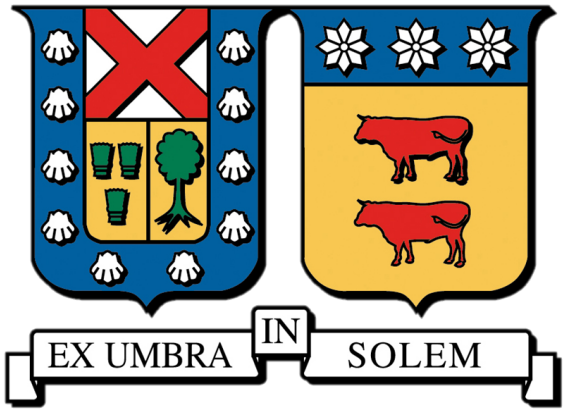
Performance of Boosted Identification Algorithms for $H \to \tau\tau$ and $HH \to b\bar{b}\tau\tau$ and New ITkStrip Bytestream Converter Developments in ATLAS

by

Daniel Torres

Supervisor: **Edson Carquin**



Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Physics

at the

Universidad Técnica Federico Santa María

March 12, 2026

CONSTANCIA DE VALIDACIÓN Y CONFIDENCIALIDAD DE MONOGRAFÍA A REPOSITORIO ACADÉMICO

1.- IDENTIFICACIÓN DEL TRABAJO ACADÉMICO

Tipo de monografía (marcar una opción): ☐ Memoria o trabajo de título ☒ Tesis de Postgrado

Título del trabajo: Performance of Boosted Identification Algorithms for $H \to \tau^+\tau^-$ and $HH \to b\bar{b}\tau^+\tau^-$ and New ITkStrip Bytestream Converter Developments in ATLAS

Nombre del candidato(a): Daniel Ignacio Torres Arza

Carrera / Grado: Magíster en ciencias mención física

Campus: Casa Central **Departamento:** Departamento de Física

2.- VALIDACIÓN DEL PROFESOR GUÍA/DIRECTOR DE TESIS

Yo, Edson Carquin, en mi calidad de profesor(a) guía/director(a) del trabajo académico mencionado anteriormente **DEJO CONSTANCIA** que:

- He revisado esta versión del documento y corresponde a la versión final aprobada del trabajo.
- El trabajo cumple con los requisitos académicos y de formato establecidos por la institución.

3.- EVALUACIÓN DE CONFIDENCIALIDAD POR PROPIEDAD INDUSTRIAL (marcar una opción)

☒ El trabajo **NO contiene** información que amerite confidencialidad y puede ser publicado de inmediato en repositorio con acceso abierto.

☐ El trabajo **CONTIENE** información con potenciales implicancias de propiedad industrial o intelectual y requiere un periodo de confidencialidad (**embargo**) por **(marcar una opción)**:

☐ 6 meses ☐ 12 meses ☐ 2 años ☐ 3 años ☐ 5 años ☐ 10 años

Fundamentación de la necesidad de confidencialidad (obligatorio si se solicita embargo):

4.- FIRMAS

Profesor(a) guía o director(a) de memoria o tesis:

Fecha: 12/03/26 **Firma:** 

Estudiante o Candidato(a):

Fecha: 12/03/26 **Firma:** 

Este formulario debe ser insertado como página 2 de la memoria o tesis, completado y firmado por estudiante y profesor(a) antes de la entrega en portal PRISMA de Biblioteca USM.

Thesis Committee

Dr. Edson Carquin
Universidad Técnica Federico Santa María

Dr. Stan Lai
Georg-August-Universität Göttingen

Dr. Nicolas Viaux
Universidad Técnica Federico Santa María

To my family,
for their unconditional support.

ABSTRACT

The discovery of the Higgs boson by the ATLAS and CMS collaborations at the Large Hadron Collider (LHC) completed the particle content of the Standard Model, while opening new directions for precision measurements of Higgs boson properties. In particular, the study of Higgs boson production in the high transverse momentum (p_T) regime, including Higgs boson pair production, poses significant challenges for object identification and triggering due to the highly collimated nature of the decay products. In parallel, the implementation of the High-Luminosity LHC requires major upgrades to the ATLAS detector, including the installation of a new Internal Tracker (ITk), as the current one is not designed to withstand HL-LHC conditions.

This thesis investigates the performance of new boosted di-τ identification algorithms in boosted $H \to \tau^+\tau^-$ and $HH \to b\bar{b}\tau^+\tau^-$ final states, and presents the development of a new trigger monitoring algorithm based on the new boosted di-τ trigger. In addition, a new ByteStream converter package for ITk Strip simulation detector is implemented, including encoding and decoding algorithms integrated into the ATLAS software framework.

Contents

CHAPTER 1

Introduction

The discovery of the Higgs boson by the ATLAS and CMS collaborations at the Large Hadron Collider (LHC) [1, 2] opened new areas of research in experimental particle physics and completed the particle content of the Standard Model (SM). Nevertheless, several Higgs boson properties remain to be precisely measured. Among them, the trilinear Higgs self-coupling (λ_{HHH}) is of particular interest, where one Higgs boson decay in a pair of Higgs boson.

The main decay channels of Higgs boson pair production have been extensively studied by the ATLAS and CMS collaborations and include the $b\bar{b}b\bar{b}$ [3–9], the $b\bar{b}VV$ [10–13], the $b\bar{b}\tau^+\tau^-$ [14–17], and the $b\bar{b}\gamma\gamma$ [18–20].

Higgs boson production, both single and pair production, becomes particularly interesting in the high transverse momentum (p_T) or boosted regime. In this region, the identification and triggering of Higgs boson decay products pose a significant challenge for the ATLAS experiment, as the decay products are highly collimated and standard reconstruction techniques lose their efficiency.

In parallel, the upcoming High-Luminosity LHC (HL-LHC) aims to reach a peak instantaneous luminosity of approximately $5 \times 10^{34}\text{cm}^{-2}\text{s}^{-1}$, significantly enhancing the sensitivity to rare processes, particularly in the high-p_T regime. The current ATLAS Inner Detector (ID) [21] was designed for operation at a luminosity of $10^{34}, \text{cm}^{-2}\text{s}^{-1}$ over a 10 year period and is therefore not suitable for HL-LHC conditions.

For this reason, a new Inner Tracker (ITk) will be installed for the HL-LHC era. The ITk consists of two main subsystems. The outermost subsystem, known as the **ITk Strip** detector [22], and the innermost subsystem, the **ITk Pixel** detector [23]. This thesis focuses on the ITk Strip detector.

This thesis is organized as follows. Chapter 2 introduces the theoretical framework relevant to this work, including an overview of the Standard Model and Higgs boson physics at the LHC. Chapter 3 describes the LHC accelerator complex, the ATLAS detector, the trigger and data acquisition system, particle reconstruction, and provides an introduction to the ITk detector. Chapter 4 presents the implementation of a new ByteStream (BS) converter package for ITk Strip simulations, explaining the concepts of ByteStream and the implemented algorithms (Encode and Decode). Chapter 5 describes the development of a new trigger

monitoring algorithm for the boosted di-τ trigger, following the Omni Tagger approach. Chapter 6 presents a study of the trigger performance of the new boosted di-τ trigger in boosted $H \to \tau^+\tau^-$ events, compared to existing triggers based on large-R jets, small-R jets, single-τ objects, and missing transverse energy (MET). Finally, Chapter 7 provides a comparison of the new boosted di-τ identification algorithms, GN2XTau and the Omni Tagger, in boosted $HH \to b\bar{b}\tau^+\tau^-$ events.

CHAPTER 2

Theoretical framework

2.1 Standard Model

The Standard Model (SM) of particle physics provides a theoretical framework that describes the fundamental constituents of matter and their interactions. This theory explains how the universe is built from *elementary particles* and how these particles interact with each other.

The SM is a gauge theory based on the symmetry group

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \tag{2.1}$$

where $SU(3)_C$ corresponds to the color symmetry of Quantum Chromodynamics (QCD), and $SU(2)_L \otimes U(1)_Y$ represents the gauge symmetry of the electroweak interaction (EW).

Within the SM, the particles are divided into two main categories, *fermions* and *bosons*. **Fermions** are spin-$\frac{1}{2}$ particles that follow Fermi-Dirac statistics, implying that two identical fermions cannot be in the same quantum state simultaneously, this is the Pauli exclusion principle.

These fundamental particles are classified into quarks and leptons, each grouped into three families, distinguished by the order of magnitude of their masses (see Table 2.1).

On the other hand, the **bosons** are particles with integer spin. These particles are responsible for mediating the fundamental interactions. First, the **electromagnetic interaction**, is mediated by the *photon* (γ), a massless and electrically neutral, which exerts its influence on all electrically charged particles, affecting both quarks, which have a fractional charge, and the W boson, which exhibits a charge of ± 1. In addition, it affects the charged leptons of the three families, composed of electron, muon, and tau, each of which carries a charge of -1.

The **weak interaction** is mediated by the massive vector bosons $W^{\pm}$ and Z^0. This interaction plays a crucial role in phenomena such as β-decay and has a very short range, being relevant only for processes that occur on the nuclear scale. Furthermore, the peculiarity of the weak force lies in its ability to interact with all particles of the SM, being the only one that interact with neutrinos.

Finally, the **strong interaction** is mediated by *gluons* (g), which are massless

particles that couple only to color-charged particles. This interaction is described by the $SU(3)_C$ symmetry of QCD. A peculiarity of the strong force is that its coupling constant α_s, varies with the **energy scale**. At high energies, the coupling constant decreases, resulting in a phenomenon known as **asymptotic freedom**, where particles appear to be "almost free". In contrast, at lower energy scales, α_s increases, generating the phenomenon of **confinement**. In this state, gluons exhibit self-interaction, making it impossible to find them as free particles. Consequently, these particles are always accompanied in such a way that the total color charge is zero. This is why the strong force is mediated by eight gluons.

	First Generation	Second Generation	Third Generation	I_3	Y	Q
Leptons	$\begin{pmatrix}\nu_e\\ e\end{pmatrix}_L$	$\begin{pmatrix}\nu_\mu\\ \mu\end{pmatrix}_L$	$\begin{pmatrix}\nu_\tau\\ \tau\end{pmatrix}_L$	$+\frac{1}{2}$ $-\frac{1}{2}$	-1	0 -1
	e_R	μ_R	τ_R	0	-2	-1
Quarks	$\begin{pmatrix}u\\ d\end{pmatrix}_L$	$\begin{pmatrix}c\\ s\end{pmatrix}_L$	$\begin{pmatrix}t\\ b\end{pmatrix}_L$	$+\frac{1}{2}$ $-\frac{1}{2}$	$+\frac{1}{3}$	$+\frac{2}{3}$ $-\frac{1}{3}$
	u_R	c_R	t_R	0	$+\frac{4}{3}$	$+\frac{2}{3}$
	d_R	s_R	b_R	0	$-\frac{2}{3}$	$-\frac{1}{3}$

Table 2.1: Lepton and quark content of the Standard Model organized by generation, weak isospin (I_3), hypercharge (Y), and electric charge (Q). Table taken from [24]

.

However, the gauge symmetries of the Standard Model require that gauge bosons have no mass. In the case of the photon and the gluon, this is not a problem. However, in the context of weak interaction, the $W^\pm$ boson has a mass of $m_w = 80.337 \pm 0.013$ GeV, and the Z^0 boson has a mass of $m_Z = 91.188 \pm 0.002$ GeV [25]. Independently between 1961 and 1964, Peter Higgs, François Englert, Robert Brout, and others proposed ideas that were later applied to solve this problem, a mechanism of spontaneous symmetry breaking (SBB).

This solution requires the introduction of a new scalar field, known as the Higgs field. Through the mechanism of spontaneous electroweak symmetry breaking, fundamental particles acquire mass via their interactions with this field, as described in Subsection 2.1.6.

The Higgs boson arises as a quantum excitation of the Higgs field around its non-zero ground state expectation value. It is an elementary scalar particle with spin 0, no electric charge, and no color charge. The Higgs boson was discovered in 2012 at the Large Hadron Collider (LHC) at CERN by the ATLAS and CMS collaborations [1, 2].

2.1.1 The Dirac Equation

The Dirac equation is based on the Klein-Gordon Equation 2.2, which is an equation arising from attempting to construct a relativistic quantum mechanics using the relation $E^2 = p^2 + m^2$. The issue of this equation, besides not including spin, is that it is a second-order equation in the derivatives, which makes it difficult to define a conserved and positively defined probability density, which in turn prevents a coherent interpretation of a single particle.

$$(\partial_\mu \partial^\mu + m^2)\psi = 0 \tag{2.2}$$

Dirac looked for a wave equation that was first order in both space and time derivatives (see Equation 2.3)

$$\hat{E}\psi = (\alpha \hat{p} + \beta m)\psi \tag{2.3}$$

$$i\frac{\partial}{\partial t}\psi = (-i\alpha_x \frac{\partial}{\partial x} - i\alpha_y \frac{\partial}{\partial y} - i\alpha_z \frac{\partial}{\partial z} + \beta m)\psi \tag{2.4}$$

However, this equation must satisfy the energy-momentum relation, just as the Klein-Gordon equation. Expanding the Dirac proposal in the Klein-Gordon equation in the same way as in [26], we obtain that the coefficients α and β must satisfy

$$\alpha_x^2 = \alpha_y^2 = \alpha_z^2 = \beta^2 = I \tag{2.5}$$

$$\alpha_j \beta + \beta \alpha_j = 0 \tag{2.6}$$

$$\alpha_j \alpha_k + \alpha_k \alpha_j = 0 \quad (j \neq k) \tag{2.7}$$

Therefore, these coefficients must be 4×4 Hermitian matrices acting on a four-component wavefunction, known as a *Dirac spinor* ψ. They can be represented in the *Dirac-Pauli representation* as

$$\beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \qquad \alpha_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix}, \tag{2.8}$$

Where I is the identity matrix 2×2 and σ_i are the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.9}$$

Finally, if we define the Dirac γ-matrices as

$$\gamma^0 \equiv \beta, \quad \gamma^1 \equiv \beta\alpha_x, \quad \gamma^2 \equiv \beta\alpha_y \quad and \quad \gamma^3 \equiv \beta\alpha_z \tag{2.10}$$

We obtain the Dirac equation in its covariant form

$$(i\gamma^\mu \partial_\mu - m)\psi = 0 \tag{2.11}$$

whose plane wave solutions describe free relativistic spin-$\frac{1}{2}$ particles. The equation admits four independent solutions, which can be written as

$$\psi_r(x) = u_r(p)e^{-ip_\mu x^\mu}, \quad \psi_r(x) = v_r(p)e^{ip_\mu x^\mu}, \quad r = 1, 2, \tag{2.12}$$

where $u_r(p)$ and $v_r(p)$ are constant spinors corresponding to positive-energy particle solutions and negative energy solutions (antiparticles), respectively. These spinors satisfy

$$(\gamma^\mu p_\mu - m)u_r(p) = 0, \quad (\gamma^\mu p_\mu + m)v_r(p) = 0. \tag{2.13}$$

A remarkable features is the presence of a solution with *negative energy*, which is a problem for the interpretation of the theory. This issue was resolved through the interpretation proposed by Stückelberg and Feynman [27], the *negative energy* could be interpreted as a particle with a negative energy propagating background in time. This implies that every charged particle has an antiparticle with the same mass and spin, but differing in its internal quantum number, such as electric charge or lepton number. For neutral particles the antiparticle is still distinct in general and differs by other internal quantum numbers. Only in special cases, such as **Majorana fermions**, can a neutral particle be identical to its antiparticle, which requires additional constraints beyond the Dirac formalism.

2.1.2 Quantum Electrodynamics (QED)

The free-fermion Lagrangian density is given by

$$\mathcal{L}_0 = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi, \; \bar{\psi} = \psi^\dagger \gamma^0 \tag{2.14}$$

This equation is invariant under a **global** $U(1)$ transformation, which is to say that the internal transformation $\psi(x) \rightarrow \psi(x)e^{i\alpha}$ leaves the Lagrangian 2.14 invariant. However, if the value of α is allowed to depend on the space-time coordinate, a **local** transformation $\psi(x) \rightarrow \psi(x)e^{i\alpha(x)}$ [28], the derivative term is no

longer invariant.

$$\partial_\mu \psi(x) \rightarrow e^{i\alpha(x)}[\partial_\mu + i(\partial_\mu \alpha(x))]\psi(x) \tag{2.15}$$

This transformation introduces an *extra term* proportional to $\partial_\mu \alpha(x)$. However, to restore local $U(1)$ invariance, we could try to introduce a new operator called the **covariant derivative**

$$D_\mu = \partial_\mu + iqA_\mu(x) \tag{2.16}$$

This will correct the local symmetry, as long as we force the new field A_μ to transform according to $A_\mu \rightarrow A_\mu - \frac{1}{q}\partial_\mu \alpha(x)$. Replacing $\partial_\mu \rightarrow D_\mu$, the Lagrangian becomes

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi = \mathcal{L}_0 - q\bar{\psi}\gamma^\mu A_\mu \psi \tag{2.17}$$

We now obtain a new Lagrangian density that is invariant under local $U(1)$ transformation. The additional term describes the interaction between the gauge field $A_\mu(x)$ and the fermionic field $\psi(x)$ and $\bar{\psi}$ [27].
Including the Lagrangian for the electromagnetic field, the complete gauge invariant Dirac Lagrangian, describing the fermions and their interactions, read

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi - q\bar{\psi}\gamma^\mu A_\mu \psi \tag{2.18}$$

Where $F_{\mu\nu}$ is the **electromagnetic field tensor**, defined as

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{2.19}$$

2.1.3 Quantum Chromodynamics (QCD)

Quantum Chromodynamics (QCD) is based on a $SU(3)$ symmetry, which assumes that the masses of the *up* (u) and *down* (d) quarks are approximately equal, while *strange* (s) quark mass is just about 100 MeV larger than u and d quarks. Compared to the typical baryon binding energies of around 1 GeV, this is relatively small [26]. However, this symmetry must be treated with care, as it is only approximate.

The group $SU(3)$ can be expressed in terms of the eight independent Hermitian generators $\hat{T}_i$. Consequently, a **global** $SU(3)$ transformation can be written as

$$\hat{U} = e^{i\alpha \cdot \hat{T}} \rightarrow \hat{T} = \frac{1}{2}\lambda \tag{2.20}$$

where the eight generators $\hat{T}_i$ are defined in terms of the *Gell-Mann* matrices λ_i. These matrices represent transformations that exchange quark flavour such as

$u \leftrightarrow d$; $u \leftrightarrow s$ and $d \leftrightarrow s$. In the flavour basis, the three quarks are written as

$$u = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad d = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad s = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \tag{2.21}$$

The complete set of $Gell - Mann$ matrices consists of eight independent matrices, given by

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{2.22}$$

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \tag{2.23}$$

$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \tag{2.24}$$

Analogously to the procedure in QED, one may ask whether QCD remains invariant under a **local** $SU(3)$ phase transformation

$$\psi(x) \rightarrow \psi'(x) = e^{ig_s \alpha(x) \cdot \hat{T}} \psi \tag{2.25}$$

where g_s is the strong coupling constant, $\hat{T} = T^a$ are the eight generators, and $\alpha^a(x)$ are eight functions of the space-time coordinate. Since the generators are 3×3 matrices acting on the wavefunction ψ, this representation must include three degrees of freedom, corresponding u, d and s quarks in $SU(3)$ flavour symmetry (see Equations 2.21). These degrees of freedom are referred to as *colour* and are represented by

$$r = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad g = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \tag{2.26}$$

Threfore, a **local** $SU(3)$ transformation corresponds to a "rotation" in this colour space, whose orientation may vary at each point in space-time. Under such a

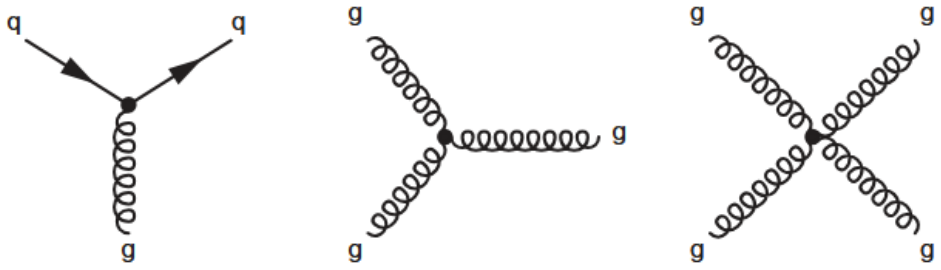


Figure 2.1: Feynman diagrams predicted by QCD, the quark-gluon interaction vertex on the left, and the three and four gluon self-interaction vertices on the middle and right. Figure taken from [26].

transformation, the Dirac equation 2.11 becomes

$$i\gamma^\mu[\partial_\mu + ig_s(\partial_\mu\alpha)\hat{T}]\psi = m\psi \tag{2.27}$$

To preserve the local gauge invariant, a new field $G^a_\mu(x)$ must be introduced, called *gluon fields*, transforming as $G^k_\mu(x) \rightarrow G^k_\mu(x) - \partial_\mu\alpha_k - g_s f_{ijk}\alpha_i G^j_\mu$. Thereby, the Dirac equation

$$i\gamma^\mu[\partial_\mu + ig_s G^a_\mu\hat{T}]\psi = m\psi \tag{2.28}$$

Here f_{ijk} are the *structure constants* of the $SU(3)$ group, which arise because QCD is a non-Abelian gauge theory, meaning that its generators do not commute $[\lambda_i, \lambda_j] = 2if_{ijk}\lambda_k$ [29].

By including the gauge field contribution, the Quantum Chromodynamics (QCD) Lagrangian density can be expressed as

$$\mathcal{L} = \bar{\psi}^f i\gamma^\mu(\partial_\mu + ig_s G^a_\mu\hat{T} - m_f)\psi^f - \frac{1}{4}G^a_{\mu\nu}G^{a\mu\nu} \tag{2.29}$$

The first term represents the quark kinetic term and its interaction with the gluon, while the second term accounts for the gluons self-interaction. The representation of the different interactions predicted by this Lagrangian are described in the figure 2.1. The diagram on the left shows the quark-gluon vertex generated by the first part of the Lagrangian. The middle and right diagrams correspond to the three and four gluon self-interactions, which are described by the expansion of the second term in the Lagrangian (see [28]).

2.1.4 Glashow–Weinberg–Salam Theory of Electroweak (EW) interaction

Before discussing electroweak (EW) theory, it is necessary to introduce the weak interaction. The phenomena of the β decay [1] suggests the existence of a new interaction, very weak and short-ranged. In 1934, Enrico Fermi [30] explained this process using a four-fermion contact interaction to describe the weak force (see Figure 2.2). However, the discovery of charge-parity (CP) violation[2] [31] led to the creation of a gauge theory for the weak interaction.

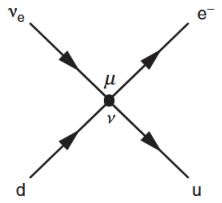


Figure 2.2: Feynman diagram of the effective four-fermion interaction of β decay in Fermi's theory. Figure taken from [26]

The weak interaction can be described by a gauge theory based on the local $SU(2)$ transformation

$$\varphi(x) \rightarrow e^{ig_W \alpha(x)\cdot \hat{T}}\varphi \tag{2.30}$$

where g_W is the weak coupling constant, $\alpha(x)$ is a local transformation in space-time coordinates, and $\hat{T}$ represent the three generators of $SU(2)$, which can be expressed in terms of the Pauli matrices(eq. 2.9)

$$T = \frac{1}{2}\sigma \tag{2.31}$$

Following the same steps as in QED and QCD, one finds that to maintain local gauge invariance it is necessary to introduce three new gauge fields W^k_μ with $k = 1, 2, 3$, corresponding to three new gauge bosons $W^{(1)}, W^{(2)}, W^{(3)}$.

Furthermore, the weak interaction couples only to *left-handed* (LH) particles and *right-handed* (RH) anti-particles [3]. Consequently, the symmetry is called $SU(2)_L$,

[1] β-decay is a weak interaction process in which a neutron transforms into a proton through the emission of a charged lepton and a neutrino $n \rightarrow p + e^- + \bar{\nu}_e$.

[2] CP violation refers to the phenomenon where the symmetry of charge conjugation (C), which swaps particles with antiparticles, and parity (P), which inverts spatial coordinates, is not conserved.

[3] The chirality of a particle, defined by the projection operators $P_r = \frac{1}{2}(1+\gamma^5)$ and $P_L = \frac{1}{2}(1-\gamma^5)$, is a property of a particle who distinguishes left- and right-handed states.

and the weak isospin doublets are composed only of the LH particle and RH antiparticle. In contrast, the RH particle and LH antiparticle are treated as singlets and remain unaffected by the $SU(2)_L$ transformation.

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \quad \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L, \quad \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L, \quad \begin{pmatrix} u \\ d \end{pmatrix}_L, \quad \begin{pmatrix} c \\ s \end{pmatrix}_L, \quad \begin{pmatrix} t \\ b \end{pmatrix}_L \tag{2.32}$$

The requirement of the local gauge invariance implies the modification of the Dirac equation 2.11 to include the new interaction term

$$ig_W T_k \gamma^\mu W_\mu^k \varphi_L = ig_W \frac{1}{2} \sigma_k \gamma^\mu W_\mu^k \varphi_L \tag{2.33}$$

The problem with a theory based only on $SU(2)_L$ is that it predict three massless gauge boson, two charged ($W^\pm$) and one neutral. One might initially identify the neutral with the Z boson, but experimental evidence shows that the Z couples to both left- and right-handed fermions, and W and Z bosons must be massive.

To solve this an electroweak unification model was proposed. This model introduces a gauge symmetry group $SU(2)_L \otimes U(1)$ that couples to left-handed fields and describes both electromagnetic and weak interactions. In this model, the $U(1)$ gauge symmetry of electromagnetism is replaced by a new local $U(1)_Y$ symmetry, where Y is the weak *hypercharge*

$$\psi(x) \to e^{ig' \frac{Y}{2} \xi(x)} \psi(x) \tag{2.34}$$

The corresponding interaction term is

$$g' \frac{Y}{2} \gamma^\mu B_\mu \psi \tag{2.35}$$

And the electroweak Lagrangian can be written as

$$\mathcal{L}_{EW} = -\frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{4} W_{\mu\nu}^a W^{a,\mu\nu} + \bar{\psi}(i\gamma^\mu D_\mu - m)\psi \tag{2.36}$$

where the covariant derivative D_μ is defined as

$$D_\mu = \partial_\mu - ig_W \frac{\sigma^a}{2} W_\mu^a - ig' \frac{Y}{2} B_\mu \tag{2.37}$$

This theory predicts four physical gauge fields

$$W^{\pm} = \frac{1}{\sqrt{2}}(W^{(1)} \pm iW^{(2)}) \tag{2.38}$$

$$Z = \cos(\theta_W)W^{(3)} - \sin(\theta_W)B \tag{2.39}$$

$$A = \cos(\theta_W)W^{(3)} + \sin(\theta_W)B \tag{2.40}$$

Where θ_W is the weak mixing angle.

Finally, the weak hypercharge of fermions can be expressed as a linear combination of the electromagnetic charge Q and the third component of weak isospin $I_W^{(3)}$ as

$$Y = 2(Q - I_W^{(3)}) \tag{2.41}$$

2.1.5 Spontaneous Symmetry Breaking

When we construct a gauge theory such as $SU(2)_L \otimes U(1)_Y$, it is built to be invariant under local gauge transformations. To maintain this invariance, the theory introduces gauge fields, in this case W_μ^a and B_μ (Subsection 2.1.4). However, if we look at the kinetic term of the gauge fields (see Eq. 2.36), we observe that no mass term can be added independently, since such a term is not invariant under local gauge transformations and would therefore break the gauge symmetry.

Nevertheless, we know that $W^{\pm}$ and Z^0 are massive. To reconcile this with the gauge principle, we require a mechanism that generates masses without explicitly breaking the symmetry, this is called spontaneous symmetry breaking (SSB). It is important to note that the SSB is not unique to the SM. A good example is the loss of ferromagnetism above Curie temperature, described by Landau theory [27].

To understand how SSB works, we could consider a Lagrangian with a *complex* scalar field $\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2)$, which can be written as

$$\mathcal{L} = (\partial_\mu \phi)^*(\partial^\mu \phi) - V(\phi) \quad \text{where} \quad V(\phi) = \mu^2(\phi^*\phi) + \lambda(\phi^*\phi)^2 \tag{2.42}$$

We are interested in the ground state, which corresponds to the configuration of the lowest energy. For the potential to be bounded from below, λ must be greater than zero. On the other hand, μ^2 does not have restriction, consequently, the shape of the potential depends on the sign of μ^2, as show in Figure 2.3:

1. For $\mu^2 > 0$, the minimum occurs at $\phi = 0$, and represent a particle with mass μ and a quartic self-interaction term proportional to λ

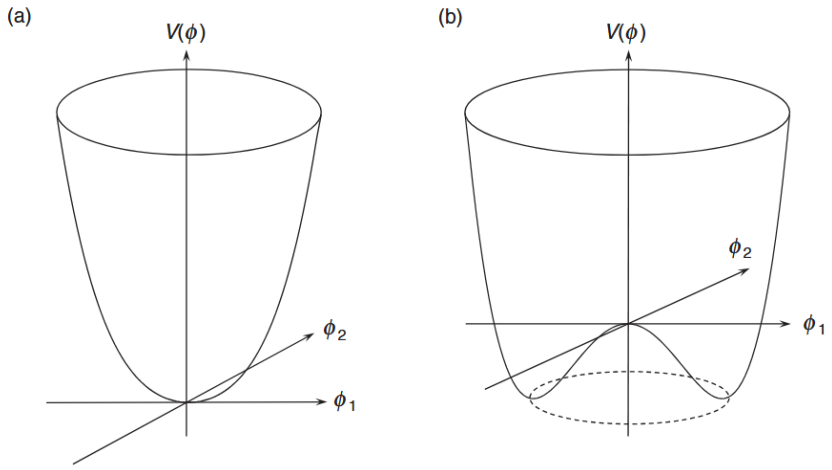


Figure 2.3: Shape of the $V(\phi)$ potential of a *complex* scalar. (a) represent the potential with $\mu^2 > 0$, (b) represent the potential with $\mu^2 < 0$. Figure taken from [26]

2. If $\mu^2 < 0$ the minimum satisfies

$$\phi = \pm \mid \sqrt{\frac{-\mu^2}{\lambda}} \mid = \pm v \tag{2.43}$$

In the second case, there are two degenerate ground states, as show in Figure 2.3(b). Choosing one of them breaks the symmetry of the Lagrangian. We can select the vacuum in the real direction, without loss of generality, $(\phi_1, \phi_2) = (v, 0)$. The fields can then be rewritten as fluctuations around the vacuum as $\phi_1(x) = \eta(x) + v$ and $\phi_2(x) = \xi(x)$. Otherwise, from equation 2.43, one finds $\mu^2 = -\lambda v$. Substituting this into the Lagrangian gives

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \eta)(\partial^\mu \eta) - \frac{1}{2} m_\eta^2 \eta^2 + \frac{1}{2}(\partial_\mu \xi)(\partial^\mu \xi) - V_{int}(\eta, \xi) \tag{2.44}$$

$$V_{int}(\eta, \xi) = \lambda v \eta^3 + \frac{1}{4}\lambda \eta^4 + \frac{1}{4}\lambda \xi^4 + \lambda v \eta \xi^2 + \frac{1}{2}\lambda \eta^2 \xi^2 \tag{2.45}$$

Which describes one massive scalar field η, with $m_\eta = \sqrt{2\lambda v^2}$ and one massless scalar field ξ. The interaction terms V_{int} correspond to triple and quartic couplings of the fields η and ξ. The massless field ξ corresponds to excitations along the direction where the potential remains unchanged, this is known as the **Goldstone boson**.

2.1.6 Higgs mechanism

Consider a local gauge symmetry $SU(2)_L \otimes U(1)_Y$. We can describe the simplest Higgs model with two *complex* scalar fields as

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{2.46}$$

and the corresponding Lagrangian

$$\mathcal{L} = (\partial_\mu \phi)^\dagger (\partial^\mu) - \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \tag{2.47}$$

For $\mu^2 < 0$, the potential has an infinite set of degenerate minima satisfying

$$\phi^\dagger \phi = -\frac{\mu^2}{2\lambda} = \frac{v^2}{2} \tag{2.48}$$

To SSB, we can choose the ground state as

$$(\phi)_0 = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}_0 = \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{2.49}$$

This choice preserves an electrically neutral vacuum, taking into accoun that the photon remains massless and the electromagnetic $U(1)_{EM}$ symmetry remains locally invariant.

The Goldstone field ξ can be removed from the Lagrangian by an appropriate gauge transformation. Since this Lagrangian is invariant under local $U(1)$ transformations, one can apply the following transformation to the field

$$\phi(x) \rightarrow \phi'(x) = e^{-i\xi(x)/v} \phi(x) \tag{2.50}$$

After the SSB and selecting *Unitary gauge*, corresponding to choosing the complex scalar field to be entirely real, the Goldston field $\xi(x)$ is eliminated. This is absorbed by the gauge fields, providing the longitudinal polarization required for a massive vector boson. The Higgs doublet is expressed as follows

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \tag{2.51}$$

Where h(x) is called the Higgs Field. Given that this Lagrangian is based on a local $SU(2)_L \otimes U(1)_Y$ symmetry, we can extend the EW Lagrangian in Eq. 2.36 to

construct the Salam–Weinberg model.

Considering that $Y = 2(Q - I_W^{(3)})$, and that the Higgs doublet is neutral with $I_W^{(3)} = -\frac{1}{2}$, we find the hypercharge is $Y = 1$. Hence, the Higgs covariant kinetic term can be written as

$$(D_\mu \phi)^\dagger (D^\mu \phi) = \frac{1}{2}\partial_\mu h \partial^\mu h - \frac{1}{8} g_W^2 (W_\mu^{(1)} - iW_\mu^{(2)})(W^{(1),\mu} - iW^{(2),\mu})(v+h)^2 \tag{2.52}$$

$$-\frac{1}{8}(g_W W_\mu^{(3)} - g' B_\mu)(g_W W^{(3),\mu} - g' B^\mu)(v+h)^2 \tag{2.53}$$

From the quadratic term in this expression, and as shown in Ref. [32], the gauge boson masses emerge like

$$m_W = \frac{1}{2} g_W v \tag{2.54}$$

$$m_Z = \frac{1}{2} v \sqrt{g_W^2 + g'^2} \tag{2.55}$$

$$m_A = 0 \tag{2.56}$$

Finnaly, the Higgs sector of the Lagrangian takes the form

$$\mathcal{L}_\phi = \frac{1}{2}\partial_\mu h \partial^\mu h^2 - \frac{1}{2} m_h^2 h - \frac{1}{2} m_Z^2 Z_\mu Z^\mu + \frac{1}{2} m_W W_\mu^\dagger W^\mu - \lambda v h^3 - \lambda h^4 + ... \tag{2.57}$$

Here $m_h = \sqrt{-2\mu^2} = \sqrt{2\lambda v^2}$ is the mass of the Higgs boson, and the cubic term h^3 corresponds to the Higgs self-interaction (trilinear coupling).

This mechanism not only explains the origin of the gauge boson masses, but also provides a description of how fermions acquire their masses in the SM. The corresponding term in the Lagrangian is given by [26]

$$\mathcal{L}_{Yukawa} = \sum_f m_f \bar{\psi}_f \psi - \frac{m_f}{v} \bar{\psi}_f \psi h \tag{2.58}$$

where m_f denotes the fermion mass. The interaction between the fermion and the Higgs field is called **Yukawa coupling** and is defined as

$$g_f = \sqrt{2}\frac{m_f}{v} \tag{2.59}$$

2.2 Higgs boson physics at the large hadron collider (LHC)

The Higgs boson can be produced through several mechanisms and has multiple decay channels. This section focuses on describing the main Higgs boson production and its corresponding decay modes, particularly at the Large Hadron Collider (LHC).

2.2.1 Higgs production modes

In proton-proton collisions, the final state particles are generated by the interaction of *partons* (quarks and gluons) within the proton, which exchange a large momentum perpendicular to the axis of the collision. It is possible to define a partonic cross section $\sigma_{ab\to X}$ for interaction between two partons a and b, describing the probability that they produce a given final state X. Then, the inclusive coss section for a hard scattering process initiated by two partons with four-momenta P_1 and P_2 can be expressed as [33]

$$\sigma(P_1, P_2) = \sum_{i,j} \int dx_1 dx_2 f_i(x_1, \mu) f_j(x_2, \mu) \hat{\sigma}_{ij}(p_1, p_2, \alpha_S(\mu), Q) \tag{2.60}$$

Where $f_i(x, \mu)$ represent the *parton distribution functions* (PDFs), that describe the probability of finding a parton i inside the proton, carrying a fraction x_i of the proton's momentum at the fatorization scale μ. The variable Q characterizes the hard scattering scale which can correspond to, for example, the mass of a weak boson, heavy quark, or the transverse momentum jet. The short-distance cross section $\hat{\sigma}_{ij}$ can be calculated perturbatively as a series expansion in the strong coupling constant α_S

$$\hat{\sigma} = c^{(0)} \alpha_S^k (1 + \sum_j^n c^{(j)} \alpha_S^j) \tag{2.61}$$

Where the coefficients $c^{(j)}$ depend on the kinematic variables.

An example of such a partonic process is Higgs production in proton-proton (pp) collisions. The cross section of the main Higgs boson production modes as a function of the centre-of mass energy are shown in Figure 2.4.

2.2.1.1 ggH production

The dominant Higgs boson production mechanism at the LHC **gluon-gluon fusion (ggH)**, this stands out to the other modes for have the largest cross-section $\sigma_{ggH} = 48.58 \pm 2.40$ pb [34], assuming a Higgs boson mass of $m_H = 125$ GeV and

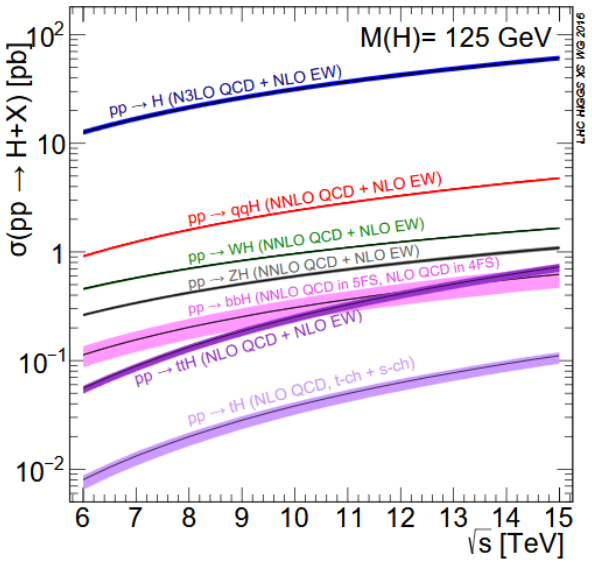


Figure 2.4: The SM Higgs boson production cross sections as a function of the LHC center of mass energy $\sqrt{s}$. Figure taken from [34]

a center of mass energy $\sqrt{s} = 13$ TeV. However, despite being the most frequent production mode, it also the process with more background, which requires careful analysis. The **leading order (LO)** Feynman diagram is shown in figure 2.5(a)

2.2.1.2 VBF production

The second most common production mode is **vector boson fusion (VBF)**. This implies the emission of two vector bosons (W or Z) by the scattering of two (anti-) quarks, which then fuse to produce a Higgs boson. The scattered quarks in the final state produce two forward jets, leading to a clean experimental signature [36]. The corresponding LO Feynman diagram is show in figure 2.5(b).

2.2.1.3 VH

The **VH** production mode involves the creation of a Higgs boson in association with a W or Z boson. The cross sections for these processes receive **next-to-leading order (NLO)** QCD corrections that modify the Drell-Yan process [4] [38–43]. Although this mode has a smaller cross section, the WH and ZH channels, especially in the boosted regime, provide a relatively clean environment for studying Higgs decays to bottom quarks [44].

[4]The Drell–Yan process describes the annihilation of a quark and an antiquark into vector bosons (γ, Z, W), which subsequently decay into leptons [37].

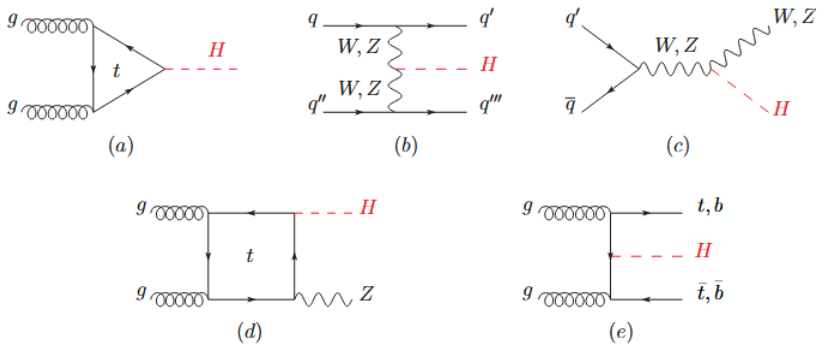


Figure 2.5: Main leading order Feynman diagrams contributing to the single Higgs boson production in (a) gluon fusion, (b) Vector-boson fusion, (c) Higgs-strahlung (or associated production with a gauge boson at tree level from a quark-quark interaction), (d) associated production with a gauge boson (at loop level from a gluon-gluon interaction), (e) associated production with a pair of top/bottom quarks. Figure taken from [35]

2.2.1.4 ttH

Another important production channel is associated production with a **top–antitop quark pair (ttH)**. The LO cross section for this production process was computed in refs. [45–49], NLO QCD [50–55] and NLO electroweak (EW) corrections [56–58] have been evaluated as well. Although the ttH channel has the smallest cross section among the main production modes, it provides a unique opportunity to directly measure the interaction between the top quark and the Higgs boson, the **top Yukawa coupling** 2.58. The LO Feynman diagram is shown in figure 2.5(e).

2.2.2 Higgs Decay Products

The Higgs boson has a decay width [5] of $\Gamma = 3.7^{+1.9}_{-1.4}$ MeV [25], which corresponds to a lifetime of approximately $O(\sim 10^{-22})$ s. Due to its short lifetime, direct observation of the particle in a detector is impossible. Therefore, we can only measure its decay products.

One way to quantify the probability that a parent particle P decays into a final state X_i is through the *branching ratio* (BR), defined as

$$BR(P \to X_i) = \frac{\Gamma(P \to X_i)}{\sum_j \Gamma(P \to X_j)} \tag{2.62}$$

where $\Gamma(P \to X_i)$ is the *partial decay width* of the process, and the denominator represents the sum over all possibles partial decays widths.

[5] The decay width is a measure of the lifetime of a particle, defined as $\Gamma = \frac{\hbar}{\tau}$, where $\hbar$ is the reduced Planck constant and τ is the particle's lifetime).

H Decay Mode	Branching Ratio (%)
WW^*	25.7 ± 2.5
ZZ^*	2.80 ± 0.30
$\gamma\gamma$	$(2.50 \pm 0.20) \times 10^{-3}$
$b\bar{b}$	53 ± 8
e^+e^-	$< 3.0 \times 10^{-4}$
$\mu^+\mu^-$	$(2.6 \pm 1.3) \times 10^{-4}$
$\tau^+\tau^-$	$6.0^{+0.8}_{-0.7}$

Table 2.2: Branching ratio percentages for the different decay channels of Higgs. Values taken from [25]

In the context of the Higgs analyses, some relevant branching ratios are shown in table 2.2. Where, two of the most relevant decay channels are $H \to b\bar{b}$ and $H \to \tau^-\tau^+$, with branching ratios of $BR(H \to b\bar{b}) = 53 \pm 8\%$ and $BR(H \to \tau^-\tau^+) = 6.0^{+0.8}_{-0.7}$ respectively. The Higgs boson decay to a bottom quark–antiquark pair, while the most frequent, is challenging to observe due to a large background from non-resonant QCD events that produce multiple jets in the detector. On the other hand, the Higgs boson decay to a tau-antitau pair is one of the most accesible leptonic channels and was key for the first observation of Higgs Yukawa couplings. Both channels have been studied by ATLAS and CMS [59–65], giving essential information about Higgs fermion interaction.

2.2.3 Higgs decay on the boosted regime

At the LHC, particles can be produced with large transverse momenta $p_T \gtrsim 600$ GeV, which causes their decay products to become highly collimated (overlapping). For this reason, identifying boosted Higgs decays is particularly challenging for ATLAS and CMS. In the resolved regime, shown in Figure 2.6(left), the decay products are well separated and can be reconstructed as two small-radius jets. In the boosted regime, however, the Higgs boson decays into two particles that carry a large fraction of the original Higgs p_T. The average angular separation between the two decay products can be approximated as $\Delta R \sim \frac{2m_H}{p_T} \lesssim 0.4$, where m_H and p_T denote the Higgs mass and its transverse momentum, respectively [5]. As a consequence, the two jets merge into a single large-radius jet (*fat jets*), as shown in Figure 2.6(right).

2.2.4 Non resonant Di-Higgs Production

The introduction of the Higgs sector on the Lagrangian (Eq. 2.57) defines the trilinear Higgs self-coupling ($\lambda_{HHH} = \frac{m_h^2}{2v^2}$). Higgs boson pair production (HH) is a

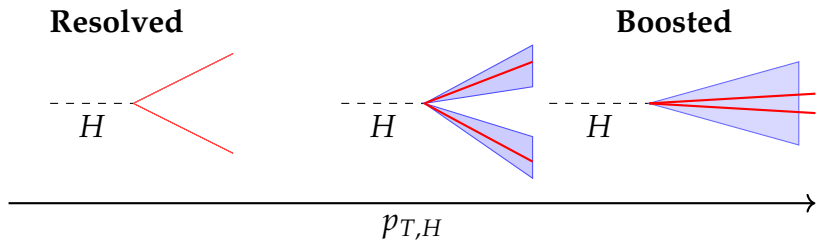


Figure 2.6: Representation of the boosted regimen for the Higgs boson decay. At low $p_{T,H}$, the decay products are separeted (resolved), while at high $p_{T,H}$, they become collimated (overlapped) inside a region (boosted)

rare process that provides direct access to this coupling.

The dominant production mechanisms are gluon-gluon fusion (ggH) and vector boson fusion (VBF). In the SM, the ggH cross section for a Higgs boson mass of 125 GeV, calculated at next-to-next-to-leading order (NNLO) precision in perturbative quantum chromodynamics (QCD), is $\sigma^{SM}_{ggH} = 31.1^{+2.1}_{-7.2}$ fb [66, 67]. This process is sensitive to the values of the trilinear coupling, λ, and the top-quark Yukawa coupling, y_t. Deviations from their SM prediction are parametrized as $\kappa_\lambda = \frac{\lambda}{\lambda_{SM}}$.

For VBF production, the cross section calculated at next-to-NNLO (N^3LO) precision, is 1.726 ± 0.036 fb [68]. This process depends on the Higgs self-coupling and the coupling of a pair of W or Z boson with either a single Higgs boson (VVH) or a pair of Higgs bosons (VVHH). The corresponding deviation from the SM expectations are parametrized as κ_V (VVH) and κ_{2V} ($VVHH$), respectively [9]. The main leading-order Feynman diagrams to HH are shown in Figure 2.7.

The three most commonly studied decay channels are shown in Figure 2.8. These channels have been extensively analyzed by ATLAS and CMS correspond to $b\bar{b}b\bar{b}$ [3–9], $b\bar{b}VV$ [10–13], $b\bar{b}\tau^-\tau^+$ [14–17] and $b\bar{b}\gamma\gamma$ [18–20].

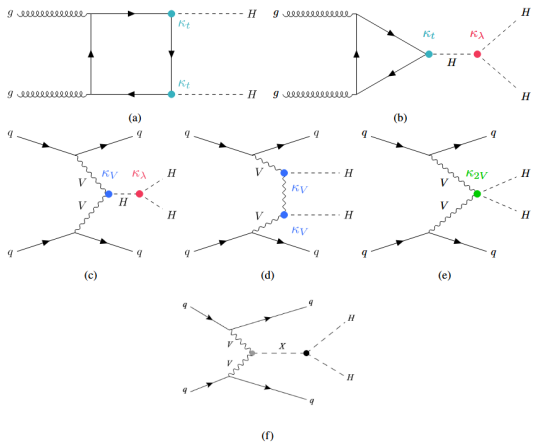


Figure 2.7: Examples of leading-order Feynman diagrams for Higgs boson pair production. For nonresonant ggF production, diagram (a) involves solely the top-quark Yukawa coupling, while diagram (b) involves the Higgs boson self-coupling. For nonresonant VBF production, diagram (c) involves the self-coupling, diagram (d) involves solely the coupling to vector bosons, and diagram (e) involves the coupling between two Higgs bosons and two vector bosons. Diagram (f) illustrates the resonant production mode. Figure taken from [7]

	bb	WW	ττ	ZZ	γγ
bb	34%				
WW	25%	4.6%			
ττ	7.3%	2.7%	0.39%		
ZZ	3.1%	1.1%	0.33%	0.069%	
γγ	0.26%	0.10%	0.028%	0.012%	0.0005%

Figure 2.8: Branching ratios for the different decay channels of non-resonant Higgs boson pair production. Figure taken from [69]

CHAPTER 3

The ATLAS experiment at the LHC accelerator complex

This chapter provides an overview of the Large Hadron Collider (LHC), its structure and the ATLAS experiment. First, we describe how the LHC is built and introduce its main experiments Section 3.1. Next, we present the ATLAS detector, describing its general design and the components relevant to particle detection, Section 3.3. The third section focused on object reconstruction and identification within ATLAS, Section 3.5. Finally, the fourth section describes the new Inner Tracker (ITk), the upgraded tracking system developed for the next phase of the ATLAS experiment, Section 3.6.

3.1 CERN accelerator complex

The Large Hadron Collider (LHC) [70] is located on the border between France and Switzerland, near Geneva. It lies 100 meters underground and forms a ring tunnel with a circumference of 27 km. This tunnel was originally built between 1984 and 1989 for the Large Electron-Positron Collider (LEP) [71]. However, in December 1994, the CERN Council decided to reconfigure and upgrade these facilities to make way for the Large Hadron Collider (LHC). The LHC was constructed to collide beams of protons at a maximum center of mass energy of $\sqrt{s} = 14$ TeV, reaching a maximum luminosity of 10^{34} cm^{-2}s^{-1}, and heavy-ions (Pb+Pb or Pb+p) at $\sqrt{s} = 5$ TeV with a maximum luminosity 10^{27} cm^{-2}s^{-1}. The main goal of these collision is to study the Higgs boson, its properties, and physics beyond the standard model (BSM).

This collider consists of four main experiments, marked as yellow circles in Figure 3.1. ALICE (A Large Ion Collider Experiment) [73], dedicated to heavy ion physics, where matter interacts strongly at extreme energy densities, LHCb (Large Hadron Collider beauty) [74], responsible for differentiating between matter and antimatter by studying *b quarks*, and finally, the best-known multi-purpose detectors, ATLAS (A Toroidal LHC ApparatuS) [75] and CMS (Compact Muon Solenoid) [76].

When a beam of particles is injected into the LHC, it is accelerated step by step, gaining more energy with each stage of the accelerator chain. For protons, this

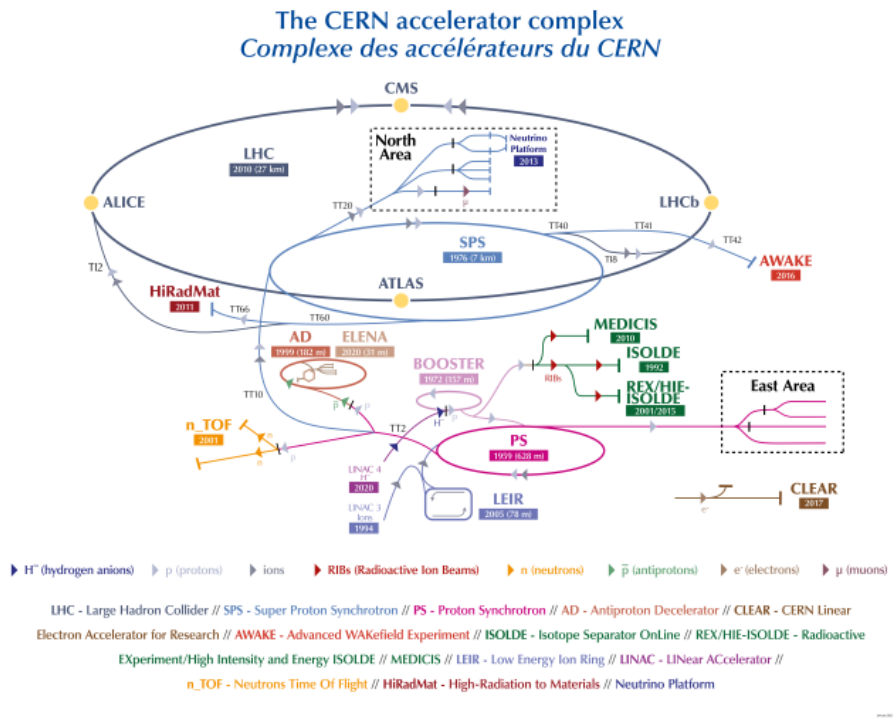


Figure 3.1: LHC accelerator complex. Figure taken from [72]

process begins in the **linear accelerator (LINAC 4)**, which accelerates negative hydrogen ions to an energy of 160 MeV. The beam is then sent to the **BOOSTER**, where the ions are stripped off their electrons, leaving only the protons, which are accelerated to 2 GeV. After this stage, the beam enters the **proton synchrotron (PS)**, where the beams are accelerated again, reaching 26 GeV. The process then continues in the **super proton synchrotron (SPS)**, which increases their energy to 450 GeV. Finally, they enter the LHC, where they reach an energy of 6.8 TeV [77]. These beams are guided by superconducting magnets operating at the extremely low temperature of 1.9 K, achieved by using liquid helium in its superfluid phase. At this temperature, the magnets become superconducting, offering no electrical resistance and producing magnetic fields of about 8.3 T.

3.2 Luminosity

At the LHC, it is crucial to know the number of collisions, which can be estimated using the **instantaneous luminosity ($\mathcal{L}$)**. This quantity provides information about the expected number of number of collision per unit of cross section in a given time. It depends solely on beam parameters and is a fundamental indicator of accelerator performance. The instantaneous luminosity is expressed in $\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

For two beams with Gaussian profiles, the instantaneous luminosity can be written as

$$\mathcal{L} = \frac{N_1 N_2 f N_b}{4\pi\sigma_x\sigma_y} \tag{3.1}$$

where N_1 and N_2 are the number of particles in each beam, f is the revolution frequency, N_b is the number of bunches and $\sigma_{x,y}$ are the standard deviations of the shapes of the Gaussian beam.

However, in a real detector, this calculation is more complex and requires other considerations, such as the details indicated here [78].

Once we know the luminosity, we can estimate the number of events of a given process produced over a period of time as

$$N_{events} = \sigma_p \cdot L \tag{3.2}$$

where $L = \int \mathcal{L}\, dt$ is the **integrated luminosity** and σ_p is the cross section of the process.

3.3 ATLAS experiment

The ATLAS detector shown in Figure 3.2, are 25 m in height and 44 m in lenght. The overall weight of the detector is approximately 7000 tones. This has been operating since 2009, it was designed to resist a maximum luminosity of $\mathcal{L} = 10^{34}\text{cm}^{-2}\text{s}^{-1}$ in proton-proton (pp) collisions with a center of mass energy of $\sqrt{s} =$ 14 TeV, and of $\mathcal{L} = 10^{27}\text{cm}^{-2}\text{s}^{-1}$ of heavy ions collisions in $\sqrt{s} = 5.5$ TeV. During the 2009–2013 period, corresponding to Run 1 of the LHC, ATLAS accumulated an integrated luminosity of 25 fb^{-1} of pp collisions at a center-of-mass energy of 7 TeV (2011) and 8 TeV (2012).

Between 2013 and 2015, collisions were halted to carry out upgrades to both the detector and the accelerator in general. In the period 2015-2018, during Run 2, ATLAS achieved an integrated luminosity of 147 fb^{-1} with pp collisions at a center-of-mass energy of 13 TeV. The maximum luminosity achieved during this time was 2.1 times greater than that originally designed for the detector. Despite this increase, the detector operated efficiently, although analysis techniques had to be adapted to maintain the physical scope during this period.

During the period 2019-2022, the detector was closed for upgrades and the connection of the new linear accelerator (LINAC 4). From 2022 to the present, Run 3 is underway, with the expectation of reaching a peak luminosity of $\mathcal{L} =$

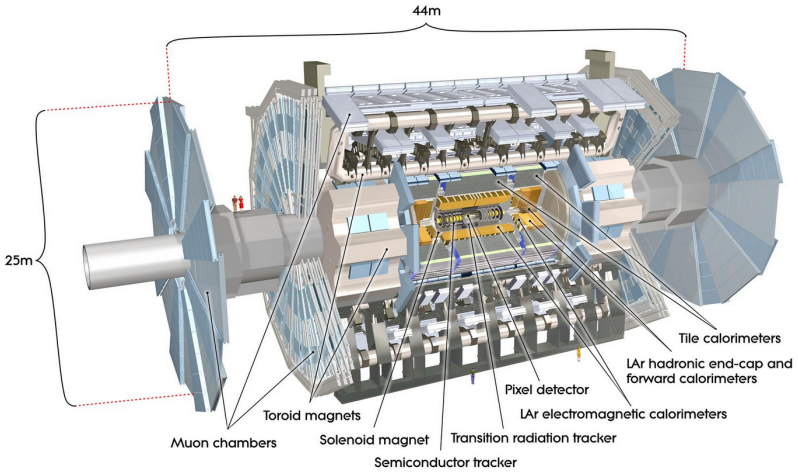


Figure 3.2: Cut-away view of the ATLAS detector. Figure taken from [79].

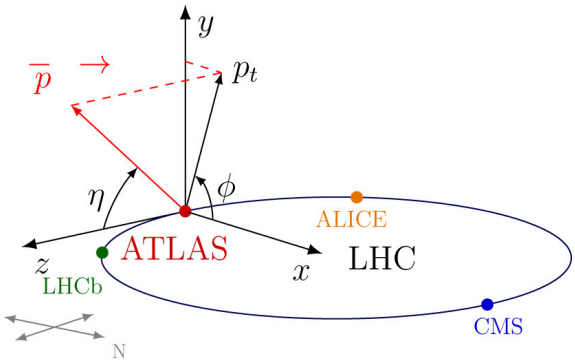


Figure 3.3: The coordinate systems used at the ATLAS experiment in the geographical context of the LHC. Figure taken from [80]

2×10^{34} cm^{-2} s^{-1} and accumulate an integrated luminosity of 250 fb^{-1} and an integrated luminosity of 250 fb^{-1} in **pp** collisions with $\sqrt{s} = 13.6$ TeV [75].

The ATLAS detector is designed to be nominally forward-backward symmetric with respect to the interaction point. This nominal interaction defines the origin of the coordinate system, in which the direction of the z-axis is aligned with the collision axis of the particles, while the x-y plane is transverse to this same axis. In this system, the positive x-axis points toward the center of the LHC and the y-axis points upward. The azimuthal angle ϕ is measured around the beam axis, and the polar angle θ is the angle from the beam axis. The pseudorapidity η, is normally used to describe the coverage of subdetectors, is define as $\eta = -\ln(\tan(\theta/2))$, as shown in the Figure 3.3.

3.3.1 Inner Detector

The Inner Detector (ID) [21] aims to precisely measure the point of origin and trajectory of charged particles. It operates within an axial magnetic field of 2 T produced by a superconducting solenoid magnet.

The ID is separated into three subsistems, as shown in the Figure 3.4 (left), the *Pixel Detector*, the *SemiConductor Tracker* (SCT) and the *Transition Radiation Tracker* (TRT).

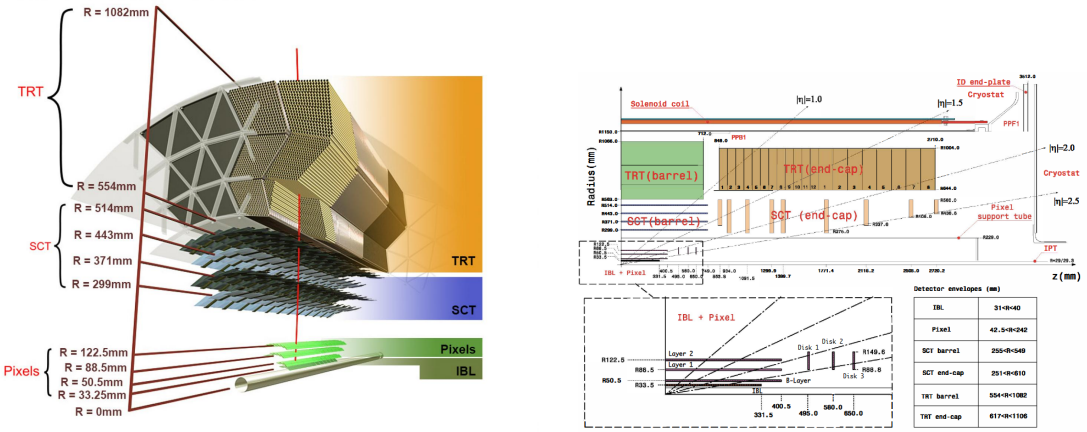


Figure 3.4: Layouts of the ATLAS Inner Detector. Figures taken from [79] (left) and [81] (right).

3.3.1.1 *Pixel Detector*

The innermost subsystem surrounding the beam pipe, which extends up to about 15 cm in radius, is the **Pixel Detector**. It was originally designed with three barrel-layers at radii of 5, 9.8, and 12.2 cm, covering a pseudorapidity of $|\eta| <= 1.7$, together with three end-cap disks on each side, providing coverage in the region $1.7 < |\eta| < 2.5$ (see Figure 3.4 (right)). During the first LHC shutdown 2013-2015, an additional innermost barrel layer was installed, the Insertable B-Layer (IBL) [81]. Positioned at a radius of 33.5 mm, extending up to $|\eta| = 2.7$.

The complete Pixel Detector is composed of 1744 *identical modules* given a sensitive area of $\sim$ 1.7 m^2 in total. Each module has an active surface of 6.08 $\times$ 1.64 cm^2 and is made of silicon sensor tile containing 47232 **pixels**. Attached to the sensor are sixteen **front-end electronics chips (FE)** each containing 2880 pixel cells with amplifier circuitry, connected to the sensor by means of fine-pitch bump bonding. A thin ($\sim$ 100 μm) double-sided printed circuit route signals and power to the electronics. The module also includes a **Module Control Chip (MMC)** located on the flex-hybrid. For barrel modules, another flexible foil, called a pigtail, provides the connection to electrical services through a micro-cable, whereas for

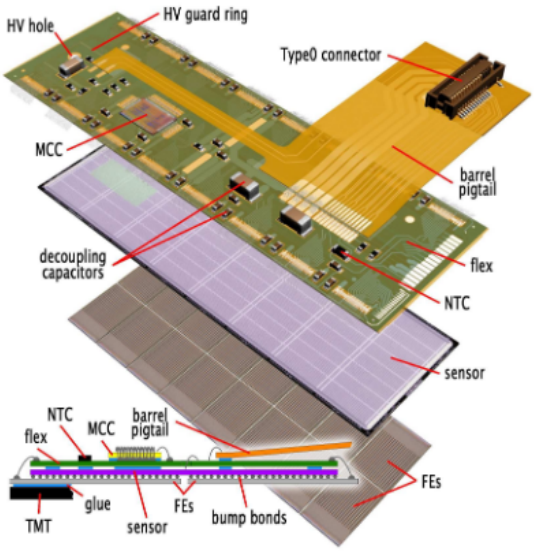


Figure 3.5: The elements of a pixel barrel module. Most of the thermal management tile (TMT) on which the module is glued is suppressed. Figure taken from [82]

disk modules the micro-cables are attached directly without the pigtail interface [82]. An schematic representation of a Pixel module is shown in Figure 3.5

3.3.1.2 SemiConductor Tracker (SCT)

The **SemiConductor Tracker (SCT)** [83] forms the middle section of the ID, covering radii from 30 to 60 cm. This subsystem is crucial for precise momentum measurement. It consists of four barrel layers located at radii of 30, 37.3, 44.7, and 52.0 cm, covering the pseudorapidity of $|\eta| = 1.4$ and nine end-cap wheels on each side that extend the tracking acceptance to $1.4 < |\eta| < 2.5$, as shown in Figure 3.4 (right).

In total, the SCT contains 4088 modules, 2112 located equally spaced over the four barrel layers Figure 3.6 (left), and 1979 over the eighteen end-cap wheels Figure 3.6 (right).

The barrel section uses four layers of silicon microstrip detectors to provide precision points in R_ϕ and z. Each silicon detector is 6.36×6.40 cm^2 with 768 readout strips each with 80 μm. Each module consists of four detectors. On each side of the module, two sensors are wire-bonded together to form 12.8 cm long strips. Two such module pairs are then glued together back-to-back at a 40 mrad angle, separated by a heat transport plate, and the electronics is mounted above the sensors on an hybrid [21].

The end-cap modules follow a similar construction principle but use tapered

strips, with one sensor set aligned radially to accommodate the disk geometry.

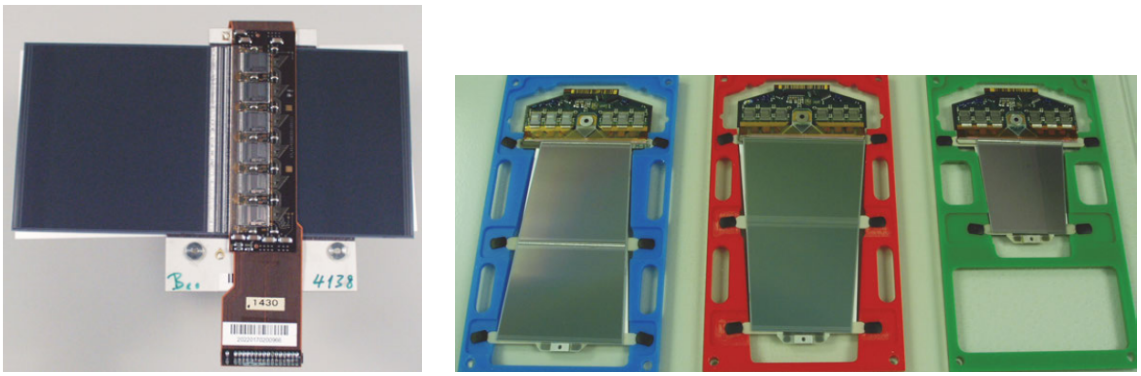


Figure 3.6: Representation of the SCT (left) barrel modules and (right) endcap modules. Figure taken from [79]

3.3.1.3 *Transition Radiation Tracker (TRT)*

The **Transition Radiation Tracker (TRT)** [84] forms the outermost part of the Inner Detector. It extends in radii from 60 to 95 cm and consists of a large number of small diameter proportional **drift tubes (straws)**, which provide good space resolution in the track bending plane and contributing strongly to the momentum resolution and the pattern recognition with multiple measurements. The TRT does not cover the full pseudorapidity range, but ends at $|\eta| \sim 2.0$ (see Figure 3.4 (right)). The main role is to contribute to momentum measurement and pattern recognition (photon conversions and long-lived particles), and to provide electron identification in addition to that from calorimeters [79].

3.3.2 The Calorimeters

The calorimeters [85] play a crucial role in the ATLAS detector. Their main function includes the accurate measurement of the energy and position of electrons and photons, the reconstruction of jets and missing transverse momentum (p_T^{miss}) and particle identification, such as distinguishing electrons and photons from hadrons and jets, or separating hadronically decay τ from jets. Additionally, they contribute to event selection at the trigger level (see Section 3.4). The calorimeter system cover a pseudorapidity of $|\eta| < 4.9$, with the central calorimeters covering approximately $|\eta| < 1.8$. A schematic diagram is shown in Figure 3.7.

The **electromagnetic calorimeter (ECAL)**, based on Lead/Liquid-Argon (LAr), is divided into a barrel section that covers $|\eta| < 1.475$, consisting of two identical half-barrels separated by a small gap, and two end-caps calorimeters covering $1.375 < |\eta| < 3.2$. Each end-cap is mechanically divided into two coaxial wheels. In addition, an LAr pre-ampler is installed in the region of $|\eta| < 1.8$ to correct for

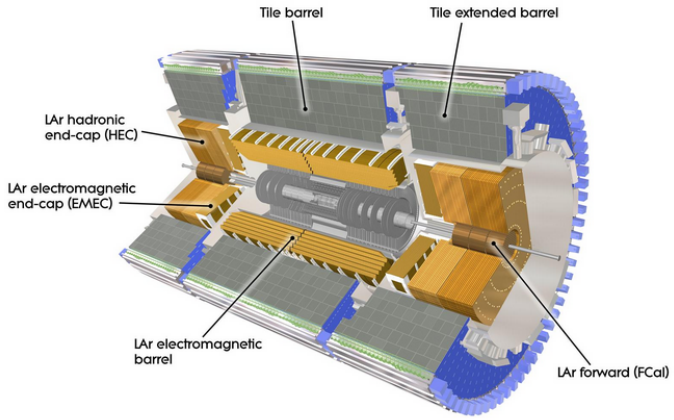


Figure 3.7: View of the ATLAS calorimeter system. Figure taken from [79].

energy loss in the material above the calorimeter.

The **hadronic calorimeter (HCAL)**, is composed of multiple subsystems. The barrel and extend-barrel calorimeters use steel absorbers and scintillator tiles and cover $|\eta| < 1.7$, while the hadronic end-cap calorimeters are based on copper/LAr covering a range of $1.5 < |\eta| < 3.2$.

In the most forward part, the **Forward Calorimeter (FCAL)** provides coverage in range $3.1 < |\eta| < 4.9$, reducing to the minimal possible level the effects of crack and dead space in the transition region around $|\eta| = 3.1$. This leads to better forward-jet tagging performance and reduces the tails of the missing transverse energy (E_T^{miss}) distribution.

3.3.3 The Muon Spectrometer

The **Muon Spectrometer (MS)** [86] is designed to measure muons (μ), which are highly penetrating charged particles capable of traversing almost all other ATLAS subdetectors (Inner detector and calorimeters) without being significantly disturbed. A schematic layout of the MS is shown in Figure 3.8.The system operates by measuring the *magnetic deflection* of muon tracks in a system of three large superconducting air-core toroid magnets. These magnets are equipped with specific **trigger chambers** and **high-precision tracking chambers**.

For the region $|\eta| \leq 1$, the magnetic bending is provided by a large barrel toroid. For $1.4 \leq |\eta| \leq 2.7$, the muon tracks are bent by two smaller end-cap toroids located at both ends of the detector. In the transition region $1 \leq |\eta| \leq 1.4$, the magnetic field is provided by a combination of barrel and end-cap magnets.

The main *high-precision tracking chambers* are the **Monitored Drift Tubes (MDT)**, covering $|\eta| < 2.7$ and the **Cathode Strip Chamber (CSCs)**, which provide addi-

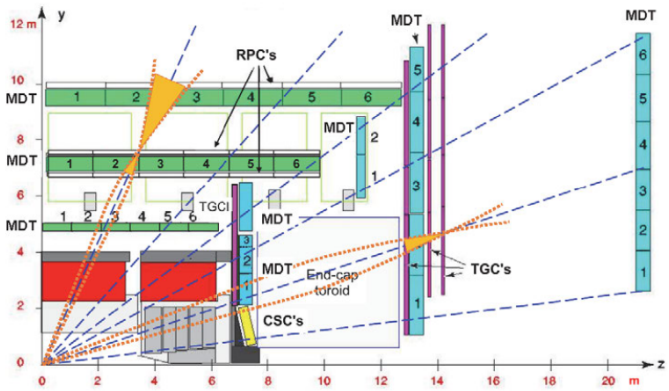


Figure 3.8: Schematic of the trigger arrangement to determine muon momentum, profiting from the bending in the magnetic field. Infinite momentum tracks are shown in blue and acceptance cones in yellow. Also includes positioning of the components (RPC, TGC, MDT and CSC). Figure taken from [79].

tional precision tracking in $2 < |\eta| < 2.7$. The *trigger system* covers $|\eta| \leq 2.4$, and consists of **Resistive Plate Chambers (RPCs)** in the barrel region and **Thin Gap Chamber (TGCs)** in the end-caps.

For Run-3, **New Small Wheels (NSWs)** [87] were installed, replacing part of the previous end-cap system. Providing improved precision tracking and trigger in the region $1.3 < |\eta| < 2.7$.

3.3.4 Forward detectors

In proton-proton collisions, there is a probability that the protons do not collide head-on, but instead interact in a more gentle manner, with a certain transverse distance between them. This causes the products to appear at a relatively small angle compared to the beam line, which means that they can go through the forward region and be lost. To recover these events, ATLAS uses *forward detector* [79] placed along the beam pipe. A schematic layout of these detectors is shown in Figure 3.9.

The closest *forward detector* to the interaction point (IP) is LUCID, which consists of several small Cherenkov detectors. It is located only 17 m from the IP, with the photomultipliers placed about 12 cm from the beamline. LUCID is used for luminosity measurement and luminosity monitoring.

The second detector is the **Zero Degree Calorimeter (ZDC)** a tungsten-quartz sampling calorimeter installed between the two LHC beam pipes. It is located 140 m downstream of the IP and is mainly used in heavy-ions collision. The ZDC detects neutral particles emitted at very high speeds and provides measurements of the energy dissipated in the collision.

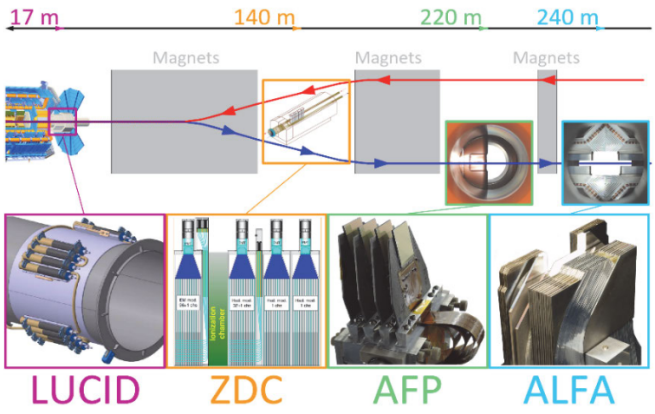


Figure 3.9: Layout of the forward detectors. Figure taken from [79].

The third forward detector is the **ATLAS Forward Proton detector (AFP)**, designed to measure protons that emerge intact from the interaction region. It is located 220 m from the IP and aims to study diffractive events, exclusive production, and photon-induced interactions.

The outermost system is **ALFA**, a scintillating-fiber tracker located 240 m downstream of the IP, on both sides of ATLAS detector. ALFA measures the protons that remain intact after the collision, mainly from elastic or diffractive scattering and provides the possibility to calibrate the luminosity measurement.

3.3.5 Future of ATLAS

The LHC does not continuously operate in Run collision periods [1], but it alternates into collision periods and **long shutdowns (LS)**, during which significant upgrades are carried out, both for the LHC and for the detectors, the full LHC timeline is shown in the Figure 3.10. The first long shutdown (LS1) was after Run 1 (2009-2012), after which the LHC entered into the Run 2 (2015-2018), operating at a center of mass energy of 13 TeV.

The most recent long shutdown (LS2) occurred from late 2018 to 2021, during which substantial upgrades were made. During Run 2 the integrated luminosity recorded by ATLAS was 2.1 times greater than originally designed. After Run 3, the LHC is expected to transition into the **High-Luminosity LHC (HL-LHC)** era, characterized by a significant increase in luminosity. The HL-LHC aims for a leveled peak luminosity of approximately 5×10^{34} $cm^{-2}s^2$, increasing its ability to generate data for further exploring rare process. It is for this reason that the ATLAS detector needs to be upgraded, specially the inner detector part, to reconstruct the do not loss efficiency in the track part, this is discussed in Section 3.6.

[1] The Run period is a specific period for data acquisition with a stable detector configuration.

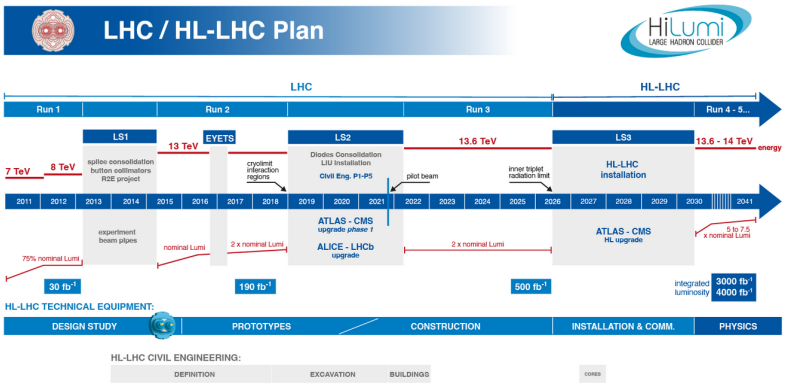


Figure 3.10: LHC and HL-LHC timeline. Figure taken from [88]

3.4 Trigger and Data acquisition (TDAQ)

The LHC is running at a collision rate of 40 MHz, and the typical event size is around 2 MB, giving an incoming data rate of $\sim$ 80 TB/s, while the ATLAS system can only store about 8 GB/s. To handle this, the ATLAS experiment relies on the Trigger and Data acquisition (TDAQ), which reduce this rate and selects only the ”interesting´´ events depending on the physics goals. An overview of the TDAQ design is shown in Figure 3.11.

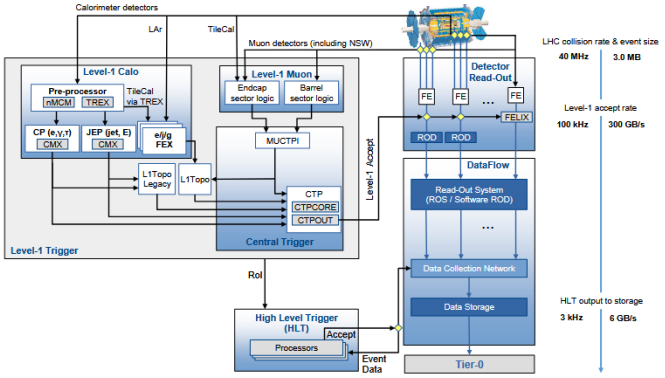


Figure 3.11: Overview of the ATLAS TDAQ system in Run 3 with emphasis on the components relevant for triggering as well as the detector read-out and data flow. Figure taken from [89].

3.4.1 Level-1 Trigger

The trigger system operates in two levels. The **Level 1 trigger** [90, 91] is the first stage and is responsible for reducing the initial 40MHz rate to less than 75 kHz. It is composed of two dedicated hardware systems: the calorimeter (L1Calo) and

muon (L1Muon) systems. The L1 topological processor (L1Topo) uses kinematic information from objects reconstructed by L1Calo and L1Muon and applies topological selections. Finally, the trigger decision is formed by the Central Trigger Processor (CTP), which receives input from both the L1Calo and L1Muon trigger systems.

3.4.1.1 Level-1 Muon

The **L1Muon** collect data from the fast muon sub-detector based on Resistive Plate Chamber (RPC), Thin Gap Chamber (TGC) and the New Small Wheel (NSW). This trigger identifies muon candidates using spatial and temporal coincidences of hits in the trigger chambers crossed by the muon trajectory (trigger roads). In the region $1.0 < |\eta| < 1.3$ signals are combined between the barrel region, the Tile Calorimeter (central region), the end-cap Inner Layer 4 and the end-cap region using the Thin Gap chambers (TGC). Another type of coincidence can be formed between the NSW and the TGC. Once the candidates are identified, the multiplicity sums are determined by the Muon-to-Central Trigger Processor Interface (MuCTPI).

3.4.1.2 Level-1 Calo

The **L1Calo** system uses energy deposits in the electromagnetic and hadronic calorimeters as input to calculate Trigger OBjects (TOBs), such as electrons, photons, taus, and small-R ($R = 0.4$) and large-R ($R = 0.8$) jets, as well as missing energy transverse.

In Run 1 and Run 2, the signal was provided through *trigger towers* with granularity varying from $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in the barrel, to 0.4×0.4 in the forward calorimeters. For Run 3, this infomation is provided by *SuperCells* (see Figure 3.12) which contain sums of four or eight calorimeter cells.

These trigger towers and SuperCells are distributed to the feature extraction (FEX) processors. Three new features extractors (FEXes) were introduced for Run 3 to identify different signatures and patters in the energy deposits, improving the particle identification. The **electromagnetic Fex (eFex)**, for identification of electron, photons, and hadronic taus. The jet Fex (jFEX), reconstruction of large-R and small-R jets, hadronic tau decays (three-prong), and missing transverse energy objects. Finally, global Fex (gFex), for processing large-R jets and missing transverse energy objects over the whole calorimeter.

3.4.1.3 Level-1 Topo

The L1Topo system receives Trigger OBjects (TOBs), which contain kinematic and further qualifying information from the L1Calo and L1Muon and applies real

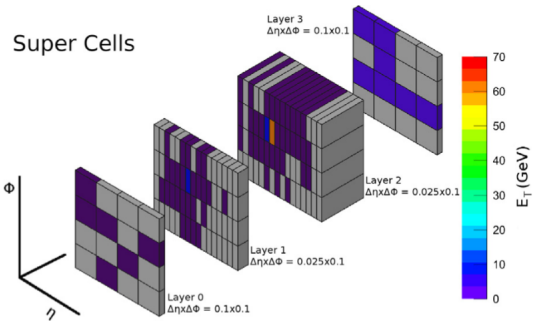


Figure 3.12: Geometrical representation of the Super Cells with the energy depositions of a simulated electron shower. Figure taken from [92].

time kinematic and angular selections.

This trigger was originally added in 2016 to lower energy threshold and increase signal purity while reducing L1 rates without losing interesting physics. For Run 3, the L1Topo hardware was upgraded with three new modules (TOPO1, TOPO2 and TOPO3) to support the new L1Calo input format. Each module contains two Field-Programmable Gate Array (FPGA) where the algorithms are executed.

The **TOPO1** module implements object counting algorithms, in which it counts the number of objects passing given energy thresholds, the resulting multiplicity bits are transmitted to the CTP. The **TOPO2** and **TOPO3** modules are used for topological selections, consistent with the functionality provided by the L1Topo system in Run 2.

3.4.2 High Level Trigger (HLT)

The **High level trigger (HLT)** receives as input the events accepted by the L1 trigger. Here, *online* reconstruction algorithms process the detector data with a much higher level of detail than L1. These algorithms can run over either the full detector or only within specific Regions-of-Interest (RoIs) identified by the CTP. The HLT uses the same reconstruction framework as the offline reconstruction, but employs faster reconstruction techniques, often guided by RoIs, while preserving precision.

After the event is reconstructed, the HLT applies physics based selection targeting specific types of objects, and applies early rejection to reduce the event rate. The output rate of the HLT is approximately 3 kHz.

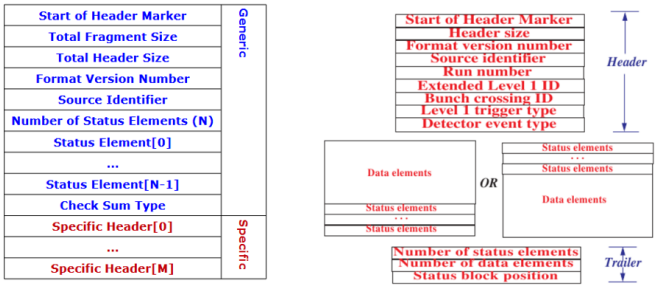


Figure 3.13: Generic Header format on left and ROD fragment format on right. Figure taken from [93]

3.4.3 Data Acquisition (DAQ) system

The **Data acquisition** (DAQ) system is responsible for supporting the operation of the two trigger levels. The ATLAS subdetector systems produce analog signals that are digitized by the **Front-End (FE)** electronics. These FE modules continuously buffer the incoming data while waiting for a decision from the Level-1 (L1) trigger. When an event is accepted by L1, the corresponding data fragments are transferred from the FE to the ReadOut Drivers (RODs). The RODs format the raw detector information into structured data fragments following the ATLAS Bytestream standard, which is based on a generic header, shown in Figure 3.13 (left), followed by a set of bits that describe the event content, as illustrated in Figure 3.13 (right).

Once the RODs have assembled the data, they forward the fragments to the **ReadOut System (ROS)**. The ROS consists of a set of commodity server machines hosting custom-built I/O cards. These servers receive the Bytestream fragments sent by the detector readout electronics and store them in internal memory buffers known as **ReadOut Buffers (ROBs)**. A schematic overview of the general event structure is shown in Figure 3.14.

The HLT processes each event by requesting only the ROB fragments necessary for specific reconstruction algorithms. The HLT decompresses the byte stream data, performs online selection, and accepts or rejects the event. Following the HLT decision, all accepted events are transferred to a dedicated server cluster, known as **Sub-Farm Outputs (SFO)**, where they are packaged, compressed, and finally sent to offline storage for permanent recording.

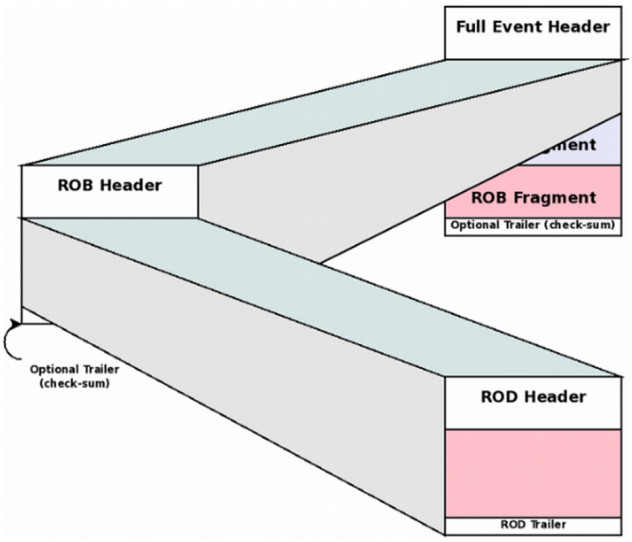


Figure 3.14: General Event Format. Figure taken from [93]

3.5 Reconstruction and Identification

3.5.1 Electrons

Electron reconstruction [94] is based on the combination of energy deposits in the electromagnetic (EM) calorimeter and charged-particle tracks in the inner detector. This is based on three fundamental components of the electron signature. First, localized clusters of energy deposits are recontructed in the EM calorimeter (**seed-cluster reconstruction**). Second, charge-particle tracks identified in the inner detector (**track reconstruction**). Finally, tracks are geometrically matched to the EM cluster in $\eta \times \phi$ space to form the final electron candidate (**electron-candidate reconstruction**).

Electron identification [94] is performed using a likelihood-based method that combines variables from both the tracking system and the calorimeter, such as shower-shape variables and tack-cluster matching quantities. Depending on the output of the likelihood discriminant, multiple operating points are provided, typically referred to as **Loose**, **Medium**, and **Tight**. These correspond to different identification efficiencies and background rejection levels.

3.5.2 Muons

Muon reconstruction [95] uses independent information from the Inner Detector (ID) and the Muon Spectrometer (MS). These two reconstructions are then combined and performed using different algorithms based on the information provided by the ID, MS, and calorimeters. Four muon *types* are defined depending

on which sub-detectors are used in the reconstruction.

For **Combined (CB) muon**, the reconstruction is performed independently in the ID and MS, producing two separate tracks. A global refit is then applied to combine the hits from both the ID and MS into a single track.

In the case of **Segment-tagged (ST)**, here the muon reconstruction is mainly based on an ID track, which is extrapolated to the MS and associated with at least one local track segment in the MDT or CSC chambers. This type of reconstruction is generally used when the muon have *low* p_T, cross only one layer of MS chambers, or transverse regions with reduced MS acceptance.

For **Calorimeter-tagged (CT) muons**, the reconstruction does not use the MS chambers. It start from an ID track and searches for energy deposits in the calorimeter compatible with a minimum-ionizing particle (MIP). This muon type has the lowest purity among all muon types and is optimized for the central region $|\eta| < 0.1$ and for a momentum range of $15 < p_T < 100$ GeV.

Finally, the **Extrapolated (ME) muons** are reconstructed using only the MS track information, together with a loose requirement of compatibility with originate from the interaction point (IP). This reconstruction is mainly used to extend the muon acceptance to the region $2.5 < |\eta| < 2.7$, which is not covered by the ID.

The identification of muons [95] is designed to distinguish *prompt* muons from background, which commonly originates from hadron decays. To achieve this, different variables are used to compare measurements from the ID and the MS. One of this is the *q/p significance*, defined as the absolute value of the difference between the charge to momentum ratios measured in the ID and the MS. Another variable is ρ', defined as the absolute value of the difference between the transverse momentum measurements in the ID and the MS, divided by the p_T of the combined track. In addition, the normalized χ^2 of the combined track fit is used.

To guaranty a robust momentum measurement, specific requirements on the number of hits in the ID and MS are applied. Based on these identification criteria, four *working points* are defined. The **Loose** working point is designed to maximise the reconstruction efficiency while still providing good-quality muon tracks, and it is optimised for the reconstruction of Higgs boson candidates in the four-lepton final state. The **Medium** working point, aims to minimise the systematic uncertainties associated with muon reconstruction and calibration. The **Tight** working point, considers only CB muons with hits in at least two MS stations that also satisfy the *Medium* selection criteria, and it is chosen to maximise the muon purity at the cost of some efficiency. Finally, the **High-p_T** working point is optimised to

maximise the momentum resolution for tracks with transverse momentum above 100 GeV and it is mainly used in searches for high-mass Z' and W' resonances.

3.5.3 Jets

The reconstruction of jets use the energy deposits in the calorimeters to form topological clusters. These clusters are grouped into jets through the **anti-k_t** algorithm [96], which is based on a sequential recombination. In this algorithm, every object (cluster or pseudojet) is iteratively combined with others according to a distance metric that depends on the transverse momentum p_T and the angular position (η, ϕ), defined as

$$d_{ij} = \min(p_{Ti}^{-2}, p_{Tj}^{-2}) \frac{\Delta R_{ij}^2}{R^2}, \tag{3.3}$$

$$d_{iB} = p_{Ti}^{-2} \tag{3.4}$$

where $\Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$ is the angular distance between two energy clusters and p_{Ti}, η_i and ϕ_i are respectively the transverse momentum, pseudo-rapidity and azimuth of the cluster i. The radius parameter R determines the characteristic size of the jet. Typically , $R = 0.4$ defines *small-R jets*, while $R = 1$ is used for *large-R jets*.

The choice of the parameter R determine how nearby energy deposits are grouped into jets. The trivial case, where a high-energy cluster has no neighbors within a distance $2R$, all particles within R around it form a single jet. Other case is two high-energy clusters satisfying $R < \Delta R_{12} < 2R$, the algorithm reconstructs two separate jets. Finally, if two high-energy clusters lie within $\Delta R_{12} < R$, the anti-k_T algorithm merges them into a single jet.

The algorithm at each iteration compares the distances d_{ij} and d_{iB}. If d_{ij} is the smallest, the clusters i and j are combined into a new pseudojet. If d_{iB} is the smallest, the cluster i is identified as a final jet. This procedure repeats until no cluster remains.

3.5.4 b-tagging

Flavour tagging, or identification of jets containing b-hadrons [97], is the key for measuring decays such as $H \to b\bar{b}$. The techniques implemented by ATLAS rely on several distinctive properties of hadrons b-quarks, in particular their relatively long lifetime, high mass, and characteristic decay multiplicity. These properties allow b-jets to be distinguished between *light*-jets and c-jets.

To identify b-jets, ATLAS employs several b-tagging algorithms that use infor-

mation from the Inner Detector (ID), the calorimeters and the muon spectrometer. The b-tagging strategy is organized into two main levels, **low-level** and *high-level taggers*.

3.5.4.1 Low-Level

The first level corresponds to the low-level b-taggers [97], which use the direct detector information to extract discriminating features of jets. Several types of algorithms are used at this stage.

The **Impact Parameter (IP) algorithms** are based on the significance of the track impact parameters. Two variants are used, the **IP2D** tagger, which uses the transverse impact parameter significance d_0/σ_{d_0} as a discriminant, and the **IP3D** which utilizes both the transverse and the longitudinal impact parameter significances, d_0/σ_{d_0} and $z_0 \sin\theta/\sigma_{z_0 \sin\theta}$, in a two-dimensional template to account for their correlation. These algorithms are based on a **Log-Likelihood Ratio (LLR)** discriminant, computed as the sum of the per-track contributions

$$\sum_{i}^{N} \log(\frac{p_b}{p_u}) \tag{3.5}$$

where N is the number of tracks associated with a given jet, and p_b, p_u are the probability density functions (PDFs) corresponding to the b-jet and light-flavour jet hypotheses, respectively. These PDFs are obtained from reference histograms of the impact parameter significances and are derived from Monte Carlo simulations.

The **Recurrent Neural Network Impact Parameter (RRNIP)** algorithm is designed to further improve the separation between b-jets and light-flavour jets. This RNN uses as input features; the transverse (S_{z0}) and longitudinal (S_{d0}) impact parameters, the fraction of transverse momentum carried by the track relative to the jet p_T (p_T^{frac}), the angular distance between the track and the jet axis ($\Delta R(\text{track}, \text{jet})$), as well as the output of the IP2D and IP3D output. The resulting probabilities for the jet to originate from a *b-jet* (p_b), *c-jet* (p_c), light-flavour jet (p_{light}), or τ-jet (p_τ). The RNN discriminant is defined as

$$D_{RNN}(b) = \ln \frac{p_b}{f_c p_c + f_\tau p_\tau + (1 - f_c - f_\tau) p_{light}} \tag{3.6}$$

where f_c and f_τ are the c-jet and τ-jet fraction, respectively, determined from the MC simulation.

Another algorithm of low-level is based on **Secondary vertex-based b-tagging**

algorithm. Two algorithms are used, one reconstructs a **Single displaced Vertex (SV1)**, while the other reconstructs the decay topology along the jet axis **(JetFitter)**. The SV1 algorithm aims to reconstruct a single displaced secondary vertex within a jet. All possible two-track vertices are rejected if they were likely to originate from the decay of a long-lived particle (K_s or Λ), photon conversions, or hadronic interactions with the detector material. An inclusive secondary vertex is then reconstructed from the remaining tracks. From this vertex, several discriminating variables are extracted, including the number of associated tracks, the invariant mass of the vertex, the fraction of the jet energy carried by the vertex, and the three-dimensional decay length significance. These variables are later used by the high-level taggers.

The **JetFitter** algorithm is an inclusive vertexing method that reconstructs the topological decay structure of heavy-flavour hadrons inside a jet. It is based on a modified Kalman filter and uses the intersections of particle tracks with the jet axis to reconstruct the full decay topology, including single-prong vertices. Variables such as track multiplicity, vertex mass, energy fraction, and three-dimensional decay length significance are extracted and used as inputs to the high-level taggers.

3.5.4.2 High-Level tagger

The high-level taggers [97] combine the outputs of the low-level algorithms into a single discriminant. ATLAS employs two main variants. The first is the **MV2** tagger, which is based on a Boosted Decision Tree (BDT) trained using the ROOT Toolkit for Multivariate Data Analysis (TMVA). In addition to the low-level tagger outputs, the jet transverse momentum p_{T} and pseudorapidity η are also used as input variables.

The second, and more recent algorithm is **DL1** [98], which is based on a deep neural network (DNN). This algorithm takes as input the outputs of the low-level taggers and provides probabilities for the jet to originate from a *b-jet* (p_b), *c-jet* (p_c), light-flavour jet (p_{light}), or τ-jet (p_τ). The corresponding discriminant is defined as

$$D_b = \log \frac{p_b}{(1 - f_c)p_{\mathrm{light}} + f_c p_c}, \tag{3.7}$$

where f_c is a free parameter that controls the balance between light-flavour and *c-jet* rejection for a given *b-jet* efficiency and can be optimised after training.

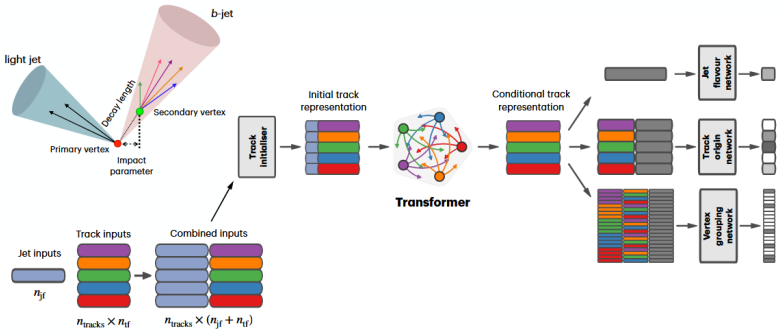


Figure 3.15: Illustration of the GN2X algorithm with jet and track input variables, discriminating between jet flavours by exploiting secondary vertices and other properties stemming from the displaced decays of b-hadrons. The jet features are copied for each track associated with the jet. The combined vectors are then fed into a per-track initialisation network, followed by a transformer encoder and a global representation of the jet. $n_{jf}(n_{tf})$ corresponds to the number of jet (track) features. The pooled jet representation and output track embeddings are provided as inputs to the three task-specific networks. Details of the GN2X architecture are summarised in the Methods section. Figure taken from [100]

3.5.4.3 *Graph Neural Network for jet tagging*

Identifying decays of high-p_T (boosted) Higgs boson into pairs of bottom or charm quarks with high efficiency and background rejection is a challenge for ATLAS, since the decay products are highly collimated and are typically reconstructed as a single large-R jet (see Subsection 2.2.3). In this regime, the tagging methods described previously lose discrimination power. The **Graph Network, 2-prong, eXtended (GN2X)** algorithm [99, 100] was developed to address this issue.

GN2X employs a **Tansformer-based neural network** architecture [101] and takes as input the properties of all reconstructed tracks associated with a jet, including impact parameters, transverse momentum, and geometrical relations with respect to the jet axis, together with global jet-level information. The jet-level features are replicated for each associated track and concatenated with the corresponding track features. The resulting jet–track sequences are processed by a per-track initialization network and subsequently passed to a *Transformer encoder* with an attention mechanism, which enables the algorithm to learn correlations among all tracks simultaneously.

Through this approach, GN2X recognizes characteristic patterns of heavy-flavour decays, such as the presence of multiple displaced tracks and topologies compatible with b-hadron decays. A schematic overview of the GN2X architecture is shown in Figure 3.15.

The network outputs a set of probability scores that quantify the likelihood for

a given jet to originate from a Higgs boson decaying into bottom quarks, a Higgs boson decaying into charm quarks, a top quark, or from multijet background processes (QCD background). These probabilities are combined into a discriminant D_{Hbb}^{GN2X}, defined as

$$D_{Hbb}^{GN2X} = \ln\left(\frac{p_{Hbb}}{f_{Hcc} \cdot p_{Hcc} + f_{top} \cdot p_{top} + (1 - f_{Hcc} - f_{top}) \cdot p_{QCD}}\right) \tag{3.8}$$

where f_{Hcc} and f_{top} are free parameters that control the relative weight of the $H(c\bar{c})$ and top-jet hypotheses with respect to the multijet background, thereby allowing the optimization of the trade-off between charm-, top-, and light-flavour jet rejection.

3.5.5 Hadronic τ leptons

The reconstruction and identification of the hadronic tau leptons in ATLAS [102] is designed to exploit the characteristic decay topology of taus, which predominantly decay into one or three charged hadrons (1-prong or 3-prongs) accompanied by neutral particles and at least one neutrino. This algorithm aim to reconstruct only the visible part of the decay ($\tau_{\text{had-vis}}$) and distinguish it from other particles with similar signatures in the detector, such as jets or electrons.

The reconstruction procedure starts by selecting $\tau_{\text{had-vis}}$ candidates seeded by jets formed using the anti-k_t algorithm (see Subsection 3.5.3) using a small radius parameter $R = 0.4$. These jets are built from calorimeter energy clusters and are required to have $p_T > 5$ GeV and $|\eta| < 2.5$.

Then, a **vertex association** is performed using the information from the ID. Since τ leptons decay travels away from the primary vertex [2] before decaying, a dedicated **Tau Jet Vertex Association (TJVA)** algorithm is implemented. This method evaluates the fraction (f_{p_T} Eq. 3.9) of the squared summed transverse momentum p_T of tracks associated with a given vertex within a cone of $\Delta R < 0.2$ around the seed jet. The vertex with the largest fraction is selected as the tau production vertex, ensuring that the reconstructed tau candidate is associated with the correct hard-scatter interaction.

$$f_{\text{p}_\text{T}} = \frac{\sum p_T^2(\text{track associated with the vertex})}{\sum p_T^2(\text{all tracks})} \tag{3.9}$$

[2]The primary vertex is defined as the vertex with the highest sum of the squared transverse momenta of the associated tracks, $\sum p_T^2(\text{track associated with the vertex})$

Not all geometrically associated tracks with a tau candidate originates from the tau decay. To address this, a **track classifier algorithm** is applied, which assigns each track to one of four categories. **Tau Tracks (TT)** correspond to charged decay products of the τ lepton and are used to determine the τ charge and prong multiplicity. **Conversion Tracks (CT)**, tracks from electrons and positrons that are created from photon conversion in the detector. The **Isolation Tracks (IT)**, tracks which are likely originating from quark or gluon jets arising from the remnants of the hard scattering interaction. Finally, the **Fake Tracks (FT)**, tracks that do not belong to the previous categories, include misreconstructed tracks and tracks from pile-up interactions.

Identification of $\tau_{\text{had-vis}}$ is performed using a **Recurrent Neural Network (RNN)** trained to discriminate the truth $\tau_{\text{had-vis}}$ from misidentified $\tau_{\text{had-vis}}$ originating from quark and gluon-initiated jets (QCD jets). The RNN exploits a combination of low-level variables that decribe individual tracks and calorimeter clusters, together with high-level observables derived from tracking and calorimeter information. The algorithm provides three working points (Loose, Medium and Tight) allowing the efficiency and background rejection to be tuned according to the needs of a given analysis.

In addition, a dedicated electron veto is applied, as energetic electrons can mimic hadronic tau signatures, particularly in the 1-prong decay mode. This veto is also based on an RNN and exploits the characteristic energy deposition in the electromagnetic calorimeter, as well as the relationship between the calorimeter energy and the momentum of the associated track.

3.5.6 Boosted Di-Tau

Reconstructing and identifying boosted decays of two hadronically decaying tau leptons is not possible using the standard algorithms described in Subsection 3.5.5. This limitation arises from the use of anti-k_t seed jets with a small radius parameter $R = 0.4$ (see Subsection 3.5.3). In boosted Higgs decays, as discussed in Subsection 2.2.3, the decay products of the two tau leptons are highly collimated and therefore tend to merge into a single small-R jet. As a consequence, a dedicated di-tau reconstruction algorithm is employed to recover sensitivity in this regime [103].

This method starts by reconstructing a large-radius seed jet using the anti-k_t algorithm with a radius parameter of $R = 1.0$. This large-R jet is interpreted as a di-tau object containing the highly collimated decay products of the two tau leptons, and it is required to have a transverse momentum greater than 300 GeV

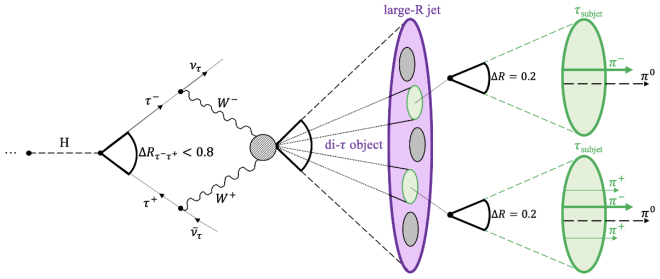


Figure 3.16: Schematic of the di-tau reconstruction algorithm. The diagram illustrates a Higgs boson decaying into a pair of tau leptons, with the resulting hadronic activity reconstructed as a large-R jet (shown in purple). Within this large-R jet, re-clustered subjets are identified, where the two leading subjets in transverse momentum (highlighted in green) are designated as the single-tau candidates. The remaining subjets (shown in gray) are located in the surrounding isolation region (which is defined as the purple area) and are not considered in the reconstruction. Figure taken from [104].

to ensure that it originates from a boosted decay. Within this large-R jet, a re-clustering procedure is performed to reconstruct smaller-radius subjets using the anti-k_t algorithm with $R = 0.2$. The two leading subjets in transverse momentum are considered as candidates for the individual hadronic tau decays and are required to have a minimum transverse momentum of $p_T > 10$ GeV. A schematic representation of the di-tau reconstruction procedure is shown in Figure 3.16.

Once the subjets are reconstructed, tracks are associated to each of them. The associated tracks are required to have a transverse momentum $p_T > 1$ GeV, at least two pixel hits, and a total of at least seven pixel and SCT hits. In addition, the transverse impact parameter and the longitudinal impact parameter must satisfy $|d_0| \leq 1.0$ mm and $|z_0 \sin\theta| \leq 1.5$ mm. To ensure correct association of tracks with the primary vertex in the presence of *pile-up* a **TJVA** algorithm is employed. This procedure guaranties that the reconstructed tracks originate from the hard-scatter interaction rather than from overlapping soft proton–proton collisions. A reconstructed di-tau candidate is required to contain at least two subjets, with the two highest-p_T subjets expected to correspond to the two hadronically decaying tau leptons.

3.5.6.1 Di-Tau object identification

The di-tau identification algorithm is designed to separate truth di-tau decays from the large contribution from multi-jet background originating from QCD processes. The initial identification strategy for boosted di-tau objects in ATLAS was based on a Boosted Decision Tree (BDT) approach [103]. In 2025, a new identifi-

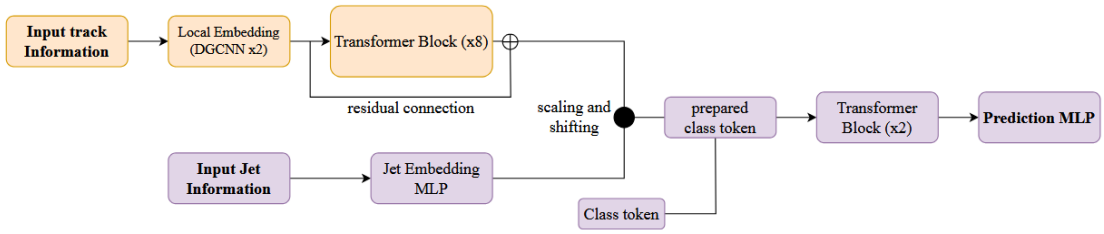


Figure 3.17: A summary of the neural network architecture used to train the Omni tagger. The orange blocks represent the main part, the PET body, and the purple blocks represent the classifier head. Figure taken from [104]

cation technique, known as the Omni tagger, was introduced [104]. This tagger is based on the **Point-Edge Transformer (PET)** architecture, which is capable of leveraging lower-level tracking information. This architecture combines elements of **Graph Neural Networks (GNNs)** with Transformer mechanisms, allowing the model to capture both local geometric features and global correlations within the jet. This capability is particularly well suited for identifying boosted di-tau decays, where the substructure of the jet plays a crucial role in discriminating signal from background.

The Omni tagger processes per-jet track information through sequential **Dynamic Graph Convolutional Neural Network (DGCNN)** layers, which establish connections between nearby tracks in the $\eta - \phi$ plane. These layers learn local geometric representations based on track-level variables. The resulting representations are then passed to a stack of Transformer blocks, which use self-attention to capture long-range correlations and dynamically weigh the importance of each track and its learned relationships.

Following the PET body, the classification head combines the track-level embeddings with global jet-level features, summarized in Table **??**, many of which are shared with the traditional BDT-based approach. These global features are used to modulate the learned representations via affine transformations. In addition, a trainable class token is introduced to act as a global information aggregator. The final embedding of this token is passed to a **Multi-Layer Perceptron (MLP)** to produce the *binary classification* output, distinguishing between truth di-tau decays and multi-jet background events. A schematic representation of the Omni tagger architecture is shown in Figure 3.17.

3.5.6.2 *Identification di-tau with Transformer Neural Network*

Another approach to reconstruct and identify boosted di-taus in ATLAS (particularly relevant for boosted final states such as $HH \rightarrow b\bar{b}\tau\tau$) use the same

Variable	Definition
$E^{s_{j1}}_{\Delta R<0.1}/E^{s_{j1}}_{\Delta R<0.2}$ and $E^{s_{j2}}_{\Delta R<0.1}/E^{s_{j2}}_{\Delta R<0.2}$	Ratios of the energy deposited in the core to that in the full cone, for the sub-jets s_{j1} and s_{j2}, respectively. Also denoted as $f^{\text{lead}}_{\text{core}}$ and $f^{\text{subl}}_{\text{core}}$.
$p_T^{s_{j2}}/p_T^{\text{LRJ}}$ and $(p_T^{s_{j1}}+p_T^{s_{j2}})/p_T^{\text{LRJ}}$	Ratio of the p_T of s_{j2} to the di-τ large-radius jet p_T and ratio of the scalar p_T sum of the two leading sub-jets to the di-τ large-radius jet p_T, respectively. Also denoted as $f^{\text{subl}}_{\text{subjet}}$ and f_{subjet}.
$\log(\sum p_T^{\text{iso-track}}/p_T^{\text{LRJ}})$	Logarithm of the ratio of the scalar p_T sum of the iso-tracks to the di-τ large-radius jet p_T. Also denoted as $f^{\text{iso}}_{\text{track}}$.
$\Delta R_{\text{max}}(\text{track}, s_{j1})$ and $\Delta R_{\text{max}}(\text{track}, s_{j2})$	Largest separation of a track from its associated sub-jet axis, for the sub-jets s_{j1} and s_{j2}, respectively. Also denoted as $R^{\text{lead}}_{\text{max}}$ and $R^{\text{subl}}_{\text{max}}$.
$\sum\left[p_T^{\text{track}}\Delta R(\text{track}, s_{j2})\right]/\sum p_T^{\text{track}}$	p_T-weighted sum of ΔR of the tracks matched to s_{j2} with respect to its axis. Also denoted as $R^{\text{subl}}_{\text{tracks}}$.
$\sum\left[p_T^{\text{iso-track}}\Delta R(\text{iso-track}, s_j)\right]/\sum p_T^{\text{iso-track}}$	p_T-weighted sum of ΔR between iso-tracks and the nearest sub-jet axis. Also denoted as $R_{\text{isotracks}}$.
$\log(m^{\text{tracks},s_{j1}}_{\Delta R<0.1})$ and $\log(m^{\text{tracks},s_{j2}}_{\Delta R<0.1})$	Logarithms of the invariant mass of the tracks in the core of s_{j1} and s_{j2}, respectively. Also denoted as $m^{\text{lead}}_{\text{core}}$ and $m^{\text{subl}}_{\text{core}}$.
$\log(m^{\text{tracks},s_{j1}}_{\Delta R<0.2})$ and $\log(m^{\text{tracks},s_{j2}}_{\Delta R<0.2})$	Logarithms of the invariant mass of the tracks with $\Delta R<0.2$ from the axis of s_{j1} and s_{j2}, respectively. Also denoted as $m^{\text{lead}}_{\text{tracks}}$ and $m^{\text{subl}}_{\text{tracks}}$.
$\log(\lvert d_0^{s_{j1},\text{lead-track}}\rvert)$ and $\log(\lvert d_0^{s_{j2},\text{lead-track}}\rvert)$	Logarithms of the closest distance in the transverse plane between the primary vertex and the leading track of s_{j1} and s_{j2}, respectively. Also denoted as $\lvert d^{\text{lead}}_{0,\text{leadtrack}}\rvert$ and $\lvert d^{\text{subl}}_{0,\text{leadtrack}}\rvert$.

Table 3.1: Jet features used in the training of Omni for di-τ identification. Here, LRJ refers to the large-radius jet associated with the di-τ object, while s_{j1} and s_{j2} denote the leading and subleading sub-jets, respectively, ordered by transverse momentum p_T. Table taken from [104].

neural network architecture as the **GN2X** b-tagging method described in Subsection 3.5.3. This approach is an extension of the jet classification network to include the $H(\tau\tau)$ final state. The algorithm is referred as *GN2XTau* [105], and exploits charge particle tracks togheter with other constituents of large-R **Unified Flow Objects (UFO)** [106] jets to identify characteristic signatures of hadronic τ decays.

As in GN2X, the GN2XTau use the same discriminant from Equation 3.8, including the $H(\tau\tau)$ probabilities. The adapted discriminant is defined as

$$D^{GN2X}_{H\tau\tau} = \ln\left(\frac{p_{H\tau\tau}}{f_{Hbb}\cdot p_{Hbb}+f_{Hcc}\cdot p_{Hcc}+f_{top}\cdot p_{top}+(1-f_{Hbb}-f_{Hcc}-f_{top})\cdot p_{QCD}}\right) \tag{3.10}$$

where $p_{H\tau\tau}$,p_{Hbb},p_H,p_{Htop} and p_{QCD} are the network output probabilities and f_{Hbb}, f_{Hcc} and f_{top} are free parameters that control the relative weight of the $H(b\bar{b})$,$H(c\bar{c})$ and top-jet hypotheses with respect to the multijet background (QCD).

3.6 New Inner Detector Tracker (ITk)

The upgrade of the LHC to the High-Luminosity LHC (HL-LHC) is expected to deliver very high instantaneous luminosities, with 25 ns between beam crossings. However, this luminosity will stabilize within a typical range of $5.0-7.5\times$

10^{34} cm^{-2}s^{-1}, corresponding to an average of about 200 proton-proton interactions per crossing.

The existing ATLAS Inner Detector was designed for 10 years of operation at a constant luminosity of 10^{34} cm^{-2}s^{-1}, with a centre-of-mass energy of $\sqrt{s} = 14$ TeV, and an average pile-up ($\langle\mu\rangle$) of 23 interactions per crossing (see Section 3.3). However, after more than a decade of LHC operation, peak luminosities and pile-up have already exceeded these values, where the current ID performs well, but this is not sufficient for the extreme conditions expected at the HL-LHC.

The current Pixel and SCT detectors were designed to withstand *radiation damage* equivalent to an integrated luminosity of 400 fb^{-1} and 700 fb^{-1}, respectively. After the installation of the Insertable B-Layer (IBL), the effective radiation tolerance in some areas, particularly the SCT end-caps, is reduced, while the IBL itself can resist up to 850 fb^{-1}. Beyond these limits, detector performance decreases.

Furthermore, the *front end-electronics* of both Pixel and SCT were designed to handle events with around 50 proton-proton interactions per crossing, corresponding to a luminosity of approximately 2×10^{34} cm^{-2}s^{-1}. Beyond this luminosity, the limited buffering of the links between the on-detector electronics (Pixel/SCT chips) and the read-out drivers (RODs) causes data losses.

With a $\langle\mu\rangle =\sim 200$ average proton-proton interactions inside the current detector, the confusion due to additional hits without the compensating increase in granularity, will compromise the efficiency of the pattern recognition and track-finding efficiency. This makes it difficult to resolve nearby tracks, such as boosted top-quark jets, significantly limiting the physics reach.

Finally, the current ID does not provide tracking information to the first-level (L1) trigger. Adding tracking information to the trigger object directly benefits the physics performance.

These limitations necessitate the complete replacement of the Inner Detector for the Phase-II upgrade of the experiment. This new Inner Tracker Detector (ITk) is designed for 10 years of operation at instantaneous luminosities of 7.5×10^{34} cm^{-2}s^{-1}, 25 ns between bunch crossings and an integrated luminosity of 3000 fb^{-1}.

The ITk consists of two main subsystems. The outermost subsystem is the **ITk Strip** [22], which consists of a four-layer barrel section and two end-caps equipped with six disks each. This subsystem covers $|\eta| \leq 2.7$, where the strip barrel extends from -1400 mm to $+1400$ mm along the z-axis. The two innermost strip layers use short strips of 24.1 mm length, while the two outer layers have longer strips with 48.2 mm. An overview of the strip geometry is shown in

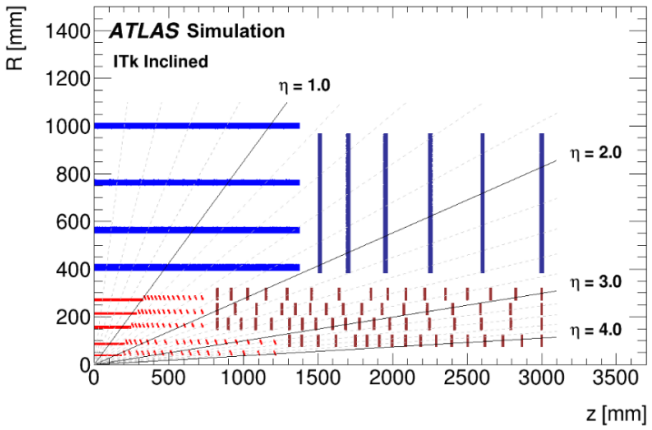


Figure 3.18: Schematic layout of the ITk for the HL-LHC phase of ATLAS. Here only one quadrant and only active detector elements are shown. The horizontal axis is the axis along the beam line with zero being the interaction point. The vertical axis is the radius measured from the IP. The outer radius is set by the bore of the solenoid. Figure taken from [22]

Figure 3.18 (blue).

The innermost subsystem is called **ITk Pixel** [23], consists of five barrel layers and four end-cap ring layers, covering up to $|\eta| = 4$. The Pixel detector will consist of five silicon tracking layers (L0-L4). The two innermost layers (L0-L1) are composed of a central region of staves and coupled-rings, designed to reduce material and service requirements in the most radiation-intense region. These layers are placed inside an Inner Support Tube (IST), allowing for their potential replacement if needed.

The other three layers (L2-L4) are made up of longer staves in the central area, surrounded by inclined half-rings in the outer regions, tilted between $55° - 67°$ with respect the beam axis. Further out, additional half-rings complete the coverage. The positions of the rings along the z-axis are optimized independently to ensure hermetic coverage for tracks coming from the collision region. An schematic layout of the ITk Pixel is shown in Figure 3.18 (red).

3.6.1 ITk Strip detector

The ITk strip detector covers a total silicon area of $\sim$ 165 m^2, it consists of a central barrel region extending between $|z| = 1.4$ m and two end-cap structures that extend to $|z| = 3$ m. The barrel composed of four cylinders, each 2.8 m long, surrounding the beam-line. On the other hand, the strip end-caps have six disks per side (end-cap A and end-cap C). A schematic layout of the ITk Strip Detector is shown in Figure 3.19.

This strip subsystem is built from **staves** in the barrel and **petals** for the end-

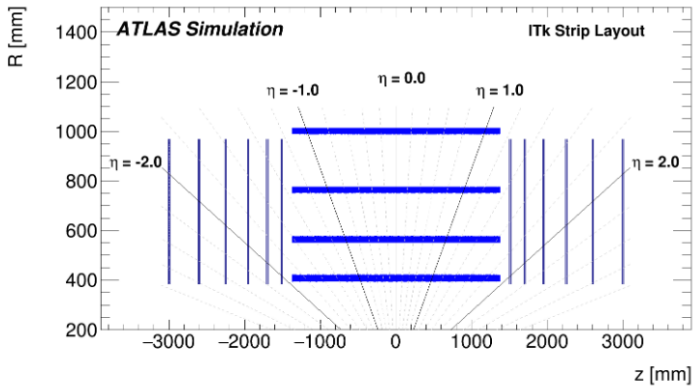


Figure 3.19: The Layout of the ITk Strip Detector. Here only the upper half of the Strip Detector is shown. The horizontal axis is the axis along the beam line with zero being the interaction point. Figure taken from [22]

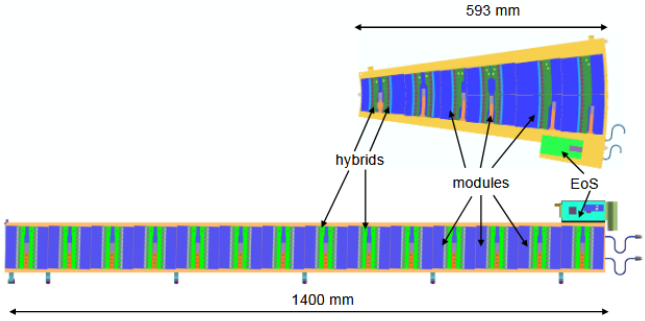


Figure 3.20: Local support components overview: End-cap petal (upper) and barrel stave (lower) components overview. In both cases modules are mounted directly onto a rigid carbon-fibre sandwich structure. The interface to the off-detector electronics is the End-of-Substructure card at one of the ends of the stave or petal. Titanium cooling pipes are directly embedded into the support structures. Figure taken from [22]

caps, which provide the local support for the modules and houses the common electrical, optical and cooling services. The four barrel layers contain 392 staves with modules on both sides (192 staves on each side of z = 0). Each stave holds 28 modules (14 on each stave side). Have two types of strips, **short-strips (SS)** in the two inner layers of the barrel, which consist of a 24.1 mm long strip, and long-strips (LS) in the two outer layers, with a length of 48.2 mm.

In the end-caps, each disk is populated with 32 identical petals. Each petal has nine modules per side, using six different sensor geometries. An overview of the stave and petal support structures is shown in Figure 3.20.

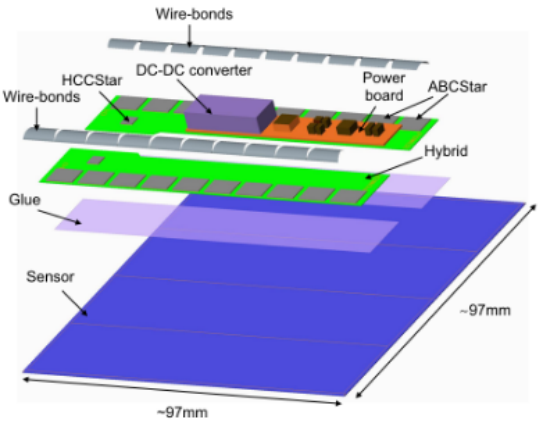


Figure 3.21: Exploded view of a short-strip barrel module with all relevant components. Long-strip modules and end-cap modules feature the same component groups. Figure taken from [22].

3.6.1.1 *Module distribution and content*

The strip system contains 17888 modules distributed in staves and petals. Each module consists of one silicon sensor and one or more low-mass printed circuit boards (PCBs), called hybrids, hosting the read-out ASICs [3] (ABCStar and HCCStar). A short-strip barrel module is presented in Figure 3.21

In the end-cap region, each petal contains nine modules per side, organized in six rings (R0-R5) (see Figure 3.22). The three inner rings (R0-R2), located closer to the beam axis, hosts one module per ring, each with one or two hybrids. The outer rings (R3-R5) contain two side-by-side modules, each read out by a single hybrid spanning the two neighboring modules. The distribution of the modules, hybrids, strip counts and ABCStars chips in the petals is in Table 3.2.

For the barrel section, there are two types of modules, **short-strip modules** in the two innermost layers, and **long-strip modules** in the two outer layers. For the innermost layers, each stave contains 14 modules per side, with two hybrids hosting 10 ABCStar chips. These modules correspond to the $|\eta| < 28$ index (excluding 0), with 1280 strips per η row. On the other hand, in the long-strip modules, each stave also contains 14 modules per side, but each module includes a single hybrid, each with 10 ABCStar chips, with a $|\eta| < 56$ index (excluding 0) and also with 1280 strips per η row.

There are two types of strips **near strips** and **far strips**, defined by the η index associated with each stave or petal. An schematic representation of these two

[3] Application-specific integrated circuit (ASIC) is a designed microchip built to perform a specific function or task.

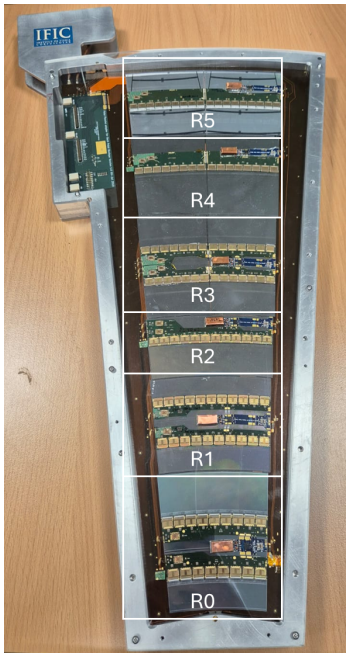


Figure 3.22: Geometry distribution of the rings R0 to R5 on the petals.

Module Ring	**Hybrid**	**Modules in ϕ**	**Disks**	**ABCStars**	**η Module**	**Strips per η**
R0	H0	32	6	8	0 (Row 0) 1 (Row 1)	1024 1024
	H1	32	6	9	2 (Row 0) 3 (Row 1)	1152 1152
R1	H0	32	6	10	4 (Row 0) 5 (Row 1)	1280 1280
	H1	32	6	11	6 (Row 0) 7 (Row 1)	1408 1408
R2	H0	32	6	12	8 (Row 0) 9 (Row 1)	1536 1536
R3	H0/H1	64	6	7	10 (Row 0) 11 (Row 1)	896 896
	H2/H3	64	6	7	12 (Row 0) 13 (Row 1)	896 896
R4	H0/H1	64	6	8	14 (Row 0) 15 (Row 1)	1024 1024
R5	H0/H1	64	6	9	16 (Row 0) 17 (Row 1)	1156 1156

Table 3.2: End-cap module configuration for rings R0–R5, showing the hybrid type, number of modules in ϕ, number of disks, ABCStars chips per module, η-module indexing, and number of strips per η row. All values are taken from [107].

types is shown in Figure 3.23.

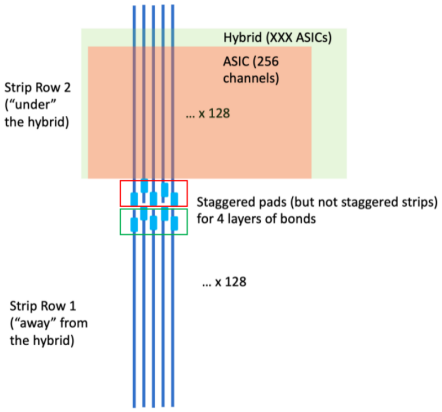


Figure 3.23: Distribution of strips within the ABCStar chips. Near strips are shown in red and far strips in blue. These are defined according to the η index associated with each stave and petal. Figure taken from [107].

3.6.1.2 *Overview of read-out chips*

When a charged particle passes through the sensor, it generates a signal charge within the silicon sensor diode. This signal is transmitted through a wire to the front-end chip, the **ATLAS Binary Chip (ABCStar)** [108], which is designed using 130 nm technology. The main read-out blocks are shown in Figure 3.24 (blue). The **Top Logic** block controls the data path, managing incoming requests such as **priority (PR)** or **low-priority (LP)**.

The **frond-end** which is optimized for 2.5 cm strips, with a shaping time of 25 ns and noise level below 1000 electrons after the complete radiation effects. This part ensures signal discrimination at each **bunch-crossing (BC)**, and the resulting binary output are stored in the **L0Buffer** for a duration determined by the programmed latency for receiving the L0 signal.

The event corresponding to the latency is assigned to the **L0tag** (A 7-bit value attached to each Level-0 Accept). The binary data is then processed by the **Cluster Finder (CF)**.

The CF creates cluster bytes for the channels that register hits. Finally the **Readout** block generates formatted packets containing the event identification and the associated cluster bytes. The data is transmitted serially at 160 Mb/s.

The ABCStar output is transferred to the **Hybrid Controller Chip (HCCStar)** [109], which is also designed using 130 nm technology. The main read-out blocks of the HCC are shown in Figure 3.24 (orange).

The **control path** receive the clock and the signal from the stave, decodes them, perform any required processing, and forwards the resulting command stream to the ABCStar read-out chips via the hybrid bus.

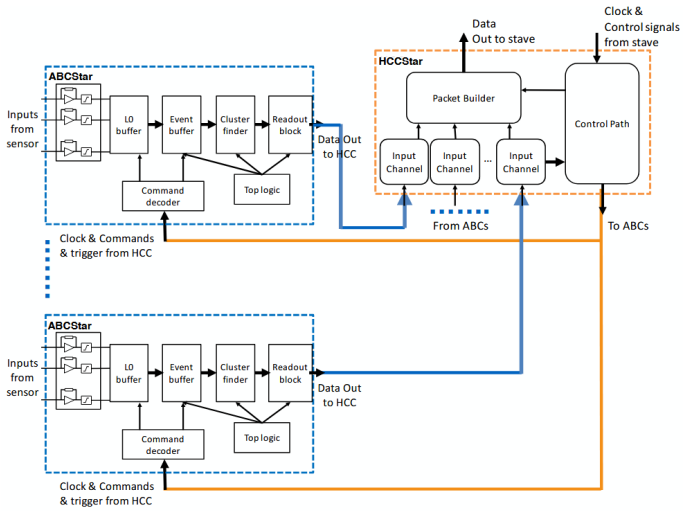


Figure 3.24: Simplistic schematic of an hybrid for the ITk strip modules. Figure taken from [22]

The data returning from the ABCStars is then read from the HCCStar's **input channels**, with one dedicated input channel per ABCStar chip (see Figure 3.25). Each channel deserializes the data coming from the ABCStar, performs some validity checks on that data and separates it based on its type (LP or PR).

Finally, the **packet builder** collects the data from various input channels and combines them into event fragment, and serializes the output packet, which is sent out across the bus tape to the End-of-Substructure (EOS).

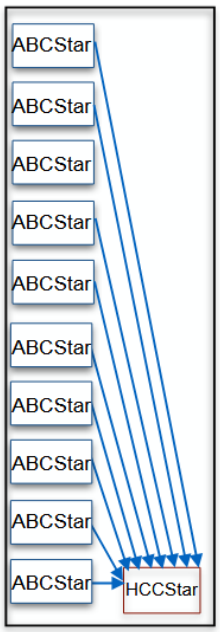


Figure 3.25: New star design for the ITk Strip Detector. Figure taken from [22].

CHAPTER 4

Inner Detector Tracker (ITk) Strip Byte-Stream package

The purpose of this chapter is to present the development of a new byte stream (BS) converter package for the simulation of the new Inner Detector Tracker (ITk), focusing on the strip component. To this end, the chapter is structured to first explain what a BS is (Section 4.1), then describe how a simulated RDO file can be transformed into a BS using the Encoding Algorithm (Section 4.3), and finally show how the BS can be decoded to recover the original RDO file using the Decoding Algorithm (Section 4.5).

4.1 What is a Byte-stream converter?

A **bytestream (BS)** in ATLAS corresponds to the raw binary data format that contains the information produced by the detector readout electronics, as introduced in Subsection 3.4.3. It is designed as a sequence of **bytes**, where each byte consists of eight bits (`0` or `1`). The BS stores the information coming from the different subdetectors, together with event-level data such as timing, bunch crossing, and trigger information.

The process of creating the BS is referred to as **Encoding**. During this stage, the information generated by the subdetector electronics including hit data, timing, trigger signals, and bunch-crossing identifiers, is collected and formatted into the binary BS structure. In the context of this work, the main objective is the development of a BS for the simulation of the new ATLAS Inner Detector, focusing on the Strip subsystem introduced in Section 3.6. In this system, the readout electronics are composed of modules containing a variable number of ABCStar chips, whose output data are transmitted to the HCCStar. An overview of the readout architecture of these chips was presented in Subsection 3.6.1.

To fully exploit the information stored in the BS, a decoding step is required. **Decoding** consists in interpreting the binary content of the BS and translating it into readable data structures. This step is essential for any subsequent reconstruction or analysis, as it allows the physical information originally recorded by the detector electronics to be accessed and processed. A schematic representation of the relationship between encoding and decoding, and their connection to the

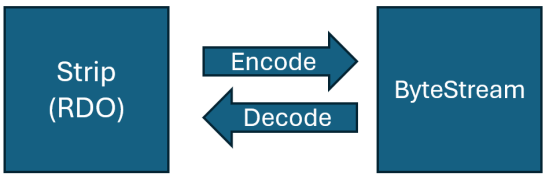


Figure 4.1: Schematic representation of the encoding and decoding processes between the Strip raw data objects (RDO) and the ATLAS bytestream format.

Strip **raw data objects (RDO)**, is shown in Figure 4.1.

4.2 Encode Packages

The first implementation of the ITk Strip BS encoding was developed within the merge request !76971 . This implementation is structured into three main classes, each responsible for a specific stage of the encoding chain.

The **ITkStripsRawContByteStreamCnv** acts as the component that takes the data stored in StoreGate and initiates its conversion into BS format. This class does not interpret the detector data itself. Instead, it serves as the entry point of the encoding process, triggering the conversion from reconstructed RDOs to the ByteStream representation within the Athena framework.

The **ITkStripsRawContByteStreamTool** implements the logic of the encoder. Its role is to organize the RDOs according to the detector topology and cabling, grouping them by ROD and preparing the corresponding input structures required for the encoding algorithm.

Finally, the **ITkStripsRodEncoder** performs the actual construction of the BS format. This class receives the grouped RDOs and translates them into binary words.

In addition to the C++ implementation, a dedicated python based continuous integration (CI) test, **testITkEncodeNewCfg.py**, was developed to validate the encoding configuration and ensure the correct functionality of the BS decoding within the Athena framework.

4.3 Encode Algorithm

4.3.1 Organization of the Hits

As described in Section 3.6, each chip contains 256 strips. There are two types of strips in each chip, *near* and *far*, which occupy the same position in the module but have different η values. Each strip is associated with an ABC logical channel

that defines its position in the logical channel map, described in Appendix A. The organization of hits is designed to build a 256-bit bitset, where each bit represents the state of one strip in the chip. This bitset is then used as input for the Cluster Finder algorithm. To construct it, we begin with the input RDO, group the hits according to their geometries, filter them by η group, identify the corresponding module and chip, and finally map them into the bitset using their logical strip positions (see the overview in Figure 4.2).

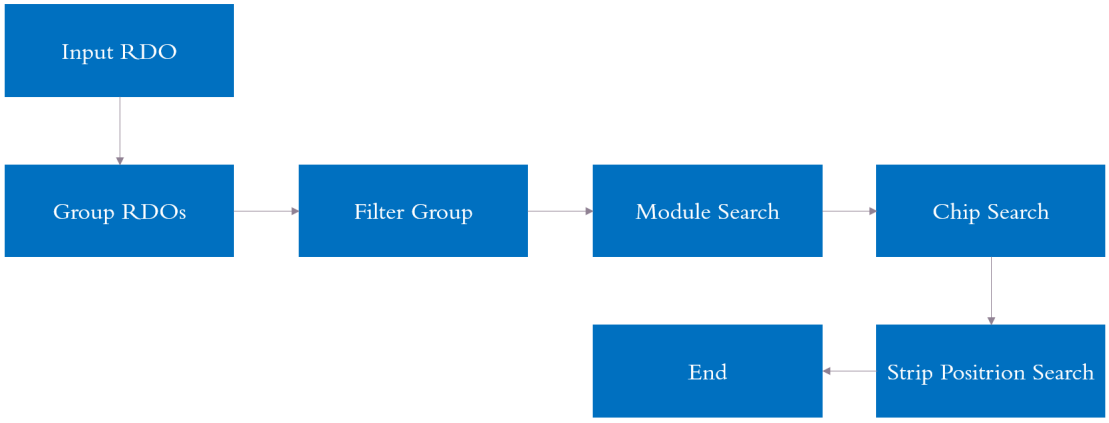


Figure 4.2: Overview of organization of 256-bit hits

4.3.1.1 *Group the RDOs*

In the petals there are nine modules per side, organized in six rings (R0-R5), according to Figure 3.22, while in the staves there are two types of geometries the short-strip (SS) modules and the long-strip modules (SS) each with 28 modules (see Subsection 3.6.1). The most straightforward way to group hits with their corresponding strips is to first determine the geometry of the strip. From this, one can identify whether the hit originates from the barrel or the endcap. Once this is established, it is possible to determine the layer (for the barrel) or the disk (for the endcap). In addition, the positions ϕ and η within the stave or petal can be retrieved. All the necessary functions to obtain these values can be found in the following file SCT_ID.h.

4.3.1.2 *Filter by Groups*

Once all hits have been grouped according to their geometry, they can be further filtered into η groups. As illustrated in Figure 3.23, the *near* and *far* strips share the same geometry. The only distinction between these two types of strips is their index η, which allows them to be combined into the same map.

In the petals, the η_{mod} values are always nonnegative and start from zero.

Therefore, the grouping can be obtained by dividing η_{mod} by two and taking the smallest integer part of the result.

On the other hand, in the staves, the values of η_{mod} are not always positive and the index does not include zero. As a result, the grouping starts from ± 1 instead of 0. To ensure consistent grouping, we add $+1$ to η_{mod}, divide by two, and take the value of the smallest integer.

With this procedure, all near and far strips can be merged into a single map, as shown in Table 4.1.

Case	η_{mod} (examples)	η_{group}
Petals	{0,1} | {2,3} | {4,5}	0 | 1 | 2
Staves	{-2,-1} | {1,2} | {3,4}	-1 | 1 | 2

Table 4.1: Example of η grouping for petals and staves.

4.3.1.3 Module Search

The staves in each module contain the same number of chips, while the petals, which have seven different geometries, contain different number of chips per module. To ensure that hits are correctly grouped, we use a map constructed from the geometry and the η grouping, called `allStripData`. From this map, we build an object called `StripData` (see Figure 4.3), which is a vector whose size is equal to the maximum chip index given by Equation 4.1. Each entry of this vector is a 256-bit `bitset`, used to store all strips belonging to that chip.

$$\text{chip}_{max} = \frac{\text{strip}_{max} + 1}{128} \tag{4.1}$$

4.3.1.4 Chip and Logical Strip position search

To determine the chip and logical channel of a given strip, we start with its absolute position within the module. Since each chip contains 256 strips (indexed from 0; 128 far strips + 128 near strips), the chip index can be obtained by dividing the strip position by 128 and taking the floor of the result. This ensures that strips $[0, 127]$ belong to chip 0, strips $[128, 255]$ to chip 1, etc.

$$\text{chip}(s) = \frac{s}{128} \tag{4.2}$$

For example, if $s = 255$, then:

$$\text{chip}(255) = \frac{255}{128} = [1.99] = 1 \tag{4.3}$$

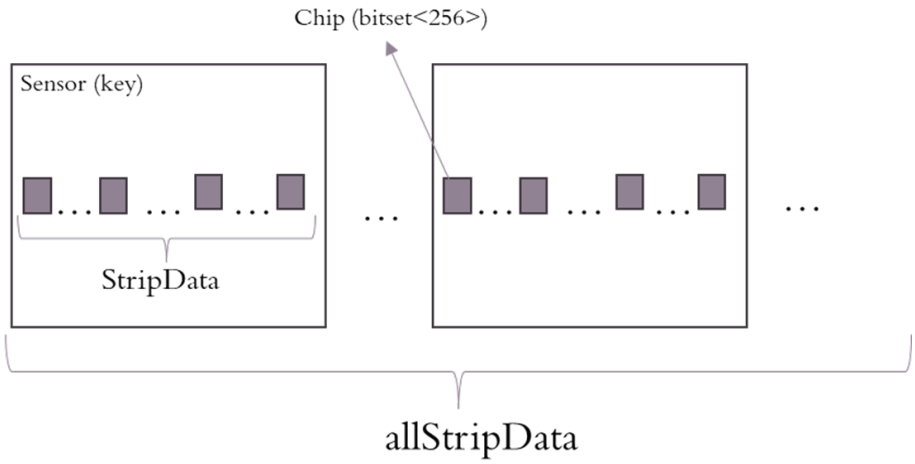


Figure 4.3: Schematic representation of the module search and strip organization into `StripData`.

which means that strip 255 belongs to chip 1.

As the *near* and *far* strips share the same physical position within the chip, they must be grouped according to their logical channel, as defined in Table A.1. In this scheme, the even logical channels correspond to *near* strips, while the odd logical channels correspond to *far* strips, depending on their η_{mod} value.

Following this mapping, all hits can be consistently assigned to their corresponding chip and logical channel. This guarantees a coherent organization of the detector readout and allows the correct association of *near* (even) and *far* (odd) strips within the same chip, as illustrated in Figure 4.4.

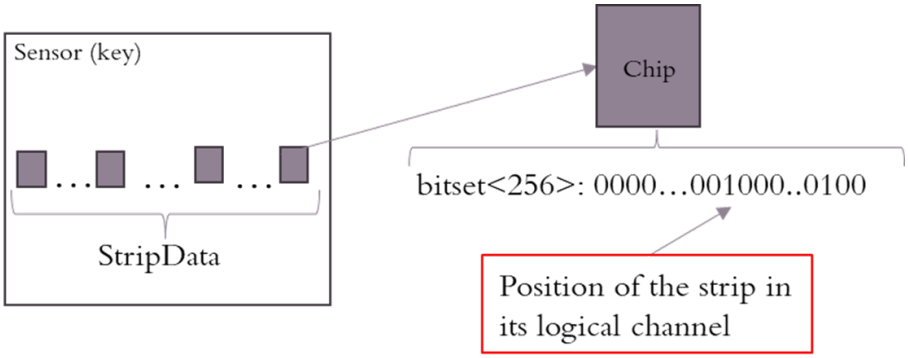


Figure 4.4: Schematic representation of the logical position of strips inside a chip, distinguishing between *near* (even) and *far* (odd) channels.

4.3.2 Cluster Finder

The Cluster Finder algorithm is designed to take as input a 256-bit strip data representation and to report 12-bit clusters. Each cluster consists of 8 bits encoding

the hit address (as defined in Table A.1), 3 bits indicating the values of the following three strips, and 1 bit specifying whether it is the last cluster or not (see Table 4.2). To achieve this, the 256-bit input is divided into two 128-bit bit-sets, denoted as *Odd* and *Even* (Figures 4.4 and 4.5), corresponding to the *far* and *near* strips, respectively. Each of these 128-bit bit-sets is then further split into two 64-bit rows, which are subsequently processed by the algorithm [108].

$$\text{dataOdd}[i] = \text{dataIn}[2i+1], \quad i = 0, 1, 2, \ldots, 127 \tag{4.4}$$

$$\text{dataEven}[i] = \text{dataIn}[2i], \quad i = 0, 1, 2, \ldots, 127 \tag{4.5}$$

The algorithm processes each of these 64-bit rows by scanning for the first hit. Once a hit is identified, it checks the following three strips to determine whether they also contain hits. These bits are then cleared, and the procedure continues iteratively until all clusters are found. For each cluster, a 12-bit word is constructed, including the hit address, the three subsequent strips, and a final flag bit marking the last cluster.

If **no hits/clusters** are found, the Cluster Finder returns a reserved 12-bit word corresponding to a "no cluster byte" value, defined as `0x3FE`.

8 bits Adress of the Hit (8 bits)	Following 3 strip values (3 bits)	1 bit (Last Cluster) (1 bit)

Table 4.2: 12-bit Cluster data

4.3.3 Physics Packet

The ABCStar chips use a 68-bit packet format. Each packet always starts with three start bits (`111` in binary) and ends with a stop bit of value `0`. There are three possible packet types sent by the ABCStar: Low Priority / Priority (LP/PR) packets, High Priority (HPR) packets, and register read packets. Since this work is based on simulation, we only consider the physical LP/PR packets, specifically those of type Priority.

The LP/PR packet is structured into 64-bits, as summarized in Table 4.3. The last four bits indicate whether the packet is of priority (`0001`) or low priority (`0010`). This is followed by a status flag bit (always set to `1`), a 7-bit L0Tag event identifier[1], a 3-bit BCID [2], and a 1-bit parity check for error detection. The remaining 48 bits are used to encode up to four cluster words of 12 bits each.

[1]The L0Tag is a 7-bit label attached to each Level-0 trigger acceptance. It is used to correlate

Frame bit	LP/PR packet field
63–64	Type$[2:3] = 0$
62	Type$[1] = 0$ (PR) or 1 (LP)
61	Type$[0] = 1$ (PR) or 0 (LP)
60	FlagBit (1 bit, always = 1)
53–59	L0Tag $[0:6]$
50–52	BCID $[0:2]$
49	Parity check (xor of BCID[7:0])
48	Cluster1_lastcluster
37–47	Cluster1 $[0:10]$
1–36	Same format for Clusters 2, 3, 4

Table 4.3: Structure of the LP/PR packet format used by the ABCStar chip.

4.3.3.1 Clusters in Physics Packet

The physics packet is constructed from a header followed by 4 clusters. The header itself is formed from the final 16 bits of the physics packet.

For events with fewer than 4 clusters, the "empty cluster byte", defined as `0x7FF`, is used.

For events with more than 4 clusters, the first packet is formed with the framing and the Physics Header (16 bits), followed by 4 cluster bytes. Additional packets are then formed, each containing the framing, the same Physics Header (16 bits), and either 4 cluster bytes or empty cluster bytes, until all clusters for the event are transmitted. A "LAST CLUSTER" bit set to one indicates the final cluster of the event.

4.3.3.2 Physics Packet Sending

The physics data are written to the byte-stream in byte format. For events with up to four clusters, each physics packet occupies 10 bytes. The packet header uses 2 bytes (16 bits), and each cluster also requires 2 bytes (16 bits). Since each cluster effectively uses only 12 bits, 4 bits per cluster remain unused, leaving a total of 16 unused bits when four clusters are present.

To make efficient use of these unused bits, the chip ID is included in each cluster, replacing the former "last cluster bit" The resulting structure of the cluster information in the byte stream is shown in Table 4.4.

An example of a BS file is shown in Figure 4.5. In the first case (green), the

the detector readout with the correct event and to avoid mixing data from different triggers.

[2]The Bunch Counter ID (BCID) is derived from the LHC bunch crossing clock. In hardware it is an 8-bit counter, but in this packet only the 3 least significant bits are transmitted. It is mainly used for synchronization between front-end and trigger signals, ensuring correct time-stamping of hits.

Start bit	4-bit Chip ID	8-bit Cluster Address	3-bit Adjacent Strips

Table 4.4: 16-bit cluster structure in the BS.

header is `0x1800`, followed by four clusters: `0x3CC8`, `0x3838`, `0x3F28`, and `0x393A`. Converting the header to binary (`0001 1000 0000 0000`) reveals that it is a PR packet (first four bits), with the flag bit set, and both the BCID and L0Tag equal to zero.

Now, consider the first cluster `0x3CC8`, which corresponds to the binary representation `0011 1100 1100 1000`.

- The first bit (`0`) marks the start.
- The next 4 bits (`0111`) indicate the chip, which is 7.
- The following 8 bits give the cluster address (`0x99`), corresponding to the far 25 strip and the logical channel 51, according to Table A.1.
- Finally, the last 3 bits (`000`) indicate that there are no additional hits in the following strips.

In contrast, the second case (red) also starts with the same header `0x1800`, but the first cluster is now `0x4BFE`, which expands to the binary `0100 1011 1111 1110`.

- The chip ID is 9.
- The cluster data `0x3FE` indicates a "no cluster" value, meaning there are no hits/clusters for this chip.
- The following three cluster entries are also empty, represented by `0x7FF`.

4.3.4 HCCStar

Now, all the data from the ABCStar chips within a hybrid are collected by a single HCCStar, as illustrated in Figure 3.25. Depending on the detector region, up to two HCCStars can be connected to the same sensor, as summarized in Table 3.2. For this reason, it is necessary to include explicit information identifying the HCC chip in the data stream. This information is introduced as a dedicated header added to every module readout.

The proposed **HCC identification scheme (HCCID)** is described in [110] and is based on a 24-bit word, divided into six nibbles, encoding the detector geometry and the HCC index. The most significant bit (MSB) corresponds to the

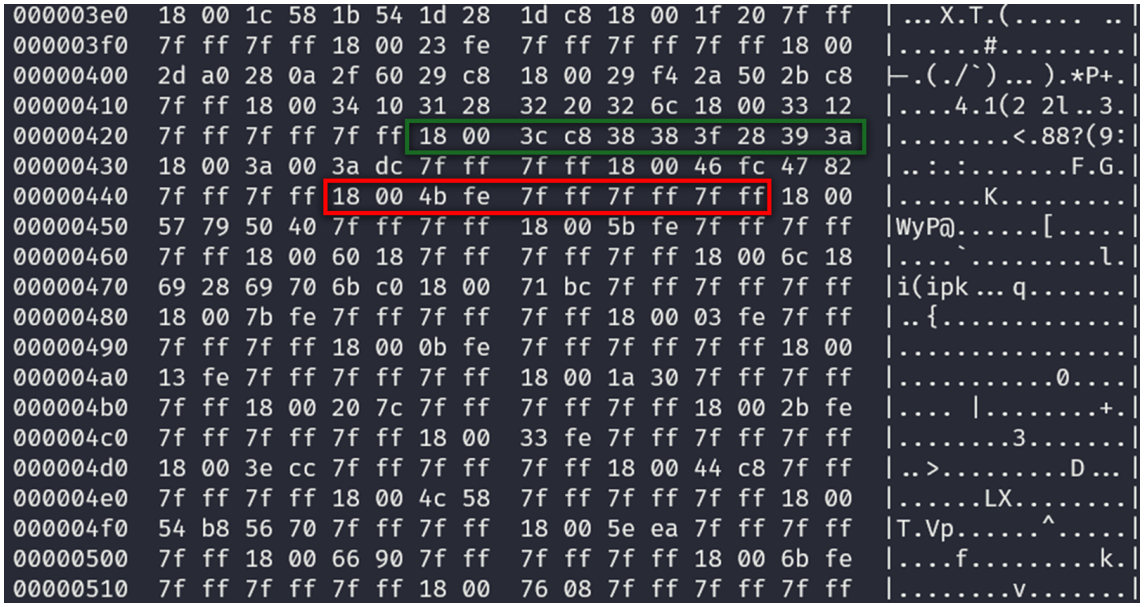


Figure 4.5: Example of a byte-stream file visualized with `hexdump`. Green highlights show valid clusters, while red highlights indicate empty and no clusters entries.

detector region, distinguishing between barrel and endcap. The A/C side of the detector is encoded in bit 20, while bits 17–19 define the layer number in the barrel or the disk number in the endcap. Bit 16 identifies whether the module is located on the inner or outer side.

Additional geometrical information is included through the azimuthal coordinates, bit 15 encodes the ϕ position within a petal (set to zero for the barrel), and bits 8–14 correspond to the global ϕ index. Bit 4 is reserved to distinguish between multiple HCC chips connected to the same sensor. Finally, the least significant bits (bits 0–3) encode the longitudinal position, corresponding to the z index in the barrel or the ring number in the endcap. Table 4.5 shows the distribution of the bits and its descriptions.

4.4 Decode Packages

The first implementation of the ITk Strip BS decoding was developed within the merge request !81433. This implementation is structured into three main classes, each responsible for different part in the decoding chain.

The **ITkStripRawDataProvider** is responsible for coordinating the decoding process. Its main role is to retrieve the relevant ROD fragments from the event store, making use of the **ITkStripCabling** service to correctly map online identifiers to offline identifiers and the **ROBDataProvider** service to access the RAW

Bit(s)	Description
23	Barrel or Endcap.
20	A/C side of the detector.
17–19	Layer number for the barrel or disk number for the endcap.
16	Inner or outer side of the detector.
15	ϕ index within a petal (always 0 for the barrel).
8–14	Global azimuthal (ϕ) index.
4	Identifier for multiple HCCStars connected to the same sensor.
0–3	z index in the barrel or ring number in the endcap.

Table 4.5: Description of the 24-bit HCC identification (HCCID) encoding scheme. Values taken from [110]

BS data.

The **ITkStripRawDataProviderTool** contains the core decoding logic. This tool iterates over the retrieved ROB fragments and manages the decoding workflow by delegating the interpretation of the binary payload to the decoder class.

Finally, the **ITkStripRodDecoder** is responsible for decoding of the BS. It parses the binary data contained in the ROD fragments, interprets the information according to the ITk Strip readout format, and produces the corresponding RDOs.

In addition to the C++ implementation, a dedicated python based continuous integration (CI) test, **testITkDecodeNewCfg.py**, was developed to validate the decoding configuration and ensure the correct functionality of the BS decoding within the Athena framework.

4.5 Decode Algorithm

The decoding algorithm is designed to interpret the content of the BS data and reconstruct the original detector information by performing the inverse operation of the encoding. The input data are read from the ROD fragment as a sequence of 32-bit words. However, since the ITk Strip online format is defined in terms of logical 16-bit words, the raw buffer is reinterpreted as a continuous sequence of bytes (8-bit).

Once the BS has been converted to a byte sequence, the decoding proceeds by looping sequentially over the data. During this loop, the algorithm continuously examines the binary patterns of the data words in order to identify the different structures defined in Section 4.3, such as headers, flags, package type, and cluster

data. Based on these patterns, the decoder transitions between different logical states corresponding to the various components of the decode chain. A schematic representation of the decoding workflow is shown in Figure 4.6.

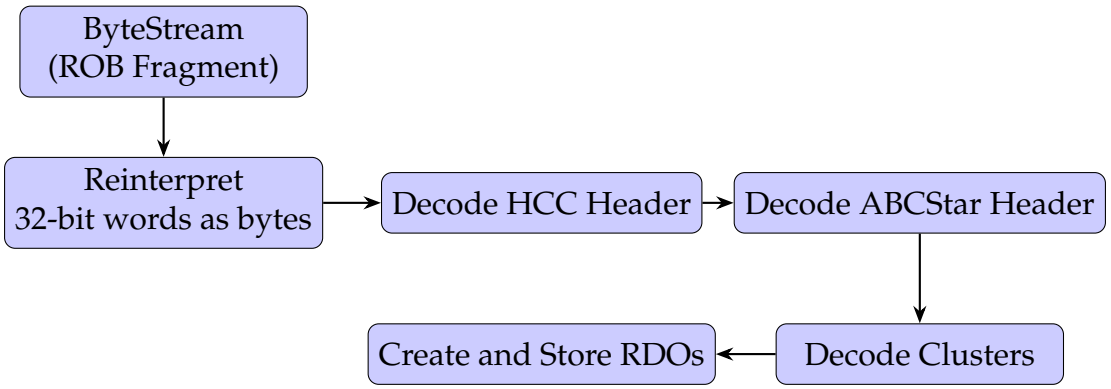


Figure 4.6: Schematic workflow of the ITk Strip ByteStream decoding algorithm.

4.5.1 Process HCC Header

The first structure decoded by the algorithm is the HCC header. In order to identify the presence of an HCC header, the decoder uses a control flag, `isHCCHeader`. This flag is set based on the most significant bits of the data words and on the current decoding state, reflecting the fact that the HCC header always appears at the beginning of a packet.

When the algorithm detects that the conditions for an HCC header are satisfied, it enters the HCC header decoding function, `processHccHeader`. Inside this function, the different geometrical bits values listed in Table 4.5 are extracted. This is achieved through explicit bit shifting and masking operations, which give us the specific bit fields corresponding to detector properties such as barrel or endcap region, side (A/C), layer or disk, petal, etc.

4.5.2 Process ABCStar Header

After the HCC header has been decoded, the algorithm proceeds to search for the ABCStar packet header. This header is identified by detecting the `PR header`, which marks the beginning of a packet associated with an individual ABCStar chip.

When this condition is satisfied, the decoder enters the ABCStar header processing function, `processHeader`. In this function, the information such as the L0 tag and the BCID are extracted from the header word.

If a PR header is not found where expected, the algorithm issues a warning message. This is implemented to prevent unexpected data packets.

4.5.3 Looking for Clusters

Once the ABCStar header has been identified and decoded, the algorithm begins searching for the cluster data. For this, each subsequent data word is passed to the `processHits` function, which is responsible for decoding the cluster information.

This function interprets the cluster address, determines the initial strip index, and identifies whether up to three consecutive strips are hit, according to Table 4.2. In addition, the algorithm track of the ABCStar chip identifier associated with each cluster.

The decoder continues processing cluster words until it encounters the special marker `0xed6f`, which indicates that no further clusters are present in the current packet. At this point, the algorithm considers the packet to be complete and prepares to finalize the decoding.

4.5.4 Make RDO

Upon exiting the main decoding loop, the algorithm performs a final consistency check to ensure that any partially accumulated cluster information is properly finalized. If such a pending cluster exists, it is converted into a Raw Data Object (RDO) through a final call to the `makeRDO` function.

Finally, all non empty RDO collections are written to the output container.

4.5.5 Validation of the Decoding Algorithm

An example of the decoded output, validated compared with the values used in the encoding is shown in Figure 4.7. The figure presents the output of the decoding algorithm, including the decoded HCC header information and the reconstructed cluster data.

The decoded values are then compared with the corresponding clusters and header used as input during the encoding, that is, before the data were converted into the ByteStream format. This comparison allows a direct verification that the decoding procedure correctly recovers the original information assigned during the encoding process.

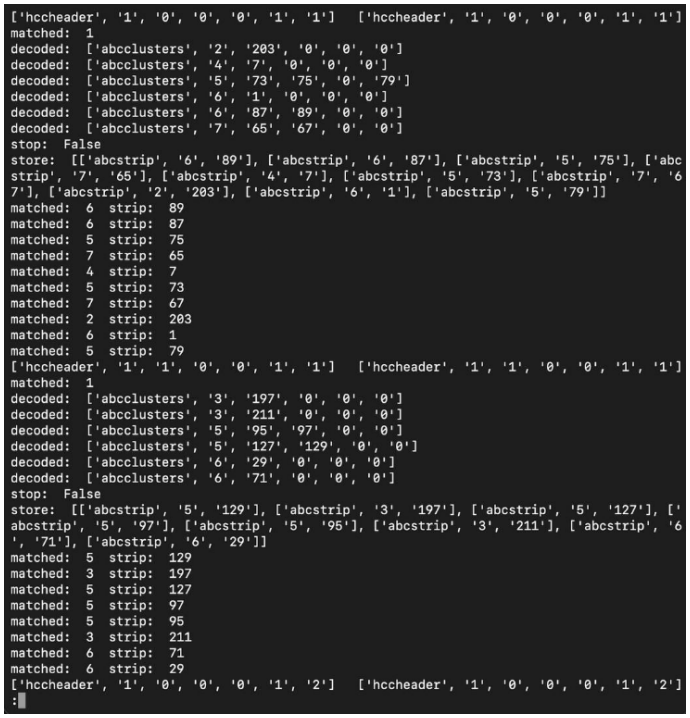


Figure 4.7: Example of the decoding validation, showing the reconstructed HCC header and cluster information compared with the values used in the encoding stage.

CHAPTER 5

Implementation of Boosted di-tau Trigger Monitoring Algorithm

The purpose of this chapter is to present the development of a new monitoring algorithm for the boosted di-τ trigger. Section 5.1 provides a brief overview of trigger monitoring and explains its importance within the ATLAS trigger framework. Section 5.2 describes the data samples and trigger chains used to test the algorithm. Section 5.3 explains how the algorithm is implemented within the ATLAS monitoring framework. Finally, Section 5.4 presents key variable distributions that are relevant for the monitoring of this new trigger.

5.1 Trigger Monitoring

The ATLAS trigger must operate continuously and efficiently during data-taking, as any failure, slowdown, or misconfiguration can lead to a loss of physics data. To address this, the ATLAS trigger monitoring system [111] is designed to detect problems promptly and allow them to be corrected as quickly as possible. Trigger monitoring is therefore divided into two complementary components: **online monitoring** and **offline monitoring**.

Online monitoring is performed during data-taking and is used to monitor the trigger system in real time. Provides visualizations of trigger performance with respect to the active trigger menu and the algorithms running online. This allows shifters and experts to monitor the system both locally and remotely.

Several tools are available for online visualization. The *Data Quality Monitoring Display* (DQMD) presents data quality results and helps shifters identify problematic histograms, which may indicate issues in the trigger system or in detector subsystems. The *Online Histogram Presenter* (OHP) is an interactive and highly configurable tool that displays distributions for different trigger signatures, such as jets, taus, electrons, and muons. These distributions are used to monitor physics-related variables and to compare the observed behavior with expectations. Finally, the *Trigger Rate Presenter* (TRP) is a specialized tool focused on trigger rates, providing information on trigger rates, rate predictions, and the status of the High-Level Trigger (HLT) computing farm.

Offline monitoring starts shortly after the end of a run and focuses on data

quality assessment and validation. The offline data quality (DQ) evaluation is organized according to trigger signatures. For each trigger signature, runs must be signed off before they can be used in physics analyses. This sign-off is usually performed using the so-called *express stream*, which is rapidly processed and becomes available a few hours after the run has finished. The express stream contains approximately 10 Hz of recorded events from the run.

Events for which the trigger could not reach a decision are recorded in the *debug stream*. The presence of an event in the debug stream usually indicates a problem, but not necessarily a software bug. Typical causes include very complex events with a large number of Regions of Interest (RoIs) or high track multiplicities, which can lead to trigger algorithms exceeding their time limits.

In the case of boosted di-τ triggers, the complex event topology and the use of dedicated reconstruction algorithms motivate the development of specialized monitoring. Therefore, a dedicated monitoring algorithm is required to ensure the correct behavior of the boosted di-τ trigger, to quickly identify potential inefficiencies or misconfigurations, and to validate its performance under different running conditions.

5.2 Data tested and Trigger Chain

To test the new boosted di-τ monitoring algorithm, a data sample of **pp** collisions collected during Run 3 in 2025 is used, at a center-of-mass energy of $\sqrt{s} = 13.6$ TeV. The specific data sample employed in this study is listed in Appendix B.1.

A single boosted di-τ trigger chain is used to validate the monitoring algorithm. This trigger corresponds to the chain listed in Table 5.1.

Trigger category	**HLT chain**
Ditau	`HLT_j250_a10sd_cssk_ditauOmni1Trk5_pf_jes_ftf_preselj200_L1jJ160`

Table 5.1: Trigger chains used to test the new boosted di-τ trigger monitoring.

5.3 Algorithm implementation

The Trigger Monitoring for Boosted DiTau is implemented within the TrigTauMonitoring package and follows the standard structure of ATLAS monitoring algorithms, but is adapted for boosted di-tau triggers.

The Diagram 5.1 illustrates the workflow, We start from AOD inputs, the algorithm retrieves the relevant HLT chain names and obtain the features of the Trigger Decision Tool (TDT). Using this information, it retrieves the `xAOD::DiTauJet`

objects. Finally, for each objet, the algorithm reads the first Ditau candidate, extracts all relevant variables, and fills histograms in the monitoring group.

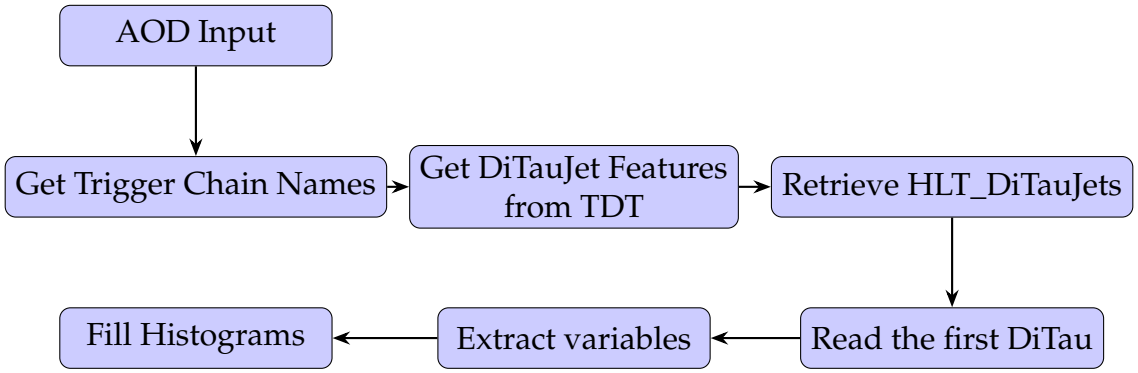


Figure 5.1: Workflow overview of the boosted di-tau monitoring algorithm.

5.4 Results

The key variables added to the new boosted di-τ trigger monitoring are observables used by the **Omni Tagger**, as listed in Table 3.1, also some kinematic and number of tracks are implemented. These variables are included to validate the behaviour of the tagger at trigger level and to provide reference distributions for future performance studies.

Figure 5.2 shows a set of general variables employed in the training of the Omni Tagger. These include the Omni score, the track-based radius R_{tracks} for the leading and subleading jets, and the corresponding f_{core} variables. The Omni score reflects the response of the tagger for boosted di-τ candidates, while the track-based observables characterise the internal structure of the reconstructed object.

Figure 5.3 presents examples of variables related to the number of tracks. The distributions shown correspond to the total number of tracks associated with the di-τ candidate, as well as the number of tracks assigned to the leading and subleading τ subjets.

Finally, Figure 5.4 shows a selection of kinematic variables of the reconstructed di-τ object, including its invariant mass, transverse momentum, pseudorapidity, and azimuthal angle.

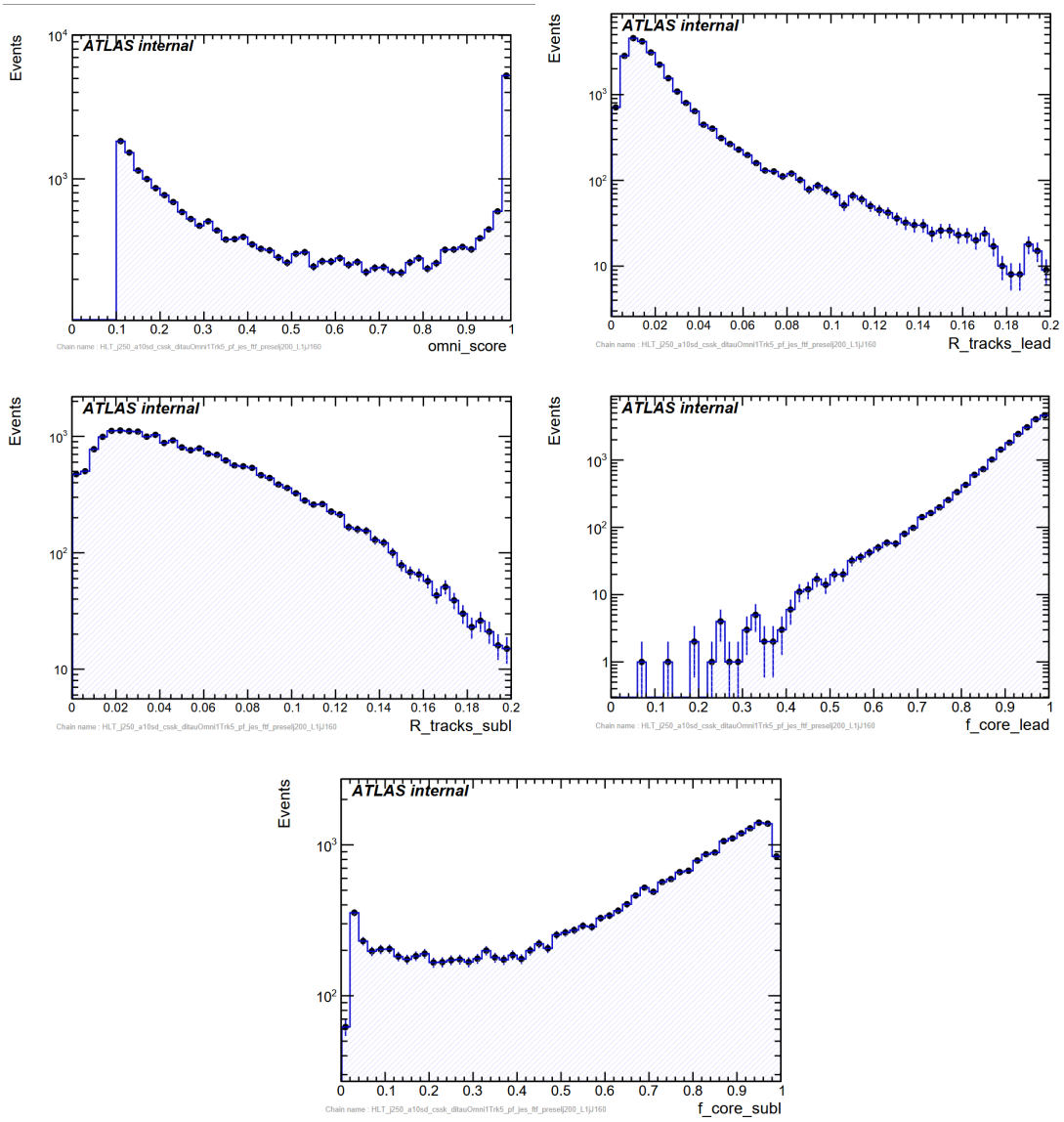


Figure 5.2: Distributions of general variables. The top left shows the Omni score, the top right the leading-track R_{tracks}, the middle left the subleading-track R_{tracks}, the middle right the leading-track f_{core}, and the bottom left the subleading-track f_{core}.

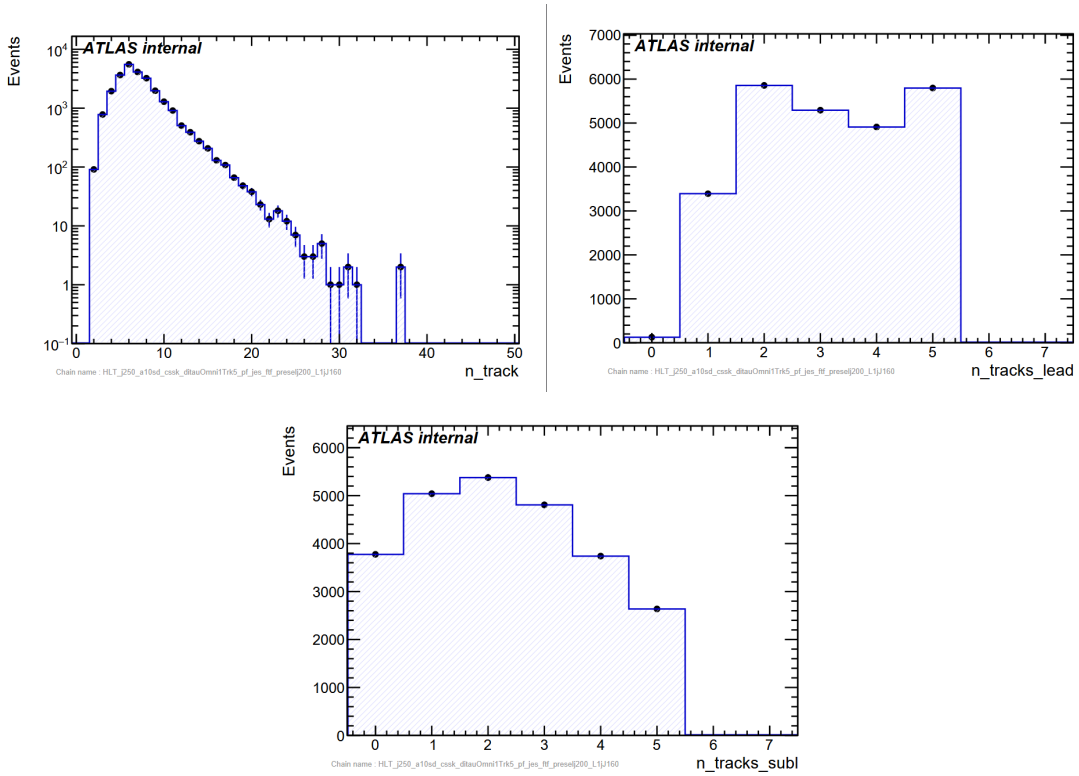


Figure 5.3: Distributions of variables related to the number of tracks. The top left shows the total number of tracks, the top right the number of tracks associated with the leading τ, and the bottom left the number of tracks associated with the subleading τ.

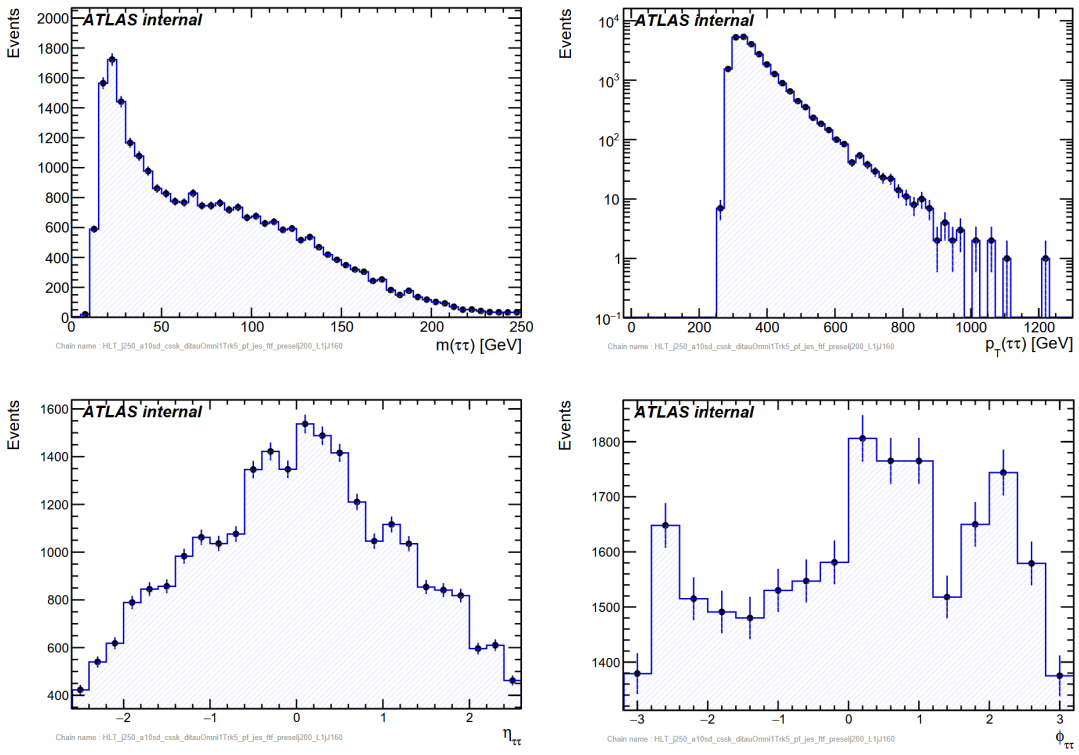


Figure 5.4: Distributions of kinematic variables. the top left shows the di-tau mass, the top right the di-tau transverse momentum, the bottom left the pseudo-rapidity (η), and the bottom right the azimuthal angle (ϕ).

CHAPTER 6

Boosted di-tau trigger performance for $H \to \tau^+\tau^-$ analysis

This chapter focused on a study of the performance of the new boosted di-τ trigger, which is based on the same boosted di-τ identification described in Subsection 3.5.6. The chapter is orginised in three sections. Section 6.1 describes the data samples used in the analysis, together with the trigger chains considered for the boosted the di-τ, large-R jet, small-R jet, single-τ or MET trigger. Section 6.2 oresebts the analysis strategy and defines the signal region selection applied in this study. Finally, Section 6.3 discusses the result obtained and compares the trigger efficiencies for all considered trigger selections.

6.1 Data-Samples and Trigger Chains

The analysis is based on NTuples produced from ATLAS PHYS derivations corresponding to pp collision partial data collected during Run 3 in 2025 at a $\sqrt{s} = 13.6$ TeV. The NTuples were produced using the **xTauFramework**. An event preselection was applied during NTuple production, consisting of an electron and muon veto and the requirement of at least one reconstructed boosted di-τ object candidate. The list of datasets used in this analysis is summarized in the Appendix B.2.

The trigger chains used for the boosted di-τ, large-R jet, small-R jet, single-τ, and E_T^{miss} triggers are listed in Table 6.1.

Trigger category	**HLT chain**
Ditau	`HLT_j250_a10sd_cssk_ditauOmni1Trk5_pf_jes_ftf_preselj200_L1jJ160`
Small-R jet	`HLT_j400_pf_ftf_preselj225_L1jJ180`
Large-R jet	`HLT_j460_a10t_lcw_jes_L1jJ160`
Single τ	`HLT_tau160_mediumGNTau_L1eTAU140`
MET	`HLT_xe65_cell_xe105_nn_L1jXE100`

Table 6.1: Trigger chains used in trigger performance on boosted $H \to \tau^+\tau^-$ analysis.

6.2 Analysis strategy

This analysis is based on a comparison of the performance of the **boosted di-τ trigger** with other trigger selections capable of identifying di-tau objects. These include **large-R jet trigger**, which are currently used in boosted di-tau analyses, as well as **small-R jet**, **single-τ**, and **missing transverse energy (MET)** triggers in the boosted $H \to \tau^+\tau^-$ regime.

A set of selection requirements is applied in order to closely reproduce the signal region defined in a previous analysis [104]. The di-τ object, is require to contain at least 2 subjets. In addition, the Omni tagger score must exceed 0.9994, corresponding to the optimized working point for separating truth di-τ decays from the multijet background. Events containing b-tagged jets are rejected in order to suppress backgrounds from $t\bar{t}$ and $Z/H \to b\bar{b}$ processes. Furthermore, the di-τ candidate is required to be separated from the leading jet by $\Delta R > 0.4$ prevent overlap between the di-tau system and other energetic jets in the event.

Additional requirements are imposed on the subjets associated with the di-τ candidate. The leading and subleading subjets must satisfy transverse momentum thresholds of $p_T^{lead} > 50$ GeV and $p_T^{subl} > 15$ GeV, respectively. An opposite-charge requirements $q_{lead} \cdot q_{subl} = -1$ is applied to be consistent selection $\tau^+\tau^-$. Moreover, the number of associated tracks for each subjet is required to lie in the range $N_{tracks}^{lead/subl} \; \epsilon \; [1, 3]$, since the decay modes of hadronic τ contain 1-prong or 3-prong.

Since hadronic τ decays involve neutrinos, a requirement of $E_T^{miss} > 60$ GeV is imposed. Finally, additional combined selection criteria are applied to ensure global consistency of the reconstructed event. These include a requirement on the visible energy fraction $E_{vis,i}/E_{sum,i} > 0$, and consistency condition on the total number of associated tracks. The selection requirements are summarised in Table 6.2.

For the triggers, an additional requirement of $p_T^{leadsubjet} > 250$ GeV is imposed for the single-τ trigger, ensuring that only boosted di-τ topologies are selected.

6.3 Results

Some key variables relevant for $H \to \tau^+\tau^-$ analysis are presented in this section. All distributions are shown together with a subplot displaying a *residual*, which quantifies the relative difference between boosted di-τ trigger and an alternative

Object	Selection
DiTau	$N_{\text{subj}} \geq 2$ Omni score ≥ 0.9994 Jet b-tagged $= 0$ $\Delta R(\text{Di}\tau, \text{lead jet}) > 0.4$
Subjets	$p_T^{\text{lead}} > 50$ GeV $p_T^{\text{subl}} > 15$ GeV $q_{\text{lead}} \cdot q_{\text{subl}} = -1$ $N_{\text{tracks}}^{\text{lead/subl}} \in [1, 3]$
MET	$E_T^{\text{miss}} > 60$ GeV
Combined	$E_{\text{vis},i} / E_{\text{sum},i} > 0$ $N_{\text{tracks}}^{\text{Di}\tau} + N_{\text{tracks}}^{\text{lead}} + N_{\text{tracks}}^{\text{subl}} = 0$

Table 6.2: Summary of event selection requirements. Table taken from [104]

trigger (large-R, small-R jet, single-τ or MET trigger). This residual is defined as

$$\text{Residuals} = \frac{N_{\text{trigger}}^{\text{Ditau}} - N_{\text{trigger}}^{\text{i}}}{N_{\text{trigger}}^{\text{Ditau}} + N_{\text{trigger}}^{\text{i}}} \tag{6.1}$$

where $N_{trigger}^{DiTau}$ and $N_{trigger}^{i}$ denote the number of events selected by the boosted di-τ trigger and the reference trigger. respectively.

The main kinematic variables of the reconstructed di-τ object, together with the MET, are shown in Figure 6.1. The top left plot presents the transverse momentum $p_T^{\text{Di}\tau}$ of the di-τ object candidate, where a clear dominance of the di-τ trigger is observed in the low-p_T region, gradually decreasing above approximately 500 GeV. The top right plot shows the E_T^{miss} distribution. The bottom-left plot displays the invariant mass of the di-τ system, exhibiting a broad peak between 100 and 200 GeV, with the di-τ trigger dominating over the full spectrum. Finally, the bottom-right plot shows the pseudorapidity η of the di-τ object.

The kinematic properties of the subjets associated with the di-τ object candidate are shown in Figure 6.2. The plot on the left shows the transverse momentum of the leading subjet, while the central plot presents the same variable for the subleading subjet. The right plot shows the angular separation ΔR between the leading and subleading subjets. The peak at small ΔR values, around 0.4–0.6, indicates that the two subjets are highly collimated, as expected for boosted $\tau^+\tau^-$ decays.

Additional discriminating variables are shown in Figure 6.3. The left and central plots display the variables x_1 and x_2, defined as the ratio between the visible energy and the total energy for the leading and subleading subjets, respectively.

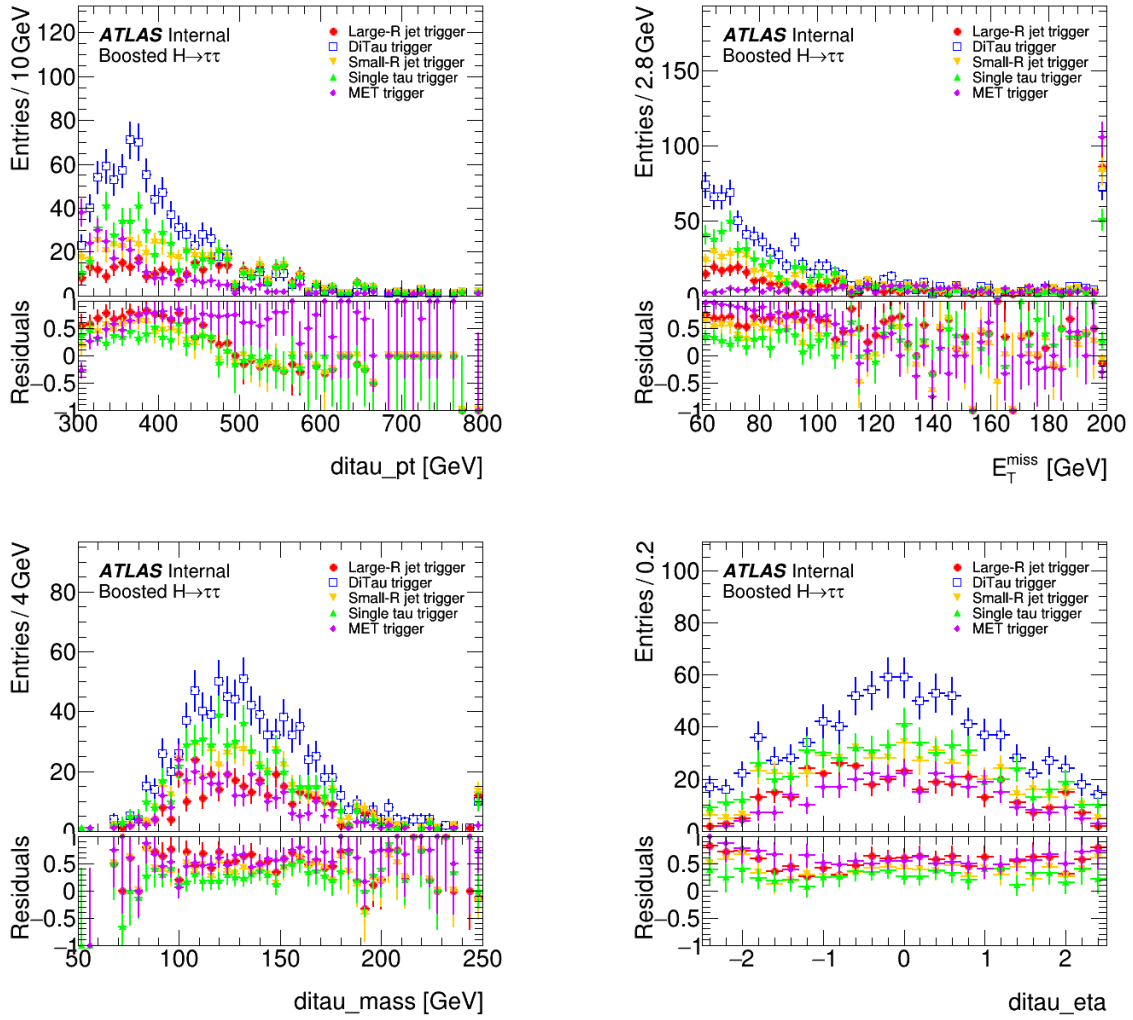


Figure 6.1: Distributions of kinematic variables of the reconstructed di-τ system for events selected by the di-τ, large-R jet, small-R jet, single-τ, and MET triggers. The top-left shows the transverse momentum of the di-τ candidate, $p_T^{\mathrm{Di}\tau}$, the top-right shows the missing transverse energy, E_T^{miss}, the bottom-left displays the visible invariant mass of the di-τ system, $m_{\mathrm{Di}\tau}$, and the bottom-right presents the pseudorapidity of the di-τ candidate, $\eta_{\mathrm{Di}\tau}$.

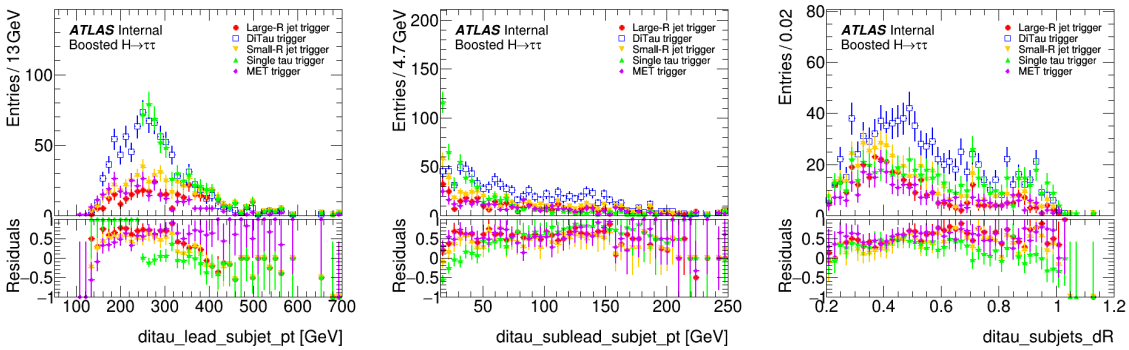


Figure 6.2: Kinematic properties of the subjets associated with the reconstructed di-τ object. The left shows the transverse momentum of the leading subjet, the center shows the transverse momentum of the subleading subjet, and the right displays the angular separation between the two subjets, $\Delta R(\text{subj}_1, \text{subj}_2)$.

These variables are sensitive to the presence of neutrinos in τ decays. The right panel shows the angular separation $\Delta R(\text{lead jet}, \text{Di}\tau)$ between the di-τ object candidate and the leading jet in the event. The distribution peaks between 2 and 4, indicating a large spatial separation of the di-τ system from other hadronic activity and supporting the isolation criteria applied.

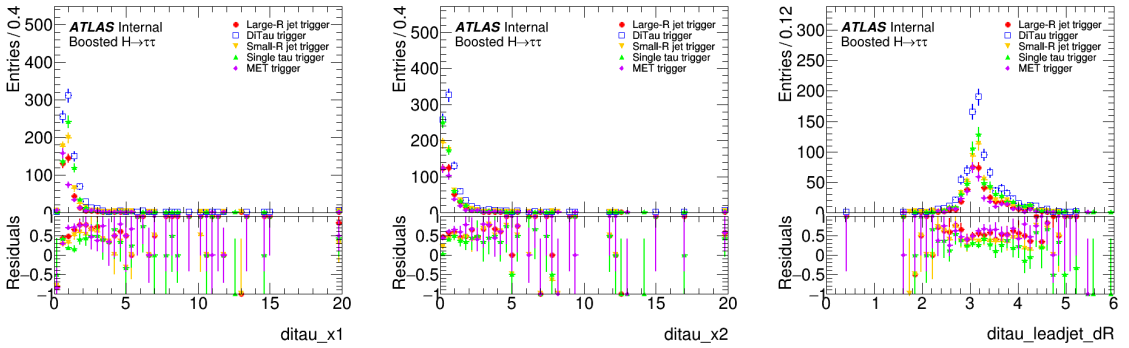


Figure 6.3: Additional variables relevant for the identification of boosted di-τ candidates. The left shows the visible to total energy fraction of the leading subjet, x_1, the center shows the corresponding fraction for the subleading subjet, x_2, and the right presents the angular separation between the di-τ object and the leading jet, $\Delta R(\text{lead jet}, \text{Di}\tau)$.

6.3.1 Efficiency

The trigger efficiency is evaluated using the offline selection described in Table 6.2. The efficiency for a given trigger is defined as

$$\epsilon = \frac{N_{OfflineSelection+trigger}}{N_{OfflineSelection+X}}, \tag{6.2}$$

$$X = \text{DiTau OR SingleTau OR LR-jet OR SR-jet OR MET} \tag{6.3}$$

Here, $N_{\text{OfflineSelection+trigger}}$ denotes the number of events that pass both the offline selection and the trigger under study (boosted di-τ, large-R jet, small-R jet, single-τ, or E_T^{miss} trigger). The denominator $N_{\text{OfflineSelection}+X}$ corresponds to the number of events that satisfy the offline selection and are accepted by at least one of the considered triggers.

This choice of denominator ensures a comparison among triggers by restricting the efficiency measurement to events that are already triggerable by the system. In particular, it avoid events that pass the offline selection but would not be selected by any trigger, such as pure QCD multijet events. Using the offline selection alone as the denominator lead to an incorrect efficiency estimate.

The resulting trigger efficiencies as a function of the di-τ transverse momentum are shown in Figure 6.4. The efficiency of the boosted di-τ trigger is shown in blue, while the efficiencies of the large-R jet, small-R jet, single-τ, and E_T^{miss} triggers are shown in red, green, orange, and purple, respectively. The boosted di-τ trigger clearly dominates in the transverse momentum range between approximately 300 and 500 GeV. At higher p_T values, the efficiency looks like start to losing efficiency. However, this region does not have to much events, as indicated by the di-τ p_T distribution shown in Figure 6.1. Therefore, the observed behavior does not necessarily indicate a efficiency loss of the di-τ trigger performance at high transverse momentum.

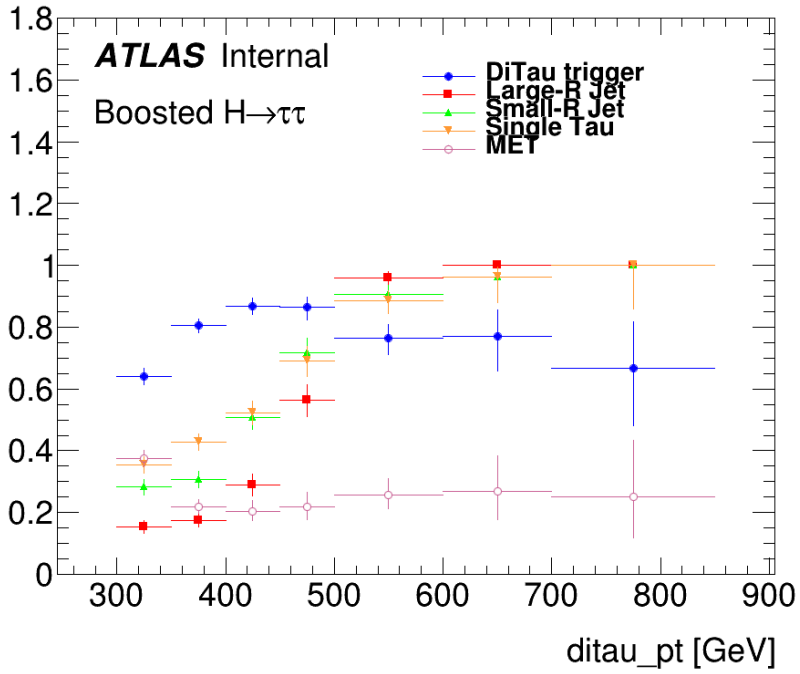


Figure 6.4: Trigger efficiency as a function of the reconstructed di-τ transverse momentum $p_T^{\mathrm{Di}\tau}$, evaluated using the offline selection defined in Table 6.2. The efficiencies of the boosted di-τ, large-R jet, small-R jet, single-τ, and E_T^{miss} triggers are shown.

CHAPTER 7

Comparison of boosted di-τ identification in boosted $HH \to b\bar{b}\tau^+\tau^-$ analysis

This chapter focuses on a comparative study of the GN2XTau and Omni taggers, both described in Subsection 3.5.6, for the identification of boosted di-τ objects in the context of the non-resonant boosted di-Higgs production analysis $HH \to b\bar{b}\tau^+\tau^-$. Section 7.1 describes the MC samples used in this analysis, together with the applied preselection criteria. Section 7.2 outlines the analysis strategy adopted to perform the object selection and truth matching. Finally, Section 7.3 presents the main results of the study, including efficiencies and the distributions of key kinematic variables.

7.1 MC samples and preselection

The analysis uses four Monte Carlo (MC) samples, shown in the Appendix B.3, all corresponding to the production of a Higgs boson pair via vector-boson fusion (VBF), with the non-resonant decay channel $HH \to b\bar{b}\tau^+\tau^-$.

The simulated samples were generated at a center of mass energy of $\sqrt{s} = 13.6$ TeV using **MadGraph5_aMC@NLO**, interfaced with **Pythia 8** for parton showering and hadronization. Generation was performed within the five-flavor scheme (5FS).

Different Higgs boson coupling configurations are considered in these samples. The first configuration corresponds to the Standard Model (SM) values, with $\kappa_\lambda = 1$, $\kappa_V = 1$, and $\kappa_{2V} = 1$. The second configuration sets $\kappa_\lambda = 0$, while the third one sets $\kappa_{2V} = 0$. Finally, a fourth configuration is considered with $\kappa_V = 0.5$.

The EasyJet Framework was used to process the samples. Events were selected using a large-R jet trigger. The preselection requires at least two reconstructed large-R jets in the event. No requirements are imposed on the presence of b-tagged small-R jets, individually reconstructed τ leptons, or isolated charged leptons (e, μ).

The choice of a VBF sample is motivated by its characteristic event topology, which typically features two large-R jet candidates that are well separated in the η–ϕ plane, shown in Figure 2.7. This distinctive topology facilitates the distinc-

Selection type	**DiTau $\to$ Large-R jet**	$b\bar{b} \to$ **Large-R jet**
Score-based	Best Omni score or best GN2XTau score	Highest GN2X score
Topological	Closest to MET: min $(\Delta\phi(\text{DiTau}_i, \text{MET}))$	Highest GN2X score
Kinematic	Highest p_T	Highest GN2X score

Table 7.1: Summary of the reconstructed object selection strategies for the di-τ and $b\bar{b}$ systems in the $HH \to b\bar{b}\tau^+\tau^-$ analysis.

tion and identification of the two large-R jets.

7.2 Analysis strategy

This analysis aims to evaluate the performance of the reconstruction by matching the reconstructed objects with their corresponding truth-level objects.

The di-τ candidate is selected using three different strategies. The first strategy is a **score-based** selection, where the di-τ object is chosen either as the candidate with the highest Omni score and subsequently matched to a large-R jet, or directly as the large-R jet with the highest GN2XTau score. The second strategy is a **topological selection**, where the di-τ candidate is chosen as the object closest in azimuth to the MET, defined by minimizing $\Delta\phi(\text{DiTau}_i, \text{MET})$. This approach is motivated by the presence of neutrinos in τ decays, which leads to MET being preferentially aligned with the di-τ system, while no such correlation is expected for the $b\bar{b}$ system. The third strategy is a **kinematic selection**, where the di-τ candidate is chosen as the object with the highest-p_T. This method is expected to be less optimal, since the presence of neutrinos in the τ decays reduces the visible momentum of the di-τ system, and therefore the $b\bar{b}$ system can have a higher p_T. A summary of the different candidate selection strategies is given in Table 7.1.

After the selection of the di-τ candidate, it is matched to a reconstructed large-R jet to ensure that the di-τ candidate and the large-R jet selected as the $b\bar{b}$ system correspond to different physical objects. The large-R jet associated with the di-τ candidate is defined as

$$\text{LR-Jet}_{\text{DiTau}} = \min(\Delta R(\text{DiTau}_{\text{selected}}, \text{LR-Jet}_i)) \tag{7.1}$$

where LR-Jet_i denotes the i-th reconstructed large-R jet in the event.

Once both reconstructed candidates, the di-τ object and the $b\bar{b}$ object are identified, they are matched to their corresponding truth-level objects. This truth matching is performed by requiring an angular separation of $\Delta R < 0.6$ between

the reconstructed object and the corresponding truth-level object. The di-τ candidate is matched to the truth visible taus, while the truth $b\bar{b}$ system is constructed from truth jets associated with the truth Higgs boson decay to $b\bar{b}$.

7.3 Results

Following the analysis strategy, Figure 7.1 shows the Omni score distributions for the three di-τ candidate selection strategies, together with the GN2XTau score for the large-R jet candidate. For the di-τ identification, values of the Omni score close to unity are expected, while for the GN2XTau discriminator values larger than 5.12 correspond to a working point with approximately 98% background rejection efficiency [105].

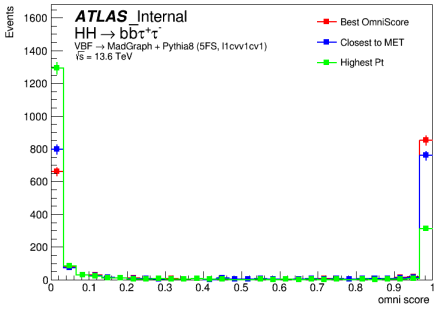
After selecting the di-τ object candidates, the matching to large-R jets is performed using Eq. 7.1. The resulting angular separation is shown in Figure 7.2. For the selection based directly on the GN2XTau score, this matching step is not required, as the di-τ candidate is already identified as a large-R jet. The distribution shows a strong correlation between the reconstructed di-τ candidate and the associated large-R jet, with $\Delta R(\text{DiTau}, \text{LR-jet}) < 0.2$ for all couplings and different di-tau selections.

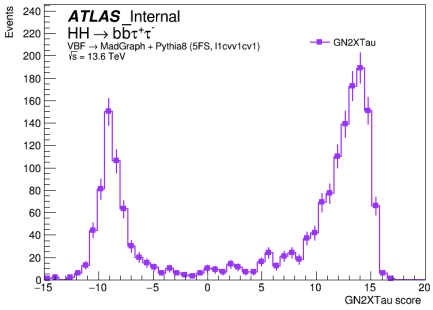
Once the large-R jet associated with the di-τ system is identified. $b\bar{b}$ candidates are selected as the large-R jet with the highest GN2X score among the remaining jets. The GN2X score distributions for the selected $b\bar{b}$ candidates are shown in Figure 7.3.

For the Standard Model and $\kappa_\lambda = 0$ scenarios, the GN2XTau di-τ selection yields the highest GN2X scores for $b\bar{b}$ candidates. In contrast, the highest-p_T selection performs worst, while the selections based on the best Omni score and on the jet closest to the missing transverse energy show very similar behavior.

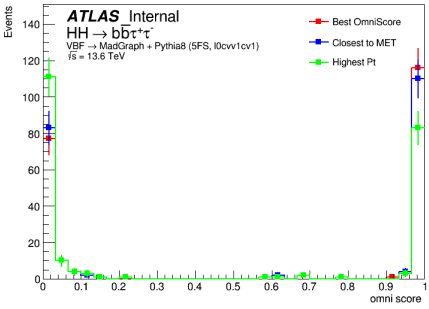
For the $\kappa_{2V} = 0$ and $\kappa_V = 0.5$ scenarios, all di-τ selection strategies exhibit very similar performance.

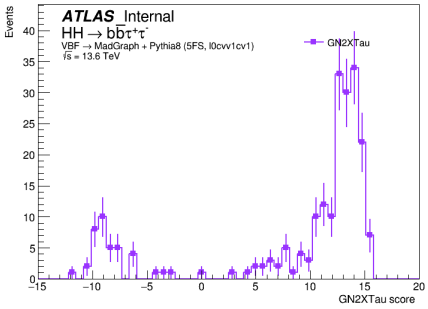
Key kinematic variables of the reconstructed di-τ and $b\bar{b}$ candidates are shown in Figures 7.4 and 7.5, respectively. The di-τ mass and transverse momentum are reconstructed from the two associated subjets. A clear correlation between the reconstructed candidates and the corresponding truth-level objects is observed. The azimuthal separation between the di-τ candidate and the MET, $\Delta\phi(\text{Di}\tau, \text{MET})$, shows similar behavior for most selections, with a peak around ± 1. In contrast, the highest-p_T selection have two peaks around ± 3, indicating a bad association between the di-τ candidate and the MET, as expected. A similar effect is observed in the $\Delta\phi(\text{b}\bar{\text{b}}, \text{MET})$ distributions, where the highest-p_T selection shows a huge



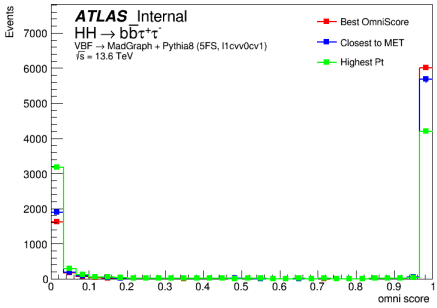


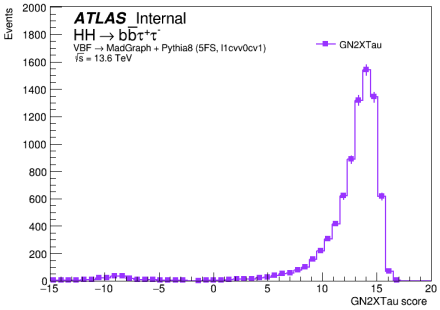
(a) $\kappa_\lambda = 1$, $\kappa_{2V} = 1$, $\kappa_V = 1$



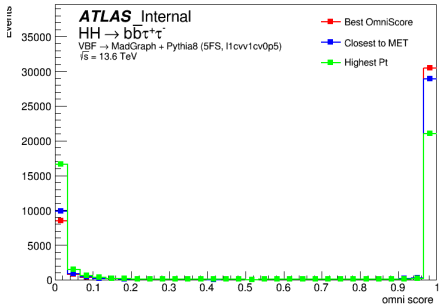


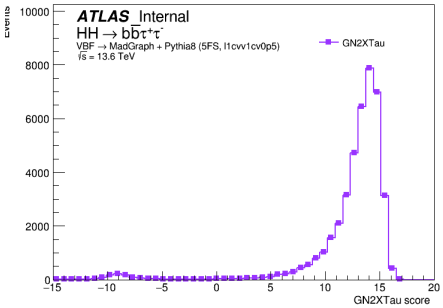
(b) $\kappa_\lambda = 0$, $\kappa_{2V} = 1$, $\kappa_V = 1$





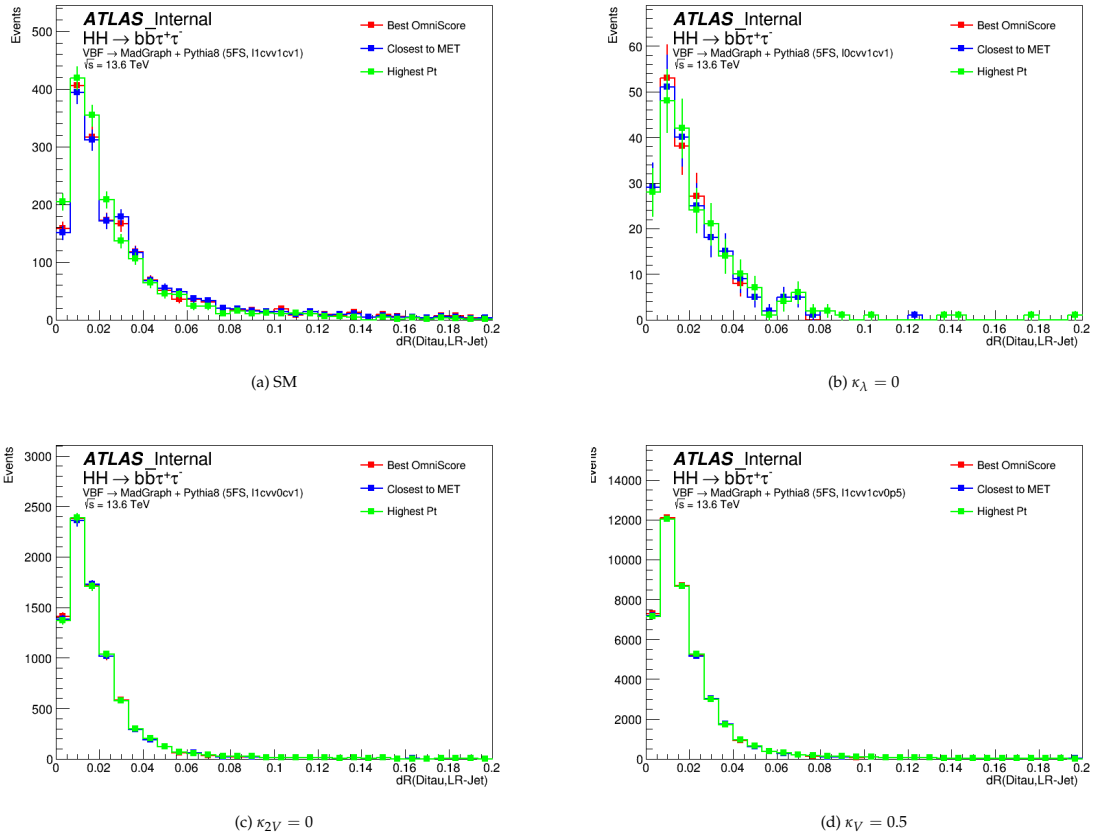
(c) $\kappa_{2V} = 0$, $\kappa_\lambda = 1$, $\kappa_V = 1$





(d) $\kappa_V = 0.5$, $\kappa_\lambda = 1$, $\kappa_{2V} = 1$

Figure 7.1: Distributions of the Omni score for reconstructed di-τ candidates (left column) and the GN2XTau score for large-R jet candidates (right column) for different coupling scenarios. The first row corresponds to the Standard Model configuration, while the lower rows show variations in κ_λ, κ_{2V}, and κ_V, respectively.



(a) SM

(b) $\kappa_\lambda = 0$

(c) $\kappa_{2V} = 0$

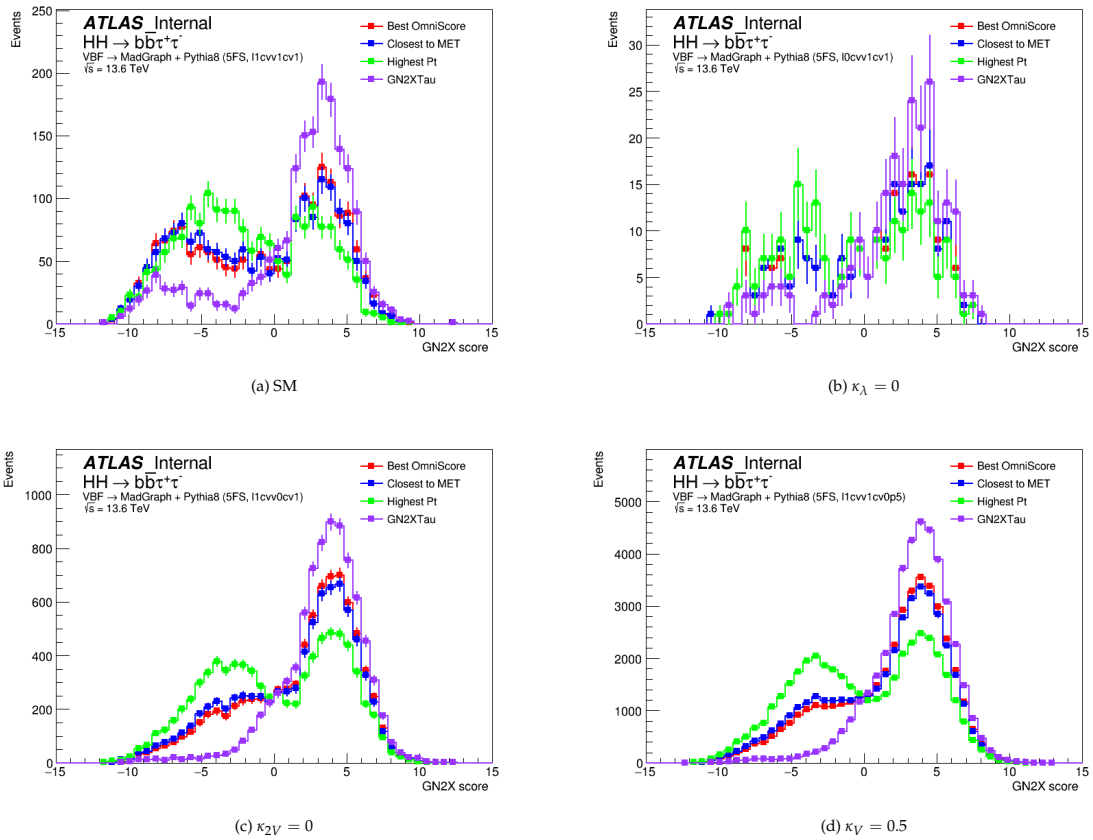
(d) $\kappa_V = 0.5$

Figure 7.2: Distributions of the angular separation ΔR between the reconstructed di-τ candidate and the matched large-R jet for different Higgs boson coupling scenarios. The Standard Model configuration is shown in the top-left panel, while the remaining panels correspond to variations in κ_λ, κ_{2V}, and κ_V.

peak at small angular separations ± 1, suggesting a higher rate of misassignment.

Figure 7.6 shows the Omni and GN2XTau score distributions after applying the truth matching requirement ($\Delta R(\text{DiTau}, \text{Truth}_{\tau\tau}) < 0.6$ and $R(b\bar{b}, \text{Truth}_{b\bar{b}}) < 0.6$). Events passing the truth match populate the high-score regions, with Omni scores above 0.9 and GN2XTau scores above 5, in agreement with the expected.

Finally, the reconstructed di-τ and $b\bar{b}$ masses and transverse momenta after truth matching are shown in Figure 7.7 and Figure 7.8. The highest-p_T selection reduced a greater number of events compared to the other strategies, while the selections based on the best Omni score and the closest to MET criterion show very similar performance.



(a) SM

(b) $\kappa_\lambda = 0$

(c) $\kappa_{2V} = 0$

(d) $\kappa_V = 0.5$

Figure 7.3: Distributions of the GN2X score for the selected $b\bar{b}$ large-R jet candidates for different Higgs boson coupling scenarios. The Standard Model configuration is shown in the top-left panel, while the remaining panels correspond to variations in κ_λ, κ_{2V}, and κ_V.

7.3.1 Efficiency

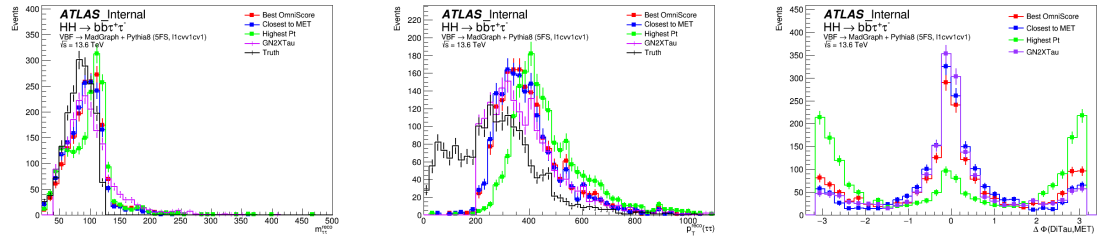
The efficiency of the reconstructed HH with the different di-τ candidate selection is defined as

$$\epsilon = \frac{N_{\text{pass}}}{N_{\text{total}}} \tag{7.2}$$

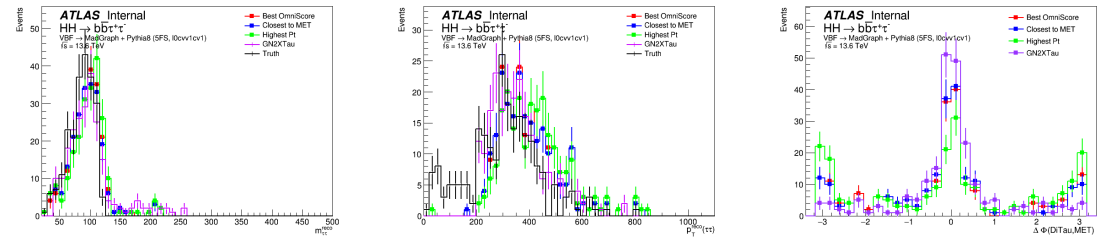
where N_{pass} corresponds to the number of events that satisfy the trut.matching requirements, and N_{total} is the total number of selected events. An event is considered as passing the truth match if both resconstructed objects are matched to their corresponding truth-level objects as

$$N_{\text{pass}} = \Delta R(\text{DiTau}, \text{Truth}_{\tau\tau}) < 0.6 \quad \text{and} \quad \Delta R(b\bar{b}, \text{Truth}_{b\bar{b}}) < 0.6 \tag{7.3}$$

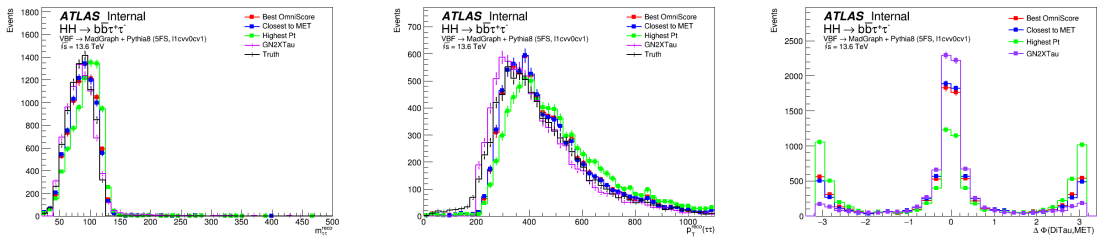
Figure 7.9 shows efficiency as a function of the truth HH invariant mass and



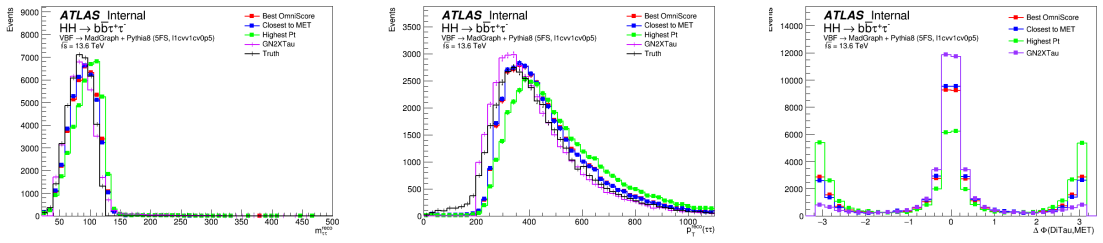
(a) SM



(b) $\kappa_\lambda = 0$, $\kappa_{2V} = 1$, $\kappa_V = 1$



(c) $\kappa_\lambda = 1$, $\kappa_{2V} = 0$, $\kappa_V = 1$

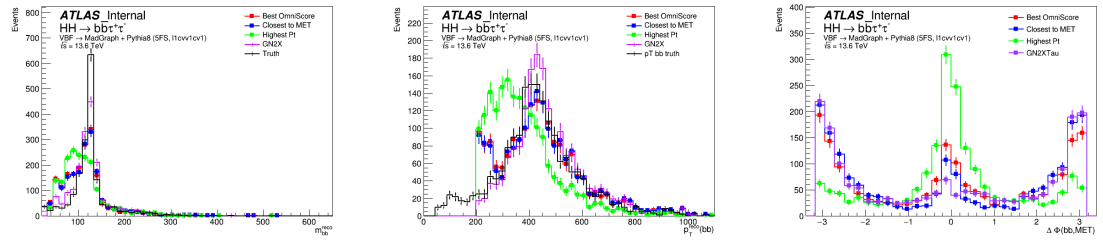


(d) $\kappa_\lambda = 1$, $\kappa_{2V} = 1$, $\kappa_V = 0.5$

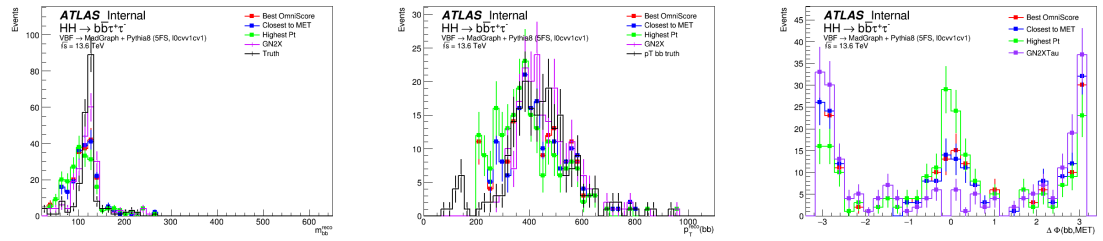
Figure 7.4: Distributions of the reconstructed di-τ invariant mass (left), transverse momentum (center), and azimuthal separation with respect to the missing transverse energy (right). Each row corresponds to a different coupling scenario.

transverse momentum. Among the different selection strategies, the highest efficiency for all coupling scenarios is obtained when the di-τ candidate is selected using the GN2XTau score. This selection is closely followed by the best Omni score and the closest to MET strategies, which exhibit very similar performance. In contrast, the highest-p_T selection shows the lowest efficiency across the full mass and p_T ranges.

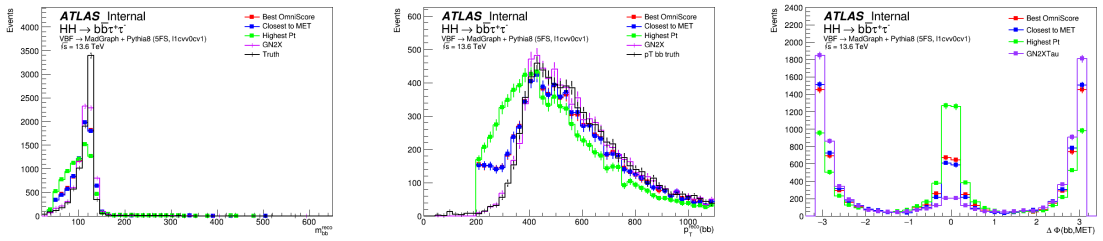
For the $\kappa_{2V} = 0$ and $\kappa_V = 0.5$ couplings, both taggers, across their different selection criteria, operate close to their maximal efficiency. In contrast, for the SM



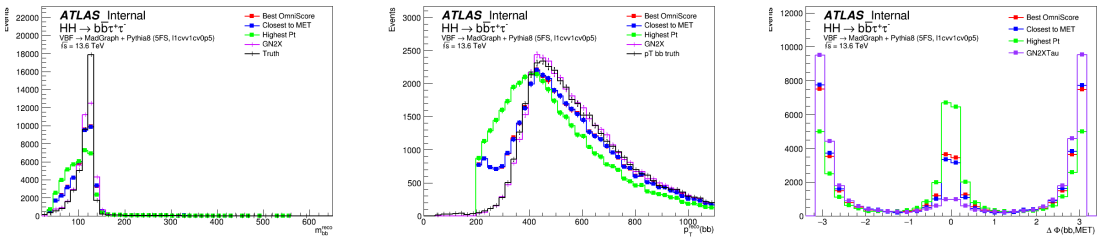
(a) SM



(b) $\kappa_\lambda = 0,\ \kappa_{2V} = 1,\ \kappa_V = 1$



(c) $\kappa_\lambda = 1,\ \kappa_{2V} = 0,\ \kappa_V = 1$



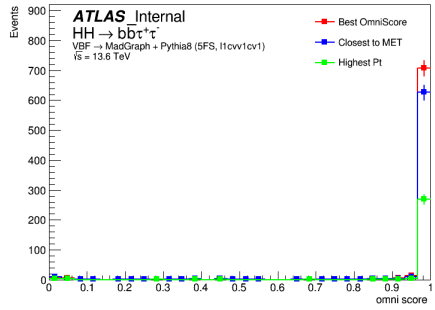
(d) $\kappa_\lambda = 1,\ \kappa_{2V} = 1,\ \kappa_V = 0.5$

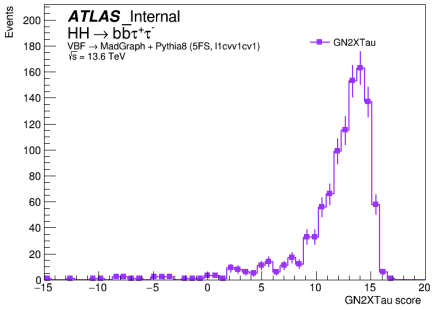
Figure 7.5: Distributions of the reconstructed $b\bar{b}$ invariant mass (left), transverse momentum (center), and azimuthal separation with respect to the missing transverse energy (right). Each row corresponds to a different coupling scenario.

and $\kappa_\lambda = 0$ scenarios, the efficiency decreases at high $m(HH)$ values and low $p_T(HH)$.

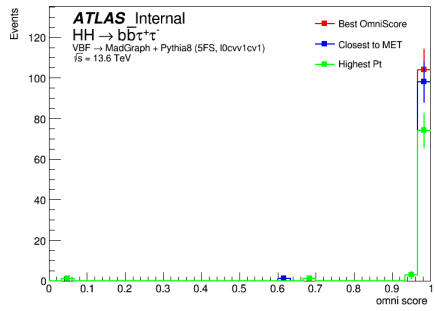
Figure 7.10 presents the number of matches between reconstructed objects and truth-level objects for each selection. In this figure, a value of 1 indicates a match of Truth$_{\tau\tau}$, a value of 2 corresponds to a match to Truth$_{b\bar{b}}$ and 0 indicates not match, all defined using a $\Delta R < 0.6$ requirement.

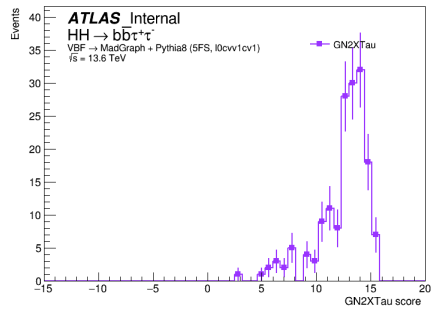
The highest-p_T selection produce the lowest number of matches to Truth$\tau\tau$ and the highest number of matches to Truth$b\bar{b}$. This behavior is consistent with



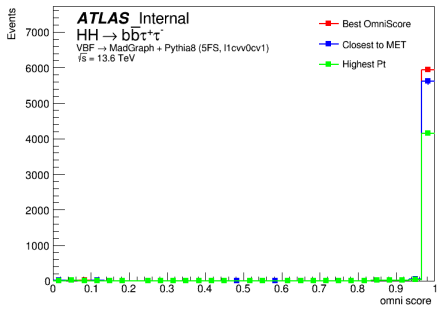


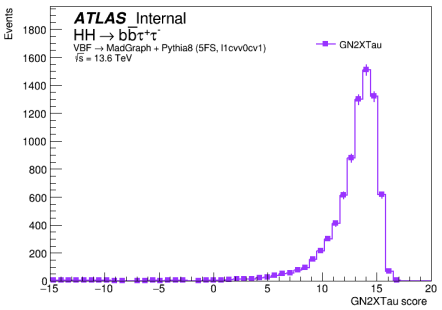
(a) $\kappa_\lambda = 1$, $\kappa_{2V} = 1$, $\kappa_V = 1$



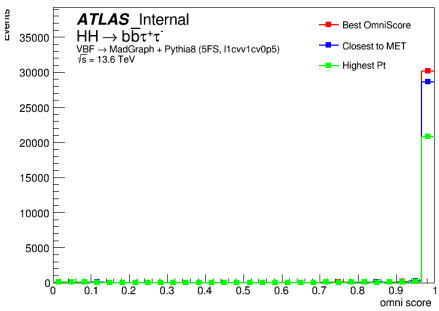


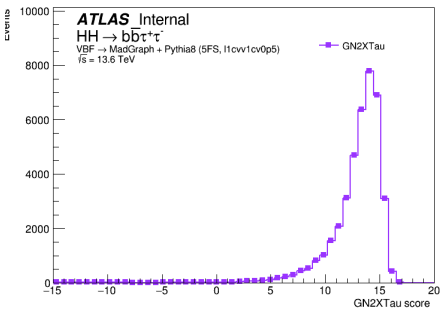
(b) $\kappa_\lambda = 0$, $\kappa_{2V} = 1$, $\kappa_V = 1$





(c) $\kappa_{2V} = 0$, $\kappa_\lambda = 1$, $\kappa_V = 1$

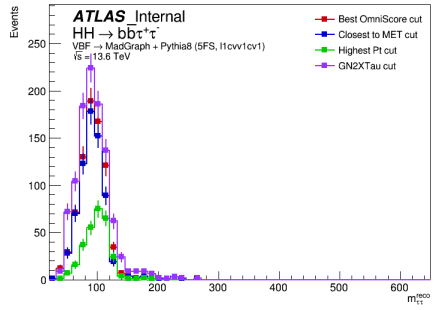


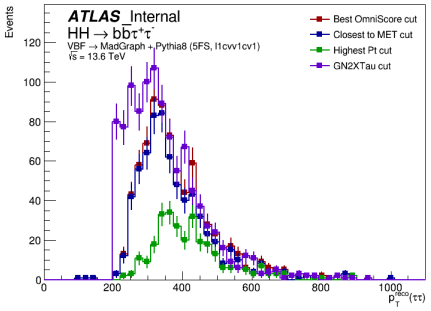


(d) $\kappa_V = 0.5$, $\kappa_\lambda = 1$, $\kappa_{2V} = 1$

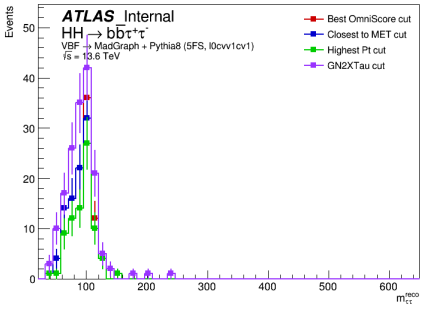
Figure 7.6: Distributions of the Omni score (left column) and GN2XTau score (right column) after applying the truth-matching requirement, shown for different coupling scenarios.

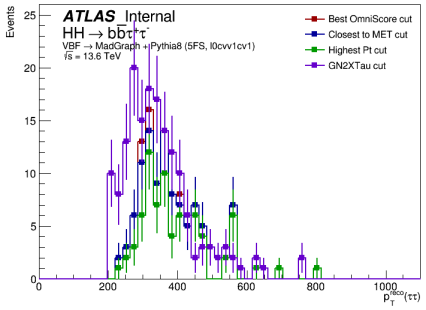
expectations, since the di-τ system is accompanied by neutrinos, which reduces



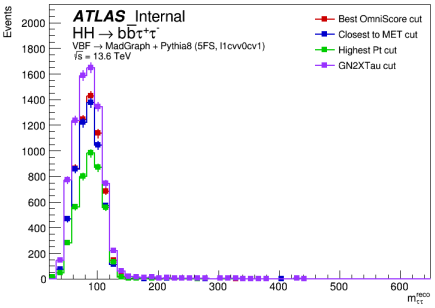


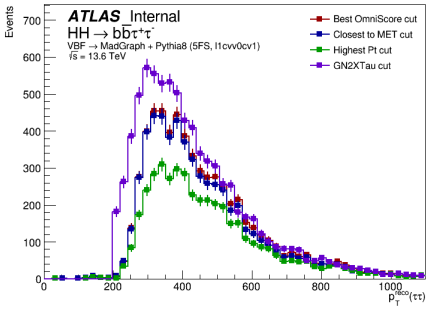
(a) $\kappa_\lambda = 1,\ \kappa_{2V} = 1,\ \kappa_V = 1$



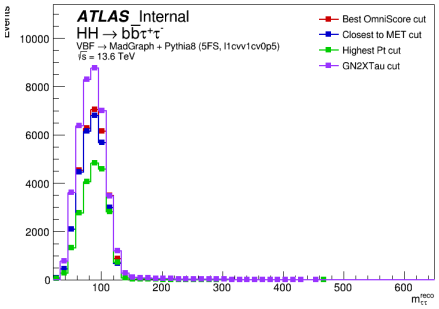


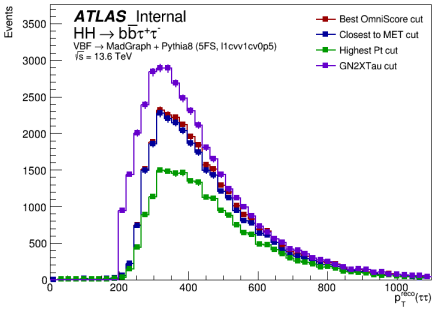
(b) $\kappa_\lambda = 0,\ \kappa_{2V} = 1,\ \kappa_V = 1$





(c) $\kappa_{2V} = 0,\ \kappa_\lambda = 1,\ \kappa_V = 1$

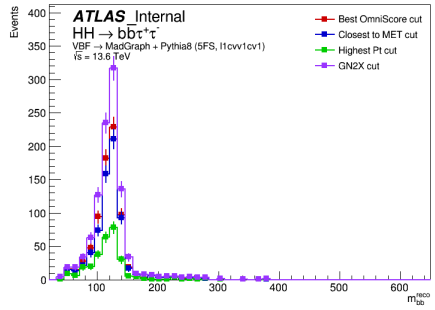


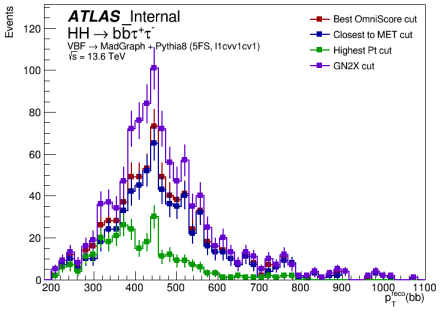


(d) $\kappa_V = 0.5,\ \kappa_\lambda = 1,\ \kappa_{2V} = 1$

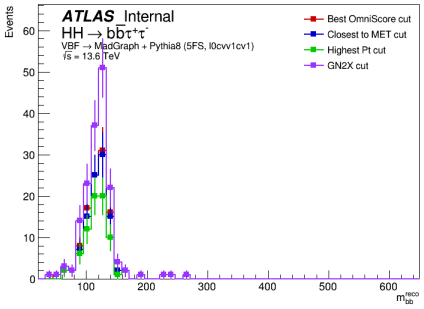
Figure 7.7: Distributions of the reconstructed di-τ masses and transverse momenta after truth-matching.

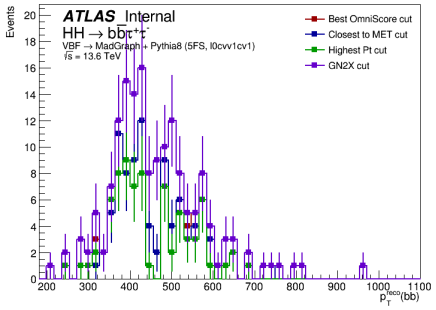
its reconstructed transverse momentum compared to the $b\bar{b}$ system. As a consequence, selecting the object with the highest p_T tends to favor the $b\bar{b}$ candidate



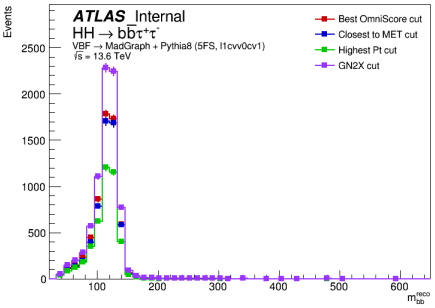


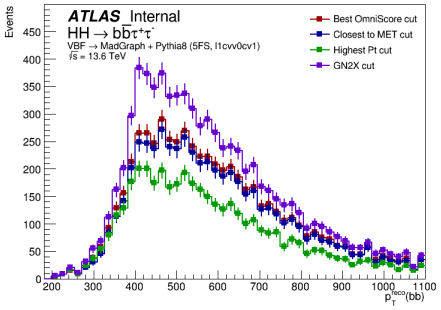
(a) $\kappa_\lambda = 1,\ \kappa_{2V} = 1,\ \kappa_V = 1$



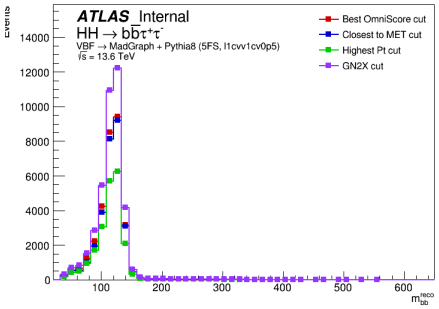


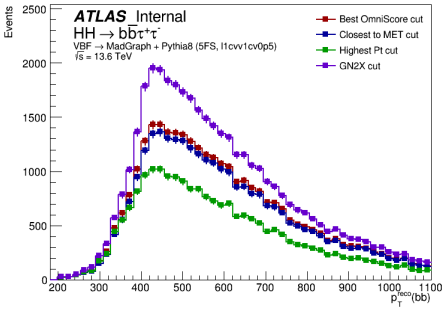
(b) $\kappa_\lambda = 0,\ \kappa_{2V} = 1,\ \kappa_V = 1$





(c) $\kappa_{2V} = 0,\ \kappa_\lambda = 1,\ \kappa_V = 1$





(d) $\kappa_V = 0.5,\ \kappa_\lambda = 1,\ \kappa_{2V} = 1$

Figure 7.8: Distributions of the reconstructed $b\bar{b}$ masses and transverse momenta after truth-matching.

rather than the di-τ.

For all coupling scenarios, the best overall performance is observed for the

GN2XTau selection, which yields the largest fraction of correct matches to Truth$_{\tau\tau}$ and only a small number of misidentified $b\bar{b}$ matches. This is followed by the best Omni score and the closest-to-MET selections, which exhibit very similar performance.

The $\kappa_{2V} = 0$ and $\kappa_V = 0.5$ coupling scenarios show the lowest mismatch rates and almost no unclassified events. In particular, the GN2XTau selection correctly classifies nearly all events as di-τ candidates.

This performance difference can be understood from the desing of the taggers. GN2XTau employd a discriminant capable of separating di-τ jets from other jet flavours, including b-jets, c-jets and light-flavour jets. In contrast, the Omni tagger use a bynary classification, identifying wheter a jet is compatible with a di-τ or not, without explicitly discriminating against other jets flavor, as discussed in Subsection 3.5.6. This can lead to a higher rate of misidentification when the Omni score is not tight. Consequently, GN2XTau provides a better performance in the identification of boosted di-τ for the $HH \to b\bar{b}\tau\tau$ analysis.

7.3.2 Comparative study

As shown in Figure 7.9, different Higgs boson coupling scenarios lead to different selection efficiencies. For the Standard Model and the $\kappa_\lambda = 0$ configuration, no significant differences are observed among the various di-τ selection strategies. In contrast, for the other coupling scenarios, all selections exhibit a higher efficiency. This behavior suggests that the event kinematics are modified by the choice of couplings and motivates a more detailed study to fully understand the underlying effects.

Figure 7.11 shows the truth-level distributions of $p_T(HH)$, $m(HH)$, and the angular separation $\Delta R(\tau, \tau)$ for the different coupling values. For the $p_T(HH)$ distribution, the Standard Model and the $\kappa_\lambda = 0$ scenarios present a similar event yield across most of the spectrum, with differences appearing only at high transverse momentum. In contrast, for the $\kappa_{2V} = 0$ and $\kappa_V = 0.5$ configurations, the event yield is enhanced at low $p_T(HH)$ (below $\sim$ 200 GeV) and decreases more rapidly at higher transverse momentum.

In the truth-level $m(HH)$ distribution, the $\kappa_{2V} = 0$ and $\kappa_V = 0.5$ scenarios show a significantly larger fraction of events at high invariant masses ($m(HH) >$ 1 TeV) compared to the Standard Model and $\kappa_\lambda = 0$ cases, which are dominated by lower-mass events. This enhancement at high $m(HH)$ can lead to more boosted and therefore more collimated decay products, as reflected in the smaller $\Delta R(\tau, \tau)$ values observed for these coupling scenarios. In contrast, the Standard

Model and $\kappa_\lambda = 0$ configurations exhibit broader $\Delta R(\tau, \tau)$ distributions, consistent with less boosted Higgs boson decays.

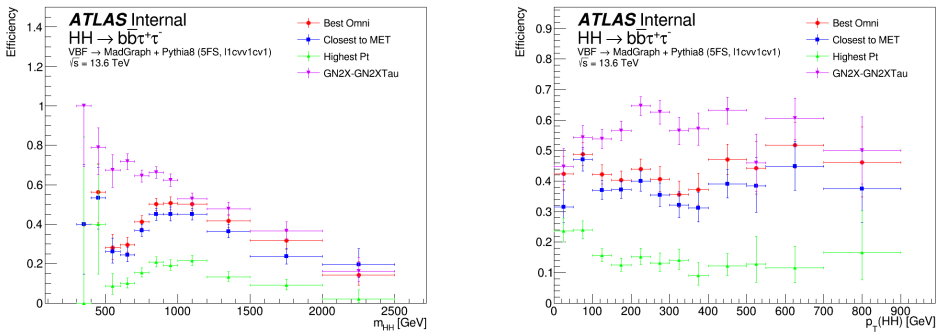
Figures 7.12 and 7.13 illustrate how the collimation of the two truth-level τ leptons is correlated with transverse momentum of the Higgs boson pair $p_T(HH)$ and invariant mass $m(HH)$. For all coupling scenarios, a high concentration of highly collimated τ pairs ($\Delta R < 0.8$) is observed in the low transverse momentum region, between 0 and 200 GeV.

In contrast, the correlation with the invariant mass shows that this collimation is predominantly associated with the high-mass region, between approximately 1 and 3 TeV. In this regime, the coupling scenarios $\kappa_{2V} = 0$ and $\kappa_V = 0.5$ clearly dominate, exhibiting a significantly larger population of collimated τ pairs. This behavior is quantified in Table 7.2, where 92.5% and 91.7% of the events are found to satisfy $\Delta R(\tau, \tau) < 0.8$ for the $\kappa_{2V} = 0$ and $\kappa_V = 0.5$ scenarios, respectively. In comparison, only 58.5% of the events in the Standard Model and 67.2% in the $\kappa_\lambda = 0$ configuration are collimated.

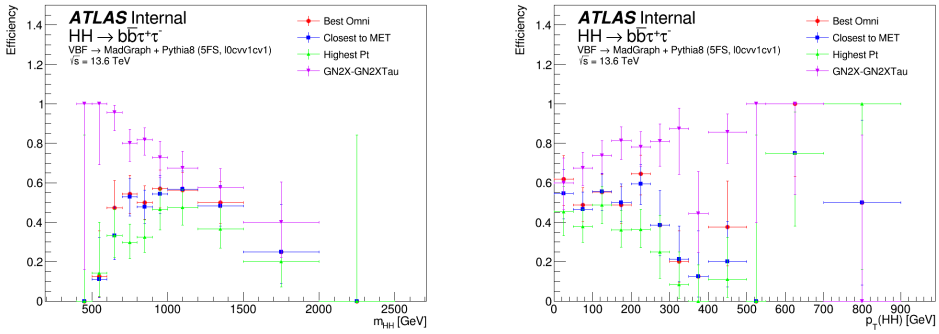
This difference in the degree of collimation has a direct impact on the selection efficiencies, as shown in Figure 7.9. Since not all events are equally boosted, the performance of the di-τ selection depends on the underlying kinematics. In particular, GN2XTau and Omniscore selections achieve higher efficiencies in scenarios with a larger fraction of boosted and collimated τ pairs, as expected.

Sample	$\Delta R(\tau, \tau) < 0.8$	$p_T(HH) > 400$ **GeV**	$m(HH) > 1000$ **GeV**
l1cvv1cv1	58.5%	41.0%	39.4%
l0cvv1cvv1	67.2%	27.7%	47.3%
l1cvv0cv1	92.5%	4.4%	83.7%
l1cvv1cv0p5	91.7%	4.7%	83.0%

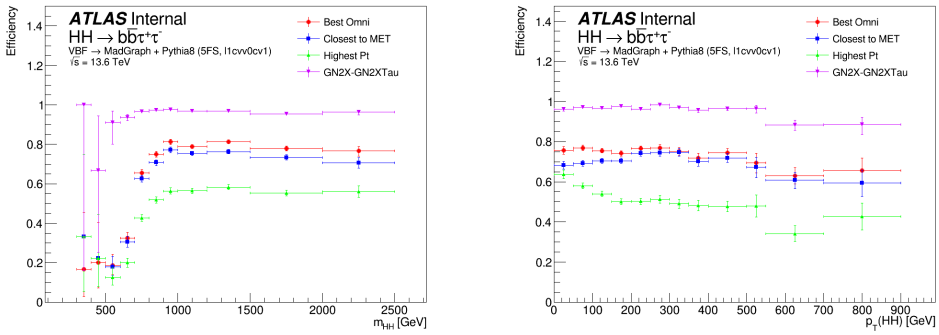
Table 7.2: Fraction of events passing different kinematic requirements for the four Higgs boson coupling scenarios. The selections correspond to $\Delta R(\tau, \tau) < 0.8$, $p_T(HH) > 400$ GeV, and $m(HH) > 1000$ GeV.



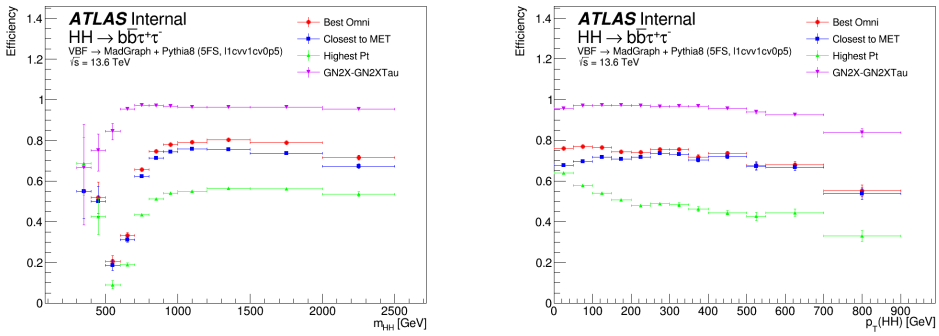
(a) $\kappa_\lambda = 1,\ \kappa_{2V} = 1,\ \kappa_V = 1$



(b) $\kappa_\lambda = 0,\ \kappa_{2V} = 1,\ \kappa_V = 1$

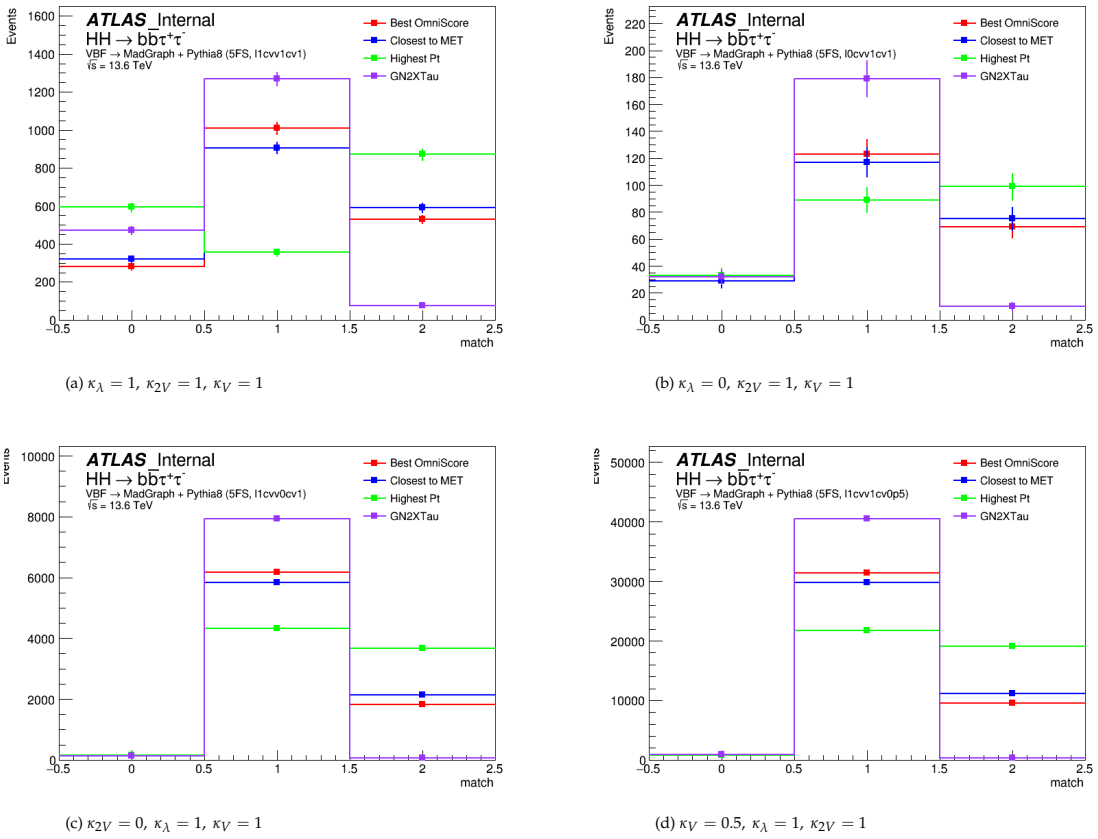


(c) $\kappa_{2V} = 0,\ \kappa_\lambda = 1,\ \kappa_V = 1$



(d) $\kappa_V = 0.5,\ \kappa_\lambda = 1,\ \kappa_{2V} = 1$

Figure 7.9: Efficiency of the HH event selection as a function of the truth-level invariant mass m_{HH} (left column) and transverse momentum p_T^{HH} (right column), shown for the different di-τ candidate selection strategies.



(a) $\kappa_\lambda = 1,\ \kappa_{2V} = 1,\ \kappa_V = 1$

(b) $\kappa_\lambda = 0,\ \kappa_{2V} = 1,\ \kappa_V = 1$

(c) $\kappa_{2V} = 0,\ \kappa_\lambda = 1,\ \kappa_V = 1$

(d) $\kappa_V = 0.5,\ \kappa_\lambda = 1,\ \kappa_{2V} = 1$

Figure 7.10: Number of truth matches for the different di-τ candidate selection strategies. A value of 1 corresponds to a match with the true di-τ system, 2 to a match with the true $b\bar{b}$ system, and 0 indicates no truth match.

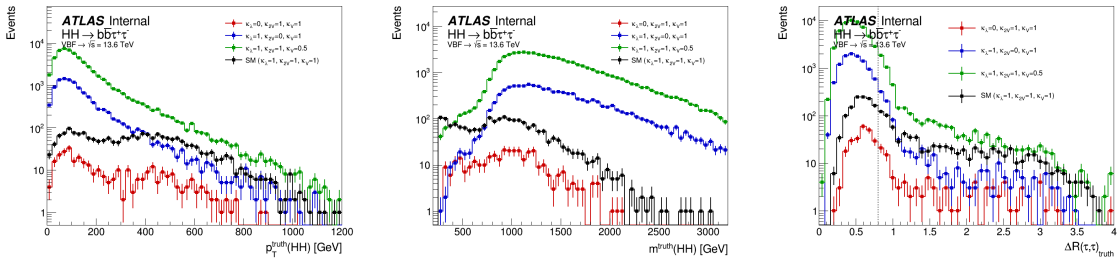


Figure 7.11: Truth-level distributions of the Higgs boson pair transverse momentum p_T^{HH}, invariant mass m_{HH}, and angular separation $\Delta R_{\tau\tau}$, shown for the different coupling scenarios.

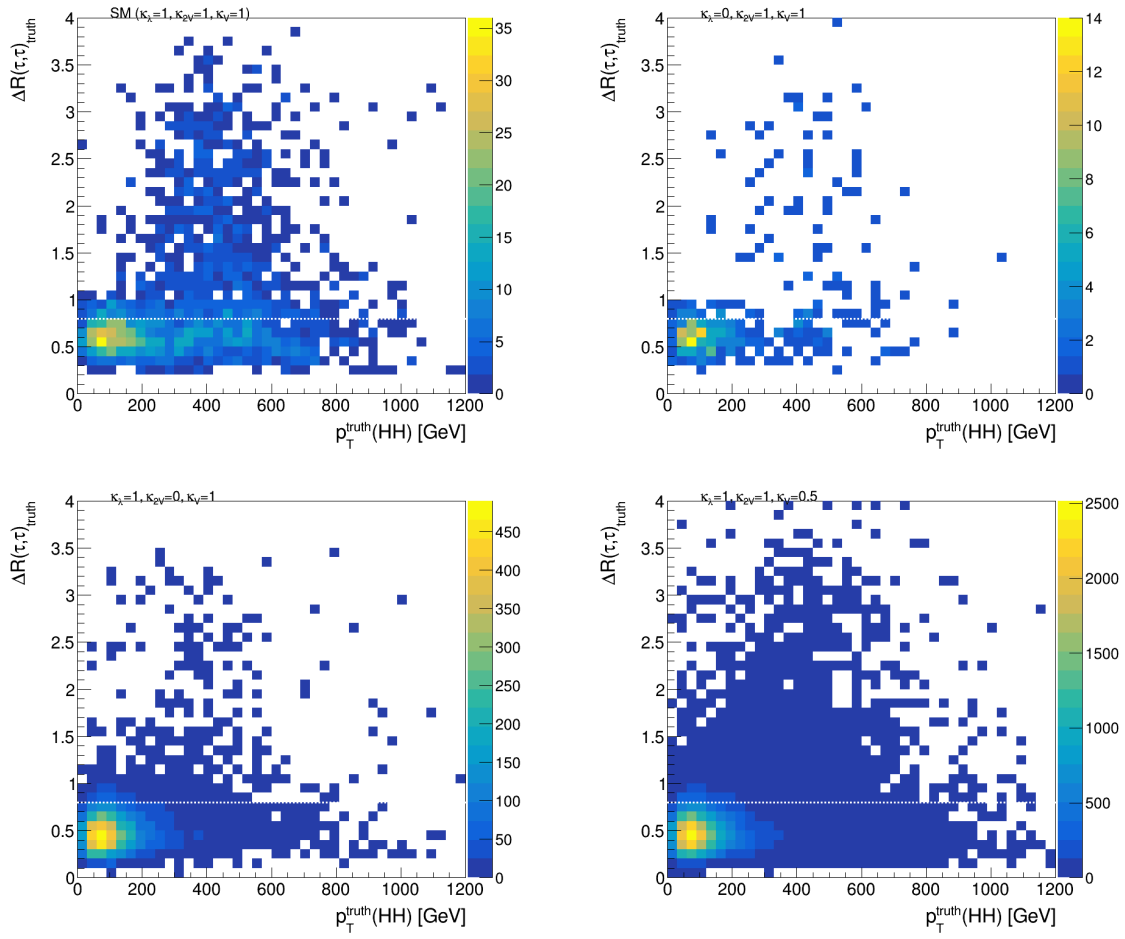


Figure 7.12: Correlation between the angular separation of the di-τ system and the transverse momentum of the Higgs boson pair, shown for the different Higgs boson coupling scenarios.

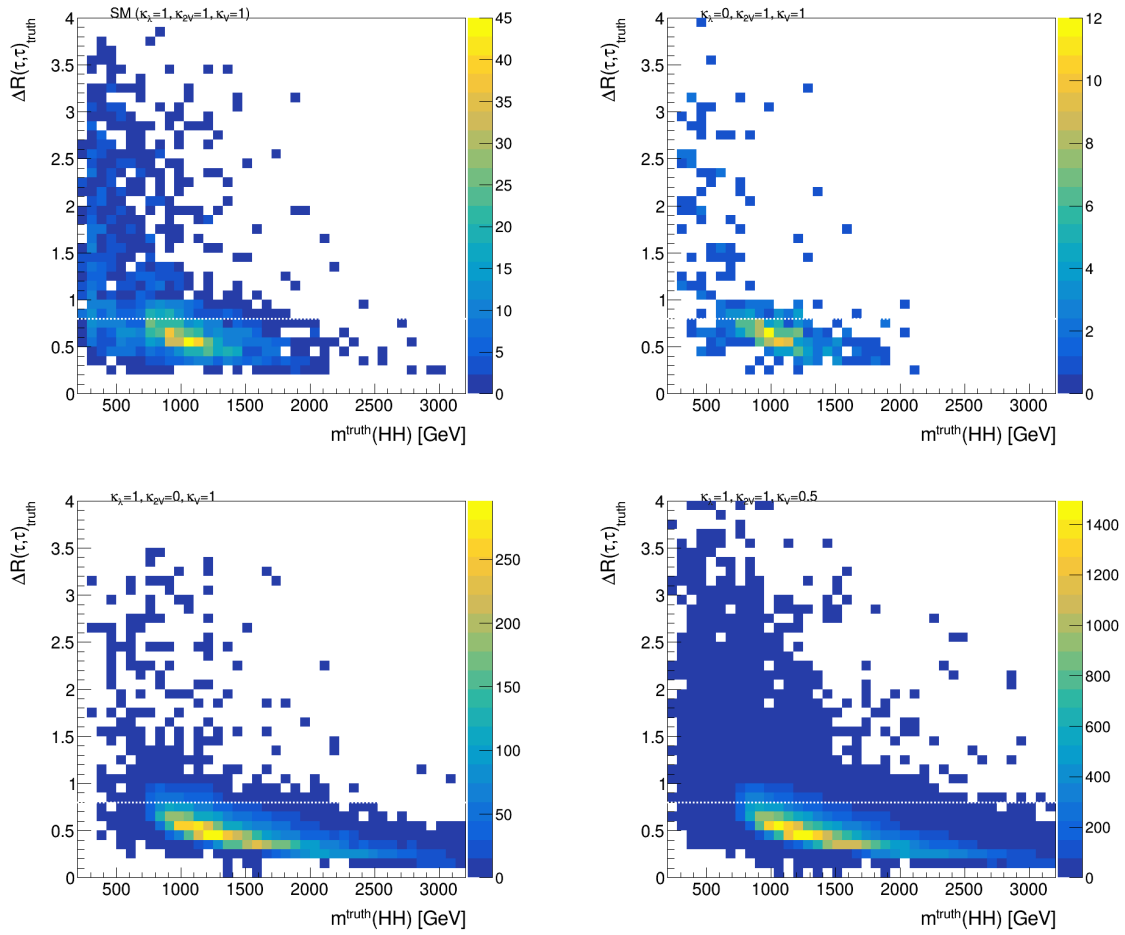


Figure 7.13: Correlation between the angular separation of the di-τ system and the mass of the Higgs boson pair, shown for the different Higgs boson coupling scenarios.

CHAPTER 8

Conclusions and Outlook

The discovery of the Higgs boson by the ATLAS and CMS collaborations at the Large Hadron Collider (LHC) [1, 2] opened new directions in experimental particle physics and completed the Standard Model (SM). Studies of Higgs boson production, both single and pair production, are particularly interesting in the high transverse-momentum (p_T), or boosted, regime. In this region, the identification and triggering of Higgs boson decay products become increasingly challenging for the ATLAS experiment, as the decay products are highly collimated and standard reconstruction techniques lose efficiency.

In parallel, the upcoming High-Luminosity LHC (HL-LHC) aims to reach a peak instantaneous luminosity of approximately $5 \times 10^{34}\text{cm}^{-2}\text{s}^{-1}$, significantly improving the sensitivity to rare processes, particularly at high p_T. However, the current ATLAS Inner Detector (ID) [21] was designed for operation at $10^{34}, \text{cm}^{-2}\text{s}^{-1}$ over a 10-year period and is therefore not suitable for HL-LHC conditions.

This thesis presents the implementation of a new ITk Strip ByteStream (BS) converter package for the simulation of the ITk Strip detector within the Athena framework. The implementation is divided into two main components. The encoding step converts Raw Data Objects (RDO) into BS format and was first developed in the merge request !76971. The decoding step, interpret the BS data and reconstructing the corresponding RDOs. Its first implementation was introduced in !81433. The algorithm and examples of this implementation are presented in Chapter 4.

In addition, Chapter 5 presents the development of a new monitoring algorithm for boosted di-τ triggers. This implementation follows the standard ATLAS trigger monitoring framework, and its first version was introduced in !81245. The algorithm extracts key di-τ variables in order to ensure correct trigger performance.

The performance of this boosted di-τ trigger was then compared with other trigger categories, including large-R jets, small-R jets, single-τ and MET triggers, in the boosted $H \to \tau^+\tau^-$ analysis. The boosted di-τ trigger shows improved performance in the range $300 < p_T < 400$ GeV, and maintains an approximately constant efficiency of about 0.7 at higher p_T, outperforming the other triggers. The full analysis is presented in Chapter 6.

Finally, this thesis compares two boosted di-τ identification strategies, a Graph Neural Network–based tagger (GN2XTau) [105] and the Omni Tagger [104]. Their performance is evaluated in the boosted $HH \rightarrow b\bar{b}\tau^+\tau^-$ channel using four VBF Monte Carlo samples corresponding to different coupling scenarios. The efficiency studies show that GN2XTau provides the best overall performance across all couplings, while the highest-p_T di-τ selection exhibits the poorest performance. Furthermore, the reconstruction efficiency is found to be close to maximal for the $\kappa_{2V} = 0$ and $\kappa_V = 0.5$ coupling scenarios, whereas for the remaining couplings the efficiency decreases at high $m(HH)$ values. A detailed comparison and discussion of these results are presented in Chapter 7.

APPENDIX A

Channels of the ABCStar chips

near strip	ABC Physical pad	ABC Logical channel	Cluster address (8b hex)	far strip	ABC Physical pad	ABC Logical channel	Cluster address (8b hex)
n0	0	0	0x00	f0	1	1	0x80
n1	2	2	0x01	f1	3	3	0x81
n2	4	4	0x02	f2	5	5	0x82
n3	6	6	0x03	f3	7	7	0x83
n4	8	8	0x04	f4	9	9	0x84
n5	10	10	0x05	f5	11	11	0x85
n6	12	12	0x06	f6	13	13	0x86
n7	14	14	0x07	f7	15	15	0x87
n8	16	16	0x08	f8	17	17	0x88
n9	18	18	0x09	f9	19	19	0x89
n10	20	20	0x0A	f10	21	21	0x8A
n11	22	22	0x0B	f11	23	23	0x8B
n12	24	24	0x0C	f12	25	25	0x8C
n13	26	26	0x0D	f13	27	27	0x8D
n14	28	28	0x0E	f14	29	29	0x8E
n15	30	30	0x0F	f15	31	31	0x8F
n16	32	32	0x10	f16	33	33	0x90
n17	34	34	0x11	f17	35	35	0x91
n18	36	36	0x12	f18	37	37	0x92
n19	38	38	0x13	f19	39	39	0x93
n20	40	40	0x14	f20	41	41	0x94
n21	42	42	0x15	f21	43	43	0x95
n22	44	44	0x16	f22	45	45	0x96
n23	46	46	0x17	f23	47	47	0x97
n24	48	48	0x18	f24	49	49	0x98
n25	50	50	0x19	f25	51	51	0x99
n26	52	52	0x1A	f26	53	53	0x9A

near strip	ABC Physical pad	ABC Logical channel	Cluster address (8b hex)	far strip	ABC Physical pad	ABC Logical channel	Cluster address (8b hex)
n27	54	54	0x1B	f27	55	55	0x9B
n28	56	56	0x1C	f28	57	57	0x9C
n29	58	58	0x1D	f29	59	59	0x9D
n30	60	60	0x1E	f30	61	61	0x9E
n31	62	62	0x1F	f31	63	63	0x9F
n32	64	64	0x20	f32	65	65	0xA0
n33	66	66	0x21	f33	67	67	0xA1
n34	68	68	0x22	f34	69	69	0xA2
n35	70	70	0x23	f35	71	71	0xA3
n36	72	72	0x24	f36	73	73	0xA4
n37	74	74	0x25	f37	75	75	0xA5
n38	76	76	0x26	f38	77	77	0xA6
n39	78	78	0x27	f39	79	79	0xA7
n40	80	80	0x28	f40	81	81	0xA8
n41	82	82	0x29	f41	83	83	0xA9
n42	84	84	0x2A	f42	85	85	0xAA
n43	86	86	0x2B	f43	87	87	0xAB
n44	88	88	0x2C	f44	89	89	0xAC
n45	90	90	0x2D	f45	91	91	0xAD
n46	92	92	0x2E	f46	93	93	0xAE
n47	94	94	0x2F	f47	95	95	0xAF
n48	96	96	0x30	f48	97	97	0xB0
n49	98	98	0x31	f49	99	99	0xB1
n50	100	100	0x32	f50	101	101	0xB2
n51	102	102	0x33	f51	103	103	0xB3
n52	104	104	0x34	f52	105	105	0xB4
n53	106	106	0x35	f53	107	107	0xB5
n54	108	108	0x36	f54	109	109	0xB6
n55	110	110	0x37	f55	111	111	0xB7
n56	112	112	0x38	f56	113	113	0xB8
n57	114	114	0x39	f57	115	115	0xB9
n58	116	116	0x3A	f58	117	117	0xBA

near strip	ABC Physical pad	ABC Logical channel	Cluster address (8b hex)	far strip	ABC Physical pad	ABC Logical channel	Cluster address (8b hex)
n59	118	118	0x3B	f59	119	119	0xBB
n60	120	120	0x3C	f60	121	121	0xBC
n61	122	122	0x3D	f61	123	123	0xBD
n62	124	124	0x3E	f62	125	125	0xBE
n63	126	126	0x3F	f63	127	127	0xBF
n64	128	128	0x40	f64	129	129	0xC0
n65	130	130	0x41	f65	131	131	0xC1
n66	132	132	0x42	f66	133	133	0xC2
n67	134	134	0x43	f67	135	135	0xC3
n68	136	136	0x44	f68	137	137	0xC4
n69	138	138	0x45	f69	139	139	0xC5
n70	140	140	0x46	f70	141	141	0xC6
n71	142	142	0x47	f71	143	143	0xC7
n72	144	144	0x48	f72	145	145	0xC8
n73	146	146	0x49	f73	147	147	0xC9
n74	148	148	0x4A	f74	149	149	0xCA
n75	150	150	0x4B	f75	151	151	0xCB
n76	152	152	0x4C	f76	153	153	0xCC
n77	154	154	0x4D	f77	155	155	0xCD
n78	156	156	0x4E	f78	157	157	0xCE
n79	158	158	0x4F	f79	159	159	0xCF
n80	160	160	0x50	f80	161	161	0xD0
n81	162	162	0x51	f81	163	163	0xD1
n82	164	164	0x52	f82	165	165	0xD2
n83	166	166	0x53	f83	167	167	0xD3
n84	168	168	0x54	f84	169	169	0xD4
n85	170	170	0x55	f85	171	171	0xD5
n86	172	172	0x56	f86	173	173	0xD6
n87	174	174	0x57	f87	175	175	0xD7
n88	176	176	0x58	f88	177	177	0xD8
n89	178	178	0x59	f89	179	179	0xD9
n90	180	180	0x5A	f90	181	181	0xDA

near strip	ABC Physical pad	ABC Logical channel	Cluster address (8b hex)	far strip	ABC Physical pad	ABC Logical channel	Cluster address (8b hex)
n91	182	182	0x5B	f91	183	183	0xDB
n92	184	184	0x5C	f92	185	185	0xDC
n93	186	186	0x5D	f93	187	187	0xDD
n94	188	188	0x5E	f94	189	189	0xDE
n95	190	190	0x5F	f95	191	191	0xDF
n96	192	192	0x60	f96	193	193	0xE0
n97	194	194	0x61	f97	195	195	0xE1
n98	196	196	0x62	f98	197	197	0xE2
n99	198	198	0x63	f99	199	199	0xE3
n100	200	200	0x64	f100	201	201	0xE4
n101	202	202	0x65	f101	203	203	0xE5
n102	204	204	0x66	f102	205	205	0xE6
n103	206	206	0x67	f103	207	207	0xE7
n104	208	208	0x68	f104	209	209	0xE8
n105	210	210	0x69	f105	211	211	0xE9
n106	212	212	0x6A	f106	213	213	0xEA
n107	214	214	0x6B	f107	215	215	0xEB
n108	216	216	0x6C	f108	217	217	0xEC
n109	218	218	0x6D	f109	219	219	0xED
n110	220	220	0x6E	f110	221	221	0xEE
n111	222	222	0x6F	f111	223	223	0xEF
n112	224	224	0x70	f112	225	225	0xF0
n113	226	226	0x71	f113	227	227	0xF1
n114	228	228	0x72	f114	229	229	0xF2
n115	230	230	0x73	f115	231	231	0xF3
n116	232	232	0x74	f116	233	233	0xF4
n117	234	234	0x75	f117	235	235	0xF5
n118	236	236	0x76	f118	237	237	0xF6
n119	238	238	0x77	f119	239	239	0xF7
n120	240	240	0x78	f120	241	241	0xF8
n121	242	242	0x79	f121	243	243	0xF9
n122	244	244	0x7A	f122	245	245	0xFA

near strip	**ABC Physical pad**	**ABC Logical channel**	**Cluster address (8b hex)**	**far strip**	**ABC Physical pad**	**ABC Logical channel**	**Cluster address (8b hex)**
n123	246	246	0x7B	f123	247	247	0xFB
n124	248	248	0x7C	f124	249	249	0xFC
n125	250	250	0x7D	f125	251	251	0xFD
n126	252	252	0x7E	f126	253	253	0xFE
n127	254	254	0x7F	f127	255	255	0xFF

Table A.1: Relationship between physical FE pads, logical channels and cluster addresses

APPENDIX B

Data samples for the Analysis

B.1 Data sample used to test the new boosted di-τ trigger monitoring algorithm

data25_13p6TeV.00502896.physics_Main.merge.AOD.f1612_m2272

B.2 Ntuples for the trigger performance for $H \to \tau^+\tau^-$ analysis

user.ademaria.Htt_D25Bst.data25_13p6TeV.00502742.physics_Main.PHYS.f1608_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00502748.physics_Main.PHYS.f1608_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00502782.physics_Main.PHYS.f1608_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00502844.physics_Main.PHYS.f1610_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00502880.physics_Main.PHYS.f1612_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00502886.physics_Main.PHYS.f1612_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00502896.physics_Main.PHYS.f1612_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00502920.physics_Main.PHYS.f1612_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00503090.physics_Main.PHYS.f1612_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00503100.physics_Main.PHYS.f1613_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00503184.physics_Main.PHYS.f1613_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00503223.physics_Main.PHYS.f1613_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00503274.physics_Main.PHYS.f1614_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00503591.physics_Main.PHYS.f1615_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00503616.physics_Main.PHYS.f1615_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00503636.physics_Main.PHYS.f1615_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00503641.physics_Main.PHYS.f1615_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00503662.physics_Main.PHYS.f1615_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00503760.physics_Main.PHYS.f1615_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00503778.physics_Main.PHYS.f1615_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00503795.physics_Main.PHYS.f1615_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00503865.physics_Main.PHYS.f1617_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00503886.physics_Main.PHYS.f1617_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00503910.physics_Main.PHYS.f1617_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00503935.physics_Main.PHYS.f1617_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00504083.physics_Main.PHYS.f1617_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00504091.physics_Main.PHYS.f1617_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00504096.physics_Main.PHYS.f1617_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00504209.physics_Main.PHYS.f1617_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00504214.physics_Main.PHYS.f1617_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00504244.physics_Main.PHYS.f1617_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00504373.physics_Main.PHYS.f1617_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00504392.physics_Main.PHYS.f1617_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00504403.physics_Main.PHYS.f1617_m2272_p6828.n_5_HS
user.ademaria.Htt_D25Bst.data25_13p6TeV.00504406.physics_Main.PHYS.f1617_m2272_p6828.n_5_HS

B.3 Monte Carlo Samples for the Boosted Di-τ Identification Study in $HH \to b\bar{b}\tau^+\tau^-$

The Monte Carlo samples used in the boosted di-τ identification study for the $HH \to b\bar{b}\tau^+\tau^-$ channel are listed below. All samples correspond to VBF HH production generated with `MGPy8EG` at $\sqrt{s} = 13.6$ TeV and are provided in the `DAOD_PHYS` format.

mc23_13p6TeV.542841.MGPy8EG_hh_bbtt_vbf_novhh_5fs_l1cvv1cv1_hh.deriv.DAOD_PHYS.e8562_e8528_a911_s4114_r15224_r15225_p6909
mc23_13p6TeV.542839.MGPy8EG_hh_bbtt_vbf_novhh_5fs_l0cvv1cv1_hh.deriv.DAOD_PHYS.e8562_e8528_s4159_s4114_r15224_r15225_p6909
mc23_13p6TeV.542840.MGPy8EG_hh_bbtt_vbf_novhh_5fs_l1cvv0cv1_hh.deriv.DAOD_PHYS.e8562_e8528_s4159_s4114_r15224_r15225_p6909
mc23_13p6TeV.542842.MGPy8EG_hh_bbtt_vbf_novhh_5fs_l1cvv1cv0p5_hh.deriv.DAOD_PHYS.e8562_e8528_a911_s4114_r15224_r15225_p6909

References

[1] ATLAS Collaboration et al., "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", arXiv:1207.7214 (2012).

[2] CMS Collaboration et al., "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", arXiv:1207.7235 (2012).

[3] A. Collaboration, "Search for pair production of higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton–proton collisions at $\sqrt{s} = 13$ tev with the atlas detector", Journal of High Energy Physics **2019**, 30 (2019).

[4] A. Collaboration, "Erratum to: search for the hh $\rightarrow b\bar{b}b\bar{b}$ process via vector-boson fusion production using proton–proton collisions at $\sqrt{s} = 13$ tev with the atlas detector", Journal of High Energy Physics **2021**, 145 (2021).

[5] A. Collaboration, "Search for resonant pair production of higgs bosons in the $b\bar{b}b\bar{b}$ final state using pp collisions at $\sqrt{s} = 13$ tev with the atlas detector", Physical Review D **105**, 092002 (2022).

[6] A. Collaboration, "Search for nonresonant pair production of higgs bosons in the $b\bar{b}b\bar{b}$ final state in pp collisions at $\sqrt{s} = 13$ tev with the atlas detector", Journal of High Energy Physics (2023).

[7] A. Collaboration, "Search for pair production of boosted higgs bosons via vector-boson fusion in the $b\bar{b}b\bar{b}$ final state using pp collisions at $\sqrt{s} = 13$ tev with the atlas detector", Physics Letters B **858**, 139007 (2024).

[8] C. Collaboration, "Search for higgs boson pair production in the four-b-quark final state in proton–proton collisions at $\sqrt{s} = 13$ tev", Physical Review Letters **129**, 081802 (2022).

[9] C. Collaboration, "Search for nonresonant pair production of highly energetic higgs bosons decaying to bottom quarks", Physical Review Letters **131**, 10.1103/PhysRevLett.131.041803 (2023).

[10] C. Collaboration, "Search for resonant and nonresonant higgs boson pair production in the $b\bar{b}\,\ell\nu\ell\nu$ final state in proton-proton collisions at $\sqrt{s} = 13$ tev", Journal of High Energy Physics **2018**, 10.1007/JHEP01(2018)054 (2018).

[11] C. Collaboration, "Search for resonant pair production of higgs bosons in the $bbZZ$ channel in proton-proton collisions at $\sqrt{s} = 13$ tev", Physical Review D **102**, 10.1103/PhysRevD.102.032003 (2020).

[12] A. Collaboration, "Search for higgs boson pair production in the $b\bar{b}WW^*$ decay mode at $\sqrt{s} = 13$ tev with the atlas detector", Journal of High Energy Physics **2019**, 10.1007/JHEP04(2019)092 (2019).

[13] A. Collaboration, "Search for non-resonant higgs boson pair production in the $bb\,\ell\nu\ell\nu$ final state with the atlas detector in pp collisions at $\sqrt{s} = 13$ tev", Physics Letters B **801**, 135145 (2020).

[14] A. Collaboration, "Search for resonant and nonresonant higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ decay channel in pp collisions at $\sqrt{s} = 13$ tev with the atlas detector", Physical Review Letters **121**, 10.1103/PhysRevLett.121.191801 (2018).

[15] A. Collaboration, "Search for resonant and non-resonant higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ decay channel using 13 tev pp collision data from the atlas detector", Journal of High Energy Physics **2023**, 10.1007/JHEP07(2023)040 (2023).

[16] A. Collaboration, "Search for the nonresonant production of higgs boson pairs via gluon fusion and vector-boson fusion in the $b\bar{b}\tau^+\tau^-$ final state in proton-proton collisions at $\sqrt{s} = 13$ tev with the atlas detector", Physical Review D **110**, 10.1103/PhysRevD.110.032012 (2024).

[17] C. Collaboration, "Search for higgs boson pair production in events with two bottom quarks and two tau leptons in proton–proton collisions at $\sqrt{s} = 13$ tev", Physics Letters B **778**, 101–127 (2018).

[18] A. Collaboration, "Search for higgs boson pair production in the $\gamma\gamma b\bar{b}$ final state with 13 tev pp collision data collected by the atlas experiment", Journal of High Energy Physics **2018**, 10.1007/JHEP11(2018)040 (2018).

[19] A. Collaboration, "Search for higgs boson pair production in the two bottom quarks plus two photons final state in pp collisions at $\sqrt{s} = 13$ tev with the atlas detector", Physical Review D **106**, 10.1103/PhysRevD.106.052001 (2022).

[20] C. Collaboration, "Search for nonresonant higgs boson pair production in final states with two bottom quarks and two photons in proton–proton collisions at $\sqrt{s} = 13$ tev", Journal of High Energy Physics **2021**, 10.1007/JHEP03(2021)257 (2021).

[21] ATLAS Collaboration, *ATLAS inner detector : Technical Design Report, 1*, Technical design report. ATLAS (CERN, Geneva, 1997).

[22] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Strip Detector*, tech. rep. (CERN, Geneva, 2017).

[23] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Pixel Detector*, tech. rep. (CERN, Geneva, 2017).

[24] C. Grefe, "Detector Optimization Studies and Light Higgs Decay into Muons at CLIC", (2014).

[25] "Review of particle physics, 2024-2025", **110**, 10.1103/PhysRevD.110.030001 (2024).

[26] M. Thomson, *Modern Particle Physics* (Cambridge University Press, 2013).

[27] T. Lancaster and S. J. Blundell, *Quantum Field Theory for the Gifted Amateur* (Oxford University Press, Oxford, 2014).

[28] F. Mandl and G. Shaw, *Quantum Field Theory*, 2nd ed. (John Wiley & Sons, Chichester; Hoboken, NJ, 2010).

[29] S. Weinberg, *The Quantum Theory of Fields, Vol. 2: Modern Applications* (Cambridge University Press, Cambridge, UK, 1996).

[30] E. Fermi, "Versuch einer Theorie der β-Strahlen. I.", Zeitschrift für Physik **88**, 161–177 (1934).

[31] I. J. R. Aitchison and A. J. G. Hey, *Gauge Theories in Particle Physics: A Practical Introduction, Volume 2 – Non-Abelian Gauge Theories*, 4th (CRC Press, 2013).

[32] S. Weinberg, "A model of leptons", Phys. Rev. Lett. **19**, 1264–1266 (1967).

[33] W. J. Stirling, "QCD and collider physics", (1991).

[34] CERN, *Cern yellow reports: monographs, vol 2 (2017): handbook of lhc higgs cross sections: 4. deciphering the nature of the higgs sector*, en, 2017.

[35] Particle Data Group, "Status of Higgs Boson Physics", Phys. Rev. Lett. (2024).

[36] P. Bolzoni et al., "Vector boson fusion at next-to-next-to-leading order in qcd: standard model higgs boson and beyond", Phys. Rev. D **85**, 035002 (2012).

[37] S. D. Drell and T.-M. Yan, "Massive lepton-pair production in hadron-hadron collisions at high energies", Phys. Rev. Lett. **25**, 316–320 (1970).

[38] S. L. Glashow, D. V. Nanopoulos, and A. Yildiz, "Associated production of higgs bosons and z particles", Phys. Rev. D **18**, 1724–1727 (1978).

[39] T. Han and S. Willenbrock, "QCD correction to the pp $\to$ WH and ZH total cross sections", Physics Letters B **273**, 167–172 (1991).

[40] H. Baer, B. Bailey, and J. F. Owens, "$O(\alpha_s)$ monte carlo approach to W+ higgs-boson associated production at hadron supercolliders", Phys. Rev. D **47**, 2730–2734 (1993).

[41] J. Ohnemus and W. J. Stirling, "Order-α_s corrections to the differential cross section for the WH intermediate-mass higgs-boson signal", Phys. Rev. D **47**, 2722–2729 (1993).

[42] A. Stange, W. Marciano, and S. Willenbrock, "Associated production of higgs and weak bosons, with h→bb¯, at hadron colliders", Physical Review D **50**, 4491–4498 (1994).

[43] A. Stange, W. Marciano, and S. Willenbrock, "Higgs bosons at the fermilab tevatron", Physical Review D **49**, 1354–1362 (1994).

[44] J. M. Butterworth et al., "Jet substructure as a new higgs-search channel at the large hadron collider", Physical Review Letters **100**, 10.1103/physrevlett.100.242001 (2008).

[45] R. Raitio and W. W. Wada, "Higgs-boson production at large transverse momentum in quantum chromodynamics", Phys. Rev. D **19**, 941–944 (1979).

[46] Z. Kunszt, "Associated production of heavy higgs boson with top quarks", Nuclear Physics B **247**, 339–359 (1984).

[47] J. N. Ng and P. Zakarauskas, "Qcd-parton calculation of conjoined production of higgs bosons and heavy flavors in $p\overline{p}$ collisions", Phys. Rev. D **29**, 876–886 (1984).

[48] J. Gunion, "Associated top-anti-top-higgs production as a large source of wh events: implications for higgs detection in the l final state", Physics Letters B **261**, 510–517 (1991).

[49] W. J. Marciano and F. E. Paige, "Associated production of higgs bosons with tt¯ pairs", Phys. Rev. Lett. **66**, 2433–2435 (1991).

[50] W. Beenakker et al., "Higgs radiation off top quarks at the tevatron and the lhc", Physical Review Letters **87**, 201805 (2001).

[51] L. Reina and S. Dawson, "Next-to-leading order results for $t\bar{t}h$ production at the tevatron", Physical Review Letters **87**, 201804 (2001).

[52] L. Reina, S. Dawson, and D. Wackeroth, "Qcd corrections to associated $t\bar{t}h$ production at the fermilab tevatron", Physical Review D **65**, 053017 (2002).

[53] S. Dawson et al., "Next-to-leading order qcd corrections to $pp \to t\bar{t}h$ at the cern large hadron collider", Physical Review D **67**, 071503 (2003).

[54] W. Beenakker et al., "Nlo qcd corrections to tH production in hadron collisions", Nuclear Physics B **653**, 151–203 (2003).

[55] S. Dawson et al., "Associated higgs boson production with top quarks at the cern large hadron collider: nlo qcd corrections", Physical Review D **68**, 034022 (2003).

[56] Y. Zhang et al., "Qcd nlo and ew nlo corrections to $t\bar{t}H$ production with top quark decays at hadron colliders", Physics Letters B **738**, 1–5 (2014).

[57] S. Frixione et al., "Weak corrections to higgs hadroproduction in association with a top-quark pair", Journal of High Energy Physics **2014**, 065 (2014).

[58] S. Frixione et al., *Electroweak and qcd corrections to top-pair hadroproduction in association with heavy bosons*, 2015.

[59] ATLAS Collaboration, "Search for the $H \to b\bar{b}$ decay of the standard model higgs boson in associated $(W/Z)H$ production with the atlas detector", Journal of High Energy Physics **2015**, 69 (2015).

[60] ATLAS and CMS Collaborations, "Measurements of the higgs boson production and decay rates and constraints on its couplings from a combined atlas and cms analysis of the lhc pp collision data at $\sqrt{s} = 7$ and 8 tev", Journal of High Energy Physics **2016**, 45 (2016).

[61] ATLAS Collaboration, "Evidence for the $H \to b\bar{b}$ decay with the atlas detector", Journal of High Energy Physics **2017**, 24 (2017).

[62] CMS Collaboration, "Evidence for the higgs boson decay to a bottom quark–antiquark pair", Physics Letters B **780**, 501–532 (2018).

[63] ATLAS Collaboration, "Evidence for the higgs-boson yukawa coupling to tau leptons with the atlas detector", Journal of High Energy Physics **2015**, 117 (2015).

[64] ATLAS Collaboration, "Test of cp invariance in vector-boson fusion production of the higgs boson using the optimal observable method in the ditau decay channel with the atlas detector", The European Physical Journal C **76**, 656 (2016).

[65] ATLAS Collaboration, "Search for lepton flavour violating decays of the higgs boson to $\mu\tau$ and $e\tau$ in proton-proton collisions at $\sqrt{s} = 13$ tev", Journal of High Energy Physics **2018**, 1 (2018).

[66] M. Grazzini et al., "Higgs boson pair production at nnlo with top quark mass effects", Journal of High Energy Physics **2018**, 59 (2018).

[67] J. Baglio et al., "$gg \to HH$: combined uncertainties", Physical Review D **103**, 056002 (2021).

[68] C. Collaboration, "Combined measurements of higgs boson couplings in proton–proton collisions at $\sqrt{s} = 13$ tev", The European Physical Journal C **79**, 421 (2019).

[69] A. Cortes Gonzalez (ATLAS), "Non-resonant di-Higgs search measurements, ATLAS", (2024).

[70] O. S. Brüning et al., *LHC Design Report*, CERN Yellow Reports: Monographs (CERN, Geneva, 2004).

[71] R. Assmann, M. Lamont, and S. Myers, "A brief history of the LEP collider", Nucl. Phys. B Proc. Suppl. **109**, edited by F. L. Navarria, M. Paganoni, and P. G. Pelfer, 17–31 (2002).

[72] E. Lopienska, *CERN graphics 2022*, CERN Document Server, `https://cds.cern.ch/images/CERN-GRAPHICS-2022-001-1`.

[73] ALICE Collaboration, "The ALICE experiment at the CERN LHC", Journal of Instrumentation **3**, S08002 (2008).

[74] LHCb Collaboration, "The LHCb Detector at the LHC", Journal of Instrumentation **3**, Also published by CERN Geneva in 2010, S08005 (2008).

[75] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider: A Description of the Detector Configuration for Run 3*, 2023.

[76] CMS Collaboration, "The CMS Experiment at the CERN LHC", Journal of Instrumentation **3**, S08004 (2008).

[77] CERN, *Accelerator complex*, https://www.home.cern/science/accelerators/accelerator-complex, n.d.

[78] W. Herr and B. Muratori, "Concept of luminosity", in CERN Accelerator School and DESY Zeuthen: Accelerator Physics (Sept. 2003), pp. 361–377.

[79] ATLAS Collaboration, *ATLAS A 25-Year Insider Story of the LHC Experiment: Volumen 30* (CERN, July 2019).

[80] G. 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", Machine Learning: Science and Technology **1**, 10.1088/2632-2153/ab983a (2020).

[81] ATLAS Collaboration, "The atlas experiment at the cern large hadron collider: a description of the detector configuration for run-3", Journal of Instrumentation **19**, P05063 (2024).

[82] ATLAS Collaboration, "ATLAS pixel detector electronics and sensors", JINST **3**, P07007 (2008).

[83] ATLAS SCT Collaboration, "The ATLAS semiconductor tracker (SCT)", Nucl. Instrum. Meth. A **541**, edited by T. Ohsugi, Y. Unno, and H. F. W. Sadrozinski, 10.1016/j.nima.2005.01.043 (2005).

[84] ATLAS TRT Collaboration, "The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: Design and performance", JINST **3**, P02013 (2008).

[85] ATLAS Collaboration, *ATLAS calorimeter performance*, Technical design report. ATLAS (CERN, Geneva, 1996).

[86] ATLAS Collaboration, *ATLAS muon spectrometer*, tech. rep. (CERN/ ATLAS Experiment, 1997).

[87] ATLAS Collaboration, *New Small Wheel Technical Design Report*, tech. rep., ATLAS New Small Wheel Technical Design Report (CERN / ATLAS Experiment, 2013).

[88] CERN HL-LHC Industry Project, *Project schedule*, https://project-hl-lhc-industry.web.cern.ch/content/project-schedule, n.d.

[89] ATLAS Collaboration, "The ATLAS trigger system for LHC Run 3 and trigger performance in 2022", Journal of Instrumentation **19**, 10.1088/1748-0221/19/06/p06029 (2024).

[90] ATLAS Collaboration, *ATLAS level-1 trigger*, Technical design report. ATLAS (CERN, Geneva, 1998).

[91] ATLAS Collaboration, *Phase-1 ATLAS Level-1 Trigger in Run-3*, tech. rep. (CERN, Geneva, 2024).

[92] ATLAS Collaboration, "The upgrade of the atlas first-level calorimeter trigger", Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **824**, 374–378 (2016).

[93] ATLAS Collaboration, *The raw event format in the ATLAS Trigger and DAQ*, tech. rep., Revised version number 5 submitted on 2016-11-03 11:47 (CERN, Geneva, 1998).

[94] 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**, 639 (2019).

[95] ATLAS Collaboration, "Muon reconstruction performance of the atlas detector in proton–proton collision data at $\sqrt{s} = 13$ tev", Eur. Phys. J. C **76**, 292 (2016).

[96] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", Journal of High Energy Physics **2008**, 063 (2008).

[97] ATLAS Collaboration, *Optimisation and performance studies of the atlas b-tagging algorithms for the 2017–18 lhc run*, tech. rep. (CERN, Geneva, 2017).

[98] ATLAS Collaboration, *Deep sets based neural networks for impact parameter flavour tagging in atlas*, tech. rep. (CERN, Geneva, 2020).

[99] ATLAS Collaboration, *Transformer neural networks for identifying boosted higgs bosons decaying into $b\bar{b}$ and $c\bar{c}$ in atlas*, tech. rep. (CERN, Geneva, 2023).

[100] ATLAS Collaboration, "Transforming jet flavour tagging at atlas", arXiv:2505.19689 [hep-ex] (2025).

[101] A. Vaswani et al., "Attention is all you need", arXiv:1706.03762 [cs.CL] (2017).

[102] ATLAS Collaboration, *Reconstruction, identification, and calibration of hadronically decaying tau leptons with the atlas detector for the lhc run 3 and reprocessed run 2 data*, tech. rep. (CERN, Geneva, 2022).

[103] ATLAS Collaboration, "Reconstruction and identification of boosted di-τ systems in a search for higgs boson pairs using 13 tev proton–proton collision data in atlas", JHEP **11**, 163 (2020).

[104] ATLAS Collaboration (Boosted Di-tau Analysis Team), *Measurement of the boosted* $H \to \tau^+_{\text{had}}\tau^-_{\text{had}}$ *cross-section with the atlas detector using full run 2 and partial run 3 datasets*, tech. rep. (CERN, Geneva, 2025).

[105] ATLAS Collaboration, *Identification of boosted higgs bosons decaying into a pair of collimated* τ *leptons decaying hadronically with transformer neural networks in atlas*, tech. rep. (CERN, Geneva, 2025).

[106] ATLAS Collaboration, "Optimisation of large-radius jet reconstruction for the atlas detector in 13 tev proton–proton collisions", Eur. Phys. J. C **81**, 334 (2021).

[107] D. Torres et al., *ITk Strips Bytestream Converter*, (Sep 2024) https://indico.cern.ch/event/1452496/contributions/6114173/attachments/2922347/5129396/ITkStripsByteStreamCnv.pdf.

[108] P. Keener, J. J. John, and J. Kaplon, *ATLAS ITk Electronics Specification: ABCStarV1*, tech. rep. (CERN / ATLAS Experiment, Feb. 2021).

[109] P. Keener, M.Newcommer, and M. Warren, *ATLAS ITk Electronics Specification: HCCStarV1*, tech. rep. (CERN / ATLAS Experiment, July 2022).

[110] Michael Hank, *HCC Chip Identifiers*, (May 2025) https://indico.cern.ch/event/1540708/contributions/6493793/subcontributions/547253/attachments/3064602/5420439/HCC%20Chip%20Identifiers.pdf.

[111] C. Bernius, *The ATLAS Trigger Menu: Design and Performance*, tech. rep. (CERN, Geneva, Switzerland, 2012).