GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN IN PUBLICA COMMODA SEIT 1737

Fakultät für
Physik

Master's Thesis

High Level Trigger Optimization Studies in the ATLAS search for Higgs Boson Pair Production in the HH $\rightarrow b\bar{b}\tau^+_{\text{had}}\tau^-_{\text{had}}$ channel

Optimierungsstudien für den High-Level-Trigger in der Suche nach Higgs-Boson-Paarproduktion im Kanal HH $\rightarrow b\bar{b}\tau^+_{\text{had}}\tau^-_{\text{had}}$ bei ATLAS

prepared by

Athul Dev Sudhakar Ponnu
Trivandrum, India

at the II. Institute of Physics

Thesis number : II.Physik-UniGö-MSc-2023/03

Thesis period : 08-12-2022 until 08-12-2023

First referee : Prof. Dr. Stanley Lai

Second referee : Prof. Dr. Arnulf Quadt

Abstract

The Standard Model of particle physics predicts the existence of the trilinear Higgs self coupling vertex. This can be studied through the Higgs boson pair production process, which is yet to be observed. The primary focus of this analysis is to optimise High Level Triggers (HLT) at ATLAS, for the Higgs boson pair production channel, with $b\bar{b}\tau^+_{had}\tau^-_{had}$ final states. The specific goal is to improve signal acceptance and background rejection in the trigger chains relevant for the study of the HH→ $b\bar{b}\tau^+_{had}\tau^-_{had}$ channel. The trigger selection criteria for two different di-τ chains and the $b + \tau$ chain have been fine-tuned to increase signal selection efficiency while maintaining or, in some cases, decreasing background rates.

Keywords: HDBS, Higgs boson pair production, Trilinear Higgs self-coupling, bbtautau, HLT, Trigger emulation, Trigger optimisation, ATLAS detector, LHC Run-3.

Contents

Chapter 1

Introduction

“नासदासीन्नो सदासीत्तदानीं नासीद्रजो नो व्योमा परो यत् । न मृत्युरासीदमृतं न तर्हि न रात्र्या अह्न आसीत्प्रकेतः ।

इयं विसृष्टिर्यत आबभूव यदि वा दधे यदि वा न । यो अस्याध्यक्षः परमे व्योमन्त्सो अङ्ग वेद यदि वा न वेद”

Not non-existent, nor existent then, there was no realm of air, no sky beyond it. What covered in, and where, and what protected? At first there was only darkness wrapped in darkness. All this was only unillumined cosmic water. That One which came to be, enclosed in nothing, arose at last, born of the power of heat.

But, after all, who knows, and who can say whence it all came, and how creation happened? The gods themselves are later than creation, so who knows truly whence it has arisen? Whence all creation had its origin, the creator, whether he fashioned it or whether he did not, the creator, who surveys it all from highest heaven, he knows - or maybe even he does not know.

- Rigveda (Nāsadīya Sūkta, 10:129) dated to 1500 BCE

Over millennia, humanity’s relentless curiosity about the origins of existence has echoed through cultures and civilisations. This intense curiosity has inspired scientific investigation, leading to groundbreaking discoveries that uncover the fundamental structure of reality. One such notable discovery was that of the electron [1], which unveiled the intricate substructure within atoms, leading to the development of quantum mechanics and its offspring, such as quantum field theory (QFT). Employing the principles of QFT, the Standard Model (SM) [2–4], a cornerstone of modern physics, encapsulates our knowledge of elementary particles and the fundamental forces that govern them. While it excels at making accurate predictions, this model is not without its imperfections, as it does not offer an explanation for perplexing mysteries such as dark matter [5], dark energy [6], and the observed matter-antimatter imbalance in the universe [7].

The SM received a significant update with the momentous discovery of the Higgs boson in 2012, a milestone in the field of particle physics. This breakthrough confirmed the existence of this elusive boson [8, 9], which plays a crucial role in imparting mass to fundamental particles [10–12]. After this groundbreaking achievement, the scientific community embarked on a quest to unravel the intrinsic properties and intricate interactions of this enigmatic particle. One particularly captivating avenue of exploration is the investigation of Higgs-boson pair production. This opens up an interesting way to probe the physics of Higgs boson self-coupling, shedding light on the mechanism responsible for electroweak symmetry breaking (EWSB) [10–12]. Among the different possibilities for pairs of Higgs bosons to decay, the HH$\rightarrow$ $b\bar{b}\tau^+\tau^-$ channel plays a significant role. This decay mode combines the two most prevalent Higgs boson decay channels, involving the $b\bar{b}$ and $\tau^+\tau^-$ final states.

The theory of Higgs and EWSB cannot be validated without the means to experimentally test it. The Large Hadron Collider (LHC) [13], a high energy proton-proton collider setup has been established for precisely this purpose. The ATLAS detector, has been carefully constructed to determine the properties of particles generated during high-energy proton-proton collisions at LHC [14]. By making use of the exhaustive datasets generated by these high-energy collisions, the ATLAS collaboration strives to extract invaluable revelations regarding the Higgs boson pair production and the underlying mechanisms governing the fundamental fabric of the universe [15–18].

Successful detection and precise measurements of these exceedingly rare processes in harsh environments, such as those of the ATLAS detector, require the use of highly sophisticated trigger strategies. The LHC is capable of operating at an unprecedented centre of mass energy($\sqrt{s}$) of 13.6 TeV with approximately billions of collisions per second, producing heaps of data that must be processed in real time. For this purpose, the ATLAS experiment relies upon a two-tiered triggering system to discriminate and select events of scientific interest from the relentless torrent of data [19].

This thesis discusses the optimisation of various trigger chains employed in the ATLAS search for Higgs boson pair production within the HH$\rightarrow$ $b\bar{b}\tau^+_{had}\tau^-_{had}$ channel. Different trigger chains with a wide range of trigger configurations were emulated and analysed using Monte Carlo (MC) simulations and Enhanced Bias (EB) datasets, using diverse data analysis techniques, aimed at determining the optimal trigger strategy. In the end, the event selection efficiency is enhanced while simultaneously curbing background rates. By increasing the trigger efficiency, this research aims to contribute to the pursuit of a more profound understanding of the Higgs potential and the nature of electroweak symmetry breaking.

Chapter 2

Theory and Phenomenology

This chapter provides a concise overview of the history of particle physics, with a primary focus on the Standard Model (SM). It explores the underlying symmetries of the SM and introduces the concept of electroweak symmetry and the Higgs field in subsequent sections. Furthermore, the discussion will move towards the properties of the Higgs boson, with a specific focus on its trilinear self-coupling, as this is crucial for understanding the context of this thesis.

2.1 The Standard Model of Particle Physics

The Standard Model of particle physics represents a comprehensive theoretical framework that has evolved over the course of more than a century of scientific inquiry and experimentation. It stands as the most precise theoretical model available for describing the fundamental constituents of matter and its associated interactions, and has been extensively tested and confirmed through experiments at various energy scales [20, 21].

The journey to the SM began with the discovery of the electron in 1897 [1], and subsequent breakthroughs, such as Einstein's theory of quantised electromagnetic radiation [22] and the emergence of Quantum Mechanics in the 1920s. Key developments included the discovery of the proton [23], the prediction and confirmation of antimatter [24], and Fermi's description of beta decay [25], laying the foundation for the theory of weak interactions. This was a crucial step in describing particle creation, annihilation, and decays within the Quantum Field Theory (QFT) framework. However, early QFTs faced challenges related to divergences in perturbative corrections. It was only in 1949 that the method of renormalisation was introduced, leading to the establishment of Quantum Electrodynamics (QED) [26].

The late 1950s witnessed a surge of discoveries of composite particles, and the concept

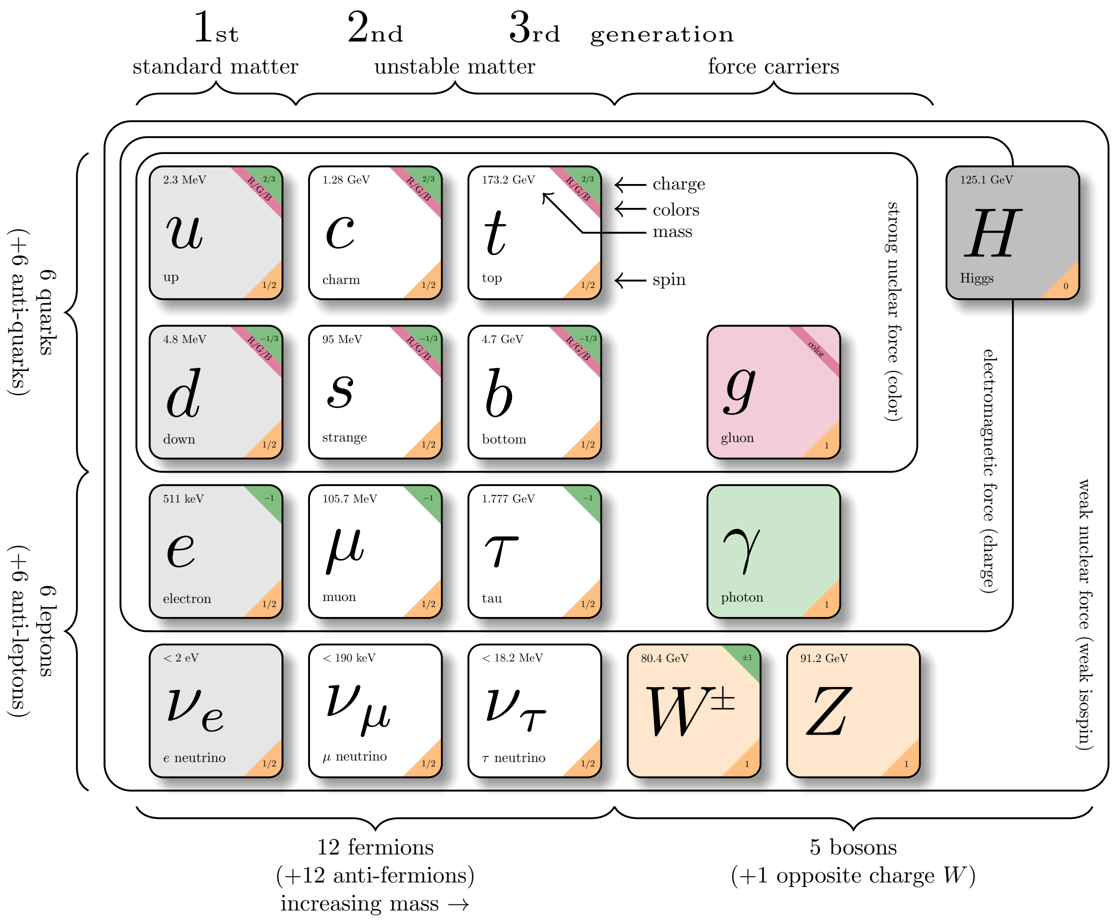


Figure 2.1: Fundamental particles of the standard model of particle physics. Adapted from graphic by © Carsten Burgard, CC BY 2.5.

of quarks was proposed by Gell-Mann [27]. They eventually were experimentally verified by the e-p scattering experiments at the Stanford Linear Accelerator Centre (SLAC) in 1968 [28]. To explain the strong interaction that exists among quarks, Quantum Chromodynamics (QCD) was developed, providing an explanation for this force as being mediated by gluons.

In 1983, a significant milestone was reached in CERN when the W and Z bosons, which serve as carriers of the weak force, were successfully discovered and their existence was confirmed [29, 30]. This momentous achievement provided crucial experimental support for the predictions of the electroweak theory, further confirming the foundational principles of the SM. The last piece of the SM puzzle fell into place in 2012 with the momentous discovery of the Higgs boson [8, 9]. These milestones transformed the SM into the comprehensive model it is today. The constituents of the Standard Model of particle physics are shown in Figure 2.1.

The fermionic part of the SM consists of quarks that interact through electromagnetic, weak, and strong forces, and leptons which interact through weak and electromagnetic forces. The left-chiral quarks and leptons form a $SU(2)$ doublet of the weak interaction, such as u and d, or ν_e and e. Fermions in the first generation are supposed to form almost all materials in the universe, since stable atoms or nuclei are composed of them.

The SM is expressed mathematically as a gauge QFT featuring global Poincaré symmetry, alongside local $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ gauge symmetry, which accounts for the various types of interactions and gives rise to bosonic force carriers. The gauge group $SU(3)_C$ corresponds to the strong interaction. It has 8 generators which correspond to the 8 types of gluons, the force carriers of the strong interaction. The electromagnetic force and the weak force are combined into the electroweak force in the $SU(2)_L \otimes U(1)_Y$ gauge theory, and it is mediated by massive weak bosons ($W^{\pm}$, Z^0) and a massless photon (γ). In order for the Lagrangian to be invariant under local transformation, it is necessary that the gauge bosons be massless, which contradicts the experimental results pointing towards a massive $W^{\pm}$ and Z^0 boson. To resolve this, the Higgs field was introduced, which will be discussed in the next section.

2.2 The Higgs Field and Symmetry Breaking

For the electroweak sector, the SM Lagrangian can be written as follows:

$$\mathcal{L}_{SM} = i\overline{\Psi_L}\gamma^{\mu}\not{D}_{\mu}\Psi_L + i\overline{\psi_R}\gamma^{\mu}\not{D}_{\mu}\psi_R - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{4}G^{i}_{\mu\nu}G^{\mu\nu i} \tag{2.1}$$

Here Ψ_L and ψ_R are the left chiral fermion doublets and the right chiral fermion singlets of the form

$$\Psi_L = \begin{pmatrix} \nu_l \\ \ell \end{pmatrix}_L, \begin{pmatrix} u \\ d \end{pmatrix}_L \quad \text{and} \quad \psi_R = \begin{pmatrix} \ell \end{pmatrix}_R, \begin{pmatrix} \nu \end{pmatrix}_R, \begin{pmatrix} u \end{pmatrix}_R, \begin{pmatrix} d \end{pmatrix}_R$$

respectively. To maintain the local gauge invariance during different symmetry transformations, the kinetic terms for the vector fields B_ν and $W^i_{\mu\nu}$ are introduced with $F^{\mu\nu}$ and $G^{\mu\nu i}$. The interaction between the vector bosons and the fermion field is explained through the covariant derivatives D_μ which is given by

$$\partial_\mu \to \not{D}_\mu = \partial_\mu + igW^i_\mu + ig'Y_L B_\mu \tag{2.2}$$

for the left chiral fermions and

$$\partial_\mu \to \not{D}_\mu = \partial_\mu + ig'Y_L B_\mu \tag{2.3}$$

for the right chiral fermions. Vector bosons and fermions do not possess associated mass terms, since they spoil the local $SU(2)_L \otimes U(1)_Y$ symmetry. To deal with this, the SM incorporates the concept of spontaneous electroweak symmetry breaking, or the Higgs mechanism, originally proposed by Higgs, Brout, Englert, Guralnik, Hagen, and Kibble [10]. This mechanism is essential for generating the masses of the weak vector bosons while satisfying the requirements of renormalizability and unitarity. The Higgs mechanism involves two complex scalar fields arranged in a weak isospin doublet.

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \tag{2.4}$$

The Lagrangian for the scalar field is given by:

$$\mathcal{L}_{\text{Higgs}} = (\not{D}_\mu \phi)^\dagger (\not{D}^\mu \phi) - V(\phi), \tag{2.5}$$

where the potential $V(\phi)$, depicted in Figure 2.2, is defined as:

$$V(\phi) = \mu^2(\phi^\dagger\phi) + \lambda(\phi^\dagger\phi)^2. \tag{2.6}$$

To ensure a finite minimum, referred to as the vacuum state, the potential is set such that $\lambda > 0$. However, μ can have either a positive ($\mu^2 > 0$) or negative ($\mu^2 < 0$) value. When $\mu^2 > 0$, the potential possesses a single vacuum state at the origin, exhibiting symmetry in all directions. In contrast, selecting $\mu^2 < 0$ yields an infinite number of vacuum states located at a distance v from the origin:

$$\Phi^\dagger\Phi = -\frac{\mu^2}{\lambda} = v^2. \tag{2.7}$$

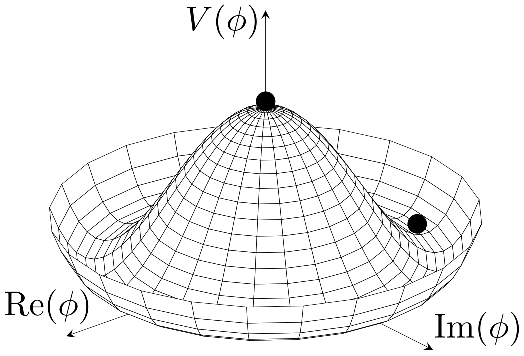


Figure 2.2: Graphical representation of the Higgs potential V(ϕ) for $\mu^2 < 0$ and $\lambda > 0$.

Once the Higgs field attains a vacuum expectation value (VEV), the symmetry is broken because the potential is no longer symmetric from any of these vacuum states. The

Higgs doublet, after attaining a vacuum expectation value and applying unitary gauge condition, can be written as

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v+h \end{pmatrix} . \tag{2.8}$$

After attaining VEV, Equation (2.6) becomes

$$\begin{aligned} \mathcal{L}_{\text{Higgs}} = & \underbrace{\frac{1}{2}\partial_\mu h \partial^\mu h - \lambda v^2 h^2}_{\text{Massive scalar}} \underbrace{-\lambda v h^3 - \frac{1}{4}\lambda h^4}_{\text{Self-interactions}} + \underbrace{\frac{1}{4} g_2^2 v^2 W_\mu^- W^{+\mu} + \frac{1}{8}(g_2^2 + g_1^2) v^2 Z_\nu Z^\nu}_{\text{Mass of gauge bosons}} \\ & \underbrace{+\frac{1}{2} g_2^2 v W_\mu^- W^{+\mu} h + \frac{1}{4}(g_2^2 + g_1^2) v Z_\nu Z^\nu h + \frac{1}{4} g_2^2 W_\mu^- W^{+\mu} h^2 + \frac{1}{8}(g_2^2 + g_1^2) Z_\nu Z^\nu h^2}_{\text{Interactions between the Higgs and gauge fields}} \end{aligned} \tag{2.9}$$

[31], where $W^\pm$, Z, and A fields are defined as

$$W_\mu^\pm = \frac{1}{\sqrt{2}}\left(W_\mu^1 \mp i W_\mu^2\right), \quad Z_\mu = \frac{g_2 W_\mu^3 - g_1 B_\mu}{\sqrt{g_2^2 + g_1^2}}, \quad A_\mu = \frac{g_1 W_\mu^3 + g_2 B_\mu}{\sqrt{g_2^2 + g_1^2}}. \tag{2.10}$$

From the quadratic field terms in the above Lagrangian, the masses of the Higgs boson and the gauge bosons can be determined to be

$$m_H = \sqrt{2\lambda v^2}, \quad m_W = \frac{gv}{\sqrt{2}}, \quad m_Z = \frac{v}{2}\sqrt{g_2^2 + g_1^2}, \quad m_A = 0 \ .$$

Thus, one can generate mass for the vector bosons without violating the local gauge symmetry. Since the Dirac mass term for the fermions violate $SU(2)_L$ symmetry, they are replaced by new mass terms, described by the Yukawa Lagrangian and it is given by

$$\begin{aligned} \mathcal{L}_Y \supset \ & g_f \left(\overline{\Psi_L}\Phi\psi_R \ + \overline{\psi_R}\Phi^\dagger\Psi_L\right) \\ & + g_f\left(\overline{\Psi_L}\Phi_c\psi_R \ + \overline{\psi_R}\Phi_c^\dagger\Psi_L\right), \end{aligned} \tag{2.11}$$

where Φ and Φ_c are the scalar Higgs doublet and its conjugate in its unitary gauge, Ψ_L and ψ_R are the left fermion doublets and the right fermion singlets of the $SU(2)_L$ group. For any fermion f, after symmetry breaking, the Yukawa Lagrangian becomes

$$\mathcal{L}_{Y_f} \supset \frac{g_f}{\sqrt{2}} v\left(\bar{f}_L f_R + \bar{f}_R f_L\right) + \frac{g_f}{2} hv\left(\bar{f}_L f_R + \bar{f}_R f_L\right), \tag{2.12}$$

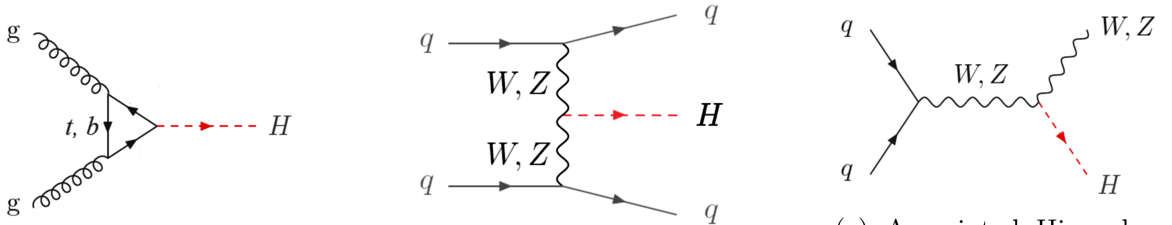
from which the mass of the fermion can be determined to be

$$m_f = \frac{g_f}{\sqrt{2}} v. \tag{2.13}$$

2.3 Higgs Boson Characteristics and Behavior

The Higgs boson was discovered in 2012 at the LHC through experiments conducted by ATLAS and CMS [8, 9]. This groundbreaking discovery was achieved by observing its decay through various channels, which includes $Z^0Z^0 \rightarrow 4\ell, W^+W^- \rightarrow \ell^+\ell^-\nu\bar{\nu}$ and $\gamma\gamma$. The Higgs boson possesses a mass of 125.10 ± 0.42 GeV [9, 32] and is characterised as a spin-0, CP-even particle with a very short lifetime of $\tau_{\text{Higgs}} \approx 10^{-22}$ seconds [33].

For the Higgs boson production at LHC, three production modes dominate, and they are represented in the Feynman diagrams in Figure 2.3. The most prominent one is gluon-gluon fusion (ggF), with a Higgs boson production cross section of $\sigma_{ggF} = 54.7$ pb at 14 TeV [33]. There are also other less prevalent modes like the vector boson fusion (VBF) mode and associated Higgs production.



(a) Gluon gluon fusion (b) Vector Boson Fusion (c) Associated Higgs boson production

Figure 2.3: Leading-order Feynman diagrams for Higgs boson production at LHC in the order of their cross-section abundance.

The Higgs boson, being an unstable particle, undergoes rapid decay shortly after its production. Although it can decay to all massive SM particles, it preferentially couples to heavier particles. It can also decay into massless particles such as photons and gluons through loops involving massive particles such as top or bottom quarks and $W^\pm$ bosons. The branching ratios of these different decay modes of the SM Higgs boson are listed in Table 2.1. Due to the high mass of the b quark and the τ lepton in comparison to other fermions, they have the highest Higgs boson branching ratio.

In addition to the Higgs boson's interactions with vector bosons and fermions, the scalar Lagrangian in Equation (2.9) also contains trilinear and quartic self-coupling terms, which correspond to the Feynman vertices in Figure 2.4. To produce two real Higgs bosons via trilinear self-coupling, the mediating Higgs boson must exist as a virtual particle with mass $m_{H^*} \geq 2m_H$. As the mass of virtual particles lacks a precise value and spans a wide range, attempts to reconstruct this mass do not yield a distinct resonance peak. Consequently, this production method is termed non-resonant. However, experimental measurements of the Higgs boson's self-interaction have not been achieved thus far [34].

Decay Channel	Branching Ratio	Relative Uncertainty
$\mathrm{H} \rightarrow b\bar{b}$	5.82×10^{-1}	$^{+1.2\%}_{-1.3\%}$
$\mathrm{H} \rightarrow \mathrm{W}^{+}\mathrm{W}^{-}$	2.14×10^{-1}	$\pm 1.5\%$
$\mathrm{H} \rightarrow \tau^{+}\tau^{-}$	6.27×10^{-2}	$\pm 1.6\%$
$\mathrm{H} \rightarrow \mathrm{c}\bar{c}$	2.89×10^{-2}	$^{+5.5\%}_{-2.0\%}$
$\mathrm{H} \rightarrow \mathrm{ZZ}$	2.62×10^{-2}	$\pm 1.5\%$
$\mathrm{H} \rightarrow \gamma\gamma$	2.27×10^{-3}	$\pm 2.1\%$
$\mathrm{H} \rightarrow \mathrm{Z}\gamma$	1.53×10^{-3}	$\pm 5.8\%$
$\mathrm{H} \rightarrow \mu^{+}\mu^{-}$	2.18×10^{-4}	$\pm 1.7\%$

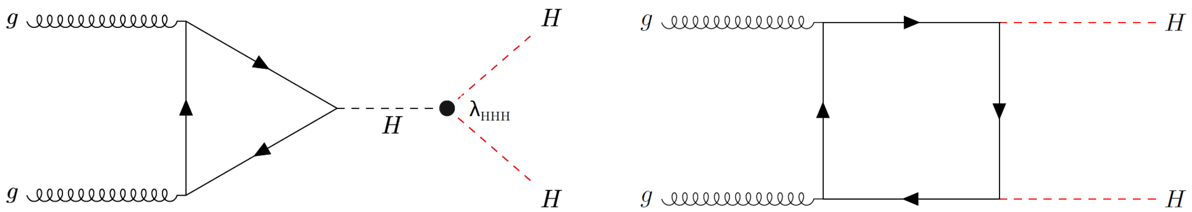
Table 2.1: Branching ratio of Higgs boson with a mass of $m_H = 125$ GeV, and their relative uncertainties [33].

(a) Trilinear self coupling (b) Quartic self coupling

Figure 2.4: Feynman vertices depicting the self-couplings of the SM Higgs boson.

At the LHC, an extensive physics program is dedicated to exploring Higgs boson pair (HH) production, which serves as an avenue to probe the Higgs boson's trilinear self-coupling, denoted as λ_{HHH}. In the SM, the self-coupling of the Higgs boson and the structure of the scalar Higgs field potential are entirely determined by the Higgs boson's mass and VEV. Any deviation from the predicted shape of the scalar potential could have profound implications on our understanding of baryogenesis [35], inflation[36], stability of the electroweak vacuum [37] and ultimately the universe's destiny. Therefore, the measurement of the Higgs boson's trilinear self-coupling holds particular significance as it provides insights into the characterization of the Higgs field potential.

Investigation of the Higgs boson trilinear self-coupling at the LHC involves an analysis of gluon-gluon fusion (ggF) and vector boson fusion (VBF) channels, depicted in Figures 2.5 and 2.6 respectively, where ggF accounts for approximately 90% of the Higgs boson pair production [33]. The expected cross section for ggF at a collision energy of 14 TeV and a Higgs boson mass of 125 GeV is $43.7^{+5.1}_{-6.0}$ fb [38] at NNLL, which is three orders of magnitude lower than the cross section for single-Higgs boson production. The process of interest is shown in Figure 2.5a, which interferes destructively with the process illustrated



(a) The ggF signal process.

(b) The ggF background process.

Figure 2.5: Leading-order Feynman diagrams for ggF non-resonant production of Higgs boson pairs. (a) illustrates the signal process, proportional to the product of the top-quark Yukawa coupling and the Higgs boson self-coupling. (b) is the background process, which is proportional to the square of the top-quark Yukawa coupling.

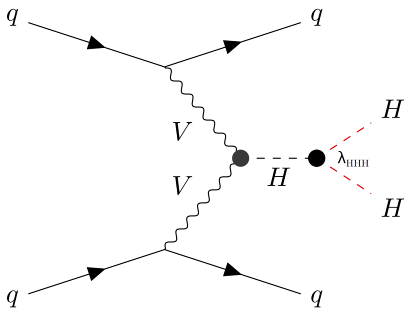


Figure 2.6: Leading-order Feynman diagram for VBF non-resonant production of the Higgs boson pairs. The trilinear self coupling is denoted by λ_{HHH}.

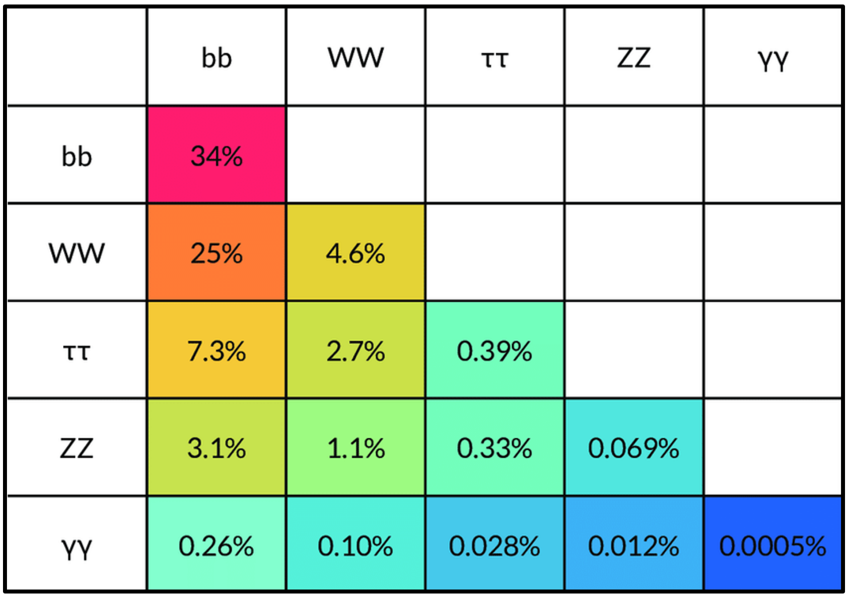
in Figure 2.5b, making it even harder to study the process.

Before leaving a detectable signature in the detector, the Higgs bosons decay into more stable particles. The branching ratio for the final state particles in the Higgs boson pair events is shown in Figure 2.7. Previous searches for non-resonant HH production were performed at the ATLAS and CMS in the $b\bar{b}\tau^+\tau^-$ [15], $\bar{b}b\gamma\gamma$ [17], $\bar{b}b\bar{b}b$ [16], $\bar{b}b\ell^+\nu\ell^-\bar{\nu}$ [18] decay channels. Analysis has also been performed by ATLAS in the $\bar{b}bq q\ell^+\ell^-$, $WW^*\gamma\gamma$ and WW^*WW^* decay channels, and by CMS in the $\bar{b}b\ell^+\ell^-\ell^+\ell^-$, WW^*WW^*, $W^*W^*\tau^+\tau^-$ and $\tau^+\tau^-\tau^+\tau^-$ decay channels.

Among these, the $b\bar{b}\tau^+\tau^-$, $b\bar{b}b\bar{b}$, $b\bar{b}\gamma\gamma$ channels are particularly favoured due to their high sensitivity. Using these channels in Run-2 ATLAS data, constraints were placed on the coupling modifier defined as the ratio of the Higgs boson self-coupling to its SM value $\kappa_\lambda = \lambda_{\mathrm{HHH}}/\lambda_{\mathrm{HHH}}^{\mathrm{SM}}$, within the range of $-0.6 < \kappa_\lambda < 6.6$ at a 95% confidence level [34].

2.4 Decay Modes for *b* Quarks and *τ* Leptons

Within the framework of the SM, particle decays represent transformation of one particle into another, often leading to the creation of lighter particles or fundamental constituents. At the heart of particle decay lies the principle of conservation laws, which encompass



	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 2.7: Branching ratio for the most important HH decay modes assuming SM couplings, calculated at NLO [39].

the preservation of various quantities such as energy, momentum, electric charge, fermion number, etc. These conservation principles intricately dictate the permissible decay channels for particles, guiding their transition pathways and contributing significantly to our comprehension of the underlying symmetries within the universe. In the context of this analysis, the decay modes of τ leptons and b quarks are of significance.

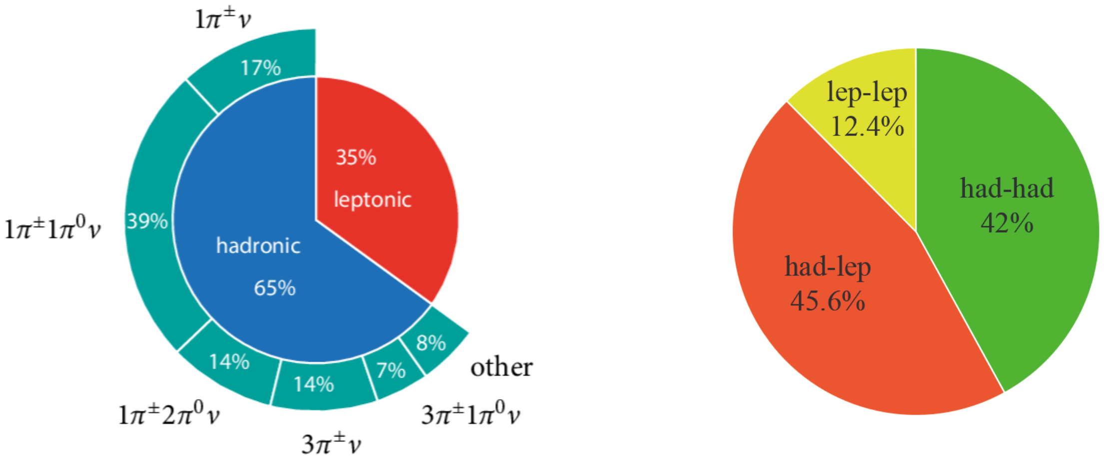
(a) Leptonic decay channel

(b) Hadronic (π^+ and π^0) decay channel

Figure 2.8: Feynman diagram depicting a potential hadronic and leptonic tau decay channel.

The tau lepton is the heaviest among the lepton family with a mass of 1776.86 ± 0.12 MeV [33], and it decays by emitting a virtual W boson, generating a ν_τ. It has a mean life time of $\tau = (2.903 \pm 0.005) \times 10^{-13}$ s . They have two types of decay modes: leptonic, as depicted in Figure 2.8a, or hadronic, an example of which is illustrated in Figure 2.8b. The corresponding branching ratio are visualised in Figure 2.9a.

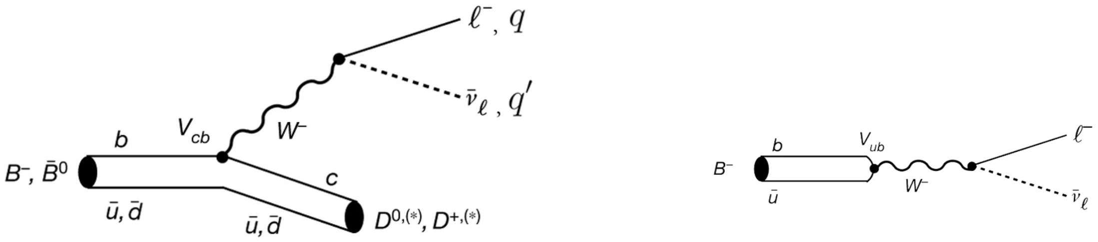
Tau leptons that undergo hadronic decay predominantly produce a mix of charged pions ($\pi^\pm$) and kaons ($K^\pm$), characterised by an odd number of charged tracks or prongs, and neutrinos, which are undetectable. These visible hadronically decaying τ lepton decay products are referred to as τ_{had}. The conservation of charge necessitates that the τ lepton decays exclusively into an odd number of charged particles, while the decay products may also contain neutral mesons such as π^0 and K^0. When we observe the decay of two τ leptons, multiple combinations of these decay processes become possible. Figure 2.9b illustrates the contributions of the different final states.



(a) Single τ branching fraction

(b) Different $\tau\tau$ final states

Figure 2.9: Figure depicting various τ and the $\tau\tau$ final states [33].



(a) B meson decaying to charm meson and lepton/quark pair.

(b) B meson decaying to purely leptonic final state.

Figure 2.10: SM Electroweak B meson decays.

The b quark, with a mass of $4.18^{+0.03}_{-0.02}$ GeV [33], hadronizes and produces jets of particles almost instantaneously with a high probability of producing a B meson in the process. The B mesons can undergo decay in several interesting ways as represented in Figure 2.10: completely leptonic, semi-leptonic with a non-prompt lepton, or fully hadronic. The B mesons have a proper lifetime of approximately 1.638 ps [33]. In the laboratory frame, B-mesons typically travel an average distance in the order of millimetres to 1 centimetre before decaying.

Chapter 3

Experimental Setup

A significant portion of this analysis relies on data collected from the ATLAS detector, a pivotal component of one of the four primary experiments conducted at CERN, which is home to the world's largest particle accelerator - the LHC. Subsequent sections in this chapter discuss the intricacies of the LHC machine, the ATLAS experiment and its detectors.

3.1 The Large Hadron Collider (LHC)

The Large Hadron Collider [13] (LHC) is a two-ring superconducting hadron accelerator and collider installed in the existing 26.7 km tunnel that was constructed between 1984 and 1989 for the electron-positron collider - LEP [40]. The LHC has a design beam energy of 7 TeV, resulting in a design collision energy of $\sqrt{s} = 14$ TeV. The LHC accommodates 2808 proton bunches, each initially containing around 1.1×10^{11} protons, spaced 25 ns apart, travelling at nearly the speed of light (c). During Run-2, which spans from 2015 to 2018, the LHC operated at $\sqrt{s} = 13$ TeV, reaching a maximum luminosity of $\mathcal{L}_{\text{peak}} = 2.14 \times 10^{34}$ cm^{-2} s^{-1} in the ATLAS experiment. As of 2023, the LHC is in its Run-3 phase operating at $\sqrt{s} = 13.6$ TeV.

The illustration depicted in Figure 3.1 shows the CERN accelerator complex, which includes the LHC and its associated accelerator chain. The proton acceleration process involves the ionisation of negative hydrogen atoms in a strong magnetic field at LINAC4 to 160 MeV. Subsequently, the hydrogen ions are stripped of their two electrons upon injection from the LINAC4 into the Proton Synchrotron Booster (PSB), thereby leaving only protons. These protons are then accelerated to 2 GeV. Further acceleration occurs within the proton-synchrotron (PS) to reach beam energies up to 26 GeV. The beam is then directed to the Super Proton Synchrotron (SPS), where particles originating from the PS

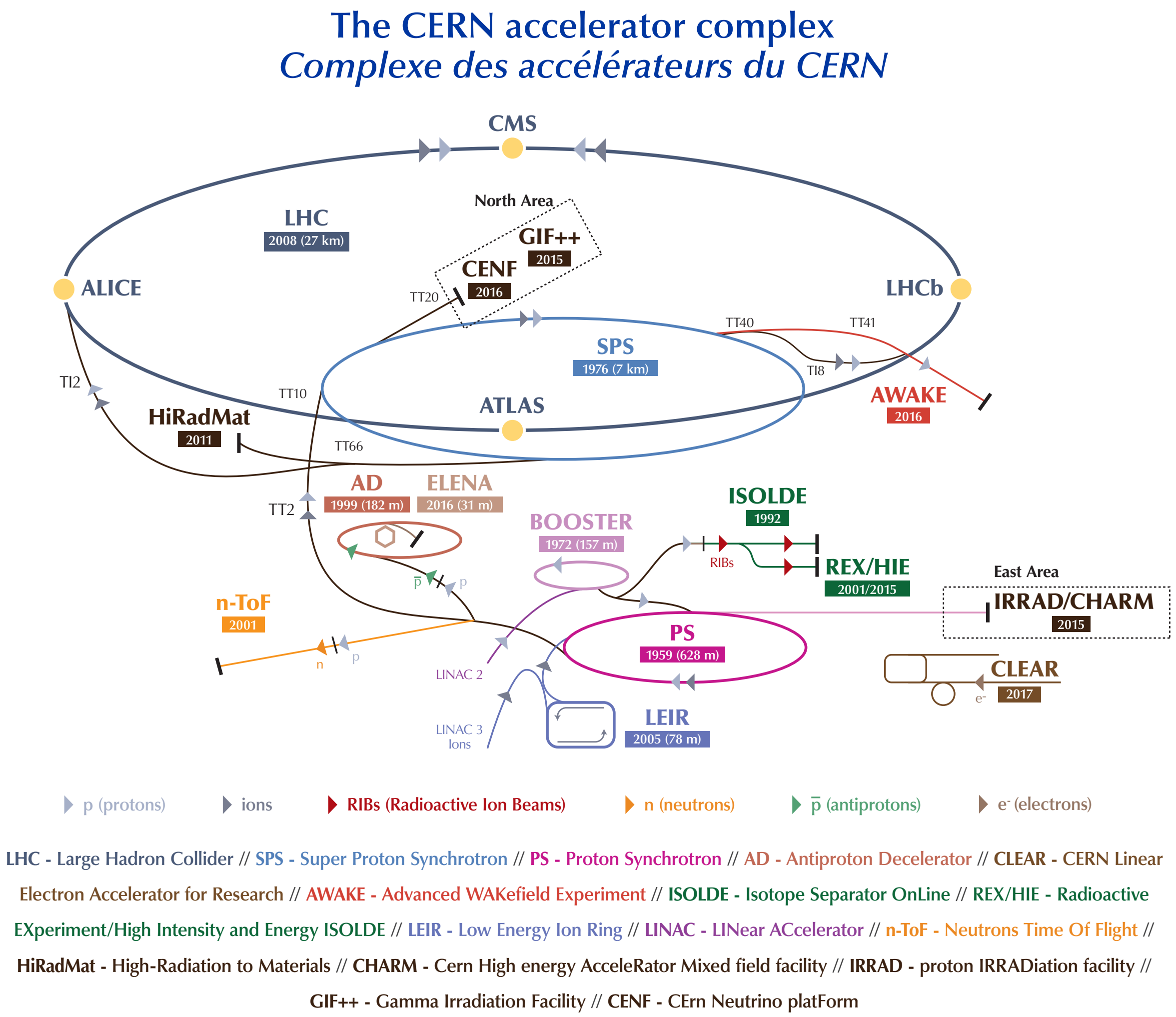


Figure 3.1: The CERN Accelerator Complex showing all experiments around the LHC (Image: CERN).

undergo further acceleration to reach an energy level of 450 GeV before being introduced into the LHC. Inside the LHC tunnel, 1232 superconducting dipole magnets, each producing an 8.33 T magnetic field, are used. These magnets are made of Niobium–titanium alloy (NbTi) and operate at temperatures below 2 K. Quadrupole and higher-order magnets fine-tune the beam path. Protons are accelerated by 16 high-frequency radio cavities with a 2 MV/m gradient. The complete acceleration process takes roughly 20 minutes.

At the LHC there are eight crossing points flanked by long straight sections for the RF cavities. However, it was decided to equip only four of the eight possible interaction regions and to suppress beam crossings in the other four regions to prevent unnecessary disruption of the beams. The experimental sites of the collider, located in the four equipped interaction regions, house complex and sophisticated detectors, such as ATLAS (A Toroidal LHC ApparatuS) [14], CMS (Compact Muon Solenoid) [41], ALICE (A Large Ion Collider Experiment) [42] and LHCb (Large Hadron Collider beauty) [43], which are used for different types of physics analysis. ATLAS and CMS are general purpose detectors designed to study a wide range of SM and Beyond the Standard Model (BSM) physics phenomena, whereas LHCb focusses on B meson physics and ALICE focuses on heavy-ion physics.

3.2 The ATLAS Detector

The ATLAS detector [14] shown in Figure 3.2 is a versatile particle detector that provides comprehensive coverage of nearly the entire solid angle surrounding the collision point. It is the largest detector ever constructed for a particle collider with 46 m in length and 25 m in diameter.

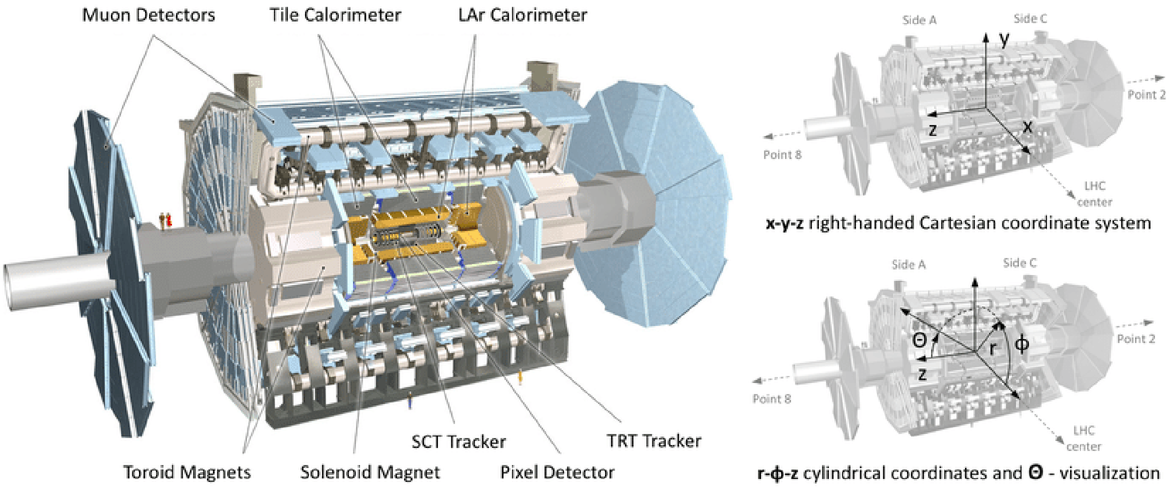


Figure 3.2: The image on the left showcases the components of the ATLAS detector. On the right side, the image defines both the cylindrical and cartesian coordinate systems of the ATLAS detector [44].

In the ATLAS detector, the nominal interaction point is defined as the origin of the coordinate system. The z axis aligns with the beam axis, and the x-y plane is transverse to the beam direction, as defined in Figure 3.2. The positive x axis is defined as pointing from the interaction point to the centre of the LHC ring, and the positive y axis is defined to be pointing upwards. The side A of the detector is defined as positive z, and the side C is negative z. The azimuthal angle ϕ is measured as usual around the beam axis, and the polar angle θ is the angle from the beam axis. R describes the distance from the beam axis. The pseudo-rapidity is defined as $\eta = -\ln(\tan(\frac{\theta}{2}))$ and in the case of massive objects such as jets, the rapidity $y = \frac{1}{2}\ln\left(\frac{(E+p_z)}{(E-p_z)}\right)$ is used. The transverse momentum p_T, the transverse energy E_T, and the missing transverse energy E^T_{miss} are defined in the x-y plane. The distance ΔR in the pseudorapidity-azimuthal angle space is defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

The inner detector lies at the heart of this setup. Beyond this lies the calorimeter section, beginning with the electromagnetic calorimeter and followed by the hadronic calorimeter. Surrounding this is the outer shell that houses the magnet system and the muon spectrometer.

3.2.1 Inner Detector

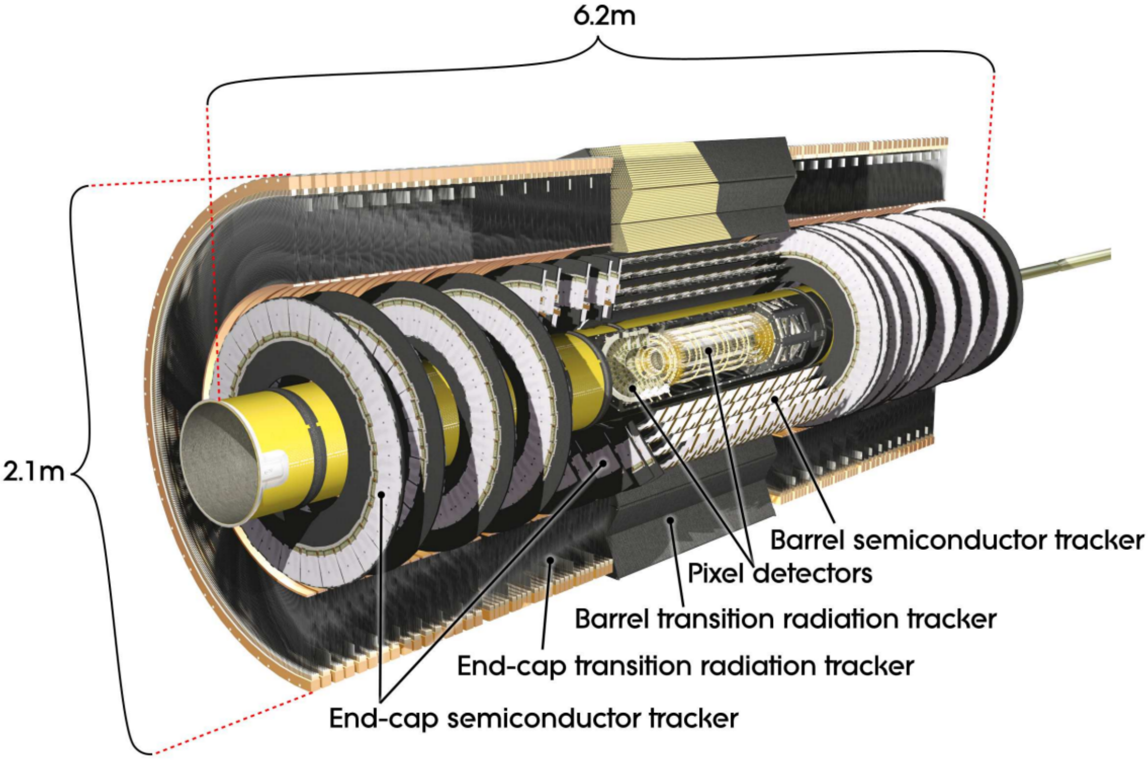


Figure 3.3: Cut-away view of the ATLAS inner detector (Image: The ATLAS collaboration).

The Inner Detector (ID) [45] within this setup comprises three key components: the pixel detector, the SemiConductor Tracker (SCT), and the transition radiation tracker (TRT). The acceptance in pseudorapidity is $|\eta| < 2.5$ for particles coming from the LHC beam-interaction region, with full coverage in ϕ. The detector has been designed to provide a transverse momentum (p_T) resolution, in the plane perpendicular to the beam axis, of $\sigma_{p_T}/p_T = 0.05\% \cdot p_T \text{ GeV} \oplus 1\%$.

The pixel detector, positioned closest to the interaction point (IP), contains the Insertable B-Layer (IBL) [46] and additional layers in the barrel and end-cap regions. With a total of approximately 92 million readout channels, it provides track measurements post-collision with a high spatial resolution with a fine pixel size of $50 \times 250\ \mu\text{m}^2$ in the IBL and $50 \times 400\ \mu\text{m}^2$ for the subsequent layers in the $R\phi \times z$ space.

The SCT, following the pixel detector, utilises 12 cm silicon strips with a pitch of 80 μm for tracking charged particles. The total number of readout channels in SCT is approximately 6.3 million. The SCT maintains spatial separation between tracks, preventing resolution degradation as the distance from the interaction point increases.

Simultaneously, the TRT aids in particle identification by detecting transition-radiation photons from the charged particle, utilising straw tubes and a xenon-based gas mixture. The crossing particles ionize the gas mixture inside the straws and create an electron drift towards the grounded wire in the center of the tube. The electrons cascade close to the wire and create a detectable signal. The total number of TRT readout channels is approximately 351,000. Together with the two end-cap regions it has an acceptance range of $|\eta| < 2.0$. These components collectively contribute to momentum measurements by utilising the Lorentz force, where charged particles' trajectories bend relative to the 2 T solenoid magnetic field. This bending allows calculation of a particle's transverse momentum (p_T) on the basis of its charge and the curvature of its path.

3.2.2 Calorimeter

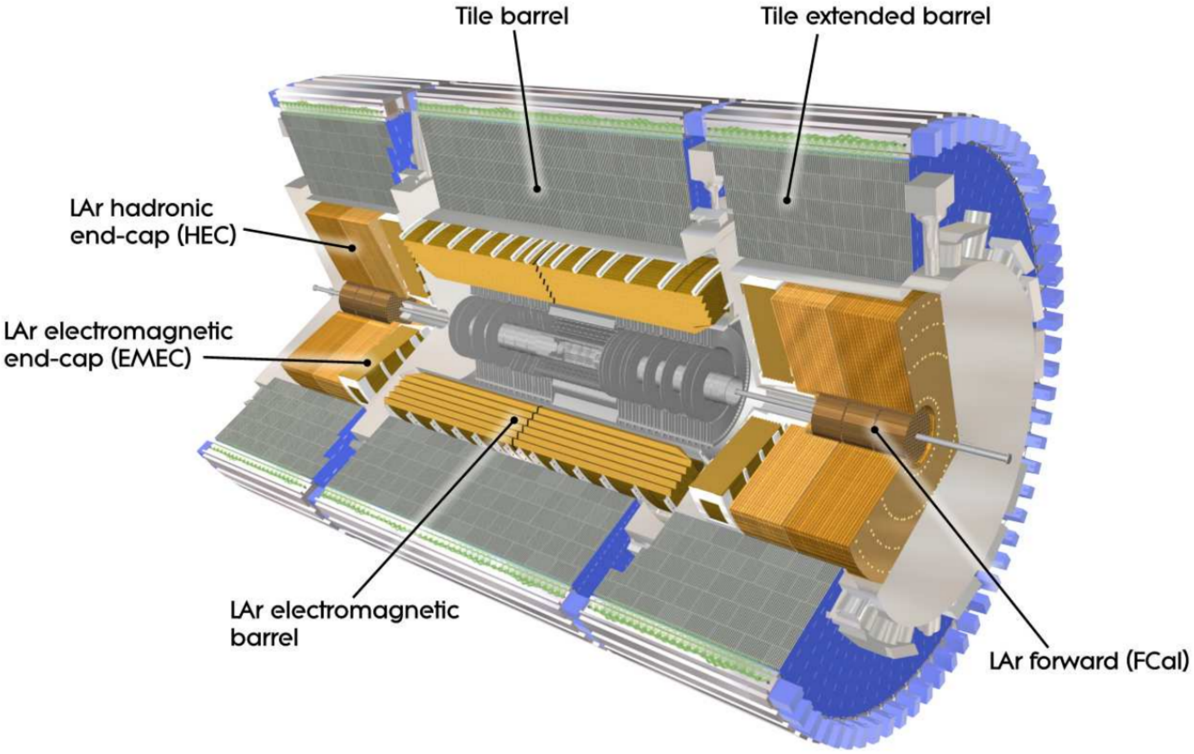


Figure 3.4: Cut-away view of the ATLAS calorimeter system (Image: The ATLAS collaboration).

The ATLAS detector's calorimeter, situated beyond the ID's solenoidal magnet, is illustrated in Figure 3.4. Calorimeters determine the particle energy through a process that destructively assesses particles as they enter, triggering a particle shower whose energy is then measured. These calorimeters are engineered to confine electromagnetic and hadronic showers while minimising their intrusion into the muon detector system.

The Electromagnetic Calorimeter (ECal) is divided into a barrel ($|\eta| < 1.475$) and two end-cap components ($1.375 < |\eta| < 3.2$). The central solenoid and liquid argon calorimeter (LAr) share a common vacuum vessel to save material costs. The ECal's design features accordion-shaped electrodes and lead absorber plates, ensuring complete ϕ symmetry without azimuthal gaps. The ECal has 101,760 read-out channels in the barrel region and 62,208 readout channels in the endcap region. Additionally, in regions of $|\eta| < 1.8$, a presampler detector corrects energy losses by electrons and photons before reaching the calorimeter, using active LAr layers of different thickness in the barrel and end cap regions. The presampler has 7808 read-out channels in the barrel region and 1536 readout channels in the end-cap region. The ECAL is designed to achieve an energy resolution of $\sigma_E/E = 10\%/\sqrt{E(\text{GeV})} \oplus 0.7\%$ for electromagnetic showers.

The Hadronic Calorimeter (HCal) consists of the Tile Calorimeter in the barrel region. It is situated just outside the EM calorimeter and uses steel as the absorber and scintillating tiles as the active material. The Tile Calorimeter covers the barrel region $|\eta| < 1$ and its two extended barrels the range $0.8 < |\eta| < 1.7$, with 5760 and 4092 readout channels, respectively.

The Hadronic End Cap Calorimeter (HEC) comprises two independent wheels per end-cap, situated behind the end-cap electromagnetic calorimeter and sharing LAr cryostats. It has 5632 readout channels in total. Extending to $|\eta| = 3.2$, it overlaps the forward calorimeter, and at $|\eta| = 1.5$, it slightly intersects the Tile Calorimeter ($|\eta| < 1.7$). Each wheel, made of 32 identical wedge-shaped modules, divides into two depth segments, totalling four layers per end cap. Using copper plates and LAr gaps, the HEC's design varies plates between 25 mm and 50 mm thickness. Meanwhile, the LAr Forward Calorimeter (FCal) with 3524 readout channels is integrated into the end-cap cryostats, as this provides clear benefits in terms of uniformity of the calorimetric coverage and reduced radiation background levels in the muon spectrometer. The HCAL and HEC have an energy resolution of $\sigma_E/E = 50\%/\sqrt{E(\text{GeV})} \oplus 3\%$, while the FCal is designed to achieve $\sigma_E/E = 100\%/\sqrt{E(\text{GeV})} \oplus 10\%$ for hadronic showers.

3.2.3 Muon Detector

The muon spectrometer shown in Figure 3.5 within the ATLAS detector is a large tracking system designed to determine the charge and momentum of charged particles that exit the calorimeters. As the name suggests, it plays a crucial role in identifying muon particles. Typically, all SM particles that interact electromagnetically or strongly, apart

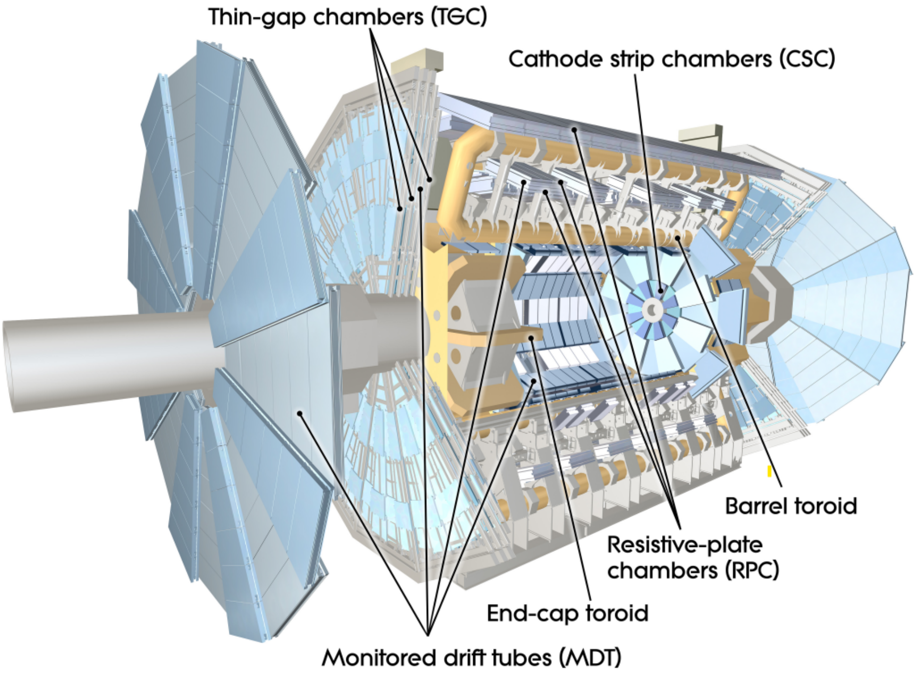


Figure 3.5: Cut-away view of the Atlas muon detector system (Image: The Atlas collaboration).

from muons, are expected to be stopped by the calorimeters, unless they possess extraordinarily high energies, enabling them to penetrate through the other detectors. The muon detector relies on superconducting air-core toroid magnets for magnetic deflection, aiding in precision tracking. Covering various η ranges, it utilizes Monitored Drift Tubes (MDT), Cathode Strip Chambers (CSC), Resistive Plate Chambers (RPC), and Thin Gap Chambers (TGC) for tracking and triggering purposes. The muons travel through the toroidal magnetic field, which bends them in curved trajectories to measure the p_T with higher precision. The barrel toroid provides 1.5 to 5.5 T·m bending power ($0 < |\eta| < 1.4$), while the end cap toroids offer 1 to 7.5 T·m ($1.6 < |\eta| < 2.7$). Over $1.4 < |\eta| < 1.6$, usually referred to as the transition region, magnetic deflection is provided by a combination of barrel and end-cap fields. The alignment precision of 30 μm within and between the chambers, monitored by optical sensors, ensures accurate momentum measurements. The muon spectrometer is designed to achieve a resolution of $\sigma_{p_T}/p_T = 10\%$ for a muon with a p_T of 1 TeV.

Chapter 4

The ATLAS Trigger system

The ATLAS experiment is designed to record a large amount of physics collisions, with a bunch crossing rate of 40MHz. Each bunch crossing involves multiple inelastic proton-proton collisions, collectively treated as a single online event, approximately 3 MB in size [47]. This substantial data size is due to the use of high-resolution detectors and

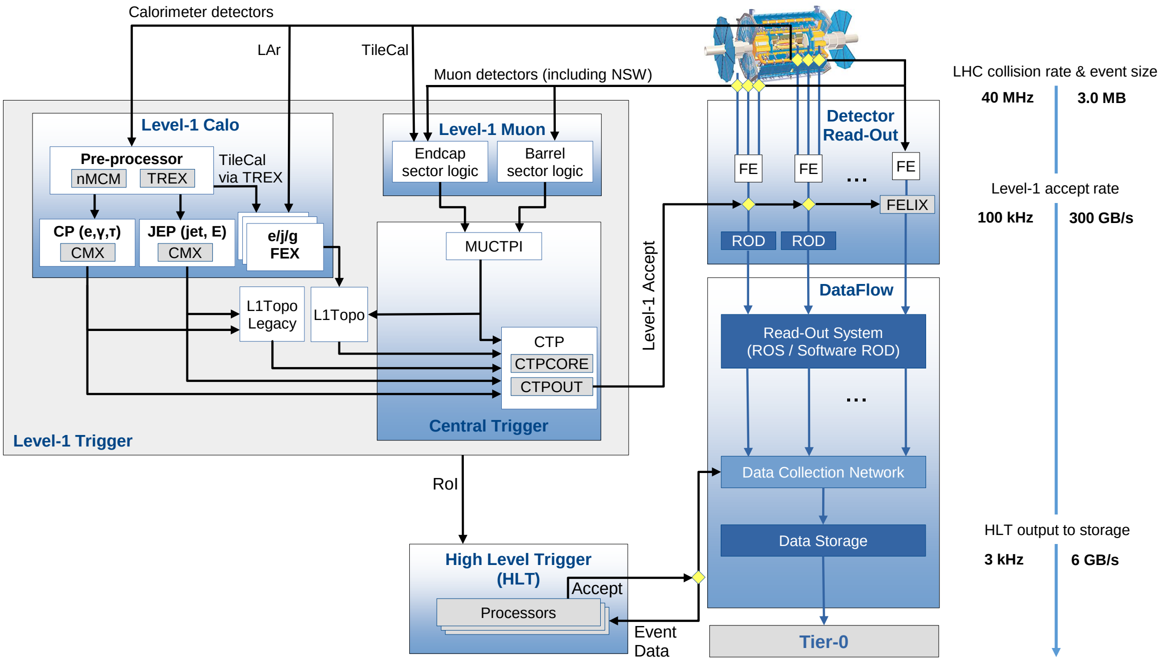


Figure 4.1: The ATLAS Trigger and Data Acquisition system in Run-3 with emphasis on the components relevant for triggering [19].

a significant number of readout channels. Consequently, the data generated at ATLAS reaches a rate in the order of TB/s. However, existing systems lack the capacity to store events at such a high rate. Therefore, it is necessary to make real-time decisions on whether to retain or discard the event. To do this, a highly sophisticated Trigger and Data AcQuisition (TDAQ) system, represented in Figure 4.1, has been implemented. A graphical representation of the ATLAS event filtering system depicting the reduction in rates after different trigger levels is provided in Figure 4.2. The TDAQ system selects an event based on a predefined set of conditions required by the trigger menu using a hardware-based Level-1 trigger (L1) and a software-based high-level trigger (HLT).

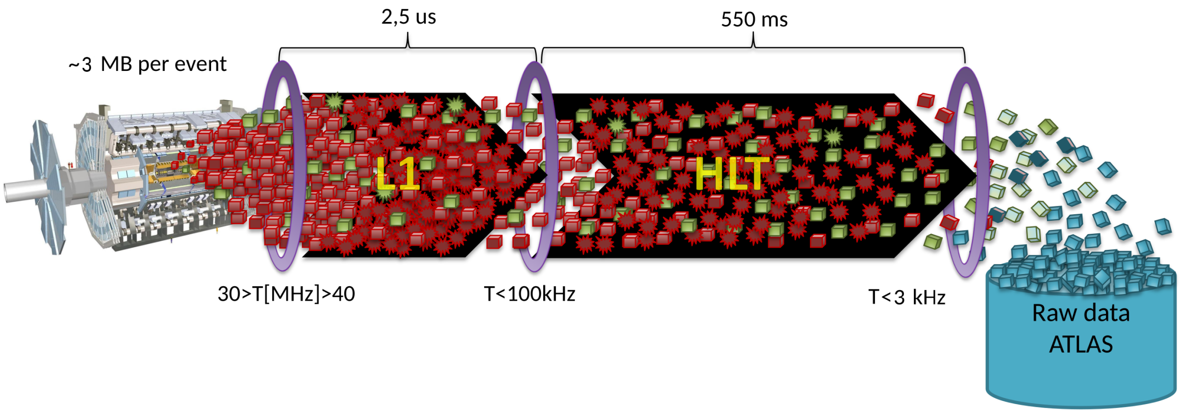


Figure 4.2: Visual representation of event filtering system at different trigger levels at ATLAS. Adapted from graphic by © C. Meyer and J. Pinto, CC BY 2.5, Introduction to e/γ Triggers.

4.1 The Level 1 Trigger System

The Level-1 (L1) trigger [48] is a hardware-based system that uses custom electronics to trigger on reduced granularity information from the calorimeter and muon detectors. Using the L1 Calorimeter (L1Calo) triggers [49], it evaluates event-level parameters such as total calorimeter energy, object counts exceeding set thresholds (e.g., muon transverse momentum), or using the L1 topological (L1Topo) trigger [50], employ topological criteria (e.g., invariant masses, angular distances).

Starting from Run 3, the upgraded L1Calo trigger will make full use of the finer granularity information from the Tile Calorimeter, using new e/γ and jet feature extraction modules, referred to as eFEX and jFEX, respectively [51]. The eFEX subsystem will employ new cluster finding algorithms on the higher granularity data to produce more refined e/γ and τ Trigger OBjects (TOBs) candidates. The jFEX subsystem identifies energetic jet candidates and also calculates the ΣE_T and E_T^{miss}.

The L1Topo system, nested within the L1 trigger, uses TOBs from the L1Calo systems as input to topological algorithms, and the results are forwarded to the Central Trigger Processor (CTP). The CTP is responsible for applying preventive dead-time. It limits the minimum time between two consecutive L1 trigger accepts (simple dead-time) to avoid overlapping readout windows and restricts the number of L1 trigger accepts allowed in a given number of bunch crossings (complex dead-time). Together, this forms the L1 trigger system. Operating within a 100 kHz read-out rate (from an original 40 MHz bunch crossing rate) with a 2.5μs latency, the L1 trigger efficiently selects events that meet defined criteria. The L1 trigger also provides region-of-interest (ROI) information of the L1 triggered events from the eFEX and jFEX subsystems to the HLT.

4.2 The High Level Trigger System

After the L1 trigger acceptance, the events are buffered in the Read-Out System (ROS) and processed by the HLT. The HLT stage reduces the event rate from 100kHz to 3kHz. This part of the trigger is software-based. A typical reconstruction sequence makes use of dedicated fast trigger algorithms to provide early rejection, followed by more precise and CPU-intensive algorithms that are similar to those used for offline reconstruction to make the final selection. Unlike the L1 trigger, it makes use of the complete granularity and precision of the calorimeter and muon chamber, and also incorporates limited tracking information from the inner detector. It is therefore able to trigger on more complex signatures, like secondary vertices from B-hadrons, over the entire event. After the events are accepted by the HLT, they are transferred to local storage at the experimental site and exported to the Tier 0 facility at the CERN computing centre for offline reconstruction.

4.3 Object Reconstruction

A large part of any physics analysis depends on the accuracy with which different objects are reconstructed. During proton-proton collisions at ATLAS, various particles emerge. With the exception of a few particles, all other SM particles possess finite lifetimes and undergo decay, transforming into other particles. Often, these lifetimes are extremely short, making it unfeasible to have detector components in close proximity to the particle created for direct detection. Therefore, only the resultant decay products are observable. The measurement process requires the interaction of the particles with the detector. Different measurement techniques have been implemented for different particles depending on their specific behaviour when interacting with matter. In the context of this thesis, the identification of τ_{had} leptons and b jets at the HLT level is of particular significance, and this section provides an extensive discussion on this matter.

4.3.1 Reconstruction of τ Leptons

Leptonic decays of the τ leptons closely resemble the direct production of lighter leptons. Therefore, τ identification algorithms usually do not consider them. As hadronic decays also contain undetectable neutrinos, reconstruction algorithms focus on the remaining decay products, denoted with the symbol τ_{had}.

τ_{had} Objects at L1 trigger

The data collected by the calorimeter first goes through the L1 trigger, where selections are made for identifying the τ_{had} candidates. Since track reconstruction takes too long at this stage, the L1 τ trigger is limited to using information from the calorimeter trigger towers as in Figure 4.3. Within the L1 Calorimeter (L1Calo) trigger system [49], each candidate is divided into two regions using trigger towers in both ECal and HCal: the core region and an isolation region around it. These trigger towers are finely divided with a granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ and cover $|\eta| < 2.5$. The core region is a square formed by 2×2 trigger towers, which translates to 0.2×0.2 in the $\Delta\eta \times \Delta\phi$ space.

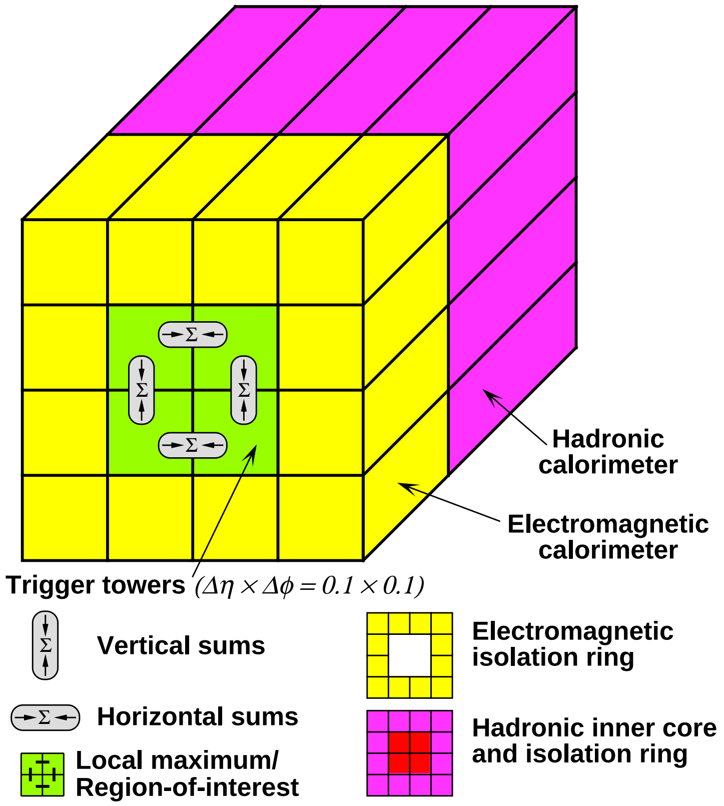


Figure 4.3: Description of the ATLAS trigger tower scheme for τ RoI.

To determine the E_T of a τ_{had} candidate in L1 trigger, the system sums the transverse energy in the two most energetic neighbouring central towers within the core region of ECal, as well as in the 2×2 towers within HCal. The Region of Interests (RoI) corresponding to the τ_{had} objects are identified by placing a lower threshold of 10 GeV on the transverse energy (E_T) sum in 2×2 clusters of ECal and HCal towers [52], and an upper threshold on the energy found in a 4×4 isolation ring[1] around the 2×2 core. The L1Calo information undergoes additional processing using the Level-1 topological trigger (L1Topo) [50], employing topological algorithms and a multiplicity trigger. These algorithms factor in the spatial configuration of trigger objects, effectively reducing the L1 trigger rate. Specifically, jets that fall within a cone of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.4$ around the axis of tau objects are linked to the τ_{had} candidates.

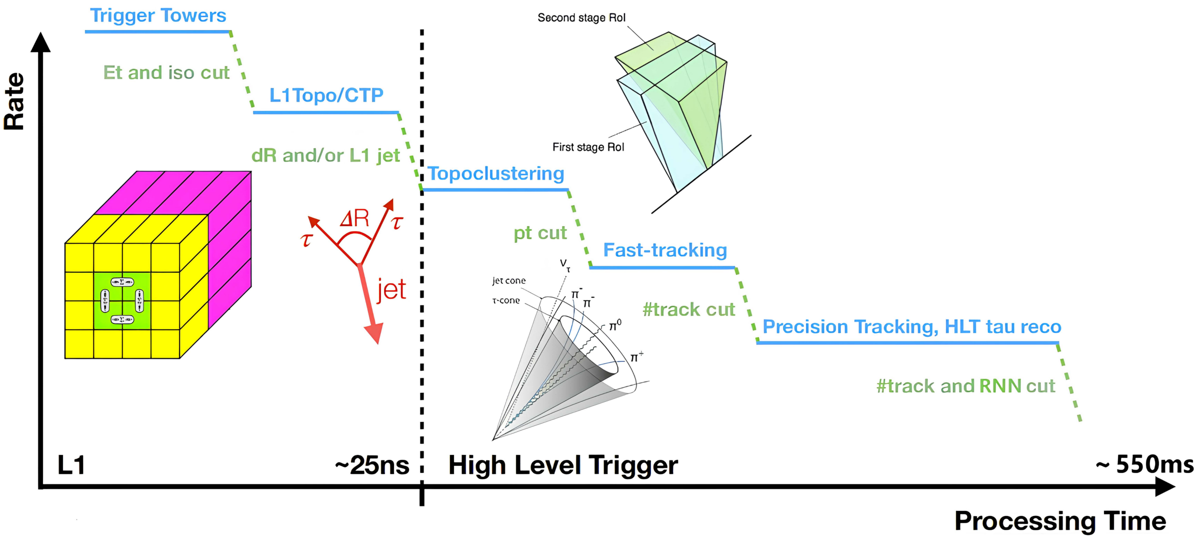


Figure 4.4: Different stages of reconstruction and selection, at different trigger levels for τ_{had} objects. Adapted from graphic by © Serhat Ördek, CC BY 2.5, PhD thesis.

τ_{had} Objects at HLT

Events that pass the L1 trigger requirements are processed by the HLT, which has access to more accurate position and energy measurements. The less expensive CPU algorithms are executed at the beginning of the HLT, whereas the more CPU intensive ones, like precision tracking, are executed in the last step and only on the candidates passing the previous selection criteria in the HLT stage. A summary of the τ lepton trigger process with respect to processing time is shown in Figure 4.4.

The ROI information from the L1Calo system is retrieved and the jets are clustered using full granularity, and these are very similar to the offline reconstructed clusters. These clusters are calibrated with the *local hadron calibration* (LC) algorithm [53] and their vector sum is used as a jet seed for the reconstruction of the τ_{had} candidate. Jets create challenges in distinguishing τ_{had} objects associated jets. The tracks of the τ_{had} jets form a tight cone, whereas the QCD jets have tracks that are more spread out. Taking

---

[1] In the L1 trigger chains, `I` or `IM` stands for loose or relatively tight version of these cuts respectively.

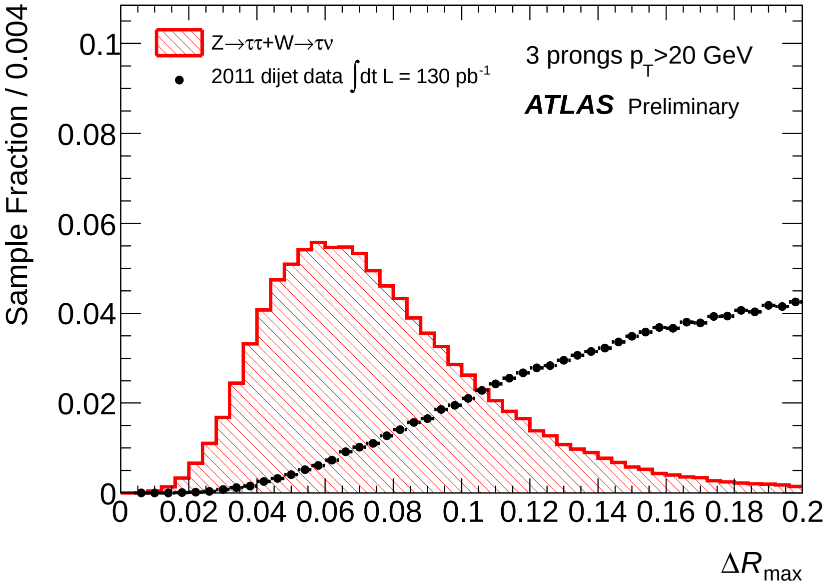


Figure 4.5: Maximal distance between a track and the axis of tau object, $\Delta R_{\max}$. MC simulated $Z \rightarrow \tau\tau$ and $W \rightarrow \tau\nu$ signal samples and a di-jet background sample selected from 2011 data.

advantage of this, an optimal cut of $\Delta R < 0.2$ is applied on the clusters to calculate E_T of the τ_{had} candidates. This is supported by Figure 4.5 which shows distribution of the maximal distance between a track and the axis of tau lepton like objects (ΔR) and of the QCD background. For this purpose, all LC calorimeter clusters are used. If it is above the E_T threshold, then the inner detector information is used for track reconstruction. An *Event Filter* (EF) algorithm [54] is used for precision tracking and requires $0 \leq N_{\text{tracks}} \leq 3$ in the core region ($\Delta R < 0.2$) and a maximum of one track in the isolation region $(0.2 < \Delta R < 0.4)^2$. In the context of trigger chains, this selection is called *tracktwoMVA*. Until Run-2, this was followed by an offline like selection using a *Boosted Decision Tree* (BDT) algorithm [55], which has been superseded by the *Recurrent Neural Networks* (RNN) [56] for Run-3.

RNNs integrate high-level variables such as Central energy fraction (f_{cent}) etc., with low-level track and calorimeter information such as Cluster depth (λ_{cluster}), transverse momentum of the seed jet ($p_{\text{T}}^{\text{seed jet}}$) etc., for the identification of τ_{had} objects, preserving sequential dependencies in the data. Using Long-Short-Term Memory (LSTM) layers, RNNs handle variable length input, capturing nuanced patterns crucial for discriminating 1-prong and multi-prong τ_{had} candidates from the background jet objects. The output score of the RNN-based tau identification for true τ_{had} is transformed to be uniform in [0, 1] in bins of reconstructed τ_{had} p_{T} and average number of interactions per bunch-crossing(μ). The flattened RNN scores for the 1 prong and the 3 prong tau objects are shown in Figure 4.6. The algorithm, the input variables, and the structure of the

---

[2] At trigger level, only tau candidates with up to 3 prongs are reconstructed to avoid QCD background.

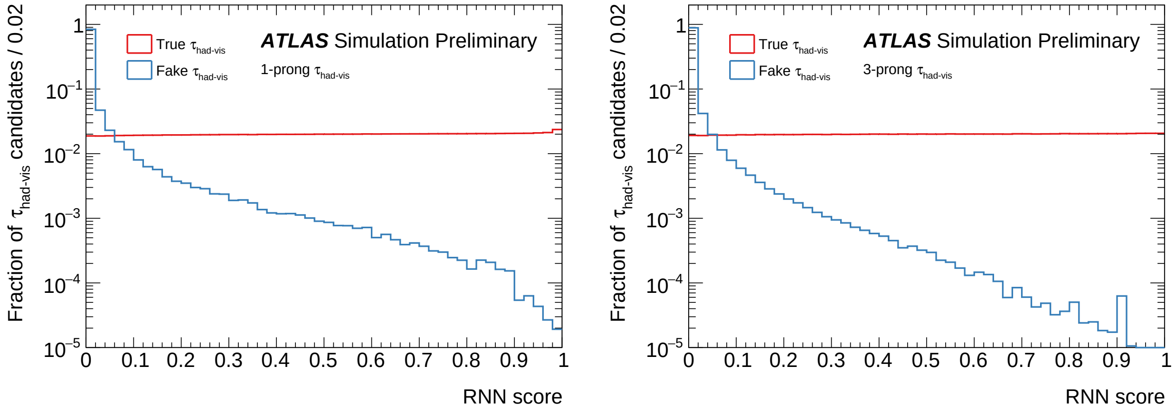


Figure 4.6: Distribution of the RNN score for true (red) and misidentified (blue) τ_{had} candidates for 1-prong (left) and 3-prong (right) τ_{had} candidates. The RNN score has been flattened such that it corresponds to the fraction of rejected true τ_{had} [56].

RNN are explained in detail in Appendix A. A comparison of the performance of these algorithms is provided in Figure 4.7, which shows that the background rejection power of the RNN-based classifier is approximately two times better than that of the BDT-based classifier in both 1-prong and multi-prong cases for any given signal selection efficiency. Table 4.1 provides the current thresholds for online working points (WP) of RNN based classifiers.

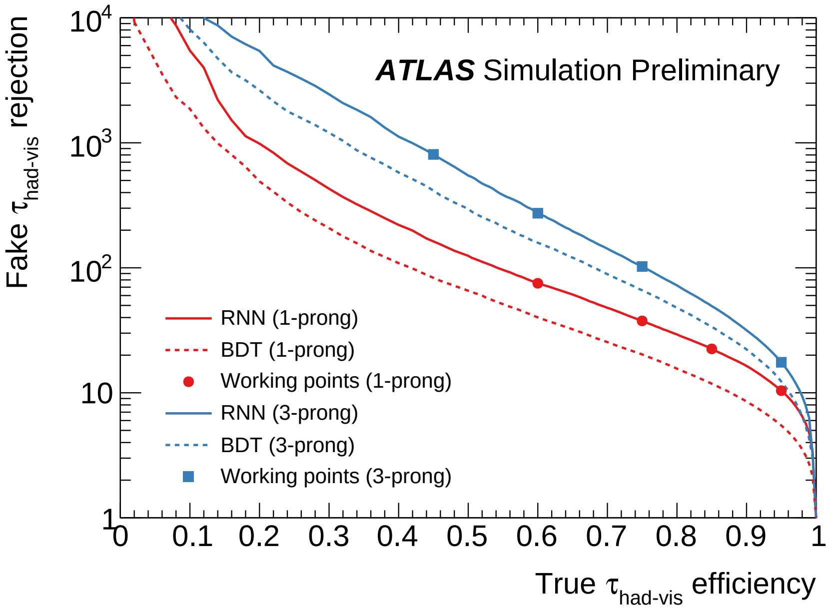


Figure 4.7: The horizontal axis shows the efficiency of correctly identifying hadronic tau jets, while the vertical axis shows the inverse of the efficiency of incorrectly identifying quark-initiated jets as hadronic tau jets. The red curves show the performance exclusively on tau jets with a single track (1-prong), while the blue curves represent tau jets with three tracks (3-prong). The RNN-based model's performance is shown in the full lines, with the dashed lines representing an algorithm with similar inputs but based on a boosted decision tree [56].

Working Point	1-Prong threshold	multi-Prong threshold
Tight	0.06	0.2
Medium	0.03	0.105
Loose	0.01	0.06
Very Loose	0.008	0.01

Table 4.1: Thresholds of the different WPs for the RNN-based online tau object classification.

4.3.2 Reconstruction of *b* Quarks

The selection of *b*-jet objects stands out as one of the most CPU-intensive procedures at the trigger level. Therefore, it is essential to streamline this procedure. The subsequent discussion will elaborate on the process of identification of the *b* jet objects at different trigger levels.

b-Jet Objects at L1 trigger

Jets leave energy in both the ECal, HCal, and FCal, and their charged constituents additionally produce tracks in the Inner Detector. Jet objects are identified by the L1Calo trigger in a 4×4 or 8×8 trigger tower cluster that includes a 2×2 local maximum that defines the RoI's coordinates. From this cluster, the E_T is calculated by summing up the cell energies. If the E_T is above a specific threshold, the L1Calo data undergoes processing through the L1Topo trigger. In addition to the p_T thresholds, the L1 selections depend on the multiplicity or the topology of the jet objects. This is determined at the L1Topo stage.

b-Jet Objects at HLT

The entire process of *b*-tagging in HLT level is illustrated in Figure 4.8. ROIs identified at L1 trigger serve as the starting point for HLT. Despite the decrease in the event rate at L1 trigger, achieving accurate *b*-tagging requires track reconstruction, a resource-intensive process that could strain the HLT farm. To address this in LHC Run-3, ATLAS has introduced a *fast b-tagger* [57] based on neural networks for a crude preselection. This *b*-tagger acts as a quick filter for high-rate L1 chains, using data from hadronic jets and tracks. Track candidates for *b*-tagging at this stage are provided by a *fast track finder* (FTF) that takes multiple L1Calo ROIs and creates a super ROI, so that the tracking algorithms do not need to run multiple times in every ROI with large overlapping regions.

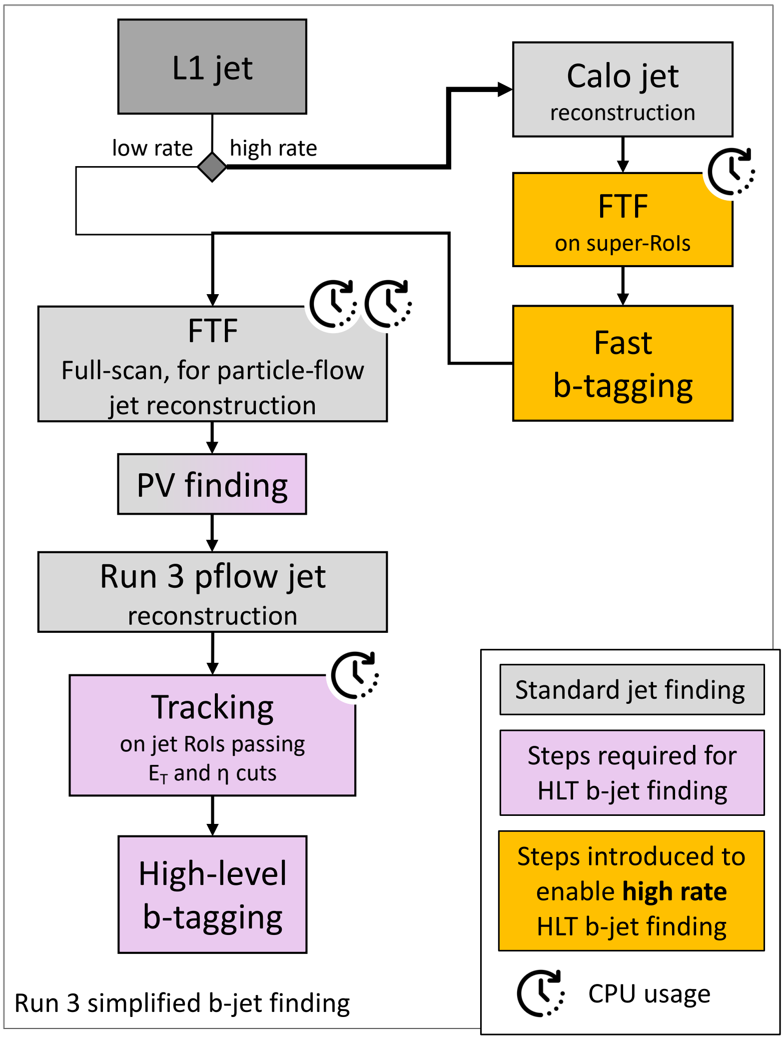


Figure 4.8: Simplified schematic of the *b*-jet trigger selections in Run-3 ATLAS trigger implementation.

Once the fast *b*-tagging selection is made, the FTF algorithm is re-run on ROIs of restricted sizes. To reduce the number of jets reconstructed as a consequence of pileup, the *Jet Vertex Tagger* (JVT) [58], which is trained on pileup-insensitive variables, is used to identify jet vertices associated with the hard-scatter interaction. Jets meeting these criteria are reconstructed through a particle flow (pflow) [59] algorithm which uses *anti-kt* [60] clustered jets. Precision tracking is then performed on these jets, which will later be used by high-level *b*-tagging algorithms.

The decay of the B meson, with its limited lifespan, is identified by the presence of a secondary vertex (SV), noticeably displaced from the primary vertex (PV), as shown in Figure 4.9. Various SV properties such as the transverse impact parameter, the distance to the SV from the PV, the number of tracks, ΔR, etc. can be used to confirm whether the identified jet is indeed a b jet.

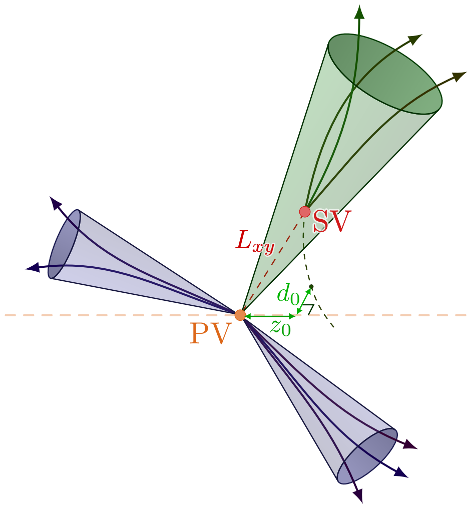


Figure 4.9: Characteristics of a jet initiated by a bottom quark with the secondary vertex, as well as a longitudinal (z_0) and transverse impact parameter (d_0) of individual tracks from their point of closest approach to the primary vertex. Adapted from graphic by © Izaak Neutelings, CC BY 2.5.

Different algorithms, such as DL1r [61] and DL1d [62] were used in the past for b-tagging purpose. The latest one uses Graph Neural Networks (GN1) [63] for the b-tagging process at ATLAS. Unlike previous approaches that relied on multiple independently optimised algorithms fed into a high-level network, GN1 uses a single neural network that directly takes tracks and jet information as input, streamlining the process, and comprehensively optimising the algorithm. A variant of GN1, called GN1 Lep, was also trained, incorporating an additional track-level input variable in the training, which indicates if the track was involved in reconstructing an electron or a muon.

All of the taggers mentioned above make predictions about the likelihood of a jet belonging to the b-, c-, and light classes. For b-tagging purposes, these probabilities are merged into a single score - D_b, which is defined as:

$$D_b = \log \frac{p_b}{(1 - f_c)p_l + f_c p_c} \tag{4.1}$$

where f_c is a free parameter that determines the relative weight of p_c to p_l in the score D_b, balancing the trade-off between the rejection of the c jet objects and the light jet objects. This parameter is usually set at a value of $f_c = 0.018$ for the DL1r model and $f_c = 0.05$ for the GN1 model [61, 63]. The final score of D_b can be used as a discriminant to separate the b jets from the c and light-flavoured jet objects, as shown in Figure 4.10. The b-tagging efficiencies of the GN1 Lep, GN1 and DL1r taggers as a function of p_T are shown in Figure 4.11.

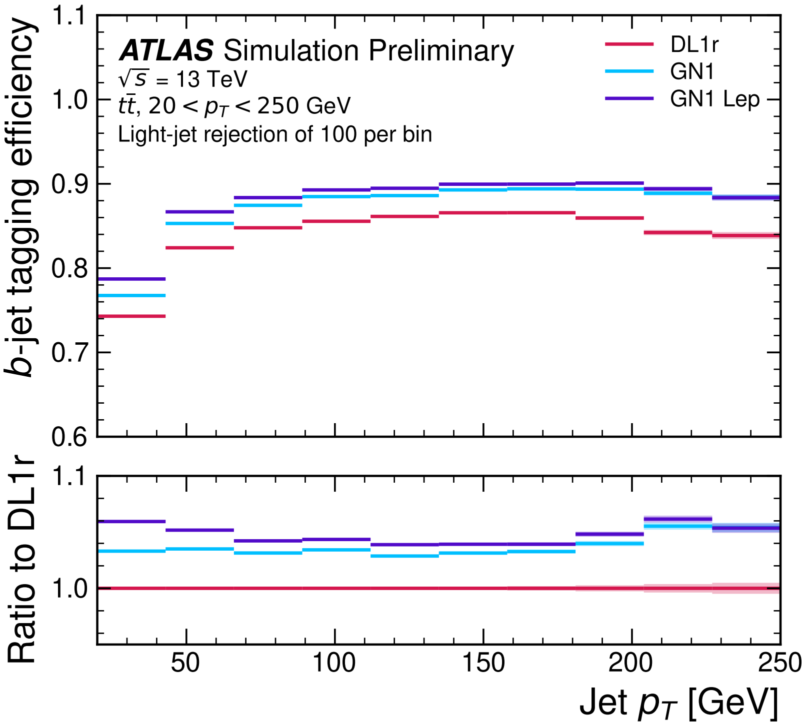


Figure 4.11: The b-jet tagging efficiency for the jets in $t\bar{t}$ sample as a function of jet p_T [63].

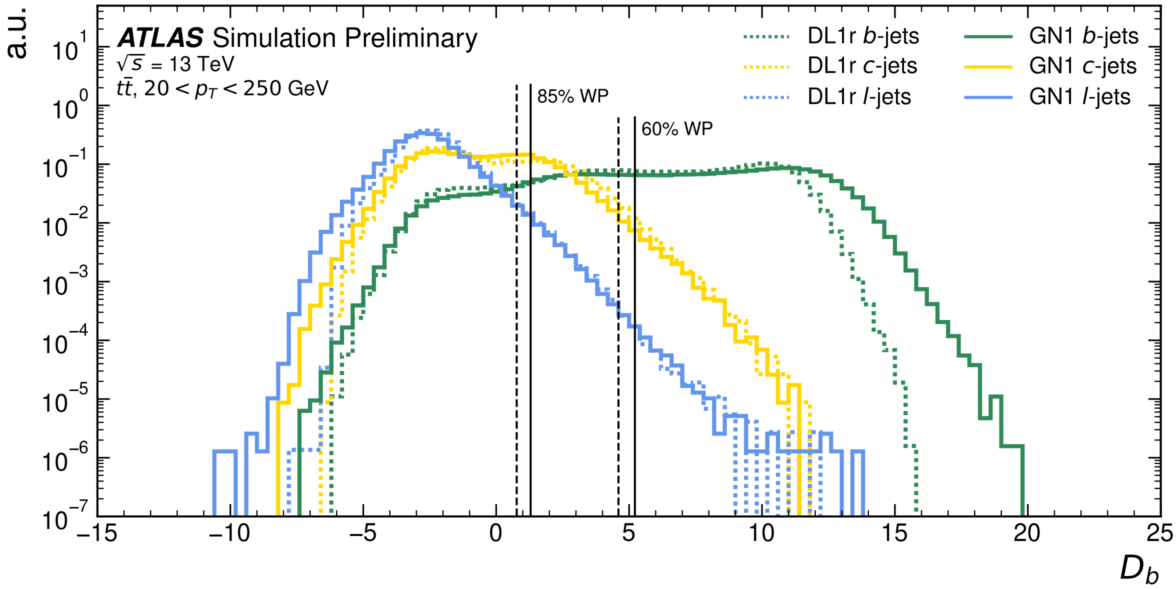


Figure 4.10: Comparison between the DL1r and GN1 btag discriminant D_b for jets in the $t\bar{t}$ sample. The 85% WP and the 60% WP are marked by the solid (dashed) lines for GN1 (DL1r), representing the loosest and tightest WPs, respectively, used by analyses. A value of $f_c = 0.018$ is used in the calculation of D_b for DL1r and $f_c = 0.05$ is used for GN1. The distributions of the different jet flavours have been normalised to unity area [63].

4.4 The Trigger Menu

The trigger menu contains a list of trigger chains used for data collection. It also includes prescales for each trigger chain which can be used to either adjust or completely disable the rate of a chain. Being able to change trigger prescales during an ongoing run is extremely valuable in providing luminosity-dependent trigger configurations during data taking and, therefore, in making optimal use of the available resources.

The physics trigger menu contains various trigger chains, each tailored to capture specific physics events or signatures. As of 2023, it includes over 2500 HLT items and more than 500 L1 trigger items.

Typically, a trigger chain in the menu begins with `HLT_` followed by conditions it must meet at the HLT stage. It concludes with the required L1 seed chain for the HLT, starting with `_L1`. L1 trigger and HLT items are written in upper-case and lower-case letters, respectively. These chains adhere to a structured naming convention, as outlined in [64]:

```
<Multiplicity><Object><Threshold>_<Additional_Params>
```

Here, `<Object>` denotes entities like jets (`j`), τ objects (`tau`), μ objects (`mu`), etc., while `<Multiplicity>` signifies the required number of these objects. `<Threshold>` specifies the threshold for transverse momentum (p_T) or transverse energy (E_T) based on object type. Sometimes, objects may have additional requirements like the Recurrent Neural Network (`RNN`) WP for tau objects or the b-tagging WP for the jet objects, etc. A chain can even have no conditions at the HLT stage, for which the chain looks like `HLT_noalg_L1<condition>`.

4.4.1 Menu items for HH $\rightarrow b\bar{b}\tau^{+}_{\mathrm{had}}\tau^{-}_{\mathrm{had}}$ channel

For effective selection of HH $\rightarrow b\bar{b}\tau^{+}_{\mathrm{had}}\tau^{-}_{\mathrm{had}}$ events, a set of di tau triggers (DTT) and single tau triggers (STT) are used in the physics trigger menu as listed in Table 4.2. Priority is given to STT events if the reconstructed τ_{had} meets the p_T threshold of the trigger (100 GeV for tau80, 140 GeV for tau125, and 180 GeV for tau160) and are geometrically matched to the HLT object that fired the trigger. If the event does not meet the STT criteria, then the DTT criteria are checked.

Within this analysis, the Run-3 version of the last two DTT triggers listed in Table 4.2 are of significance, denoted as $\mathtt{HLT}_{\mathtt{L1Topo}}$ and $\mathtt{HLT}_{\mathtt{4J12}}$ respectively. The $\mathtt{HLT}_{\mathtt{L1Topo}}$ is defined as:

Single τ_{had} triggers (STT)
`HLT_tau80_medium1_tracktwo_L1TAU60`
`HLT_tau125_medium1_tracktwo`
`HLT_tau160_medium1_tracktwo`
`HLT_tau160_medium1_tracktwo_L1TAU100`
`HLT_tau160_medium1_tracktwoEF_L1TAU100`
`HLT_tau160_mediumRNN_tracktwoMVA_L1TAU100`
Di τ_{had} triggers (DTT)
`HLT_tau35_medium1_tracktwo_tau25_medium1_tracktwo_L1TAU20IM_2TAU12IM`
`HLT_tau35_medium1_tracktwo_tau25_medium1_tracktwo`
`HLT_tau35_medium1_tracktwo_tau25_medium1_tracktwo_L1TAU20IM_2TAU12IM_4J12`
`HLT_tau35_medium1_tracktwo_tau25_medium1_tracktwo_L1DR-TAU20ITAU12I-J25`
`HLT_tau35_medium1_tracktwoEF_tau25_medium1_tracktwoEF_L1TAU20IM_2TAU12IM_4J12.0ETA23`
`HLT_tau35_medium1_tracktwoEF_tau25_medium1_tracktwoEF_L1DR-TAU20ITAU12I-J25`
`HLT_tau35_mediumRNN_tracktwoMVA_tau25_mediumRNN_tracktwoMVA_L1TAU20IM_2TAU12IM_4J12.0ETA23`
`HLT_tau35_mediumRNN_tracktwoMVA_tau25_mediumRNN_tracktwoMVA_L1DR-TAU20ITAU12I-J25`

Table 4.2: Triggers used for ATLAS Run-2 data taking in the $\tau_{\text{had}}\tau_{\text{had}}$ channel [65].

`HLT_tau30_mediumRNN_tracktwoMVA_tau20_mediumRNN_tracktwoMVA`
`_03dRAB30_L1DR-TAU20ITAU12I-J25`

and $\texttt{HLT}_{\texttt{4J12}}$ is defined as:

`HLT_tau30_mediumRNN_tracktwoMVA_tau20_mediumRNN_tracktwoMVA`
`_03dRAB_L1TAU20IM_2TAU12IM_4J12p0ETA23`.

Both chains have different L1 seed triggers and require two tau objects reconstructed using *tracktwoMVA* algorithm with a minimum p_T of 30 and 20 GeV, respectively. This pair of τ leptons should also pass the medium RNN condition. The `mediumRNN` tag in the $\texttt{HLT}_{\texttt{L1Topo}}$ and $\texttt{HLT}_{\texttt{4J12}}$ is a combination of multiple WPs in different p_T region of the tau leptons.

- Tau leptons with $p_T < 280$ GeV must pass a medium RNN ID cut.
- Tau leptons with $280 < p_T < 440$ GeV must pass a loose RNN ID cut.
- Tau leptons with $p_T > 440$ GeV are exempt from the RNN ID requirement.

In addition to this, the same tau object pair should be separated by an angle $0.3 < \Delta R < 3$ to pass $\texttt{HLT}_{\texttt{L1Topo}}$ and $0.3 < \Delta R$ to pass $\texttt{HLT}_{\texttt{4J12}}$.

A $\texttt{HLT}_{\texttt{b+}\tau}$ chain defined as :

`HLT_tau20_mediumRNN_tracktwoMVA`
`_j75c_020jvt_j50c_020jvt_j25c_020jvt_j20c_020jvt`

```
_SHARED_j20c_020jvt_bgn185_pf_ftf_presel3c20XX1c20b85
                _L1J45p0ETA21_3J15p0ETA25
```

is also proposed for 2024 Run-3 data taking and is under investigation. This HLT chain is seeded by an `L1` which requires a jet with a $p_T > 20$ GeV and $0 < \eta < 2.1$, and three jets with $p_T > 15$ GeV and $0 < \eta < 2.5$. The HLT chain requires at least one τ jet, and four **`pf_ftf`** jets, meaning the chain reconstructs particle flow [59] jets which need full-scan fast track finding (FTF). The `c` (central jet) after the jet p_T implies that the jets should satisfy a pseudorapidity condition $0 < \eta < 2.4$. The τ leptons, reconstructed using *tracktwoMVA* algorithm must have a minimum p_T of 20 GeV with a medium RNN ID WP. The four jets should have a minimum p_T of 75 GeV, 50 GeV, 25 GeV and 20 GeV, respectively, with a JVT threshold of 0.2 for all the jets. Among these jets, one must meet the *Graph Neural Network* (gn1) b-tagging WP requirement of 85% efficiency. There is also a preselection condition that is used to reduce the CPU costs for tracking, hence applied right after the L1 trigger condition. The preselection condition requires one central jet with a fast b-tagging [57] selection applied with 85% efficiency and three central jets, all with $p_T > 20$ GeV.

Chapter 5

Sample Selection and Data Exploration

The ATLAS data processing chain is a complicated procedure [66] that reconstructs detector signatures into a format that physicists can analyse. This entire process is implemented in the ATLAS bulk data processing framework called Athena [67].

The reconstruction workflow begins by obtaining event information from the detector components that are accepted by the triggers. This online byte stream information is stored in RAW format. Reconstruction algorithms are run on the RAW event and a structured Event Summary Data (ESD) format is produced. This contains all the reconstructed objects such as tracks, jets, muons, electrons, etc. It also contains low-level information including detector hits and calorimeter cells.

A subset of the ESD, termed Analysis Object Data (AOD), focuses solely on the reconstructed objects that are predominantly used in any physics analysis. AODs, often stored in ROOT [68] format, focus on physics objects such as electrons, muons, jets, or photons and provide access to their properties in a readily accessible form for physics analysis. Subsequently, this information is further condensed into analysis-specific data known as Ntuples, a more compact representation utilised for most physics analyses.

For Monte Carlo (MC) simulated events, the reconstruction process mirrors the standard procedure, with an additional digitisation step before reconstruction. Details regarding the reprocessing history, software versions, etc. that are used to produce the samples are logged via the ATLAS Metadata Interface (AMI) tags [69]. The resulting derivations are stored in various grid sites. These derivations can be accessed through Rucio, a Distributed Data Management system [70], through which all samples used in this thesis were accessed. Specifically, for trigger optimisation purposes, two types of samples are used, which will be elaborated upon in this chapter.

5.1 Monte Carlo Simulation of HH $\rightarrow b\bar{b}\tau^{+}_{\text{had}}\tau^{-}_{\text{had}}$ channel

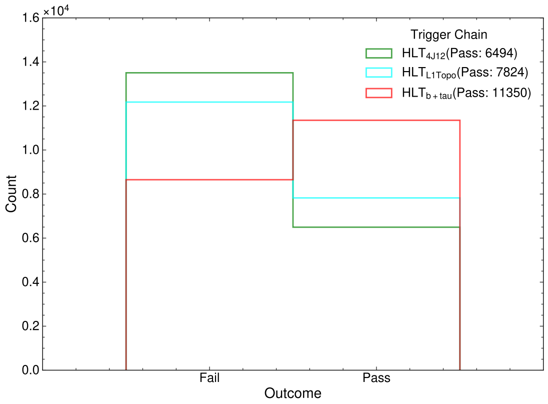
Monte Carlo (MC) methods are algorithms that use random sampling. In many cases, they provide approximate solutions for problems where exact solutions are hard to determine. In high-energy physics, MC methods are crucial for simulating particle collisions and interactions with detectors. They are used to generate detailed event data mimicking real collider and detector observations. MC data provides precise details about particles, including their identity, kinematics of particles, positions of decay vertices, and the relationship between parent and daughter particles. This truth information is essential for studying how detectors respond to specific particles and for identifying signals of new physics processes.

In this analysis, an MC sample representing the HH→ $b\bar{b}\tau^{+}\tau^{-}$ process, where both tau leptons decay hadronically, is used. This particular signal sample is identified by the string:

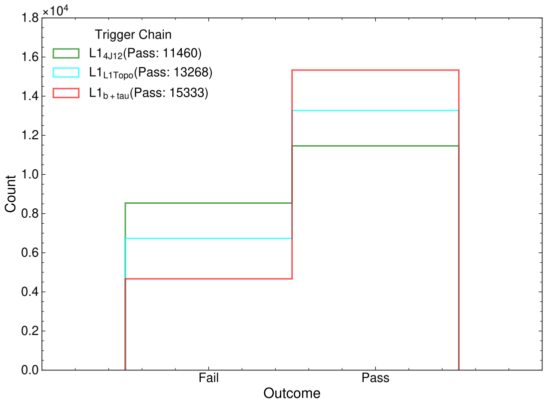
```
mc23_13p6TeV.601477.PhPy8EG_HHbbtautauHadHad_cHHH01d0.
            -recon.AOD.e8514_s4111_r14668.
```

Breaking down the string, `mc23` signifies the MC campaign number. The part `13p6TeV` stands for the centre of mass energy of the simulation. Next comes the data set number, followed by abbreviated generator names. For example, in this scenario, POWHEG [71] is used to generate events up to the parton level. The process is completed by parton shower generation, hadronization, and decay simulations using PYTHIA7 [72] and EVTGEN [73]. The specific physics process that is simulated is indicated by the subsequent part of the string, `HHbbtautauHadHad_cHHH01d0`, which represents the simulation of $HH \rightarrow b\bar{b}\tau^{+}_{\text{had}}\tau^{-}_{\text{had}}$, with the trilinear self-coupling modifier set to 1 as in SM. The dataset is in AOD format with 20000 events.

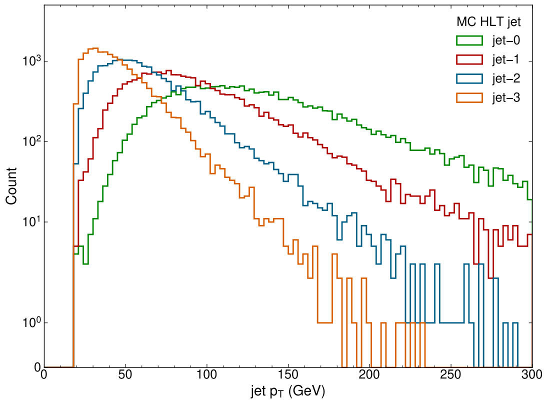
For this analysis, several key variables carry importance. These include the p_T, number of prongs and the RNN score for tau objects, the JVT, p_T, and D_b score for jet objects, the invariant mass spectrum and events accepted by various trigger chains. The distributions of these variables are shown in Figure 5.1 for the MC signal sample. The analysis uses tau objects reconstructed at HLT, sourced from the `HLT_TrigTauRecMerged_MVA` and `TauJets` collections, representing online and offline reconstructed tau objects, respectively. The objects in these collections can be matched with the $\texttt{HLT}_{\texttt{L1Topo}}$ and $\texttt{HLT}_{\texttt{4J12}}$ triggers to obtain the respective trigger-matched tau objects. The HLT jet collection used for the $\texttt{HLT}_{\texttt{b+}\tau}$ trigger is `HLT_AntiKt4EMPFlowJets_subresjesgscIS_ftf_bJets`.



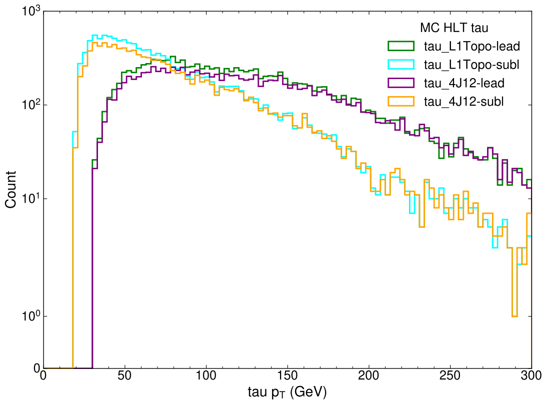
(a) Number of events that pass/fail $\texttt{HLT}_{\texttt{L1Topo}}$, $\texttt{HLT}_{\texttt{4J12}}$ and $\texttt{HLT}_{\texttt{b+}\tau}$ chains.



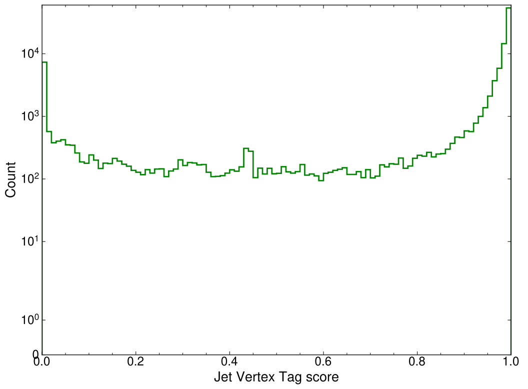
(b) Number of events that pass/fail the L1 seed of $\texttt{HLT}_{\texttt{L1Topo}}$, $\texttt{HLT}_{\texttt{4J12}}$ and $\texttt{HLT}_{\texttt{b+}\tau}$ chains.



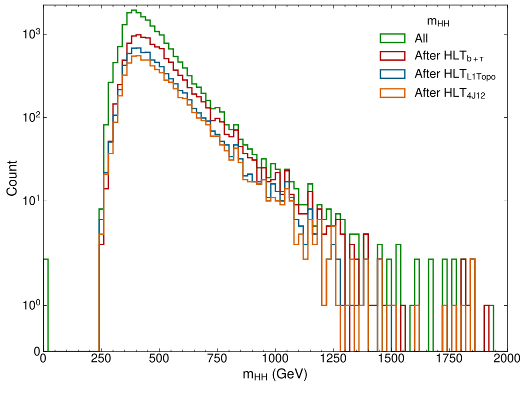
(c) Jet p_T spectra for the first four jet objects in each event.



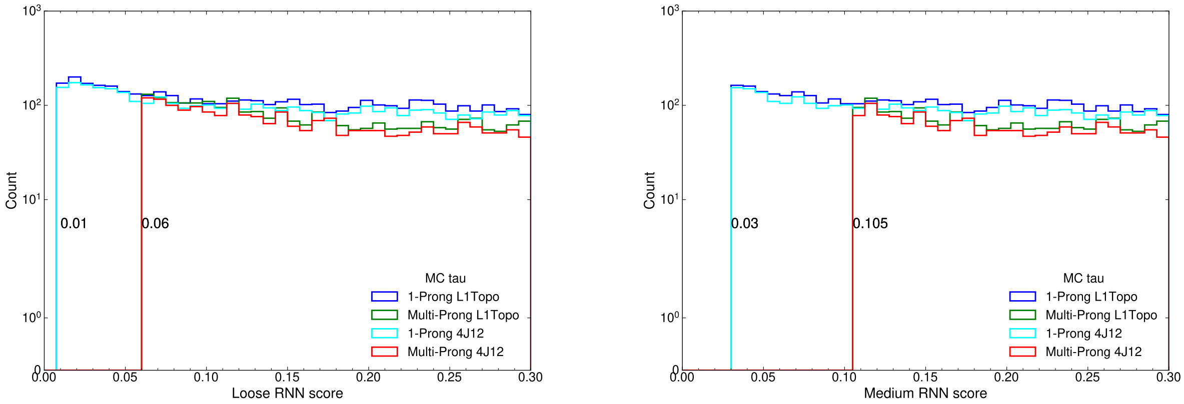
(d) τ_{p_T} spectra with objects matched to $\texttt{HLT}_{\texttt{L1Topo}}$ and $\texttt{HLT}_{\texttt{4J12}}$ chains in each event.



(e) Jet Vertex Tagger score for all HLT jets in the sample.

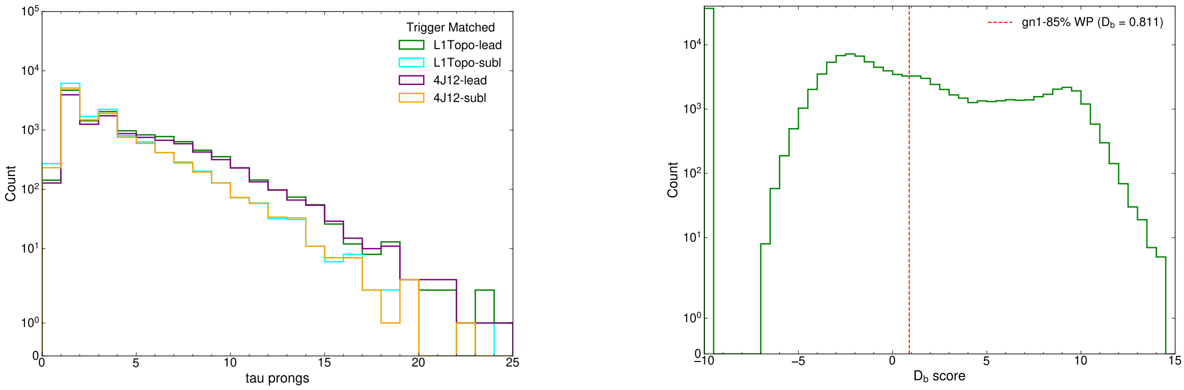


(f) Invariant truth mass of the di-Higgs system after different trigger selections.



(g) RNN score of 1-prong ($\mathrm{RNN}^{\mathrm{loose}}_{\mathrm{1p}}$) and multi-prong ($\mathrm{RNN}^{\mathrm{loose}}_{\mathrm{mp}}$) τ leptons matched to $\mathtt{HLT}_{\mathtt{L1Topo}}$ and $\mathtt{HLT}_{\mathtt{4J12}}$ chains with *loose* working point.

(h) RNN score of 1-prong ($\mathrm{RNN}^{\mathrm{medium}}_{\mathrm{1p}}$) and multi-prong ($\mathrm{RNN}^{\mathrm{medium}}_{\mathrm{mp}}$) τ leptons matched to $\mathtt{HLT}_{\mathtt{L1Topo}}$ and $\mathtt{HLT}_{\mathtt{4J12}}$ chains with *medium* working point.



(i) τ_{prong} distribution with objects matched to $\mathtt{HLT}_{\mathtt{L1Topo}}$ and $\mathtt{HLT}_{\mathtt{4J12}}$ chains in each event.

(j) D_b score of all HLT jets calculated using Eq. (4.1).

Figure 5.1: Distributions of key parameters obtained from the Monte Carlo signal sample of HH→ $b\bar{b}\tau^{+}_{\mathrm{had}}\tau^{-}_{\mathrm{had}}$ events. The tau objects are trigger matched to the $\mathtt{HLT}_{\mathtt{L1Topo}}$ and $\mathtt{HLT}_{\mathtt{4J12}}$ trigger chains.

5.2 The Enhanced Bias Dataset

Apart from the physics data stream employing the physics trigger menu, there are several other streams dedicated to other important studies in the trigger menu. To measure the anticipated rate of a given HLT algorithm, in principle an unbiased sample of events (zero bias) is needed. Most trigger algorithms aim to select rare events, requiring an extensive set of unbiased events to gather adequate statistics in the phase space where these rare events occur.

In such scenarios, Enhanced Bias (EB) triggers prove beneficial [74]. These data sets consist of events influenced solely by the L1 decision, by selecting a higher fraction of high p_T triggers and other interesting physics objects that would not be selected in a zero bias sample. This is demonstrated in the graph shown in Figure 5.2. The EB data can be inverted to generate a minimum bias dataset that preserves statistical information across all p_T scales. This inversion process involves applying a per-event weight that effectively eliminates the online prescale used for event selection.

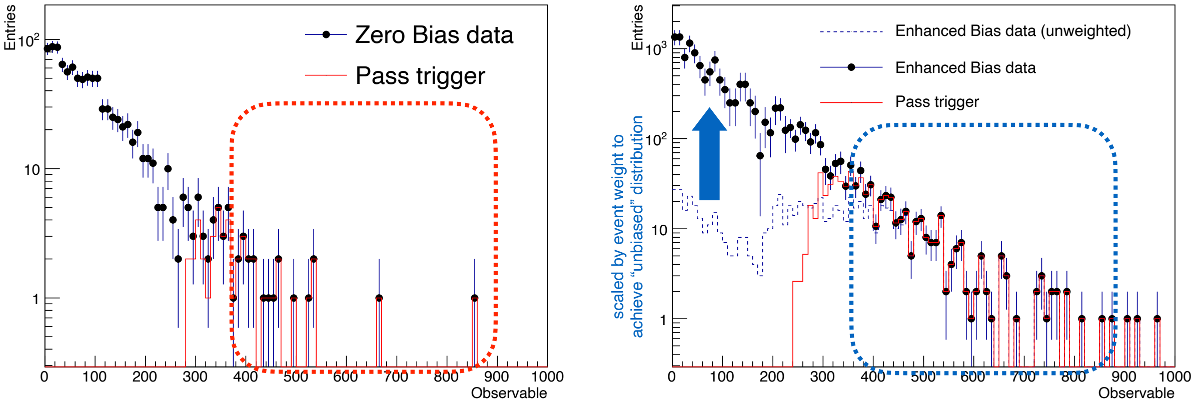


Figure 5.2: A graphical comparison of simulated zero bias and EB data. It can be seen that the EB data has better statistics in the tail regions compared to zero bias data. Adapted from graphic by © Takuya Nobe, CC BY 2.5, Offline reprocessing shifter tasks, 2023.

The collection of an EB dataset occurs in parallel to the standard data collection using the physics trigger menu. Collecting the EB dataset involves using a distinct trigger menu that comprises a variety of L1 trigger items, ranging from high-p_T primary triggers to low-p_T L1 triggers. Additionally, a random trigger item is incorporated to introduce a zero bias element, particularly for processes with very high cross sections. More information on EB triggers can be found in the Appendix B. In this analysis, we used the EB sample obtained from Run-440499[1], which took place on November 25, 2022. This RAW dataset is first reprocessed to a more analysis-friendly AOD format. During the reprocessing phase, a developer trigger menu containing physics menu items and additional test triggers chains from different physics groups was applied to these unbiased RAW events to assess CPU costs and rates for different trigger chains. The specific reprocessed AOD used for this analysis is identified by the string:

```
data22_13p6TeV.00440499.physics_EnhancedBias.recon.
-AOD.r14669_r14670_tid33426406_00
```

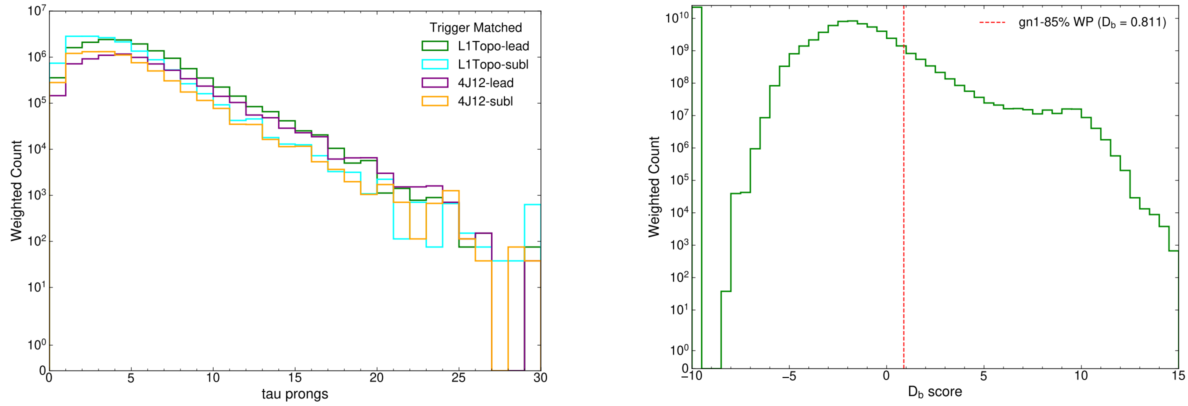
During this run, a total of 1,030,938 events, amounting to 1.4 TB, were triggered in the EB stream. Similar to the MC sample, `data22_13p6TeV` denotes the data campaign

---

[1] For further details about the run, refer to `https://atlas-runquery.cern.ch`.

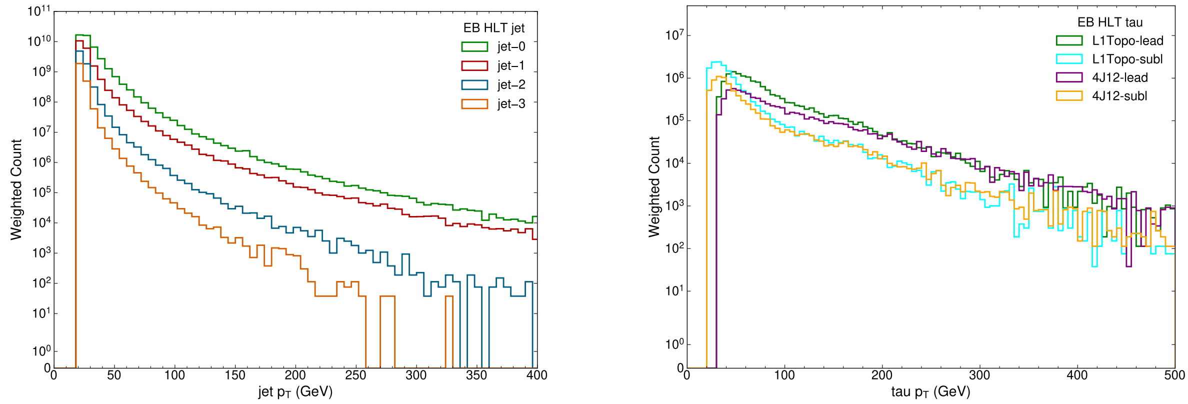
number followed by the centre of mass energy[2]. The run number corresponding to the EB data follows. This information is followed by the data stream. The subsequent part of the string denotes the data stream and the final part indicates the file type, i.e., `AOD`.

The Enhanced Bias sample includes all the information found in an MC sample, except the truth information of various particles. As of Release 21, the b-tagging information link was removed from the `HLT_AntiKt4EMPFlowJets_subresjesgscIS_ftf_bJets` collection. To access the b-tagging information, a separate collection, `HLT_AntiKt4EMPFlowJets_subresjesgscIS_ftf_BTagging` was used[3]. Apart from this, all other variables can be accessed similarly to an MC sample. The distributions of variables that are relevant for this analysis such as p_T, the number of prongs, and the RNN score for tau objects, the JVT, p_T and D_b score for jet objects, are shown in Figure 5.3. Of 1,030,938 unweighted events, the number of events that satisfied the L1 seed criteria of the $\texttt{HLT}_{\texttt{L1Topo}}$, $\texttt{HLT}_{\texttt{4J12}}$ and $\texttt{HLT}_{\texttt{b+}\tau}$ chains were 30512, 18100 and 44433 events respectively. A



(a) τ_{prong} distribution with objects matched to $\texttt{HLT}_{\texttt{L1Topo}}$ and $\texttt{HLT}_{\texttt{4J12}}$ chains in each event.

(b) D_b score of all HLT jets calculated using Eq. (4.1).



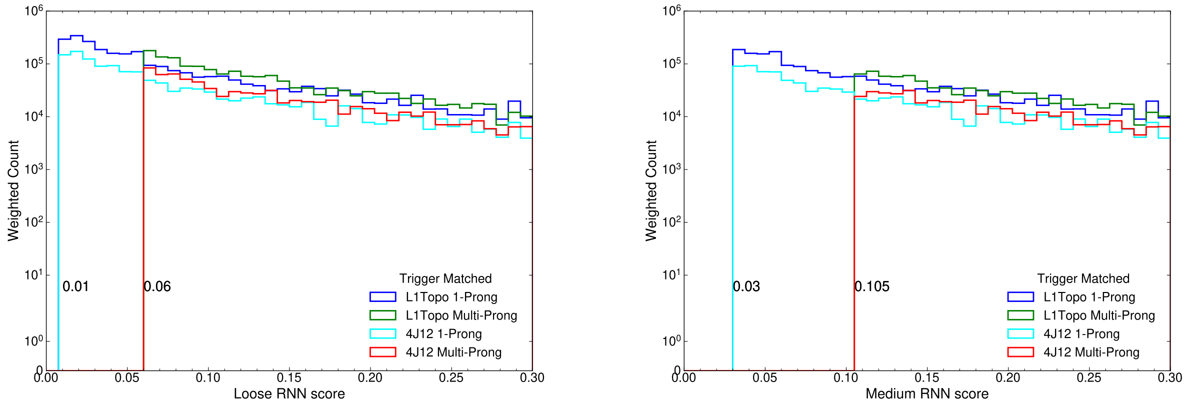
(c) Jet p_T spectra for the first four jets in each event.

(d) τ_{p_T} spectra with objects matched to $\texttt{HLT}_{\texttt{L1Topo}}$ and $\texttt{HLT}_{\texttt{4J12}}$ chains in each event.

---

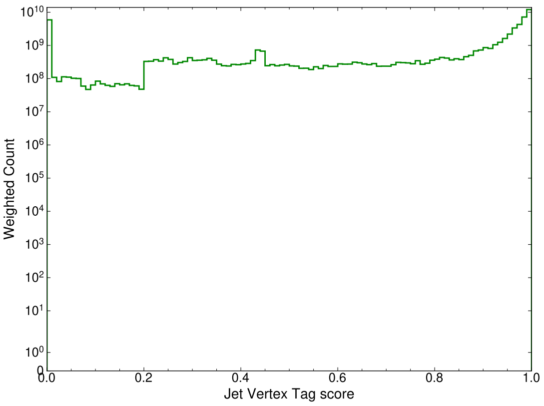
[2] For more information about the reprocessing, see https://its.cern.ch/jira/browse/ATR-27517.

[3] For more information on this issue, see https://its.cern.ch/jira/browse/ATR-26904.



(e) RNN score of 1-prong ($\mathrm{RNN}_{\mathrm{1p}}^{\mathrm{loose}}$) and multi-prong ($\mathrm{RNN}_{\mathrm{mp}}^{\mathrm{loose}}$) τ leptons matched to $\mathtt{HLT}_{\mathtt{L1Topo}}$ and $\mathtt{HLT}_{\mathtt{4J12}}$ chains with *loose* working point.

(f) RNN score of 1-prong ($\mathrm{RNN}_{\mathrm{1p}}^{\mathrm{medium}}$) and multi-prong ($\mathrm{RNN}_{\mathrm{mp}}^{\mathrm{medium}}$) τ leptons matched to $\mathtt{HLT}_{\mathtt{L1Topo}}$ and $\mathtt{HLT}_{\mathtt{4J12}}$ chains with *medium* working point.



(g) Jet Vertex Tagger score for all HLT jets.

Figure 5.3: Distributions of key parameters obtained from the Enhanced Bias data sample from Run-440499. The tau objects are trigger matched to the $\mathtt{HLT}_{\mathtt{L1Topo}}$ and $\mathtt{HLT}_{\mathtt{4J12}}$ trigger chains.

total of 430 events satisfied the $\mathtt{HLT}_{\mathtt{L1Topo}}$ criteria, 161 events satisfied the $\mathtt{HLT}_{\mathtt{4J12}}$ criteria, and 1258 events satisfied the $\mathtt{HLT}_{\mathtt{b+\tau}}$ criteria.

Chapter 6

High Level Trigger Optimization for HH $\to b\bar{b}\tau^{+}_{\text{had}}\tau^{-}_{\text{had}}$ channel

This chapter focuses on techniques for emulating and improving trigger performance. Various HLT parameters of the $\texttt{HLT}_{\texttt{L1Topo}}$, $\texttt{HLT}_{\texttt{4J12}}$ and $\texttt{HLT}_{\texttt{b+}\tau}$ chains are varied and emulated to determine the efficiency and rates of the trigger chain, a quantity that is important to evaluate the performance of the trigger chain. For this purpose, a trigger emulation framework[1] has been developed. This framework is currently being adapted by the Run-3 HH→ $b\bar{b}\tau^{+}_{\text{had}}\tau^{-}_{\text{had}}$ trigger analysis group to emulate the rates of different trigger chains. At the end of this chapter, a set of optimal parameters is presented for each of those chains.

6.1 Efficiency and Rate Emulation Procedures

The efficiency of a trigger denotes its ability to select events that are potentially interesting for further physics analysis. It is usually quantified as the fraction of signal events that are accepted by the trigger chain. For instance, $\texttt{HLT}_{\texttt{L1Topo}}$, $\texttt{HLT}_{\texttt{4J12}}$ and $\texttt{HLT}_{\texttt{b+}\tau}$ chains are introduced in the menu to study the HH→ $b\bar{b}\tau^{+}_{\text{had}}\tau^{-}_{\text{had}}$ channel. Consequently, the efficiencies of these chains depend on the fraction of HH→ $b\bar{b}\tau^{+}_{\text{had}}\tau^{-}_{\text{had}}$ events that they manage to accept.

To calculate the efficiency of the di-tau triggers, a selection cut is first applied to the truth-matched offline tau objects to simulate the L1 condition. In the case of $\texttt{HLT}_{\texttt{b+}\tau}$triggers, the HH→ $b\bar{b}\tau^{+}_{\text{had}}\tau^{-}_{\text{had}}$ trigger analysis group made the choice to use the absolute efficiency of the trigger, which avoids the selection process prior to efficiency calculation. The selection criteria for the $\texttt{HLT}_{\texttt{L1Topo}}$ and $\texttt{HLT}_{\texttt{4J12}}$ chains are as follows:

---

[1] https://gitlab.cern.ch/asudhaka/hlt_bbtautau_trigger_emulation

1. The event should have at least two offline truth matched tau leptons.
2. The leading and subleading offline truth matched tau objects should have a minimum p_T of 20 GeV and 12 GeV, respectively.
3. The leading and subleading offline truth matched tau objects should pass the loose RNN criteria.

The HLT emulation is then performed to determine the fraction of events that meet the HLT criteria. To accept an event, the $\texttt{HLT}_{\texttt{L1Topo}}$ and $\texttt{HLT}_{\texttt{4J12}}$ emulations will verify the following criteria for events reconstructed in the HLT stage:

1. The event should contain at least two online tau objects.
2. These tau objects should have a maximum of 3 prongs to mimic the TracktwoMVA condition.
3. The leading and subleading tau objects should satisfy a minimum p_T requirement of $<\tau^{\rm lead}_{p_T}>$ and $<\tau^{\rm subl}_{p_T}>$, respectively.
4. The same pair of tau objects should either have a minimum 1-prong medium RNN threshold of $<\mathrm{RNN}^{\rm 1p}_{\rm medium}>$, a multi-prong medium RNN threshold of $<\mathrm{RNN}^{\rm mp}_{\rm medium}>$, or a zero-prong medium RNN threshold of 0.35[2].
5. The same pair of tau objects should be separated by an angle (ΔR) ranging between 0.3 and 3 radians for the $\texttt{HLT}_{\texttt{L1Topo}}$chain and more than 0.3 radians for the $\texttt{HLT}_{\texttt{4J12}}$ chain.

Both the $\texttt{HLT}_{\texttt{L1Topo}}$ and $\texttt{HLT}_{\texttt{4J12}}$ triggers have their existing threshold values set for $<\tau^{\rm lead}_{p_T}>$ and $<\tau^{\rm subl}_{p_T}>$ at 30 GeV and 20 GeV, respectively. The $<\mathrm{RNN}^{\rm 1p}_{\rm medium}>$ and $<\mathrm{RNN}^{\rm mp}_{\rm medium}>$, thresholds are set to 0.03 and 0.105 respectively.

For an event to be accepted by the emulated $\texttt{HLT}_{\texttt{b+}\tau}$ chain, the following conditions for events reconstructed in the HLT stage should be satisfied:

1. The event should satisfy the L1 condition of the $\texttt{HLT}_{\texttt{b+}\tau}$ chain.
2. The event should have at least one tau object (τ^0) and 4 jet objects (j^0, j^1, j^2, j^3).
3. The tau object should have a minimum p_T requirement of $<\tau^{0}_{p_T}>$ GeV.
4. The tau object should pass the <`RNN-WP`> criteria and should have a maximum of three prongs.

---

[2] The medium RNN condition in $\texttt{HLT}_{\texttt{L1Topo}}$ and $\texttt{HLT}_{\texttt{4J12}}$ chains require different WP in different p_T ranges as described in subsection 4.4.1.

5. The four jet objects should satisfy a minimum p_T requirement of $<j^0_{p_T}>$, $<j^1_{p_T}>$, $<j^2_{p_T}>$, and $<j^3_{p_T}>$, respectively.

6. The four jets should have a minimum JVT score of 0.2.

7. One of the four jets should have a GN1 WP efficiency of <`GN1-WP %`>.

In the proposed version of the $\texttt{HLT}_{\texttt{b+}\tau}$ chain, the p_T thresholds for the four jet objects were 75 GeV, 50 GeV, 25 GeV and 20 GeV, and the p_T threshold for the tau object was 20 GeV. The tau objects had to satisfy the medium RNN criteria and one of the four jets should have an 85% efficiency for the GN1 b-tagging WP.

Using the number of events that satisfy these criteria, the efficiency ϵ can be computed using the formula:

$$\epsilon = \frac{\text{Number of HLT pass events}}{\text{Number of HLT input events}} \tag{6.1}$$

The (background) rate refers to the frequency of background events selected by a trigger chain, typically measured in Hz, and depends on the luminosity. Assessing rates for a trigger chain that is not present in the developer menu during reprocessing is time-intensive. To include a chain in the developer menu, a comprehensive efficiency study and rationale for its addition must be provided. This menu will subsequently be used in the forthcoming weekly reprocessing, enabling the acquisition of rates from it.

Given that this study extensively scans the trigger parameter space, evaluating the rates for thousands of trigger chains becomes impractical using the conventional approach. Therefore, to determine the rates, a rate emulation was performed using the EB sample for the variants of $\texttt{HLT}_{\texttt{L1Topo}}$, $\texttt{HLT}_{\texttt{4J12}}$ and $\texttt{HLT}_{\texttt{b+}\tau}$ chains. The steps for rate emulation are as follows:

1. Confirm the presence of a loose variant of the trigger chain that is emulated in the EB reprocessing trigger menu. This chain, called the standard trigger, should be looser than all triggers to be emulated to obtain an unbiased estimate of the rate and also serves as a reference point for standardising the obtained rates. The rate of the standard trigger chain, denoted as Rate_{std}, should be accessible from EB reprocessing.

2. Select events passing through the standard trigger chain using the corresponding flag. Calculate the total number of weighted events that pass through the standard trigger chain (N_{std}).

3. Implement HLT emulation for stricter variations of the corresponding standard trigger chain and count the weighted number of events that meet the emulation

criteria ($\mathrm{N_{emu}}$).

4. Compute the emulated rate ($\mathrm{Rate_{emu}}$) using the formula:

$$\mathrm{Rate_{emu}} = \mathrm{N_{emu}} \times \frac{\mathrm{Rate_{std}}}{\mathrm{N_{std}}} \tag{6.2}$$

6.2 Emulated Efficiencies and Rates Validation

Ensuring the accuracy of the emulated efficiencies and rates in comparison to the actual trigger performance is crucial. The emulated efficiencies of the trigger chains have been validated by comparing them with efficiencies calculated using the associated trigger flag (referred to as standard efficiency) to the $\mathtt{HLT_{L1Topo}}$ chain, $\mathtt{HLT_{4J12}}$ chain, and various variants of the $\mathtt{HLT_{b+\tau}}$ chain. The results of this comparison are presented in Figure 6.1.

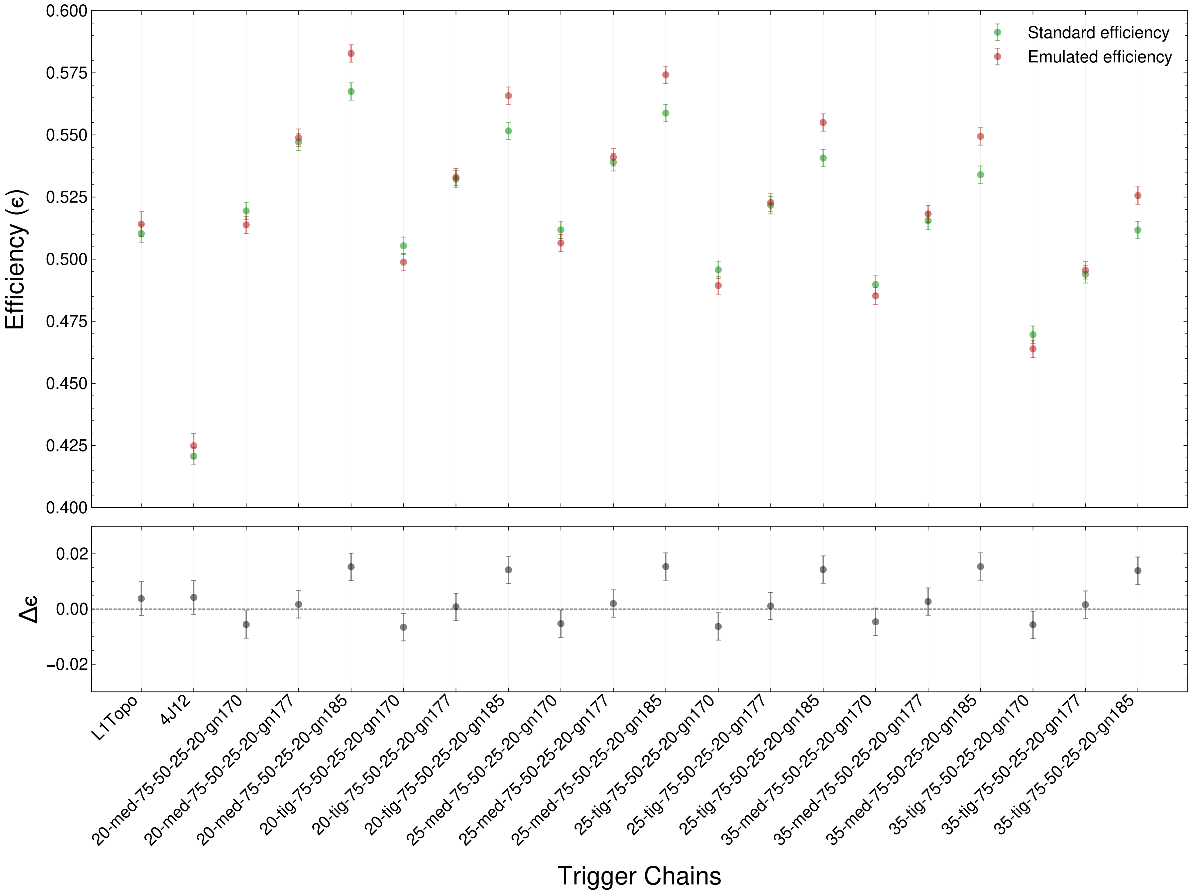


Figure 6.1: The plot illustrates the comparison between emulated efficiencies and standard efficiencies. The lower pane shows the difference in the emulated and standard efficiency. The first two trigger chains correspond to $\mathtt{HLT_{L1Topo}}$ and $\mathtt{HLT_{4J12}}$, respectively. The subsequent chains represent variants of the $\mathtt{HLT_{b+\tau}}$ chain, labeled in the format: $<\tau^{\mathrm{lead}}_{p_T}>$-$<\tau^{\mathrm{lead}}_{\mathrm{RNN\text{-}ID}}>$-$<j^{0}_{p_T}>$-$<j^{1}_{p_T}>$-$<j^{2}_{p_T}>$-$<j^{3}_{p_T}>$-$<\mathtt{GN1\text{-}WP}>$.

To validate the accuracy of the emulated rates, the emulated rates of the $\texttt{HLT}_{\texttt{L1Topo}}$, $\texttt{HLT}_{\texttt{4J12}}$ and some variants of $\texttt{HLT}_{\texttt{b+}\tau}$ chains which were on the EB reprocessing trigger menu were compared to the standard reprocessing rates. The results are shown in Figure 6.2.

It can be seen that the emulated efficiencies and rates are closely aligned with the standard efficiencies and rates for the $\texttt{HLT}_{\texttt{L1Topo}}$ and $\texttt{HLT}_{\texttt{4J12}}$ trigger chains. However, the deviations were significantly higher in variants of the $\texttt{HLT}_{\texttt{b+}\tau}$ trigger chain, especially those that used the loosest b-tagging GN1 WP (`gn185`). This discrepancy likely originated from challenges in emulating the preselection condition required by the $\texttt{HLT}_{\texttt{b+}\tau}$ chain. To obtain a rough approximation of the efficiency, these preselection conditions were ignored. This omission could explain the deviations observed in the $\texttt{HLT}_{\texttt{b+}\tau}$ chains, particularly in the loosest chains. These discrepancies are less noticeable in tighter chains due to their potential overlaps with the preselection condition. Therefore, the emulation results for the $\texttt{HLT}_{\texttt{b+}\tau}$ chain should be interpreted with caution.

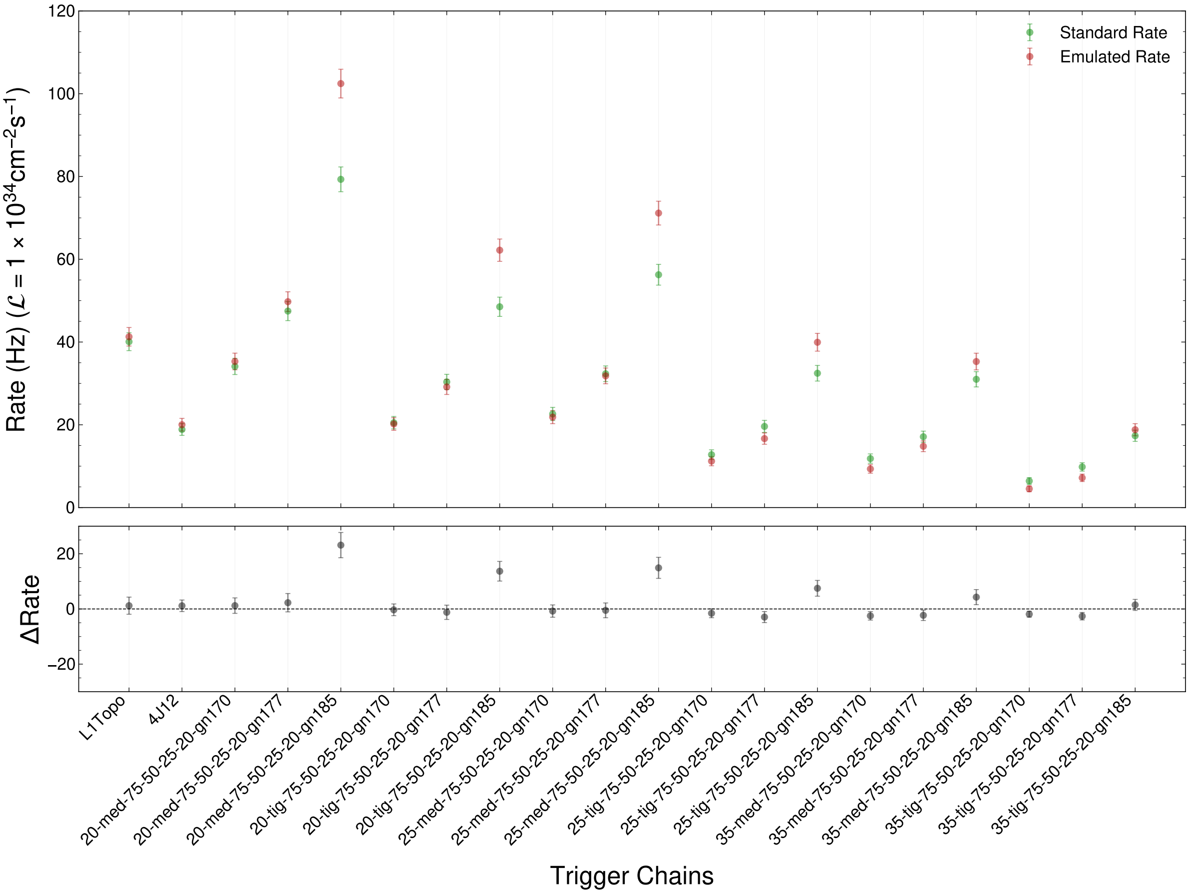


Figure 6.2: The plot illustrates the comparison between emulated rates and standard rates. The lower pane shows the difference in the emulated and standard rate. The first two trigger chains correspond to $\texttt{HLT}_{\texttt{L1Topo}}$ and $\texttt{HLT}_{\texttt{4J12}}$, respectively. The subsequent chains represent variants of the $\texttt{HLT}_{\texttt{b+}\tau}$ chain, labeled in the format: $<\tau_{p_T}^{\text{lead}}>$-$<\tau_{\text{RNN-ID}}^{\text{lead}}>$-$<j_{p_T}^{0}>$-$<j_{p_T}^{1}>$-$<j_{p_T}^{2}>$-$<j_{p_T}^{3}>$-$<\texttt{GN1-WP}>$.

6.3 Investigating Emulation Results

Trigger emulation is then performed for the $\texttt{HLT}_{\texttt{L1Topo}}$, $\texttt{HLT}_{\texttt{4J12}}$ and $\texttt{HLT}_{\texttt{b+}\tau}$ chains. A total of 1128 variants of the $\texttt{HLT}_{\texttt{b+}\tau}$ chain and 1375 variants of the $\texttt{HLT}_{\texttt{L1Topo}}$ and $\texttt{HLT}_{\texttt{4J12}}$ chains were emulated. The parameters that can be optimised, their current thresholds, and the thresholds with which they are emulated are listed in Table 6.1.

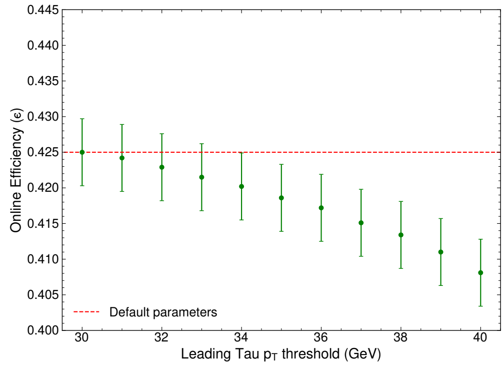
Trigger chain	Parameter	Current Threshold	Emulated threshold
$\texttt{HLT}_{\texttt{L1Topo}}$ and $\texttt{HLT}_{\texttt{4J12}}$	$\tau^{\mathrm{lead}}_{p_T}$	`30 GeV`	`[30,40]` in steps of `1 GeV`
	$\tau^{\mathrm{subl}}_{p_T}$	`20 GeV`	`[20,22]` in steps of `0.5 GeV`
	$\mathrm{RNN}^{\mathrm{medium}}_{\mathrm{1p}}$	`0.03`	`[0.01,0.05]` in steps of `0.01`
	$\mathrm{RNN}^{\mathrm{medium}}_{\mathrm{mp}}$	`0.105`	`[0.085,0.125]` in steps of `0.01`
$\texttt{HLT}_{\texttt{b+}\tau}$	$\tau^{\mathrm{lead}}_{p_T}$	`20 GeV`	`[35,30,25,20]`
	$\tau^{\mathrm{lead}}_{\mathrm{RNN}}$	`Medium`	`[Medium,Tight]`
	$j^{0}_{p_T}$	`75 GeV`	`[40,60,75,80,100]`
	$j^{1}_{p_T}$	`50 GeV`	`[30,40,50,60,70]`
	$j^{2}_{p_T}$	`25 GeV`	`[20,25,30]`
	$j^{3}_{p_T}$	`20 GeV`	`[20,25,30]`
	$\mathrm{GN1}_{\mathrm{WP}}$	`85%`	`[70%,77%,85%]`

Table 6.1: List of all the trigger parameters for the $\texttt{HLT}_{\texttt{L1Topo}}$, $\texttt{HLT}_{\texttt{4J12}}$, and $\texttt{HLT}_{\texttt{b+}\tau}$ chains, their current thresholds, and emulated thresholds.

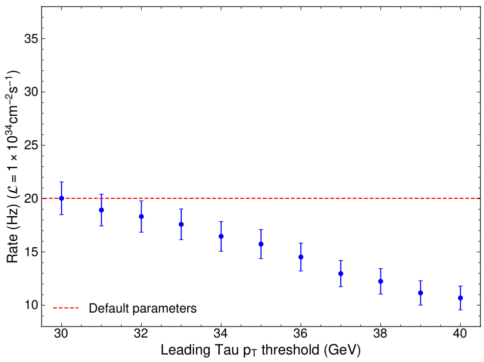
Di-Tau Triggers

The default rate for the emulated $\texttt{HLT}_{\texttt{L1Topo}}$ and $\texttt{HLT}_{\texttt{4J12}}$ chains are 41.27 ± 2.24 Hz and 20.02 ± 1.53 Hz respectively. The rates are determined for the luminosity of $\mathcal{L} = 1 \times 10^{34}\ \mathrm{cm}^{-2}\mathrm{s}^{-1}$. The default efficiency for the emulated $\texttt{HLT}_{\texttt{L1Topo}}$ and $\texttt{HLT}_{\texttt{4J12}}$ chains are 51.41 ± 0.50 % and 42.5 ± 0.47 % respectively. The efficiency and rates of the $\texttt{HLT}_{\texttt{L1Topo}}$ and $\texttt{HLT}_{\texttt{4J12}}$ chains are influenced by changes in trigger parameters such as p_T thresholds of tau objects and thresholds for 1-prong and multi-prong RNN. Upon examination of the $\texttt{HLT}_{\texttt{L1Topo}}$ and $\texttt{HLT}_{\texttt{4J12}}$ chains, it was evident that both chains exhibit similar trends in efficiency and rates. This alignment is expected given their similar selection criteria with the exception in their L1 seed chains and the angular separation condition for pairs of tau objects. For the $\texttt{HLT}_{\texttt{4J12}}$ chain, these variations are shown in Figure 6.3. The results for the $\texttt{HLT}_{\texttt{L1Topo}}$ chains are available in Appendix C.1 for further comparative analysis.

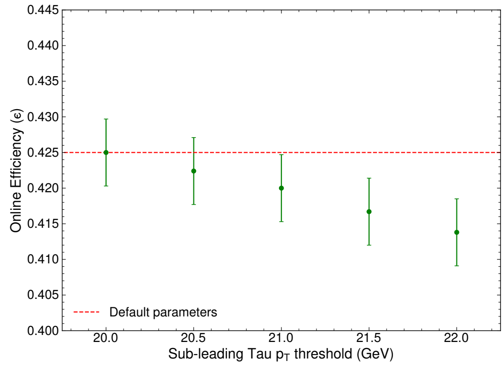
Upon analysing the results, it was evident that varying the leading p_T threshold demonstrates a comparatively smaller impact on efficiency compared to varying the subleading p_T threshold of tau objects. At the same time, the variations in rates are relatively smaller with changes in the subleading p_T threshold. Compared to the 1-prong medium RNN threshold, changes to the multi-prong medium RNN thresholds exhibit a lesser impact on both efficiency and rate.



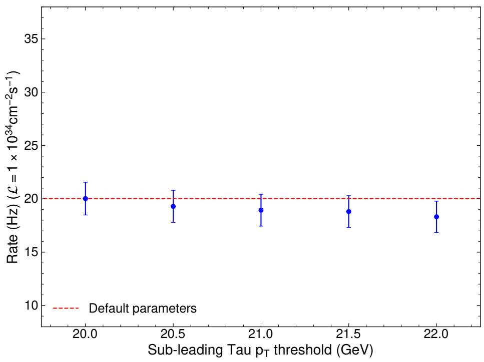
(a) Correlation between efficiency and leading p_T threshold requirement of tau objects.



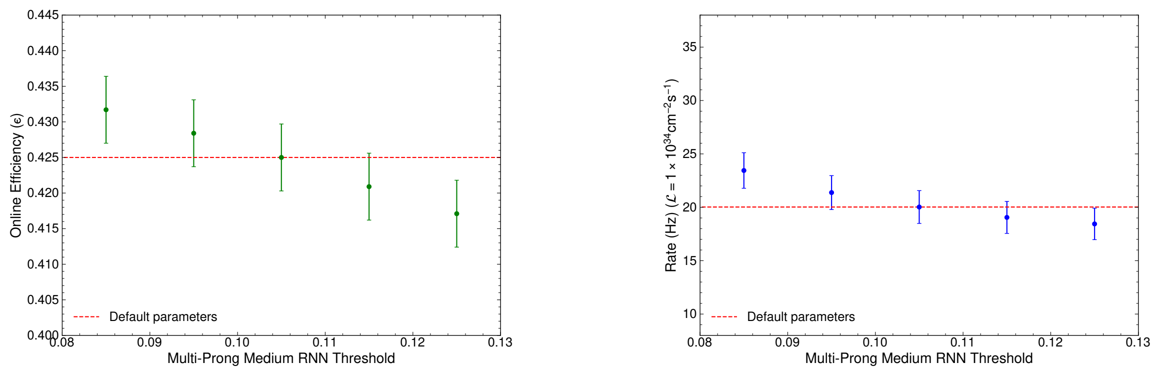
(b) Correlation between rate and leading p_T threshold requirement of tau objects.



(c) Correlation between efficiency and subleading p_T threshold requirement of tau objects.

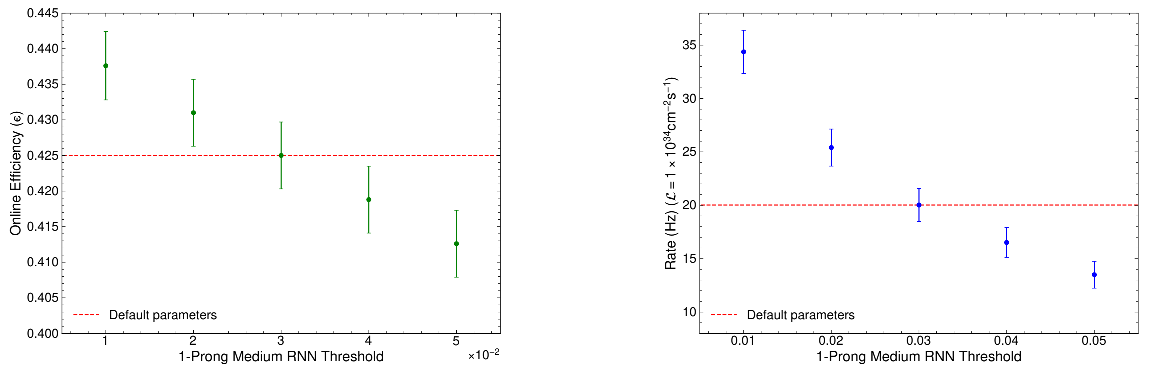


(d) Correlation between rate and subleading p_T threshold requirement of tau objects.



(e) Correlation between efficiency and the multi-prong medium RNN threshold requirement of tau objects.

(f) Correlation between rate and the multi-prong medium RNN threshold requirement of tau objects.



(g) Correlation between efficiency and the 1-prong medium RNN threshold requirement of tau objects .

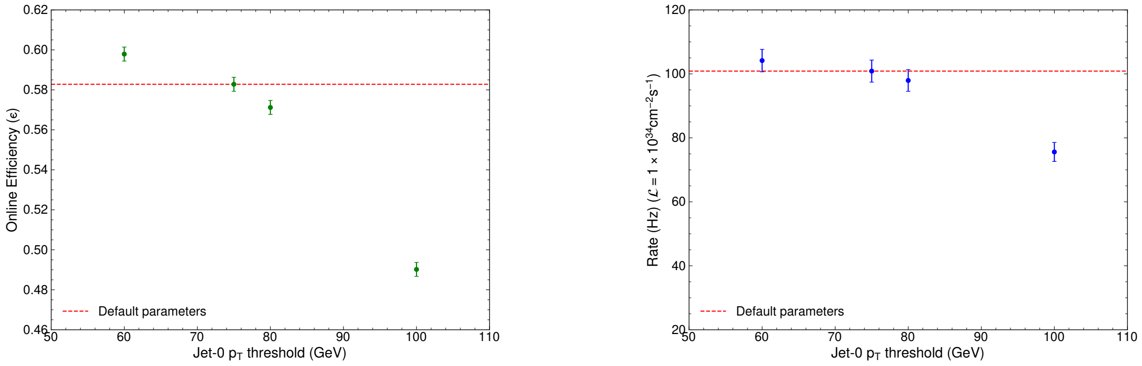
(h) Correlation between rate and the 1-prong medium RNN threshold requirement of tau objects .

Figure 6.3: Comparative analysis showing efficiency variations on the left and rate variation on the right, with respect to individual trigger parameters. The red dashed line denotes the efficiency of the $\mathtt{HLT}_{\mathtt{4J12}}$ chain with default parameters. All parameters, except the variable under observation, remain constant. The y-axis range for the rates was constrained between 8 Hz and 37 Hz. Meanwhile, for efficiencies, the y-axis range was constrained between 0.4 and 0.445.

$b + \tau$ Triggers

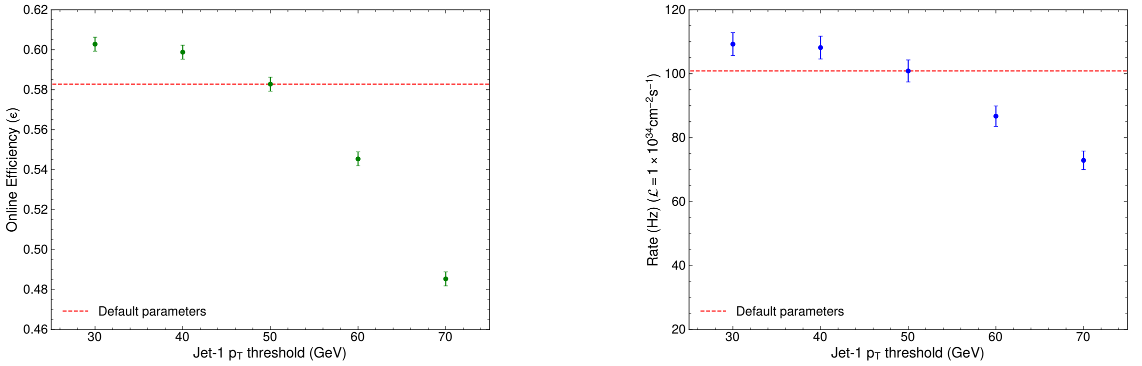
The default rate for the emulated $\mathtt{HLT}_{\mathtt{b+}\tau}$ chain is 100.36 ± 3.44 Hz. The rates are determined for the luminosity of $\mathcal{L} = 1 \times 10^{34}$ $\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The default efficiency for the emulated $\mathtt{HLT}_{\mathtt{b+}\tau}$ chain is 58.28 ± 0.35 %. In the case of $\mathtt{HLT}_{\mathtt{b+}\tau}$ chain, the influence of the p_T threshold of the three leading jet objects, the p_T threshold and the RNN WP of the tau object and the GN1-WP thresholds on the efficiencies and rates was examined and visualised in Figure 6.4.

Upon analysis, it was observed that variations in the p_T threshold of the second jet object had a negligible impact on both efficiency and rate compared to other variables. The p_T thresholds of the leading jet (j^0) and the subleading jet (j^1) objects notably influenced efficiency, while the impact of the p_T thresholds of tau objects was comparatively lower. In terms of rates, the trends reversed. The p_T thresholds of the tau objects significantly affected the rate compared to those of the jet objects. Additionally, both the GN1 WP and the RNN WP for jet and tau objects displayed a relatively substantial impact on the rate, and the RNN WP of tau objects has a comparatively lower influence on efficiency compared to the GN1 WP.



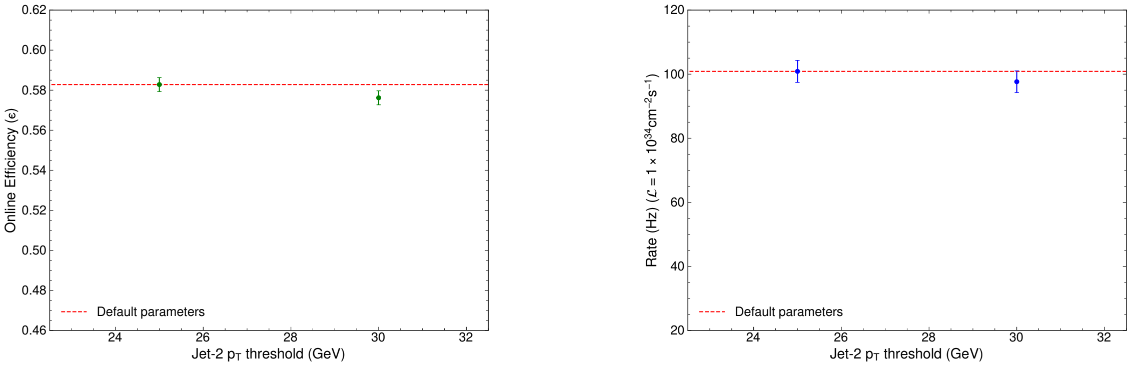
(a) Correlation between efficiency and the p_T threshold requirement of zeroth jet object.

(b) Correlation between rate and the p_T threshold requirement of zeroth jet object.



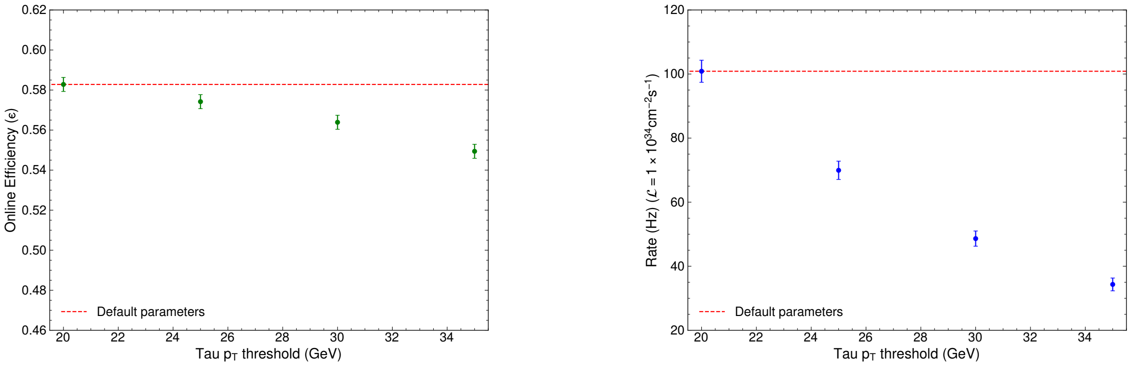
(c) Correlation between efficiency and the p_T threshold requirement of first jet object.

(d) Correlation between rate and the p_T threshold requirement of first jet object.



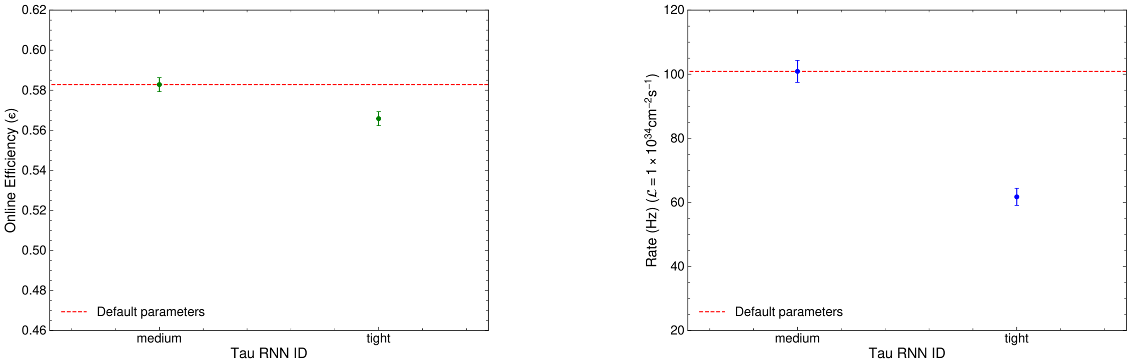
(e) Correlation between efficiency and the p_T threshold requirement of second jet object.

(f) Correlation between rate and the p_T threshold requirement of second jet object.



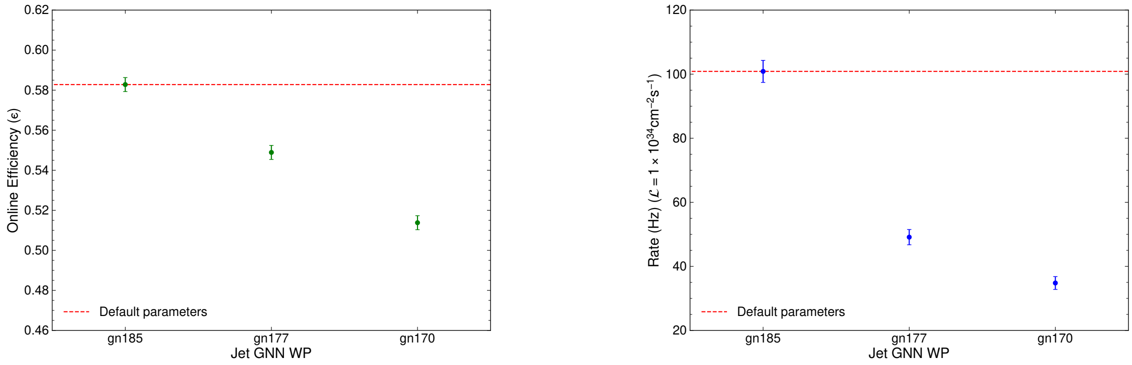
(g) Correlation between efficiency and the p_T threshold requirement of tau object.

(h) Correlation between rate and the p_T threshold requirement of tau object.



(i) Correlation between efficiency and RNN ID requirement of tau object.

(j) Correlation between rate and RNN ID requirement of tau object.



(k) Correlation between efficiency and GN1 WP requirement of the jet.

(l) Correlation between rate and GN1 WP requirement of the jet.

Figure 6.4: Comparative analysis showing efficiency variations on the left and rate variation on the right, with respect to individual trigger parameters. The red dashed line denotes the efficiency of the $\texttt{HLT}_{\texttt{b+}\tau}$ chain with the default parameters. All parameters, except the variable under observation, remain constant. The y-axis range for the rates was constrained between 20 Hz and 120 Hz. Meanwhile, for efficiencies, the y-axis range was constrained between 0.46 and 0.62.

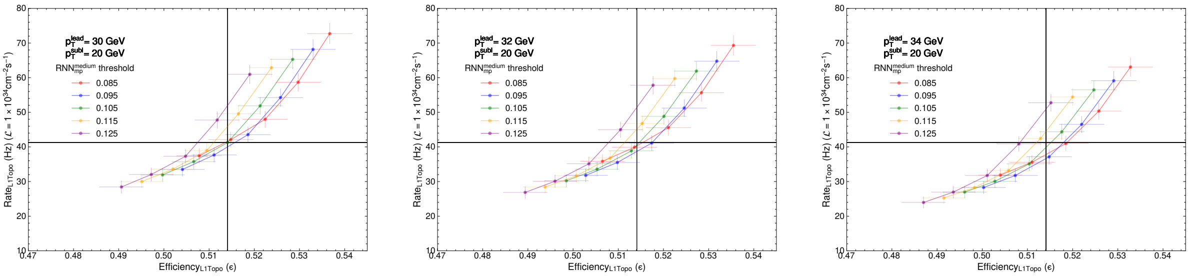
6.4 Optimal Parameters

To optimise a trigger chain, it is necessary to determine the set of trigger parameters that has better signal acceptance and background rejection. Analysing the efficiency versus rate plot significantly streamlines this task. The efficiency versus rate plots representing all the parameter combinations for the $\texttt{HLT}_{\texttt{L1Topo}}$ $\texttt{HLT}_{\texttt{4J12}}$ and $\texttt{HLT}_{\texttt{b+}\tau}$ triggers are available in Appendix C.2. Since multiple parameters are involved for each trigger, the optimisation process becomes more difficult. Hence, the variables are systematically fixed one at a time, carefully examined in a step-by-step fashion. Further details on this process are elaborated upon in this section.

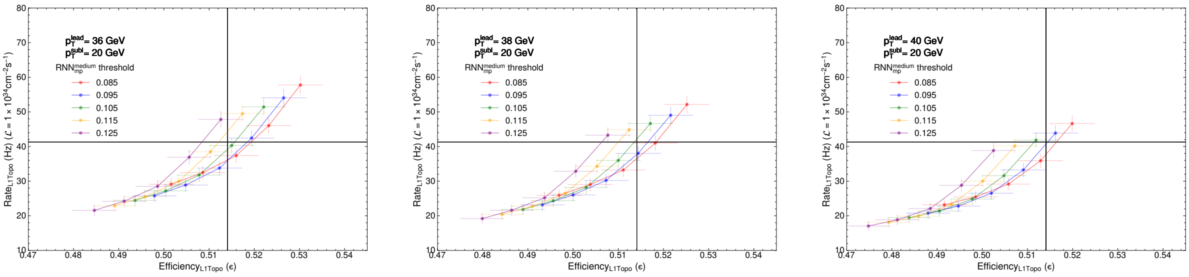
6.4.1 $\texttt{HLT}_{\texttt{L1Topo}}$ Chain

In the context of $\texttt{HLT}_{\texttt{L1Topo}}$, the p_T threshold for the subleading tau object has been set to the loosest point at 20 GeV. This decision was made due to its minimal effect on the trigger rate, with a marginally higher impact on efficiency. The remaining parameters, i.e. the leading p_T threshold of the tau objects, the multi-prong medium RNN thresholds and the 1-prong medium RNN threshold, were varied. The results are depicted in Figure 6.5.

The outcomes clearly demonstrate that the contour characterised by the loosest multi-prong medium RNN threshold of 0.085 has better efficiencies and manageable rates com-



(a) $\tau_{p_T}^{\text{lead}} = 30$ GeV (b) $\tau_{p_T}^{\text{lead}} = 32$ GeV (c) $\tau_{p_T}^{\text{lead}} = 34$ GeV



(d) $\tau_{p_T}^{\text{lead}} = 36$ GeV (e) $\tau_{p_T}^{\text{lead}} = 38$ GeV (f) $\tau_{p_T}^{\text{lead}} = 40$ GeV

Figure 6.5: The plots show the efficiencies and rates of the $\texttt{HLT}_{\texttt{L1Topo}}$ trigger when altering the thresholds for the leading p_T, $\text{RNN}^{\text{medium}}_{\text{1p}}$, and $\text{RNN}^{\text{medium}}_{\text{mp}}$ associated with the tau objects. The p_T^{subl} is set at a constant 20 GeV. Each plot varies the $\tau_{p_T}^{\text{lead}}$ threshold. The intersection point of the horizontal and vertical lines in each plot represents the rate and efficiency of the default variant of the $\texttt{HLT}_{\texttt{L1Topo}}$ trigger. Different coloured contours on each plot indicate different multi-prong thresholds, while different points on the contour correspond to various 1-prong WPs. The y-axis range for the rates was constrained between 10 Hz and 80 Hz. Meanwhile, for efficiencies, the x-axis range was constrained between 0.470 and 0.545.

pared to other multi-prong medium RNN thresholds. Hence, the $\text{RNN}^{\text{medium}}_{\text{1p}}$ threshold was set at 0.085. This is illustrated in Figure 6.6. As a result of this analysis, a specific set of parameters that have been considered optimal are detailed in Table 6.2.

Marker	$\text{RNN}^{\text{medium}}_{1p}$	$\text{RNN}^{\text{medium}}_{mp}$	$\tau_{p_T}^{\text{lead}}$ (GeV)	$\tau_{p_T}^{\text{subl}}$ (GeV)	Efficiency (ϵ)	Rate (Hz)
▲	0.02	0.085	34	20	52.58 ± 0.50 %	50.35 ± 2.48
▼	0.02	0.085	36	20	52.32 ± 0.50 %	46.03 ± 2.37
▲	0.03	0.085	34	20	51.85 ± 0.50 %	40.98 ± 2.23
Default	0.03	0.105	30	20	51.41 ± 0.50 %	41.27 ± 2.24

Table 6.2: List of optimal parameters for the $\texttt{HLT}_{\texttt{L1Topo}}$ trigger chain, and their rates and efficiencies. The default set of parameters is provided at the bottom along with their rates and efficiencies for comparison. The green colour signifies increase in efficiency or decrease in rate while the red colour signifies decrease in efficiency or increase in rate compared to the default.

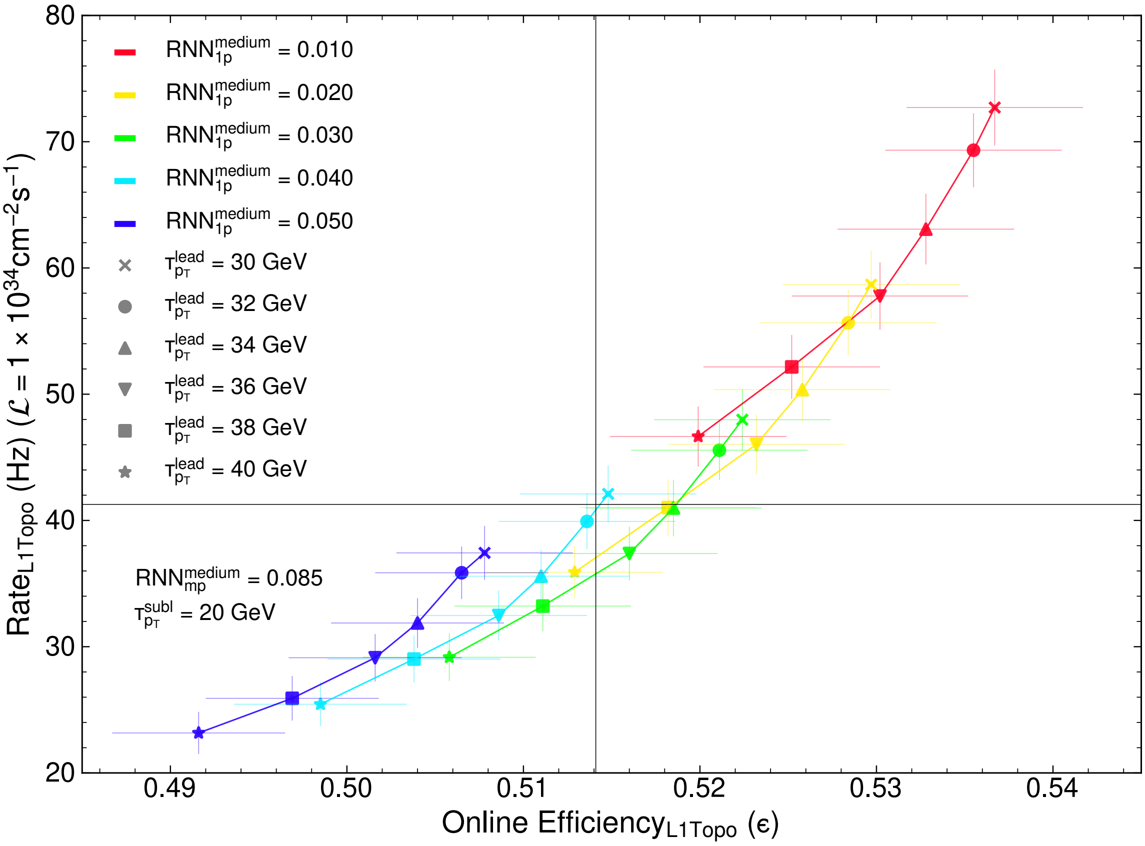
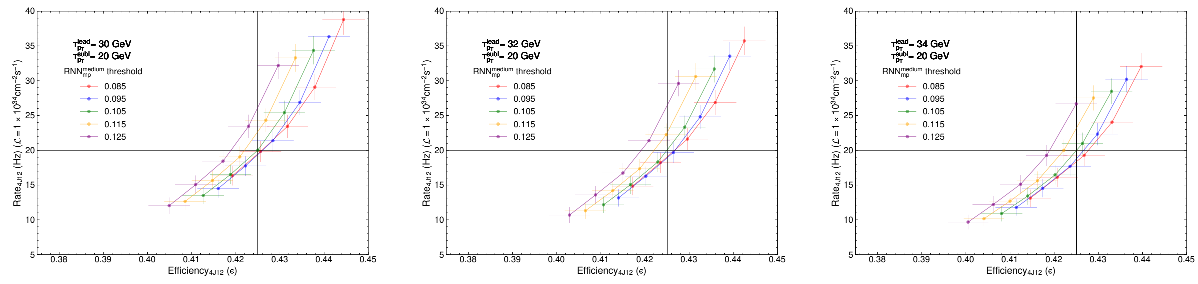


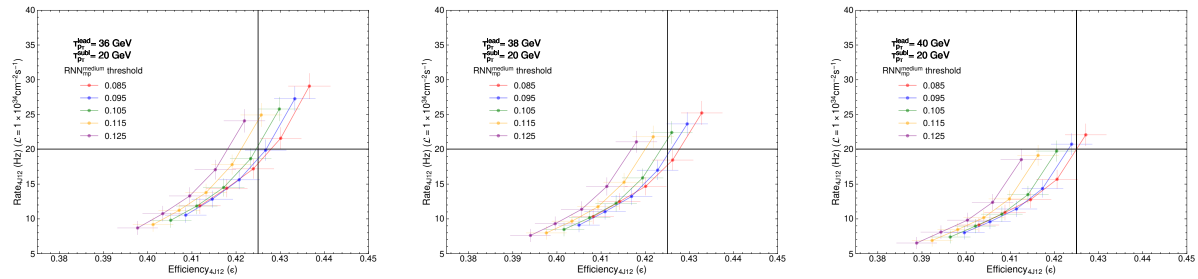
Figure 6.6: Plot displaying the variations in efficiencies and rates of the $\mathtt{HLT}_{\mathtt{L1Topo}}$ trigger when altering the $\tau^{\mathrm{lead}}_{p_T}$ and $\mathrm{RNN}^{\mathrm{medium}}_{\mathrm{1p}}$ thresholds associated with the tau objects. The $\tau^{\mathrm{subl}}_{p_T}$ and $\mathrm{RNN}^{\mathrm{medium}}_{\mathrm{mp}}$ thresholds of the tau objects are set to a constant 20 GeV and 0.085 respectively. The intersection point of the horizontal and vertical lines represents the rate and efficiency of the default variant of the $\mathtt{HLT}_{\mathtt{L1Topo}}$ trigger. The different coloured contours on each plot indicate different $\mathrm{RNN}^{\mathrm{medium}}_{\mathrm{1p}}$ thresholds, while different markers correspond to various leading p_T thresholds.

6.4.2 $\mathtt{HLT}_{\mathtt{4J12}}$ Chain

A similar analysis was performed for the $\mathtt{HLT}_{\mathtt{4J12}}$ chain. The optimisation process began by examining the efficiency versus rate plots shown in Figure 6.7, using a fixed subleading p_T threshold of 20 GeV for the tau objects. Using a reasoning analogous to that of $\mathtt{HLT}_{\mathtt{L1Topo}}$, the threshold for the multi-prong medium RNN was set at 0.085. The resulting plots of rate versus efficiency are illustrated in Figure 6.8. From this analysis a specific set of optimal parameters was derived, which is detailed in Table 6.3.



(a) $\tau_{p_T}^{\mathrm{lead}} = 30$ GeV (b) $\tau_{p_T}^{\mathrm{lead}} = 32$ GeV (c) $\tau_{p_T}^{\mathrm{lead}} = 34$ GeV



(d) $\tau_{p_T}^{\mathrm{lead}} = 36$ GeV (e) $\tau_{p_T}^{\mathrm{lead}} = 38$ GeV (f) $\tau_{p_T}^{\mathrm{lead}} = 40$ GeV

Figure 6.7: Plots that show the variations in the efficiencies and rates of the $\mathtt{HLT}_{\mathtt{4J12}}$ trigger when altering the thresholds for the leading p_T, $\mathrm{RNN}_{\mathrm{1p}}^{\mathrm{medium}}$, and $\mathrm{RNN}_{\mathrm{mp}}^{\mathrm{medium}}$ associated with the tau objects. The $\tau_{p_T}^{\mathrm{subl}}$ is set at a constant 20 GeV. Each plot varies the $\tau_{p_T}^{\mathrm{lead}}$ threshold. The intersection point of the horizontal and vertical lines in each plot represents the rate and efficiency of the default variant of the $\mathtt{HLT}_{\mathtt{L1Topo}}$ trigger. Different coloured contours on each plot indicate different multi-prong thresholds, while different points on the contour correspond to various 1-prong WPs. The y-axis range for the rates was constrained between 5 Hz and 40 Hz. For efficiencies, the x-axis range was constrained between 0.375 and 0.450.

Marker	$\mathrm{RNN}_{\mathrm{1p}}^{\mathrm{medium}}$	$\mathrm{RNN}_{\mathrm{mp}}^{\mathrm{medium}}$	$\tau_{p_T}^{\mathrm{lead}}$ (GeV)	$\tau_{p_T}^{\mathrm{subl}}$ (GeV)	Efficiency (ϵ)	Rate (Hz)
●	0.02	0.085	32	20	43.59 ± 0.48 %	26.87 ± 1.78
▼	0.02	0.085	36	20	43.01 ± 0.47 %	21.59 ± 1.59
▲	0.03	0.085	34	20	42.68 ± 0.47 %	19.27 ± 1.51
Default	0.03	0.105	30	20	42.50 ± 0.47 %	20.02 ± 1.53

Table 6.3: List of optimal parameters for the $\mathtt{HLT}_{\mathtt{4J12}}$ trigger chain, and their rates and efficiencies. The default set of parameters are provided at the bottom along with their rates and efficiencies for comparison. The green colour signifies increase in efficiency or decrease in rate while the red colour signifies decrease in efficiency or increase in rate compared to the default.

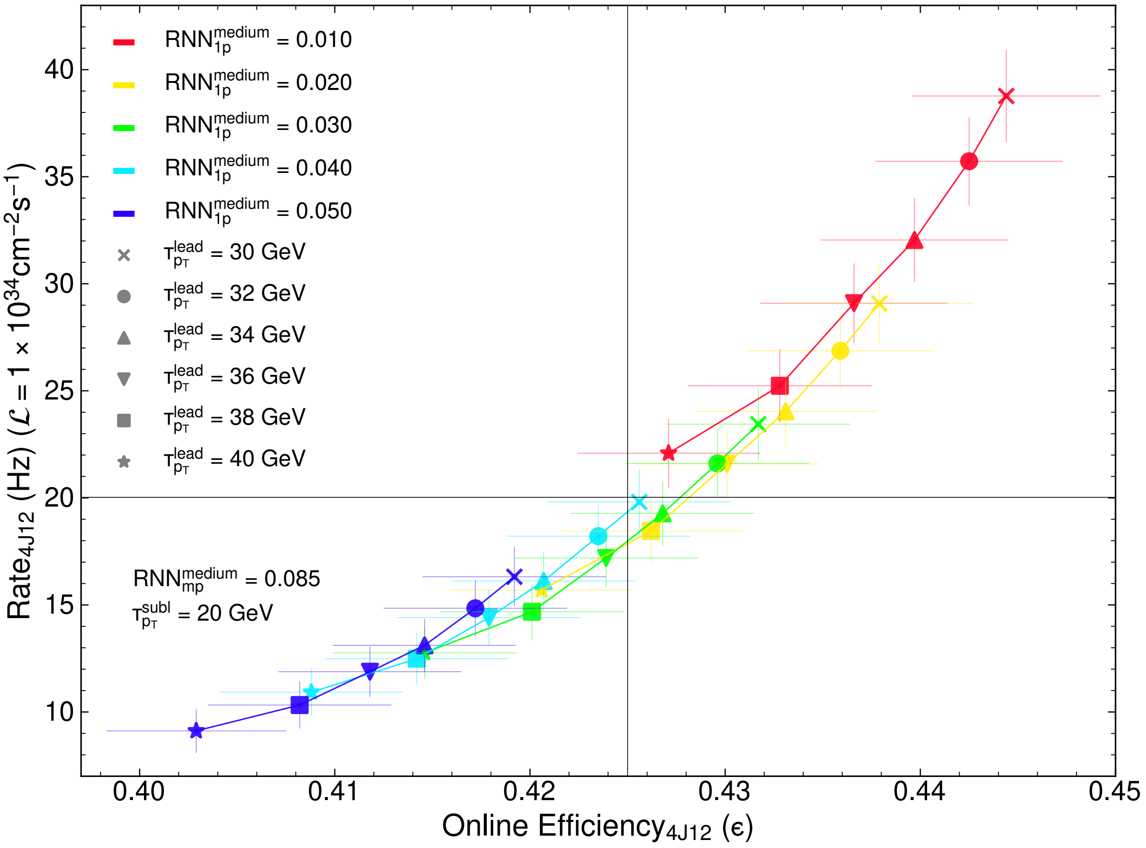
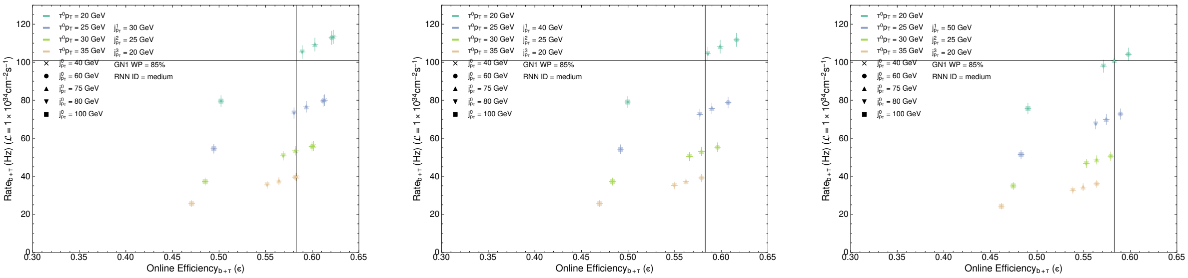


Figure 6.8: Plot displaying the variations in efficiencies and rates of the $\texttt{HLT}_{\texttt{4J12}}$ trigger when altering the $\tau^{\text{lead}}_{p_T}$ and $\text{RNN}^{\text{medium}}_{\text{1p}}$ thresholds associated with the tau objects. The $\tau^{\text{subl}}_{p_T}$ and $\text{RNN}^{\text{medium}}_{\text{mp}}$ thresholds of the tau objects are set to a constant 20 GeV and 0.085 respectively. The intersection point of the horizontal and vertical lines represents the rate and efficiency of the default variant of the $\texttt{HLT}_{\texttt{4J12}}$ trigger. Different colored contours on each plot indicate different $\text{RNN}^{\text{medium}}_{\text{1p}}$ thresholds, while different markers correspond to various leading p_T thresholds.

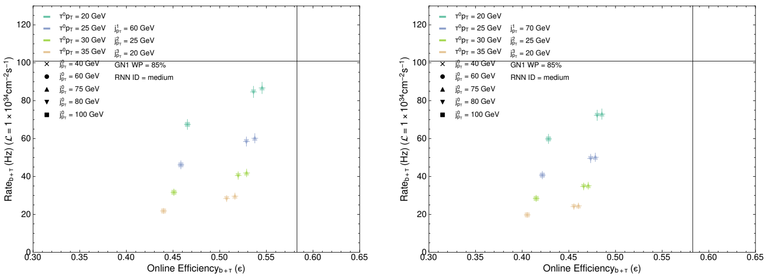
6.4.3 $\texttt{HLT}_{\texttt{b+}\tau}$ Chain

Regarding the $\texttt{HLT}_{\texttt{b+}\tau}$ chains, a basic analysis was performed, although it was not initially included in the thesis plan. This task emerged in the last months of the thesis period when the HH $\rightarrow b\bar{b}\tau^{+}_{\text{had}}\tau^{-}_{\text{had}}$ trigger analysis group requested a study on optimising the newly introduced $\texttt{HLT}_{\texttt{b+}\tau}$ chain. Therefore, at the time of writing this thesis, it was not feasible to present the optimisation process for all variables.

In the analysis, it was found that the p_T thresholds of the second jet ($j^2_{p_T}$) objects and the third jet ($j^3_{p_T}$) objects had a relatively minor effect on both efficiencies and rates. Consequently, these thresholds were set at 25 GeV and 20 GeV, respectively. The RNN ID requirements for the tau objects were fixed to medium WP. The p_T thresholds of the tau objects ($\tau^0_{p_T}$), the leading jet ($j^0_{p_T}$) objects and the subleading jet ($j^1_{p_T}$) objects were varied for different GN1 WP. The results for GN1 WP fixed at 85% are presented in Figure 6.9. The efficiency versus rate plots for the GN1 WP fixed to 77% and 70% can be found in Appendix C.3.



(a) $j^1_{p_T} = 30$ GeV (b) $j^1_{p_T} = 40$ GeV (c) $j^1_{p_T} = 50$ GeV



(d) $j^1_{p_T} = 60$ GeV (e) $j^1_{p_T} = 70$ GeV

Figure 6.9: Plots illustrating the variations in absolute efficiencies and rates of the $\texttt{HLT}_{\texttt{b+}\tau}$ trigger with GN1 working point set at 85%. The subleading jet object's p_T threshold is varied in different plots. The thresholds of the $j^2_{p_T}$ and $j^3_{p_T}$ objects are set at a constant of 25 GeV and 20 GeV, respectively. In each plot, the intersection point of the horizontal and vertical lines signifies the rate and absolute efficiency of the default version of the $\texttt{HLT}_{\texttt{b+}\tau}$ trigger. Different colours are used to represent different p_T thresholds of the tau object, while different markers denote varying leading jet p_T thresholds. The y-axis range for the rates was constrained between 0 Hz and 130 Hz. For efficiencies, the range of the x-axis was constrained between 0.30 and 0.65.

From the analysis, it is evident that the rates are clearly dominated by the p_T thresholds of the tau objects, while the efficiencies are dominated by the leading jet p_T thresholds. For many configurations, 85% GN1 WP has better efficiencies with similar rates, compared to 70% and 77% GN1 WP. Consequently, the GN1 WP was fixed at 85%. For similar reasons, the p_T thresholds of the subleading jet objects ($j^1_{p_T}$) were fixed to 30 GeV. Configurations with p_T thresholds for tau objects with 20 GeV and leading jet objects with 100 GeV were removed due to their high background rates and low efficiencies, respectively. For combinations of these parameters, the absolute efficiency versus rate plots are provided in Figure 6.10.

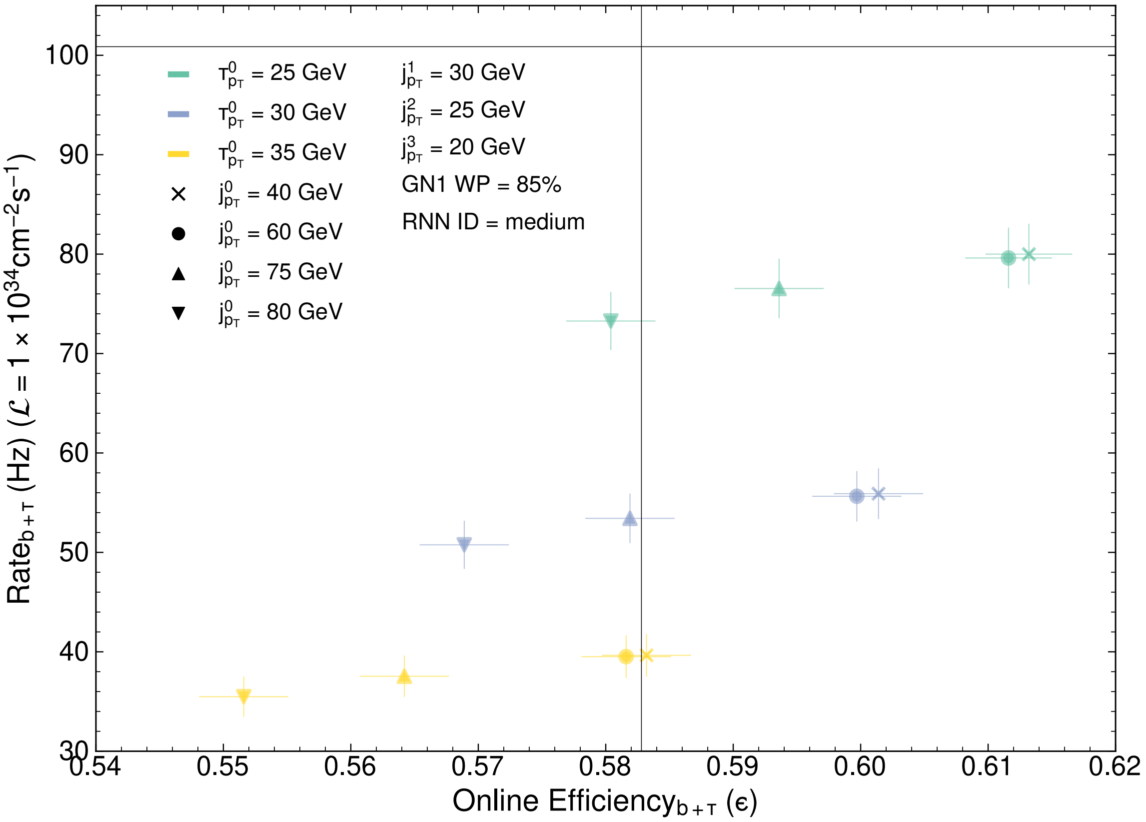


Figure 6.10: Plot displaying the variations in absolute efficiencies and rates of the `HLT`$_{\mathtt{b+}\tau}$ trigger on changing the p_T thresholds of the tau objects and leading jet objects. The p_T threshold of the tau objects changes with different colours, and the leading jet objects with different markers. The GN1 WP and the RNN WP is fixed to 85% and medium respectively. The p_T thresholds of the first ($j^1_{p_T}$), second ($j^2_{p_T}$) and third ($j^3_{p_T}$) jet objects are fixed to 30, 25 and 20 GeV respectively. The intersection point of the horizontal and vertical lines denotes the rate and absolute efficiency of the default variant of the `HLT`$_{\mathtt{b+}\tau}$ trigger.

An optimal set of parameters was determined from this analysis and are presented in Table 6.4. While the emulation framework may overstate the rates and efficiencies for the `HLT`$_{\mathtt{b+}\tau}$ chain with GN1-85% configuration, the relative trends remain accurate. Although actual improvements may not be as significant, the suggested choices still outperform the default configurations by a considerable margin.

This particular set of parameters, while showing improved efficiency gains with a moderate increase in rate, is not the exclusive optimal configuration. It is important to note the potential existence of alternative parameter sets that could further enhance efficiency but could also lead to higher rates. As the HH→ $b\bar{b}\tau^+\tau^-$ channel holds high priority at ATLAS, more resources could be allocated if a significant physics gain is demonstrated. Consequently, there is a degree of flexibility with respect to the rate of the trigger chain.

Marker	$\tau^{0}_{p_T}$ (GeV)	RNN WP	$j^{0}_{p_T}$ (GeV)	j^{1}_{pT} (GeV)	j^{2}_{pT} (GeV)	j^{3}_{pT} (GeV)	GNN WP	Absolute Efficiency (ϵ)	Rate (Hz)
×	25	medium	40	30	25	20	85%	$61.32 \pm 0.34\%$	79.99 ± 3.05
×	30	medium	40	30	25	20	85%	$60.14 \pm 0.35\%$	55.90 ± 2.55
×	35	medium	40	30	25	20	85%	$58.32 \pm 0.35\%$	39.64 ± 2.14
Default	20	medium	75	50	25	20	85%	$58.28 \pm 0.35\%$	100.86 ± 3.44

Table 6.4: List of optimal parameters for the $\texttt{HLT}_{\texttt{b+}\tau}$ trigger chain, with their corresponding rates and efficiencies. The initial set of parameters proposed for the $\texttt{HLT}_{\texttt{b+}\tau}$ chains is listed in the last row. The colors green signifies enhancement in efficiency and decrease in rate compared to the default.

Taking this into account, the configurations with the highest efficiency are chosen from this optimised set of parameters for the two di-tau triggers, to determine their combined absolute efficiency[3] for online event selection. This showed an improvement from $36.30 \pm 0.34\%$ to $37.20 \pm 0.34\%$. When these two di-tau triggers are combined with the optimised $\texttt{HLT}_{\texttt{b+}\tau}$ chain, the absolute efficiency increased from 67.80 ± 0.33 % to 69.26 ± 0.32 %.

---

[3] Absolute efficiencies do not use the selection events for efficiency emulation.

Chapter 7

Conclusion and Outlook

Examining the Higgs boson pair production offers a direct way to effectively probe the Higgs boson trilinear self-coupling. Among the various decay channels for the pair of Higgs bosons, the $b\bar{b}\tau^{+}_{\text{had}}\tau^{-}_{\text{had}}$ channel is noteworthy because of its high sensitivity and relatively lower background compared to other channels such as HH→ $b\bar{b}b\bar{b}$. However, given the rarity of this process, the rate at which the Higgs boson pair production signal events are collected may not suffice to conclusively confirm the production of Higgs pairs during Run-3 of the LHC. Hence, enhancing the efficiencies of the trigger chains to accept these rare signal events as much as possible, while managing the background rate, becomes crucial.

A study was carried out to optimise the existing di-tau trigger chains, $\texttt{HLT}_{\texttt{L1Topo}}$ and $\texttt{HLT}_{\texttt{4J12}}$. An attempt was also made to optimise the recently proposed $\texttt{HLT}_{\texttt{b+}\tau}$ chain. This involved modifying various parameters of these trigger chains and identifying the optimal configuration that increased signal acceptance and background rejection. Estimating background rates for a trigger chain that was not present in the trigger menu during reprocessing is a time-consuming process. To address this challenge, a trigger emulation framework was developed that estimated trigger chain rates in a few seconds. Given its significant utility, this framework is now adopted by the ATLAS experiment's Run-3 HH→ $b\bar{b}\tau^{+}_{\text{had}}\tau^{-}_{\text{had}}$ analysis group for rates emulation.

Prior to implementing the trigger emulation framework, validation was conducted for both efficiency and rate emulation results. Satisfactory agreement was noted for the di-tau triggers. However, notable deviations were observed for the $\texttt{HLT}_{\texttt{b+}\tau}$ triggers, particularly in variants that enforce the 85% b-tagging WP requirement, where the trigger rates were overestimated. This deviation is likely attributed to the omission of emulation of the preselection condition within the $\texttt{HLT}_{\texttt{b+}\tau}$ chain.

After optimising the $\texttt{HLT}_{\texttt{L1Topo}}$ and $\texttt{HLT}_{\texttt{4J12}}$ chains, multiple trigger variants were recommended. With a manageable relative increase in the rate of 22.02% and 34.18%, an improvement in efficiency of up to 1.17% and 1.09% was observed in the $\texttt{HLT}_{\texttt{L1Topo}}$ and $\texttt{HLT}_{\texttt{4J12}}$ chains, respectively, compared to their current parameter configuration. For the recently proposed $\texttt{HLT}_{\texttt{b+}\tau}$ chains, three different sets of trigger parameters were suggested. Of these, the trigger configuration with the highest signal acceptance had an absolute efficiency of $61.32 \pm 0.34\%$ with a background rate of 79.9 ± 3.4 Hz.

By combining all three triggers with their optimal configuration in a logical `OR` resulted in an absolute efficiency of $69.26 \pm 0.32\%$. This is an improvement compared to the existing absolute efficiency of $67.80 \pm 0.33\%$, derived from the initially proposed parameters for the $\texttt{HLT}_{\texttt{b+}\tau}$ chain and the current configuration of the $\texttt{HLT}_{\texttt{L1Topo}}$ and $\texttt{HLT}_{\texttt{4J12}}$ chains.

Before proceeding with the modification of these parameters within the trigger menu, a more comprehensive examination of certain aspects of the trigger chains is crucial. While the individual rates of trigger chains have been scrutinised thus far, it is equally important to assess the unique rate of each trigger chain to estimate potential overlaps in the background event selection with other trigger chains. If the unique rate is small, the trigger parameter under analysis could be loosened further.

Additionally, as part of optimising these configurations, it is essential to ensure that the CPU usage associated with these optimised triggers remains within manageable limits. Moreover, an investigation of potential improvements in offline efficiency and signal sensitivity is also necessary. By improving the signal acceptance of these trigger chains, this study makes the discovery of the Higgs pair production process during the Run 3 slightly more achievable and ultimately assists with validating the SM Higgs hypothesis.

Bibliography

[1] J. J. Thomson, *XL. Cathode Rays*, Lond. Edinb. Dublin philos. mag. j. sci. 44.269 (1897), p. 293, https://doi.org/10.1080/14786449708621070.

[2] S. L. Glashow, *Partial-symmetries of weak interactions*, Nucl. Phys. 22 (1961), pp. 579–588, https://doi.org/10.1016/0029-5582(61)90469-2.

[3] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. 19 (1967), pp. 1264–1266, https://doi.org/10.1103/PhysRevLett.19.1264.

[4] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. C 680519 (1968), pp. 367–377, https://doi.org/10.5170/CERN-19680904.367.

[5] V. C. Rubin and W. K. F. Jr., *Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions*, ApJ 159 (1970), p. 379, https://doi.org/10.1086/150317.

[6] S. Perlmutter et al., *Measurements of Ω and Λ from 42 High-Redshift Supernovae*, ApJ 517 (1999), p. 565, https://doi.org/10.1086/307221.

[7] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, Sov. Phys. Usp. 34 (1967), p. 392, https://doi.org/10.1070/PU1991v034n05ABEH002497.

[8] The Atlas collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the Atlas detector at the Lhc*, Phys. Lett. B 716.1 (2012), pp. 1–29, https://doi.org/10.1016/j.physletb.2012.08.020.

[9] The Cms Collaboration, *Observation of a new boson at a mass of 125 GeV with the Cms experiment at the Lhc*, Phys. Lett. B 716 (2012), p. 30, https://doi.org/10.1016/j.physletb.2012.08.021.

[10] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys. Rev. 145 (4 1966), p. 1156, https://doi.org/10.1103/PhysRev.145.1156.

[11] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. 13 (9 1964), p. 321, https://doi.org/10.1103/PhysRevLett.13.321.

[12] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. 13 (1964), p. 585, https://doi.org/10.1103/PhysRevLett.13.585.

[13] L. Evans and P. Bryant, *LHC Machine*, JINST 3 (2008), S08001, https://doi.org/10.1088/1748-0221/3/08/S08001.

[14] The ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST 3 (2008), S08003, https://doi.org/10.1088/1748-0221/3/08/S08003.

[15] The ATLAS 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*, JHEP 2023 (2023), p. 40, https://doi.org/10.1007/jhep07(2023)040.

[16] The ATLAS 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*, Phys. Rev. D 105 (2022), p. 092002, https://doi.org/10.1103/PhysRevD.105.092002.

[17] The ATLAS 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*, Phys. Rev. D 106 (2022), p. 052001, https://doi.org/10.1103/physrevd.106.052001.

[18] The ATLAS 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*, Phys. Lett. B 801 (2020), p. 135145, https://doi.org/10.1016/j.physletb.2019.135145.

[19] The ATLAS Collaboration, *Operation of the ATLAS trigger system in Run 2*, JINST 15 (2020), P10004, https://doi.org/10.1088/1748-0221/15/10/P10004.

[20] P. S. Wells, *Experimental tests of the Standard Model*, Eur. Phys. J. C 33 (2004), s5–s20, https://doi.org/10.1140/epjcd/s2004-03-1692-8.

[21] T. Kobayashi, *Experimental verification of the Standard Model of particle physics*, PJA, Series B 97 (2021), p. 211, https://doi.org/10.2183/pjab.97.013.

[22] A. Einstein, *On a Heuristic Viewpoint Concerning the Production and Transformation of Light*, Ann. Phys. 322 (1905), p. 132, https://doi.org/10.1002/andp.19053220607.

[23] E. Rutherford, *The Scattering of* α *and* β *Particles by Matter and the Structure of the Atom*, Philos. Mag. 21 (1911), p. 669, https://doi.org/10.1080/14786440508637080.

[24] P. A. M. Dirac, *The Quantum Theory of the Electron*, Proc. R. Soc. Lond. A 117 (1928), p. 610, https://doi.org/10.1098/rspa.1928.0023.

[25] E. Fermi, *An attempt of a theory of beta radiation.* I. Z. Physik 88 (1934), p. 161, https://doi.org/10.1007/BF01351864.

[26] J. Schwinger, *On Quantum-Electrodynamics and the Magnetic Moment of the Electron*, Phys. Rev. 73 (1948), p. 416, https://doi.org/10.1103/PhysRev.73.416.

[27] M. Gell-Mann, *A schematic model of baryons and mesons*, Phys. Lett. 8 (1964), p. 214, https://doi.org/10.1016/S0031-9163(64)92001-3.

[28] M. Breidenbach et al., *Observed Behavior of Highly Inelastic Electron-Proton Scattering*, Phys. Rev. Lett. 23 (1969), p. 935, https://doi.org/10.1103/PhysRevLett.23.935.

[29] The UA1 Collaboration, *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at* $\sqrt{s} = 540\ GeV$, Phys. Lett. B 122 (1983), p. 103, https://doi.org/10.1016/0370-2693(83)91177-2.

[30] The UA1 Collaboration, *Experimental observation of lepton pairs of invariant mass around 95* GeV/c^2 *at the CERN SPS collider*, Phys. Lett. B 126 (1983), p. 398, https://doi.org/10.1016/0370-2693(83)91123-7.

[31] A. Djouadi, *The anatomy of electroweak symmetry breaking: Tome I: The Higgs boson in the Standard Model*, Phys. Rep. 457 (2008), p. 1, https://doi.org/10.1016/j.physrep.2007.10.004.

[32] The ATLAS Collaboration, *Measurement of the Higgs boson mass in the* $H \rightarrow ZZ^* \rightarrow 4\ell$ *and* $H \rightarrow \gamma\gamma$ *channels with* $\sqrt{s} = 13\ TeV$ *pp collisions using the ATLAS detector*, Phys. Lett. B 784 (2018), p. 345, https://doi.org/10.1016/j.physletb.2018.07.050.

[33] Particle Data Group, *Review of Particle Physics*, PTEP 2022 (2022), p. 083C01, https://doi.org/10.1093/ptep/ptac097.

[34] The ATLAS Collaboration, *Constraints on the Higgs boson self-coupling from single and double-Higgs production with the ATLAS detector using pp collisions at* $\sqrt{s} = 13\ TeV$, Phys. Lett. B 843 (2023), p. 137745, https://doi.org/10.1016/j.physletb.2023.137745.

[35] M. E. Shaposhnikov, *Possible appearance of the baryon asymmetry of the universe in an electroweak theory*, JETP Lett. 44 (1986), p. 465, https://ui.adsabs.harvard.edu/abs/1986JETPL..44..465S.

[36] S. Alexander, *Inflation and electroweak symmetry breaking*, JCAP (2023), p. 008, https://doi.org/10.1088/1475-7516/2023/09/008.

[37] P. Q. Hung, *Vacuum Instability and New Constraints on Fermion Masses*, Phys. Rev. Lett. 42 (1979), p. 873, https://doi.org/10.1103/PhysRevLett.42.873.

[38] D. de Florian and J. Mazzitelli, *Higgs pair production at next-to-next-to-leading logarithmic accuracy at the LHC*, JHEP 2015 (2015), p. 53, `https://doi.org/10.1007/JHEP09(2015)053`.

[39] The ATLAS Collaboration, *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*, CERN Yellow Reports: Monographs, CERN, 2011, `https://doi.org/10.5170/CERN-2011-002`.

[40] E. Picasso and G. Plass, *The Machine Design*, Europhys. News 20 (1989), p. 80, `https://doi.org/10.1051/epn/19892006080`.

[41] The CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST 3 (2008), S08004, `https://doi.org/10.1088/1748-0221/3/08/S08004`.

[42] The ALICE Collaboration, *The ALICE experiment at the LHC*, JINST 3 (2008), S08002, `https://doi.org/10.1088/1748-0221/3/08/S08002`.

[43] The LHCb Collaboration, *The LHCb Detector at the LHC*, JINST 3 (2008), S08005, `https://doi.org/10.1088/1748-0221/3/08/S08005`.

[44] K. Fabian, *Signal Formation Processes in Micromegas Detectors and Quality Control for large size Detector Construction for the ATLAS New Small Wheel*, 2017, arXiv: `1708.01624 [physics.ins-det]`.

[45] The ATLAS Collaboration, *The ATLAS Inner Detector commissioning and calibration*, Eur. Phys. J. C 70 (2010), p. 787, `https://doi.org/10.1140/epjc/s10052-010-1366-7`.

[46] The ATLAS Collaboration, *Production and integration of the ATLAS Insertable B-Layer*, Journal of Instrumentation 13.05 (2018), T05008, `https://doi.org/10.1088/1748-0221/13/05/t05008`.

[47] Wainer Vandelli, *ATLAS DataFlow Infrastructure: Recent results from ATLAS cosmic and first-beam data-taking*, J. Phys. Conf. Ser. 219 (2010), p. 022047, `https://doi.org/10.1088/1742-6596/219/2/022047`.

[48] R. Simoniello, *The ATLAS Level-1 Topological Processor: From design to routine usage in Run-2*, Technical report ATL-DAQ-PROC-2018-044, CERN, 2019, `https://doi.org/10.1109/NSSMIC.2018.8824280`.

[49] The ATLAS Collaboration, *The ATLAS Level-1 Calorimeter Trigger*, JINST 3 (2008), P03001, `https://doi.org/10.1088/1748-0221/3/03/P03001`.

[50] The ATLAS Collaboration, *Performance of the ATLAS Level-1 topological trigger in Run 2*, Eur. Phys. J. C 82 (2022), p. 7, `https://doi.org/10.1140/epjc/s10052-021-09807-0`.

[51] R. Schwienhorst, *The Phase-1 upgrade of the ATLAS first level calorimeter trigger*, JINST 11 (2016), p. C01018, `https://doi.org/10.1088/1748-0221/11/01/C01018`.

[52] The 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*, Technical report ATL-PHYS-PUB-2022-044, CERN, 2022, https://cds.cern.ch/record/2827111.

[53] The ATLAS Collaboration, *Local Hadronic Calibration*, Technical report ATL-LARG-PUB-2009-001-2, CERN, 2008, https://cds.cern.ch/record/1112035.

[54] The ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2010*, Eur. Phys. J. C 72 (2012), p. 1849, https://doi.org/10.1140/epjc/s10052-011-1849-1.

[55] The ATLAS Collaboration, *The ATLAS Tau Trigger in Run 2*, Technical report ATLAS-CONF-2017-061, CERN, 2017, https://cds.cern.ch/record/2274201.

[56] The ATLAS Collaboration, *Identification of hadronic tau lepton decays using neural networks in the ATLAS experiment*, Technical report ATL-PHYS-PUB-2019-033, CERN, 2019, https://cds.cern.ch/record/2688062.

[57] The ATLAS Collaboration, *Fast b-tagging at the high-level trigger of the ATLAS experiment in LHC Run 3*, 2023, arXiv: 2306.09738 [hep-ex].

[58] The ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, Technical report ATLAS-CONF-2014-018, CERN, 2014, https://cds.cern.ch/record/1700870.

[59] The ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. C 77 (2017), p. 466, https://doi.org/10.1140/epjc/s10052-017-5031-2.

[60] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-kt jet clustering algorithm*, JHEP 04 (2008), p. 063, https://doi.org/10.1088/1126-6708/2008/04/063.

[61] The ATLAS Collaboration, *ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset*, Eur. Phys. J. C 83 (2023), p. 681, https://doi.org/10.1140/epjc/s10052-023-11699-1.

[62] The ATLAS Collaboration, *Neural Network Jet Flavour Tagging with the Upgraded ATLAS Inner Tracker Detector at the High-Luminosity LHC*, Technical report ATL-PHYS-PUB-2022-047, CERN, 2022, https://cds.cern.ch/record/2839913.

[63] The ATLAS Collaboration, *Graph Neural Network Jet Flavour Tagging with the ATLAS Detector*, Technical report ATL-PHYS-PUB-2022-027, CERN, 2022, https://cds.cern.ch/record/2811135.

[64] T. Hryn'ova and K. Nagano, *Trigger Menu Strategy for Run 2*, Technical report ATL-COM-DAQ-2014-054, CERN, 2014, https://cds.cern.ch/record/1703730.

[65] The Atlas Collaboration, *Searches for Higgs boson pair production in the $b\bar{b}\tau^{+}\tau^{-}$ final state with 139 fb^{-1} of pp collision data with the Atlas detector*, Technical report ATL-COM-PHYS-2020-766, Cern, 2020, https://cds.cern.ch/record/2743097.

[66] The Atlas Collaboration, *Prompt data reconstruction at the Atlas experiment*, Technical report ATL-SOFT-PROC-2012-062, Cern, 2012, https://doi.org/10.1088/1742-6596/396/2/022049.

[67] P. Calafiura et al., *The Athena Control Framework in Production, New Developments and Lessons Learned* (2005), https://doi.org/10.5170/CERN-2005-002.456.

[68] R. Brun and F. Rademakers, *ROOT — An object oriented data analysis framework*, NIM-A 389 (1997), New Computing Techniques in Physics Research V, p. 81, https://doi.org/10.1016/S0168-9002(97)00048-X.

[69] The Atlas Collaboration, *The Atlas Metadata Interface (AMI), a generic metadata framework*, Technical report ATL-SOFT-PROC-2017-007, Cern, 2017, https://doi.org/10.1088/1742-6596/898/6/062001.

[70] M. Barisits et al., *Rucio: Scientific Data Management*, Comput. Softw. Big Sci. 3 (2019), https://doi.org/10.1007/s41781-019-0026-3.

[71] K. Hamilton et al., *A positive-weight Next-to-Leading Order Monte Carlo simulation for Higgs boson production*, JHEP 2009 (2009), p. 116, https://doi.org/10.1088/1126-6708/2009/04/116.

[72] M. Bertini et al., *Pythia version 7-0.0 – a proof-of-concept version*, Comput. Phys. Commun. 134 (2001), p. 365, https://doi.org/10.1016/s0010-4655(00)00206-x.

[73] D. J. Lange, *The EvtGen particle decay simulation package*, NIM-A 462 (2001), p. 152, https://doi.org/10.1016/S0168-9002(01)00089-4.

[74] The Atlas Collaboration, *Trigger monitoring and rate predictions using Enhanced Bias data from the Atlas Detector at the Lhc*, Technical report ATL-DAQ-PUB-2016-002, Cern, 2016, https://cds.cern.ch/record/2223498.

Appendix A

Algorithms for τ lepton identification

The BDT algorithm for offline like tau lepton identification proceeds as outlined below:

A precision-tracking algorithm similar to the offline one is run using tracks identified by the FTK, which act as seeds to measure the tau TOBs' properties more precisely. Using these tracks and calorimeter information, the input variables are computed for the identification of τ_{had} objects. Two sets of variables for 1-prong and multi-prong τ_{had} candidates are used for the BDT selection as listed in Figure A.1.

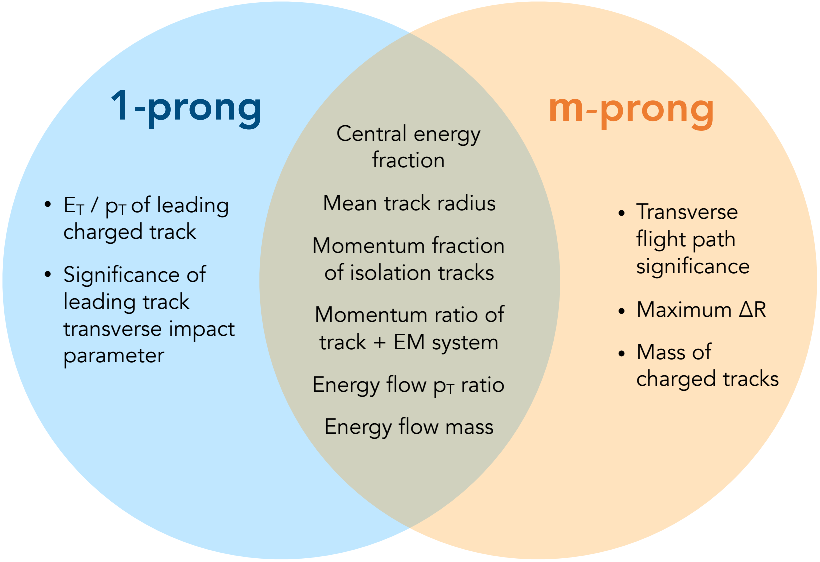


Figure A.1: Discriminating variables used as input to the BDT algorithm for 1-prong and multi-prong τ_{had} candidates.

As the input variables in the BDTs are all high-level, there is a risk of losing some information. To address this, Recurrent Neural Networks (RNN) integrate low-level input variables alongside the BDT variables. Details on the specific variables used in the RNN are available in Figure A.2.

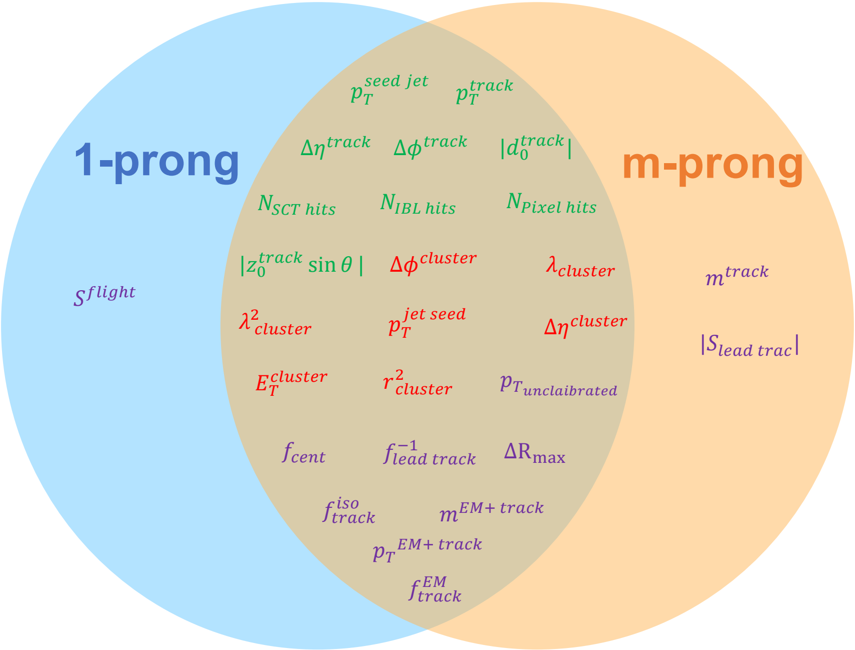


Figure A.2: List of input variables for the RNN 1-prong and multi-prong (multi-prong) τs. The track inputs in represented in green. The cluster inputs are represented in red and the high level inputs are represented in purple. The explanation for these variables can be found in [56].

To facilitate the use of charged particle tracks, clusters in the calorimeter, and high level variables, the network is split into three dedicated branches for each type of input. This structural division is visually depicted in Figure A.3. The high-level observables associated with each τ_{had} candidate is processed through a sequence of three fully connected layers with 128, 128 and 16 neurons, each. The sequences of charged particle tracks and calorimeter clusters are ordered by decreasing p_T. Shared dense layers with shared weights process these sequences, creating intermediate representations for each track or cluster. These representations are then fed into recurrent layers using the long-short-term memory (LSTM) architecture, known for its ability to maintain information over sequences. The first LSTM layer maps the input sequence to an output sequence, enabling information from earlier elements to influence later outputs due to the LSTM's internal state. The second LSTM layer summarises this sequence into a single vector, condensing the information for further computations.

The results of the computation in the three branches are combined by concatenating their output vectors in a merge layer. Further computation on the resulting vector is performed by three fully connected layers with 64, 32 and 1 neurons each. After applying a sigmoid activation function, the final output corresponds to the RNN tau identification score.

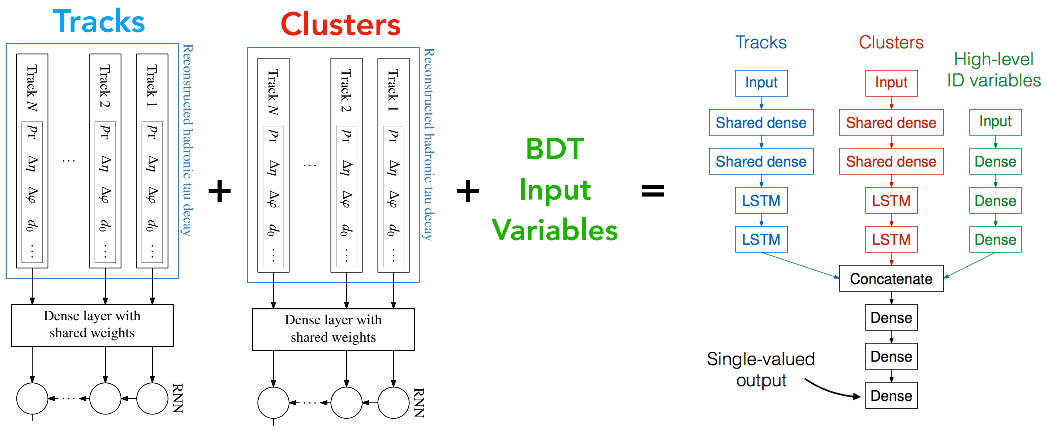


Figure A.3: Structure of the RNN architecture. Adapted from graphic by © Mariel Pettee, CC BY 2.5, 2018 US LHC Users Association Meeting.

Appendix B

Enhanced Bias Triggers

All the triggers in the Enhanced Bias trigger menu during the Run 3 data-taking phase have been listed below:

- `HLT_eb_low_L1RD2_FILLED` - which is seeded from random filled `L1_RD2_FILLED` and then selects events in the HLT with an `OR` of a list of L1 items. The reason for this construction is to maintain the correlations between the different chains so that their overlaps can be correctly calculated. If the `OR` were to be implemented after the prescales were applied the correlation would be suppressed by the smaller of the prescales. Using the `L1_RD2_FILLED` as a seed and then applying a prescale after the `OR` maintains the correlations.
- `HLT_eb_medium_L1RD2_FILLED` is the same as `HLT_eb_low_L1RD2_FILLED` but with a tighter set of L1 items in the `OR`.
- `HLT_noalg_L1PhysicsHigh_noPS` is seeded from an `OR` of many L1 items which were not prescaled in L1. This allows us to collect high statistics for the tightest L1 items. Since seeds are not prescaled at L1, the correlations are maintained without need for the random seed mechanism. This chain runs at highest rate as it contains primary unprescaled triggers.
- `HLT_noalg_L1PhysicsVaryHigh_noPS` is like `HLT_noalg_eb_L1PhysicsHigh_noPS`, but for the very highest p_T L1 objects, to obtain good statistics for these too. It runs at a much lower prescale at the HLT.
- `HLT_noalg_L1RD3_FILLED` is an `L1_RD3_FILLED` seeded chain used to fill in the thresholds below those included in `HLT_eb_low_L1RD2_FILLED`. When calculating rates offline, HLT chains that are explicitly seeded from random filled `L1` chains, are forced offline to only use events selected by this trigger to avoid double counting through inclusion of the biased sample.

- `HLT_noalg_L1RD3_EMPTY` is defined analogous to `HLT_noalg_eb_L1RD3_FILLED` but for empty bunches.

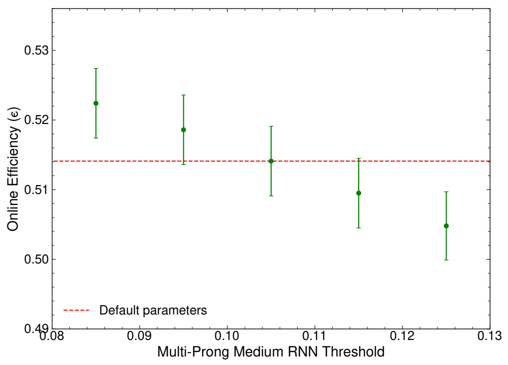
- `HLT_noalg_noPS_EMPTY`, `HLT_noalg_noPS_FIRSTEMPTY`, `HLT_noalg_noPS_UNPAIRED_ISO`, `HLT_noalg_noPS_UNPAIRED_NONISO`, `HLT_noalg_noPS_ABORTGAPNOTCALIB` are defined analogous to `HLT_noalg_eb_L1PhysicsLow_noPS` but for bunches other than filled.

The Run-3 EB trigger chain information is obtained from https://twiki.cern.ch/twiki/bin/view/Atlas/EnhancedBiasData

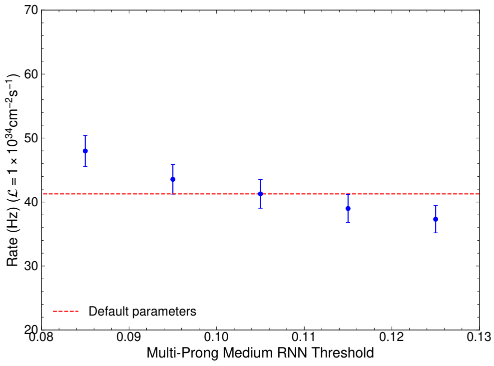
Appendix C

Additional Figures

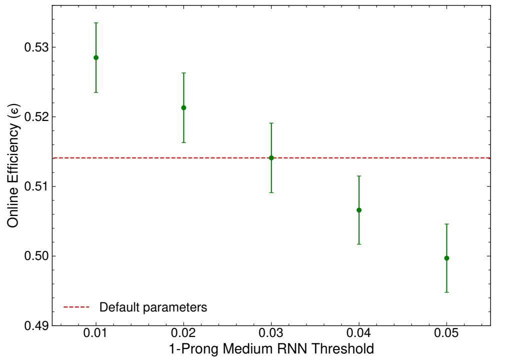
C.1 Efficiency and Rates Emulation results - $\texttt{HLT}_{\texttt{L1Topo}}$



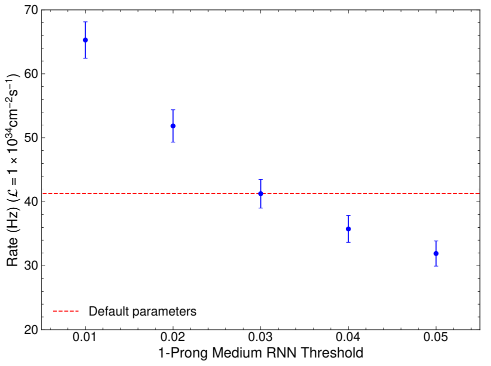
(a) Correlation between efficiency and the multi-prong medium RNN threshold requirement of tau objects.



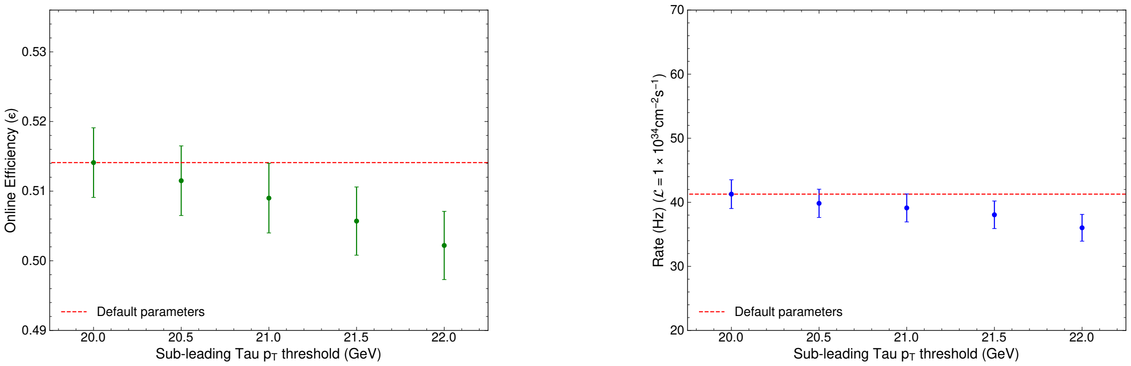
(b) Correlation between rate and the multi-prong medium RNN threshold requirement of tau objects.



(c) Correlation between efficiency and the 1-prong medium RNN threshold requirement of tau objects .

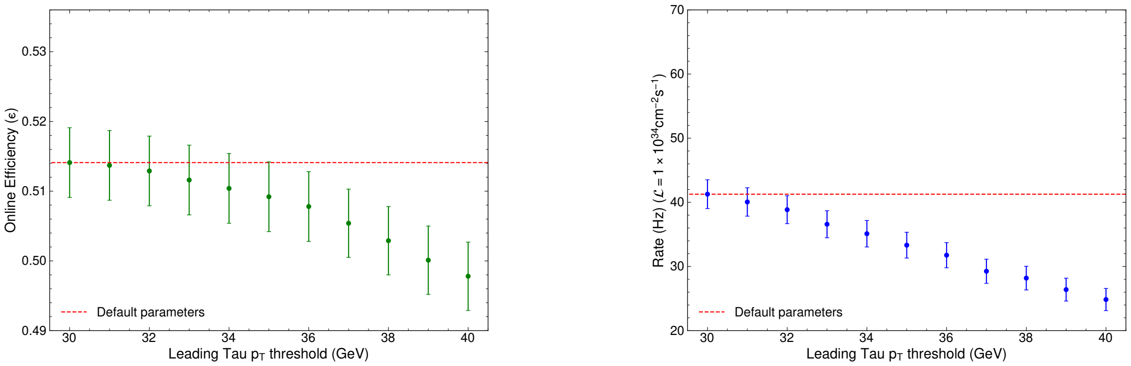


(d) Correlation between rate and the 1-prong medium RNN threshold requirement of tau objects .



(e) Correlation between efficiency and subleading p_T threshold requirement of tau objects.

(f) Correlation between rate and subleading p_T threshold requirement of tau objects.



(g) Correlation between efficiency and leading p_T threshold requirement of tau objects.

(h) Correlation between rate and leading p_T threshold requirement of tau objects.

Figure C.1: Comparative analysis showing efficiency variations on the left and rate variation on the right with respect to individual trigger parameters. The red dashed line denotes the efficiency of the $\texttt{HLT}_{\texttt{L1Topo}}$ chain with default parameters. All parameters, except the variable under observation, remain constant. The y-axis range for the rates was constrained between 20 Hz and 70 Hz. Meanwhile, for efficiencies, the y-axis range was constrained between 0.49 and 0.536.

C.2 Efficiency versus Rates plots for all parameter combinations

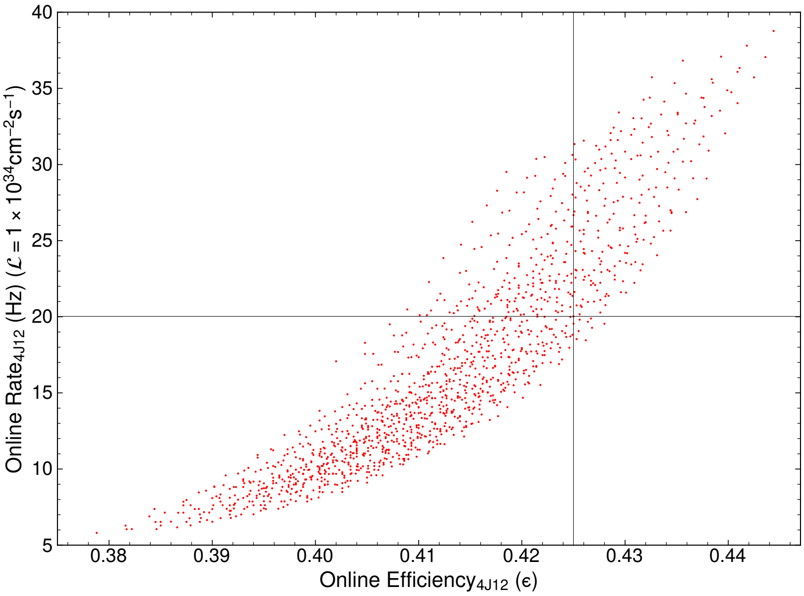


Figure C.2: Efficiencies and rates of the `HLT`$_{\texttt{4J12}}$ chain generated by varying the thresholds of different trigger parameter. The intersection point corresponds to the rate and efficiency of the default trigger parameters of the `HLT`$_{\texttt{4J12}}$ chain.

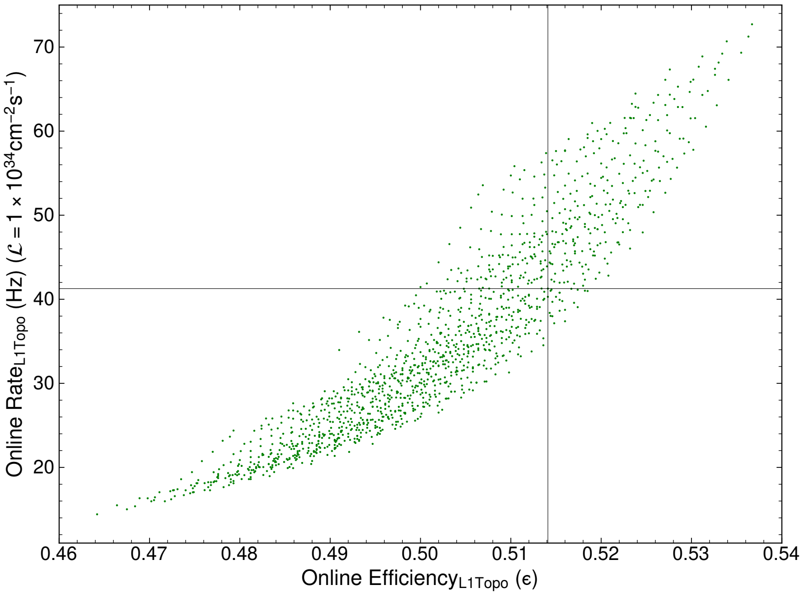


Figure C.3: Efficiencies and rates of the `HLT`$_{\texttt{L1Topo}}$ chain generated by varying the thresholds of different trigger parameter. The intersection point corresponds to the rate and efficiency of the default trigger parameters of the `HLT`$_{\texttt{L1Topo}}$ chain.

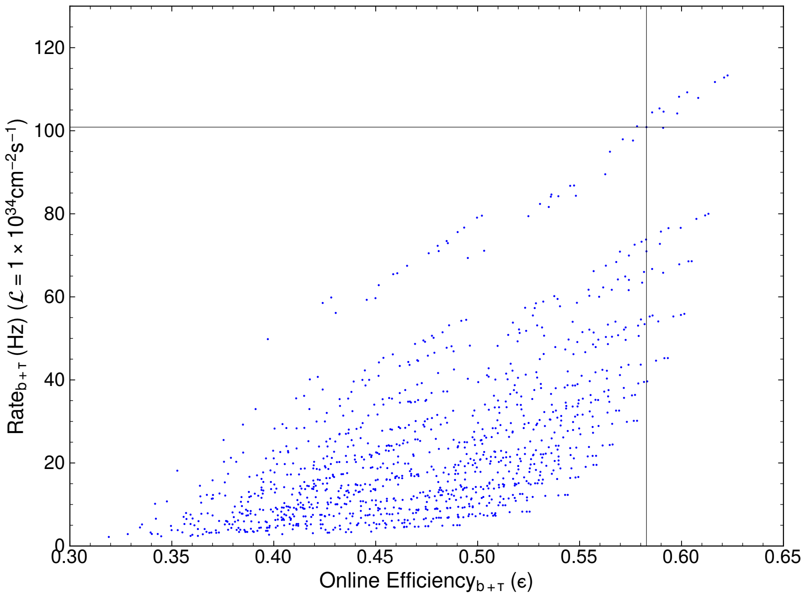
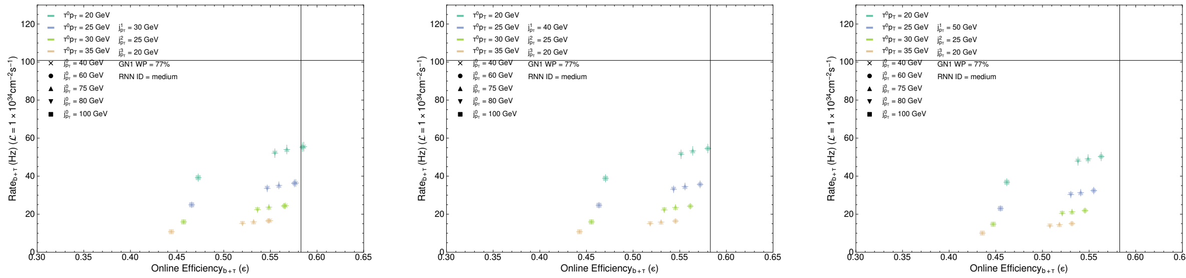
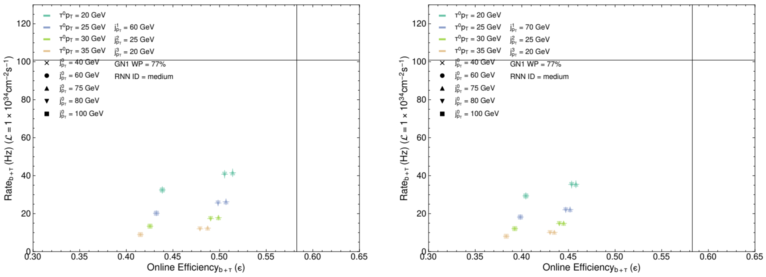


Figure C.4: Efficiencies and rates of the $\texttt{HLT}_{b+\tau}$ chain generated by varying the thresholds of different trigger parameter. The intersection point corresponds to the rate and efficiency of the default trigger parameters of the $\texttt{HLT}_{b+\tau}$ chain.

C.3 Efficiency versus Rate plots for $\mathtt{HLT}_{\mathtt{b+}\tau}$ chain with fixed GN1 WP

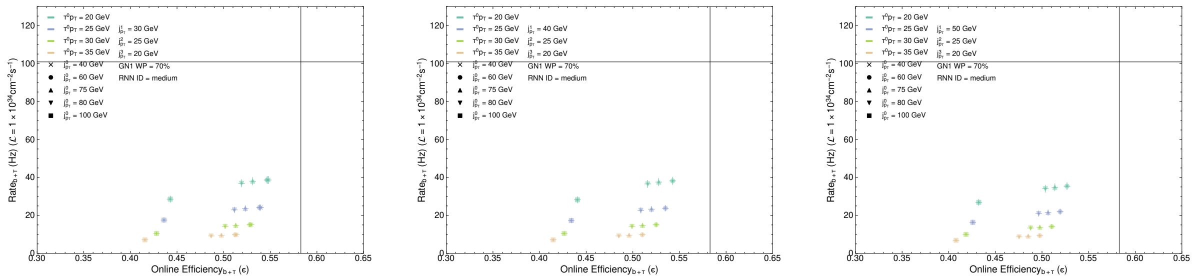


(a) $j^1_{p_T} = 30$ GeV (b) $j^1_{p_T} = 40$ GeV (c) $j^1_{p_T} = 50$ GeV

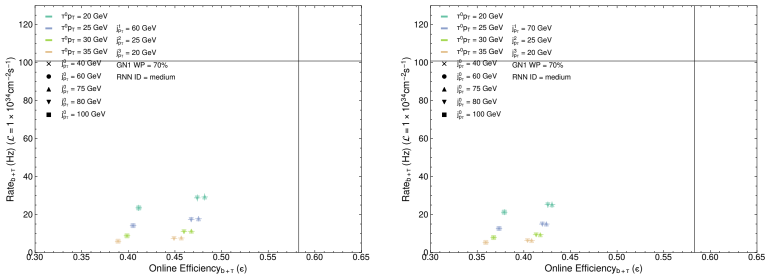


(d) $j^1_{p_T} = 60$ GeV (e) $j^1_{p_T} = 70$ GeV

Figure C.5: Plots illustrating the variations in efficiencies and rates of the $\mathtt{HLT}_{\mathtt{b+}\tau}$ trigger with GN1 working point set at 77%. The subleading jet object's p_T threshold is varied in different plot. The $j^2_{p_T}$ and $j^3_{p_T}$ thresholds are set at a constant 25 GeV and 20 GeV, respectively. In each plot, the intersection point signifies the rate and efficiency of the default version of the $\mathtt{HLT}_{\mathtt{b+}\tau}$ trigger. Different colours are used to represent different p_T thresholds of the tau object, while different markers denote varying leading jet p_T thresholds.



(a) $j^1_{p_T} = 30$ GeV (b) $j^1_{p_T} = 40$ GeV (c) $j^1_{p_T} = 50$ GeV



(d) $j^1_{p_T} = 60$ GeV (e) $j^1_{p_T} = 70$ GeV

Figure C.6: Plots illustrating the variations in efficiencies and rates of the $\mathtt{HLT}_{\mathtt{b+}\tau}$ trigger with GN1 working point set at 70%. The subleading jet object's p_T threshold is varied in different plot. The $j^2_{p_T}$ and $j^3_{p_T}$ thresholds are set at a constant 25 GeV and 20 GeV, respectively. In each plot, the intersection point signifies the rate and efficiency of the default version of the $\mathtt{HLT}_{\mathtt{b+}\tau}$ trigger. Different colours are used to represent different p_T thresholds of the tau object, while different markers denote varying leading jet p_T thresholds.

Danksagung

First, a heartfelt shout out to my incredible parents and my amazing little sister! Their unwavering belief in me and their constant support have been the wind beneath my wings. None of these would have been possible without their sacrifices. Reflecting on the past two years, my time in Germany has been a roller coaster of challenges and incredible experiences. Throughout it all, I have had an amazing support crew.

My deepest appreciation goes to Stan, my thesis supervisor, whose exceptional lectures and unwavering encouragement inspired me to pursue particle physics. Your guidance in presenting my thoughts, continuous motivation during the most demanding circumstances, and mentorship have significantly improved my skills as a researcher.

I extend my sincere thanks to Andrés for being my primary point of contact during this project addressing my queries and providing assistance with proofreading. I owe a great deal of thanks to Philipp, Rishabh, and Saidev for all those insightful discussions about my research and everything in between. Thanks to Katharine, Liaoshan, and Song Ming for their invaluable help in clarifying all the technical questions related to my research. Prof. Dr. Arnulf Quadt, I extend my gratitude for agreeing to be my second referee. And this section would not be complete without acknowledging the support, motivation, and delicious food offered by Anjana.

To everyone who contributed to this journey, whether through support, wisdom, or simply good company: thank you from the bottom of my heart! Your contributions made this experience an unforgettable ride.

Erklärung nach §17(9) der Prüfungsordnung für den Bachelor-Studiengang Physik und den Master-Studiengang Physik an der Universität Göttingen:

Hiermit erkläre ich, dass ich diese Abschlussarbeit selbständig verfasst habe, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe und alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen wurden, als solche kenntlich gemacht habe.

Darüberhinaus erkläre ich, dass diese Abschlussarbeit nicht, auch nicht auszugsweise, im Rahmen einer nichtbestandenen Prüfung an dieser oder einer anderen Hochschule eingereicht wurde.

Göttingen, den April 5, 2024

(Athul Dev Sudhakar Ponnu)