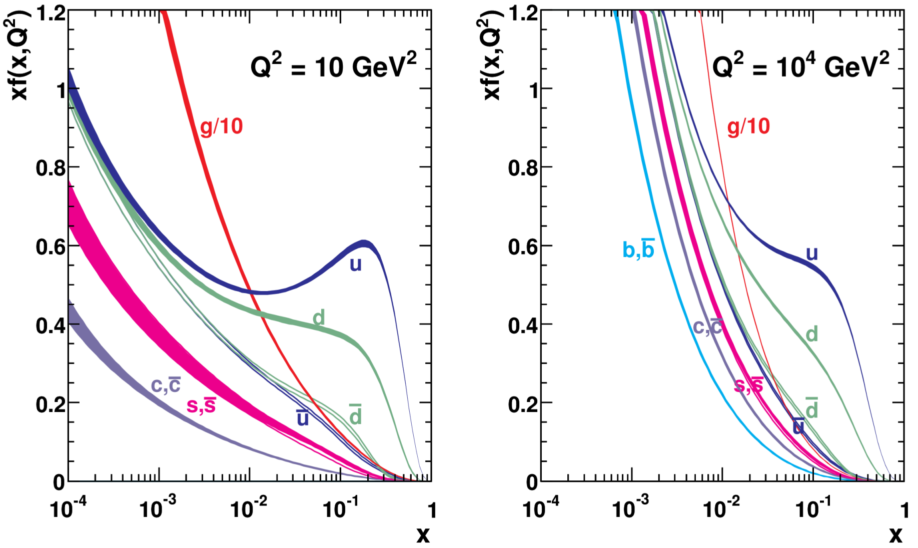

Figure 1.1: Parton Distribution Function (PDF) of partons at two energy scales, $\sqrt{10}$ GeV (left) and $\sqrt{10^4}$ GeV (right), calculated to Leading Order (LO) [15].

N.b. to LO this scaling effect need not necessarily depend on absolute energies of different colliders, but more so on dimensionless quantities like the ratio of energy to momentum transfer or the scattering angle. At the LHC, when two protons undergo $\sim 10\,\text{TeV}$ collisions, they pass through one another at $\sim 10^{-29}\,\text{s}$ timescales. Conversely, parton creation and annihilation dynamics inside protons have an upper limit of $\sim 10^{-24}\,\text{s}$. This discrepancy implies that pp collisions at the LHC are effectively parton-parton scatterings which may be described with the aid of PDFs. Figure 1.1 shows the PDF of partons for two energy scales. The up and down quarks dominate at high x whilst for low x all partons tend to diverge (more quickly for the high energy scale). For high energy ranges, there are considerable discrepancies between light quarks and antiquarks whereas the heavy quarks and antiquarks have a shared distribution throughout the range, for both scales. This results from a difference between valence quarks and sea quarks, where the latter constitute quark-antiquark pairs produced through gluon splitting which do not bear an influence on the composition of hadrons. Across both scales the gluons dominate for $x \lesssim 10^{-2}$.

The nature of QCD [1,14] limits the energy scales for which perturbation theory is applicable. Since the strong force is associated with a running coupling, $\alpha_\text{s}(k^2)$, interaction energies ought to be high enough for $\alpha_\text{s}(k^2)$ to be low enough. For instance, $\alpha_\text{s}(m_s \sim 0.1\ \text{GeV}) = 1.7$ although $\alpha_\text{s}(\sim 1000\ \text{GeV}) = 0.09$, such that at LHC energies where $\sqrt{s} \sim 10$ TeV, α_s is sufficiently-small to be treated by perturbation theory, *i.e.* a parameter in which a series expansion is valid. However, when dealing with hadrons the case is subtler. In proton-proton collisions, calculation is divided into a perturbative part formalised by the partonic portion of the $\text{p}(p_A) + \text{p}(p_B) \rightarrow X + Y$ cross-section, $\hat{\sigma}$, where X and Y are final states and a non-perturbative part contained in the PDF. The two are jointly computed by a convolution:

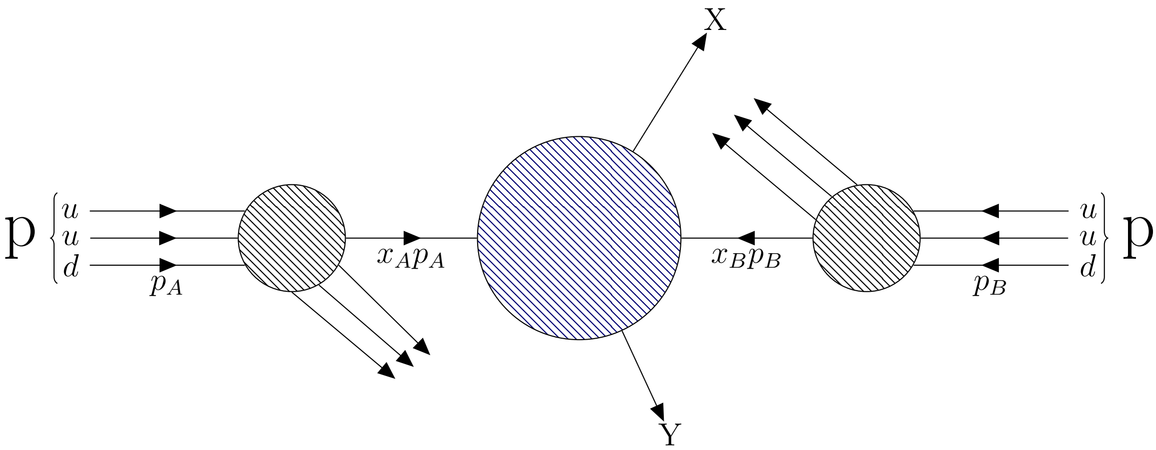

Figure 1.2: Schematic proton-proton collision Feynman diagram. At high $|Q|$ the valence quarks of hadrons A and B can be treated as asymptotically-free partons which effectively participate in the scattering. Each interacting parton carries momentum $x_i p_i \mid i \in \{A, B\}$ as indicated by the Parton Distribution Function (PDF). The final-state particles X and Y are determined by the cross-section, depending on the residue energy.

$$\sigma(\mathrm{p}(p_A) + \mathrm{p}(p_B) \to X + Y) = \int_0^1 \mathrm{d}x_A \mathrm{d}x_B \sum_{i,j} f_i(x_A, Q^2) f_j(x_B, Q^2) \hat{\sigma}(\psi_i(x_A p_A) + \psi_j(x_B p_B) \to X + Y), \tag{1.7}$$

where the sum runs over all possible initial states. $\hat{\sigma}$ is often referred to as the hard part and the PDF as the normalisation part. The PDFs cannot be calculated but may only obtained by experimental measurements. In order for the momentum sum-rule $\sum_i \int_0^1 x p_i \mathrm{d}x$ to yield unity, QCD splitting *i.e.* gluon interchange amongst participating quarks, must be considered. The complete picture of momentum distribution amongst partons present in a *pp* collision may not neglect gluons, which in total can carry up to nearly 50% of the overall proton momentum. A Feynman diagram schematically depicting such a process where the hard scatter is distinctly separated from the accelerated protons is given in Figure 1.2.

Therefore, for scale separation of hadron collisions one should factorise the non-perturbative dynamics from the perturbative hard cross-section. Nominally, this separation is determined by a factorisation scale 1 GeV$^2 < \mu^2 < Q^2$. In a similar vein with running coupling, μ^2 can be regarded a running parameter. Its variation can be related to the energy scale and PDF via renormalisation group equations, analogous to Eq. (1.6). These relations are described by the Dokshizer–Gribov–Lipatov–Altarelli–Parisi (DGLAP) evolution equations (not shown) [16–18]. Although *pp* collisions are well-described via parton-parton scattering, outgoing final-state partons are yet to be directly observed. As partons travel away from the interaction point (IP), radiations of quarks and gluons are formed, otherwise known as parton showers. Once enough energy has dissipated and these showers are of energy scales $\lesssim 1$ GeV, α_s grows greater and confinement dominates over asymptotic freedom. This process leads to hadronisation, manifested in a collimated spray coined hadronic jet. At the LHC, jets are common due to the quark and gluon radiation from the accelerated protons and more importantly from the outgoing partons generated by the hard scatter of the pp collisions. Jets embed a cumbersome radiation pattern which will be discussed in subsequent chapters.

1.2 Physics Beyond the Standard Model

However far-reaching and accurate, the SM does not provide a complete description of Nature. The previous section discussed the SM and showed its mathematical illustration of the workings of the electromagnetic, weak and strong interactions, as well as how fermions and gauge bosons acquire mass through the Higgs mechanism. There exists several observations in the Universe which are beyond the scope of the SM whose explanation entails BSM physics. The salient phenomena of such nature are briefly noted below.

Neutrino masses

As discussed in the previous section and in [11], observation of neutrino oscillations [19] provides evidence of massive neutrinos which contradicts the zero-mass neutrinos notion predicted by the SM.

Dark Matter

Primarily stemming from astrophysical observations, there exists an abundance of evidence to suggest that the majority of the matter content of the Universe is currently unaccounted for. This content is a non-luminous, weakly-interacting matter type called Dark Matter (DM) [20]. Many models seeking to incorporate DM into the a theory of particle and forces have been put forward, prominent amongst which is the existence of Weekly-Interacting Massive Particles (WIMPs). Numerous endeavours in high energy physics (HEP) are trying to find particles which either themselves constitute DM or mediators thereof.

Quantum Gravity

A fundamental physical sector not included in the SM is gravity. Hitherto, a coherent quantum theory of gravity has not been proposed. For all intents and purposes, the SM augmented with General Relativity is the current description of Nature to the most accurate level known. There exists new theoretical routes from QFT which might unify particle physics with quantum gravity. However, presently these are far from ready to be searched-for experimentally.

Matter-antimatter asymmetry

The only SM parameter permitting CP violation is the complex phase in the CKM matrix, as given in Eq. (1.5). Nonetheless, these falls short from allowing the amount of CP violation required in order to account for the rate of matter-antimatter asymmetry in the Universe. Hence, BSM CP-violating phenomena are needed to explain the baryonic asymmetry dominating the Universe and perhaps answer profound questions in early-Universe cosmology.

The Hierarchy Problem

Unlike fermions and gauge bosons, the Higgs boson has got no gauge symmetry to protect its mass from growing higher by quantum corrections. Effectively, $m_h << M_\mathrm{P}$ where $M_\mathrm{P} \sim 1 \times 10^{19}\,\mathrm{GeV}$ is the Planck mass, although this is not an expected behaviour: a particle's mass is calculated to an arbitrary order in perturbation theory; every quantum correction beyond tree-level should add to its mass. The Higgs mass can gain contributions from high-mass corrections such as m_t, the mass of the top quark. However, $M_\mathrm{P}/m_h \sim 10^{16}$ constitutes an extreme disparity, fundamental to our understanding of the SM's reach and the profound reason for the Higgs mass to be $\mathcal{O}(10^2)\,\mathrm{GeV}$ although unbounded from above. If anticipated quantum corrections to m_h were somehow cancelled by an unknown field, this may be regarded as hints for New Physics (NP). The Hierarchy Problem [21] maintains a potential of substantial influence and might be related to notions such as supersymmetry and quantum gravity.

Chapter 2

Experimental Apparatus

The work reported herein is based upon data collected in the particle physics laboratory at the European Organization for Nuclear Research (CERN), located in Geneva, Switzerland. CERN is associated with over 13,000 researches in experimental HEP, representative of more than 110 nationalities, who are employed by universities or laboratories across over 70 countries. These researches typically work within any independent experimental collaboration along the LHC beamline shown in FIG 2.1 or on projects directly related to the accelerator apparatus shared by the experiments. There are four large experimental collaborations currently active in CERN, all along the LHC beamline: ALICE, LHC*b*, CMS and ATLAS. Whilst CMS and ATLAS constitute multi-purpose detectors which serve a wide variety of physics programmes, ALICE and LHC*b* are specialised experiments for the study of heavy-ion collisions and *b*-quark physics, respectively. Below, a brief summary of the LHC in Section 2.1 will be followed by a discussion of the different parts and purposes of the ATLAS detector in Section 2.2.

2.1 The Large Hadron Collider

The LHC [23] is a 27 km-long circular particle accelerator located at an average 100 m depth below the surface of the Earth. It is primarily intended for generating pp collisions via two counter-rotating proton beams which collide head-to-head at different interaction points (IP) along the ring. The LHC has reached a maximal centre-of-mass (CoM) proton acceleration energy of $\sqrt{s} = 13\,\text{TeV}$.

Since the LHC is built in the already-existing tunnels of the previous Large Electron-Positron (LEP) Collider, in order to reach CoM collision energies of $\mathcal{O}(10)$ TeV with a fixed accelerator radius (R) it was required to increase the magnetic field strength applied on the beam (B). The magnetic fields for the LHC are achieved via novel superconducting niobium-titanium (Nb-Ti) electromagnets, kept at sub-10 K temperatures. These cryogenic temperatures are accomplished by submerging the Nb-Ti alloy in superfluid Helium, of which over 120×10^3 kg are used at the LHC.

Since the LHC is a particle-particle collider it requires two oppositely-polarised magnetic fields to rotate one of its beams in the clockwise direction and the other in the anti-clockwise direction. This led to a so-called "2-in-1" design of the LHC's

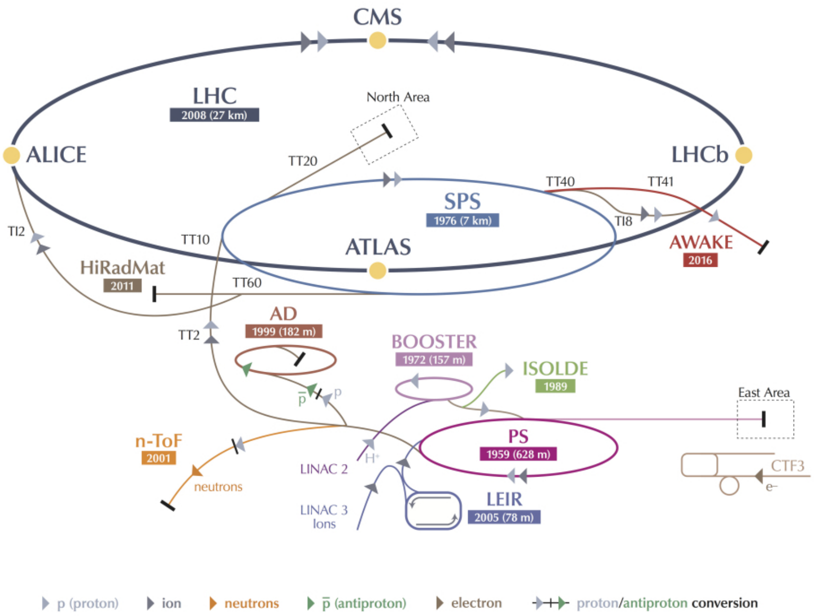

Figure 2.1: Illustration of CERN's Accelerator Complex. The constituent beamlines, accelerators, storage rings and independent experiments are marked. The accelerated protons which eventually partake in the $p-p$ collisions at the IP are trajected along a path: LINAC 2 → Booster → Proton Synchotron → Super Proton Synchotron (SPS) → LHC [22].

bending magnets which has the two beam pipes contained in the same cooling chambers (cryostats) with the counter-rotating beams. The construction relies on the 3.7 m-wide LEP tunnel, where for the LHC each of the 15 m-long superconducting dipole electromagnets responsible for maintaining the proton beams in orbit. There is a 11850 A current flowing through the magnets which produce a field strength of 8.33 T.

Hence the LHC is effectively a chain of extremely-cooled superconducting dipole magnets bending two opposite proton beams. The ring is of a double-octagonal structure, partitioned such that each of the eight octants are straight at their centre. The straight sections house the interaction regions, referred to as Points. Interaction Points (IP) 1, 2, 5 and 8 are the locations of the four large experiments, respectively: ATLAS, CMS, ALICE and LHC*b*, where the counter-rotating proton beams collide. Other points host various service areas necessary for LHC maintenance and operation. An illustration of the LHC layout is shown in FIG. 2.2.

LHC service points include *e.g.* Point 4 which hosts the superconducting radio-frequency (RF) systems which accelerate the proton beams to their nominal collision energies. Additionally, Point 6 is the location of the beam dump where, via kicker magnets, the accelerated beams can be diverted outside of the LHC ring safely, should a proton refill be necessary or in cases of instability.

The LHC injection chain [25] is responsible for generating the proton source inserted into the collider ring and includes a series of pre-acceleration stages which intensify the initially-low-energy protons up to energies sufficient for them to embark upon their course about the ring. Depicted in FIG. 2.1, the protons sourced

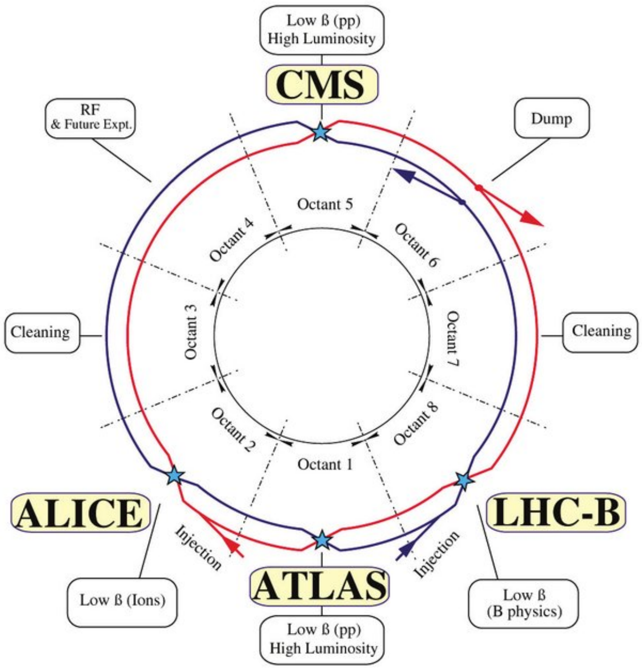

Figure 2.2: Layout of the LHC. The two counter-rotating beams are given: Beam 1 is in blue and rotates counter-clockwise whereas Beam 2 is in red and rotates clockwise. At the centre of each octant the LHC ring is straight, where experimental caverns or beam facilities are housed. The octants boundaries are characterised by curved regions which are primarily composed of dipole magnets. [24].

from H atoms are released at Linac 2, a linear accelerator through which an energy of 50 MeV is gained. Then the protons enter the Booster, an apparatus made up of four stacked rings which circularly accelerates them to 1.4 GeV. Subsequently, they are accelerated further to 25 GeV by the Proton Synchotron (PS) and yet further to 450 GeV by the Super Proton Synchotron (SPS). At this point the protons are adequately energised to be injected into the LHC ring. This is done from two injection points, one for the counter-clockwise-rotating beam and the other for the clockwise-rotating beam (blue and red beams in FIG. 2.2, respectively). The injection intervals are called fills and until all protons pertaining to a single fill travel into the ring, they are kept circulating at the injection energy, 450 GeV. Once a fill is completed, the superconducting RF cavities at Point 4 accelerate the protons up to their final collision energy.

Protons travelling through the LHC ring are arranged into bunches by the RF cavities. These bunches, amounting to $\sim 10^{11}$ protons each, are shaped by the RF field such that the nose and tail of each bunch are slowed-down or sped-up, respectively, in order to maintain the desired bunch structure throughout the collider. The RF cavities exert a field oscillating at a frequency of 400 MHz, according to which the boundaries of the proton bunches are contained, otherwise known as RF buckets, which are limited to to about 35640 in number. A a minimal separation of 10 bucket is kept between the LHC bunches (effectively only 2808 bunches are filled). This corresponds to a minimal time between bunches — the bunch spacing — of 25 ns (equivalent to ~ 7.5 m spacing) corresponding to a maximum collision rate of 40 MHz.

The instantaneous luminosity [26] of the collider, the rate of scattered events per scattering cross section. It is approximated by

$$\mathcal{L} = \frac{N^2 n_b f}{4\pi\sigma_x\sigma_y} S, \tag{2.1}$$

where N is the number of particles per bunch, n_b the number of colliding bunches, f the bunch frequency, $\sigma_{x,y}$ the Gaussian-approximated transverse beam widths and S the reduction parameter accounting for geometrical factors such as a non-zero crossing angle between the colliding beams. The instantaneous luminosity has units of $1/\text{cm}^2\text{s}$. The LHC is designed to provide an instantaneous luminosity of $\mathcal{L} = 10^{34}\,\text{cm}^{-2}\text{s}^{-1}$.

An accurate knowledge of $\mathcal{L}$ is important for any measurement at collider experiments due to its direct relation to the predicted number of events of interest, $N_{\text{coll.}}$,

$$N_{\text{coll.}} = \sigma_{\text{coll.}} \int \mathcal{L}\text{d}t = \sigma_{\text{coll.}} L, \tag{2.2}$$

where $\sigma_{\text{coll.}}$ is the cross-section of the process of interest and L the time-integrated luminosity units of which are cm^{-2}. In particle physics, integrated luminosity is given in units of inverse barns, defined such that $1 \times b = 10^{-24}\text{cm}^{-2}$, where b is a single barn.

LHC running commenced in 2009 and has remained in stable operation since. Its operation cycles are named Runs, multi-year periods of continuous data collection. During Runs, there exist few-months-long spans during which maintenance

Run			1		2			
Year			2011	2012	2015	2016	2017	2018
Parameter	Notation	Units						
Collision energy	$\sqrt{s}$	TeV	7	8	13			
Peak luminosity	$\mathcal{L}$	$\mathrm{cm}^{-2}\mathrm{s}^{-1} \times 10^{34}$	0.36	0.77	0.5	1.4	2.1	2.1
Integrated luminosity	L	fb^{-1}	5.5	20.2	3.2	33.0	44.3	59.9
Mean interactions per crossing	$\langle\mu\rangle$	N/A (tally)	9.0	20.7	13.4	25.1	37.8	36.1

Table 2.1: Summary of key parameters for data collection periods of the LHC. For purposes of data analysis, the integrated luminosity is most relevant. This parameter varies when measured in the LHC generally compared with its values when delivered to specific experiments due to stringent quality requirements.

and upgrades take place. In between Runs, large-scale projects of repair and improvements to the LHC and experiments therein take place, otherwise known as Long Shutdowns. Hitherto, there have been two LHC runs, namely Run 1 and Run 2, which lasted between 2009–2012 and 2015–2018, respectively. This thesis was written during Long Shutdown 2 and is based on data taken during Run 2. The different Runs reflect increasing luminosities and energies of the LHC, as can be appreciated through Table 2.1, along with other parameters such as the mean number of interactions per bunch crossing, $\langle\mu\rangle$, which is especially relevant to the calculation of pileup, caused by overlapping *pp* collisions within the same or neighbouring bunch crossing(s) at the IP.

2.2 The ATLAS Detector

Multi-purpose, large and complex, the ATLAS detector is located at Point 1 of the LHC ring *cf.* FIG. 2.2. It is cylindrically shaped, 44 m long, 25 m tall and coaxial with the LHC beam pipe. ATLAS detector possesses a nearly-4π solid angle coverage around the *pp* IP. In order to identify a vast range of different particles, leptons and hadrons alike, it consists of various subdetectors, each of which specifically sensitive to a particular set of particles and interactions. The geometrical layout is often described as onion-shaped: cylindrical subdetectors symmetrically layered about the IP, such that *pp* collisions' final states are detected by one or more of the subdetectors, regardless of their trajectory's direction. This design is portrayed in an overview of the ATLAS detector and its subdetectors in FIG 2.3.

2.2.1 Coordinate System

ATLAS is based upon a right-handed coordinate system whose origin is located at the IP, the centre of the detector, with the z-axis along the beam pipe, for which the positive (negative) direction is referred to as the A (C) side of the detector. The x-axis points from the IP to the centre of the LHC ring and the y-axis points

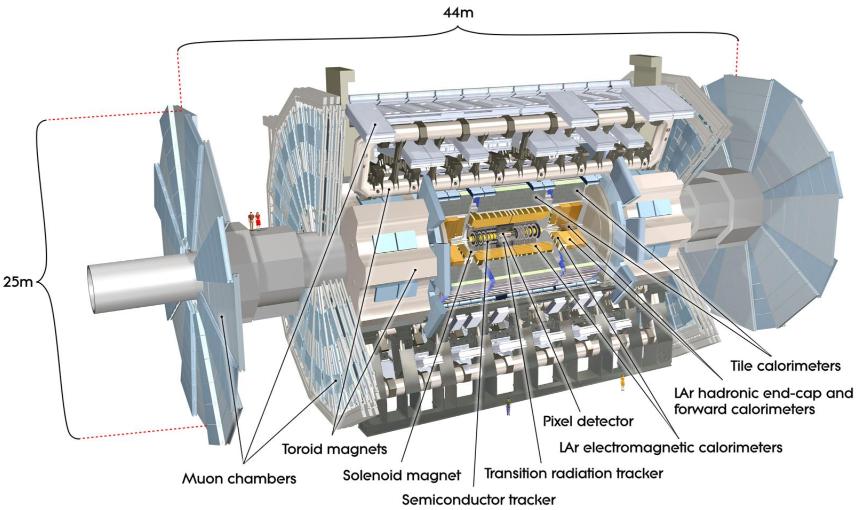

Figure 2.3: Computer-generated overview of the ATLAS detector and its subsystems. For comparison to average-height humans, figures are shown standing on the forward shielding between the Big Wheels of the muon system (in red and crimson) and by the detecor's feet (in yellow) [27].

upward, away from the centre of the Earth. Cylindrical polar coordinates (r, ϕ, z) are used in the transverse plane, ϕ being the azimuthal angle around the beam pipe such that $\phi = 0$ would be along the positive x-axis. The pseudorapidity is defined as $\eta = -\ln[\tan(\theta/2)]$ where the polar angle θ is set such that $\theta = 0$ indicates parallel to the beam pipe and $\theta = \pi/2$ directs in the x-y plane transverse to the beam pipe. Large (small) η values correspond to the forward (central) region of the detector. Transverse momentum, $p_{\mathrm{T}} = p \sin\theta$, and transverse energy, $E_{\mathrm{T}} = E \sin\theta$, are defined where p and E are the magnitudes of the total momentum and energy, respectively. The rapidity, y, is related to η and given by $y = \frac{1}{2}\ln[(E + p_z)(E - p_z)]$, such that $y = \eta$ for massless or ultra-relativistic particles. Trajectories through the ATLAS detector are often described using $(p_{\mathrm{T}}, \eta, \phi)$, whereas distance between two trajectories is typically provided by $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$, referred to as the boost-invariant angular distance.

2.2.2 Inner Detector

Examining ATLAS from the inside out, in the innermost layer the Inner Detecor (ID) [28] is placed, covering the region of $|\eta| < 2.5$. In order of increasing radial distance from the beam pipe (outward), the ID consists of the silicon Pixel detector, the silicon-strip Semiconductor Tracker (SCT) and the Transition Radiation Tracker (TRT), as seen schematically in FIG. 2.3. These subdetectors reconstruct tracks associated with charged particles emerging from bunches undergoing pp collisions every 25 ns. An illustratin of the ID is given in FIG. 2.4.

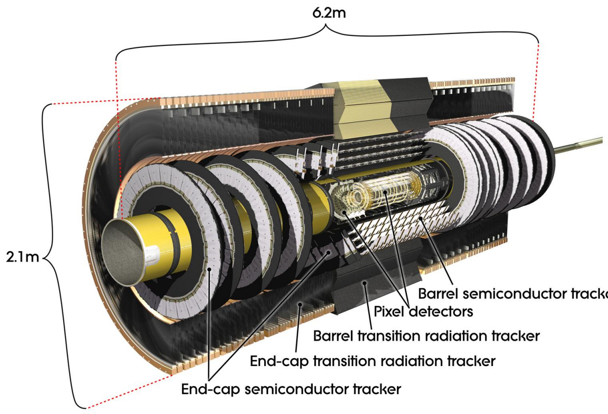

Figure 2.4: Depiction of the ATLAS Inner Detector and its subdetectors from a cross-sectional view. The barrel and end-cap segments of the Pixel Semiconductor Tracker (SCT) and Transition Radiation Tracker (TRT) detectors are shown. [29].

Pixel Detector

The innermost subdetector in the ID, Pixel surrounds the beamline at a close distance. Its design consists of a barrel section and two end-cap sections. The former is a four-layered portion of cylindrical geometry and the latter are three-layered disk-shaped parts positioned on the beam pipe. The barrel section of the ID is depicted in FIG 2.5.

Being so near the IP, the Pixel detector is subject to the greatest particle fluxes of any other subsection of ATLAS is. Resultantly, its design is characterised by fine granularity attained by 250 μm thick segments composed of $50 \times 400\,\mu\text{m}^2$ large silicon semiconductor pixels. Thereby, the pixel detector achieves a remarkably high spatial hit resolution of 10 μm in the r-ϕ plane and 115 μm along the z-axis. Within the Pixel detector's barrel section, the innermost layer was installed during Run 2 and is referred to as Insertable B-Layer (IBL). Possessing 8×10^6 readout channels of the Pixel detector, the IBL is located at a radial distance of 33 mm from the centre of the beamline. This leads to its extraordinarily precise spatial hit resolution of 8 μm in the r-ϕ plane and 40 μm along the z-axis. In addition to its enhancement of overall measurement and reconstruction of charged particle trajectories, the IBL is aimed to improve the reconstruction of secondary vertices, an important feature for the identification of hadronic jets stemming from b-quark decays.

Semiconductor Tracker

Similarly to the Pixel detector, SCT also employs silicon sensing devices. As shown in FIG. 2.5, the SCT surrounds the Pixel detector, where they possess comparable geometry. The SCT's barrel portion consists of four cylindrically shaped layers and nine disk-like end-cap layers. Its silicon sensing devices are 80 μm wide strip-shaped elements. Those in the barrel are parallel to the beamline whereas those in the end-caps are perpendicular to the beamline, positioned along the radial direction. Both the barrel and end-caps sections of the SCT contain thin stereo strips which enable measurements of r, ϕ and z. The SCT reaches a spatial resolution of 17 μm

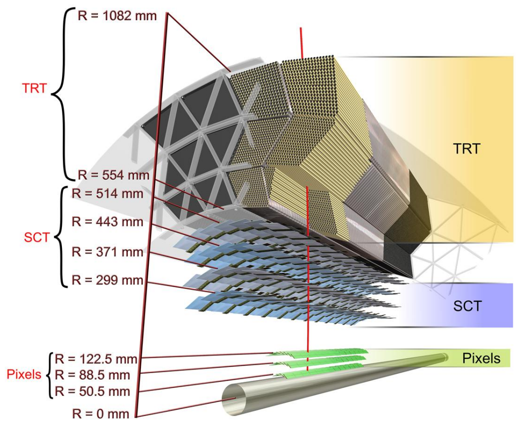

Figure 2.5: Cut-away illustration of the barrel section of the ATLAS Inner Detector. The three constituent subdetectors, Pixel, Semiconductor Tracker (SCT) and Transition Radiation Tracker (TRT) are marked in respective order from the beamline outward. [29].

in the r-ϕ plane and 580 μm along the z-axis.

Transition Radiation Tracker

Surrounding the SCT, the outermost layer of the ID is the TRT. A tracking device, the TRT is based upon 4 mm-wide polyimide proportional drift tubes. Its barrel contains 73 layers of 144 cm-long tubes and runs parallel to the beamline whilst its (two) end-caps have 160 tube planes each. Every such plane is composed of 37 cm-long tubes, formed circularly into wheels as can be seen in the top section of FIG. 2.5. The TRT's drift tubes are coined straws, each of which contains a 31 μm-thick wire coated with gold at its centre. These wires perform as anodes and are grounded, whilst the inner walls of each straw are subject to an electric potential of ~ 1.5 kV. The straw volume is filled with a gas mixture of 70% Xe-27%-CO_2-3% O_2 such that an indication of a charged particle hit through a straw is the outcome of ionisation and the subsequent drift of electrons towards the central wire. The drifting charge yields an electrical current which is thereafter converted to a readout signal.

The TRT only provides hit information in the r-ϕ plane with a spatial resolution of 130 μm per straw. Although pertaining to a meagre precision compared to the silicon-based subdetector systems, the TRT compensates by being able to measure extensively longs tracks: a single charged-particle trajectory hits the TRT 36 times on average. Hence the longest track reconstruction through the ID is provided by the TRT. Moreover, the polypropylene fibres by which the straws are separated generate transition-radiation photons. The particular photon pattern relates to the mass of the hitting particle such that the TRT adds to the discriminative qualities of the ATLAS detector. One such example is the separation between electrons and charged pions which is usually obtained from the calorimeter subsystems but also usefully complemented by the TRT.

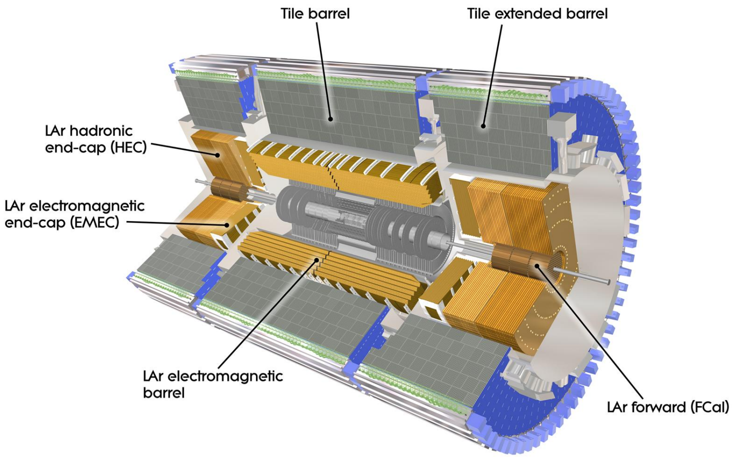

Figure 2.6: Illustration of the ATLAS calorimeter with indications of liquid-argon (LAr) and tile scintillator subsystems. [32].

2.2.3 Calorimeters

ATLAS includes two calorimeter detectors, namely the electromagnetic calorimeter (ECAL) and the hadronic calorimeter (HCAL). They are positioned outside the ID and responsible for measurement of showers of electrically-charged as well as neutral, particles. The ECAL covers an η range matching the ID's coverage and is optimally-designed for precision measurements of electrons and photons. A subdetector of coarser-granularity, the HCAL is especially-purposed for jet measurements where the readout of transverse momentum is key. Apart from discerning between calorimeter types based upon the particles they are measure, ATLAS calorimeter subsystems may also be classified according to their detection technology (active medium in use): one class [30] has gaps of cooled liquid argon (LAr) installed whereas the other [31] utilises scintillation tiles. An outline of the calorimeter, including the latter categorisation, is given in FIG. 2.6.

Electromagnetic Calorimeter

The ECAL is a high-granularity lead (PB)/LAr subdetector located outside the ID. Accordion-like shaped to provide a full ϕ coverage, the ECAL's barrel and end-cap sections cover the region $|\eta| < 3.2$ and are indicated in FIG. 2.6. The cooled LAr fills the gaps between the ECAL layers. Particles trajecting through the layers undergo bremsstrahlung due to the lead absorbers and thus produce electrons and photons which ionise the LAr. In a similar manner to the mechanism in the ID's TRT, ionisation in the ECAL induces drift of electrons whose signals are intercepted by the readout electrodes. The drift chamber in use is 2.1 mm-wide and subject to operating voltage of about 2 kV. Due to the fine granularity of the readout, individual photons generated through pion decays, $\pi^0 \rightarrow \gamma\gamma$, can be individually detected. This capacity of the ECAL to distinguish individual photon pairs is deeply important *e.g.* for analyses of the Higgs boson decay mode $h \rightarrow \gamma\gamma$.

An additional subdetector named presampler is installed within the ECAL and

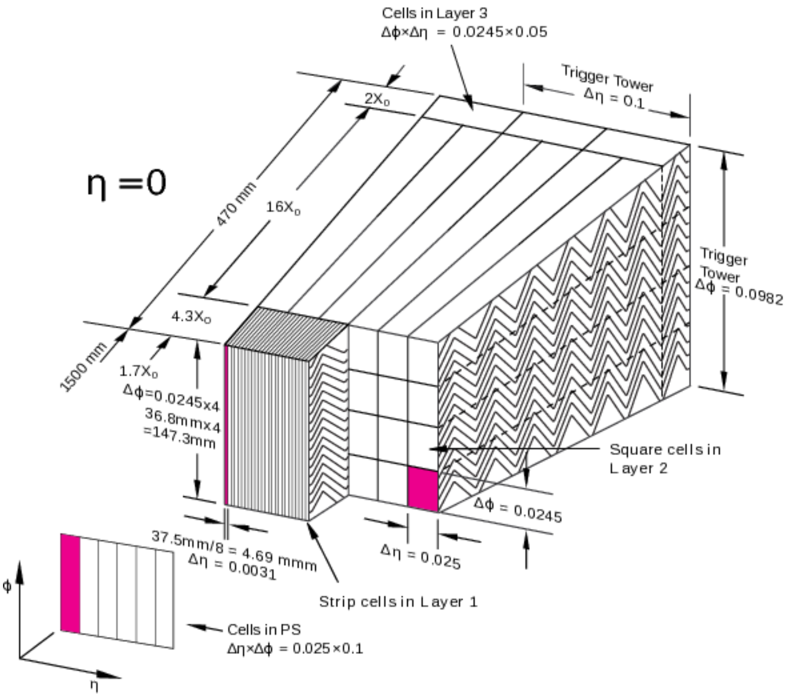

Figure 2.7: Computer-generated image of the ATLAS electromagnetic calorimeter's (ECAL) barrel section. An accordion like-design, the ECAL's absorption features are shown through the three sampling layers. Additionally, the $\Delta\phi \times \Delta\eta$ readout precision, *i.e.* granularity is marked in each subsection. [30].

covers the region where $|\eta| < 1.8$. Its purpose is to provide energy corrections compensating for electron and photon loss due to material interactions taking place in the ECAL. It consists of a single LAr layer, 1.1 cm long in the barrel and 0.5 cm wide in the end-caps. The barrel section of the ECAL is depicted in FIG 2.7.

Hadronic Calorimeter

Positioned beyond the ECAL, the HCAL is a lead scintillator-tile detector in its barrel portion and a copper/LAr-based device in its end-cap section. The tile calorimeter part covers the region where $|\eta| < 1.0$ in its barrel section and $0.8 < |\eta| < 1.7$ through two additional barrel section extensions, as noticeable in FIG. 2.6. A sampling calorimeter, the HCAL employs scintillation polymers as active media and steel as a passive absorber. The scintillating tiles of the tile calorimeter are placed in steel absorbers on the r-ϕ plane. In the radial direction the HCAL is partitioned into three layers whilst the scintillation light readout, performed via wavelength-shifting optial fibres stuffed into photomultiplier tubes (PMTs) is designed projectively. This geometry is portrayed through the sketch in FIG. 2.8.

The end-cap portiones of the HCAL, often abbreviated as HEC, are composed of two wheels each and placed behind the ECAL end-caps, as may be seen in FIG. 2.6. They cover the $1.5 < |\eta| < 3.2$ region and made of copper (Cu) layers interweaved with 8.5 mm-wide LAr divisions which constitute the active medium of the HEC. These LAr divisions are partitioned into drift areas, each uniquely related to readout electrodes, organised in a projective geometry.

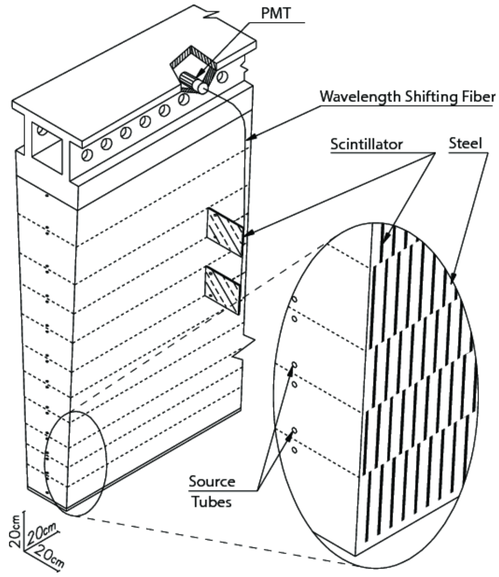

Figure 2.8: Schematic illustration of the ATLAS tile calorimeter. The steel absorbers, scintillating tiles and photomultiplier tubes (PMTs) are marked. Also shown are the source tubes used for calibration [31].

Forward Calorimeter

In order to cover the high-$|\eta|$ region of $3.1 < |\eta| < 4.9$ there is the forward calorimeter (FCal), which is indicated in FIG. 2.6. It is made of three layers in the z-axis direction, namely one electromagnetic layer followed by two hadronic layers. All three use LAr as active medium although employ different passive media: the electromagnetic layer uses copper as an absorber which promotes heat removal. The hadronic layers have tungsten instead, in order to attain high containment and minimise lateral spread of hadronic showers. In the FCal, the passive media are interlaced with readout tubes which run parallel to the beamline and filled with the mutually active medium, cooler LAr.

2.2.4 The Muon Spectrometer

Installed around the calorimeter systems, the muon spectrometer (MS) [33] is designated for the detection of energetic muons originating from pp collisions at the IP. Fundamentally, it relies upon deflection of muon trajectories using magnetic fields and thus measuring their momenta. This deflection is achieved via the ATLAS magnet system, which provides an average field strength of 4 T via toroid magnets installed in the barrel section covering the $|\eta| < 1.4$ range as well as in the end-cap sections covering the $1.6 < |\eta| < 2.7$ range. As may be inferred from FIG. 2.3, the vast portion of volume outside the calorimeter systems within the ATLAS detector is occupied by the MS. Indeed, the overall design of the toroid structure is governed by the immense size of the MS, which also heavily influences the design of the other subdetectors.

The MS employs four types of gaseous radiation detectors which are designed in chambers, based on the notion of projective towers. These chambers are arranged

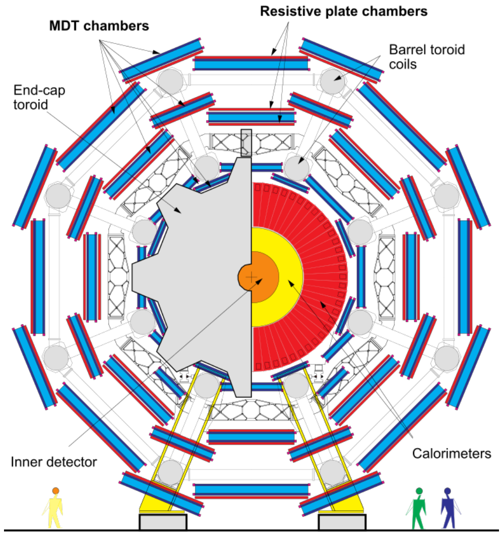

Figure 2.9: Cross-sectional view of the barrel muon spectrometer perpendicular to the beam axis (non-bending plane). The monitored drift tube (MDT) chambers are shown in sky blue and the resistive plate chambers(RPC) chambers in red. The eight coils are marked in grey [33].

accordingly with the toroid magnet structure and have a circular 16-fold segmentation. They are partitioned into large and small sectors, where the former cover the regions between the toroid coils and the latter cover the range where the coils are installed. The detector types may be discriminated into two groups, namely precision or trigger chambers. The former are built from Monitored Drift Tube (MDT) and Cathode Strip Chamber (CSC). These precision chambers provide measurements of muon tracks as they propagate through the MS, specifically their detection and precise measurement in the plane onto which they are magnetically bent and thus attain accurate measurements of muon momenta via their curvature. The trigger chambers are composed of Resistive Plate Chamber (RPC) and Thin Gap Chamber (TGC) subdetectors and yield fast signal readout times which enable accurate relation between muon tracks through the MS to a specific pp bunch crossing. Both the precision and trigger chambers are installed in the barrel and end-cap sections of the MS. Throughout the entire $|\eta|$ range, there are three layers of precision chambers which are required for pointed measurements of specific muon tracks to determine their momenta. An illustration of the MS with its detection subsystems is given in FIG. 2.9.

The combination of small and large sectors in the MS allows for a complete azimuthal coverage. Nevertheless, in the $|\eta| \sim 0$ and $|\eta| \sim 1.2$ regions there are noticeable drops in chamber coverage. The former is applied to enable ID and calorimeter services and the latter accounts for the barrel/end-cap transition regions. Conversely, the regions where there are overlaps between large and small sectors have an increase of muon hits. A view of the MS subsystems is given in FIG. 2.10. The different detector technologies operating in the barrel and end-cap portions of the MS are succinctly described below.

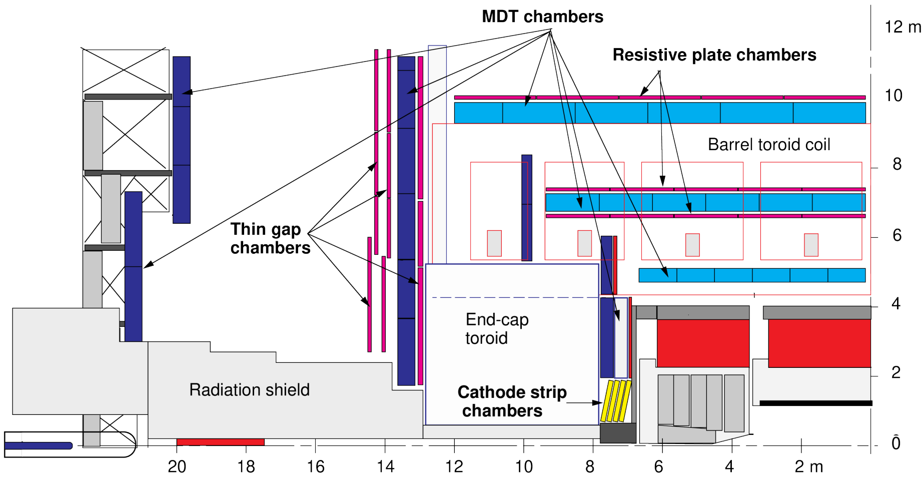

Figure 2.10: Computer-generated illustration of a quadrant of the muon spectrometer (MS) in the r-z plane with indications of all four detection technologies employed. The monitored drift tube (MDT) chambers are marked in sky blue at the barrel and in purple at the end-caps. In pink, the resistive plate chambers (RPCs) are indicated at the barrel and the thin gap chambers (TGCs) at the end-cap. Shown in red are the barrel toroid coils and, in yellow, the CSCs. The light grey rectangles at $6\,\mathrm{m} < r < 9\,\mathrm{m}$ are indicative of the barrel toroid structures' locations [33].

Barrel

Rectangular muon chambers are arranged in three cylindrical shells concentric about the beamline at a radial distance if 5 m, 7.5 m and 10.5 m, *cf.* FIG. 2.9. The precision chambers in the barrel include MDT chambers with tubes perpendicular to the beamline and parallel to the magnetic field, thereby enabling precision measurements along $|\eta|$. These tubes are 3 cm-wide in diameter and contain a 93% Ar − 7% CO_2 gas mixture and a wire operating under 3 kV voltage. Passing minimum-ionising particles (MIP) ionise the contained gas and the induced signal is read-out. The resulting spatial resolution of an MDT is less than 100 μm. The chambers housing the MDTs are stacked as layers, a design by which the spatial resolution reaches as fine as 50 μm, once data stemming from different individual layers are combined. Responsible for constructing muon triggers in the MS barrel are the RPCs. They are 2 mm-gapped chambers filled with tetrafluorethane ($C_2H_2F_4$) and lined with parallel plate electrodes subject to a voltage of 9.8 kV. This combination of a high-voltage gas chamber attain a timing resolution of 2 ns. Strips perpendicular to the beam-axis read-out the charge yielded from the ionising events in the gas gap and thus give additional spatial information which add to the construction of tracks and triggers in the MS. A side-view of an RPC indicative of its *modus operandi* is given in FIG. 2.11.

End-caps

Installed at $1.0 < |\eta| < 2.7$, the end-cap muon chambers are arranged in four wheels at $z \approx 7.5$ m, 10 m, 14 m and 22 m from the IP, all concentric with the beam-axis. The first is coined Small Wheel, whilst those located at a distance of $z > 10$ m from

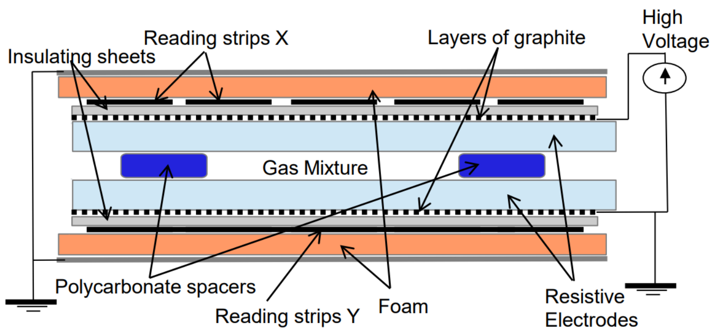

Figure 2.11: Illustration of a resistive plate chamber (RPC) in the muon spectrometer (MS). The ionising particle traverses through the plate, thereby ionising the gaseous volume within it. The strong electric field due to the high voltage on the plates causes avalanche multiplication of the primary ionisation initiated by the incident particle. [34].

the IP are referred to as Big Wheels. The second wheel, at $z \approx 10\,\text{m}$ consists of MDT chambers and provides lower coverage than those yielded by the Small and Big Wheels.

Similarly to the barrel, also at the end-caps MS MDTs are used as key precision measurement tools, installed on all four wheels. The MDTs are positioned azimuthally as to attain precision measurements in η. The caveat to this is in the $2 < |\eta| < 2.7$ region, the innermost muon station at the end-caps which is subject to the highest background rates, wherein muon precision measurements are obtained by the CSCs. Multi-wire proportional chambers, the CSC detectors contain cathode strips oriented perpendicularly to anode wires and are filled with a Ar-CO_2-CF_4 gas mixture. Traversing MIPs generate ionisation signals which pass through the strips and wires and are eventually read-out. The CSCs provide spatial resolution both in η and ϕ as a result of their high granularity design which can endure high background rates, as common in a very-forward detector region. Four-layered detectors, the CSCs reach a spatial resolution of $60\,\mu$m.

In the MS end-caps, the trigger chambers are made of TGC-type detectors. Also a multi-wire proportional chamber detector, the TGC is filled with n-pentane (n-C_5H_{12}) gas. In it, graphite cathodes and wires, separated by 1.4 mm, are held at a 2.9 kV potential difference. Like the RPCs in the barrel, the use of high voltage and particular design (the anode/cathod geometry) allows for a fast signal readout of 4 ns. The signals collected along the wires from the drift electrons as well as the strip charge are read-out as to grant a two-dimensional spatial information used both for track and trigger construction.

2.2.5 Trigger and Data Acquisition

During Run 2 of the LHC, ATLAS was delivered with pp collisions at instantaneous luminosity of the order of $10^{34}\,\text{cm}^{-2}\text{s}^{-1}$, at a 25 ns bunch spacing which resulted in a rate of about 10^9 interaction per second. However, neither the ATLAS detector records each one of these interactions, nor it is sought from a physics point of view. Indeed, a considerable portion of these events are the results of unwanted, soft collision processes which are unlikely to hint at processes of interest *e.g.* BSM physics and Higgs decays. ATLAS trigger system is an online filtering approach to pick candidate events with a higher potential of indicating at phenomena of interest, which are then delivered to permanent storage. Contrarily, an offline environment would mean that data are investigated after being saved into permanent storage.

In ATLAS, the trigger system is partitioned into two layers, namely a hardware-based low-level trigger coined Level-1 (L1) trigger and a software-based high-level trigger (HLT). L1 trigger employs calorimeter and MS measurements, upon which it performs the initial selection, thereby reducing the first event input rate of 40 MHz to 100 kHz. Electrons and photons induce triggers through energy deposits on the ECAL systems whereas hadronic jet candidates are delivered to the L1 trigger system by the HCAL. At L1, calorimeter-based trigger objects are constructed through the calculation of energy sums collected from calorimeter cells in η and ϕ. L1 triggers initiated by muons stem from hit coincidences along the MS.

Once candidate events are selected by the L1 trigger, they are delivered to the HLT, which consists of a Level-2 (L2) trigger system and an event filer (EF). The former performs preciser measurements of the initial objects which had been accepted by the L1 trigger. The EF is a software which employs the ATLAS Athena reconstruction framework as well as algorithms akin to such used in offline analysis environment in order to carry-out object reconstruction and identification. The HLT acceptance rate is of the order of 1 kHz and effectively constitutes the last stage before data are transferred to permanent storage.

Chapter 3

Hadronic Jets in ATLAS

By virtue of QCD confinement, partonic pp collisions final states do not exist as free particles for observable timescales and thus do not deposit meaningful signatures in the detector. Conversely, radiation of additional quarks and gluons collinear with the initial coloured particles is ubiquitous. This radiation is governed by formation of colour-neutral quark-binding combinations, a process otherwise known as hadronisation which eventuates in the object of hadronic jets [35]. These hadron sprays constitute macroscopically-detectable signatures of the quarks and gluons produced in pp collisions. In order to transform jets to analysable objects, two processes take place: reconstruction and calibration. The former is any method for clustering hadronisation end-products such that the physical properties of the initiating partons may be deduced from the formed object. Calibration relates to techniques of converting reconstructed jets into accurate representations of the initiating partons such that they can be employed relevantly within physics analyses. A schematic illustration of a typical jet creation process is given in Figure 3.1.

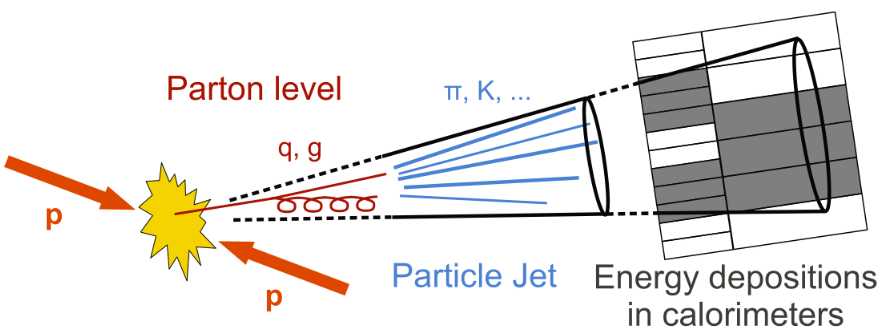

Figure 3.1: Depiction of a jet formation process from a pp collision. Stemming from the initial partons which eventually hadronise, in ATLAS jets are distinguishable by tracking detectors within the Inner Detector and calorimeters in which they deposit energy [36].

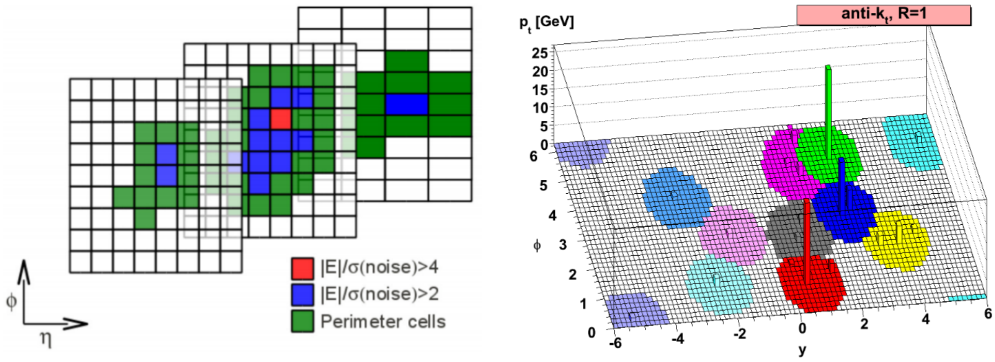

Figure 3.2: (Left) Topological cluster formation on grids representative of cells upon three hadronic calorimeter barrel layers [38]. (Right) Parton-level event where the anti-k_{T} algorithm is employed to cluster both hard and soft constituents into hadronic jets; energetic constitutents are marked with different colours. The filled and coloured circle-like shapes indicate jet catchment areas, where constituents are aggregated into jets and sizes of which are determined by the R parameter inputted into the algorithm [39].

3.1 Jet Reconstruction

3.1.1 Topological Clustering

Reconstruction of hadronic jets initiates with clustering of low-level signals read-out through energy depositions in cells connected to channels in the LAr and tile calorimeter systems. In ATLAS, this clustering is achieved via a topological clustering algorithm [37]. The high-granularity calorimeter detectors permits the collected topological clusters (often abbreviated as topo-clusters) to include detailed information relating to the energy flow of jets. The evolution of topo clusters originates from primary seed cells, where a high signal-to-noise ratio of $|S/N| > 4$ is mandated. Here, $S = |E|$, the magnitude of calorimeter-cell energy measurement, and N is the quadrature sum of the RMS electronic and pileup noise. In three dimensions, seed-neighbouring cells are defined as either having bordering calorimeter cells with the seed cells on the same calorimeter layer or having partial overlap in the η-ϕ plane on adjacent calorimeter layers. With more-lenient requirements, the seed-neighbouring cells are subject to $|S/N| > 2$ and also contained within the topo-cluster. Lastly, the cells surrounding the cells which had already been collected in each layer, otherwise known as perimeter cells or basic threshold cells, are also joint whilst merely having $|S/N| \geq 0$. A topological clustering process is illustrated in FIG. 3.2.

3.1.2 Jet Finding

In order to choose which inputs are amassed into single jets, a jet-finding algorithm is required. Typical in LHC experimets is a sequential-recombination algorithm known as anti-k_{T} [39]. This algorithm is infrared and collinear (IRC) safe and produces circular jets in the $\eta - \phi$ plane. Anti-k_{T} computes the following quantities for all jet constituents:

$$
\begin{aligned}
d_{ij} &= \min\Big(\frac{1}{k_{\mathrm{T}i}^2}, \frac{1}{k_{\mathrm{T}j}^2}\Big)\frac{\Delta R_{ij}^2}{R^2}, \\
d_{iB} &= \frac{1}{k_{\mathrm{T}i}^2}
\end{aligned}
\tag{3.1}
$$

where $k_{\mathrm{T}i}$ is the transverse momentum of constituent i, R the chosen jet radius parameter and $\Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$. d_{ij} is the generalised distance used by the algorithm between constituents i and j whilst d_{iB} the distance between constituent i and the beam, used in order to distinguish between constituents originating in the hard-scatter process and those coming from proton remnants.

The anti-k_{T} clustering algorithm thereafter calculates all distances between constituents i and j and identifies the smallest of them, as to have the most energetic particles clustered first. If that distance yields $d_{ij} < d_{iB}$, the algorithm recombines constituents i and j. Otherwise, if the smallest distance is d_{iB}, it refers to i as a jet and subsequently omits it from the list of constituents. After recombination, the distances are recalculated for remaining constituents repeatedly until none are left. This process may be understood via an illustration of clustering of hard and soft particles into jets via anti-k_{T}, as shown in FIG. 3.2 (Right). The main analysis object used in the study reported in this thesis is hadronic jets reconstructed via the anti-k_{T} algorithm with $R = 1.0$.

3.2 Jet Calibration

Referred to at this stage as EM-scale jet, jets reconstructed as outlined in Section 3.1 are clustered at the electromagnetic (EM) scale and thus ascribed correct energy measurements relating to EM showers. These are calibrated to the hadronic scale as described below. Attributing energy and momentum measurements which correspond to energies and momenta of the underlying particle-level jets to the reconstructed jets, a step known as jet energy scale (JES) calibration, is called for.

In ATLAS topo cluster are calibrated to the hadronic scale with the local-cluster weighting (LCW) method. Firstly, the LCW calibration determines whether topo-clusters are EM or hadronic, based on their measured energy densities and longitudinal shower depths. According to this classification, it then applies energy corrections to each topo-cluster using Monte Carto (MC) simulations of hadrons which typically participate in jet formation, like charged and neutral pions. Moreover, this calibration introduces additional corrections such as for energy loss in uninstrumented regions of the detector. Jets calibrated using LCW are referred to as LCW jets, whilst those built from non-calibrated clusters are usually called EM jets.

Once jets are reconstructed, their calibration process [40] consists of several stages, as in the flowchart given in FIG. 3.3. The salient steps will be described in the following subsections.

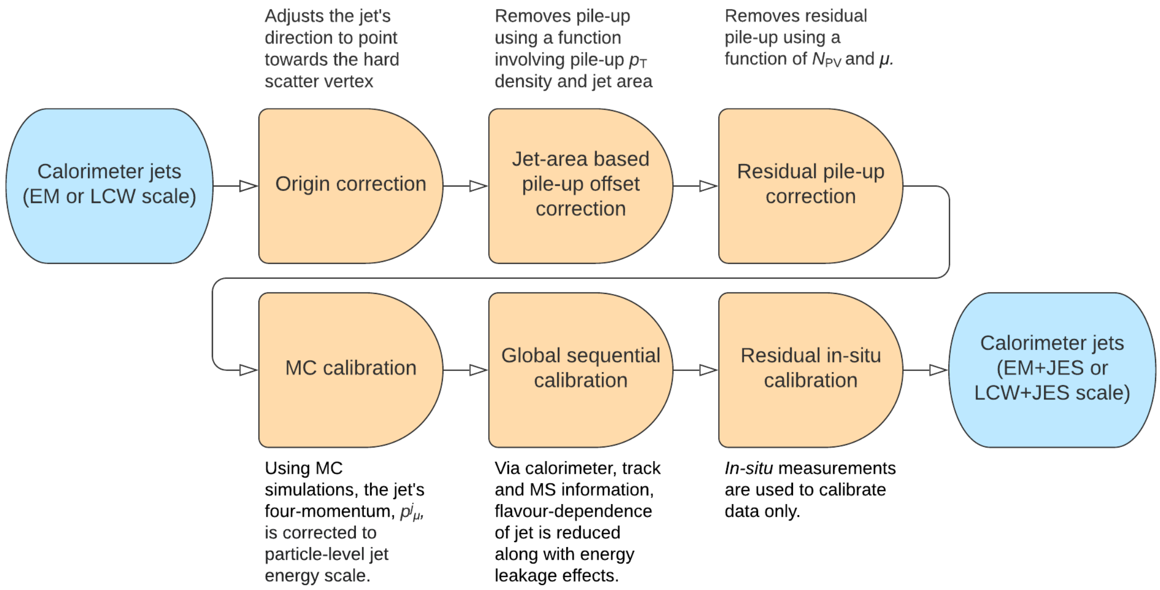

Figure 3.3: Flowchart outlining the steps undertaken in jet calibration.

3.2.1 Origin Correction

The calorimeter jets, produced with the premise they stemmed from ATLAS' geometric centre, undergo an adjustment such that their direction is changed as to point towards the primary hard-process vertex (PV) from which the initiating partons originate. This stage constitutes a recalculation of the jet four-momentum which also improves its η resolution. Herein, the jet energy remains unchanged whilst its p_T response improves.

3.2.2 Pile-up corrections

Presence of pile-up events may offset the measured energy of a given jet. An initial correction is calculated for the transverse momentum for a given event,

$$p_T^{\text{corr}} = p_T^{\text{reco}} - \rho \cdot A - \alpha \cdot (N_{PV} - 1) - \beta \cdot \langle \mu \rangle, \tag{3.2}$$

where ρ is the pile-up energy density defined as the median of the p_T/A distribution, A the jet area calculated via ghost-matching (scanning a single infinitely-soft particle over its rapidity-azimuth cylinder and determining the region over which it is clustered within the jet) [41], α the coefficient of the residual correction due to N_{PV} and β the coefficient of the residual correction due to out-of-time pileup [38]. The second term in the RHS of Eqs. (3.2) represents area-based corrections and the third term accounts for residual corrections. Dependence upon residual pile-up is typically greater in the forward region of the detector (high $|\eta|$) where pile-up and background activities are largest.

3.2.3 Jet energy scale and η calibration

EM/LCW-scale reconstructed ("reco") jets are corrected to the true energy scale based upon particle-level rather than detector-level jets. Therefore, this calibration solely relies on MC simulation of particle-level jets, which are clustered via the anti-k_T algorithm using particles generated post-hadronisation as input constituents, thereby deferring from reco jets (the former and latter are reconstructed before and after detector interaction, respectively) *cf.* FIG. 3.1. An important feature of this calibration is that it compensates for overlooking inactive material in the detector where radiation not accounted for in the reco calorimeter-based jets is dissipated. Furthermore, this step corrects jets for out-of-cone radiation disregarded by the clustering algorithm, as well as for effects related to varying detector geometry and different technologies used within it.

The calibration parameters are derived from simulated reco jets to particle level jets via the average energy response, $E^{\text{reco}}/E^{\text{truth}}$, where E^{truth} is the particle-level jet energy. Moreover, η calibration refers to areas where reco jets are biased due to their surrounding by two calorimeter regions or technologies. The η difference between the reco and particle-level ("truth") jets is computed and used to further calibrate the reco jet. When this stage is finalised, jets which were initially referred to as reco jets are said to be of (EM/LCW)+JES scale.

3.2.4 Global Sequential Calibration

Jets undergoing Global Sequential Calibration (GSC) are already (EM/LCW)+JES-scaled. Here they are inspected for fluctuations in their flavour composition, dependence on shower shapes and inter-jet energy distribution. One noteworthy example of fluctuations fixed at this stage relates to the initiating particles of the jets, *e.g.* quarks or gluons. A quark-initiated jet would usually be highly collimated consisting of high-p_T hadronic constituents. On the other hand, a gluon-initiated jet would be relatively less collimated and contain many more low-p_T constituents. A common outcome of these distinctnesses is that jets of the latter type do not traject as far into the calorimeter as jets of the former type do.

Apart from employing similar responses as described in §3.2.3, GSC uses observables related to the jet shower profile, growth through the calorimeter and number of associated tracks. Further, information from the MS regarding associated tracks is instrumental in improving the energy response of jets which are not fully-contained in the calorimeter but are accounted for in the MS, often referred to as punch-through jets. These jets traverse through the HCAL, induce a hadronic shower and thereby produce high-momentum particles which dissipate detectable marks in the muon chambers. Often, such particle are misattributed to muon tracks unrelated to punch-through jets and thus impede the correct computation of the four-momentum of such jets. Such cases are often dealt with via GSC which ingeniously use various information from different sections of the detector, combined with novel algorithms to properly assign the correct energy to punch-through jets.

3.2.5 *In-situ* Corrections

The last step in hadronic jet calibration is assessing data-to-MC discrepancies via *in-situ* techniques which uniquely use data rather than MC, contrarily to all other calibrations stages. This method inspects the p_{T} balance between a jet and well-measured photons, Z bosons or jets and only applies to data in the aim of restoring the energy of reco jets to that expected from MC simulations. Initially, a substage named η-intercalibration is performed. The average response of forward jets $(0.8 < |\eta| < 4.5)$ to that of central jets is corrected using dijet events wherein the two jets recoil against each other $(|\phi| \sim \pi)$. Secondly, the (EM/LCW)+JES-scale jet is balanced against a well-measured object like the ones mentioned above. This substage uses only very central jets $(|\eta| < 0.8)$ whereas the particular reference object yields the p_{T} scale to which the response correction applies. Typically, response adjustments for jets of $p_{\mathrm{T}} \leq 950\,\mathrm{GeV}$ rely on reference photons and Z bosons undergoing leptonic decays, $Z \to ee$ and $Z \to \mu\mu$. Higher-p_{T} jets are corrected via lower-p_{T} multijet events as reference objects.

3.3 Jet Substructure Observables

Categorisation of jets into different kinds requires the usage of a set of variables related to their substructure. Typically, they are aimed to separate quark from gluon-initiated jets or identify $W/Z/H/t$ signal jets from ubiquitous QCD background jets emanating from "standard" quarks and gluons [35]. These variables are associated with the jet's kinematic and topological substructure as well as with its relation to subjets of its composition. Below several substructure observables are introduced.

Energy correlation functions (ECFs) are useful jet observables. Although they can be generalised to any n-point function, the 2-point and 3-point functions, $e_2^{(\beta)}$ and $e_3^{(\beta)}$ respectively, are not only the typical ones computed but also those whose ratio is regularly used. They are defined as

$$\begin{aligned} e_2^{(\beta)} &= \frac{1}{p_{\mathrm{T}\,J}^2} \sum_{1 \leq i < j \leq n_J} p_{\mathrm{T}\,i} p_{\mathrm{T}\,j} \Delta R_{ij}^{\beta}, \\ e_3^{(\beta)} &= \frac{1}{p_{\mathrm{T}\,J}^3} \sum_{1 \leq i < j \leq n_J} p_{\mathrm{T}\,i} p_{\mathrm{T}\,j} p_{\mathrm{T}\,k} \Delta R_{ij}^{\beta} \Delta R_{ik}^{\beta} \Delta R_{jk}^{\beta}, \end{aligned} \tag{3.3}$$

where $p_{\mathrm{T}J}$ is the transverse momentum of the jet, $p_{\mathrm{T}i}$ the transverse momentum of particle i, n_J the number of particles in the jet and the boost-invariant angle, $\Delta R_{ij}^2 = (\phi_i - \phi_j)^2 + (\eta_i - \eta_j)^2$, the Euclidean distance in the azimuth-rapidity plane. The exponent $\beta > 0$ is determined on a par with avoiding the presence of soft particles inside the jet, otherwise known as maintaining infrared safety [1].

This sort of ECFs is handy for algorithms known as prong-taggers, purposed for the construction of hard subjets inside a large-R jet and comparing their properties. Ratios of ECFs are used to separate boosted massive particles from background QCD

[1] β is a free parameter but usually $\beta = 1$ is implicitly assumed [35]

jets and indeed constitute effective two-prong taggers. A couple of these ratios are often employed, namely

$$C_2^{(\beta)} = \frac{e_3^{(\beta)}}{\left(e_2^{(\beta)}\right)^2}, \qquad\qquad D_2^{(\beta)} = \frac{e_3^{(\beta)}}{\left(e_2^{(\beta)}\right)^3}, \tag{3.4}$$

where D_2 bears a stronger discriminating power at the expense of being more sensitive to model-dependent soft contamination in the jet stemming from *e.g.* pileup or the underlying event (UE). UE is the diffuse radiation from disintegrating remnants of the colliding hadrons.

Another helpful observable is a jet-shape property coined N-subjettiness, $\tau_N^{(\beta)}$, used to classify jets on a par with the number of N subjets they are composed of. To that end, a set of axes $a_1, \ldots, a_N$ is defined such that the subjettiness is given by

$$\tau_N^{(\beta)} = \frac{1}{d_0} \sum_{i\,\in\,\text{jet}} p_{\text{T}i} \min(\Delta R_{ia_1}^{\beta}, \ldots, R_{ia_N}^{\beta}), \tag{3.5}$$

where a_i runs over constituent particles in a given jet. The normalisation factor, d_0, is taken as

$$d_0 = \sum_k p_{\text{T}k} R_0, \tag{3.6}$$

where R_0 is the characteristic jet radius used in the original jet clustering algorithm. τ_N may be regarded as the degree to which a jet is N-subjetty *i.e.* to what extent it can be treated as a jet comprised of N subjets. Jets of $\tau_N \approx 0$ are such where all of their radiation is aligned with the candidate subjet directions and hence contain no more than N subjets. Conversely, jets with $\tau_N \gg 0$ have a large fraction of their energy dispersed afar from their candidate subjets and as such associate with no less than $N + 1$ subjets. Plainly, τ_N does not provide a great discriminating power for boosted objects but may be very helpful when used in conjunction with complementary substructure variables [42].

Another standard variable relates to the constituent hadrons in 2-pronged events otherwise known as Fox–Wolfram moment, defined as

$$R_2^{\text{FW}} = \sum_{i,j} \frac{|\mathbf{p}_i||\mathbf{p}_j|}{E_{\text{Event}}^2} P_\ell(x) \cos\theta_{ij}, \tag{3.7}$$

where θ is the opening angle between hadrons i and j, E_{Event} the total detected event energy and $P_\ell(x)$ the Legendre polynomials. A second, called planar flow [43] and labelled $\mathcal{P}$, measures the extent to which the energy flow inside the jet is linear or planar. A shape tensor is given by

$$I_w^{kl} = \frac{1}{m_J} \sum_i \frac{1}{E_i}{}_{\text{T}i,k} p_{\text{T}i,l}, \tag{3.8}$$

where m_J is the jet mass, E_i the energy of the i^{th} jet-constituent particle and $p_{\text{T}i,k}$ the k^{th} component of the i^{th} constituent transverse momentum relative to the jet axis. Thereby the planar flow is defined as

$$\mathcal{P} = \frac{4\det(I_E)}{[\text{tr}(I_E)]^2} = \frac{4\lambda_1\lambda_2}{(\lambda_1+\lambda_2)^2}, \tag{3.9}$$

where $\lambda_{1,2}$ are the eigenvalues of I_w. For linear shapes $\mathcal{P} \to 0$ whereas for isotropic energy depositions $\mathcal{P} \to 1$. This shape-analysis substructure variable is usually complemented with jet mass cuts which empower its discriminating power.

Sensitivity to the symmetry of the energy flow inside a jet is gained through a class of observables called angularities. A general formula is formulated,

$$a_n = \frac{1}{M}\sum_i E_i \sin^n\theta_i[1-\cos\theta_i]^{1-n}, \tag{3.10}$$

such that a is a parameter indicating whether the radiation is near the edges ($a < 0$) or core ($a > 0$) of the jet and θ_i the i^{th} jet constituent's angle with respect to the jet axis [44]. All other parameters have their usual meaning. In the small-angle radiation limit ($\theta_i \ll 1$), the angularity is approximated by

$$a_n \simeq \frac{2^{(n-1)}}{M}\sum_i E_i\theta_i^{(2-n)}, \tag{3.11}$$

where the angularity a_3 is often used in two-pronged jet analyses. In a two-body phase-space the angularity should peak about $a_n^{\text{peak}} \simeq (\frac{M}{2p_{\text{T}}})^{1-n}$ which corresponds to the two hard jet constituents being in a symmetrical p_{T} configuration around the jet axis. Moreover, an approximation for the distribution's maximum in the small-angle limit yields $a^{\max_n} \simeq (\frac{2}{R})^n(\frac{M}{2p_{\text{T}}})$ which is on a par with a hard constituent close to the jet axis and a soft constituent near the jet edge.

An additional jet-shape substructure observable is aplanarity. Like other shape variables, it is calculated from energy clusters deposited in the jet at its CoM frame. One defines the sphericity tensor,

$$S^{\alpha\beta} = \frac{\sum_i p_i^{\alpha_i} p_i^{\beta_j}}{\sum_i |\mathbf{p}_i|^2}, \tag{3.12}$$

where $\alpha, \beta \in \{x, y, z\}$ are momentum components of each energy cluster in the jet. Diagonalisation of $S^{\alpha\beta}$ yields eigenvalues such that $\lambda_1 \geq \lambda_2 \geq \lambda_3$ where $\lambda_1+\lambda_2+\lambda_3 = 1$. By this a jet's aplanarity is given by

$$A = \frac{3\lambda_3}{2}, \tag{3.13}$$

and constrained within $0 \leq A \leq 1/2$. A strongly directional jet distribution yields $A \to 0$ whereas an isotropic distribution corresponds to $A \to 1/2$.

One other substructure variables category relates to splitting. Generally, these are observables relating between kinematic properties of a jet's daughter partons, say A and B, to the final jet collection through calorimetry. One splitting measure

Variable	Type
C_2, D_2	Energy correlation ratios
τ_{21}	N-subjettiness
R_2^{FW}	Fox–Wolfram moment
$\mathcal{P}$	Planar flow
a_3	Angularity
A	Aplanarity
Z_{cut}, $\sqrt{d_{12}}$	Splitting scales
KtDR	k_{T}-subjet ΔR

Table 3.1: Summary of jet substructure observables discussed in the text.

is commonly employed by the k_{T} clustering algorithm (a discussion of a similar algorithm is given in §3.1.2) and defined as

$$d_{AB} = \min(p_{\mathrm{T}A}^2, p_{\mathrm{T}B}^2)\frac{\Delta R_{AB}^2}{R^2}, \tag{3.14}$$

which can be optimised to account for the exact collinear limit and then labelled d_{cut}. In the clustering sequence, the associated jet observable is given as $\sqrt{d_{12}}$. This is then used to define another splitting variable called energy sharing generally formalised as

$$z_M = \frac{\min(E_A, E_B)}{E_M}, \tag{3.15}$$

where M is the mother parton in the process $M \to AB$ [45]. This variable can also be taken to the collinear limit via d_{cut},

$$z_{\mathrm{cut}} = \frac{d_{\mathrm{cut}}}{d_{\mathrm{cut}} + Q_M^2} \to \frac{\min(E_A, E_B)}{E_A + E_B}, \tag{3.16}$$

which is often useful for highly-boosted events where A and B are collinear in order to account for soft singularities. Together, these two splitting variables measure both the soft ($z \to 0$) and collinear ($Q^2 \to 0 \Rightarrow z \to 1$) singularities.

The last observable is referred to as k_{T}-subjet ΔR and labelled KtDR. It is effectively the angular distance ΔR_{ij} between objects i and j corresponding to the final clustering step within k_{T} algorithm families. As expected, the higher the boost, the narrower the distance between i and j *i.e.* smaller KtDR is associated with more collimated jet-constituent quarks. Therefore, KtDR possesses a large discriminating power against a low-p_{T} background where the distance between constituents is relatively huge. A summary of the jet substructure variables described in this section is given in Table 3.1.

Chapter 4

Estimation of the Sensitivity to Vector Bosons through an ISR-$\gamma +$ large-R Jet Channel

Numerous models have proposed [46–48] new particles with considerable couplings to partons, as well as BSM resonances with additional couplings to DM particles. Searches for resonant excesses in the invariant mass distributions of monojet and multijet events are an indispensable part of the LHC physics programme. Within such searches, although attaining high integrated luminosities, LHC experiments have been limited at lower masses due to a large multijet background. These events are produced at rates beyond the bandwidth limitations of LHC experiments thresholds which therefore have had a poor sensitivity to masses below 1 TeV in searches for pure hadronic resonances. Hence, minimum transverse momentum cuts are set on triggers using jets to avoid a flooding of the event collection.

These challenges are tackled through several routes by ATLAS and CMS where novel techniques to overcome the lower-bound limitations due to trigger requirements in searches for new low-mass resonances have been introduced [49, 50]. A key technique uses a high-p_{T} initial-state radiation (ISR) object, a jet or a photon, against which the new particle recoils. The ISR may also be used for triggering [51]. In such an event topology, a light resonance is highly boosted when recoiling against the ISR object and thereafter may result in a single large-radius (R) jet. Another method called trigger-level analysis averts trigger bandwidth restrictions by recording minimal event information, albeit at a substantially high rate.

A typical benchmark [46] of the new light resonance is a leptophobic boson, Z', with axial-vector couplings to SM quarks and to a Dirac fermion DM candidate. Interactions between the spin-1 mediator Z' and SM and DM fields are modelled through Lagrangians for the vector and axial-vector couplings. Respectively,

$$
\begin{aligned}
\mathcal{L}_V &= \mathrm{g}_{\mathrm{SM}} \sum_{q\in\{u,d,s,c,b,t\}} Z'_\mu \bar{q}\gamma^\mu q + \mathrm{g}_{\mathrm{DM}} Z'_\mu \bar{\chi}\gamma^\mu \chi \\
\mathcal{L}_A &= \mathrm{g}_{\mathrm{SM}} \sum_{q\in\{u,d,s,c,b,t\}} Z'_\mu \bar{q}\gamma^\mu \gamma^5 q + \mathrm{g}_{\mathrm{DM}} Z'_\mu \bar{\chi}\gamma^\mu \gamma^5 \chi,
\end{aligned}
\tag{4.1}
$$

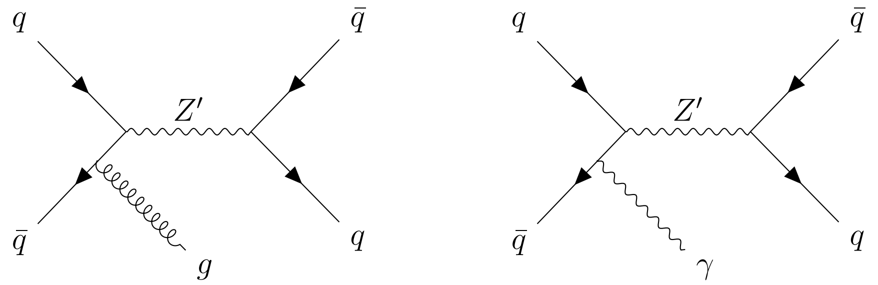

Figure 4.1: Feynman diagrams exemplifying a generic new resonance Z' produced in association with an ISR object, a jet (left) or a photon (right), and decaying to a diquark final state, which may be reconstructed as a single large-R jet or two small-R jets.

where χ is the DM field. This model contains several parameters:

$$\{g_{SM}, g_{DM}, m_\chi, m_{Z'}\}, \tag{4.2}$$

where g_{SM} is the flavour-universal coupling of Z' to SM quarks, g_{DM} the coupling of Z' to the DM fermion of mass m_χ and $m_{Z'}$ the Z' mass. Searches of this kind employ SM signatures which are identifiable by LHC experiments and hence focus on the g_{SM} and $m_{Z'}$ parameters. The mass of the DM fermion candidate is large such that $m_\chi >> m_{Z'}$, thereby constraining the decay width of Z' to DM to zero. The signal choice relates to the scope searchable through an ISR+jet/s combination, which thus limits the interactions between DM and SM. The Z' stems from minimal SM extensions, namely an additional $U(1)'$ gauge symmetry and a Dirac fermion DM particle which is only charged under this new group. Given that SM quarks are also charged under $U(1)'$, the BSM Z' can mediate interactions between SM and DM particles.

In case the new resonance were of a higher mass, too heavy to produce a strongly-boosted decay, its final state would be collected as two small-R jets in addition to the ISR object. When the light resonance is transversely boosted following a recoil against a high-p_T ISR photon [52,53] it is collected by a large-R jet. Despite resulting in a low signal acceptance, requiring a high-p_T photon allows a collection of signal events with low resonant masses. Feynman diagrams illustrating processes which include a production of a generic new vector boson resonance, Z', in association with an ISR object are given in FIG. 4.1.

The Z' resonances is explored via several channels. One is the γ + dijet, where the trigger object is the ISR photon and the final-state quarks are collected as two distinguishable small-R jets. Parallel to this is the trijet channel, where the ISR is a jet and the final state is also composed of two small-R jets. All channels where the final state is reconstructed from a pair of nearby anti-k_T jets are categorised as resolved. Another two channels where the final-state is reconstructed as a single large-R jet and the ISR is either a photon or a small-R jet are collectively known as boosted. A highly boosted mediator causes a significant overlap between the jet cones of the hadronic final-state which are eventually reconstructed as a single fat jet.

The analysis presented herein belongs to the boosted ISR-γ+large-R jet channel. It uses the ISR photon as trigger and collects events which include a single photon in association with a fat jet. The dominant background stems from QCD events where light quarks and gluons result in a single fat jet in association with a photon radiation. Events where SM vector bosons decay to a quark pair in association with a photon constitute an additional background to the Z' study but used as signal in the context of a control study. The events are filtered through an optimised event selection which is validated using truth-level labelling. An array of cuts uses the ISR-jet event topology to reconstruct the signal where also jet substructure observables are assessed. The sensitivity to vector resonances is estimated from the invariant mass distribution of the final-state jet.

The study begins with the validation of the MC simulation against data in terms of tally, kinematic and angular distributions of the ISR photon and final-state jet. Following the high-p_{T} single-photon trigger verification, the event selection employs kinematic, topological cuts in conjunction with truth-level labelling to optimise the MC simulation and estimate the sensitivity to massive SM vector resonances. By this it is ensured that the events taken are the ones where the ISR γ recoils against the large-R jet, where the latter contains a V boson and matches both the truth jet and particle. Thereafter, the performance of a machine learning (ML) tagging tool in background discrimination is evaluated. Finally, the sensitivity of the whole program, coined a tagger of V bosons through an ISR-$\gamma + J$ channel, is evaluated.

4.1 Data Sets

4.1.1 Collision Data

The analysis described in this thesis uses data from pp collisions at a centre-of-mass energy of $\sqrt{s} = 13\,\mathrm{TeV}$ collected by the ATLAS detector. The sample includes data from all periods throughout 2015–2016 with stable LHC beam conditions, when all relevant subdetectors were operational. The total integrated luminosity at the time was $\mathcal{L}^{\mathrm{int}} = 36.0\,\mathrm{fb}^{-1}$ with mean interactions per bunch crossing being $\langle\mu\rangle = 13.4$ and $\langle\mu\rangle = 25.1$ for 2015 and 2016, respectively [26]. Data is used in the context of MC validation and control study. This analysis is a part of an ongoing search and therefore only a subset of the available LHC data is used for those purposes.

4.1.2 Monte Carlo Simulations

QCD Background

The dominant background of this search consists of boosted events where gluons and light quarks decayed to a single large-R jet after recoiling against a single photon. This background is referred to as QCD background. Its simulation is based upon events [1] generated via Sherpa v2.1.1 software [54]. Using built-in tree-level matrix element generators, Sherpa simulates parton-level events where decays of heavy resonances occur. These include automated methods for phase-space integrations as well as subtraction of IR divergences in NLO calculations. Furthermore, Sherpa merges multijet production processes at NLO based on truncated parton showers

[1] Samples from the MC16 production campaign are used

Quark-jet selection	p_T^γ range [GeV]	σ	ϵ_{filter}	$\sigma_{\text{Eff.}}$	N_{Events}
c-veto, b-veto		3.129	0.399	1.802	41650
c-filter, b-veto	70–140	3.131	0.482	2.174	57301
b-filter		3.132	0.117	0.529	35149
c-veto, b-veto		0.247	0.392	0.139	7672782
c-filter, b-veto	140–280	0.247	0.477	0.170	4716954
b-filter		0.249	0.129	0.046	2356725
c-veto, b-veto		0.014	0.386	0.007	3971043
c-filter, b-veto	280–500	0.014	0.474	0.009	1488446
b-filter		0.014	0.140	0.003	297818
c-veto, b-veto		9×10^{-4}	0.378	5×10^{-4}	799983
c-filter, b-veto	500–1000	9×10^{-4}	0.471	6×10^{-4}	598989
b-filter		9×10^{-4}	0.149	2×10^{-4}	120000
c-veto, b-veto		2×10^{-5}	0.370	1×10^{-5}	500000
c-filter, b-veto	1000–2000	2×10^{-5}	0.471	1×10^{-5}	250000
b-filter		2×10^{-5}	0.161	4×10^{-6}	70000
c-veto, b-veto		8×10^{-8}	0.368	4×10^{-8}	100000
c-filter, b-veto	2000–4000	8×10^{-8}	0.466	5×10^{-8}	150000
b-filter		8×10^{-8}	0.168	2×10^{-8}	10000
c-veto, b-veto		2×10^{-12}	0.422	1×10^{-12}	1000
c-filter, b-veto	4000 <	2×10^{-12}	0.425	1×10^{-12}	1000
b-filter		2×10^{-12}	0.152	5×10^{-13}	1000

Table 4.1: Samples used for background simulation, computed to next-to-leading order via Sherpa v2.1.1. Each sample of the three included per p_T^γ slice is calculated with a different c/b-quark jet selection scheme. The effective cross-section, $\sigma_{\text{Eff.}}$, is computed as $\sigma_{\text{Eff.}} = \sigma \times \epsilon_{\text{filter}} \times k$-factor, where k-factor $= 1.44$ is common to all samples.

and simulates parton hadronisation via a cluster fragmentation model. Each event includes at least a single photon, used also as the HLT object, as well as a fat jet ($R = 1.0$). The hard-scatter portions of the events are simulated via CT10 PDFs [55]. The pile-up conditions are on a par with those of the LHC during 2015–2016 [26]. Full samples are partitioned into p_T slices according to the transverse momentum of the ISR photon and begin from 70 GeV. In order to accurately model different parts of the phase space, each p_T slice is generated three-fold: once without c and b jets ("c-veto, b-veto"), once with selections on c jets and without b jets ("c-filter, b-veto") and once with selections on b jets only ("b-filter"). All events are fully simulated for detector response via GEANT 4 (see description later in this section). In this stage of the analysis, background estimation is carried out via MC simulations which are sufficient in describing the data in the context of a preliminary sensitivity estimation. As the search progresses, more advanced, data-driven techniques will be used to estimate the background, in a similar fashion to the work conducted in the previous search. Details of the MC samples used are summarised in Table 4.1.

Feynman diagrams exemplifying QCD processes involving a photon radiation in association with a diquark state which may be collected as a single large-R jet are given to LO in FIG. 4.2.

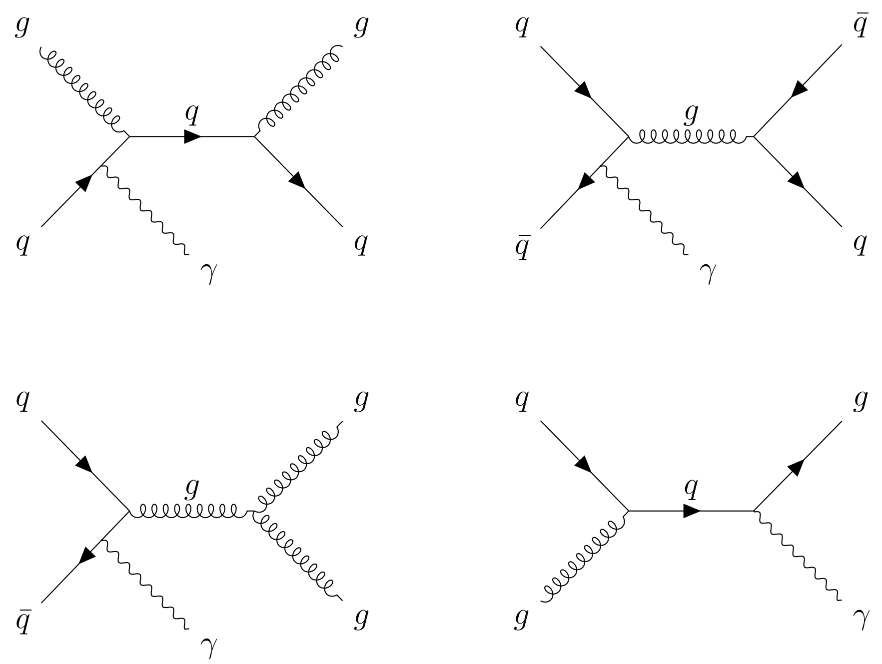

Figure 4.2: Feynman diagrams exemplifying generic quantum chromodynamics background which consists of a photon radiation and a fat jet at leading order. Many more such diagrams are possible.

SM $Vqq + \gamma$

The overall background also includes SM processes where a boosted V boson is produced in association with a photon and decays to a large-R jet. Although included in the background, such events [2] are also used as signal against the dominant QCD background in the context of a control study. Simulation of events where a boosted massive SM vector boson, $W^{\pm}$ or Z^0, is produced in association with a photon and decays to a collimated diquark final state is computed to LO via Sherpa v2.1.1. Like the QCD background samples, these are also partitioned into p_{T} slices according to the transverse momentum of the photon which is also used for HLT, starting from 140 GeV. The hard-scatter sections are modelled via the CT10 PDFs. A summary of the samples used to model the SM $Vqq + \gamma$ events is given in Table 4.2.

Truth Levels *vs.* Reconstruction Level

The simulation consists of three layers, otherwise known as fiducial volumes, each models a different level where interactions can be analysed. From the bottom up, the most fundamental level is the (truth) parton level at which a process is estimated upon the resonance decay to its final state. Only hard-scatter considerations are included as to model the event up to the final-state partons. It is based upon the final-state quarks to which the V boson decayed and the analysis object it provides is named truth particle. The middle level is the (truth) particle level, extending from final-state partons to stable hadrons where hadronisation processes

[2]Samples from the MC16 production campaign are used

Simulated process	$Wqq + \gamma$		$Zqq + \gamma$	
p_{T}^{γ} range [GeV]	$\sigma_{\mathrm{Eff.}}$ [pb]	N_{Events}	$\sigma_{\mathrm{Eff.}}$ [pb]	N_{Events}
140–280	1.7	174140	0.7	515351
280–500	0.2	97740	0.07	199242
500–1000	0.04	100000	0.01	299994
1000–2000	2×10^{-3}	80000	4×10^{-4}	80000
2000 <	2×10^{-5}	10000	4×10^{-6}	10000

Table 4.2: Samples used for signal $Vqq + \gamma$ simulation produced via Sherpa 2.1.1. These samples are calculated to leading order throughout and hence the k-factor is unity across the lot of them. Since also the filter efficiency for this production equals one, the cross-sections and effective cross-sections are equal.

are modelled. Thereby, this level consists of an interaction simulation covering truth particle production until its collection as truth jets, which are the analysis objects provided following quark and gluon fragmentation. Detector-stable hadrons can be used as inputs from which truth jets are produced.

Detector level is the basis for reconstructed hadronic jets to be used as analysis objects. These jets pertain to experimental features in addition to perturbative computations, particle shower, hadronisation and fragmentation phenomenology. Naturally it is the only level comparable to data collected by the ATLAS detector. The program which the detailed detector response is simulated by is the GEometry ANd Tracking (GEANT) 4 platform [56]. Shortly, it is a toolkit designed to model the passage of particles through matter, adaptable to any geometry, apparatus and media, amongst which are ATLAS subdetectors. The physics processes offered in GEANT 4 include electromagnetic, hadronic and optical interactions. These match the tracker and calorimeter subsystems utilised in order to reconstruct photons and hadronic jets in ATLAS. Event passing those filters can be validated via truth-level labelling. Once the reconstructed object is chosen as to optimally reconstruct the decaying resonance probed by truth particles and jets, event selection is concluded and inferred to the data as well. The different fiducial volumes and their partitioning into analysis levels are shown schematically in FIG 4.3.

Therefore, there are three levels of information produced within the MC simulation, namely detector level, particle level and parton level. The objects pertaining to detector level are the photons and LC topo jets, reconstructed and calibrated as discussed in Sections 3.1–3.2. Objects used at particle level are truth W/Z fat jets, inclusive of particle shower and hadronisation simulation but not of detector response simulation. Lastly, at parton level the most fundamental objects simulated are truth photons and truth W/Z bosons. Since the MC compared to data is at detector level, the particle and parton levels serve as aids to improve the purity and focus of the simulated reconstructed objects to better model the sought resonances.

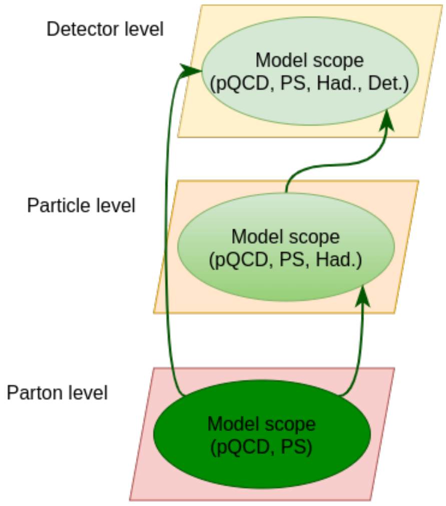

Figure 4.3: Fiducial volumes at different information levels. Below the top level, the detector level at which objects are reconstructed, the particle and parton truth levels are presented. Commonly, where theoretical predictions at the fundamental interaction level are corrected through the particle level to the detector level an analysis process known as folding is taking place, illustrated by the green lines up-streaming on the LHS. The parton level models purturbative quantum chromodynamics and parton shower, the particle level additionally simulates hadronisation and the detector level adds the detector response simulation on top of the latter three, using the GEANT 4 platform in conjunction with the ATLAS detector geometry.

BSM $Z'qq + \gamma$ Signal

The benchmark signal simulation is produced [3] via a combination of generators namely MadGraph5 v2.3.3 [57], Pythia8 v8.210 [58] and EvtGen v1.2.0 [59]. Using the NNPDF23LO PDF with A14 tune [60], MadGraph5 simulates the hard scatter to LO and Pythia8 models the particle shower, hadronisation and fragmentation processes which result in the final-state fat jets. The hadronic final state is inclusive in quark flavour where EvtGen complements the simulation with final states composed of b and c jets. To LO, it models processes where a BSM resonance Z' which has a universal axial-vector coupling to quarks with a constant $g_q = 0.5$ is produced in association with a photon, used as the trigger object. The new boson is boosted and recoils against the ISR photon as to decay into a final state composed of two collimated quarks. The resonance is modelled with a set of masses, $m_{Z'} \in [100, 220]\,\text{GeV}$. MC events simulating different $m_{Z'}$ values have unique ISR filters, where the minimum photon transverse momentum is 50% greater than the respective resonance mass and thus increases linearly with it. Details of the samples used for the Z' analysis are summarised in Table 4.3. The effective cross section, $\sigma_{\text{Eff.}}(qq + \gamma \to Z' \to J)$, is plotted as a function of $m_{Z'}$ as can be seen in FIG. 4.4.

[3]Samples from the MC16 production campaign are used

$m_{Z'}$ [GeV]	σ [pb]	ϵ_{filter}	$\sigma_{\text{Eff.}}$ [pb]	$p^{\gamma}_{\text{T, filter}}$ [GeV]	N_{Events}
100	6.74	0.291	1.97	150	50000
130	5.86	0.138	0.81	195	50000
160	5.04	0.076	0.38	240	50000
190	4.32	0.046	0.20	285	50000
220	3.70	0.030	0.11	330	50000

Table 4.3: Samples used for signal $Z'qq + \gamma$ simulation, computed to LO via MadGraph5, Pythia8 and EvtGen.

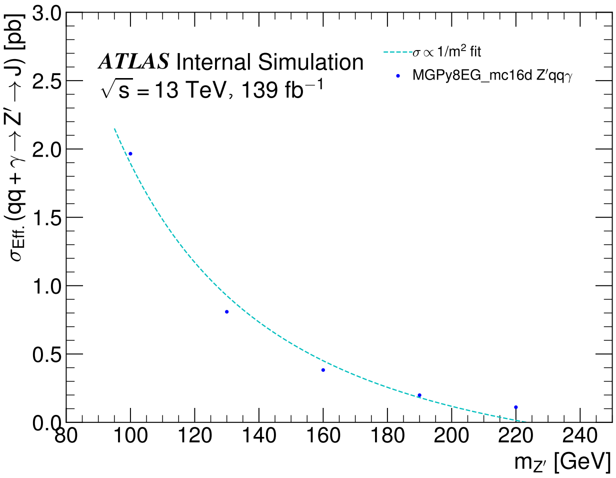

Figure 4.4: Effective cross-section modelling the production of a boosted Z' resonance in association with an initial-state radiation photon decaying to a diquark final state. Each sample pertains to a lower photon transverse momentum filter where $p^{\gamma}_{\text{T, filter}}/m_{Z'} = 1.5$

4.2 Object Definition

4.2.1 Photons and Trigger

The photon energy and direction are reconstructed using the longitudinal segmentation of the calorimeters along the shower axis [49]. Candidates are restricted to the regions where $|\eta| < 2.37$, $|\eta| \notin [1.37, 1.52]$ to exclude extremely-forward photons as well as to avoid the barrel/end-cap transition region which ensures they stem from well-calibrated sections of the calorimeter. A reference choice coined loose [61] is applied on the trigger ISR photon. It is based upon shower shapes in the second layer of the ECAL. It is chosen in order to allow greater trigger efficiency as opposed to the next reference cut named tight [61]. The latter employs information from the finely segmented strip layer of the calorimeter on top of the information obtained for the loose selection. In addition to the photon identification selection, photon isolation is also taken under consideration. A tight isolation criterion uses calorimeter and tracking information to verify that photons are separated from nearby event activity. Only photon candidates meeting the loose identification and tight isolation criteria are considered. Each event typically contains 0–4 photons in descending order of transverse momentum, p_{T}^{γ}, such that the zeroth photon is coined leading, the first subleading *etc.* The photon which satisfies the object definition criteria with the greatest p_{T}^{γ} is taken.

The lowest unprescaled single loose photon trigger is used where $p_{\mathrm{T},\gamma}^{\mathrm{HLT}} \geq 140\,\mathrm{GeV}$ for the SM simulations and data [4]. In addition, an offline cut on the photon transverse momentum of 150 GeV is placed such that the trigger efficiency is maintained high. For the BSM simulations, a larger $p_{\mathrm{T\ filter}}^{\gamma} = 1.5 \times m_{Z'}$ is used due to the requirement on the jet in the event selection, compatible with different resonances masses in use. The different trigger photon transverse momentum filters are listed in Table 4.3. Conversely to a jet trigger, the photon trigger allows for lower-p_{T} jets to be collected from which lighter resonances can be reconstructed. Previous analyses which used the jet portion of the event as the trigger object *e.g.* [49, 51] were subject to saturation of the trigger scheme with the ubiquitous multijet background resulting from pp collisions at the LHC. This limited their ability to reduce the lower bound of the resonant mass they were sensitive to. Triggering on the ISR photon enhances the flexibility to collect a wider range of jets from which the Z' is sought.

4.2.2 Jets

Jets analysed are LC-topo reconstructed using the anti-k_{T} algorithm [39] with radius $R = 1.0$ from clusters of energy deposits in the calorimeters. Basic requirements are applied in order to filter-out events containing spurious jets from detector noise and out-of-time energy deposits in the calorimeter from cosmic rays or other non-collision sources [62]. The large-R jets undergo trimming which mitigates pile-up and soft radiation effects. In this, original jet constituents are reclustered using the k_{T} algorithm with a smaller radius parameter, $R_{\mathrm{subjet}} = 0.2$, to produce a collection of subjets. A subject is disregarded if it carries less than $f_{\mathrm{cut}} = 5\%$

[4] `HLT_g140_loose`

of the original jet p_{T}. Jet energies are calibrated accordingly with the LCW-JES scheme as elaborated in Section 3.2. As per the large-R jet transverse momentum reconstruction threshold, jets are reconstructed from $p_{\mathrm{T}}^{J} = 200\,\mathrm{GeV}$ upwards. For the signal, the final-state quarks are the result of a boosted decay and thus collimated as to be merged into a single, large-R jet. Containment of the hadronic final state within the collected jet is ensured via taking jets for which $p_{\mathrm{T}}^{J} \geq 2 \times m_J$, on a par with $R = 1.0$ [63]. Like photons, in each event jets are arranged in descending order of p_{T}^{J} such that the zeroth jet is the leading jet with the highest transverse momentum. Jets are required not to overlap with the chosen photon *i.e.* satisfy the boost-invariant distance requirement $\Delta R_{\gamma,J} \geq 1.0$. Additionally, to take jets from the barrel, a candidate is required to be within $|\eta_J| < 2.0$. The jet with the highest transverse momentum satisfying the object definition requirements is chosen as an analysis candidate.

4.3 Truth Labelling

The truth fat jets (particle level) simulating the V boson signals can be probed for their reconstruction of the truth V boson (parton level). A useful technique used here on the MC jets is ghost-association to match between detector-level jets and parton-level objects. During ghost-association, the V bosons in the event are considered as ghosts *i.e.* assigned zero p_{T} and included as inputs in the jet reconstruction process. When a ghost W/Z boson is a constituent of a jet, a matching parameter, `nWBosons` or `nZBosons`, is none-zero.

In SM $W/Z + \gamma$ events, the V bosons are often produced close to the γ and therefore the leading jet in the event may contain off-resonance activity. Requiring that the jet recoils back-to-back against the photon ensures that it reconstructs an isolated V boson. The mass responses of the truth-level jets against the parton-level V bosons in SM $W/Zqq+\gamma$ events are given in FIG. 4.5 when a recoil requirement or ghost-matching is applied on the reconstructed jet. In this case, truth-level jets are associated with reconstruction-level jets via maintaining $\Delta R_{\mathrm{reco}J,\mathrm{truth}J} \leq 1.0$. The mass response of the reconstruction jet to its associated truth jet is shown in FIG. 4.6.

The V-boson match selection yields 45% and 49% efficiencies for the $Wqq + \gamma$ and $Zqq + \gamma$ signal simulation respectively. This is expected since in approximately 50% of $Vqq + \gamma$ events where the γ has $p_{\mathrm{T}} \geq 150\,\mathrm{GeV}$, the V bosons are produced with $p_{\mathrm{T}} < 200\,\mathrm{GeV}$.

The photon-jet recoil requirements pertain to efficiencies of 25% and 30% for the $Wqq+\gamma$ and $Zqq+\gamma$ signal simulation respectively, compared to the same reference. This is due to the significant amount of events where the leading jet and photon are not well-separated. Combining the ghost-association and recoil requirement has efficiencies of 12% and 21% for the $Wqq + \gamma$ and $Zqq + \gamma$ simulations respectively, for the same reference.

Once the particle-level truth jets are attuned with the parton-level truth W/Z boson, it is understood that the sought object is at hand. Later in this section it is described how the photon-jet recoil requirement, $\Delta\phi_{\gamma,J}$ cut, is optimised to yield the highest sensitivity in the context of a control study which takes the W/Z resonances

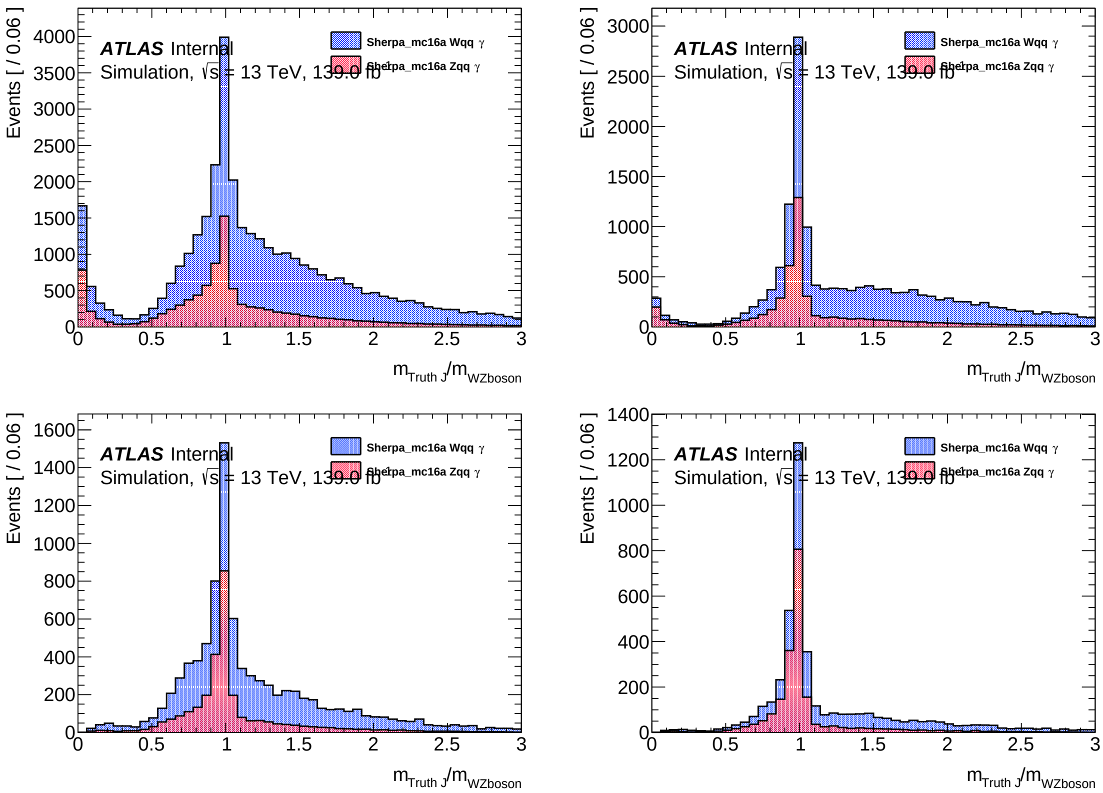

Figure 4.5: Mass responses of the truth fat jet (particle level) against the truth W/Z bosons (parton level), for the $Wqq+\gamma$ and $Zqq+\gamma$ MC simulation with different cuts set upon the reconstructed jet. (Top left) No recoil requirement nor V-boson match selection applied. (Bottom left) Recoil requirement applied without V-boson match selection. (Top right) Recoil requirement applied; leading fat jet matches a V boson in the event. (Bottom right) both recoil requirement and V-boson match selection are applied. The tally and kinematic cuts on both the photon and reconstructed jet are common to all subfigures. The recoil requirement is highly restrictive in order to ensure beyond-overlap removal and back-to-back recoil of the photon and jet: $\Delta R_{\gamma,J} \geq 1.4, \Delta\phi_{\gamma,J} \geq 0.9\pi$. Truth jets are matched to reconstructed jets via $\Delta R_{\mathrm{reco}J,\mathrm{truth}J} \leq 1.0$.

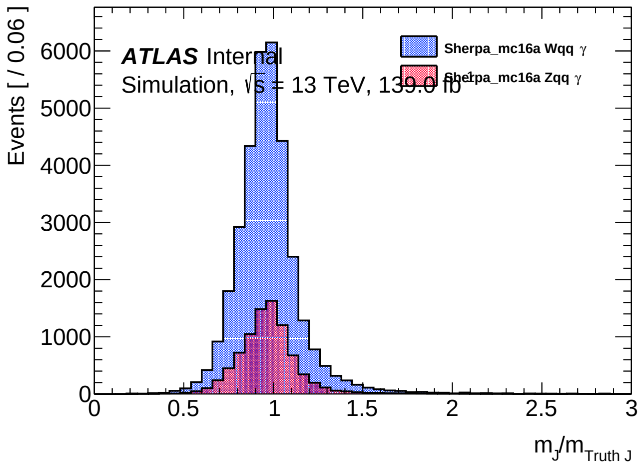

Figure 4.6: Mass response of the reconstructed fat jet (detector level) against its associated truth fat jet, for the $Wqq + \gamma$ and $Zqq + \gamma$ MC simulations. No recoil requirement nor V-boson match selection applied. Truth jets are matched to reconstructed jets via $\Delta R_{\mathrm{reco}J,\mathrm{truth}J} \leq 1.0$.

as signals.

4.4 Monte Carlo Simulation Validation

4.4.1 Standard Model Processes

Several distributions of the ISR photon and final-state large-R jet chosen accordingly with the object definition requirements as resulting from the MC background simulations are compared against the data in use. The primary aims of this probe are to compare the shapes of the distributions as well as estimate the tally in order to verift that the MC is comparable to data. The tally is used here to verify that correct cross-sections are associated with each of the MC samples compared with the data and normalise the simulation as needed. The single-photon trigger requires at least one candidate with $p_{\mathrm{T},\gamma}^{\mathrm{HLT}} \geq 140\,\mathrm{GeV}$ whereas the offline selection on the photon is $p_{\mathrm{T}}^{\gamma} \geq 150\,\mathrm{GeV}$ in order to ensure the trigger is efficient beyond a rate of 99%. In terms of tally, each event is obliged to contain at least a single photon and a single large-R jet. The LC topo large-R jets used for this analysis are reconstructed with a minimum transverse momentum of 200 GeV and hence the offline cut on them is $p_{\mathrm{T}}^{J} \geq 200\,\mathrm{GeV}$. The cut on the jet is applied in order to match the reconstruction requirement and avoid below-200 GeV fluctuations. The event selection algorithm scans the available photons and selects the highest-p_{T} candidate which satisfies the object definition. Thereafter, the jets are scanned where the highest-p_{T} candidate which meets the object definition requirements is compared against the photon. If it overlaps with the photon, the subleading candidate is checked *etc.* A series of distributions for the validation of data *vs.* MC simulations with the cuts and selection described above is given in FIG. 4.7–4.8.

The starting points in the distributions of the photon and large-R jet indicate

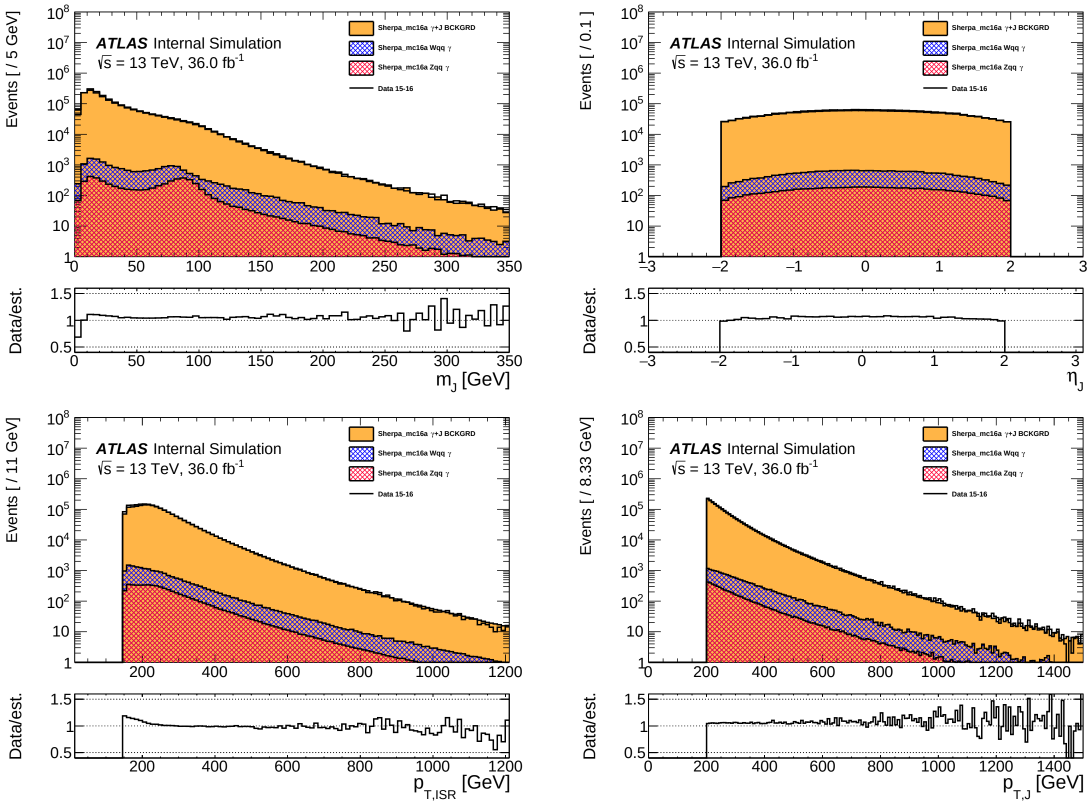

Figure 4.7: Validation of background and signal simulations against data. (Top Left) Invariant mass distribution of the selected fat jet. (Top Right) Eta angle distribution of the selected fat jet. (Bottom Left) Transverse momentum distribution of the selected initial-state radiation photon. (Bottom Right) Transverse momentum distribution of the selected fat jet. The bottom pane shows the ratio of data to the total Monte Carlo simulations, which are scaled to data at a total integrated luminosity of $\mathcal{L} = 36.0\,\text{fb}^{-1}$.

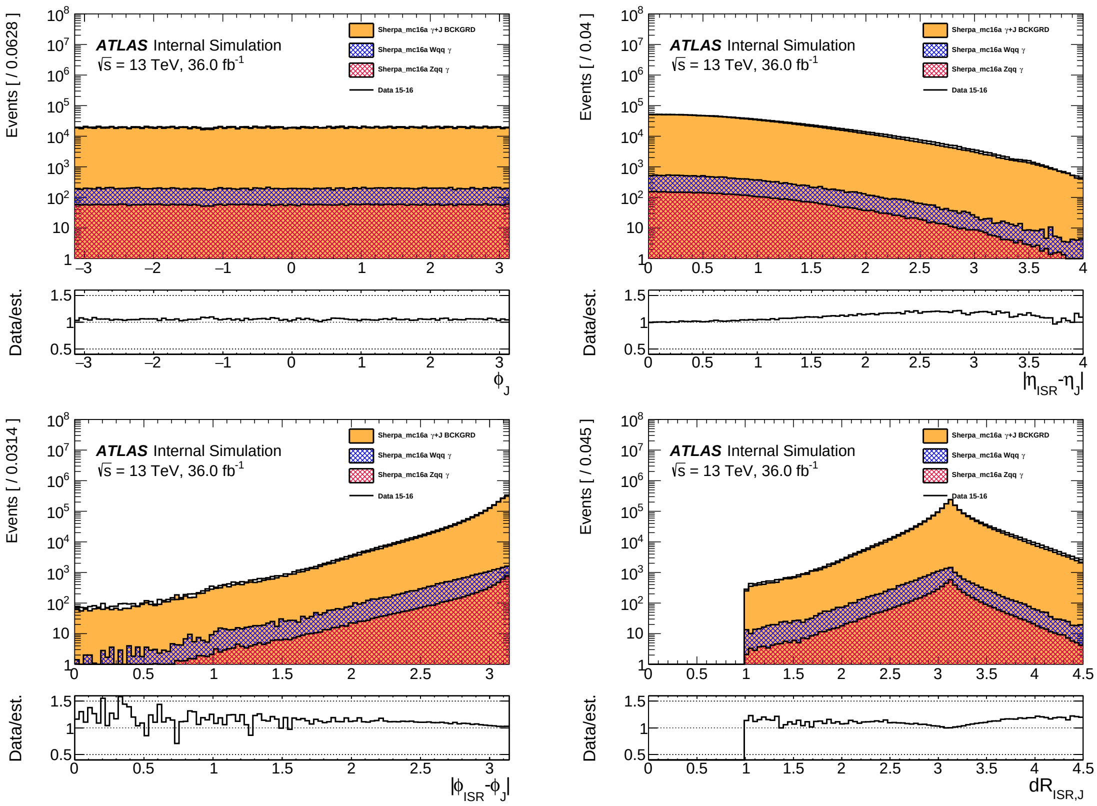

Figure 4.8: Validation of background and signal simulations against data. (Top Left) Phi angle distribution of the selected fat jet. (Top Right) Eta angle difference between the selected jet and initial-state radiation (ISR) photon. (Bottom Left) Phi angle difference between the selected jet and ISR photon. (Bottom Right) Boost-invariant angular distance between the selected jet and ISR photon. The bottom pane shows the ratio of data to the total Monte Carlo simulations, which are scaled to data at a total integrated luminosity of $\mathcal{L} = 36.0\,\text{fb}^{-1}$.

that the transverse momentum behaviours are due to the offline selections. The invariant jet mass plot shows the $J + \gamma$ background as well as signal distributions where the SM V boson resonances at $m_W = 80\,\text{GeV}$ and $m_Z = 91\,\text{GeV}$ are noticeable. Also evident are the low-mass peaks in both samples. These may stem from low containment of the final-state quarks in the large-R jet but other possible explanations include low-angular resolution regions, where final-state partons are trajected such that the scattering angle is very sharp, as common in highly-boosted interactions. To account for the low-mass peaks in the QCD background a Sudakov peak, typical in high-p_T QCD jets could be valid [64]. Due to the difficulty in calibrating jets of masses below 50 GeV, a cut is added to the event selection. The η_J distribution indicates that most jets are collected at the barrel, as often is the case in high-Q^2 transfer processes. Also apparent is that the p_T^γ remain almost constant at transverse momentum higher than the selection (that of the $Wqq + \gamma$ MC drops more suddenly than the rest) up until $220\,\text{GeV} \lesssim p_T^\gamma \lesssim 240\,\text{GeV}$. Conversely, in the p_T^J distribution, an immediate drop above $\min(p_T^J) = 200\,\text{GeV}$ is apparent throughout all of the samples. That of the $Wqq + \gamma$ MC decreases more moderately than that of the $Zqq + \gamma$ MC. The plateau in $150\,\text{GeV} \leq p_\mathrm{T}^\gamma \leq 240\,\text{GeV}$ is induced by the kinematic correlation with the jet selection.

Distributions given in FIG. 4.8 testify to the angular nature of the analysis objects. The $\Delta R_{\gamma,J}$ cutoff at 1.0 is due to the overlap definition. The $|\Delta\phi_{\gamma,J}|$ distribution yields that following object definition, the two objects recoil in the majority of SM $Vqq + \gamma$ events albeit with additional activity at lower separation which may reduce the sensitivity in the W/Z signal region. The $J + \gamma$ background yields a smoother distribution with more events at low azimuthal separation. This will be improved with further selections. The bottom panes throughout the subplots given in FIG. 4.7–4.8 yield a strong match between data and simulation, where the ratio nears unity almost constantly. Some discrepancies are seen *e.g.* in η_J and $\Delta R_{\gamma,J}$ but these do not bear a considerable influence on the analysis' sensitivity estimation. The tally comparisons show a mismatch in normalisation of 1/0.68 which is used for MC rescaling henceforth.

4.4.2 Z' Signal

The MC simulations are validated against SM MC in a similar fashion to that shown in §4.4.1 albeit not to data (comparison to data of $139\,\text{fb}^{-1}$ luminosity is due at a later stage of the analysis and not covered in this thesis). The kinematic and angular distributions match the SM simulations, as expected. Two key observations relate to the invariant Z' jet mass and the angular distribution of the reconstructed jets, compared to truth jets. The mass is the basis upon which the sensitivity is computed whereas the truth-jet ΔR matching is prone to testify whether the detecor-level and particle-level objects share propagation axes and thus whether the former is a valid reconstruction of the resonance. The two plots are shown in FIG. 4.9 for all Z' mass points.

The m_J distribution shows a spread about respective $m_{Z'}$ values which grows wider with increasing mass, a factor which might affect the sensitivity using a signal window about the resonances. Detector-level jets show a good match to particle-

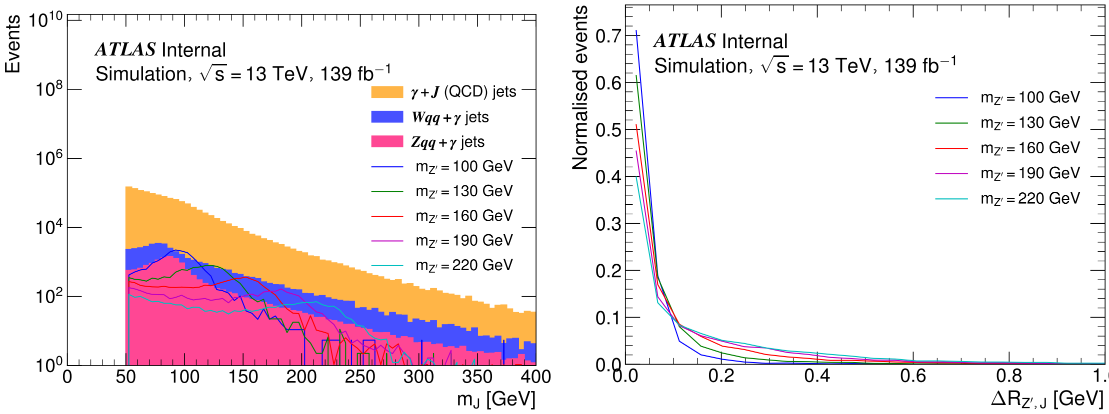

Figure 4.9: Kinematic validation of the $Z'qq+\gamma$ simulation. (Left) Invariant mass distribution of the chosen large-radius jet. (Right) Boost-invariant angle difference distribution between the reconstructed Z' jet and the truth Z' jet.

level (truth) jets, which validates the reconstructed jets as candidates from which the Z' signal may be probed.

4.5 Event Selection

From §4.3 it is inferred how the resonance can be reconstructed to best match the truth parton and particle levels whilst reducing the background at a minimal efficiency cost. The event selection includes kinematic and topological selections on top of trigger and object definition.

Events are collected by the lowest unprescaled photon trigger (140 GeV). They are required to have at least one photon of $p_T^\gamma \geq 150$ GeV and at least one non-overlapping large-R jet. The jet satisfies $p_T^J \geq 200$ GeV, $m_J \geq 50$ GeV and $p_T^J \geq 2 \times m_J$ for $R = 1.0$. The leading jet meeting the above criteria is considered as the resonance candidate. Both objects are chosen at leading transverse momentum.

In the process of interest, the resonance recoils against the photon. Topological cuts require the photon and resonance candidate jet to be separated azimuthally, where $\Delta\phi_{\gamma,J} = 0$ would mean they share a trajectory axis and $\Delta\phi_{\gamma,J} = \pi$ that they perfectly recoil back-to-back. It is deduced from the truth objects that $\Delta\phi_{\gamma,J} < 0.5\pi$ is dominated by background and therefore the range $\Delta\phi_{\gamma,J} \in [0.5\pi, \pi]$ is tested. In the purpose of maximising the sensitivity to the W/Z resonances through the jet invariant mass distribution, it is found that $\Delta\phi_{\gamma,J} = 0.5\pi$ is optimal. *N.b.* in this, not only the best signal reconstruction is considered, but also background reduction, such that the two must be chosen in conjunction recalling the sensitivity is defined $S/\sqrt{B}$.

4.5.1 Adversarial Neural Network Tagging

Due to the considerable background of non-resonant jets, this search relies upon the identification of boosted hadronically decaying resonances, reconstructed as a large-R jet. Jet substructure variables can be used to indicate the differences in the radiation pattern of jets resulting from resonant and off-resonance decays and be applied in order to separate between the two [65].

Multivariate analysis (MVA) taggers attribute much weight to the jet mass when they are trained on substructure variables due to the non-linear correlations of the decaying resonance with the jet mass [66]. Therefore, cuts on MVA classifier score often lead to a distortion of the background jet mass distributions and hence, its resemblance to that of the resonance jet mass, a phenomenon also referred to as sculpting. To avoid this whilst maintaining the benefits of jet substructure taggers, it is necessary to decorrelate the MVA classification from the jet mass.

Adversarial Neural Networks (ANNs) [67] are commonly used for classifying whilst decorrelating chosen parameters from the entire learning. In the context of jet tagging in ATLAS, ANNs constitute one of the techniques used for jet mass decorrelation [65]. Through MVA, jet substructure variables are employed to identify hadronically-decaying resonances whilst avoiding distortions to the jet mass. The simulated LC topo jets analysed in this study are assigned an ANN score similarly to [66]. This section describes the ANN's configuration and workings. The ANN's efficacy in discriminating background jets in the context of this analysis is also assessed.

Configuration and Architecture

In order to make the classification of jets independent of the jet mass, adversarial training is applied. The classifier intakes the 10 substructure variables discussed in Section 3.3 and outputs a jet tagging variable, `zANN` $\in [0, 1]$. Initially, it is trained whilst aiming to minimise the classification loss function, $L_{\text{clf.}}$, through which the correlation with m_J is introduced. Thereupon a second net, the adversary, is employed to deduce m_J from the first net's output whilst minimising its own loss function, $L_{\text{adv.}}$. The combined ANN is trained through a min-max optimisation which balances the classification against the decorrelation,

$$\min_{\theta_{\text{clf.}}} \max_{\theta_{\text{adv.}}} L_{\text{clf.}}(\theta_{\text{clf.}}) - \lambda L_{\text{adv.}}(\theta_{\text{clf.}}, \theta_{\text{adv.}}), \tag{4.3}$$

where θ denotes the weights of a net. The internal optimisation finds the $\theta_{\text{adv.}}$ for a given $\theta_{\text{clf.}}$ to maximise $L_{\text{adv.}}(\theta_{\text{clf.}}, \theta_{\text{adv.}})$, as the adversary net's goal is contingent upon the classifier's configuration. The external optimisation finds the weights, $\min_{\theta_{\text{clf.}}}$, which minimise the effective loss, $L_{\text{clf.}}(\theta_{\text{clf.}}) - \lambda L_{\text{adv.}}(\theta_{\text{clf.}}, \theta_{\text{adv.}})$.

The balance between the classifier and adversary is determined by λ: $\lambda \to 0$ yields the standard neural network (NN) whereas $\lambda \to \infty$ grants that the classifier solely retains information uncorrelated with the jet mass.

The ANN architercture, inclusive of the classifier and adversary nets, is presented in FIG. 4.10. $\log(p_{\text{T}}^J)$ is used as an auxiliary parameter (a in the figure) to the adversary net in order to maintain decorrelation throughout a wide p_{T} range.

Impact of the ANN Classifier on the Jet Mass Distribution

The ANN classifier score, `zANN`, is expected to be large for resonance jets and small for background jets. A selection on `zANN` scores reduces the background without

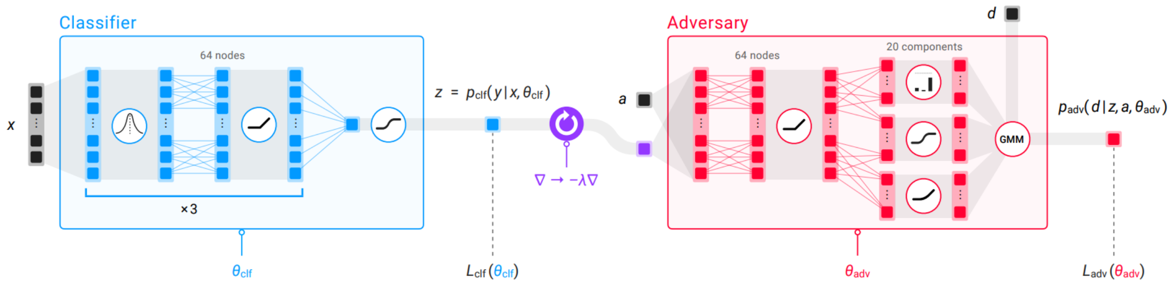

Figure 4.10: Adversarial neural network architecture. Based on 10 input jet substructure variables (denoted x) summarised in Table 3.1, the classifier network predicts jet labels (denoted y) and eventually outputs a tagger variable (denoted z). Thereupon the adversary network deduces the values of variables *e.g.* m_J, from which it decorrelated the classifier (denoted d). Auxiliary jet features *e.g.* $\log(p_{\mathrm{T}}^{J})$ (denoted a) may improve the adversary's ability in the decorrelation task. The adversarial training employs a gradient-reversal layer which parameterises the counteraction between $L_{\mathrm{clf.}}$ and $L_{\mathrm{adv.}}$, the classifier and adversary loss functions respectively, controlled by λ. [65].

sculpting the jet mass distribution and keeps a large fraction of resonant jets. To observe this, the jet mass distribution of the $Wqq + \gamma$ and $Zqq + \gamma$ simulations are evaluated against the $J + \gamma$ background for different `zANN` cuts, as given in FIG 4.11.

In FIG. 4.11, the W/Z jets are clearly identified by the ANN: non-resonance jets are reduced and the mass peaks about m_W and m_Z become sharper as the minimal `zANN` increases, seen especially in the bottom plots. The bottom right subfigure where min(`zANN`) = 0.7 shows a narrower distribution about m_Z than about m_W. In the same plot, an undesirable high-mass tale is apparent in both signal samples. With regard to the background, a reduction is visible at about an order of magnitude between min(`zANN`) = 0.5 and min(`zANN`) = 0.7 cut and about half an order of magnitude between the lower cuts. The jet mass distribution is nearly non-sculpted up until high `zANN` values. The cutflow in Table 4.4 shows the number of events remaining and corresponding efficiency following these selections.

Criterion	$\gamma + J$ background		$Wqq + \gamma$ signal		$Zqq + \gamma$ signal	
	Events remaining	% remaining	Events remaining	% remaining	Events remaining	% remaining
`zANN` ≥ 0.0	1212067	100	35478	100	12361	100
`zANN` ≥ 0.1	947616	78	29020	82	10542	85
`zANN` ≥ 0.3	541883	48	18284	52	7138	58
`zANN` ≥ 0.5	251680	21	10169	29	4093	33
`zANN` ≥ 0.7	70996	5.9	3900	11	1679	14

Table 4.4: Cutflow for Monte Carlo simulation of the $\gamma + J$ background and both the $Vqq + \gamma$ Standard Model signal samples. The total initial samples, corresponding to 100% in the table, is the one following object definition and baseline selections (including $m_J \geq 50\,\mathrm{GeV}$) mentioned in the text in addition to a recoil requirement of $\Delta\phi_{\gamma,J} \geq 0.7\pi$.

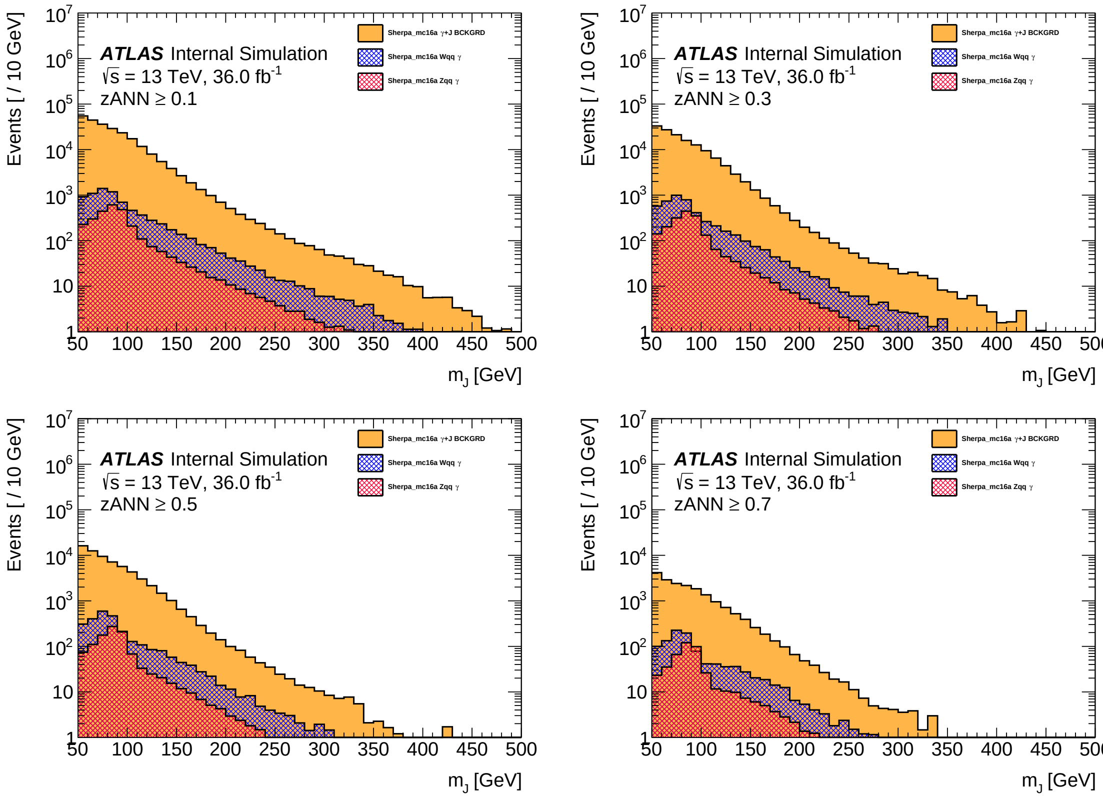

Figure 4.11: Jet mass distribution of the Standard Model $Vqq + \gamma$ simulations against the single large-R jet+γ background, evaluated at different minimal ANN scores. (Top left) `zANN` ≥ 0.1. (Top right) `zANN` ≥ 0.3. (Bottom left) `zANN` ≥ 0.5. (Bottom right) `zANN` ≥ 0.7. A photon-jet recoil requirement of $\Delta R_{\gamma,J} \geq 1.0$ and $\Delta\phi_{\gamma,J} \geq 0.7\pi$ is common to all subplots. High-level trigger requirement on the initial-state radiation photon and offline transverse momenta selections are applied as outlined in the text.

Classification Performance in the Search

In [65] the performance of the ANN is assessed and the results obtained there are used as a reference. The reference `zANN` distribution as well as the receiver operating characteristic (ROC) curve are given in FIG. 4.12. A defining feature apparent in the left subfigure of FIG. 4.12 is the multitude of signal events in high-`zANN` regions against a decreasing amount of background events. Conversely, the number of signal events in the low-`zANN` regions is low, compared with background events. This is indicative of the ANN's success in tagging signal *vs.* background events.

A similar assessment is performed in this search. The performance of the ANN is tested following object definition, trigger, tally and kinematic selections elaborated previously. In order to test the ANN performance whilst "feeding" it purified signal as input, the W/Z events are selected stringently: ISR-jet topology maintains $\Delta R_{\gamma,J} > 1.4$ and $\Delta\phi_{\gamma,J} > 0.9\pi$. The signal region is taken as 60 GeV–100 GeV and 70 GeV–110 GeV for the W and Z bosons, respectively, and corresponding cuts are applied on the reconstructed jet mass. Every selected event has its reconstructed leading fat jet ghost-matched to a truth V boson. Moreover, the reconstructed and truth fat jets are constrained to be within $\Delta R \leq 1.0$ of each other's axis. The re-

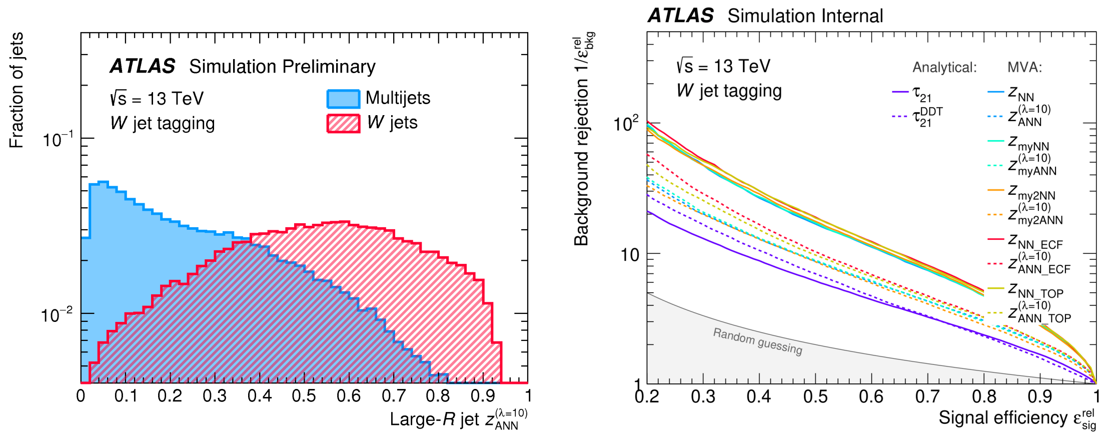

Figure 4.12: Reference performance of the Adversarial Neutral Network (ANN) on W jets against simulated Standard Model background. (Left) Distribution of the `zANN` score variable, the eventual output of the tagger tool following learning based on 10 jet substructure variables described in the text. (Right) Receiver operating characteristic (ROC) curve showing the background rejection through the signal efficiency spectrum for different taggers. The MVA tagger relating to the ANN is shown in dashed sky blue [66].

sponse between the truth fat jet (particle level) and the truth V boson (parton level) is kept within 0.8–1.2. Thereby, the reconstructed jets provide a reliable account of the real resonance simulation, at truth parton and particle levels. Given these selections, the ANN score distributions of the background and signal simulation are given in FIG. 4.13.

The histograms in FIG. 4.13 show that the ANN rightly ascribes low scores to background events, which peak around $0 < \texttt{zANN} < 0.1$. Nevertheless, it can be seen that relatively too many signal events are assigned with low ANN scores despite cuts around the resonance regions. Hence, too-few signal events are awarded high `zANN` values which, in turn, hinders the overall background rejection obtained by the tagger. This can be illustrated through the ROC curve corresponding to this ANN score distribution, as shown in FIG. 4.14.

The multiplicity of signal events in the low-`zANN` regions reduce the background rejection attained at *e.g.* $\epsilon_{\text{sig.}} = 0.5$ and above. When examining the ANN score distributions in FIG. 4.13, it can be inferred that a signal efficiency of $\sim 50\%$ is attained after reaching a low ANN score, where the background rejection is low. At high ANN scores, where there is very little amount of background events left, the signal efficiency is already very high. This negates a desired tagging performance where a 50% signal efficiency is reached at higher ANN scores which would, in turn, result in an improved background rejection. Although the selection is purified by taking events where the V boson recoils against the ISR photon, matching the reconstructed jet to a truth V boson as well as maintaining it geometrically in-line with the truth jet, the ANN did not produce desired outcomes.

The poor performance is also observable when comparing the ROC curves shown in FIG. 4.14 to those in FIG. 4.12. Where the reference yields a background rejection of approximately 30 and 10 for signal efficiencies of 25% and 50% respectively, the ANN operating in the context of this analysis gains respective background rejections

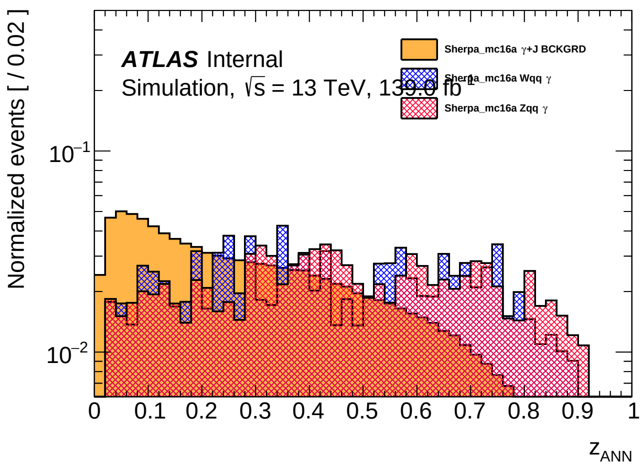

Figure 4.13: Adversarial Neural Network (ANN) score of the Standard Model $Vqq+\gamma$ simulations and $J+\gamma$ background simulations for optimal event selection with jet mass windows of $60\,\mathrm{GeV} \leq m_J \leq 100\,\mathrm{GeV}$ and $70\,\mathrm{GeV} \leq m_J \leq 110\,\mathrm{GeV}$ for $Wqq+\gamma$ and $Zqq+\gamma$ signal events, respectively. The histograms are scaled to an integrated luminosity of $\mathcal{L} = 139.0\,\mathrm{fb}^{-1}$ and normalised to unity.

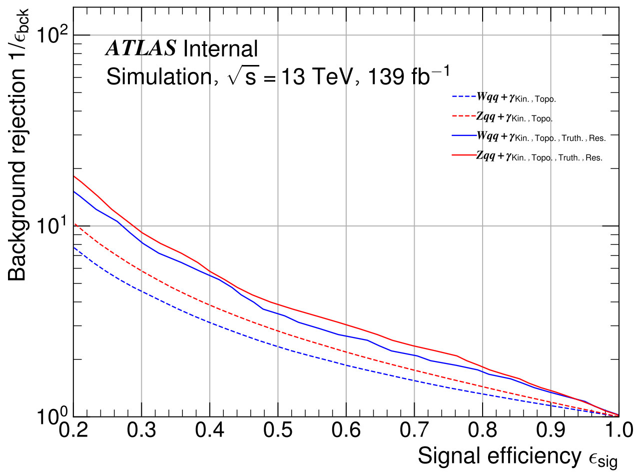

Figure 4.14: Receiver operating characteristic (ROC) curve corresponding to the Adversarial Neural Network (ANN) score of the Standard Model $Vqq+\gamma$ simulations and $\gamma+J$ background simulations. Curves representing tightening events selections starting from basic tally and kinematic cuts, through ISR-jet topology to truth-level-aided selection and m_J cuts around the resonances are shown. As the events taken become a purer representation of the signal resonance, the ANN's background rejection increases.

<table>
<tr><th>Selection type</th><th>Analysis object/ cut parameter</th><th>ISR-γ (HLT)</th><th>ISR-γ</th><th>Reco J (detector level)</th></tr>
<tr><td>Tally</td><td>$=$</td><td>≥ 1</td><td>≥ 1</td><td>≥ 1</td></tr>
<tr><td rowspan="2">Kinematics</td><td>p_{T} [GeV]</td><td>≥ 140</td><td>≥ 150</td><td>$\geq 200\,\&\geq 2\times m_J$</td></tr>
<tr><td>m [GeV]</td><td rowspan="4">N/A</td><td>N/A</td><td>≥ 50</td></tr>
<tr><td rowspan="3">Topology</td><td>$|\eta|$</td><td>< 2.47; $\notin (1.36, 1.52)$</td><td>< 2.0</td></tr>
<tr><td>$\Delta\phi$</td><td colspan="2">$\geq 0.5\pi$</td></tr>
<tr><td>ΔR</td><td colspan="2">≥ 1.0</td></tr>
</table>

Table 4.5: Schematic representation of the different cuts used *en route* to the final event selection. N/A — not applicable; N/S — not selected.

of 12 and 4.2 for the W jets and 12.1 and 4.0 for the Z jets. The similarities of the background zANN suggest that the tagger correctly classifies background jets, unlike its tagging of signal jets which does not work as expected.

One may check whether the ANN tagging improves the sensitivity of a search ($S/\sqrt{B}$) by means of an enhancement parameter, defined as

$$\alpha = \frac{\epsilon_{\mathrm{SGNL}}}{\sqrt{\epsilon_{\mathrm{BCKG}}}}, \tag{4.4}$$

where $S, \epsilon_{\mathrm{SGNL}}$ and $B, \epsilon_{\mathrm{BCKGRD}}$ are the baseline number of events and efficiency for the signal and background, respectively. Therefore, if $\alpha \leq 1$ the tagger is not useful and *vice versa*. The values for the enhancement parameter gained whilst using the pure signal described above are 0.812 and 0.87 at 25% signal efficiency for the W and Z jets respectively in addition to 0.92 and 0.99 at 50% signal efficiency for the W and Z jets, respectively. Therefore, it is deemed that in the context of this analysis program, the ANN does not possess an added value on top of other event selection methods. Due to a lower performance than had been expected, in future analysis separate distributions of substructure observables used as inputs to the ANN will be investigated further. Such work may resolve a possible behaviour, currently unknown, which is hindering the tagger to seize its full potential in amplifying the overall sensitivity to the signal.

4.5.2 Final Event Selection

The mass-decorrelated tagger is not yet useful in enhancing the sensitivity but might well become so in the future. Final event selection includes trigger, object definition and multiplicity requirements, as well as the topology and kinematic cuts as explained previously. A summary of the event selection is given in Table 4.5.

The number of background events left through the selection process is described in Table 4.6 for the SM $J + \gamma$ background and in Table 4.7 for the SM $Vqq + \gamma$. Event selection is equal for the $J + \gamma$ and $Vqq + \gamma$ simulations. These are scaled to $139.0\,\mathrm{fb}^{-1}$ in order to provide the expected background for the full analysis which will use the latest total integrated LHC luminosity. In the control analysis described in Section 4.6, the $Vqq+\gamma$ portion of the background is treated as a reference signal to estimate the performance of this analysis program to vector resonances using a

Criterion	Single photon trigger	
	Events remaining	% remaining
All	20736588	100 %
Trigger	12572643	61 %
$n_\gamma \geq 1$	9799430	47 %
$p_{\mathrm{T}}^{\gamma} \geq 150\,\mathrm{GeV}$	9310281	45 %
$\Delta R_{\gamma,J} \geq 1.0$	5128595	25 %
$n_J \geq 1$	5128595	25 %
$p_{\mathrm{T}}^{J} \geq 200\,\mathrm{GeV}$	5128595	25 %
$m_J \geq 50\,\mathrm{GeV}$	1310425	6.3 %
$p_{\mathrm{T}}^{J} \geq 2 \times m_J$	1258646	6.0 %
$\Delta\phi_{\gamma,J} \geq 0.5\pi$	1212067	5.8 %

Table 4.6: Cutflow for $\gamma + J$ background simulation with samples scaled to total integrated luminosity of $\mathcal{L} = 139\,\mathrm{fb}^{-1}$. Percentages remaining are relative to the total sample. The topological cuts are done independently of the jet mass cuts.

part of the total luminosity for data.

4.5.3 Z' Event Selection and Cutflow

The event selection optimised for the SM W/Z control study is applied with a small caveat: rather than using the 140 GeV HLT ISR photon, each Z' mass pertains to a different filter transverse momentum at which the offline cuts is applied too. As shown below, this costs no more than 1% in efficiency reduction. Here it is also expected that the dominant background stems from the common $J + \gamma$. In addition the SM $Vqq + \gamma$ events are included as part of the background, which amounts to less than 4% of the total background expectation.

Upon the the Z' MC simulation samples described in §4.1.2 an event selection is applied and summarised in Tables 4.8–4.10 through cutflow.

The cutflows described above are scaled to a total integrated luminosity of $139\,\mathrm{fb}^{-1}$ in order to infer the expected sensitivity potential when the full ATLAS data available are used. The selections on the different mass point samples yield a relatively-stable decline where the jet mass cut and azimuthal separation selection filter-out about 50% of the remaining sample each. These are in turn used to estimate the acceptances needed for the sensitivity estimation of ATLAS to a new resonance of such sort. They are summarised and compared in the inclusive acceptance plot shown in FIG. 4.15. The near-stagnant bottom yellow line shows the overall stable acceptance, subjecting the lot of samples to the same analysis scheme.

Criterion	$Wqq+\gamma$ signal		$Zqq+\gamma$ signal	
	Events remaining	% remaining	Events remaining	% remaining
All	174269	100 %	54231	100 %
Trigger	160708	92 %	50246	93 %
$n_\gamma \geq 1$	118477	68 %	38815	72 %
$p_{\mathrm{T}}^{\gamma} \geq 150\,\mathrm{GeV}$	112572	65 %	37398	69 %
$\Delta R_{\gamma,J} \geq 1.0$	78111	45 %	22936	42 %
$n_J \geq 1$	78111	45 %	22936	42 %
$p_{\mathrm{T}}^{J} \geq 200\,\mathrm{GeV}$	78111	45 %	22936	42 %
$m_J \geq 50\,\mathrm{GeV}$	41093	24 %	13783	25 %
$p_T^J \geq 2 \times m_J$	39239	23 %	13231	24 %
$\Delta\phi_{\gamma,J} \geq 0.5\pi$	38250	22 %	13035	24 %

Table 4.7: Cutflow for $Vqq+\gamma$ signal simulation using a single photon trigger. Percentages remaining are relative to the total sample.

Criterion / $m_{Z'}$ [GeV]	100 ($p_{\mathrm{T,\,filter}}^{\gamma} = 150$)		130 ($p_{\mathrm{T,\,filter}}^{\gamma} = 195$)	
	Events remaining	% remaining	Events remaining	% remaining
All	144192	100 %	103748	100 %
$n_\gamma \geq 1$	114057	79 %	84255	81 %
$p_{\mathrm{T}}^{\gamma} \geq p_{\mathrm{T,\,filter}}^{\gamma}$	113806	79 %	82557	80 %
$n_J \geq 1$	113014	78 %	82197	79 %
$\Delta R_{\gamma,J} \geq 1.0$	57133	40 %	42363	41 %
$p_{\mathrm{T}}^{J} \geq 200\,\mathrm{GeV}$	57133	40 %	42363	41 %
$m_J \geq 50\,\mathrm{GeV}$	36735	25 %	22296	21 %
$p_T^J \geq 2 \times m_J$	35265	24 %	20695	20 %
$\Delta\phi_{\gamma,J} \geq 0.5\pi$	17758	12 %	10221	10 %

Table 4.8: Cutflow for Z' signal simulation. Events remaining are computed via $\mathcal{L} \times \sigma \times \epsilon$ where $\mathcal{L} = 139\,\mathrm{fb}^{-1}$. Percentages remaining are relative to the total sample.

Criterion / $m_{Z'}$ [GeV]	160 ($p_{\mathrm{T,\ filter}}^{\gamma} = 240$) Events remaining	% remaining	190 ($p_{\mathrm{T\ ,\ filter}}^{\gamma} = 285$) Events remaining	% remaining
All	50844	100 %	26943	100 %
$n_{\gamma} \geq 1$	41411	81 %	22046	82 %
$p_{\mathrm{T}}^{\gamma} \geq p_{\mathrm{T,\ filter}}^{\gamma}$	40426	80 %	21565	80 %
$n_J \geq 1$	40366	79 %	21549	80 %
$\Delta R_{\gamma,J} \geq 1.0$	27913	55 %	17518	65 %
$p_{\mathrm{T}}^{J} \geq 200\,\mathrm{GeV}$	27913	55 %	17518	65 %
$m_J \geq 50\,\mathrm{GeV}$	13163	26 %	7865	29 %
$p_T^J \geq 2 \times m_J$	12244	24 %	7289	27 %
$\Delta\phi_{\gamma,J} \geq 0.5\pi$	6127	12 %	3620	13 %

Table 4.9: Cutflow for Z' signal simulation. Events remaining are computed via $\mathcal{L} \times \sigma \times \epsilon$ where $\mathcal{L} = 139\,\mathrm{fb}^{-1}$. Percentages remaining are relative to the total sample.

Criterion / $m_{Z'}$ [GeV]	220 ($p_{\mathrm{T,\ filter}}^{\gamma} = 330$) Events remaining	% remaining
All	15179	100 %
$n_{\gamma} \geq 1$	12505	82 %
$p_{\mathrm{T}}^{\gamma} \geq p_{\mathrm{T,\ filter}}^{\gamma}$	12197	80 %
$n_J \geq 1$	12194	80 %
$\Delta R_{\gamma,J} \geq 1.0$	11099	73 %
$p_{\mathrm{T}}^{J} \geq 200\,\mathrm{GeV}$	11099	73 %
$m_J \geq 50\,\mathrm{GeV}$	4790	32 %
$p_T^J \geq 2 \times m_J$	4460	29 %
$\Delta\phi_{\gamma,J} \geq 0.5\pi$	2179	14 %

Table 4.10: Cutflow for Z' signal simulation. Events remaining are computed via $\mathcal{L} \times \sigma \times \epsilon$ where $\mathcal{L} = 139\,\mathrm{fb}^{-1}$. Percentages remaining are relative to the total sample.

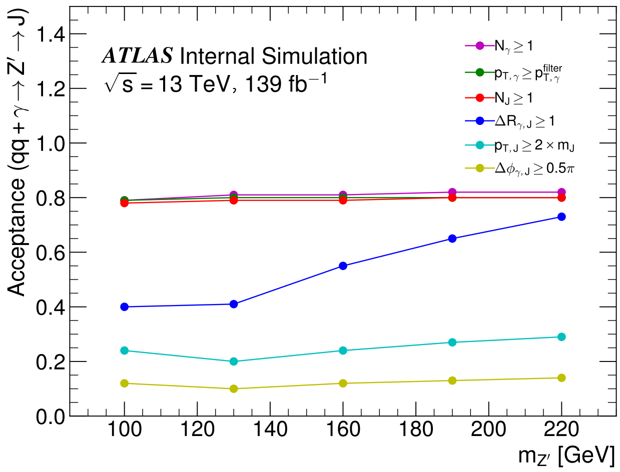

Figure 4.15: Acceptance of the simulation modelling a $qq+\gamma \to Z' \to J$ process as a function of the event selection applied upon different $m_{Z'}$ samples.

4.6 Control Region

4.6.1 Background Expectation

Within the BSM Z' analysis, the main background consists of any QCD process which produces a photon and a large-R jet which is not mediated by a V boson. In addition, two massive SM resonances, W and Z, each produced boosted in association with a photon and decays into a collimated quark pair contribute slightly to the background. Referred to jointly as $Vqq+\gamma$, these simulations are first used as test signals in the context of a control analysis where the background is only $J+\gamma$. Upon estimating the sensitivity to new vector resonances, the $Vqq+\gamma$ simulations constitute a fraction of the total background estimation. The samples used for background simulation and their key details are listed in Tables 4.1–4.2. Estimating the sensitivity to a vector boson signal is initially conducted upon the SM particles. Events are analysed via the algorithm employing the event selection scheme outlined in the previous section.

4.6.2 Control Analysis

In order to evaluate the program's ability to discover new vector boson resonances within a known mass régime, the SM $Wqq+\gamma$ and $Zqq+\gamma$ simulations are taken as signal and compared to the $J+\gamma$ background; the latter constitutes the preeminent background in the $Z'qq+\gamma$ investigation. The event selection described in Section 4.5 which finalises in the one elaborated in §4.5.2 at reconstruction level is the result of optimising the sensitivity to the SM bosons within this controlled study. Data of 36.0 fb^{-1} integrated luminosity subject to the same event selection are used. The MC is scaled to the luminosity and then rescaled to data by 1/0.68 on a par with the discrepancy found in §4.4.1. It is sought to find excesses in the invariant mass distribution of the final-state large-R jet against the background in addition to observe the behaviour of data in the W and Z jets signal regions. A plot comparing

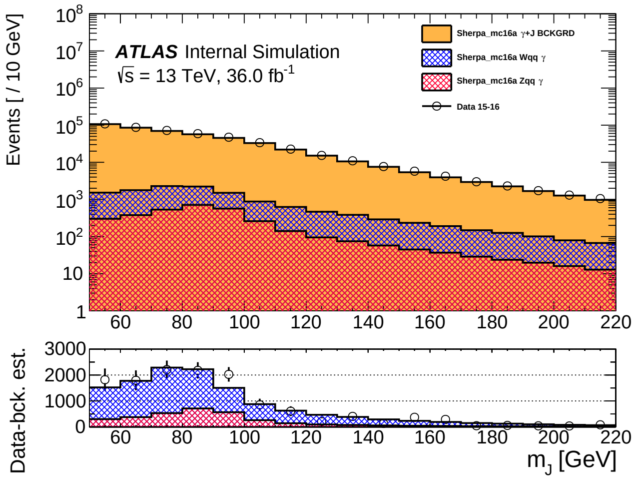

Figure 4.16: Distribution of the chosen fat jet invariant mass, m_J, following optimised tally, kinematic and topological event selections. By object definition, the photon and jet are separated and the two recoil ($\Delta\phi_{\gamma,J} \geq 0.5\pi$). The trigger efficiency of the photon is complete via an offline p_{T}^{γ} cut, 10 GeV above the high level trigger threshold. The fat jet is selected at minimal reconstruction transverse momentum. To maintain a collimated fat jet wherein the two final-state quarks are contained, a $p_{\mathrm{T}}^{J} \geq 2 \times m_J$ cut is applied. The simulation is scaled to the data produced at $\mathcal{L} = 36.0\,\mathrm{fb}^{-1}$ and thereafter rescaled by 1/0.68 accordingly with the discrepancy noted upon basic validation. The bottom pane shows the data after background estimation subtraction, compared against the two $Wqq+\gamma$ and $Zqq+\gamma$ simulations, the former stacked on top of the latter.

these is given in FIG. 4.16.

The match between the excesses in m_V signal regions in the MC and data is apparent. Sensitivity[5] attained by applying the kinematic and topological selections described in Tables 4.6–4.7 (with $m_J \geq 50\,\mathrm{GeV}$) at 90% of the signal from the peak are 31.33 and 10.86 corresponding to ranges of 50 GeV–155 GeV and 50 GeV–135 GeV for the $Wqq+\gamma$ and $Zqq+\gamma$ simulations, respectively. These significances correspond to 34159 and 1188656 events for the $Wqq+\gamma$ signal and $\gamma + J$ background simulations respectively, and to 11676 and 1156172 events for the $Zqq+\gamma$ signal and $\gamma + J$ background simulations, respectively. Event numbers are calculated accordingly with

$$N_{\mathrm{Events}} = \sigma \times \mathcal{L} \times \mathcal{A} \times \epsilon, \tag{4.5}$$

where σ is the cross section, $\mathcal{L}$ the total integrated luminosity, $\mathcal{A}$ the acceptance and ϵ the efficiency. Therefore, in terms of a discovery potential and compared to the SM, this analysis scheme is sensitive to processes including a production of a SM V boson in association with an ISR photon where the resonance is reconstructed from a single large-R jet at a cross-section as small as $5/31.33 \times \sigma(q\bar{q}+\gamma \to W \to J)$ and $5/10.86 \times \sigma(q\bar{q}+\gamma \to Z \to J)$, referring the the individual significances separately. Succinctly, these discovery potentials can be summarised as

[5]The sensitivity is defined as $S/\sqrt{B}$ where S and B are the number of signal and background events in the simulation samples described in the text

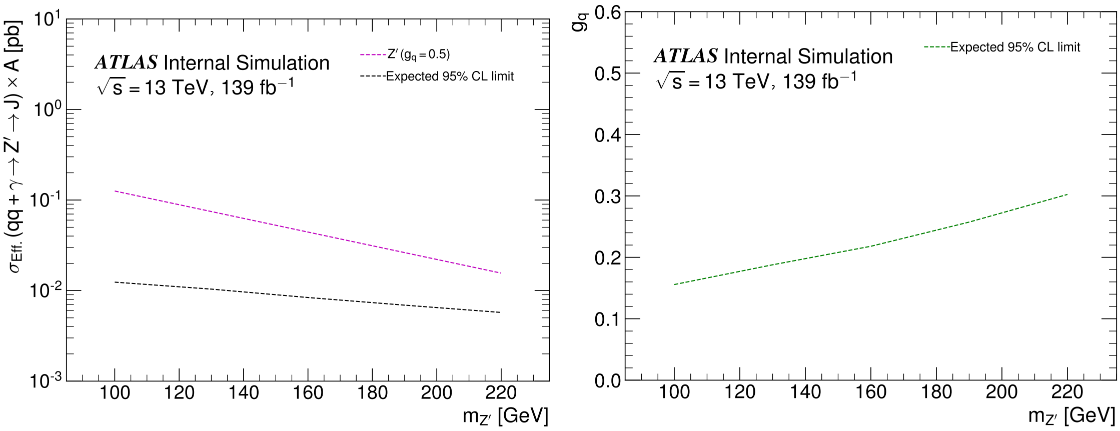

Figure 4.17: Predicted sensitivity to a boosted BSM vector boson Z' produced in association with and ISR photon and decays to a diquark final state, collected as a single large-radius jet, as a function of $m_{Z'}$. (Left) Production cross-section as predicted by theory and its limit at 95% CL interval, both calculated with the acceptance. (Right) Coupling to SM quarks limit at 95% CL interval.

$$
\begin{aligned}
5/(S/\sqrt{B}) \times \sigma(q\bar{q} + \gamma \to W \to J) &= 0.16 \times \sigma(q\bar{q} + \gamma \to W \to J) \\
5/(S/\sqrt{B}) \times \sigma(q\bar{q} + \gamma \to Z \to J) &= 0.46 \times \sigma(q\bar{q} + \gamma \to Z \to J),
\end{aligned}
\tag{4.6}
$$

where a discovery is taken as sensitivity greater or equal to 5σ. Based on these results, further exploration can be carried out to be applied on new vector bosons of various masses decaying to SM quarks.

4.6.3 Sensitivity Limits to Beyond the Standard Model Vector Bosons

Using the acceptances as calculated in §4.5.3, the cross-sections are rescaled per each mass. Recalling the production cross-section shown in FIG. 4.4 and defining the sensitivity cross-section as

$$
\sigma(qq + \gamma \to Z' \to J) \times \epsilon_{\text{filter}} \times \mathcal{A}, \tag{4.7}
$$

where the effective production cross-section, $\sigma_{\text{Eff.}}(qq + \gamma \to Z' \to J) = \sigma(qq + \gamma \to Z' \to J) \times \epsilon_{\text{filter}}$, is multiplied by the acceptance defined as the portion left of the original samples following the applied event selection. Recalling $\sigma(qq + \gamma \to Z' \to J) \propto g_q^2$, the sensitivity to a new boosted Z' resonance produced in association with an ISR photon and decaying to a diquark final-state can also be evaluated in terms of its coupling to SM quarks. The two are given to 95% CL in the limit plots shown in FIG. 4.17. The estimated background to the Z' resonances includes all simulated SM processes in the control region, namely the QCD $J + \gamma$ and $Vqq + \gamma$ simulations.

These limit plots show a slightly-improved sensitivity at certain mass points than the reference study [53] which also did not benefit greatly from the jet-substructure-based Designed Decorrelated Tagger it used. The sensitivity to cross-section rejects mostly all events above $\sigma_{\text{Eff.}} \times \mathcal{A} \sim 10^{-2}$ which constitutes a near-order-of-magnitude improvement compared to the theoretical line in the low-mass régime. The cross-section limit does however decline moderately as $m_{Z'}$ grows which in turn does not produce a significant sensitivity enhancement. This is compared to the reference where a sharper descent is observed. In terms of the coupling limit, the low mass yield a sensitivity to a lower coupling by 25%. The sensitivity to coupling quickly reduces with increasing Z' mass where a sharper rise than that seen in the reference is present. If a more-moderate rise for masses above 100 GeV were kept, an improved coupling sensitivity would have been attained throughout the $m_{Z'}$ spectrum. In the current analysis, no systematic uncertainty calculation is carried out however one will be conducted at a later stage of the search. It is expected that systematic uncertainties will reduce the sensitivity of the analysis to the sought BSM resonances by 15%–25% in the γ channel.

4.6.4 Prospective Paths

Two factors are destined to enhance the sensitivity to a Z' resonance considerably. The first is a mass-decorrelated tagger working as expected with a background rejection of the rate seen in FIG. 4.12. This would have improved the sensitivity to SM V resonances and accordingly to similar BSM extensions. Additionally, recently-developed reconstruction-level analysis objects named Unified Flow Object (UFO) jets are due to amplify the sensitivity to boosted massive propagators strongly. These jets optimise tracking information in conjunction with energy depositions from the HCAL. The UFO jets combine Particle Flow Object (PFO)-type jets with Track Calo Cluster (TCC)-type jets. The former exploits the better energy resolution at lower transverse momentum attained by the tracker whilst the latter uses the improved energy resolution at higher transverse momentum common in the calorimeter. Further, the UFO uses the higher spatial resolution in the tracker, not reached when TCC jets are used alone. A newly-designed mass-decorrelated tagger adapted to UFO jets seem promising, with a considerably higher background rejection than the LC topo jets enable. The improved tagging performance of the ANN applied on UFO jets are shown for W jets and compared to other jet reconstruction algorithms in FIG. 4.18.

The sensitivity calculations for the higher jet masses are subject to considerably wider distributions than the lower masses are, as predicted by the MC. If the limits are further attuned to the varying mass distribution width, an improved cross-section sensitivity can be validly expected, along with a finer coupling limit, rejecting couplings above $g_q = 0.2$ across the spectrum. These of course come on top of using the full data set at $\mathcal{L} = 139.0\,\text{fb}^{-1}$.

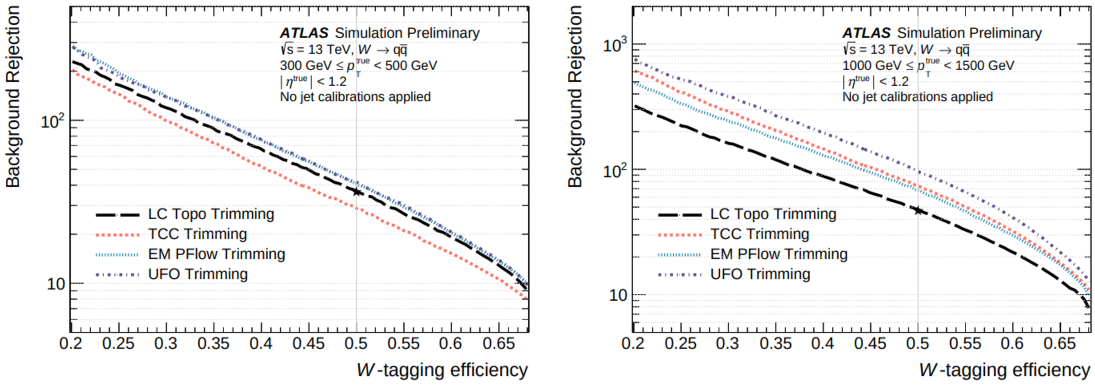

Figure 4.18: Receiver operating characteristic curve showing the background rejection as a function of signal efficiency for a tagger using the jet mass and D_2 for W boson jets at two transverse momentum regions. Improved taggings with different reconstruction algorithms, maximising with Unified Flow Object (UFO) jets are shown. (Left) $300\,\text{GeV} \leq p_{\text{T}} \leq 500\,\text{GeV}$. (Right) $1000\,\text{GeV} \leq p_{\text{T}} \leq 1500\,\text{GeV}$ [68].

Chapter 5

Conclusions

Exploring new routes to extend the SM is strongly braided with the quest to find evidence of DM in the context of particle physics. An accessible way to investigate such avenues relates to mediators, BSM particles which interact with SM and may well lead nearer to direct DM detection. In this, there are clear levers to be pushed forward, associated with the experimental techniques exercised in pp collisions-based experiments such as ATLAS. The high luminosity gained at the end of Run 2 of the LHC alone is due to attain increased sensitivities to such mediators throughout the invariant jet mass of their hadronic final state. On top of this, many methodological tools are redefined and some developed anew, aimed to enhance the sensitivity of the reconstructed jet mass to such resonances. The approach which uses an ISR object as a means to avoid the prevalent jet background is eminently helpful in enhancing the final-state resolution to new resonances. In this realm, different channels serve as routes to search the same phenomenon, both through a single large-R jet or two small-R jets final state whether the ISR is a jet or a photon.

In the boosted régime, the ISR photon channel enables a lower mass bound due to the smaller $p_{\mathrm{T}}^{\mathrm{HLT}}$ requirement as opposed to a small-R jet, as the trigger object. This is complemented by the large-R jet final state which contains the complete hadronic decay products of the recoiling resonance. This channel was firstly examined using SM boson decays as signal against the $J + \gamma$, mostly stemming from QCD events. The event selection was increasingly improved as to ensure optimal photon and jet candidates are chosen. After a 99%-efficient offline photon selection, every chosen jet was of transverse momentum greater or equal the reconstruction threshold at 200 GeV. Within object definition, the photon and jet were required to maintain a boost-invariant angle separation of at least 1.0, in order to refrain from collecting events with overlapping objects. Ensuring the diquark final state was contained within the $R = 1.0$ jet span, a $p^J \geq 2 \times m_J$ cut was applied. Removing low-mass activity was achieved by a $m_J \geq 50\,\mathrm{GeV}$ cut, which also reduced the Sudakov peak in the QCD background. The azimuthal separation required in order to reconstruct the resonance produced recoiling back-to-back against the ISR was chosen as to maximise the sensitivity to the SM V through the jet invariant mass. A side process substantiating those selections included truth-levels comparisons. Via responses of the jet mass between different simulated fiducial volumes, reconstruction-level cuts were confirmed such that the final event selection could be said to provide the best sensitivity available to the SM bosons through the final-state jets.

A control study treated the W and Z bosons as signals whilst using the general $J+\gamma$ simulation as background. The event selection described above was applied on the MC as well as on $36.0\,\mathrm{fb}^{-1}$ data. Subtraction of the background from the data and comparison to the signal showed a clear agreement with the signal simulation. The latter was subsequently assessed where the discovery potential yielded

$$\begin{aligned} &0.16 \times \sigma(q\bar{q} + \gamma \rightarrow W \rightarrow J) \\ &0.46 \times \sigma(q\bar{q} + \gamma \rightarrow Z \rightarrow J), \end{aligned} \tag{5.1}$$

which confirmed this analysis scheme may bring about progress in sensitivity to BSM resonances too. The efficacy of a mass-decorrelated tagger tool in the form of an ANN was assessed. Adding its discriminating score as a variable to the event selection seemed to show increased signal resolution in conjunction with background reduction. However, appraising the actual performance of the tool in terms of a ROC curve yielded that the background rejection at different signal efficiency values was insufficient: it did not contribute to the overall sensitivity obtained with the optimised event selection. This was the result even when pure signal was injected following the strictest cuts which included truth-level selections in addition to narrow jet mass windows about the W and Z resonances. The reason for the low tagging performance is yet unknown. Basic inspection of the jet substructure observables upon which it is based did not signify any clear trace. This point constitutes work-in-progress which shall continue in the near future.

The final part of the analysis estimated the sensitivity to a boosted Z' resonance decaying to a fat jet and produced in association with an ISR photon, based on the selections used for the SM control study. Following kinematic validation, the sensitivity was computed from the invariant jet mass using the same background in addition to the SM $Vqq + \gamma$ estimation. These were given in terms of cross section and coupling limits at 95% CL. In this, the MC was scaled to $139.0\,\mathrm{fb}^{-1}$ which this analysis is due to use. Already the results showed an improvement from the reference by rejecting lower cross sections and couplings, especially at $m_{Z'} = 100\,\mathrm{GeV}$, where the coupling limit yielded 0.156. However, the coupling rejection decreased considerably with growing mass values, which might be due to the widening $m_{Z'}$ distributions. The two main tools with which this sensitivity is to grow exceedingly finer are the UFO jets and its tagging by an especially-designed ANN. These object exemplify the current optimisation of combined calorimeter and tracker jet reconstruction. Nonetheless, resolving the ANN-based tagging of LC topo jets used in this thesis might well also improve the sensitivity to a Z' resonance across the mass spectrum even before the new jets are at hand. Further investigation of the SM-DM mediators with the power bestowed by superior jets and ML tagging tools through a combination of boosted and resolved channels induces hopefulness for the lever to be found farther away before this analysis is concluded.

Bibliography

[1] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, Perseus Press **1** (1995) ISBN: 0367320568.

[2] M. Sredniki, *Quantum Field Theory*, Cambridge University Press **1** (2008) ISBN: 0511813917

[3] M. D. Schwartz, *Quantum Field Theory and the Standard Model*, Cambridge University Press **1** (2013) ISBN: 1107034736

[4] M. E. Peskin, *Concepts of Elementary Particle Physics*, Oxford Master Series in Physics **26**, Oxford University Press (2019)

[5] C. P. Burgess and G. D. Moore, *The standard model : a primer*, Cambridge University Press **1** (2006).

[6] M. Thomson, *Modern Particle Physiscs*, Cambridge University Press **1** (2013).

[7] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (16): 508–09 (1964).

[8] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (9): 321–23 (1964).

[9] C. Patrignani *et. al.* (Particle Data Group), Chin. Phys. C, **40**, 100001 (2016) and 2017 update.

[10] M. Kobayashi and T. Maskawa, *CP-violation in the renormalizable theory of weak interaction*, Prog. Theor. Phys. **49** 2: 652–657 (1973).

[11] M.C. Gonzalez-Garcia and Y. Nir, *Neutrino Masses and Mixing: Evidence and Implications*, Rev. Mod. Phys. **75** 345–402 (2003)

[12] C. G. Callan, *Broken Scale Invariance in Scalar Field Theory*, Phys. Rev. D. **2** (8): 1541–1547.

[13] K. Symanzik, *Small distance behaviour in field theory and power counting*, Commun. Math. Phys. **18** (3): 227–246 (1970).

[14] J. Campbell, J. Huston and F. Krauss, *The Black Book of Quantum Chromodynamics: A Primer for the LHC Era* , Oxford University Press **1** (2018).

[15] A. D. Martin, W. J. Stirling, R. S. Thorne and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. C **63** 189–285 (2009).

[16] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and $e^+ e^-$ Annihilation by Perturbation Theory in Quantum Chromodynamics*, Sov. Phys. JETP **46** (1977).

[17] V. N. Gribov and L. N. Lipatov, *Deep inelastic e p scattering in perturbation theory*, Sov. J. Nucl. Phys. **15** (1972).

[18] G. Altarelli, R. D. Ball and S. Forte, *Resummation of singlet parton evolution at small x*, Nucl. Phys. B. **575** (2000).

[19] The Super-Kamiokande Collaboration, M. Koshiba, Y. Fukuda *et al.*, *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81**, 1562–1567 (1998)

[20] S. Profumo, *An Introduction To Particle Dark Matter*, **1** WSPC (2017) ISBN: 1786340011.

[21] S. Biondini, D. Boer and and R. Peeters, *The hierarchy problem and fine-tuning in a decoupling approach to multi-scale effective potentials*, arXiv:2012.08441 [hep-ph].

[22] J. Haffner of CERN, *CERN's Accelerator Complex*, public website, `http://cds.cern.ch/images/OPEN-PHO-ACCEL-2013-056-1` (2013), from URL on 15 August 2020.

[23] L. Evand and P. Bryant, *LHC Machine*, J. Instrum. **3** (2008).

[24] M. Gyr, *Polarities of LHC septum magnets MSI and MSD*, LHC-DC-ES-0001-30-20 (2014).

[25] M. Benedikt, *LHC Design Report: The LHC Injector Chain* **3** CERN Report, DOI: 10.5170/CERN-2004-003-V-3 (2004).

[26] The ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13\,TeV$ using the ATLAS detector at the LHC*, ATLAS Conf. Note, ATLAS-CONF-2019-021 (2019)

[27] J. Pequenao of CERN, *Computer-generated image of the whole ATLAS detector*, public website, `https://cds.cern.ch/images/CERN-GE-0803012-01` (2008), from URL on 21 August 2020.

[28] S. Haywood *et. al.*, *ATLAS inner detector, Technical Design Report*, **2**, CERN-LHCC-97-17 (1997).

[29] J. Pequenao of CERN, *Computer-generated image of the ATLAS inner detector*, public website, `https://cds.cern.ch/images/CERN-GE-0803014-01` (2008), from URL on 21 August 2020.

[30] The ATLAS Collaboration, *ATLAS liquid-argon calorimeter : Technical Design Report*, CERN-LHCC-96-041 (1996).

[31] The ATLAS Collaboration, *ATLAS tile calorimeter : Technical Design Report*, CERN-LHCC-96-042 (1996).

[32] J. Pequenao of CERN, *Computer Generated image of the ATLAS calorimeter*, public website, `https://cds.cern.ch/record/1095927` (2008), from URL on 22 August 2020.

[33] The ATLAS Collaboration, *ATLAS muon spectrometer : Technical Design Report*, CERN-LHCC-97-022 (1997).

[34] L. Pizzimento, *New generation of RPC detector for the ATLAS Muon detector upgrade*, ATL-MUON-SLIDE-2019-789 (2019).

[35] S. Marzani, G. Soyez and M. Spannowsky, *Looking inside jets: an introduction to jet substructure and boosted-object phenomenology*, Lect. Notes Phys. **958**, Springer (2019).

[36] The CMS Collaboration, public website, `http://cms.web.cern.ch/news/jets-cms-and-determination-their-energy-scale` (2012), from URL on 15 August 2020.

[37] G. Aad *et. al.* (The ATLAS Collaboration), *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. C **77** 490 (2017).

[38] J. M. Berlingen, *Search for new physics in $t\bar{t}$ final states with additional heavy-flavor jets with the ATLAS detector*, Universidad Autónoma de Barcelona, CERN-THESIS-2015-140 (2015).

[39] M. Cacciari and G. P. Salam, *The anti-k_{T} jet clustering algorithm*, JHEP **0804** 063 (2008).

[40] M. Asboud *et. al.* (The ATLAS Collaboration), *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at* $\sqrt{s} = 13\,TeV$ with the ATLAS detector, Phys. Rev. D **96** 072002 (2017).

[41] M. Cacciari, G. P. Salam and G. Soyez, *The catchment area of jets*, JHEP **04** 005 (2008).

[42] J. Thaler and K. van Tilburg, *Identifying Boosted Objects with N-subjettiness*, JHEP **03** 015 (2011).

[43] L. G. Almeida, S. J. Lee, G. Perez, I. Sung and J. Virzi, *Top Jets at the LHC*, Phys. Rev. D **79** 074012 (2009).

[44] ATLAS collaboration, *ATLAS measurements of the properties of jets for boosted particle searches*, Phys. Rev. D **86** 072006 (2012).

[45] J. Thaler and L.-T. Wang, *Strategies to Identify Boosted Tops*, JHEP **07** 092 (2008).

[46] D. Abercrombie, N. Akchurin, E. Akilli *et. al.*, *Dark Matter Benchmark Models for Early LHC Run-2 Searches: Report of the ATLAS/CMS Dark Matter Forum*, Phys. Dark Univ. **26** 100371 (2019).

[47] A. Alves, S. Profumo and F. S. Queiroz, *The dark Z′ portal: direct, indirect and collider searches*, JHEP **1404** 063 (2014).

[48] M. Chala, F. Kahlhoefer, M. McCullough, G. Nardini and K. Schmidt-Hoberg, *Constraining dark sectors with monojets and dijets*, JHEP **07** 089 (2015).

[49] M. Aabod, *et. al.* (The ATLAS Collaboration), *Search for low-mass resonances decaying into two jets and produced in association with a photon using pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Phys. Lett. B **795** 56 (2019).

[50] CMS Collaboration, *Search for low mass vector resonances decaying into quark-antiquark pairs in proton-proton collisions at* $\sqrt{s} = 13\,TeV$, J. High Energy Phys. **01** 097 (2017).

[51] J. Alison, S. Barnes, A. Boveia *et. al.* (within the ATLAS Collaboration), *Search for resolved dijet resonances produced in association with a photon*, ATLAS Internal Note, ATL-COM-PHYS-2018-228 (2018).

[52] C. Shimmin and D. Whiteson, *Boosting low-mass hadronic resonances*, Phys. Rev. D **94** 055001 (2016).

[53] M. Aabod, *et. al.* (The ATLAS Collaboration), *Search for light resonances decaying to boosted quark pairs and produced in association with a photon or a jet in proton-proton collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Phys. Lett. B **788** 316 (2019).

[54] E. Bothmann, G. S. Chahal, S. Höche *et. al.*, *Event Generation with Sherpa 2.2*, SciPost Phys. **7** 034 (2019).

[55] H. Lai, M. Guzzi, J. Huston *et. al.*, *New parton distributions for collider physics*, Phys. Rev. D **82** 074024 (2010).

[56] S. Agostinelli, J. Allison, K. Amako *et. al.*, *GEANT4 — a simulation toolkit*, Nucl. Instrum. Meth. A **506** 250-303 (2003).

[57] J. Alwall *et. al.*, *MadGraph 5: going beyond*, J. High Energy Phys. **06** 128 (2011).

[58] T. Sjostrand, S. Mrenna and P. Z. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** 852 (2008).

[59] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** 152–155 (2001).

[60] The ATLAS collaboration, *ATLAS Pythia 8 tunes to* 7 *TeV data*, ATL-PHYS-PUB-2014-021 (2014).

[61] The ATLAS collaboration, *Performance of the ATLAS trigger system in 2015*, Eur. Phys. J. C **77** 317 (2017).

[62] The ATLAS collaboration, *Selection of jets produced in 13 TeV proton-proton collisions with the ATLAS detector*, ATLAS Conference Note, ATLAS-CONF-2015-029 (2015).

[63] J. Shelton, *TASI Lectures on Jet Substructure*, arXiv:1302.0260v2 [hep-ph].

[64] M. Dasgupta, *Jet masses for high p_{T} QCD jets at all orders*, PoS **104** (2010).

[65] The ATLAS collaboration, *Performance of mass-decorrelated jet substructure observables for hadronic two-body decay tagging in ATLAS*, ATLAS-PHYS-PUB-2018-014 (2018).

[66] The ATLAS collaboration, *Identification of Hadronically-Decaying W Bosons and Top Quarks Using High-Level Features as Input to Boosted Decision Trees and Deep Neural Networks in ATLAS at $\sqrt{s}$ = 13 TeV*, ATLAS-PHYS-PUB-2017-004 (2017).

[67] I. Goodfellow, J. Pouget-Abadie, M. Mirza *et. al.*, *Generative Adversarial Nets*, Advances in Neural Information Processing Systems **27** 2672 (2014)

[68] The ATLAS collaboration, *Optimisation of large-radius jet reconstruction for the ATLAS detector in 13 TeV proton–proton collisions*, Submitted to EPJC, arXiv:2009.04986 [hep-ex] (2020).

אומדן הרגישות לחלקיקים חדשים הדועכים לצמד קווארקים אנרגטיים ונוצרים בהקשר לפוטון בניסוי אטלס

חיבור על מחקר

מוגש לשם מילוי חלקי
אחר הדרישות לקבלת תואר
מגיסטר למדעים בפיסיקה

רועי שימל ברנר

מוגש לסנט טכניון – מכון טכנולוגי לישראל
כסלו התשפ"א, חיפה, דצמבר 2020

המחקר נעשה בהנחיית ד"ר אנריקה קחמוביץ' בפקולטה לפיסיקה

אני מודה לטכניון על התמיכה הכספית הנדיבה בהשתלמותי

תקציר

התורה המבוססת ביותר של פיסיקת החלקיקים הינה הדגם התקני אשר כולל שלושה מתוך ארבעת ההידודים הידועים ביקום, לאמור אלקטרומגנטי, חלש וחזק. אע"פ שתחזיותיו של הדגם התקני נבחנות תדיר לרמת דיוק ראויה לציון, ישנן סיבות מהותיות להאמין כי הוא אינו מהווה תורה שלמה של הטבע. ביניהן נמנות אי־יכולתו של הדגם לכלול כבידה בתיאורו את הכוחות היסודיים, לחזות מסות נייטרינו שונות מאפס ולהעיד על קיומו של חומר אפל. הרחבות רבות הקרויות פיסיקה מעבר לדגם התקני מוצעות ונחקרות באמצעות ניסויים בנוגש ההדרונים הגדול, היכן שגבולות מוטלים עליהן. המחקר המפורט בחיבור זה מתייחס למתווכים בין חלקיקי דגם תקני לחלקיקי חומר אפל ובוחן שיטות חדשניות על־ידן עשויים מתווכים שכאלו להיחקר בעזרת גלאי אטלס.

הדגם התבחיני מציע בוזון ווקטורי לפטופובי Z' טעון תחת סימטריית כיול חדשה $U(1)'$ אשר מצומד לקווארקי דגם תקני ולפרמיוני דיראק של חומר אפל. הידודיו עם חומר אפל נדירים מאוד כך שחיפושים המערבים מתווכים שכאלו מתרכזים בראש ובראשונה בדעיכותיהם לחלקיקי דגם תקני הניתנים לזיהוי באמצעות גלאי נוגש ההדרונים הגדול. סוג זה של חלקיקים מעבר לדגם התקני הדועכים הדרונית עודנם חמקניים לחיפושים עדכניים הודות לרקע רב־סילוני מרובה המציף את תכנית ההדק. חיבור זה מאמץ שיטה חילופית.

שיטה זו עושה שימוש באירועים בהם פוטון קרינת מצב התחלתי הינו מושא ההדק ואילו החלקיק הנרתע נבנה מהמצב הסופי ההדרוני כסילון יחיד בעל מחוג רחב. סילונים אלו נבנים בשימוש הקלורימטר ההדרוני ומדידות האנרגיה שלהם בעלות דיוק של 3 אחוזים לערך. לאור הרתע של החלקיק בחיפוש כנגד הפוטון, החלקיק נתון להגברת לורנץ גבוהה. בוזונים הדועכים הדרונית עם הגברות לורנץ גבוהות מסווגים ע"י תתי־מבנים מדידים, כאשר במוקד עומדת תכונתם הדו־חודית האופיינית.

ע"י הדמיות מונטה קרלו, תבחיני בחירת האירועים ומושאי המחקר, סילונים ופוטונים, ממוטבים. בחירת האירועים כוללת בין היתר הפרדה זוויתית בין הפוטון לסילון הנרתע אשר מחד גיסא מדמה את התהליך הרצוי ומאידך גיסא שואפת להשגת רגישות מיטבית. הדמיית המונטה קרלו עושה שימוש בהדמיות רמות־אמת ולפיכך מתבצעת בשלושה מישורים, לאמור דרגת הפרטון, דרגת החלקיק ודרגת הגלאי. רמות שונות של הדמייה אלו מספקות תבחינים יסודיים יותר מרמת הגלאי על־פיהם האות מנוקה מהרקע והתהליך המהותי מטוהר. באופן זה מאומתת בחירת האירועים ברמת הגלאי המושווית לבסוף לנתוני הניסוי. בנוסף לכך, שימושיותו של כלי לימוד מכונה המבצע סיווג מעין זה עומדת למבחן. בתוך כך, רשת עצבית מלאכותית לומדת כיצד לסווג את האות המוגדר כסילון רחב־מחוג המהווה תוצאת דעיכה של החלקיק הווקטורי שבחיפוש. באמצעות 10 משתנים של תת־מבנה הסילון, כלי זה לומד את חתימת האות האופיינית ותו"כ מבצע למידה נגדית של מסת הסילון ע"מ להימנע מפיסול הרקע כצורת האות.

בהמשך לבחירה הממוטבת, הדמיות מונטה קרלו וחקר תחום מנוצלים ע"מ לאמוד את רגישותו של ניסוי אטלס ל־Z' באמצעות ערוץ זה. במסגרת החקר התחום, 36.0 פמטו בארן הופכי של נתוני גלאי אטלס מושווים להדמיות סילון רחב־מחוג בנוסף לפוטון וכן להידודים בהם אירע $q\bar{q}+\gamma \rightarrow W/Z \rightarrow q\bar{q}$. נאמד כי בתכנית זו צפונה היכולת לגלות תהליכים הדומים לאלו במסגרתם בוזון דגם תקני ווקטורי נוצר בהקשר לפוטון ודועך למצב סופי של צמד קווארקים לחתך פעולה קטן עד כדי 16 אחוזים ו־46 אחוזים מחתכי הפעולה לתוצרת W ו־Z, בהתאמה. לסיכום, באמצעות הדמיות מונטה קרלו של תהליכים בהם $q\bar{q} + \gamma \rightarrow Z' \rightarrow q\bar{q}$ עבור טווח מסה שבין 100 ל־220 ג'יגה אלקטרון וולט, הרגישות לחלקיק החדש נאמדת ומוצגת במונחי רווח בר־סמך של חתך הפעולה והצימוד.