

Graph Neural Network Flavour Tagging and Boosted Higgs Measurements at the LHC

Samuel John Van Stroud
University College London

Submitted to University College London in fulfilment
of the requirements for the award of the
degree of **Doctor of Philosophy**

August 10, 2022

Declaration

I, Samuel John Van Stroud confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Samuel Van Stroud

Abstract

Here some useful packages are demonstrated. In particular, the `hepunit` package which adds additional units to `SIUnit`. A variety of jet measurements are made using data collected during the first year of 7 TeV proton-proton collisions from the general-purpose ATLAS experiment at the LHC. no more than 300 words

Impact Statement

impact statement 500 words [link to ucl info](#)

Acknowledgements

Here is an example of how to declare commands for use in a single file that will not be needed elsewhere. Additionally, it serves to illustrate the chapter referencing system.

Perhaps you might want to point out that Peter Higgs provided help

This thesis was made in L^AT_EX 2_ε using the “hepthesis” class [\[1\]](#).

Brian Moser provided insight on the fit.

Contents

1	Theoretical Framework	9
1.1	The Standard Model	10
1.1.1	Quantum Electrodynamics	11
1.1.2	Quantum Chromodynamics	12
1.1.3	The Electroweak Sector	13
1.2	The Higgs Mechanism	14
1.2.1	Electroweak Symmetry Breaking	15
1.2.2	Fermionic Yukawa Coupling	18
1.2.3	Higgs Phenomenology	19
2	The Large Hadron Collider and the ATLAS Detector	21
2.1	Coordinate System & Collider Definitions	22
2.1.1	ATLAS Coordinate System	22
2.1.2	Collider Definitions	23
2.2	The ATLAS Detector	25
2.2.1	The Inner Detector	25
2.2.2	Calorimeters	29
2.2.3	Muon Spectrometer	30
2.2.4	The Trigger	31
2.3	Reconstructed Physics Objects	32
2.3.1	Tracks	32
2.3.2	Vertices	32
2.3.3	Jets	33
2.3.4	Leptons	34
3	Investigating Tracking Improvements	35
3.1	b -hadron Reconstruction	35
3.1.1	b -hadron Decay Topology	35

3.1.2	<i>b</i> -hadron Decay Track Reconstruction	37
3.2	Pseudotracks and Ideal Tracks	40
3.3	Investigating Improvements for High p_T B Tracking	40
3.3.1	Looser Track Cuts & Track Refit Procedure	40
3.3.2	Region of Interest Optimisation	41
3.3.3	Fit Quality as a Discriminant for Wrong Hits	42
3.3.4	Conclusion	42
3.4	Global χ^2 Fitter Outlier Removal	44
3.4.1	Cut Optimisation	45
3.5	Tracking software validation	46
4	Track Classification MVA	47
4.1	Machine Learning Background	47
4.2	Track Truth Origin Labelling	47
4.3	Fake Track Identification Tool	47
4.3.1	<i>b</i> -hadron Decay Track Identification Tool	47
4.4	General Track Origin Classifier Tool	47
4.5	Conclusion	48
5	Graph Neural Network Flavour Tagger	49
5.1	Motivation	50
5.2	Graph Neural Network Theory	53
5.3	Experiemental Setup	53
5.3.1	Datasets	53
5.4	Model Architecture	54
5.4.1	Model Inputs	54
5.4.2	Auxiliary Training Objectives	56
5.4.3	Architecture	57
5.4.4	Training	61
5.5	Results	62
5.5.1	<i>b</i> -tagging Performance	63
5.5.2	<i>c</i> -tagging Performance	67
5.5.3	Ablations	69
5.5.4	Inclusion of Low-Level Vertexing Algorithms	69
5.5.5	Vertexing Performance	73
5.5.6	Track Classification Performance	76

5.6	Conclusion	79
5.7	Extensions	79
5.7.1	Looser Track Selection	79
6	VHbb Boosted Analysis	80
6.1	Overview	81
6.2	Introduction	82
6.3	Modelling Work	83
6.3.1	Background	83
6.3.2	Vector Boson + Jets Modelling	87
6.3.3	Diboson Modelling	89
6.4	Fit Studies	89
6.4.1	Fit Model	89
6.5	Conclusion	90
7	Conclusion	91
A	Combining Multiple Triggers	92
	Bibliography	93
	List of figures	101
	List of tables	108

Chapter 1

Theoretical Framework

- Introduce sm
- brief history
- current areas of study
- Reference relevance to rest of thesis (studying hbb)

The Standard Model (SM) of particle physics is the theory describing all known elementary particles and their interactions via three of the four fundamental forces. Developed by merging the successful theories of quantum mechanics and relativity in the second half of the 20th century, the SM's position today at the centre of our understanding of the nature of the universe is firmly established by an unparalleled level of agreement between the predictions from the model and experimental results [2, 3].

The SM has predicted the discovery of the top and bottom quarks [4–6], the W and Z bosons [7], and the tau neutrino [8]. The last missing piece of the SM to be discovered was the Higgs boson, first theorised in the 1960s [9–11], and eventually observed at the LHC in 2012 [12, 13]. After its discovery, much ongoing work has been carried out performing detailed measurements of its mass and interactions with other particles.

This thesis describes research to improve Higgs analysis to b quarks, by improving algorithm methods...

1.1 The Standard Model

The SM is formulated in the language of Quantum Field Theory (QFT). In this framework, particles are localised excitations of corresponding quantum fields, which are operator-valued distribution across spacetime.

Central to QFT is the Lagrangian density which describes the kinematics and dynamics of a field. Observations of conserved quantities are linked, via Noether's theorem, to symmetries which are expressed by the Lagrangian. Alongside Global Poincaré symmetry, the SM Lagrangian observes a local non-Abelian $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ gauge symmetry. Gauge symmetries leave observable properties of the system unchanged when certain gauge transformations are applied to the fields. The full Lagrangian of the SM can be broken up into distinct terms corresponding to the different sectors. An overview of each sector is given in the following chapters.

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{EW}} + \mathcal{L}_{\text{QCD}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}} \quad (1.1)$$

The SM provides a mathematical description of how the four fundamental forces interact with the matter content of the universe. The SM contains 12 spin- $1/2$ fermions, listed in Table 1.1, and 5 bosons listed in Table 1.2.

	Leptons			Quarks		
Generation	Flavour	Mass [MeV]	Charge [e]	Flavour	Mass [MeV]	Charge [e]
First	e	0.511	-1	u	2.16	$2/3$
	ν_e	$< 1.1 \times 10^{-6}$	0	d	4.67	$-1/3$
Second	μ	105.7	-1	c	1.27×10^3	$2/3$
	ν_μ	< 0.19	0	s	93.4	$-1/3$
Third	τ	1776.9	-1	t	173×10^3	$2/3$
	ν_τ	< 18.2	0	b	4.18×10^3	$-1/3$

Table 1.1: The half-integer spin fermions of the SM [14]. Three generations of particles are present. Also present (unlisted) are the antiparticles, which are identical to the particles up to a reversed charge sign.

Name	Symbol	Mass [GeV]	Charge [e]	Spin
Photon	γ	$< 1 \times 10^{-27}$	$< 1 \times 10^{-46}$	1
Weak boson	W^\pm	80.377 ± 0.012	± 1	1
Weak boson	Z	91.1876 ± 0.0021	0	1
Gluon	g	0	0	1
Higgs	H	125.25 ± 0.17	0	0

Table 1.2: The integer spin bosons of the SM [14]. The photon, weak bosons and gluons are gauge bosons arising from gauge symmetries, and carry the four fundamental forces of the SM.

1.1.1 Quantum Electrodynamics

Consider a Dirac spinor field $\psi = \psi(x)$ and its adjoint $\bar{\psi} = \psi^\dagger \gamma^0$, where ψ^\dagger denotes the Hermitian conjugate of ψ . The field ψ describes fermionic spin-1/2 particle, for example an electron. The Dirac Lagrangian density is

$$\mathcal{L}_{\text{Dirac}} = \bar{\psi}(i\cancel{\partial} - m)\psi, \quad (1.2)$$

where $\cancel{\partial} = \gamma^\mu \partial_\mu$ denotes the contraction with the Dirac gamma matrices γ^μ (summation over up-down pairs of indices is assumed). Application of the Euler-Lagrange equation on Eq. (1.2) yields the Dirac equation

$$(i\cancel{\partial} - m)\psi = 0. \quad (1.3)$$

Suppose some fundamental symmetry that requires invariance under a local $U(1)$ gauge transformation

$$\psi \rightarrow \psi' = \psi e^{-iq\alpha(x)}, \quad (1.4)$$

where α varies over every spacetime point x . Under this transformation, the Dirac equation transforms as

$$(i\cancel{\partial} - m)\psi e^{-iq\alpha(x)} + q\cancel{\partial}\alpha(x)\psi e^{-iq\alpha(x)} = 0. \quad (1.5)$$

For the Dirac equation to remain invariant under the transformation in Eq. (1.4), a new field A_μ which transforms as $A_\mu \rightarrow A'_\mu = A_\mu + \partial_\mu \alpha(x)$ must be added. The

mention
EoMs?

check consistency with
 $Y=1/2$ used
later

transformed interaction term

$$-q\cancel{A}\psi \rightarrow -q\cancel{A}\psi e^{-iq\alpha(x)} - q\cancel{\partial}\alpha(x)\psi e^{-iq\alpha(x)} \quad (1.6)$$

will then cancel the asymmetric term in Eq. (1.5) as required. The $U(1)$ invariant Lagrangian can therefore be constructed by adding an interaction between ψ and A_μ to Eq. (1.2). The kinetic term for the new field A_μ is also added in terms of $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, which is trivially invariant under the transformation in Eq. (1.4). The interaction term is absorbed into the covariant derivative $D_\mu = \partial_\mu + iqA_\mu$, thus named as it transforms in the same way as the field ψ . This yields the QED Lagrangian

$$\mathcal{L}_{\text{QED}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\not{D} - m)\psi, \quad (1.7)$$

A quadratic term $A_\mu A^\mu$ is not invariant and therefore the field A_μ must be massless. Requiring invariance under local $U(1)$ gauge transformations necessitated the addition of a new field A_μ , corresponding to photons, which interact with charged matter.

improve interpretation

1.1.2 Quantum Chromodynamics

Quantum Chromodynamics (QCD) is the study of quarks, gluons and their interactions. Quarks and gluons carry colour charge, which comes in three kinds, called red, green and blue. While the $U(1)$ symmetry group in Section 1.1.1 was Abelian, the QCD Lagrangian is specified by requiring invariance under transformations from the non-Abelian $SU(3)$ group, making it a Yang-Mills theory [15] which requires the addition of self-interacting gauge fields. The infinitesimal $SU(3)$ group generators are given by $T_a = \lambda_a/2$, where λ_a are the eight Gell-Mann matrices. These span the space of infinitesimal group transformations and do not commute with each other, instead satisfying the commutation relation

$$[T_a, T_b] = if_{abc}T_c, \quad (1.8)$$

where f_{abc} are the group's structure constants. Consider the six quark fields $q_k = q_k(x)$. Each flavour of quark q_k transforms in the fundamental triplet representation, where

each component of the triplet corresponds to a different value of the colour quantum number. $G_{\mu\nu}^a$ are the eight gluon field strength tensors, one for each generator T_a , defined as

$$G_{\mu\nu}^a = \partial_\mu A_\nu - \partial_\nu A_\mu - g_s f^{abc} A_\mu^b A_\nu^c, \quad (1.9)$$

where A_μ^a are the gluon fields and g_s is the strong coupling constant. The covariant derivative is written as

$$D_\mu = \partial_\mu + ig_s T_a A_\mu^a. \quad (1.10)$$

The full QCD Lagrangian is then given by

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu} + q_k (i \not{D} - m_k) q_k. \quad (1.11)$$

Cubic and quartic terms of the gauge fields A_μ^a appear in the Lagrangian, leading to the gluon's self interaction.

The QCD coupling constant g_s varies, or “runs”, with energy. At lower energy scales (and corresponding larger distance scales) the interaction is strong. This leads to quark confinement, whereby an attempt to isolate individual colour-charged quarks requires so much energy that additional quark-antiquark are produced. At higher energy scales (and corresponding smaller distance scales), asymptotic freedom occurs as the interactions become weaker, allowing perturbative calculations to be performed. Hadrons are bound states of quarks. They are invariant under $SU(3)$ gauge transformations (i.e. are colour-charge neutral, or colourless).

1.1.3 The Electroweak Sector

The weak and electromagnetic forces are unified in the Glashow-Weinberg-Salam (GWS) model of electroweak interaction [16–18]. The Lagrangian is specified by requiring invariance under the symmetry group $SU(2)_L \otimes U(1)_Y$, as motivated by a large amount of experimental data. Here, $SU(2)_L$ is referred to as weak isospin and $U(1)_Y$ as weak hypercharge. (the product of the isospin and hypercharge groups.).

The generators of $SU(2)_L$ are $T_a = \sigma_a/2$, where σ_a are the three Pauli spin matrices which satisfy the commutation relation

$$[T_a, T_b] = i\varepsilon_{abc}T_c. \quad (1.12)$$

The generator of $U(1)_Y$ is $Y = 1/2$. Each generator corresponds to a gauge field, which, after symmetry breaking (discussed in Section 1.2), give rise to the massive vector bosons, W^\pm and Z , and the massless photon. The massive vector bosons are the carriers of the weak force, and are unique to the weak sector. Due to the mass of the force carriers, the weak force has a short range and so it appears weak even though its intrinsic strength is comparable to that of QED. The gauge bosons acquire mass through the Higgs mechanism, discussed in Section 1.2.

why not
mass by
hand?

The weak force violates parity conservation [19–21], i.e. invariance under parity transformations (mirror reflections). Only left handed fermions participate in the weak interaction. There are no right handed neutrinos in the standard model. Furthermore, the weak sector exhibits CP violation. CP violation is one of the three necessary Sakharov conditions required to produce baryon asymmetry in the universe. Since the SM alone does not appear to have enough CP violation to generate the cosmologically observed matter-antimatter asymmetry, looking for signs of more experimental CP violation is considered to be a promising way to discover new physics.

add parity
operator?

The charge operator can be written

put some-
where else?

$$Q = T_3 + Y. \quad (1.13)$$

1.2 The Higgs Mechanism

The Brout-Englert-Higgs mechanism (henceforth “Higgs mechanism”) is the mechanism through which the fundamental particles of the SM acquire mass [9–11]. Experimentally it was known that the weak force had a weak effective strength, which was suggestive of a massive mediating gauge particle. However, directly adding mass to the weak gauge bosons violates the non-Abelian symmetry of the SM. Instead, the gauge bosons can gain mass through the interaction with a scalar

Higgs field which results from the spontaneous breakdown of symmetry as discussed in Section 1.2.1. Similarly, the Higgs mechanism gives mass to the fermions, as discussed in Section 1.2.2.

1.2.1 Electroweak Symmetry Breaking

Spontaneous symmetry breaking (SSB) is a key part of the Higgs mechanism. It is the transition of a physical system from a state of manifest symmetry to a state of hidden, or “broken”, symmetry. In particular, this applies to physical systems where the Lagrangian observes some symmetry, but the lowest energy vacuum states do not exhibit that same symmetry. In other words, the symmetry is broken for perturbations around the vacuum state.

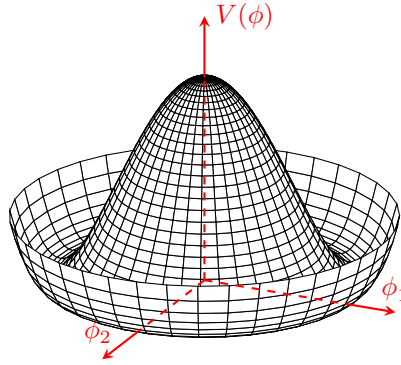


Figure 1.1: The Higgs potential $V(\phi)$ of the complex scalar field singlet $\phi = \phi_1 + i\phi_2$, with a choice of $\mu^2 < 0$ leading to a continuous degeneracy in the true vacuum states. A false vacuum is present at the origin. The SM Higgs mechanism relies on a complex scalar doublet with a corresponding 5-dimensional potential.

Consider gauge fields from the local $SU(2)_L \otimes U(1)_Y$ symmetry group discussed in Section 1.1.3 coupled to a complex scalar field $\phi = \phi(x)$. The scalar field ϕ transforms as a weak isospin doublet. Omitting the kinetic term of the gauge fields, and writing $\phi^2 \equiv \phi^\dagger \phi$, the Lagrangian is

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \phi)^\dagger (D^\mu \phi) - [\mu^2 \phi^2 + \lambda \phi^4], \quad (1.14)$$

where the covariant derivative is given by

$$D_\mu = \partial_\mu + ig A_\mu^a T^a + ig' B_\mu, \quad (1.15)$$

and T^a are the generators of $SU(2)$. The potential term $V(\phi)$ is made up of a quadratic and quartic term in the scalar field ϕ , which contain an arbitrary parameter, respectively λ and μ . The quartic term gives the field self-interaction, and cannot be negative as this would lead to a potential that was unbounded from below. The quadratic term can be positive or negative. In the case where the quadratic term is positive, it is interpreted as a mass term for the scalar field. By choosing $\mu^2 < 0$ the field becomes unphysical due to its negative mass. In order to obtain a physical interpretation of the Lagrangian in Eq. (1.14) for the case where $\mu^2 < 0$, the field ϕ is expanded around the vacuum state. The vacuum expectation value (VEV) is the expected value of the field ϕ which minimises the potential $V(\phi)$ (equivalently the expected value of the field operator ϕ when the system is in a vacuum state, $|\langle\phi\rangle_0|^2 \equiv |\langle 0|\phi|0\rangle|^2 \equiv \phi_0^2$). Minimising the potential gives a VEV of

$$\phi_0^2 = -\mu^2/\lambda = v^2. \quad (1.16)$$

Due to the shape of the potential in Fig. 1.1, there is degeneracy in the direction that the complex doublet ϕ points. As all the different vacuum states minimise the potential and therefore yield identical physics, so we can arbitrarily choose the state to lie along the second component of the doublet. Application of Eq. (1.13) shows this choice is manifestly invariant under, unbroken subgroup under which the ground state is invariant is $U(1)_Q$, generated by Q .

Adding the particle content back to the theory by expanding the field around the vacuum state, and making a transformation to the unitary gauge to remove unphysical would-be Nambu-Goldstone modes (which arise in the context of global symmetries [22, 23]) yields

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix}, \quad (1.17)$$

Where H is a real scalar field, the “true vacuum” Higgs field. Substituting this into Eq. (1.14) and identifying physical fields as quadratic terms of linear combinations, one can write the physical fields W_μ^\pm , Z_μ and A_μ in terms of the original fields A_μ^a

and B_μ . This gives

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(A_\mu^1 \mp iA_\mu^2) \quad \begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ A_\mu^3 \end{pmatrix}. \quad (1.18)$$

where θ_W is the weak mixing angle defined by

$$\cos \theta_W = \frac{g}{\sqrt{g^2 + g'^2}}. \quad (1.19)$$

The corresponding masses of the now massive vector bosons can be read off as

$$m_W = \frac{1}{2}gv \quad m_Z = \frac{m_W}{\cos \theta_W}, \quad (1.20)$$

while the photon remains massless. The Higgs mass is $m_H = v\sqrt{\lambda} = \mu$.

This is the Higgs mechanism. It maintains the renormalisability and unitarity of the SM whilst allowing the weak vector bosons to acquire mass. In summary, an unphysical complex scalar field ϕ with a nonzero VEV leads to spontaneous symmetry breaking from the full SM symmetry group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ to $SU(3)_C \otimes U(1)_\gamma$. Due to the non-Abelian symmetry breaking, would-be massless Nambu-Goldstone modes, which arise after expansion around the true vacuum state, are exactly cancelled out by making a local gauge transformation to the unitary gauge, and instead are absorbed by the vector bosons, allowing them to acquire mass.

This sector of the SM contains four fundamental parameters that must be taken from experiment. These can be specified by the Lagrangian parameters g , g' , v and λ or the physically measurable parameters m_Z , $\sin \theta_W$, m_H and e . In the local neighbourhood around the true vacuum, the macroscopic symmetry of the system is not realised, and therefore the physical particles do not obey the original symmetry. However, information about the symmetry is retained through some additional constraints on the particles and their interactions. Prior to symmetry breaking, the potential contained two terms and two constants. After symmetry breaking we have three terms but still only two constants that relate these terms. This is the vestige of the original symmetry.

Spontaneous symmetry breaking has taken us from $SU(2)_L \times U(1)_Y \rightarrow U(1)_\gamma$. Three broken generators of the symmetry group $U(2) = SU(2) \otimes U(1)$ have been absorbed

into the definition of the physical weak vector bosons, giving them mass. The same methodology can be used to generate the fermion masses, as shown in the next section.

1.2.2 Fermionic Yukawa Coupling

Adding the masses of the fermions by hand breaks the gauge invariance of the theory. Instead, we can use a Yukawa coupling of the fermion fields with the Higgs field in order to generate mass terms after spontaneous electroweak symmetry breakdown [17]. In this way, the fermion masses are determined by the respective couplings to the Higgs field, and the VEV of the Higgs field itself, which sets the basic mass scale of the theory.

If the $T = 1/2$ doublet ϕ has weak hypercharge $Y = 1/2$, then there is a unique renormalisable and gauge invariant coupling given by

$$\mathcal{L}_{\text{Yukawa}} = -G_f(\bar{\psi}_L\phi\psi_R + \bar{\psi}_R\phi\psi_L) \quad (1.21)$$

The Yukawa terms mix quarks of different generations. The indices on the Λ matrix stand for the generations, and the elements of Λ are complex valued.

The physical particles are detected in their mass eigenstates (but transform under $SU(2)$ according to their weak eigenstates). In other words, the physical particles diagonalise the mass matrix. It therefore makes sense to work in a basis that diagonalises the mass matrix. These are the mass eigenstates, as opposed to the weak eigenstates. Quark mixing can be expressed using the CKM matrix.

$$\begin{pmatrix} \tilde{d} \\ \tilde{s} \\ \tilde{b} \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad (1.22)$$

The CKM matrix is a result of the transformation between mass and gauge eigenstates of the quarks. It is a unitary matrix which contains information on the strength of flavour-changing weak decays. The \tilde{q} represent gauge or weak eigenstates, whereas q represent mass eigenstates. The size of the elements $|V_{pq}|^2$ measures the probability of a transition between states p and n .

We start with the most general form of the coupling of the $T = 1/2$, $Y = 1/2$ Higgs field ϕ to the quark fields. From ϕ we can form a conjugate field ϕ^c which transforms under $SU(2) \otimes U(1)_Y$ as a doublet with $Y = -1/2$,

$$\phi^c = i\sigma_2\phi^* = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \phi_1^\dagger \\ \phi_2^\dagger \end{pmatrix} = \begin{pmatrix} \phi_2^\dagger \\ -\phi_1^\dagger \end{pmatrix}. \quad (1.23)$$

We can see (???) that ϕ^c transforms under $SU(2)$ in the same way as ϕ . The ground state of this conjugate field is

$$\phi_0^c = \frac{1}{\sqrt{2}} \begin{pmatrix} v \\ 0 \end{pmatrix}. \quad (1.24)$$

The general gauge invariant expression for the coupling of the Higgs field to the quark fields is then of the form (???)

$$\mathcal{L}_{\phi,\text{quarks}} = -\sqrt{2} [\bar{L}_N f_{NM}^- \phi R_{M-} + \bar{L}_N f_{NM}^+ \phi^c R_{M+} + \text{h.c.}], \quad (1.25)$$

1.2.3 Higgs Phenomenology

- Production mechanisms
- Decay modes
- Relevance to analysis Vhbb

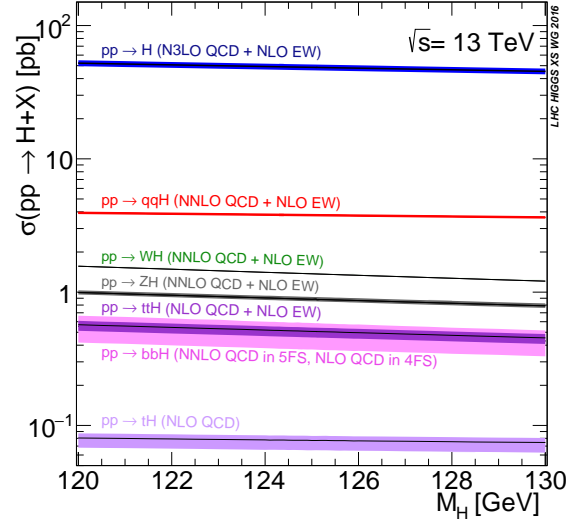


Figure 1.2: Higgs production cross sections at $\sqrt{s} = 13$ TeV [24].

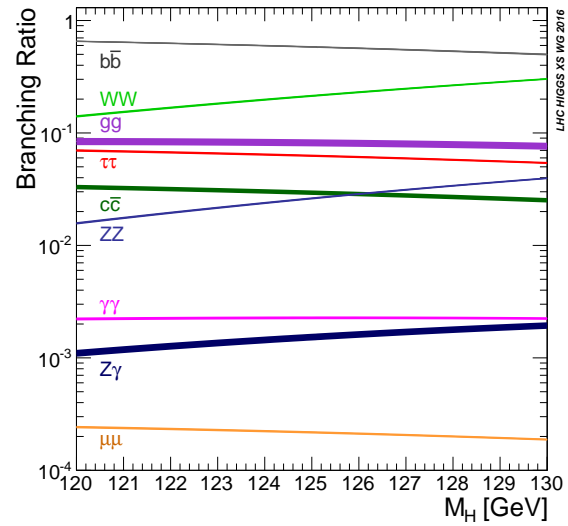


Figure 1.3: Higgs branch ratios at $\sqrt{s} = 13$ TeV [24].

Chapter 2

The Large Hadron Collider and the ATLAS Detector

The Large Hadron Collider (LHC) at CERN has extended the frontiers of particle physics through its unprecedented energy and luminosity. The LHC accelerates protons around a 27 km ring until they are travelling just 3 m s^{-1} slower than the speed of light, at which point they are made to collide. The protons travel round the ring 11 thousand times per second in two concentric beams, which are guided by superconducting magnets cooled using liquid helium to -271.3°C (1.9 K). The beams travel in opposite directions and are crossed at four locations so that collisions between protons can take place. Around these collision points four specialised detectors, ALICE [25], CMS [26], LHCb [27] and ATLAS [28], are located to capture information about the products of the collisions.

The LHC is operated in *runs* during which beams of protons are actively being circulated and collided. Between runs there are periods of shutdown while the accelerator and detector machinery is maintained and upgraded. In 2010, the LHC collided proton bunches, each containing more than 10^{11} particles, 20 million times per second, providing 7 TeV proton-proton collisions at instantaneous luminosities of up to $2.1 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$. Run 2, which began in 2015, increased the proton-proton collision energy to 13 TeV. The bunch spacing was also reduced, leading to a collision rate of 40 MHz. Over the course of Run 2 a total integrated luminosity of 146.9 fb^{-1} was recorded. 2022 marked the beginning of Run 3 which, with a higher center of mass energy and peak luminosity, is expected to culminate in the approximate doubling of the dataset size.

Period	Year	\sqrt{s} [TeV]	$\langle\mu\rangle$	Bunch spacing [ns]	Luminosity [$\text{cm}^{-2} \text{s}^{-1}$]
Run 1	2011–2012	7–8	18	50	8×10^{33}
Run 2	2015–2018	13	34	25	$1\text{--}2 \times 10^{34}$
Run 3	2022–2025	13.6	50	25	2×10^{34}

Table 2.1: Overview of the different LHC runs [29,30]. The average number of interactions per bunch-crossing is denoted as $\langle\mu\rangle$ (see Section 2.1.2). Numbers for Run 3 are preliminary and are only provided to give an indication of expected performance.

2.1 Coordinate System & Collider Definitions

2.1.1 ATLAS Coordinate System

ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point in the centre of the detector and the z -axis along the beam pipe. The x -axis points from the interaction point to the centre of the LHC ring, and the y -axis points upwards. Cylindrical coordinates (r, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the z -axis. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln \tan(\theta/2)$. Angular distance is measured in units of $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

this is from
pubcom

The transverse plane lies in x - y while the longitudinal plane lies along the z -axis.

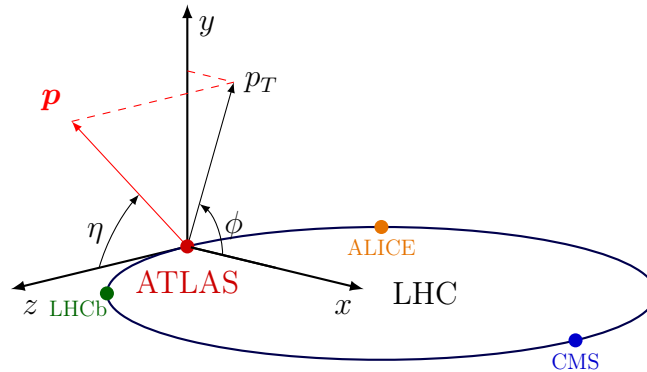


Figure 2.1: The coordinate system used at the ATLAS detector, showing the locations of the four main experiments located at various points around the LHC. Reproduced from Ref. [31].

The polar angle θ is commonly specified in terms of the pseudorapidity η , defined as

$$\eta = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right]. \quad (2.1)$$

Differences in η are invariant under Lorentz boosts.

The transverse momentum p_T

$$p_T = \sqrt{p_x^2 + p_y^2}. \quad (2.2)$$

Impact parameters d_0 and z_0 are defined by

2.1.2 Collider Definitions

Luminosity

The luminosity is defined by

xy full def

$$L = \frac{1}{\sigma} \frac{dN}{dt}, \quad (2.3)$$

where N is the number of individual proton-proton collisions and σ is the cross-sectional area of the beams as they cross.

Pile-up

At the centre of the ATLAS detector, bunches of more than 10^{11} protons meet head on. Each bunch-crossing is called an *event*. There is generally at most one hard proton-proton scatter per event. Additional interactions are typically relatively soft and are known as *pile-up*. Pile-up complicates the reconstruction of the hard scatter event as results of the interactions of different proton-proton interactions have to be separated. Pile-up from interactions within the same bunch-crossing is known as *in-time* pile-up while residual signatures from other bunch-crossings is known as *out-of-time* pile-up. The number of pile-up interactions is denoted μ , which is often given as a time-averaged value $\langle \mu \rangle$. The average number of pile-up interactions for different years during Run 2 is given in Fig. 2.3.

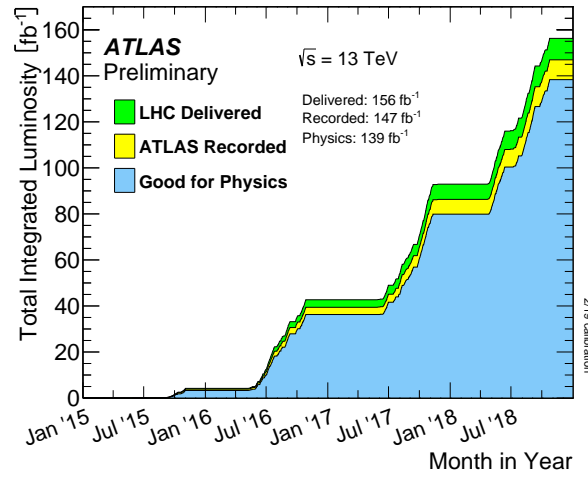


Figure 2.2: Delivered, recorded, and usable integrated luminosity as a function of time during Run 2 [30].

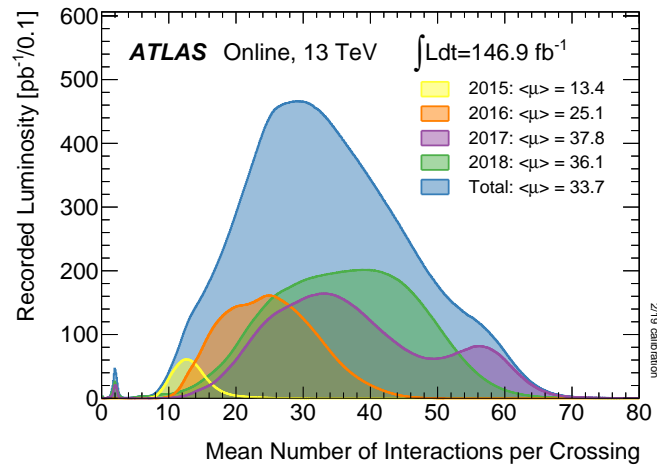


Figure 2.3: Average pile-up profiles measured by ATLAS during Run 2 [30]. During Run 3, even higher levels of pile-up are expected.

2.2 The ATLAS Detector

The ATLAS¹ detector at the LHC covers nearly the entire solid angle around the collision point. It consists of an inner tracking detector surrounded by a thin superconducting solenoid, electromagnetic and hadronic calorimeters, and a muon spectrometer incorporating three large superconducting air-core toroidal magnets.

The detector is made up of several specialised sub-detectors as shown in Fig. 2.4. In this section a condensed overview of each sub-detector is given, in order of increasing radial distance from the point of collision. A more complete picture can be found in Ref. [28], or in the technical design reports (TDRs) of the individual sub-detectors.

suggestion:
make the
first sen-
tence more
of a intu-
itive under-
standing,
and then go
into barel
eta encap
jargon

okay this
is the stan-
dard pub-
com text?

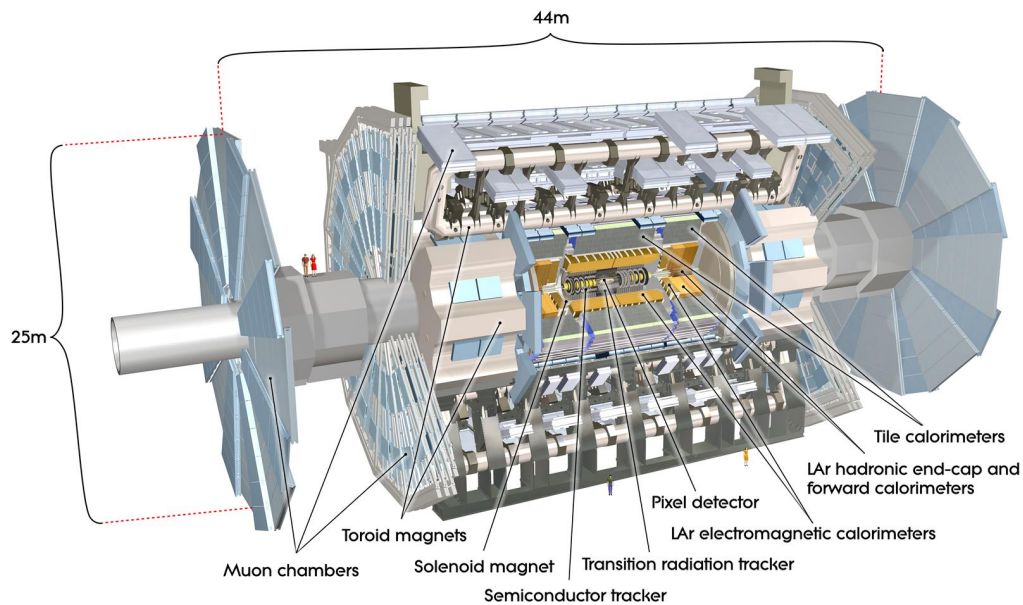


Figure 2.4: A 3D model of the entire ATLAS detector. Cutouts through the detector

2.2.1 The Inner Detector

The inner-detector system (ID) provides high-resolution charged-particle trajectory tracking in the range $|\eta| < 2.5$. The ID is immersed in a 2 T axial magnetic field, produced by a superconducting solenoidal magnet, which enables the measurement

¹A Toroidal Lhc ApparatuS.

of particle transverse momentum and charge². After Run 3, the ID will be replaced by the ITk [32].

The inner detector is made up of several sub-systems, shown in Figs. 2.5 and 2.6. The high-granularity silicon pixel detector covers the vertex region and typically provides four spacepoint measurements per track. It is followed by the silicon microstrip tracker (SCT), which usually provides a further four spacepoint measurements per track. These silicon detectors are complemented by the Transition Radiation Tracker (TRT), which enables radially extended track reconstruction up to $|\eta| = 2.0$.

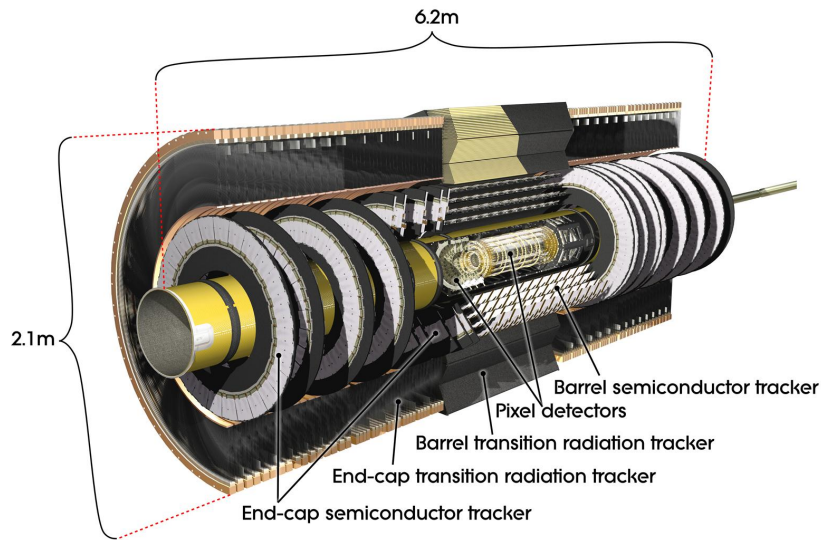


Figure 2.5: A 3D model of the ATLAS ID showing the barrel layers and end-cap disks.

Pixel Detector

The silicon pixel detector is comprised of four cylindrical barrels at increasing radii from the beamline, and four disks on each side. The innermost barrel layer is the insertable B-layer (IBL), which was installed before Run 2 [33, 34]. The specification of the pixel detector determines the impact parameter resolution and the ability to reconstruct primary and secondary vertices. The detector is required to have a high granularity (i.e. resolution) to maintain the low occupancy required to resolve nearby particles. Individual pixels are $50\,\mu\text{m}$ in the transverse direction $R\phi$ and $400\,\mu\text{m}$ in

²Reconstructed charged particles are assumed to have a charge of ± 1 .

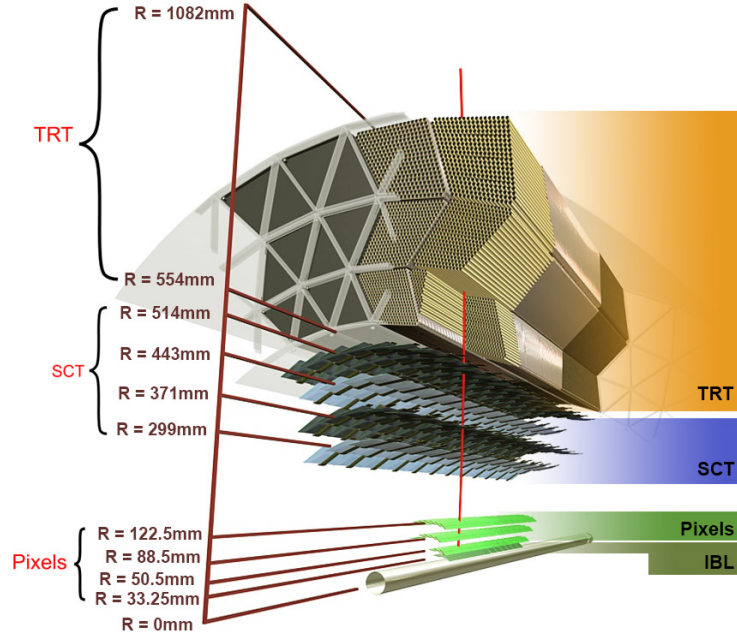


Figure 2.6: A cross-sectional view of the ATLAS ID, with the radii of the different barrel layers shown.

the longitudinal z direction ($250\text{ }\mu\text{m}$ for the IBL). Cluster positions have a resolution of approximately $10\text{ }\mu\text{m}$ in $R\phi$ and $100\text{ }\mu\text{m}$ in z .

Semi-Conductor Tracker (SCT)

The SCT is made up of four concentric barrel layers in the central region, and nine disks in each end-cap. Each layer is itself made of a pair of silicon microstrip layers, with a small stereo angle (20 mrad) between the two layers enabling the z -coordinate measurement from a pair of strip measurements. The SCT typically provides four precision spacepoint measurements (eight strip measurements) per track in the barrel region. These have intrinsic uncertainties of $17\text{ }\mu\text{m}$ in the transverse direction $R\phi$, and $580\text{ }\mu\text{m}$ in the longitudinal direction z . The measurements provide a key contribution to the measurement of charged-particle momentum and impact parameter, along with vertex position. Charge-particle tracks can be distinguished if separated by more than $\sim 200\text{ }\mu\text{m}$. Hits are registered as binary signals if the pulse height in a channel exceeds a certain threshold.

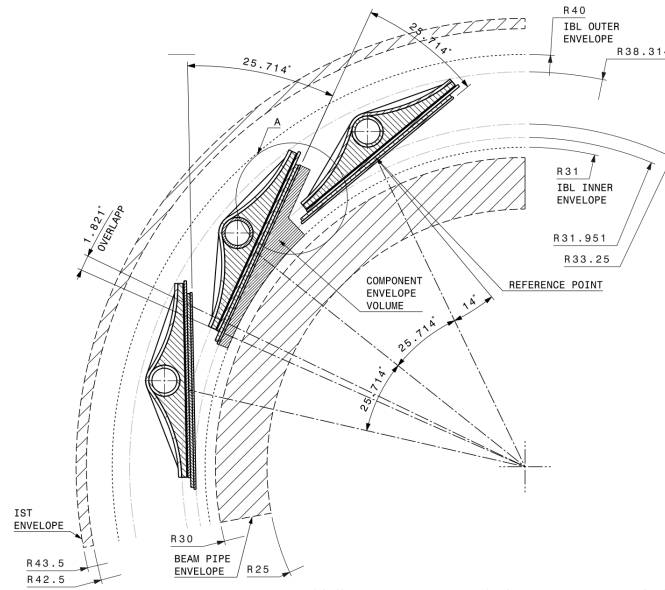


Figure 2.7: A cross-sectional view of the ATLAS IBL.

Transition Radiation Tracker (TRT)

The TRT is a straw-tube tracker which complements the higher-resolution silicon-based tracks by offering a larger number of hits per track (typically around 30) and a long lever arm, which aids the accurate measurement of particle momentum. It is made up of approximately 300 000 drift tubes with a diameter of 4 mm which are filled with xenon gas. The walls of each tube are electrically charged, and a thin conducting wire runs along the center. When a charged particle traverses a tube, it ionises the xenon and the resulting liberated electrons drift along the electric field to the wire, where an associated charge is registered. In the barrel the straws run parallel to the z -axis and therefore the TRT only provides tracking information in $R\phi$. Straws are arranged radially in the end-caps. The resulting two-dimensional spacepoints have a resolution of approximately $120\ \mu\text{m}$. The spaces between the straws are filled with a polymer which leads to the emission of transition radiation, aiding electron identification.

2.2.2 Calorimeters

The calorimeter system measures the energy of incident particles over the range $|\eta| < 4.9$. There are two main sub-systems: the electromagnetic calorimeter (ECal), which focuses on the measurement of electrons and photons, and the hadronic calorimeter (HCal), which measures the energy of hadrons. Upon entering the calorimeter, incident particles will interact with the detector material to produce a shower of secondary particles with reduced energies. The charge deposited in this process is measured to reconstruct the energy of the initial incident particle. The two calorimeter sub-systems must provide strong containment of showering particles to prevent punch-through of EM and non-muon particles to the HCal and muon system respectively.

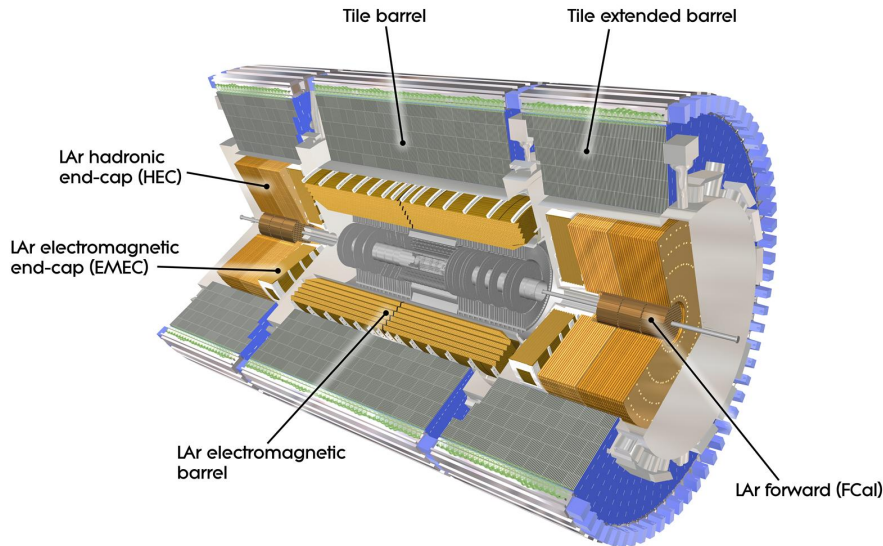


Figure 2.8: The ATLAS calorimeters. The ECal (orange) and HCal (grey, brown).

Liquid Argon (LAr) Electromagnetic Calorimeter

The more granular lead/liquid-argon ECal covers the region $|\eta| < 3.2$ and is split into barrel (covering $|\eta| < 1.475$) and end-cap (covering $1.375 < |\eta| < 3.2$) regions. EM calorimetry works by encouraging electrons and photons to interact with electrically charged particles in detector material via bremsstrahlung ($e \rightarrow e\gamma$) and pair

production ($\gamma \rightarrow e^+e^-$). The EM calorimeter uses lead absorber plates to initiate EM showers, resulting in secondary particles which ionise the surrounding liquid argon. The charge is collected on copper electrodes and read out. The accordion geometry of the ECal allows for a full coverage in ϕ without any azimuthal cracks.

Hadronic Tile Calorimeter

In the central barrel region with $|\eta| < 1.7$, the HCal uses a tile calorimeter with steel as an absorbing material, and scintillating tiles as the active material. Two copper/liquid-argon calorimeter end-caps are also used. Incident hadrons interact via the strong and electromagnetic forces with the absorber material, mainly losing energy due to multiple inelastic nuclear collisions. The active material captures the resulting electrons and photons to measure the energy of the incident hadron.

2.2.3 Muon Spectrometer

Due to their higher mass, muons easily pass unimpeded through the ID and calorimeters and therefore require specialised detectors for their measurement. The Muon Spectrometer (MS) is made up of dedicated tracking and triggering hardware. The precision tracking system uses three layers of monitored drift tubes with a barrel region covering $|\eta| < 1.2$ and end-caps covering $1 < |\eta| < 2.7$. The inner layers of the end-caps use cathode strip chambers to better cope with the high occupancy in the forward region. Precision tracking resolution is approximately $50\,\mu\text{m}$. The trigger system is comprised of resistive plate chambers in the barrel region covering $|\eta| < 1.0$ and thin gap chambers in the end-cap regions covering $1 < |\eta| < 2.4$. A set of three toroidal magnets, each made up of eight coils, is used in each of the barrel and end-caps to deflect the muons as they pass through the MS, allowing their momentum and charge to be measured from the direction and magnitude of curvature. The toroidal magnets generate a field which is largely orthogonal to the muon trajectories which allows for maximum deflection.

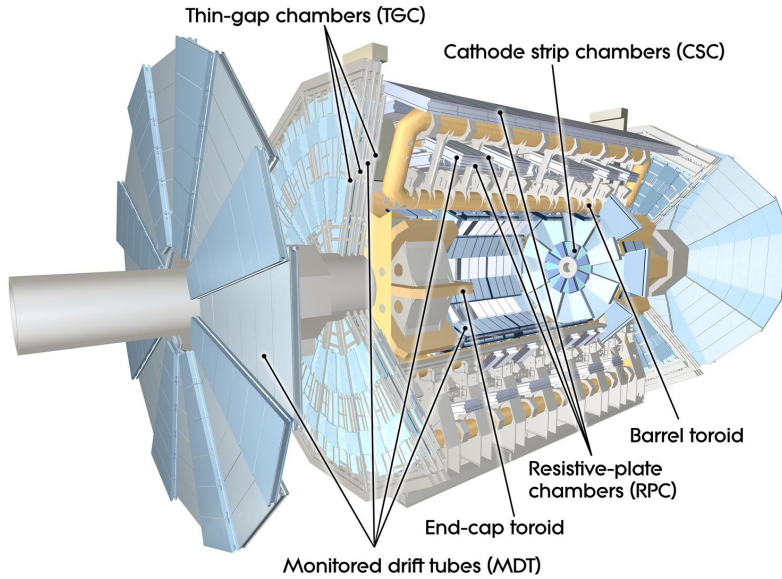


Figure 2.9: The ATLAS muon spectrometer.

2.2.4 The Trigger

The 2.5 ns bunch spacing used over the course of Run 2 corresponds to a bunch-crossing or event rate of 40 MHz (see Table 2.1). If the full information for the detector was written out for each event, this would correspond to the generation of 60 TB of data each second. This is more than can be feasibly processed and stored, requiring the use of a trigger system which quickly makes a decision about whether or not an event is potentially interesting and should be kept for further analysis. The trigger system is comprised of two levels which search for signs of electrons, muons photons, taus and jets, as well as events with high total or missing transverse energy. The hardware-based Level-1 (L1) trigger uses coarse information from the calorimeters and MS to accept events at a rate of 100 kHz within 2.5 μ s of the event. After the L1 trigger, the software-based High Level Trigger (HLT) makes use of 40 000 CPU cores to make a final selection on surviving events within a few hundred milliseconds. The final event read-out rate is approximately 1.2 kHz, corresponding 1.2 GB s^{-1} of permanent data storage. More information is provided in [35].

2.3 Reconstructed Physics Objects

Events passing the trigger selection (described in Section 2.2.4) undergo full offline reconstruction, which makes use of the full information from the detector. Reconstruction and analysis of events relies on the extensive ATLAS software stack, see Ref. [36] for more information.

Reconstruction is the process of analysing the raw signals from the detector to determine which particles and their properties.

The reconstructed event provides information about the underlying physics process that led to the observable final state.

2.3.1 Tracks

The trajectories of charged particles are reconstructed as tracks from the energy depositions (hits) of the particles as they traverse the sensitive elements of the inner detector. Track selection follows the loose selection described in Ref. [37] and outlined in Table 2.2, which was found to improve the flavour tagging performance compared to previous tighter selections, whilst ensuring good resolution of tracks and a low fake rate [38]. The transverse IP d_0 and longitudinal IP z_0 are measured with respect to the hard scatter primary vertex, defined as the reconstructed primary vertex (PV) with the largest sum of the transverse momentum (p_T) of the associated tracks squared, $\sum p_T^2$.

2.3.2 Vertices

Primary Vertices

Each proton-proton interaction happens at a *primary vertex*. Using tracks, all vertices are reconstructed iteratively. The primary vertex, or **hard scatter** of an event is chosen as the interaction whose associated tracks have the highest sum of squared p_T .

Table 2.2: Quality selections applied to tracks, where d_0 is the transverse IP of the track, z_0 is the longitudinal IP with respect to the PV and θ is the track polar angle. Shared hits are hits used on multiple tracks which have not been classified as split by the cluster-splitting neural networks [38]. Shared hits on pixel layers are given a weight of 1, while shared hits in the SCT are given a weight of 0.5. A hole is a missing hit, where one is expected, on a layer between two other hits on a track.

Parameter	Selection
p_T	$> 500 \text{ MeV}$
$ d_0 $	$< 3.5 \text{ mm}$
$ z_0 \sin \theta $	$< 5 \text{ mm}$
Silicon hits	≥ 8
Shared silicon hits	< 2
Silicon holes	< 3
Pixel holes	< 2

Secondary Vertices

SV1 JF

2.3.3 Jets

Jets are reconstructed from particle-flow objects [39] using the anti- k_T algorithm [40] with a radius parameter of 0.4. The jet energy scale is calibrated according to Ref. [41]. Jets are also required not to overlap with a generator-level electron or muon from W boson decays. All jets are required to have a pseudorapidity $|\eta| < 2.5$ and $p_T > 20 \text{ GeV}$. Additionally, a standard selection using the Jet Vertex Tagger (JVT) algorithm at the tight working point is applied to jets with $p_T < 60 \text{ GeV}$ and $|\eta| < 2.4$ in order to suppress pile-up contamination [42]. Tracks are associated to jets using a ΔR association cone, the width of which decreases as a function of jet p_T , with a maximum cone size of $\Delta R \approx 0.45$ for jets with $p_T = 20 \text{ GeV}$ and minimum cone size of $\Delta R \approx 0.25$ for jets with $p_T > 200 \text{ GeV}$. If a track is within the association cones of more than one jet, it is assigned to the jet which has a smaller $\Delta R(\text{track}, \text{jet})$.

Jet flavour labels are assigned according to the presence of a truth hadron within $\Delta R(\text{hadron}, \text{jet}) < 0.3$ of the jet axis. If a b -hadron is found the jet is labelled a b -jet. In the absence of a b -hadron, if a c -hadron is found the jet is called a c -jet. If no b - or c -hadrons are found, but a τ is found in the jet, it is labelled as a τ -jet, else it is labelled as a light-jet.

- Jet finding algorithms

2.3.4 Leptons

Chapter 3

Investigating Tracking Improvements

Todo:

- Check all info wrt to [this PDG review](#)

3.1 b -hadron Reconstruction

3.1.1 b -hadron Decay Topology

B-hadrons (or b -hadrons) refer collectively to B mesons $\bar{b}q$ and hadrons with $bq\bar{q}$, where q is a quark lighter than b . In essence then, b -hadrons are bound states of a bottom quark and lighter quarks. Due to the smallness of V_{cb} , in the CKM matrix, they live more than a trillionth of a second before disintegrating (relatively long time). Due to the existence of an internal structure, the hadron can be in distinct energy states in the same way as an atomic nucleus. Different combinations of spin and orbital angular momentum give rise to these excited states. The excited states will quickly de-excite by photon emission (or similar?). When the b quark itself decays the B-hadron is likely to become a D-hadron, which contain c type quarks.

b -hadrons are quasi-stable bound states of quarks, where one of the quarks is a bottom quark (b quark). The proper lifetimes τ of the various b -hadrons are similar and relatively long, with $\tau \sim 10^{-12}$ s. This lifetime corresponds to a proper decay length $c\tau \sim 300 \mu\text{m}$. In the rest frame of the detector, the typical b -hadron travels a distance $d = \beta\gamma c\tau$ before decaying, where at high energies $\gamma \sim E_B/m_B$. For a 1

TeV b -hadron, this gives $d \sim 60$ mm - well beyond the radius of the first pixel layer (IBL) at 33 mm. At the LHC, b quarks are generated in the hard scattering

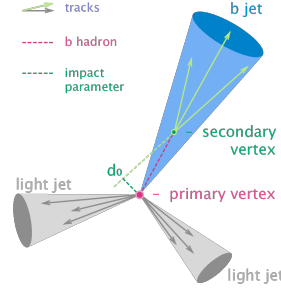


Figure 3.1: Diagram of a typical b -jet (blue) which has been produced along with two light jets (grey). The b -hadron has travelled a significant distance from the primary interaction point (pink dot) before its decay. The large transverse impact parameter d_0 is a characteristic property of the trajectories of b -hadron decay products.

of proton-proton (pp) collisions. They quickly hadronize into a b -hadron, which is often initially in an excited state due to the high energies of the pp collisions at the LHC ($\sqrt{s} = 13$ TeV). The hadronisation process is hard - around 70-80% of the b quark's momentum goes into the b -hadron, with the rest being radiated as other particles. The excited b -hadron will quickly fragment (i.e. de-excite) by radiating particles, which are prompt (they are formed close to the primary vertex). These fragmentation particles have an increasing multiplicity and collimation to the b -hadron axis as the p_T of the b -hadron increases. The de-excited b -hadron subsequently weakly decays to on average 4 or 5 particles (the multiplicity of the decay products of the weak decay of the b -hadron is unaffected by increases in the b -hadron p_T).

Due to their lifetimes, energetic b -hadrons can travel a significant distance from the primary pp interaction point before decaying to a spray of collimated stable particles. This signature is registered in the detector as a displaced jet. Due to the elements of the CKM matrix, b -hadrons decay with a high probability to D hadrons (which contain a c quark), which also have significant lifetimes - this can lead to reconstructed tertiary vertices in the jet core. The typical features of a b -jet, and in particular the large track impact parameter d_0 which can result from displaced decays, are shown in Fig. 3.1. Many ATLAS analyses rely on a method of tagging jets instantiated by b quarks and rejecting jets created from other quarks (c and light flavours u, d, s). These “ b -tagging” algorithms work by discriminating against

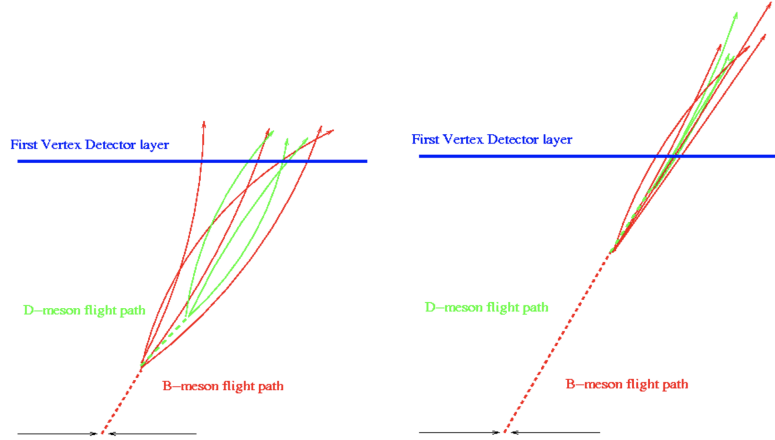


Figure 3.2: As b -hadron p_T increases, the time of flight of the B increases, so tracks will have less room to diverge before reaching detector elements. To compound the problem, the collimation of the tracks increases. The detector may then be unable to resolve individual tracks.

the unique signatures of b -jets discussed above. b -tagging relies on the efficient and accurate reconstruction the tracks corresponding to the b -hadron decay products. These tracks are then used as inputs to vertex reconstruction algorithms and jet making algorithms.

3.1.2 b -hadron Decay Track Reconstruction

A necessary requirement for successful jet b -tagging is the efficient and accurate reconstruction of the charged particle trajectories in the jet. For high p_T jets ($p_T > 200$ GeV) this task becomes difficult due to a combination of effects. As the jet energy increases, the track multiplicity of the jet increases due to the presence of additional fragmentation tracks. Tracks in the jet also become increasingly collimated as their inherited transverse momentum increases. Together, these two effects lead to a very high density of charged particles in the jet core, making reconstruction difficult. At high energies, the increased decay length of B (and D) hadrons means that decay products have less of an opportunity to diverge before reaching the first tracking layers of the detector. If the decay takes place very close to a detector layer, or if the decays are sufficiently collimated, hits left by nearby particles may not be resolved individually, leading to merged clusters (shown in Fig. 3.2). Shared hits generally predict bad tracks. As such, shared hits are heavily penalised during

reconstruction (and in particular as part of ambiguity solving). However, in the core of high p_T b -jets, where decay particles are displaced from the primary vertex and are highly collimated, the density of particles is high enough that the probability of clusters being merged increases dramatically. The presence of merged clusters requires that the corresponding tracks share hits (if they are to be reconstructed successfully), which may end up impairing the successful reconstruction of the track. Furthermore, decays may also take place inside the tracking detectors themselves, which can lead to missing or wrong innermost cluster assignment. The combination of effects described above makes reconstructing tracks in the core of high p_T b -jets particularly challenging.

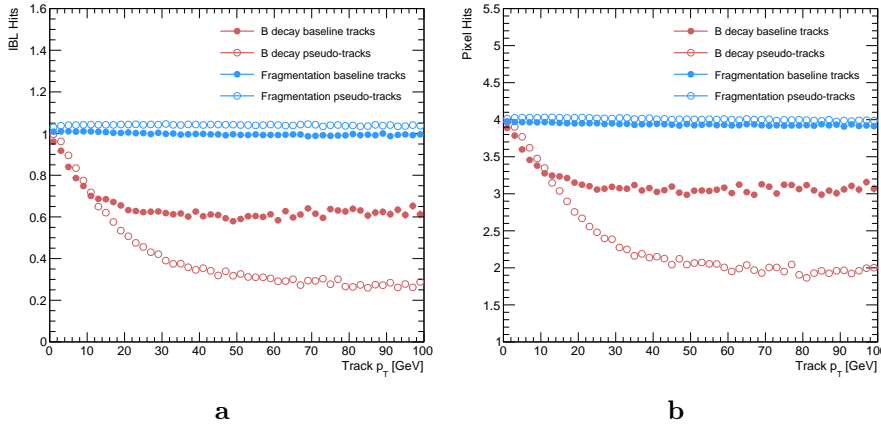


Figure 3.3: Hit multiplicities on the IBL (3.3a) and the pixel layers (3.3b) as a function of the p_T of the reconstructed track. Tracks from the weak decay of the b -hadron are shown in red, while fragmentation tracks (which are prompt) are in blue. For each of these, standard tracks and pseudo-tracks are plotted. Hit multiplicities on the pseudo-tracks at high p_T due to the increased flight of the b -hadron. The baseline tracks have more hits than the pseudo-tracks, indicating that they are being incorrectly assigned additional hits.

Figure 3.4: Track reconstruction efficiency from b -hadron decay products for baseline ATLAS tracking (black), Bcut+Refit procedures applied (green), pseudo-tracking (blue), and for tracking where the ambiguity solver has been manually removed (orange).

Concretely, then, the issues relating to high p_T b -hadron tracking can be factorised into two parts. The first part is a drop in track reconstruction efficiency. As mentioned, tracks originating from high energy b -hadron decay products can have a high rate of shared hits due to the number of particles present in a high p_T b -jet

Figure 3.5: The total number of pixel hits on tracks from b -hadron decays as a function of the production radius of the decay product. An excess of hits is assigned to the standard tracks in comparison to the ideal pseudo-tracks.

and their relative collimation. Additionally, tracks may be missing hits on the inner layers of the detector. This occurs primarily when the decay b -hadron decays inside the detector. These features of can make it difficult for B decay tracks to meet the ambiguity solver’s stringent track quality requirements. As a result, many B decay tracks are rejected in the ambiguity solving stage, leading to a severe drop in tracking reconstruction efficiency. This is shown by the severe decrease in reconstruction efficiency visible when comparing baseline tracking with the ideal pseudo-tracks in Fig. 3.4. This situation presents a problem: relaxing cuts on shared hits significantly degrades the ambiguity solver’s power to reject bad tracks. However for b -hadron decay tracks it seems these same restrictions on shared hits are seriously impairing the reconstruction efficiency of good tracks. The second part of the problem is that, due to the high density of clusters available for assignment in the vicinity of the typical high energy b -hadron decay track, and also given the strong positive bias of the ambiguity solver towards those tracks with precise pixel measurements (especially the innermost IBL measurement), many b -hadron decay tracks are assigned incorrect inner layer hits. This is only a problem for those decay products which were produced inside the pixel detector as a result of a long-flying b -hadron, and so do not have a correct hit available for assignment (evidenced in Fig. 3.8b). The incorrect hits may skew the parameters of the track, which can in turn mislead b -tagging algorithms. In particular, b -tagging algorithms rely heavily on the transverse impact parameter significance $d_0/\sigma(d_0)$ of the track. The quality of this measurement is expected to be adversely affected by wrong inner-layer hits on the track. This combination of reduced reconstruction efficiency and incorrectly assigned hits is thought to be the cause of the observed drop in b -tagging efficiency at high energies , although it is not clear which effect may dominate.

3.2 Pseudotracks and Ideal Tracks

Pseudotracking and ideal tracking are used as benchmarks of the best tracking possible given the ATLAS detector. Both pseudotracks and ideal tracks are constructed using truth information to group combinations of hits that have been left by the same truth particle. As a result, hit-to-track association and track reconstruction efficiency are both ideal (given the ATLAS detector). Ideal tracks represent a yet more idealised tracking scenario by correcting the cluster positions based on truth information, and smearing the cluster position based on the detector resolution.

When pseudotracking is run alongside standard tracking, those clusters which are shared on the reconstructed tracks run through the cluster splitting machinery. If a cluster is found to be compatible with being split, its definition is changed, and the pseudotracks use this definition too. As a result, pseudotracks can have split clusters.

3.3 Investigating Improvements for High p_T B Tracking

An investigation into

3.3.1 Looser Track Cuts & Track Refit Procedure

A solution for the problem of wrong inner-layer hits on B tracks had previously been developed. This solution selects tracks which pass a b -jet Region of Interest (ROI) selection, and then removes the innermost hits on these tracks based on the result of a “refit” procedure. The refit procedure runs as follows. Each track is refitted without the innermost hit, and if there is a significant improvement in the fit quality (the χ^2 of the track fit divided by the number of degrees of freedom on the track n), the innermost hit is rejected and the new track is replaces the old. If the fit quality does not improve by a certain amount, the initial track is kept. This procedure is recursively applied. The b -jet ROI selection selects tracks that are matched within $dR < 0.14$ ($|\eta| < 0.1$, $|\phi| < 0.1$) of a CaloCluster with $E_T > 150$

GeV. The track itself must also pass a transverse momentum cut with $p_T > 15$ GeV. The refit procedure was previously shown to lead to a reduction in the rate of wrongly assigned IBL hits on B decay tracks (see Fig. 3.8b). However, this apparent improvement did not lead to an increase in b -tagging performance. It was found that the refit procedure also removed unacceptable numbers of good hits, degrading the quality of un-problematic tracks, shown in Fig. 3.8a. This is likely the cause of the underwhelming b -tagging performance improvement.

The performance of both the ROI, and the hit removal using track fit information, is examined, and an attempt at improving the performance of the refit procedure is made. Results are discussed in the following two sections.

3.3.2 Region of Interest Optimisation

Selection cuts for the b -jet ROI were determined on a largely ad-hoc basis. An effort was made to systematically optimise the selection cuts. The decay tracks of B hadrons are tightly collimated with the B itself, with most decay products satisfying $dR(B, \text{track}) < 0.02$, as shown in Fig. 3.6a. Meanwhile, calorimeter clusters relating to the B hadrons are generally found within $dR < 0.05$ of the B Fig. 3.6b. In total, then, B decay tracks will usually be found within $dR < 0.07$ of the relevant calorimeter cluster, which suggests that the current $dR < 0.14$ is loose by a factor of two. Similar analysis of cluster and track energy distributions found that the related cuts were also loose, and so they were modified from $E_T > 150$ GeV to $E_T > 300$ GeV, and from $p_T > 15$ GeV to $p_T > 30$ GeV.

Additionally examined in the course of this work was the fake rate of the b -jet ROI. The distributions in Fig. 3.7a demonstrate that most of clusters passing the $E_T > 150$ GeV selection were unable to be matched to a nearby B hadron using truth information. Clusters that pass the selection but do not correspond to energy depositions from B hadrons lead to fake ROIs. As a consequence of these distributions, tracks selected by the ROI are largely impure in the desired B hadron tracks.

The modified ROI was used to re-run the refit procedure. A comparison of of “standard” and “optimised” (using the optimised b -jet ROI) refit procedures is found in Fig. 3.8. These results show that whilst tighter selection cuts did lead to a

recovery of some good hits (Fig. 3.8a), performance with respect to the baseline is still significantly degraded.

a b

Figure 3.6: Distributions of angular distance dR between B hadrons and their weak decays and other fragmentation tracks (Fig. 3.6a), and the distribution of angular distance dR between B hadrons and the calorimeter clusters in the hadronic calorimeter (Fig. 3.6b). In Fig. 3.6a, the tracks from the weak decay of the B are significantly more collimated to the B than the other fragmentation tracks.

3.3.3 Fit Quality as a Discriminant for Wrong Hits

As mentioned, tracks selected by the ROI are refitted without their innermost hit, and, if an improvement in fit quality is observed, the hit is rejected. In order to test the effectiveness of this procedure, a dataset of two sets of tracks was produced. The first set contained unmodified baseline-reconstructed tracks. The second contained the same tracks as the first, but modifications made during reconstruction removed the innermost hit on each track. Then, using Monte Carlo (MC) truth information, a track-by-track fit quality comparison was made for tracks with good and wrong innermost hits.

It is clear from the distributions in Fig. 3.7b that the fit quality improvement (measured by fractional change in χ^2/n of the track before and after the innermost hit is removed) is not a discriminating variable for wrong hits, and indeed attempted optimisations of the of the refit procedure based on these distributions were found to be ineffectual. While wrong hits are likely to degrade the track fit, it is also true that any additional measurement, good or wrong, constrains the track, and therefore removal of that measurement will be likely to lead to an increase in the χ^2/n of the track. Removing hits in this way is therefore problematic.

3.3.4 Conclusion

The work outlined in the two preceding sections has uncovered issues with both the b -jet ROI, and the methodology of identification and removal of wrong hits on tracks inside a given ROI. Attempts were made to optimise the selection cuts of the ROI,

a b

Figure 3.7: The distribution of cluster transverse momentum, in Fig. 3.7a for both clusters that were able (orange) and unable (blue) to be matched to a B hadron using MC truth information. The normalisation shows that the majority of clusters are not matched to B hadrons, resulting in fake ROIs. In Fig. 3.7b, the fractional improvement in track fit quality (χ^2/n) is shown for all track (blue), tracks with good IBL hits (green), and tracks with wrong IBL hits (orange). The distributions are overlapping, suggesting that the χ^2/n improvement is not a good discriminator of good and wrong hits.

however the large background of energetic phenomena produced in collisions that are not B hadron related means that the ROI is largely unsuccessful in selecting a pure sample of likely B hadron candidates. An additional effort was made to improve the removal of wrong hits using other information in addition to the track fit improvement. Information such as the type and locations of its, and track d_0 were considered. While progress here was not insignificant, without substantial overhaul of the ROI to improve B purity, the results were not strong enough to demonstrate any viable solutions that would successfully target and then improve B hadron decay tracks. Alongside the refit procedure, a “Bcut” cut scheme was suggested in order to improve

a b

Figure 3.8: Distributions of good (Fig. 3.8a) and wrong (fig:refit optimisation results sub2) hit assignment rates on the IBL for tracks using baseline tracking (black), the original unmodified refit procedure (green), and the refit procedure with an optimise set of ROI selection cuts (blue). The IBL lies at a radius of 33 mm from the beam pipe. Hence, particles produced with a production radius greater than this cannot leave good hits on the IBL.

reconstruction performance. This consisted primarily of loosening the shared hit cuts in the ambiguity solver. While this did lead to a measurement increase in track reconstruction efficiency (see Fig. 3.4), it was determined that the corresponding increase in fake tracks (i.e. those tracks for which the majority of hits do not come from a single truth particle) was too large to justify the implementation of the “Bcut” scheme. In conclusion, then, a different approach is required to address the problems discussed.

3.4 Global χ^2 Fitter Outlier Removal

This section documents ongoing progress into improving hit assignments using the Global χ^2 Fitter (GX2F) to prevent wrong hits from being assigned to tracks during the track fit. This is in contrast to the approach discussed in `sec:refit`, which attempts to identify and remove wrong hits after the reconstruction of the track (of which the track fit is a part). As part of the track fit, an outlier removal procedure is run, in which suspicious hits are identified and removed. The GX2F code, as a relatively low-level component of track reconstruction, has not undergone significant modification for several years. During this time, a new tracking sub-detector, the IBL, was installed, and subsequently precise detector alignments have been derived. The motivation for looking at the GX2F is that these changes may require re-optimisation of the GX2F code, and in particular the outlier removal procedures. Further motivation for this approach comes from the low rate of labelled outliers in baseline tracking. For example, while approximately 15% of B hadron decay tracks have a wrong IBL hit (a value which only increases with the p_T of the B), less than 1% of these tracks have had their IBL hit labelled and removed as an outlier.

Implementation

The outlier removal procedure for the pixel detector is described in this section. The states (also called measurements, or hits) on the track are looped over in order of increasing radial distance to the beam pipe. For each state, errors $\sigma(m_i)$ on the measurement of the transverse and longitudinal coordinates are calculated. These errors are dependent on the sub-detector which recorded the measurement (as some sub-detectors are more precise than others). Additionally, a residual displacement r_i between the predicted position of the track x_i (inclusive of the current measurement), and the position of the measurement itself, m_i , is calculated. The pull p_i on the track state due to the current measurement is calculated according to

$$p_i = \frac{m_i - x_i}{\sqrt{\sigma(m_i)^2 - \sigma(x_i)^2}}, \quad r_i = m_i - x_i. \quad (3.1)$$

This pull is computed for the transverse and longitudinal coordinates of the measurement, and the maximum of the two is selected and checked to see if it exceeds a

certain threshold. If it does, the hit will be removed, after some additional checks are made to confirm or deny the presence of the outlier. The threshold is set as a member variable `m_outlcut`. The results of varying this cut are described in Section 3.4.1.

3.4.1 Cut Optimisation

A systematic variation of the cut point `m_outlcut` has been carried out. The results, demonstrating a reduction in wrong hit assignment whilst keeping virtually all good hits assigned to tracks, are shown in Fig. 3.9. The rate of wrong hits assigned to tracks decreases from 0.32 to 0.28 at the highest energies (12.5% reduction). Moreover, this result is obtained looking at all tracks inclusively, and the demonstrated improvement removes the need for a specific b -jet ROI (a requirement which led to problems outlined in Section 3.3.2). These results hold when looking exclusively at B decay tracks. The fact that, as shown in Fig. 3.8a, virtually all correctly assigned hits are

a b

Figure 3.9: Profiles, as a function of parent B hadron p_T , of good (Fig. 3.9a) and wrong (Fig. 3.9b) hit assignment rates on the IBL for tracks using baseline tracking (black), and various looser values of the outlier cut.

retained suggests that it may possible to relax this cut further. Tests are ongoing which will confirm this. The current GX2F treats all layers in the pixel detector in the same way - applying the same cut to each. While Fig. 3.8a shows no adverse affects for hits on the IBL, when relaxing `m_outlcut` to a value of 1, some small reduction in good hit assignment efficiency was observed in other layers of the pixel detector, which are less precise. This difference in precision motivates the need to treat different layers in the pixel detector differently. To this end, layer-specific cutting capabilities for the GX2F are under development, which will allow each pixel layer to have their own cut point for outlier removal. Layer specific cuts will then be optimised to see if greater numbers of wrong hits can be successfully identified as outliers and removed, while maintaining high good hit assignment efficiency.

3.5 Tracking software validation

- tracking validation
- qspi validation

Chapter 4

Track Classification MVA

4.1 Machine Learning Background

4.2 Track Truth Origin Labelling

4.3 Fake Track Identification Tool

Probably talk about this model as a stepping stone to the general classifier

4.3.1 *b*-hadron Decay Track Identification Tool

Maybe don't need this section since it was talked about less

4.4 General Track Origin Classifier Tool

Culmination of this work in the general tool Martino has implemented

Applications:

- Frack to jet association
- Fake track studies (removal and for recommendations)

4.5 Conclusion

Improved with GNNs

Chapter 5

Graph Neural Network Flavour Tagger

Flavour tagging, the identification of jets originating from b - and c -quarks, is a critical component of the physics programme of the ATLAS experiment at the Large Hadron Collider. Current flavour tagging algorithms rely on the outputs of several low-level algorithms, which reconstruct various properties of jets using charged particle tracks, that are then combined using machine learning techniques. In this note a new machine learning algorithm based on graph neural networks, GN1, is introduced. GN1 uses information from a variable number of charged particle tracks within a jet, to predict the jet flavour without the need for intermediate low-level algorithms. Alongside the jet flavour prediction, the model predicts which physics processes produced the different tracks in the jet, and groups tracks in the jet into vertices. These auxiliary training objectives provide useful additional information on the contents of the jet and improve performance. GN1 compares favourably with the current ATLAS flavour tagging algorithms. For a b -jet efficiency of 70%, the light (c)-jet rejection is improved by a factor of ~ 1.8 (~ 2.1) for jets coming from $t\bar{t}$ decays with transverse momentum $20 < p_T < 250$ GeV. For jets coming from Z' decays with transverse momentum $250 < p_T < 5000$ GeV, the light (c)-jet rejection improves by a factor ~ 6 (~ 2.8) for a comparative 30% b -jet efficiency.

5.1 Motivation

Flavour tagging, the identification of jets originating from b - and c -quarks, is a critical component of the physics programme of the ATLAS experiment [28] at the Large Hadron Collider (LHC) [43]. It is of particular importance for the study of the Standard Model (SM) Higgs boson and the top quark, which preferentially decay to b -quarks [44, 45], and additionally for several Beyond Standard Model (BSM) resonances that readily decay to heavy flavour quarks [46]. The significant lifetime of b -hadrons, approximately 1.5 ps [47], provides the unique signature of a secondary decay vertex which has a high mass and is significantly displaced from the primary vertex. Additional signatures of b -hadrons are the tertiary decay vertex, resulting from $b \rightarrow c$ decay chains, and the reconstructed trajectories of charged particles (henceforth simply referred to as tracks) with large impact parameters¹ (IPs). These signatures are primarily identified using tracks associated to jets. As such, efficient and accurate track reconstruction is essential for high performance flavour tagging.

This note introduces a novel algorithm, GN1, which uses Graph Neural Networks (GNNs) [48] with auxiliary training objectives, to aid the primary goal of classifying whether jets originate from b - or c -quarks (referred to as a flavour tagger). The concept is illustrated in Fig. 5.1. The use of GNNs offers a natural way to classify jets with variable numbers of unordered associated tracks, while allowing for the inclusion of auxiliary training objectives [49, 50].

The current ATLAS flavour tagger, DL1r [51], is a deep neural network which takes the outputs of a number of independently optimised “low-level” algorithms [52] as inputs. Each of these low-level algorithms makes use of tracks to reconstruct a particular aspect of the experimental signature of heavy flavour jets. The low-level algorithms can be manually optimised reconstruction algorithms, for example the SV1 and JetFitter algorithms that reconstruct displaced decay vertices, or trained taggers such as RNNIP and DIPS that use the IPs of a variable number of tracks to identify the flavour of the jet [37, 52–54]. In contrast GN1 utilises a single neural network, which directly takes the tracks and some information about the jet as inputs. As such, it does not depend on any other flavour tagging algorithm, and a single training of the GN1 fully optimises all aspects of the algorithm.

¹The distance of closest approach from a track to the primary vertex.

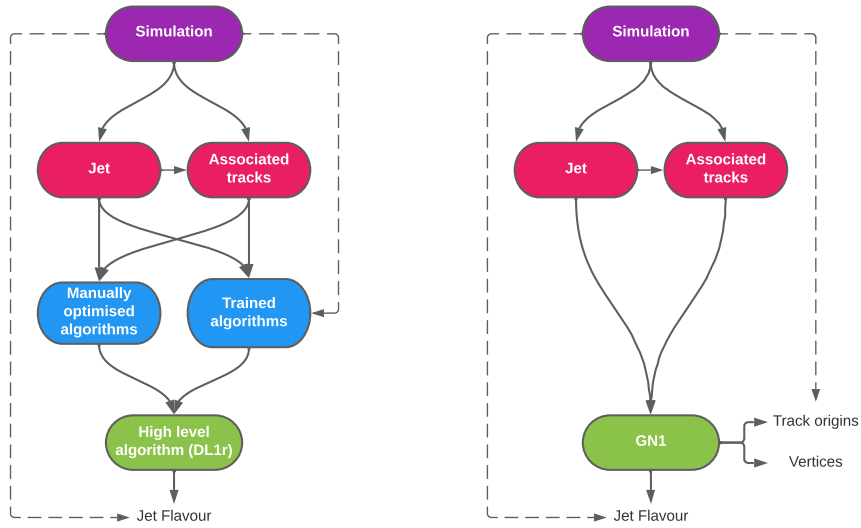


Figure 5.1: Comparison of the existing flavour tagging scheme (left) and GN1 (right). The existing approach utilises low-level algorithms (shown in blue), the outputs of which are fed into a high-level algorithm (DL1r). Instead of being used to guide the design of the manually optimised algorithms, additional truth information from the simulation is now being used as auxiliary training targets for GN1. The solid lines represent reconstructed information, whereas the dashed lines represent truth information.

GN1 is trained to understand the internal structure of the jet through the use of two auxiliary training objectives: the grouping of tracks originating from a common vertex, and the prediction of the underlying physics process from which each track originated. These auxiliary objectives are meant to guide the neural network towards a more complete understanding of the underlying physics, removing the need for the low-level algorithms, and therefore simplifying the process of optimising the tagger for new regions of phase space (e.g. c -tagging or high- p_T b -tagging), or when the detector or charged particle reconstruction algorithms are updated. The training targets for the primary and auxiliary objectives are extracted from “truth information”, i.e. information only available in simulation, as opposed to reconstructed quantities available in both collision data and simulation.

In this note, the following benefits of this approach will be shown:

1. Improved performance with respect to the current ATLAS flavour tagging algorithms, with larger background rejection for a given signal efficiency.
2. The same network architecture can be easily optimised for a wider variety of use cases (e.g. c -jet tagging and high- p_T jet tagging), since there are no low-level algorithms to retune.
3. There are fewer flavour tagging algorithms to maintain.
4. Alongside the network’s prediction of the jet flavour, the auxiliary vertex and track origin predictions provide more information on why a jet was (mis)tagged or not. This information can also have uses in other applications, for instance to explicitly reconstruct displaced decay vertices or to remove fake tracks.²

This note is organised as follows: a brief description of the ATLAS detector, object definitions and selections, and samples are provided in Section 5.3; details about the model architecture and training procedure are given in Section 5.4; and results are discussed in Section 5.5.

²A fake track is defined as a track with a truth-matching probability less than 0.5, where the truth-matching probability is defined in Ref. [38].

5.2 Graph Neural Network Theory

5.3 Experimental Setup

5.3.1 Datasets

To train and evaluate the model, simulated SM $t\bar{t}$ and BSM Z' events initiated by proton-proton collisions at a center of mass energy $\sqrt{s} = 13$ TeV are used. The Z' sample is constructed in such a manner that it has a relatively flat jet p_T spectrum up to 5 TeV and decays to an equal numbers of b -, c - and light- jets. The generation of the simulated event samples includes the effect of multiple pp interactions per bunch crossing with an average pileup of $\langle\mu\rangle = 40$, which includes the effect on the detector response due to interactions from bunch crossings before or after the one containing the hard interaction.

The $t\bar{t}$ events are generated using the POWHEGBOX [55–58] v2 generator at next-to-leading order with the NNPDF3.0NLO [59] set of parton distribution functions (PDFs). The h_{damp} parameter³ is set to 1.5 times the mass of the top-quark (m_{top}) [60], with $m_{\text{top}} = 172.5$ GeV. The events are interfaced to PYTHIA 8.230 [61] to model the parton shower, hadronisation, and underlying event, with parameters set according to the A14 tune [62] and using the NNPDF2.3LO set of PDFs [63]. Z' events are generated with PYTHIA 8.2.12 with the same tune and PDF set. The decays of b - and c -hadrons are performed by EVTGEN v1.6.0 [64]. Particles are passed through the ATLAS detector simulation [65] based on GEANT4 [66].

For the $t\bar{t}$ events, at least one W boson from the top quark decay is required to decay leptonically. Truth labelled b -, c - and light- jets are kinematically re-sampled in p_T and η to ensure identical distributions in these variables. The resulting dataset contains 30 million jets, 60% of which are $t\bar{t}$ jets and 40% of which are Z' jets. While DL1r uses 70% $t\bar{t}$ jets and 30% Z' jets, the change in sample composition did not affect the final performance of GN1. To evaluate the performance of the model, 500k jets from both the $t\bar{t}$ and Z' samples, which are statistically independent from the

³The h_{damp} parameter is a resummation damping factor and one of the parameters that controls the matching of POWHEG matrix elements to the parton shower and thus effectively regulates the high- p_T radiation against which the $t\bar{t}$ system recoils.

training sample, are used. Track- and jet-level inputs are scaled to have a central value of zero and a variance of unity before training and evaluation.

5.4 Model Architecture

5.4.1 Model Inputs

GN1 is given two jet variables and 21 tracking related variables for each track fed into the network. The jet transverse momentum and signed pseudorapidity constitute the jet-level inputs, with the track-level inputs listed in Table 5.1. If a jet has more than 40 associated tracks, the first 40 tracks with the largest transverse IP significance⁴ $s(d_0)$ are selected as inputs. Full track parameter information and associated uncertainties, along with detailed hit information, carry valuable information about the jet flavour. In the dense cores of high- p_T jets, tracks are highly collimated and separation between tracks can be of the same order as the active sensor dimensions, resulting in merged clusters and tracks which share hits [38]. Due to the relatively long lifetimes of b -hadrons and c -hadrons, which can traverse several layers of the ID before decaying and have highly collimated decay products, the presence of shared or missing hits is a critical signature of heavy flavour jets.

Dependence on the absolute value of the azimuthal jet angle ϕ is explicitly removed by providing only the azimuthal angle of tracks relative to the jet axis. The track pseudorapidity is also provided relative to the jet axis.

Since heavy flavour hadrons can decay semileptonically, the presence of a reconstructed lepton in the jet carries discriminating information about the jet flavour. In addition to the baseline GN1 model, the GN1 Lep variant includes an additional track-level input, leptonID, which indicates if the track was used in the reconstruction of an electron, a muon or neither. The muons are required to be combined [68], and the electrons are required to pass the *VeryLoose* likelihood-based identification working point [69].

⁴Impact parameter significances are defined as the IP divided by its corresponding uncertainty, $s(d_0) = d_0/\sigma(d_0)$ and $s(z_0) = z_0/\sigma(z_0)$. Track IP significances are lifetime signed according to the track's direction with respect to the jet axis and the primary vertex [67].

Table 5.1: Input features to the GN1 model. Basic jet kinematics, along with information about the reconstructed track parameters and constituent hits are used. Shared hits, are hits used on multiple tracks which have not been classified as split by the cluster-splitting neural networks [38], while split hits are hits used on multiple tracks which have been identified as merged. A hole is a missing hit, where one is expected, on a layer between two other hits on a track. The track leptonID is an additional input to the GN1 Lep model.

Jet Input	Description
p_T	Jet transverse momentum
η	Signed jet pseudorapidity
Track Input	Description
q/p	Track charge divided by momentum (measure of curvature)
$d\eta$	Pseudorapidity of the track, relative to the jet η
$d\phi$	Azimuthal angle of the track, relative to the jet ϕ
d_0	Closest distance from the track to the PV in the longitudinal plane
$z_0 \sin \theta$	Closest distance from the track to the PV in the transverse plane
$\sigma(q/p)$	Uncertainty on q/p
$\sigma(\theta)$	Uncertainty on track polar angle θ
$\sigma(\phi)$	Uncertainty on track azimuthal angle ϕ
$s(d_0)$	Lifetime signed transverse IP significance
$s(z_0)$	Lifetime signed longitudinal IP significance
nPixHits	Number of pixel hits
nSCTHits	Number of SCT hits
nIBLHits	Number of IBL hits
nBLHits	Number of B-layer hits
nIBLShared	Number of shared IBL hits
nIBLSplit	Number of split IBL hits
nPixShared	Number of shared pixel hits
nPixSplit	Number of split pixel hits
nSCTShared	Number of shared SCT hits
nPixHoles	Number of pixel holes
nSCTHoles	Number of SCT holes
leptonID	Indicates if track was used to reconstruct an electron or muon

5.4.2 Auxiliary Training Objectives

In addition to the jet flavour classification, two auxiliary training objectives are defined. Each auxiliary training objective comes with a training target which, similar to the jet flavour label, are truth labels derived from the simulation. The presence of the auxiliary training objectives improves the jet classification performance as demonstrated in Section 5.5.3.

The first auxiliary objective is the prediction of the origin of each track within the jet. Each track is labelled with one of the exclusive categories defined in Table 5.2 after analysing the particle interaction that led to its formation. Since the presence of different track origins is strongly related to the flavour of the jet, training GN1 to recognise the origin of the tracks may provide an additional handle on the classification of the jet flavour. This task may also aid the jet flavour prediction by acting as a form of supervised attention [70] - in detecting tracks from heavy flavour decays the model may learn to pay more attention to these tracks.

Table 5.2: Truth origins which are used to categorise the physics process that led to the production of a track. Tracks are matched to charged particles using the truth-matching probability [38]. A truth-matching probability of less than 0.5 indicates that reconstructed track parameters are likely to be mismeasured and may not correspond to the trajectory of a single charged particle. The “OtherSecondary” origin includes tracks from photon conversions, K_S^0 and Λ^0 decays, and hadronic interactions.

Truth Origin	Description
Pileup	From a pp collision other than the primary interaction
Fake	Created from the hits of multiple particles
Primary	Does not originate from any secondary decay
fromB	From the decay of a b -hadron
fromBC	From a c -hadron decay, which itself is from the decay of a b -hadron
fromC	From the decay of a c -hadron
OtherSecondary	From other secondary interactions and decays

Displaced decays of b - and c -hadrons lead to secondary and tertiary vertices inside the jet. Displaced secondary vertices can also occur in light-jets as a result of material interactions and long-lived particle decays (e.g. K_S^0 and Λ^0). The second auxiliary objective is the prediction of track-pair vertex compatibility. For each pair of tracks

in the jet, GN1 predicts a binary label, which is given a value 1 if the two tracks in the pair originated from the same point in space, and 0 otherwise. To derive the corresponding truth labels for training, truth production vertices within 0.1 mm are merged. Track-pairs where one or both of the tracks in the pair have an origin label of either Pileup or Fake are given a label of 0. Using the pairwise predictions from the model, collections of commonly compatible tracks can be grouped into vertices. The addition of this auxiliary training objective removes the need for inputs from a dedicated secondary vertexing algorithm.

Both auxiliary training objectives can be considered as “stepping stones” on the way to classifying the flavour of the jet. By requiring the model to predict the truth origin of each track and the vertex compatibility of each track-pair, the model is guided to learn representations of the jet which are connected to the underlying physics and therefore relevant for classifying the jet flavour.

5.4.3 Architecture

As discussed above, the GN1 model combines a graph neural network architecture [71] with auxiliary training objectives in order to determine the jet flavour. Coarse optimisation of the network architecture hyperparameters, for example number of layers and number of neurons per layer, has been carried out to maximise the tagging efficiency.

The model architecture is based on a previous implementation of a graph neural network jet tagger [50]. As compared to the previous approach, GN1 uses a only a single graph neural network and makes use of a more sophisticated graph neural network layer [72], described below. These changes yield improved tagging performance and a significant reduction in training time with respect to the previous approach.

The model takes jet- and track-level information as inputs, as detailed in Section 5.4.1. The jet inputs are concatenated with each track’s inputs, as shown in Fig. 5.2. The combined jet-track vectors are then fed into a per-track initialisation network with three hidden layers, each containing 64 neurons, and an output layer with a size of 64, as shown in Fig. 5.3. The track initialisation network is similar to a Deep Sets model [73], but does not include a reduction operation (mean or summation) over the output track representations.

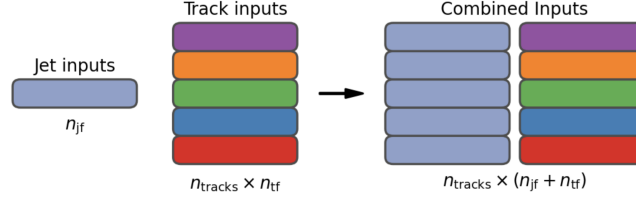


Figure 5.2: The inputs to GN1 are the two jet features ($n_{jf} = 2$), and an array of n_{tracks} , where each track is described by 21 track features ($n_{tf} = 21$). The jet features are copied for each of the tracks, and the combined jet-track vectors of length 23 form the inputs of GN1.

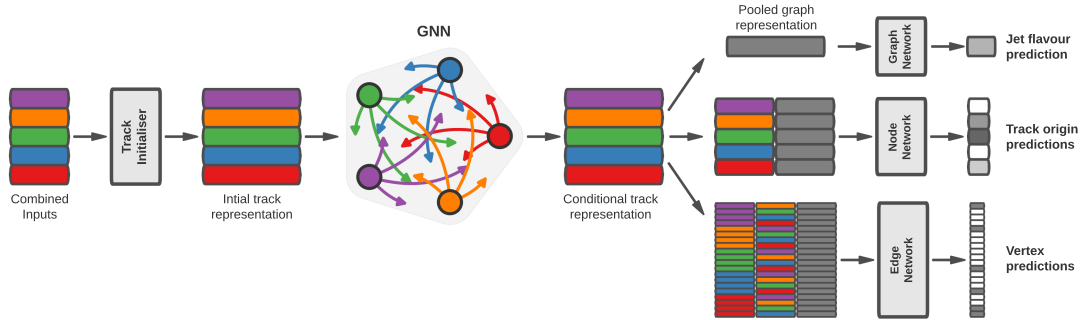


Figure 5.3: The network architecture of GN1. Inputs are fed into a per-track initialisation network, which outputs an initial latent representation of each track. These representations are then used to populate the node features of a fully connected graph network. After the graph network, the resulting node representations are used to predict the jet flavour, the track origins, and the track-pair vertex compatibility.

A fully connected graph is built from the outputs of the track initialisation network, such that each node in the graph neighbours every other node. Each node h_i in the graph corresponds to a single track in the jet, and is characterised by a feature vector, or representation. The per-track output representations from the initialisation networks are used to populate the initial feature vectors of each node in the graph. In each layer of the graph network, output node representations h'_i are computed by aggregating the features of h_i and neighbouring nodes \mathcal{N}_i as described in Ref. [72]. First, the feature vectors of each node are fed into a fully connected layer \mathbf{W} , to produce an updated representation of each node $\mathbf{W}h_i$. These updated feature vectors are used to compute edge scores $e(h_i, h_j)$ for each node pair,

$$e(h_i, h_j) = \mathbf{a}^\top \theta [\mathbf{W}h_i \oplus \mathbf{W}h_j], \quad (5.1)$$

where \oplus denotes vector concatenation, θ is a non-linear activation function, and \mathbf{a} is a second fully connected layer. These edge scores are then used to calculate attention weights a_{ij} for each pair of nodes using the softmax function over the edge scores

$$a_{ij} = \text{softmax}_j [e(h_i, h_j)]. \quad (5.2)$$

Finally, the updated node representation h'_i is computed by taking the weighted sum over each updated node representation $\mathbf{W}h_i$, with weights a_{ij}

$$h'_i = \sigma \left[\sum_{j \in \mathcal{N}_i} a_{ij} \cdot \mathbf{W}h_j \right]. \quad (5.3)$$

The above set of operations constitute a single graph network layer. Three such layers are stacked to construct the graph network, representing a balance between achieving optimal performance and preventing overtraining. The final output node feature vectors from the network are representations of each track that are conditional on the other tracks in the jet. The output representation for each track is combined using a

weighted sum to construct a global representation of the jet, where the attention weights for the sum are learned during training. Three separate fully connected feedforward neural networks are then used to independently perform the different classification objectives of GN1. Each of the objectives makes use of the global representation of the jet. A summary of the different classification networks used for the various training objectives is shown in Table 5.3.

Table 5.3: A summary of GN1’s different classification networks used for the different training objectives. The hidden layers column contains a list specifying the number of neurons in each layer.

Network	Hidden layers	Output size
Node classification network	128, 64, 32	7
Edge classification network	128, 64, 32	1
Graph classification network	128, 64, 32, 16	3

A node classification network, which takes as inputs the features from a single output node from the graph network and the global jet representation, predicts the track truth origin, as defined in Table 5.2. This network has three hidden layers containing 128, 64 and 32 neurons respectively, and an output size of seven, corresponding to the seven different truth origins.

An edge classification network, which takes as inputs the concatenated representations from each pair of tracks and the global jet representation, is used to predict whether the tracks in the track-pair belong to a common vertex. The edge network has three hidden layers containing 128, 64 and 32 neurons respectively, and a single output, which is used to perform binary classification of the track-pair compatibility. These predictions are used for the auxiliary training objectives discussed in Section 5.4.2.

A graph classification network takes only the global jet representation as an input, and predicts the jet flavour. The graph classification network is comprised of four fully connected hidden layers with 128, 64, 32 and 16 neurons respectively, and has three outputs corresponding to the b -, c - and light- jet classes.

5.4.4 Training

The full GN1 training procedure minimises the total loss function L_{total} , defined in Eq. (5.4). This loss is composed of three terms: L_{jet} , the categorical cross entropy loss over the different jet flavours; L_{vertex} , the binary track-pair compatability cross entropy loss averaged over all track-pairs; and L_{track} , the categorical cross entropy loss for the track origin prediction. L_{vertex} is computed by averaging over all track-pairs in the batch, and L_{track} is computed by averaging over all tracks in the batch.

$$L_{\text{total}} = L_{\text{jet}} + \alpha L_{\text{vertex}} + \beta L_{\text{track}} \quad (5.4)$$

The different losses converge to different values during training, reflective of differences in the relative difficulty of the various objectives. As such, L_{vertex} and L_{track} are weighted by $\alpha = 1.5$ and $\beta = 0.5$ respectively to ensure they converge to similar values, giving them an equal weighting towards L_{total} . The values of α and β also ensure that L_{jet} converges to a larger value than L_{vertex} and L_{track} , reflecting the primary importance of the jet classification objective. In practice, the final performance of the model was not sensitive to modest variations in the loss weights α and β , or to pre-training using L_{total} and fine tuning on the jet classification task only. As there was a significant variation in the relative frequency of tracks of different origins, the contribution of each origin class to L_{track} was weighted by the inverse of the frequency of their occurrence. In L_{vertex} , the relative class weight in the loss for track-pairs where both tracks are from either a b - or c -hadron is increased by a factor of two as compared with other track-pairs.

The track classification and vertexing objectives are supplementary to the jet classification objective and trainings can be performed with either the node or edge networks, or both, removed, as discussed in Section 5.5.3. In these cases, the corresponding losses L_{vertex} and L_{track} are removed from the calculation of L_{total} . The resulting trainings demonstrate how useful the different auxiliary training objectives are for the primary jet classification objective.

GN1 trainings are run for 100 epochs on 4 NVIDIA V100 GPUs, taking around 25 mins to complete each epoch over the training sample of 30 million jets described in Section 5.3.1. The Adam optimiser [74] with an initial learning rate of $1\text{e-}3$,

and a batch size of 4000 jets (spread across the 4 GPUs) was used. Typically the validation loss, calculated on 500k jets, stabilised after around 60 epochs. The epoch that minimized the validation loss was used for evaluation. GN1 has been integrated into the ATLAS software [36] using ONNX [75], and jet flavour predictions for the test sample are computed using the ATLAS software stack.

5.5 Results

The performance of the GN1 tagger is evaluated for both b -tagging and c -tagging use cases, and for both jets with $20 < p_T < 250$ GeV from the $t\bar{t}$ sample and jets with $250 < p_T < 5000$ GeV from the Z' sample. Performance is compared to the DL1r tagger [51], which has been retrained on 75 million jets from the same samples as GN1. The input RNNIP tagger [54] to DL1r has not been retrained.

The taggers predict the probability that a jet belongs to the b -, c - and light- classes. To use the model for b -tagging, these probabilities are combined into a single score D_b , defined as

$$D_b = \log \frac{p_b}{(1 - f_c)p_l + f_cp_c}, \quad (5.5)$$

where f_c is a free parameter that determines the relative weight of p_c to p_l in the score D_b , controlling the trade-off between c - and light-jet rejection performance. This parameter is set to a value of $f_c = 0.018$ for the DL1r model, obtained through an optimisation procedure designed to maximise the c - and light-jet rejection of DL1r [51]. For the GN1 models a value of $f_c = 0.05$ is used, based on a similar optimisation procedure. The choice of f_c is arbitrary, with the different optimised values reflecting the relative c - versus light-jet rejection performance of the various taggers. A fixed-cut working point (WP) defines the corresponding selection applied to the tagging discriminant D_b in order to achieve a given inclusive efficiency on the $t\bar{t}$ sample.

The technical implementation of GN1 results in any jet with no associated tracks or exactly one associated track to be classified as a light-jet. The impact of this on the

tagging performance of GN1 was found to be negligible, with 0.12% of b -jets in the $t\bar{t}$ sample and 0.02% of b -jets in the Z' sample affected. Of those, 89% of the b -jets in the $t\bar{t}$ sample and 98% of the b -jets in the Z' sample are classified as light-jets by DL1r at the 70% $t\bar{t}$ WP.

A comparison of the b -tagging discriminant D_b between DL1r and GN1 is given in Fig. 5.4. The shapes of the distributions are broadly similar for b -, c - and light-jets, however, the GN1 model shifts the b -jet distribution to higher values of D_b in the regions with the best discrimination. The GN1 c -jet distribution is also shifted to lower values of D_b when compared with DL1r, enhancing the separation and indicating that GN1 will improve c -jet rejection when compared with DL1r.

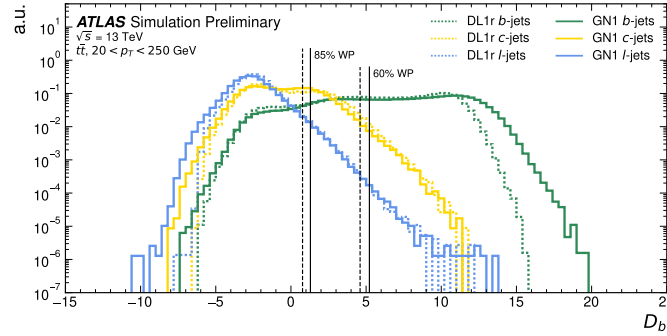


Figure 5.4: Comparison between the DL1r and GN1 b -tagging discriminant D_b for jets in the $t\bar{t}$ sample. The 85% WP and the 60% WP are marked by the solid (dashed) lines for GN1 (DL1r), representing respectively the loosest and tightest WPs used by analyses. A value of $f_c = 0.018$ is used in the calculation of D_b for DL1r and $f_c = 0.05$ is used for GN1. The distributions of the different jet flavours have been normalised to unity area.

5.5.1 b -tagging Performance

The performance of a b -tagging algorithm is quantified by its power to reject c - and light-jets for a given b -jet tagging efficiency, or WP. In order to compare the b -tagging performance of the different taggers for the b -jet tagging efficiencies in the range typically used by analyses, the corresponding c - and light-jet rejection rates are displayed in Figs. 5.5 and 5.6 for jets in the $t\bar{t}$ and Z' samples respectively. Four standard WPs with b -jet tagging efficiencies of 60%, 70%, 77% and 85% are used by physics analyses depending on their specific signal and background requirements. These WPs are defined using jets in the $t\bar{t}$ sample only. The b -jet tagging efficiencies

for jets in the Z' sample are lower than the corresponding WPs calculated in the $t\bar{t}$ sample, due to the much higher jet p_T range in the Z' sample. For instance the WP defined to provide a 70% b -jet tagging efficiency on the $t\bar{t}$ sample results in a b -jet tagging efficiency of $\sim 30\%$ on the Z' sample. To account for this, the range of b -jet tagging efficiencies displayed in Fig. 5.6 is chosen to span the lower values achieved in the Z' sample.

For jets in the $t\bar{t}$ sample with $20 < p_T < 250$ GeV, GN1 demonstrates considerably better c - and light-jet rejection compared with DL1r across the full range of b -jet tagging efficiencies probed. The relative improvement depends on the b -jet tagging efficiency, with the largest improvements found at lower values. At a b -jet tagging efficiency of 70%, the c -rejection improves by a factor of ~ 2.1 and the light-jet rejection improves by a factor of ~ 1.8 with respect to DL1r. For high- p_T jets in the Z' sample with $250 < p_T < 5000$ GeV, GN1 also brings considerable performance improvements with respect to DL1r across the range of b -jet tagging efficiencies studied. Again, the largest relative improvement in performance comes at lower b -jet tagging efficiencies. At a b -jet tagging efficiency of 30%, GN1 improves the c -rejection by a factor of ~ 2.8 and the light-jet rejection by a factor of ~ 6 . An increasing statistical uncertainty due to the high rejection of background affects the comparison at lower b -jet tagging efficiencies. It is estimated that for a b -jet tagging efficiency of 70% in the $t\bar{t}$ sample, $\sim 5\%$ ($\sim 30\%$) of the relative improvement in the c -jet (light-jet) rejection comes from loosening the track selection and for a b -jet tagging efficiency of 30% in the Z' the corresponding number is $\sim 10\%$ for both c -jets and light-jets. Given the sophisticated exploitation of low-level information, further studies are needed to confirm if the performance gain is also observed in experimental data.

The GN1 Lep variant shows improved performance with respect to the baseline GN1 model, demonstrating the additional jet flavour discrimination power provided by the leptonID track input. For jets in the $t\bar{t}$ sample, the relative c -rejection improvement with respect to DL1r at the 70% b -jet WP increases from a factor of ~ 2.1 for GN1 to a factor of ~ 2.8 for GN1 Lep. The improvement in light-jet rejection also increases from a factor of ~ 1.8 to ~ 2.5 at this WP. For jets in the Z' sample, the relative c -rejection (light-jet rejection) improvement with respect to DL1r increases from a factor of ~ 2.8 to ~ 3 (~ 6 to ~ 7.5) at a b -jet tagging efficiency of 30%. As shown in Fig. 5.7, the greatest improvement of GN1 Lep over GN1 is seen at low p_T .

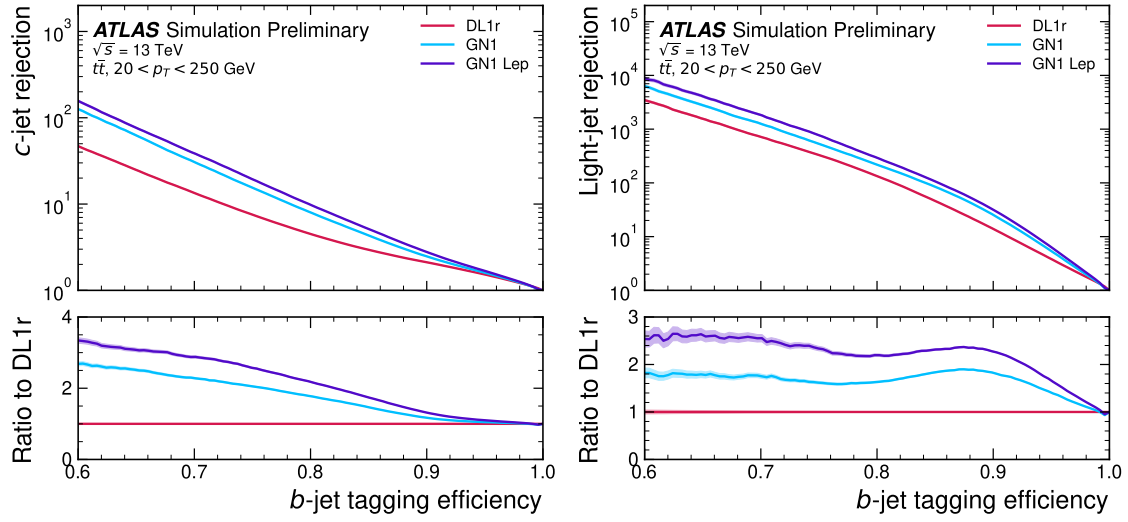


Figure 5.5: The c -jet (left) and light-jet (right) rejections as a function of the b -jet tagging efficiency for jets in the $t\bar{t}$ sample with $20 < p_T < 250$ GeV. The ratio with respect to the performance of the DL1r algorithm is shown in the bottom panels. A value of $f_c = 0.018$ is used in the calculation of D_b for DL1r and $f_c = 0.05$ is used for GN1 and GN1 Lep. Binomial error bands are denoted by the shaded regions. At b -jet tagging efficiencies less than $\sim 75\%$, the light-jet rejection becomes so large that the effect of the low number of jets is visible. The lower x -axis range is chosen to display the b -jet tagging efficiencies usually probed in these regions of phase space.

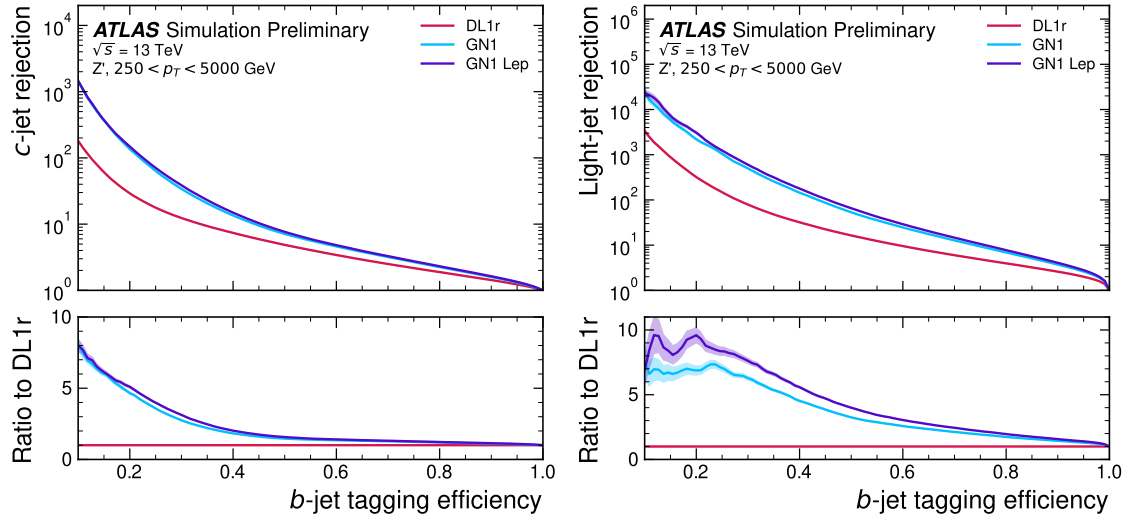


Figure 5.6: The c -jet (left) and light-jet (right) rejections as a function of the b -jet tagging efficiency for jets in the Z' sample with $250 < p_T < 5000$ GeV. The ratio with respect to the performance of the DL1r algorithm is shown in the bottom panels. A value of $f_c = 0.018$ is used in the calculation of D_b for DL1r and $f_c = 0.05$ is used for GN1 and GN1 Lep. Binomial error bands are denoted by the shaded regions. At b -jet tagging efficiencies less than $\sim 20\%$, the light-jet rejection becomes so large that the effect of the low number of jets is visible. The lower x -axis range is chosen to display the b -jet tagging efficiencies usually probed in these regions of phase space.

The performance of the taggers is strongly dependent on the jet p_T . Charged particle reconstruction is particularly challenging within high- p_T jets [38]. The multiplicity of fragmentation particles increases as a function of p_T , while the number of particles from heavy flavour decays stays constant. Collimation of particles inside the jet increases and approaches the granularity of the tracking detectors, making it difficult to resolve the trajectories of different particles. Furthermore, at high p_T , heavy flavour hadrons will travel further into the detector before decaying. For hadrons which traverse one or more layers of the ID before decaying, the corresponding decay tracks may pick up incorrect hits, left by the hadron itself or fragmentation particles, in the inner layers of the detector, reducing the accuracy of the reconstructed track parameters. These factors contribute to a reduced reconstruction efficiency for heavy flavour tracks, and a general degradation in quality of tracks inside the core of a jet, which in turn reduces the jet classification performance.

In order to study how the b -jet tagging efficiency of the taggers varies as a function of jet p_T , the b -jet tagging efficiency as a function of p_T for a fixed light-jet rejection of 100 in each bin is shown in Fig. 5.7. For jets in the $t\bar{t}$ sample, at a fixed light-jet rejection of 100, GN1 improves the b -jet tagging efficiency by approximately 4% across all jet p_T bins. GN1 Lep shows improved performance with respect to GN1, in particular at lower p_T , with the relative increase in the b -jet tagging efficiency going from 4% to 8%. For jets in the Z' sample, GN1 has a higher b -jet tagging efficiency than DL1r across the p_T range, with the largest relative improvement in performance, approximately a factor of 2, found at jet $p_T > 2$ TeV. GN1 outperforms DL1r across the entire jet p_T spectrum studied. The performance was also evaluated as a function of the average number of pileup interactions in an event, and was found to have no significant dependence on this quantity.

5.5.2 c -tagging Performance

Since GN1 does not rely on any manually optimised low-level tagging algorithms, which may not have been optimised for c -tagging, tagging c -jets presents a compelling use case for GN1. To use the model for c -tagging, the output probabilities are combined into a single score D_c , defined similarly to Eq. (5.5) as

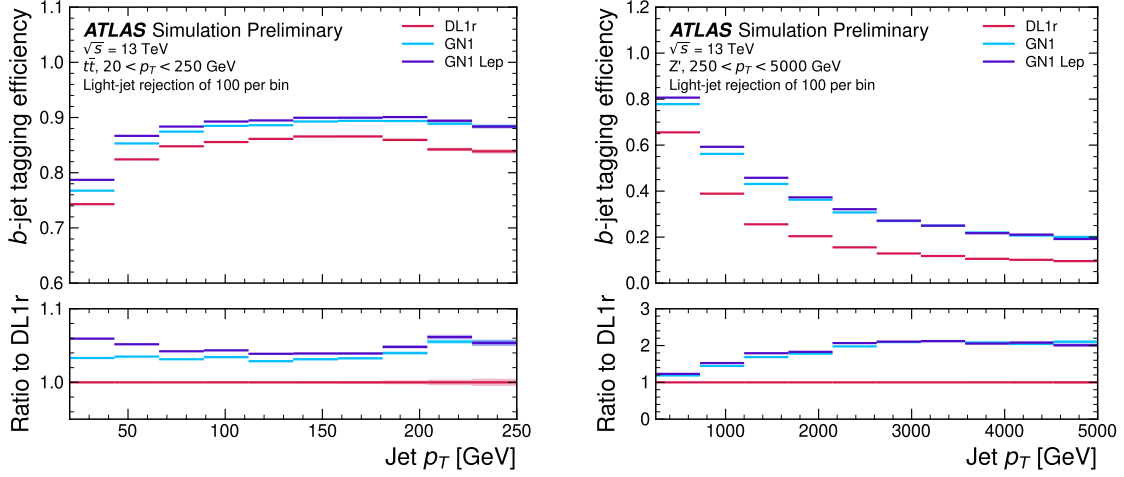


Figure 5.7: The b -jet tagging efficiency for jets in the $t\bar{t}$ sample (left) and jets in the Z' sample (right) as a function of jet p_T with a fixed light-jet rejection of 100 in each bin. A value of $f_c = 0.018$ is used in the calculation of D_b for DL1r and $f_c = 0.05$ is used for GN1 and GN1 Lep. Binomial error bands are denoted by the shaded regions.

$$D_c = \log \frac{p_c}{(1 - f_b)p_l + f_b p_b}. \quad (5.6)$$

A value of $f_b = 0.2$ is used for all models. Similar to Section 5.5.1, performance of the different taggers is compared by scanning through a range of c -jet tagging efficiencies and plotting the corresponding b - and light-jet rejection rates. As in Section 5.5.1, WPs are defined using jets in the $t\bar{t}$ sample. Standard c -jet tagging efficiency WPs are significantly lower in comparison with the b -tagging WPs in order to maintain reasonable b - and light-jet rejection rates. This is reflected in the range of c -jet tagging efficiencies used in Figs. 5.8 and 5.9. In Fig. 5.8, which displays the c -tagging performance of the models on the jets in the $t\bar{t}$ sample, GN1 performs significantly better than DL1r. The b - and light-jet rejection improve most at lower c -jet tagging efficiencies, with both background rejections increasing by a factor of 2 with respect to DL1r at a c -jet tagging efficiency of 25%. GN1 Lep outperforms GN1, with the b -rejection (light-jet rejection) relative improvement increasing from a factor of 2 to 2.1 (2 to 2.3) at the 25% c -jet WP. Fig. 5.9 shows the c -tagging performance on the jets in the Z' sample. Both GN1 and GN1 Lep perform similarly,

improving the b -rejection by 60% and the light-jet rejection by a factor of 2 at the 25% c -jet WP.

5.5.3 Ablations

Several ablations, the removal of components in the model to study their impact, are carried out to determine the importance of the auxiliary training objectives of GN1 to the overall performance. The “GN1 No Aux” variant retains the primary jet classification objective, but removes both track classification and vertexing auxiliary objectives (see Section 5.4.2) and as such only minimises the jet classification loss. The “GN1 TC” variant includes track classification but not vertexing, while “GN1 Vert” includes vertexing, but not track classification.

For jets in both the $t\bar{t}$ and Z' samples, the models without one or both of the auxiliary objectives display significantly reduced c - and light-jet rejection when compared with the baseline GN1 model, as shown in Figs. 5.10 and 5.11. For jets in the $t\bar{t}$ sample, the performance of GN1 No Aux is similar to DL1r, while GN1 TC and GN1 Vert perform similarly to each other. For jets in the Z' sample, the GN1 No Aux model shows a clear improvement in c - and light-jet rejection when compared with DL1r at lower b -jet tagging efficiencies. Similar to jets in the $t\bar{t}$ sample, GN1 TC and GN1 Vert perform similarly, and bring large gains in background rejection when compared with GN1 No Aux, but the combination of both auxiliary objectives yields the best performance.

It is notable that the GN1 No Aux model matches or exceeds the performance of DL1r without the need for inputs from the low-level algorithms. This indicates that the performance improvements enabled by GN1 appear to be able to compensate for the removal of the low-level algorithm inputs. The GN1 TC and GN1 Vert variants each similarly outperform DL1r, demonstrating that both contribute to the overall high performance of the baseline model.

5.5.4 Inclusion of Low-Level Vertexing Algorithms

GN1 does not include inputs from low-level tagging algorithms, including the vertexing tools SV1 and JetFitter [52]. Since these algorithms are known to improve the

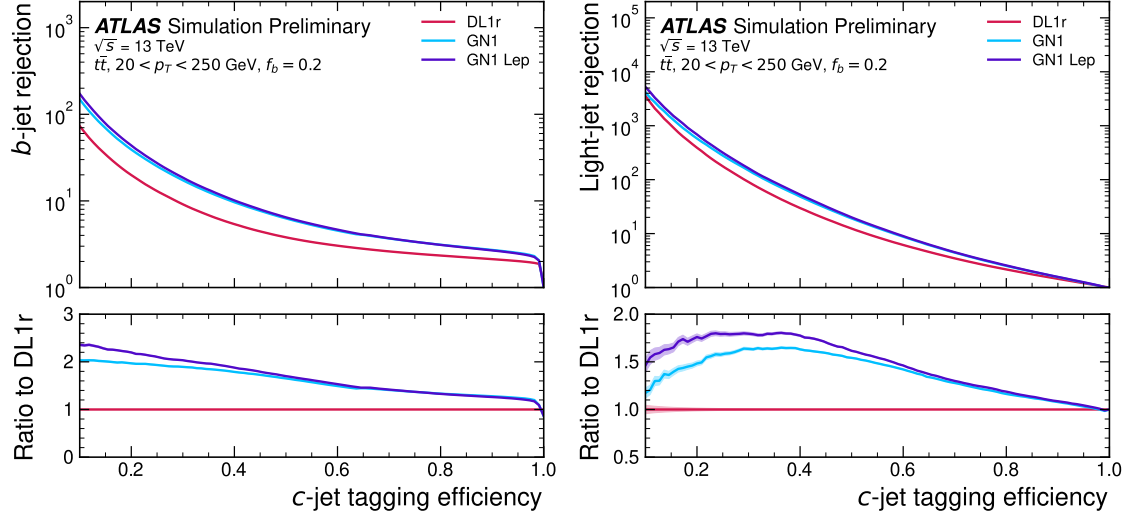


Figure 5.8: The b -jet (left) and light-jet (right) rejections as a function of the c -jet tagging efficiency for $t\bar{t}$ jets with $20 < p_T < 250$ GeV. The ratio to the performance of the DL1r algorithm is shown in the bottom panels. Binomial error bands are denoted by the shaded regions. At c -jet tagging efficiencies than $\sim 25\%$, the light-jet rejection becomes so large that the effect of the low number of jets is visible. The lower x -axis range is chosen to display the c -jet tagging efficiencies usually probed in these regions of phase space.

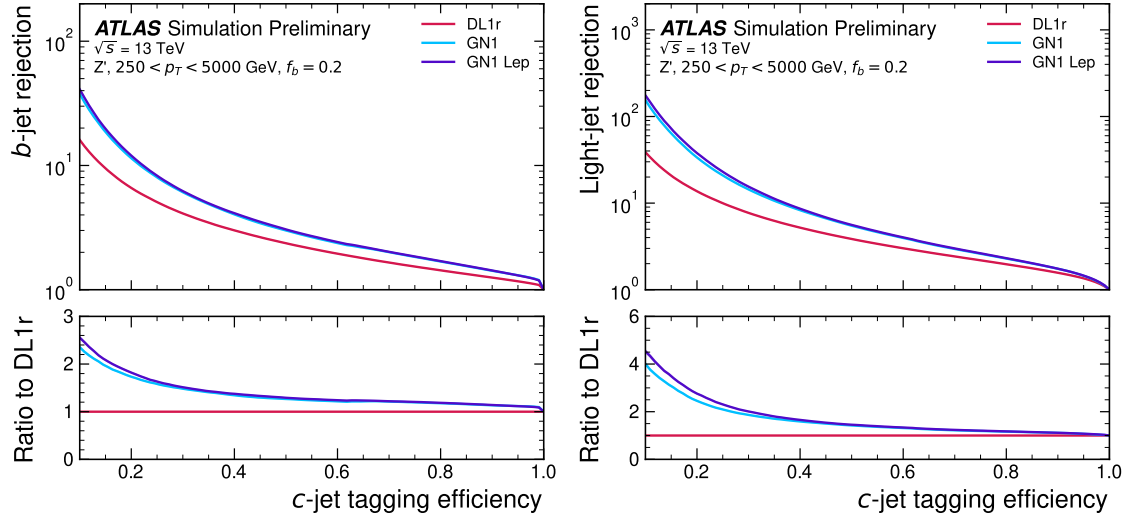


Figure 5.9: The b -jet (left) and light-jet (right) rejections as a function of the c -jet tagging efficiency for Z' jets with $250 < p_T < 5000$ GeV. The ratio to the performance of the DL1r algorithm is shown in the bottom panels. Binomial error bands are denoted by the shaded regions. The lower x -axis range is chosen to display the c -jet tagging efficiencies usually probed in these regions of phase space.

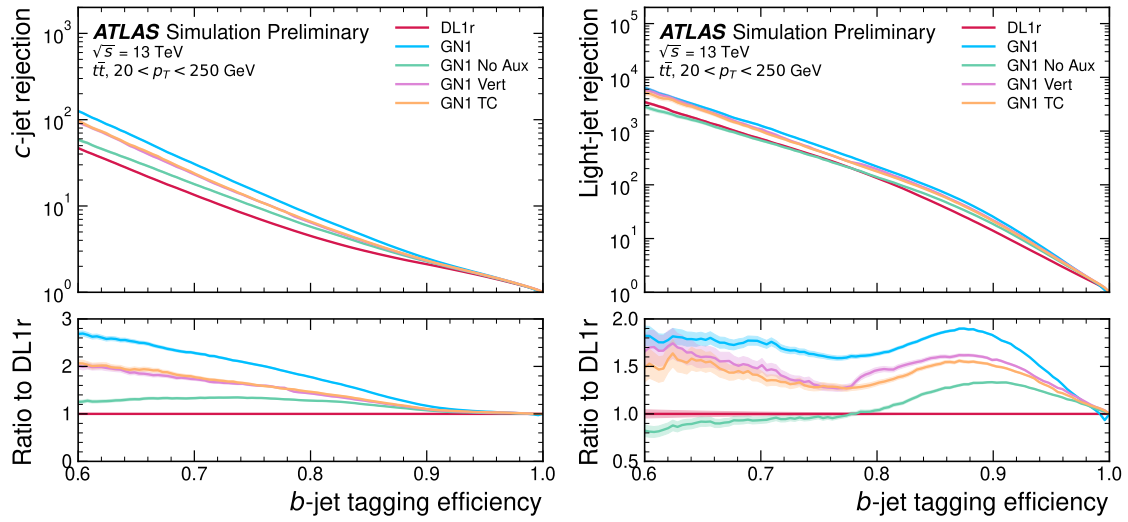


Figure 5.10: The c -jet (left) and light-jet (right) rejections as a function of the b -jet tagging efficiency for $t\bar{t}$ jets with $20 < p_T < 250$ GeV, for the nominal GN1, in addition to configurations where no (GN1 No Aux), only the track classification (GN1 TC) or only the vertexing (GN1 Vert) auxiliary objectives are deployed. The ratio to the performance of the DL1r algorithm is shown in the bottom panels. A value of $f_c = 0.018$ is used in the calculation of D_b for DL1r and $f_c = 0.05$ is used for GN1. Binomial error bands are denoted by the shaded regions. At b -jet tagging efficiencies less than $\sim 65\%$, the light-jet rejection become so large that the effect of the low number of jets are visible. The lower x -axis range is chosen to display the b -jet tagging efficiencies usually probed in these regions.

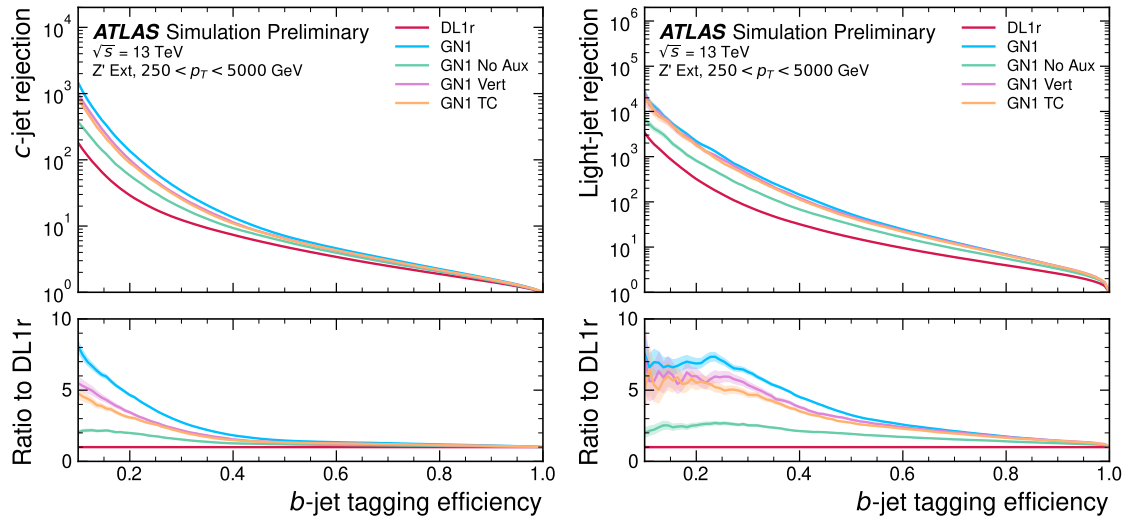


Figure 5.11: The c -jet (left) and light-jet (right) rejections as a function of the b -jet tagging efficiency for Z' jets with $250 < p_T < 5000$ GeV, for the nominal GN1, in addition to configurations where no (GN1 No Aux), only the track classification (GN1 TC) or only the vertexing (GN1 Vert) auxiliary objectives are deployed. The ratio to the performance of the DL1r algorithm is shown in the bottom panels. A value of $f_c = 0.018$ is used in the calculation of D_b for DL1r and $f_c = 0.05$ is used for GN1. Binomial error bands are denoted by the shaded regions. At b -jet tagging efficiencies less than $\sim 25\%$, the light-jet rejection become so large that the effect of the low number of jets are visible. The lower x -axis range is chosen to display the b -jet tagging efficiencies usually probed in these regions.

performance of DL1r, it was feasible that their inclusion in GN1 may further improve on the performance of the GN1 models. In a dedicated training of GN1 the SV1 and JetFitter tagger outputs were added to the GN1 jet classification network as an input, similar to their use in DL1r. These outputs include information on the reconstructed vertices, including the number of vertices, the vertex mass, displacement, and other properties. In addition, the index of the reconstructed SV1 or JetFitter vertices were included as two track-level inputs to GN1. The jet classification performance of this GN1 model was not significantly different to the baseline model, and in some cases the performance was slightly reduced. A dedicated look at the vertexing performance of GN1 with some comparisons to SV1 and JetFitter is found in Section 5.5.5

5.5.5 Vertexing Performance

From the track-pair vertex prediction described in Section 5.4.2, tracks can be partitioned into compatible groups representing vertices (see [50]). As such, GN1 is able to be used to perform vertex “finding”, but not vertex “fitting”, i.e. the reconstruction of a vertex’s properties, which currently still requires the use of a dedicated vertex fitter. In order to study the performance of the different vertexing tools inside b -jets, the truth vertex label of the tracks, discussed in Section 5.4.2, are used. To estimate the efficiency with which GN1 manages to find vertices inclusively, vertices from GN1 containing tracks identified as coming from a b -hadron are merged together and compared to the inclusive truth decay vertices that result from a b -hadron decay (where if there are multiple distinct truth vertices from a b -hadron decay they are also merged together). Vertices are compared with the target truth vertex and the number of correctly and incorrectly assigned tracks is computed. Since secondary vertex information is only recovered for reconstructed tracks, an efficiency of 100% here denotes that all possible secondary vertices are recovered given the limited track reconstruction efficiency. A vertex is considered matched if it contains at least 65% of the tracks in the corresponding truth vertex, and has a purity of at least 50%. GN1 manages to achieve an inclusive reconstruction efficiency in b -jets of $\sim 80\%$, demonstrating that it effectively manages to identify the displaced vertices from b -hadron decays.

More detail

In order to study the performance of the different vertexing tools inside b -jets, the truth vertex label of the tracks, discussed in Section 5.4.2, is used. The reconstructed vertices from GN1, SV1 and JetFitter are compared to the target truth vertices in order to calculate the efficiencies of the different vertexing tools. Since secondary vertex information is only recovered for reconstructed tracks, an efficiency of 100% here denotes that all possible secondary vertices are recovered given the limited track reconstruction efficiency.

There are several caveats to a comparison of the vertexing tools which are a result of the different approaches they take to vertexing. SV1 and JetFitter are designed to only find secondary vertices in the jet, whereas GN1 is also trained to determine which tracks in the jet belong to the primary vertex (the vertex of the hard scatter pp interaction). To account for this the GN1 vertex with the largest number of predicted primary tracks is excluded from the vertex finding efficiency calculation. While JetFitter and GN1 aim to resolve each displaced vertex inside the jet, such that secondary vertices from b -hadron decays are found separately to tertiary vertices from $b \rightarrow c$ decay chains, SV1 by design attempts to find a single inclusive vertex per jet. This inclusive vertex groups inclusive b -hadron decays. These are tracks from the b -hadron decay itself (FromB) and tracks from $b \rightarrow c$ decays (FromBC). In order to fairly compare the performance of the different tools, both the exclusive and inclusive vertex finding efficiency is studied. For the exclusive vertex finding case JetFitter and GN1 can be directly compared, while a comparison with SV1 is not possible due to aforementioned design constraints. The inclusive vertex finding performance of all three tools can be compared using the procedure outlined below.

The starting point for the secondary vertex finding efficiency in both the exclusive and inclusive cases is to select truth secondary vertices are those containing only inclusive b -hadron decays to be considered as initial targets. For exclusive vertex finding, these truth secondary vertices can be used directly as the denominator for the efficiency calculation. Meanwhile for the inclusive efficiency all such truth secondary vertices in the jet are merged into a single inclusive target vertex. Correspondingly, for the inclusive vertex finding case, the vertices found by JetFitter are merged into a single vertex, and the vertices found by GN1 with at least one predicted inclusive

b -hadron decay track are also merged similarly. SV1 does not require any vertex merging.

Next, in both cases for each truth secondary vertex, vertices in the jet found by the different vertexing tools are compared with the target truth vertex. The number of correctly and incorrectly assigned tracks is computed. In order to call a vertex efficient, it is required to contain at least 65% of the tracks in the corresponding truth vertex, and to have a purity of at least 50%. Single track vertices are required to have a purity of 100%. Additionally, for GN1 only, at least one track in the vertex is required to have a predicted heavy flavour origin.

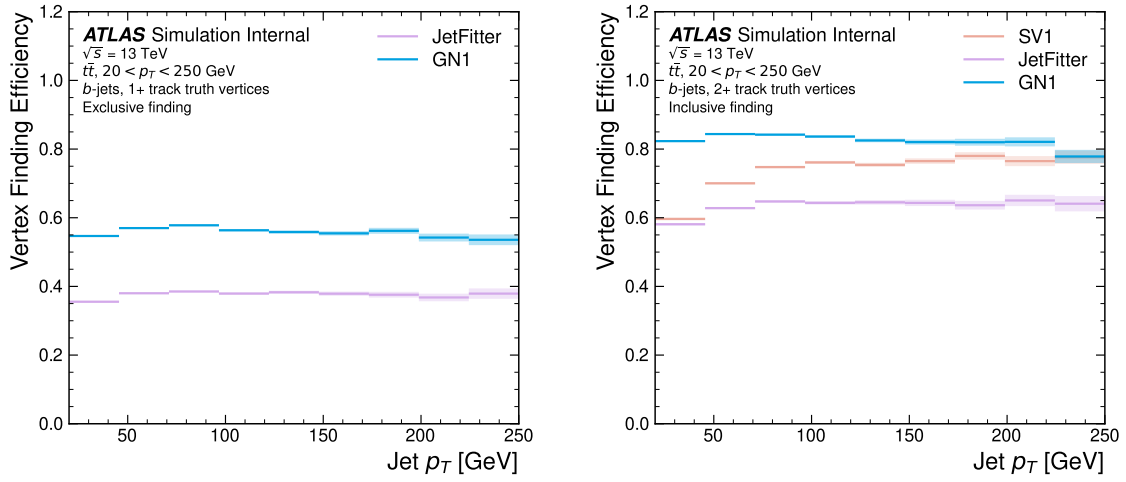


Figure 5.12: Vertex finding efficiency as a function of jet p_T for b -jets in the $t\bar{t}$ sample using the exclusive (left) and inclusive (right) vertex finding approaches. Efficient vertex finding requires the recall of at least 65% of the tracks in the truth vertex, and allows no more than 50% of the tracks to be included incorrectly. Binomial error bands are denoted by the shaded regions.

Vertex finding efficiencies for b -jets in the $t\bar{t}$ sample are displayed as a function of p_T separately for the inclusive and exclusive approaches in Fig. 5.12. For b -jets in the $t\bar{t}$ sample with $20 < p_T < 250$ GeV, the exclusive vertex finding efficiency of JetFitter and GN1 is relatively flat as a function of p_T . Of the truth secondary vertices in this p_T region, JetFitter efficiently finds approximately 40% and GN1 finds approximately 55%. When finding vertices inclusively the vertex finding efficiency is generally higher. An increased dependence on p_T is also visible for JetFitter and SV1. As the jet p_T increases from 20 GeV to 100 GeV, the efficiency of JetFitter increases from 55% to 65%. In the same range, the efficiency of SV1 increases from 55% to 75%. GN1

displays less dependence on p_T than JetFitter and SV1, efficiently finding upwards of 80% of vertices in b -jets in this p_T region. For b -jets with $p_T > 100$ GeV, JetFitter finds approximately 65% of vertices, SV1 finds approximately 75% of vertices, and GN1 finds approximately 80% of vertices.

For b -jets in the Z' sample, the vertex finding efficiency drops steeply with increasing p_T up until $p_T = 3$ TeV. GN1 outperforms SV1 and JetFitter across the p_T spectrum. In the first bin, the efficiency of GN1 is 75%, while the efficiencies of SV1 and JetFitter are around 60%. The efficiency of SV1 drops rapidly to almost zero above 3 TeV, while JetFitter and GN1 retain approximately 30% efficiency. Fig. 5.13 compares the exclusive vertex finding efficiencies of JetFitter and GN1 for multitrack vertices. JetFitter finds 45-50% of vertices in b -jets in the $t\bar{t}$ sample, while GN1 finds 60-65%. For b -jets in the Z' sample, JetFitter finds 35% of vertices in the first bin, dropping to 20% of vertices above 2 TeV. GN1 finds 55% of vertices in the first bin, dropping to 30% above 2 TeV.

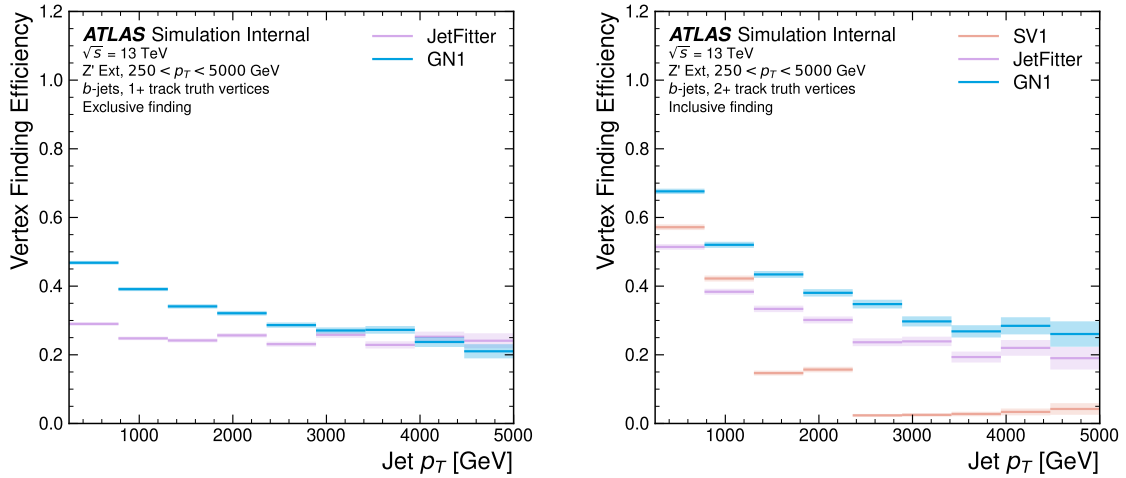


Figure 5.13: Inclusive vertex finding efficiency for multitrack truth vertices in b -jets in the $t\bar{t}$ sample (left) and jets in the Z' sample (right) as a function of jet p_T . Efficient vertex finding requires the recall of at least 65% of the tracks in the truth vertex, and allows no more than 50% of the tracks to be included incorrectly.

5.5.6 Track Classification Performance

As discussed in Section 5.4.2, one of the auxiliary training objectives for GN1 is to predict the truth origin of each track in the jet. Since the equivalent information is

not provided by any of the existing flavour tagging tools, as a benchmark a multi-class classification multilayer perceptron (MLP) is trained on the same tracks used for the baseline GN1 training. The model uses the same concatenated track-and-jet inputs as GN1 (see Section 5.4.1), but processes only a single track at a time. The model is comprised of five densely connected layers with 200 neurons per layer, though the performance was not found to be strongly sensitive to changes in the network structure. To measure the track classification performance, the area under the curve (AUC) of the receiver operating characteristic (ROC) curve is computed for each origin class using a one versus all classification approach. The AUCs for the different truth origin classes are averaged using both an unweighted and a weighted approach. The unweighted mean treats the performance of each class equally, while the weighted mean uses the fraction of tracks from each origin as a weight. As seen in Table 5.4, GN1 outperforms the MLP, both at $20 < p_T < 250$ GeV for jets in the $t\bar{t}$ sample, and at $250 < p_T < 5000$ GeV for jets in the Z' sample. For tracks in jets in the $t\bar{t}$ sample, GN1 can reject 65% of fake tracks while retaining more than 99% of good tracks. The GN1 model has two advantages over the MLP which can explain the performance improvement. Firstly, the mixing of information between tracks, enabled by the fully connected graph network architecture as discussed in Section 5.4.3, is likely to be beneficial since the origins of different tracks within a jet are to some extent correlated. Secondly, the jet classification and vertexing objectives can be considered auxiliary to the track classification task, and may bring improved track classification performance with respect to the standalone MLP.

Fig. 5.14 shows the track origin classification ROC curves for the different track origins for jets in both the $t\bar{t}$ and Z' samples. In order to improve legibility of the figure, the heavy flavour truth origins have been combined weighted by their relative abundance, as have the Primary and OtherSecondary labels. In jets in both the $t\bar{t}$ and Z' samples, the AUC of the different (grouped) origins is above 0.9, representing good classification performance. Fake tracks, followed by pileup tracks, are the easiest to classify in both samples.

Table 5.4: The area under the ROC curves (AUC) for the track classification from GN1, compared to a standard multilayer perceptron (MLP) trained on a per-track basis. The unweighted mean AUC over the origin classes and weighted mean AUC (using as a weight the fraction of tracks from the given origin) is provided. GN1, which uses an architecture that allows track origins to be classified in a conditional manner as discussed in Section 5.4.3, outperforms the MLP model for both $t\bar{t}$ and Z' jets.

		AUC	
		Mean	Weighted
$t\bar{t}$	MLP	0.87	0.89
	GN1	0.92	0.95
Z'	MLP	0.90	0.94
	GN1	0.94	0.96

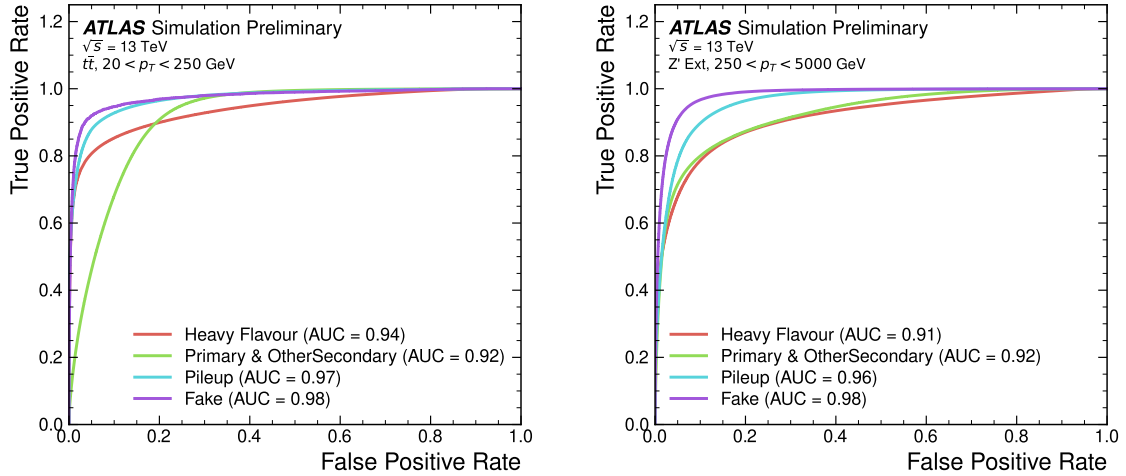


Figure 5.14: ROC curves for the different groups of truth origin labels defined in Table 5.2 for jets in the $t\bar{t}$ sample (left) and jets in the Z' sample (right). The FromB, FromBC and FromC labels have been combined, weighted by their relative abundance, into the Heavy Flavour category, and the Primary and OtherSecondary labels have similarly been combined into a single category. The mean weighted area under the ROC curves (AUC) is similar for both samples.

5.6 Conclusion

A novel jet tagger, GN1, with a graph neural network architecture and trained with auxiliary training targets, is presented and now fully implemented in the ATLAS software. GN1 is shown to improve flavour tagging performance with respect to DL1r, the current default ATLAS flavour tagging algorithm, when compared in simulated collisions. GN1 improves c - and light-jet rejection for jets in the $t\bar{t}$ sample with $20 < p_T < 250$ GeV by factors of ~ 2.1 and ~ 1.8 respectively at a b -jet tagging efficiency of 70% when compared with DL1r. For jets in the Z' sample with $250 < p_T < 5000$ GeV, GN1 improves the c -rejection by a factor of ~ 2.8 and light-jet rejection by a factor of ~ 6 for a comparative b -jet efficiency of 30%. Previous multivariate flavour tagging algorithms relied on inputs from low-level tagging algorithms, whereas GN1 needs no such inputs, making it more flexible. It can be easily fully optimised via a retraining for specific flavour tagging use cases, as demonstrated with c -tagging and high- p_T b -tagging, without the need for time-consuming retuning of the low-level tagging algorithms. The model is also simpler to maintain and study due to the reduction of constituent components. GN1 demonstrates improved track classification performance when compared with a simple per-track MLP and an efficiency of $\sim 80\%$ for inclusive vertex finding in b -jets. The auxiliary track classification and vertex finding objectives are shown to significantly contribute to the performance in the jet classification objective, and are directly responsible for the improvement over DL1r. Further studies need to be undertaken to verify the performance of GN1 on collision data.

5.7 Extensions

5.7.1 Looser Track Selection

Chapter 6

VHbb Boosted Analysis

Paper abstract :

The associated production of a Higgs boson with a W or Z boson decaying into leptons and where the Higgs boson decays to a $b\bar{b}$ pair is measured in the high vector-boson transverse momentum regime, above 250 GeV, with the ATLAS detector. The analysed data, corresponding to an integrated luminosity of 139 fb^{-1} , were collected in proton–proton collisions at the Large Hadron Collider between 2015 and 2018 at a centre-of-mass energy of $\sqrt{s} = 13\text{ TeV}$. The measured signal strength, defined as the ratio of the measured signal yield to that predicted by the Standard Model, is $0.72^{+0.39}_{-0.36}$ corresponding to an observed (expected) significance of 2.1 (2.7) standard deviations. Cross-sections of associated production of a Higgs boson decaying into b quark pairs with a W or Z gauge boson, decaying into leptons, are measured in two exclusive vector boson transverse momentum regions, 250–400 GeV and above 400 GeV, and interpreted as constraints on anomalous couplings in the framework of a Standard Model effective field theory.

Final states containing zero, one and two charged leptons (electrons or muons) are considered. targeting the decays $Z \rightarrow \nu\nu$, $W \rightarrow \ell\nu$ and $Z \rightarrow \ell\ell$. The $H \rightarrow b\bar{b}$ decay is reconstructed using a jet with transverse momentum $p_{\text{T}} > 250\text{ GeV}$, found with the anti- k_{\perp} $R = 1.0$ jet algorithm, and groomed to remove soft and wide-angle radiation and to mitigate contributions from the underlying event and additional proton-proton collisions. Higgs bosons are identified using b -tagged $R = 0.2$ track-jets matched to the groomed large- R calorimeter jets.

6.1 Overview

Lifted from old Vhbb preamble chapter:

The Higgs boson, discovered at the LHC in 2012, is predicted by the standard model to decay primarily to two b quarks, with a branching factor of 0.582 ± 0.007 [24]. Observation of this decay mode was recently reported by ATLAS [44]. Whilst the dominant Higgs production mode at the LHC is gluon-gluon fusion, this mode has an overwhelming QCD multijet background and so sensitivity to the Higgs is low. The $H \rightarrow b\bar{b}$ observation therefore searched for Higgs bosons produced in association with a vector boson (W or Z). This production mechanism results in leptonic final states from the decay of the vector boson, allowing for leptonic triggering, whilst at the same time significantly reducing the multi-jet background.

A closely related analyses now searches for the $H \rightarrow b\bar{b}$ decay of the Higgs boson, produced in association with a vector boson, when the vector boson and Higgs are highly boosted. The full Run-2 dataset is used for a total integrated luminosity of 139 fb^{-1} . The analysis is split into 0-, 1- and 2-lepton channels depending on the number of selected electrons and muons, to target the $ZH \rightarrow \nu\nu b\bar{b}$, $WH \rightarrow \ell\nu b\bar{b}$, $ZH \rightarrow \ell\ell b\bar{b}$ processes, respectively, where ℓ is an electron or muon. In all channels, events are required to have exactly two b -tagged jets, which form the Higgs boson candidate. At least one of the b -tagged jets is required to have p_T greater than 45 GeV. Events are further split into 2-jet or 3-jet categories depending on whether additional, untagged jets are present.

In the 0- and 1-lepton channels, the analysis is further split into signal and control regions. To leading order, there are no additional b -jets in the event other than the two coming from the reconstructed Higgs candidate. For this reason, there is a signal region veto (i.e. events are not accepted into the signal region) for events with additional b -tagged jets in the event. Events with additional b -tagged jets are included in the control region, which is highly pure in $t\bar{t}$ events. The control region is used to constrain the normalisation of the $t\bar{t}$ background.

6.2 Introduction

Lifted from paper:

Since the discovery of the Higgs boson (H) [9–11, 76] with a mass of around 125 GeV [77] by the ATLAS and CMS Collaborations [12, 13] in 2012, the analysis of proton–proton (pp) collision data at centre-of-mass energies of 7 TeV, 8 TeV and 13 TeV delivered by the Large Hadron Collider (LHC) [43] has led to precise measurements of the main production cross-sections and decay rates of the Higgs boson, as well as measurements of its mass and its spin and parity properties. In particular, the observation of the decay of the Higgs boson into b -quark pairs provided direct evidence for the Yukawa coupling of the Higgs boson to down-type quarks [44, 78]. Finally, a combination of 13 TeV results searching for the Higgs boson produced in association with a leptonically decaying W or Z boson established the observation of this production process [44]. A first cross-section measurement as a function of the vector-boson transverse momentum was also carried out by the ATLAS Collaboration [79].

The previous ATLAS analyses [44, 79] in this channel were mainly sensitive to vector bosons with transverse momentum (p_T) in the range of approximately 100–300 GeV. These analyses considered a pair of jets with radius parameter of $R = 0.4$, referred to as small-radius (small- R) jets, to reconstruct the Higgs boson. For higher Higgs boson transverse momenta, the decay products can become close enough that they cannot be reconstructed with two small- R jets. To explore this ‘boosted’ regime, the Higgs boson is reconstructed as a single large- R jet with $R = 1.0$ [80]. This high- p_T regime is particularly interesting due to its sensitivity to physics beyond the Standard Model [81].

This Letter presents a measurement of cross-sections for the associated production of a high transverse momentum Higgs boson that decays into a $b\bar{b}$ pair with a leptonically decaying W or Z boson. The analysis uses pp collision data recorded between 2015 and 2018 by the ATLAS detector [28] during Run 2 at the LHC. This dataset corresponds to an integrated luminosity of 139 fb^{-1} . Events are selected in 0-, 1- and 2-lepton channels, based on the number of reconstructed charged leptons, ℓ (electrons or muons), in the final state to explore the $ZH \rightarrow \nu\nu b\bar{b}$, $WH \rightarrow \ell\nu b\bar{b}$ and $ZH \rightarrow \ell\ell b\bar{b}$ signatures, respectively. The Higgs boson is reconstructed as a single

large- R jet and the b -quarks from its decay as a pair of jets, reconstructed with a p_T -dependent radius parameter, associated with the large- R jet and identified as containing a b -hadron.

The analysis using small- R jets and focusing on slightly lower Higgs boson transverse momentum regions was recently updated with the complete Run 2 dataset [82]. The large- R jet analysis significantly overlaps with the small- R jets analysis. The two results can therefore not be straightforwardly combined.

The dominant background processes after the event selection correspond to the production of $V + \text{jets}$, where V refers to either a W or Z boson, $t\bar{t}$, single-top and dibosons. The signal is extracted from a combined profile likelihood fit to the large- R jet mass, using several signal and control regions. The yield of diboson production VZ with $Z \rightarrow b\bar{b}$ is also measured using the same fit and provides a validation of the analysis. The cross-section measurements are performed within the simplified template cross-section (STXS) framework [24, 83]. These measurements are then used to constrain anomalous couplings in a Standard Model effective field theory (SMEFT) [84].

6.3 Modelling Work

6.3.1 Background

Alternative Samples

As mentioned, alternative samples of $V + \text{jets}$ events was generated using MADGRAPH5_AMC@NLO+PYTHIA8, and the results are compared with the nominal SHERPA 2.2.1 samples. This allows for a comparison of different parton showering and underlying event models, and derivation of the systematic uncertainties on the nominal choice of models.

Source of Uncertainty	Implementation
Renormalisation scale (μ_R)	Internal weights
Factorisation scale (μ_F)	Internal weights
PDF set	Internal weights
α_S value	Internal weights
Parton Shower (PS) models	Alternative samples
Underlying Event (UE) models	Alternative samples
Resummation scale (QSF)	Parameterisation
CKKW merging scale	Parameterisation

Table 6.1: Different sources of uncertainty (i.e. variations in the model) considered for V+jets background, and the corresponding implementation. For each uncertainty, acceptance and shape uncertainties are derived.

Internal Weight Variations

Nominal signal samples generated with SHERPA 2.2.1 include systematic variations of certain modelling parameters which are stored as alternative event weights. The samples contain event weight variations which correspond to variations of renormalisation scale μ_R , and factorisation scale μ_F , of 0.5 and 2 times the nominal value. Additionally stored is event weight variations corresponding to 30 different variations on the PDF and two variations of the strong coupling constant α_S . Variations of α_S were found to have negligible impact on the results of the analysis, and are not discussed further.

Parameterisation Methods

While the inclusion of internal weight variation in MC event generators has decreased simulation times and increased available statistics, there are in SHERPA 2.2.1 currently some sources of systematic uncertainty that are unable to be stored as internal weight variations due to technical limitations. Two such systematics relate to the choice of CKKW matrix element merging scale, and resummation scale (QSF). The generation of high statistics alternative samples is a time consuming process, as is typically not done for all samples for every new generator release. A method to parameterise the systematic variation using one sample, and to then apply this parameterisation to another sample, has been developed by the ATLAS SUSY group [85]. This method

was used to derive CKKW and QSF uncertainties for the nominal SHERPA 2.2.1 sample, using a previous (lower statistic) SHERPA 2.1 alternative sample. The resulting uncertainties were studied and found to be negligible in comparison with systematics from other sources.

Shape Uncertainties

In order to derive shape uncertainties (which as the name suggests affect shapes but not overall normalisations of distributions), the following procedure is carried out. Normalised distributions of the reconstructed Higgs candidate mass m_J are compared for the nominal sample and variations. For each variation, the ratio of the variation to nominal is calculated, and an analytic function is fit to those sources of variation which have a ratio deviating from unity. If different analysis regions or channels show the same pattern of variation, a common uncertainty is assigned. An example of a significant source of uncertainty, arising from choice of factorisation scale μ_R is shown in Fig. 6.1. An exponential function has been fitted to the ratio of the normalised distributions. Two different analysis regions (medium and high p_{T^V} bins) are shown. The difference of the shape of the variation means that two separate uncertainties have to be added in the fit, and applied individually in each p_{T^V} region.

Acceptance Uncertainties

Several different types of acceptance uncertainties have been calculated. These are implemented as nuisance parameters in the fit and for the most part account for the migration of events between different analysis regions. The list acceptance uncertainties relevant to the V+jets processes are given summarised below.

- **Overall normalisation:** only relevant where normalisation cannot be left floating (i.e. determined in the fit).
- **SR-to-CR relative acceptance:** the uncertainty on the normalisation of the signal region due to events migrating between the signal and control regions.

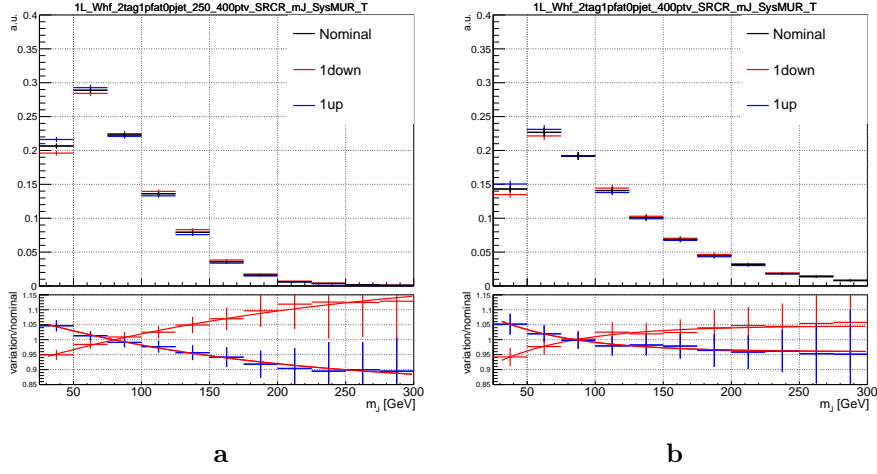


Figure 6.1: Normalised distributions of leading fat jet mass m_J for medium (6.1a) and high (6.1b) p_{T^V} analysis regions for W+heavy-flavour-jets in the 0 lepton channel. Merged in heavy flavours, high and low purity signal regions. The renormalisation scale μ_R has been varied by a factor of 2 (“1up”) and 0.5 (“1down”). An exponential function has been fit to the ratio.

- **HP-to-LP relative acceptance:** the uncertainty on the normalisation of the high-purity (HP) signal region due to events migrating between the high- and low-purity signal regions.
- **Medium-to-high p_{T^V} relative acceptance:** describes any ‘shape’ effect in p_{T^V} distribution, given that the analysis only uses two p_{T^V} bins (medium and high).
- **Flavour relative acceptance:** for each flavour Vxx , where $xx \in \{bc, bl, cc\}$ the ratio of Vxx/Vbb events is calculated. This corresponds to the uncertainty of Vbb events due to the miss-tagging of other flavours Vxx .

The uncertainties on different systematics are summed in quadrature to give a total uncertainty on each region. A summary of the different acceptance uncertainties that were derived in this way and subsequently applied in the fit are given in Table 6.2. An effort has been made, wherever possible, to harmonise similar uncertainties across different analysis regions and channels.

6.3.2 Vector Boson + Jets Modelling

The background processes involving W or Z boson decays into leptons (including those in which the W boson arises from a top-quark decay) are collectively referred to as electroweak (EW), or V+jets, backgrounds. W +jets events are most relevant to the 1-lepton channel via the leptonic decay of $W \rightarrow \ell \nu$. In the event of $W \rightarrow \tau \nu$, and subsequent decay of the τ , or the lack of the successful reconstruction of the e or μ , W +jets can also contribute to the 0-lepton channel. Meanwhile, Z +jets contributes primarily to the 0- and 2-lepton channels via the processes $Z \rightarrow \nu \nu$ and $Z \rightarrow \ell \ell$ respectively.

Modelling is used to predict the outcomes of the analysis and to assess the impact of sources of different systematic uncertainty. Signal and background modelling has primarily consisted of using Monte Carlo (MC) generators to produce simulated events. The uncertainties on the simulated output must be well understood to perform a successful analysis. To achieve this, a set of “nominal” samples are first defined as a reference to which different variations can be compared. The nominal samples are chosen as the best possible representation of the underlying physical process. “Alternative” samples are used to understand the systematic uncertainties on the nominal samples. To generate an alternative sample, some aspect of the model is varied, and the simulation is re-run. A comparison back to the nominal sample gives a handle on the systematic uncertainty associated with the model parameter which was changed. Detailed information can be found in [86]. In order to access uncertainties associated with the use of MC generators, variations of the data are produced using alternative generators or variation of nominal generator parameters. The variation of nominal generator parameters can in certain cases be implemented using internal weight variations stored alongside the nominal events, and in other cases a new independent sample must be generated. The nominal generator used for V+jets events is SHERPA 2.2.1, while MADGRAPH5_AMC@NLO+PYTHIA8 (which uses different parton showering models) is used as an alternative generator. As production of large MC samples is computationally expensive, a feature of state of the art simulation packages is to store some sources of variation as internal event weights, which can be generated alongside the nominal samples, saving computation time. Several sources of uncertainty, summarised in Table 6.1, have been assessed.

V+jets Acceptance Uncertainties				
Boson	W		Z	
Channel	0L	1L	0L	2L
Vbb Norm.	30%	-	-	-
SR/CR	90% [†]	40% [†]	40%	-
HP/LP	18%		18%	-
High/Medium p_T^V	30%	10%*	10%	
Channel Extrap.	20%	-	16%	-
Vbc/Vbb	30%			
Vbl/Vbb	30%			
Vcc/Vbb	20%			
Vcl Norm.	30%			
Vl Norm.	30%			

Table 6.2: V+jets acceptance uncertainties. W +jets SR and CR uncertainties marked with a superscript [†] are correlated. The 1L W +jets H/M uncertainty marked by * is applied as independent and uncorrelated NPs in both HP and LP signal regions. The 0L W +jets Wbb Norm uncertainty is only applied when a floating normalisation for Wbb cannot be obtained from the 1L channel. A 30% uncertainty for Zbb norm is applied in the 1L channel when a floating normalisation for Zbb cannot be obtained from the 0L or 2L channels.

6.3.3 Diboson Modelling

6.4 Fit Studies

6.4.1 Fit Model

A global profile likelihood fit is used to extract the signal strength μ and its significance from the data. This statistical setup treats each bin as a Poisson counting experiment. The combined likelihood over N bins, without considering sources of systematic uncertainty, is given by

$$\mathcal{L}(\mu) = \prod_{i=1}^N \frac{(\mu s_i + b_i)^{n_i}}{n_i!} \exp[-(\mu s_i + b_i)], \quad (6.1)$$

where s_i (b_i) is the expected number of signal (background) events in bin i , and n_i is the number of events observed in data in bin i . The presence of systematic uncertainties which can affect the expected numbers of signal and background events necessitates the addition of nuisance parameters (NPs), θ , to the likelihood. Each source of systematic uncertainty for V+jets samples discussed in the previous section was implemented as a NP θ_j in the fit. The presence of NPs modifies the likelihood as

$$\mathcal{L}(\mu) \rightarrow \mathcal{L}(\mu, \theta) = \mathcal{L}(\mu) \times \mathcal{L}(\theta), \quad s_i \rightarrow s_i(\theta), \quad b_i \rightarrow b_i(\theta), \quad (6.2)$$

where

$$\mathcal{L}(\theta) = \prod_{\theta_j \in \theta} \frac{\exp[-\theta_j^2/2]}{\sqrt{2\pi}}. \quad (6.3)$$

Post-fit m_J distributions in the high-purity medium p_{T^V} regions for the 0- and 2-lepton channels are shown in Fig. 6.2. The plots show large falling backgrounds, predominantly made up of W +jets and Z +jets events, and a signal distribution corresponding to the Standard Model Higgs boson peaking around $m_H = 125$ GeV.

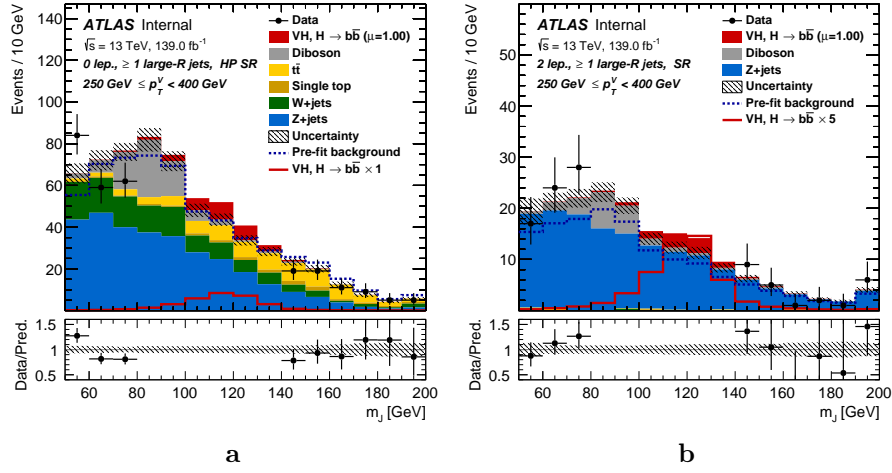


Figure 6.2: Post-fit distributions for the 0-lepton (6.2a) and 2-lepton (6.2b) channels in the high purity medium p_{T^V} region, obtained in the combined conditional $\mu = 1$ fit to data. The last bin of each plot is an overflow bin.

6.5 Conclusion

Work has been carried out as part of the boosted VHbb analysis group to understand, and implement in the global profile likelihood fit, systematic uncertainties on V+jets samples. This background modelling work is an essential part of the success of the analysis. So far the fit has proved stable with the inclusion of the V+jets uncertainties, and detailed studies are now underway to determine the causes behind any observed pulls of the added NPs. Additional work is ongoing to help with the derivation of uncertainties on diboson samples, another important background. The analysis is already advanced, and is now progressing into its final stages. Publication is expected in the new year.

Chapter 7

Conclusion

Appendix A

Combining Multiple Triggers

Bibliography

- [1] A. Buckley, *A class for typesetting academic theses*, 2010.
- [2] L. Morel, Z. Yao, P. Cladé and S. Guellati-Khélifa, *Determination of the fine-structure constant with an accuracy of 81 parts per trillion*, *Nature* **588** (2020) 61.
- [3] T. Sailer, V. Debierre, Z. Harman, F. Heiße, C. König, J. Morgner et al., *Measurement of the bound-electron g -factor difference in coupled ions*, *Nature* **606** (2022) 479.
- [4] CDF collaboration, *Observation of top quark production in $\bar{p}p$ collisions*, *Phys. Rev. Lett.* **74** (1995) 2626 [[hep-ex/9503002](#)].
- [5] D0 collaboration, *Observation of the top quark*, *Phys. Rev. Lett.* **74** (1995) 2632 [[hep-ex/9503003](#)].
- [6] S. W. Herb et al., *Observation of a Dimuon Resonance at 9.5-GeV in 400-GeV Proton-Nucleus Collisions*, *Phys. Rev. Lett.* **39** (1977) 252.
- [7] UA1 collaboration, *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV*, *Phys. Lett. B* **122** (1983) 103.
- [8] DONUT collaboration, *Observation of tau neutrino interactions*, *Phys. Lett. B* **504** (2001) 218 [[hep-ex/0012035](#)].
- [9] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (1964) 321.
- [10] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964) 508.

- [11] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, *Phys. Rev. Lett.* **13** (1964) 585.
- [12] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett. B* **716** (2012) 1 [1207.7214].
- [13] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys. Lett. B* **716** (2012) 30 [1207.7235].
- [14] PARTICLE DATA GROUP collaboration, *Review of Particle Physics*, *PTEP* **2022** (2022) 083C01.
- [15] C. N. Yang and R. L. Mills, *Conservation of isotopic spin and isotopic gauge invariance*, *Phys. Rev.* **96** (1954) 191.
- [16] S. L. Glashow, *Partial Symmetries of Weak Interactions*, *Nucl. Phys.* **22** (1961) 579.
- [17] S. Weinberg, *A Model of Leptons*, *Phys. Rev. Lett.* **19** (1967) 1264.
- [18] A. Salam, *Weak and Electromagnetic Interactions*, *Proceedings of the 8th Nobel symposium, Ed. N. Svartholm, Almquist & Wiskell, 1968*, **Conf. Proc. C680519** (1968) 367.
- [19] T. D. Lee and C. N. Yang, *Question of parity conservation in weak interactions*, *Phys. Rev.* **104** (1956) 254.
- [20] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes and R. P. Hudson, *Experimental test of parity conservation in beta decay*, *Phys. Rev.* **105** (1957) 1413.
- [21] R. L. Garwin, L. M. Lederman and M. Weinrich, *Observations of the failure of conservation of parity and charge conjugation in meson decays: the magnetic moment of the free muon*, *Phys. Rev.* **105** (1957) 1415.
- [22] Y. Nambu, *Quasi-particles and gauge invariance in the theory of superconductivity*, *Phys. Rev.* **117** (1960) 648.
- [23] J. Goldstone, *Field theories with « superconductor » solutions*, *Il Nuovo Cimento (1955-1965)* **19** (1961) 154.

- [24] LHC HIGGS CROSS SECTION WORKING GROUP collaboration, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, [1610.07922](#).
- [25] T. A. Collaboration, *The ALICE experiment at the CERN LHC*, *Journal of Instrumentation* **3** (2008) S08002.
- [26] CMS Collaboration, *The CMS experiment at the CERN LHC*, *JINST* **3** (2008) S08004.
- [27] T. L. Collaboration, *The LHCb detector at the LHC*, *Journal of Instrumentation* **3** (2008) S08005.
- [28] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003.
- [29] A. Collaboration, *Integrated luminosity summary plots for 2011-2012 data taking*, .
- [30] A. Collaboration, *Public ATLAS Luminosity Results for Run-2 of the LHC*, .
- [31] G. C. Strong, *On the impact of selected modern deep-learning techniques to the performance and celerity of classification models in an experimental high-energy physics use case*, [2002.01427](#).
- [32] ATLAS collaboration, *Technical Design Report for the ATLAS Inner Tracker Pixel Detector*, tech. rep., CERN, Geneva, Sep, 2017.
10.17181/CERN.FOZZ.ZP3Q.
- [33] ATLAS Collaboration, *ATLAS Insertable B-Layer: Technical Design Report*, .
- [34] B. Abbott et al., *Production and integration of the ATLAS Insertable B-Layer*, *JINST* **13** (2018) T05008 [[1803.00844](#)].
- [35] ATLAS Collaboration, *Performance of the ATLAS trigger system in 2015*, *Eur. Phys. J. C* **77** (2017) 317 [[1611.09661](#)].
- [36] ATLAS Collaboration, *The ATLAS Collaboration Software and Firmware*, .
- [37] ATLAS Collaboration, “Deep Sets based Neural Networks for Impact Parameter Flavour Tagging in ATLAS.” ATL-PHYS-PUB-2020-014, 2020.

- [38] ATLAS Collaboration, *Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2*, *Eur. Phys. J. C* **77** (2017) 673 [[1704.07983](#)].
- [39] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, *Eur. Phys. J. C* **77** (2017) 466 [[1703.10485](#)].
- [40] M. Cacciari, G. P. Salam and G. Soyez, *The anti- k_t jet clustering algorithm*, *JHEP* **04** (2008) 063 [[0802.1189](#)].
- [41] 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** (2017) 072002 [[1703.09665](#)].
- [42] ATLAS Collaboration, “Tagging and suppression of pileup jets with the ATLAS detector.” ATLAS-CONF-2014-018, 2014.
- [43] L. Evans and P. Bryant, *LHC Machine*, *JINST* **3** (2008) S08001.
- [44] ATLAS Collaboration, *Observation of $H \rightarrow b\bar{b}$ decays and VH production with the ATLAS detector*, *Phys. Lett. B* **786** (2018) 59 [[1808.08238](#)].
- [45] ATLAS Collaboration, *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*, *Phys. Lett. B* **784** (2018) 173 [[1806.00425](#)].
- [46] ATLAS Collaboration, *Search for new resonances in mass distributions of jet pairs using 139 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *JHEP* **03** (2020) 145 [[1910.08447](#)].
- [47] PARTICLE DATA GROUP collaboration, *Review of particle physics*, *Phys. Rev. D* **98** (2018) 030001.
- [48] P. W. Battaglia, J. B. Hamrick, V. Bapst, A. Sanchez-Gonzalez, V. Zambaldi, M. Malinowski et al., *Relational inductive biases, deep learning, and graph networks*, *arXiv preprint arXiv:1806.01261* (2018) .
- [49] J. Shlomi, S. Ganguly, E. Gross, K. Cranmer, Y. Lipman, H. Serviansky et al., *Secondary vertex finding in jets with neural networks*, *The European Physical Journal C* **81** (2021) .

- [50] H. Serviansky, N. Segol, J. Shlomi, K. Cranmer, E. Gross, H. Maron et al., *Set2graph: Learning graphs from sets*, 2020.
- [51] ATLAS Collaboration, “Optimisation and performance studies of the ATLAS b -tagging algorithms for the 2017-18 LHC run.” ATL-PHYS-PUB-2017-013, 2017.
- [52] ATLAS Collaboration, *ATLAS b -jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **79** (2019) 970 [[1907.05120](#)].
- [53] ATLAS Collaboration, “Secondary vertex finding for jet flavour identification with the ATLAS detector.” ATL-PHYS-PUB-2017-011, 2017.
- [54] ATLAS Collaboration, “Identification of Jets Containing b -Hadrons with Recurrent Neural Networks at the ATLAS Experiment.” ATL-PHYS-PUB-2017-003, 2017.
- [55] P. Nason, *A new method for combining nlo qcd with shower monte carlo algorithms*, *Journal of High Energy Physics* **2004** (2004) 040–040.
- [56] S. Frixione, G. Ridolfi and P. Nason, *A positive-weight next-to-leading-order monte carlo for heavy flavour hadroproduction*, *Journal of High Energy Physics* **2007** (2007) 126–126.
- [57] S. Frixione, P. Nason and C. Oleari, *Matching nlo qcd computations with parton shower simulations: the powheg method*, *Journal of High Energy Physics* **2007** (2007) 070–070.
- [58] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing nlo calculations in shower monte carlo programs: the powheg box*, *Journal of High Energy Physics* **2010** (2010) .
- [59] NNPDF collaboration, *Parton distributions for the LHC run II*, *JHEP* **04** (2015) 040 [[1410.8849](#)].
- [60] ATLAS Collaboration, “Studies on top-quark Monte Carlo modelling for Top2016.” ATL-PHYS-PUB-2016-020, 2016.
- [61] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159

- [1410.3012].
- [62] ATLAS Collaboration, “ATLAS Pythia 8 tunes to 7 TeV data.” ATL-PHYS-PUB-2014-021, 2014.
- [63] R. D. Ball et al., *Parton distributions with LHC data*, *Nucl. Phys. B* **867** (2013) 244 [1207.1303].
- [64] D. J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth. A* **462** (2001) 152.
- [65] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, *Eur. Phys. J. C* **70** (2010) 823 [1005.4568].
- [66] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4 – a simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250.
- [67] ATLAS Collaboration, *Performance of b-jet identification in the ATLAS experiment*, *JINST* **11** (2016) P04008 [1512.01094].
- [68] ATLAS Collaboration, “Muon reconstruction performance in early $\sqrt{s} = 13$ TeV data.” ATL-PHYS-PUB-2015-037, 2015.
- [69] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **79** (2019) 639 [1902.04655].
- [70] D. Hwang, J. Park, S. Kwon, K.-M. Kim, J.-W. Ha and H. J. Kim, *Self-supervised auxiliary learning with meta-paths for heterogeneous graphs*, 2020. 10.48550/ARXIV.2007.08294.
- [71] J. Shlomi, P. Battaglia and J.-R. Vlimant, *Graph neural networks in particle physics*, *Machine Learning: Science and Technology* **2** (2021) 021001.
- [72] S. Brody, U. Alon and E. Yahav, *How Attentive are Graph Attention Networks?*, *arXiv e-prints* (2021) arXiv:2105.14491 [2105.14491].
- [73] M. Zaheer, S. Kottur, S. Ravanbakhsh, B. Póczos, R. Salakhutdinov and A. Smola, *Deep sets*, 2018.
- [74] D. P. Kingma and J. Ba, *Adam: A method for stochastic optimization*, .

- [75] J. Bai, F. Lu, K. Zhang et al., “Onnx: Open neural network exchange.” <https://github.com/onnx/onnx>, 2019.
- [76] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, *Phys. Lett.* **12** (1964) 132.
- [77] ATLAS and CMS Collaborations, *Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments*, *Phys. Rev. Lett.* **114** (2015) 191803 [[1503.07589](#)].
- [78] CMS Collaboration, *Observation of Higgs Boson Decay to Bottom Quarks*, *Phys. Rev. Lett.* **121** (2018) 121801 [[1808.08242](#)].
- [79] ATLAS Collaboration, *Measurement of VH , $H \rightarrow b\bar{b}$ production as a function of the vector-boson transverse momentum in 13 TeV pp collisions with the ATLAS detector*, *JHEP* **05** (2019) 141 [[1903.04618](#)].
- [80] J. M. Butterworth, A. R. Davison, M. Rubin and G. P. Salam, *Jet Substructure as a New Higgs Search Channel at the Large Hadron Collider*, *Phys. Rev. Lett.* **100** (2008) 242001 [[0802.2470](#)].
- [81] K. Mimasu, V. Sanz and C. Williams, *Higher order QCD predictions for associated Higgs production with anomalous couplings to gauge bosons*, *JHEP* **08** (2016) 039 [[1512.02572](#)].
- [82] ATLAS Collaboration, *Measurements of WH and ZH production in the $H \rightarrow b\bar{b}$ decay channel in pp collisions at 13 TeV with the ATLAS detector*, *Eur. Phys. J. C* **81** (2021) 178 [[2007.02873](#)].
- [83] J. R. Andersen et al., *Les Houches 2015: Physics at TeV Colliders Standard Model Working Group Report*, in *9th Les Houches Workshop on Physics at TeV Colliders (PhysTeV 2015) Les Houches, France, June 1-19, 2015*, 2016, [1605.04692](#).
- [84] R. Contino, M. Ghezzi, C. Grojean, M. Mühlleitner and M. Spira, *Effective Lagrangian for a light Higgs-like scalar*, *JHEP* **07** (2013) 035 [[1303.3876](#)].
- [85] J. K. Anders and M. D’Onofrio, *V +Jets theoretical uncertainties estimation via a parameterisation method*, Tech. Rep. ATL-COM-PHYS-2016-044, CERN, Geneva, 1, 2016.

- [86] A. S. Bell and F. Lo Sterzo, *Signal and Background Modelling Studies for the Standard Model $VH, H \rightarrow b\bar{b}$ Analysis*, Tech. Rep. ATL-COM-PHYS-2018-505, CERN, Geneva, 5, 2018.

List of figures

1.1	The Higgs potential $V(\phi)$ of the complex scalar field singlet $\phi = \phi_1 + i\phi_2$, with a choice of $\mu^2 < 0$ leading to a continuous degeneracy in the true vacuum states. A false vacuum is present at the origin. The SM Higgs mechanism relies on a complex scalar doublet with a corresponding 5-dimensional potential.	15
1.2	Higgs production cross sections at $\sqrt{s} = 13$ TeV [24].	20
1.3	Higgs branch ratios at $\sqrt{s} = 13$ TeV [24].	20
2.1	The coordinate system used at the ATLAS detector, showing the locations of the four main experiments located at various points around the LHC. Reproduced from Ref. [31].	22
2.2	Delivered, recorded, and usable integrated luminosity as a function of time during Run 2 [30].	24
2.3	Average pile-up profiles measured by ATLAS during Run 2 [30]. During Run 3, even higher levels of pile-up are expected.	24
2.4	A 3D model of the entire ATLAS detector. Cutouts through the detector	25
2.5	A 3D model of the ATLAS ID showing the barrel layers and end-cap disks.	26
2.6	A cross-sectional view of the ATLAS ID, with the radii of the different barrel layers shown.	27
2.7	A cross-sectional view of the ATLAS IBL.	28

2.8	The ATLAS calorimeters. The ECal (orange) and HCal (grey, brown).	29
2.9	The ATLAS muon spectrometer.	31
3.1	Diagram of a typical b -jet (blue) which has been produced along with two light jets (grey). The b -hadron has travelled a significant distance from the primary interaction point (pink dot) before its decay. The large transverse impact parameter d_0 is a characteristic property of the trajectories of b -hadron decay products.	36
3.2	As b -hadron p_T increases, the time of flight of the B increases, so tracks will have less room to diverge before reaching detector elements. To compound the problem, the collimation of the tracks increases. The detector may then be unable to resolve individual tracks.	37
3.3	Hit multiplicities on the IBL (3.3a) and the pixel layers (3.3b) as a function of the p_T of the reconstructed track. Tracks from the weak decay of the b -hadron are shown in red, while fragmentation tracks (which are prompt) are in blue. For each of these, standard tracks and pseudo-tracks are plotted. Hit multiplicities on the pseudo-tracks at high p_T due to the increased flight of the b -hadron. The baseline tracks have more hits than the pseudo-tracks, indicating that they are being incorrectly assigned additional hits.	38
3.4	Track reconstruction efficiency from b -hadron decay products for baseline ATLAS tracking (black), Bcut+Refit procedures applied (green), pseudo-tracking (blue), and for tracking where the ambiguity solver has been manually removed (orange).	38
3.5	The total number of pixel hits on tracks from b -hadron decays as a function of the production radius of the decay product. An excess of hits is assigned to the standard tracks in comparison to the ideal pseudo-tracks.	39

3.6	Distributions of angular distance dR between B hadrons and their weak decays and other fragmentation tracks (Fig. 3.6a), and the distribution of angular distance dR between B hadrons and the calorimeter clusters in the hadronic calorimeter (Fig. 3.6b). In Fig. 3.6a, the tracks from the weak decay of the B are significantly more collimated to the B than the other fragmentation tracks.	42
3.7	The distribution of cluster transverse momentum, in Fig. 3.7a for both clusters that were able (orange) and unable (blue) to be matched to a B hadron using MC truth information. The normalisation shows that the majority of clusters are not matched to B hadrons, resulting in fake ROIs. In Fig. 3.7b, the fractional improvement in track fit quality (χ^2/n) is shown for all track (blue), tracks with good IBL hits (green), and tracks with wrong IBL hits (orange). The distributions are overlapping, suggesting that the χ^2/n improvement is not a good discriminator of good and wrong hits.	43
3.8	Distributions of good (Fig. 3.8a) and wrong (fig:refit optimisation results sub2) hit assignment rates on the IBL for tracks using baseline tracking (black), the original unmodified refit procedure (green), and the refit procedure with an optimise set of ROI selection cuts (blue). The IBL lies at a radius of 33 mm from the beam pipe. Hence, particles produced with a production radius greater than this cannot leave good hits on the IBL.	43
3.9	Profiles, as a function of parent B hadron p_T , of good (Fig. 3.9a) and wrong (Fig. 3.9b) hit assignment rates on the IBL for tracks using baseline tracking (black), and various looser values of the outlier cut.	45
5.1	Comparison of the existing flavour tagging scheme (left) and GN1 (right). The existing approach utilises low-level algorithms (shown in blue), the outputs of which are fed into a high-level algorithm (DL1r). Instead of being used to guide the design of the manually optimised algorithms, additional truth information from the simulation is now being used as auxiliary training targets for GN1. The solid lines represent reconstructed information, whereas the dashed lines represent truth information.	51

- 5.2 The inputs to GN1 are the two jet features ($n_{\text{jf}} = 2$), and an array of n_{tracks} , where each track is described by 21 track features ($n_{\text{tf}} = 21$). The jet features are copied for each of the tracks, and the combined jet-track vectors of length 23 form the inputs of GN1. 58
- 5.3 The network architecture of GN1. Inputs are fed into a per-track initialisation network, which outputs an initial latent representation of each track. These representations are then used to populate the node features of a fully connected graph network. After the graph network, the resulting node representations are used to predict the jet flavour, the track origins, and the track-pair vertex compatibility. . . 58
- 5.4 Comparison between the DL1r and GN1 b -tagging discriminant D_b for jets in the $t\bar{t}$ sample. The 85% WP and the 60% WP are marked by the solid (dashed) lines for GN1 (DL1r), representing respectively the loosest and tightest WPs used by analyses. A value of $f_c = 0.018$ is used in the calculation of D_b for DL1r and $f_c = 0.05$ is used for GN1. The distributions of the different jet flavours have been normalised to unity area. 63
- 5.5 The c -jet (left) and light-jet (right) rejections as a function of the b -jet tagging efficiency for jets in the $t\bar{t}$ sample with $20 < p_T < 250$ GeV. The ratio with respect to the performance of the DL1r algorithm is shown in the bottom panels. A value of $f_c = 0.018$ is used in the calculation of D_b for DL1r and $f_c = 0.05$ is used for GN1 and GN1 Lep. Binomial error bands are denoted by the shaded regions. At b -jet tagging efficiencies less than $\sim 75\%$, the light-jet rejection becomes so large that the effect of the low number of jets is visible. The lower x -axis range is chosen to display the b -jet tagging efficiencies usually probed in these regions of phase space. 65

- 5.6 The c -jet (left) and light-jet (right) rejections as a function of the b -jet tagging efficiency for jets in the Z' sample with $250 < p_T < 5000$ GeV. The ratio with respect to the performance of the DL1r algorithm is shown in the bottom panels. A value of $f_c = 0.018$ is used in the calculation of D_b for DL1r and $f_c = 0.05$ is used for GN1 and GN1 Lep. Binomial error bands are denoted by the shaded regions. At b -jet tagging efficiencies less than $\sim 20\%$, the light-jet rejection becomes so large that the effect of the low number of jets is visible. The lower x -axis range is chosen to display the b -jet tagging efficiencies usually probed in these regions of phase space. 66
- 5.7 The b -jet tagging efficiency for jets in the $t\bar{t}$ sample (left) and jets in the Z' sample (right) as a function of jet p_T with a fixed light-jet rejection of 100 in each bin. A value of $f_c = 0.018$ is used in the calculation of D_b for DL1r and $f_c = 0.05$ is used for GN1 and GN1 Lep. Binomial error bands are denoted by the shaded regions. 68
- 5.8 The b -jet (left) and light-jet (right) rejections as a function of the c -jet tagging efficiency for $t\bar{t}$ jets with $20 < p_T < 250$ GeV. The ratio to the performance of the DL1r algorithm is shown in the bottom panels. Binomial error bands are denoted by the shaded regions. At c -jet tagging efficiencies than $\sim 25\%$, the light-jet rejection becomes so large that the effect of the low number of jets is visible. The lower x -axis range is chosen to display the c -jet tagging efficiencies usually probed in these regions of phase space. 70
- 5.9 The b -jet (left) and light-jet (right) rejections as a function of the c -jet tagging efficiency for Z' jets with $250 < p_T < 5000$ GeV. The ratio to the performance of the DL1r algorithm is shown in the bottom panels. Binomial error bands are denoted by the shaded regions. The lower x -axis range is chosen to display the c -jet tagging efficiencies usually probed in these regions of phase space. 70

- 5.10 The c -jet (left) and light-jet (right) rejections as a function of the b -jet tagging efficiency for $t\bar{t}$ jets with $20 < p_T < 250$ GeV, for the nominal GN1, in addition to configurations where no (GN1 No Aux), only the track classification (GN1 TC) or only the vertexing (GN1 Vert) auxiliary objectives are deployed. The ratio to the performance of the DL1r algorithm is shown in the bottom panels. A value of $f_c = 0.018$ is used in the calculation of D_b for DL1r and $f_c = 0.05$ is used for GN1. Binomial error bands are denoted by the shaded regions. At b -jet tagging efficiencies less than $\sim 65\%$, the light-jet rejection become so large that the effect of the low number of jets are visible. The lower x -axis range is chosen to display the b -jet tagging efficiencies usually probed in these regions. 71
- 5.11 The c -jet (left) and light-jet (right) rejections as a function of the b -jet tagging efficiency for Z' jets with $250 < p_T < 5000$ GeV, for the nominal GN1, in addition to configurations where no (GN1 No Aux), only the track classification (GN1 TC) or only the vertexing (GN1 Vert) auxiliary objectives are deployed. The ratio to the performance of the DL1r algorithm is shown in the bottom panels. A value of $f_c = 0.018$ is used in the calculation of D_b for DL1r and $f_c = 0.05$ is used for GN1. Binomial error bands are denoted by the shaded regions. At b -jet tagging efficiencies less than $\sim 25\%$, the light-jet rejection become so large that the effect of the low number of jets are visible. The lower x -axis range is chosen to display the b -jet tagging efficiencies usually probed in these regions. 72
- 5.12 Vertex finding efficiency as a function of jet p_T for b -jets in the $t\bar{t}$ sample using the exclusive (left) and inclusive (right) vertex finding approaches. Efficient vertex finding requires the recall of at least 65% of the tracks in the truth vertex, and allows no more than 50% of the tacks to be included incorrectly. Binomial error bands are denoted by the shaded regions. 75

5.13	Inclusive vertex finding efficiency for multitrack truth vertices in b -jets in the $t\bar{t}$ sample (left) and jets in the Z' sample (right) as a function of jet p_T . Efficient vertex finding requires the recall of at least 65% of the tracks in the truth vertex, and allows no more than 50% of the tracks to be included incorrectly.	76
5.14	ROC curves for the different groups of truth origin labels defined in Table 5.2 for jets in the $t\bar{t}$ sample (left) and jets in the Z' sample (right). The FromB, FromBC and FromC labels have been combined, weighted by their relative abundance, into the Heavy Flavour category, and the Primary and OtherSecondary labels have similarly been combined into a single category. The mean weighted area under the ROC curves (AUC) is similar for both samples.	78
6.1	Normalised distributions of leading fat jet mass m_J for medium (6.1a) and high (6.1b) p_{TV} analysis regions for W+heavy-flavour-jets in the 0 lepton channel. Merged in heavy flavours, high and low purity signal regions. The renormalisation scale μ_R has been varied by a factor of 2 (“1up”) and 0.5 (“1down”). An exponential function has been fit to the ratio.	86
6.2	Post-fit distributions for the 0-lepton (6.2a) and 2-lepton (6.2b) channels in the high purity medium p_{TV} region, obtained in the combined conditional $\mu = 1$ fit to data. The last bin of each plot is an overflow bin.	90

List of tables

1.1	The half-integer spin fermions of the SM [14]. Three generations of particles are present. Also present (unlisted) are the antiparticles, which are identical to the particles up to a reversed charge sign. . . .	10
1.2	The integer spin bosons of the SM [14]. The photon, weak bosons and gluons are gauge bosons arising from gauge symmetries, and carry the four fundamental forces of the SM.	11
2.1	Overview of the different LHC runs [29,30]. The average number of interactions per bunch-crossing is denoted as $\langle\mu\rangle$ (see Section 2.1.2). Numbers for Run 3 are preliminary and are only provided to give an indication of expected performance.	22
2.2	Quality selections applied to tracks, where d_0 is the transverse IP of the track, z_0 is the longitudinal IP with respect to the PV and θ is the track polar angle. Shared hits are hits used on multiple tracks which have not been classified as split by the cluster-splitting neural networks [38]. Shared hits on pixel layers are given a weight of 1, while shared hits in the SCT are given a weight of 0.5. A hole is a missing hit, where one is expected, on a layer between two other hits on a track.	33

5.1	Input features to the GN1 model. Basic jet kinematics, along with information about the reconstructed track parameters and constituent hits are used. Shared hits, are hits used on multiple tracks which have not been classified as split by the cluster-splitting neural networks [38], while split hits are hits used on multiple tracks which have been identified as merged. A hole is a missing hit, where one is expected, on a layer between two other hits on a track. The track leptonID is an additional input to the GN1 Lep model.	55
5.2	Truth origins which are used to categorise the physics process that led to the production of a track. Tracks are matched to charged particles using the truth-matching probability [38]. A truth-matching probability of less than 0.5 indicates that reconstructed track parameters are likely to be mismeasured and may not correspond to the trajectory of a single charged particle. The “OtherSecondary” origin includes tracks from photon conversions, K_S^0 and Λ^0 decays, and hadronic interactions.	56
5.3	A summary of GN1’s different classification networks used for the different training objectives. The hidden layers column contains a list specifying the number of neurons in each layer.	60
5.4	The area under the ROC curves (AUC) for the track classification from GN1, compared to a standard multilayer perceptron (MLP) trained on a per-track basis. The unweighted mean AUC over the origin classes and weighted mean AUC (using as a weight the fraction of tracks from the given origin) is provided. GN1, which uses an architecture that allows track origins to be classified in a conditional manner as discussed in Section 5.4.3, outperforms the MLP model for both $t\bar{t}$ and Z' jets.	78
6.1	Different sources of uncertainty (i.e. variations in the model) considered for V+jets background, and the corresponding implementation. For each uncertainty, acceptance and shape uncertainties are derived.	84

6.2	V+jets acceptance uncertainties. W +jets SR and CR uncertainties marked with a superscript \dagger are correlated. The 1L W +jets H/M uncertainty marked by $*$ is applied as independent and uncorrelated NPs in both HP and LP signal regions. The 0L W +jets Wbb Norm uncertainty is only applied when a floating normalisation for Wbb cannot be obtained from the 1L channel. A 30% uncertainty for Zbb norm is applied in the 1L channel when a floating normalisation for Zbb cannot be obtained from the 0L or 2L channels.	88
-----	---	----