SAPIENZA
UNIVERSITÀ DI ROMA

The Beauty and Charm Yukawa Couplings of the Higgs Boson with the ATLAS Detector at the LHC

4D Tracking, Particle Flow and Jet Flavour Reconstruction Algorithms Development

Scuola di Dottorato in Scienze Astronomiche Chimiche Fisiche e Matematiche "Vito Volterra"
Dottorato di Ricerca in Fisica (XXXVI cycle)

Lorenzo Santi
ID number 1700424

Advisors	Co-Advisor
Local - Prof. Marumi Kado	Dr. Francesco Di Bello
CERN - Dr. Valentina Cairo	

Academic Year 2023/2024

The Beauty and Charm Yukawa Couplings of the Higgs Boson with the ATLAS Detector at the LHC
PhD thesis. Sapienza University of Rome

This thesis has been typeset by LaTeX and the Sapthesis class.

Author's email: l.santi@cern.ch

Non al denaro non all'amore né al cielo
- F. De Andre

Abstract

This dissertation presents two analyses of Higgs boson processes within the ATLAS experiment at the Large Hadron Collider (LHC). The primary analysis focuses on the legacy $VH(b\bar{b}/c\bar{c})$ process, examining the decays of the Higgs boson into bottom quarks ($H \to b\bar{b}$) and charm quarks ($H \to c\bar{c}$). Utilizing data from Run 2, a multi-variate approach with improved b-tagging algorithms was employed to increase sensitivity. The secondary analysis explores the Di-Higgs process, specifically $HH(b\bar{b})H(\gamma\gamma)$, investigating the simultaneous production of two Higgs bosons decaying into bottom quarks and photons. This analysis aims to provide insights into Higgs self-coupling, contributing to our understanding of the Higgs potential and Electroweak symmetry breaking.

The performance enhancements discussed are crucial to the success of these analyses. Key improvements include advanced Flavour Tagging techniques, a study on the potential impact of 4D tracking, and a Machine Learning-based approach to the Global Particle Flow (PFlow) algorithm.

Flavour Tagging, essential for identifying jets originating from heavy quarks, has seen significant advancements with the DL1 series tagger algorithms and Graph Neural Networks (GNNs), which provide a more precise classification of jet flavors.

4D tracking, incorporating precise timing information in the tracking system, shows that the object reconstruction of particles can significantly impact vertexing and b-tagging algorithms for HL-LHC and beyond.

Global PFlow algorithms, which combine information from various detector components, enable a more accurate reconstruction of the overall event topology.

The dissertation shows that these performance improvements directly reflect on the physics outcomes of the described analyses, bringing us closer to understanding the most fundamental laws of the universe.

Contents

Preface

During my Ph.D., I had the opportunity to work on several aspects of High Energy Physics. As a member of the ATLAS Collaboration at CERN, most of the work presented in this dissertation is the outcome of teamwork projects in which I had a central role. Teamwork is crucial for advancing the knowledge in High Energy Physics. Achieving such results wouldn't have been possible without the contributions of everyone involved, demonstrating the importance of collaborative work in reaching significant milestones.

To become an Author for the ATLAS Collaboration, I performed my *Qualification Task* in 2021 within the Flavour Tagging Group. I measured the b-tagging efficiency, this measurement can also be used as a data-to-Monte Carlo calibration, during the second data-taking period of the LHC (Run2: 2015-2018) using an alternative method that allows to calibrate the efficiency at higher transverse momentum with semi-leptonically decaying top-quarks. I then performed a statistical combination of this calibration with the standard method used in ATLAS. Within the Flavour Tagging group I am one of the Offline Data Quality experts with the task of assessing the quality of the ongoing data taking (Run 3: 2022-2026) in terms of b-tagging performance.

Simultaneously, I joined the Higgs Group, where I was one of the primary analyzers for the VH legacy Run 2 analysis. This analysis aims to simultaneously measure the $H \to b\bar{b}$ and $H \to c\bar{c}$ decay channels of the Higgs boson. The legacy analysis utilizes the same Run 2 data but incorporates new techniques and exploits correlations between different decay channel uncertainties. The main results of this analysis include the first observation of the $WH, H \to b\bar{b}$ process, the best ATLAS limits set on the cross-sections of the $VH, H \to c\bar{c}$ process, and the first observation of the $VZ, Z \to c\bar{c}$ cross-check process. My contributions included incorporating a new region at low transverse momentum of the Higgs boson in the WH channel, deriving the QCD background modelling using data-driven methods, working on the Multi Variate Analysis for signal-background discrimination, and contributing to the development of the statistical analysis for the combined fit.

During the last year of my Ph.D., I joined the Di-Higgs Group, where I am one of the main analyzers for the search that considers two b-quarks and two photons in the final state ($H(b\bar{b})H(\gamma\gamma)$) exploiting on Run 2 and partial Run 3 data. New algorithms and analysis techniques are being developed, and my contributions focus on the impact of the improved b-tagging, particularly at low transverse momentum.

During my second year if my Ph.D, I was awarded a scholarship as a CERN Doctoral Student to work on the potential impact of a 4-Dimensional tracker in ATLAS. We know that after about $2ab^{-1}$ of data at the HL-LHC the innermost part of the ATLAS Inner Tracker (ITk) will be replaced, this leads to a unique opportunity in terms of technological advancements. For this reason we studied the feasibility of such a timing detector in terms of performance, to assess what could be the impact on physics. This work is the first study in ATLAS and it is documented in a PUB Note [1], for which I was the primary author. In this work, I assessed the impact of a potential timing detector, with a spatial and temporal resolution of $\mathcal{O}(10\mu m) \times \mathcal{O}(10ps)$, at the HL-LHC on physics object reconstruction, such as vertexing and b-tagging, demonstrating how performances could improve if ATLAS would implement a timing detector with hermetic coverage in the tracking system.

In parallel with my Ph.D., as a continuation of my Master's work [2] on simultaneous reconstruction using Machine Learning methods, I was part of a team outside ATLAS investigating the potential of Graph Neural Networks in Global Object Reconstruction (Global Particle Flow) in collider experiments. This project culminated in two papers. The first describes a configurable detector simulation of a calorimeter designed for Machine Learning studies [3], and the second introduces a novel advanced technique for Global Particle Flow using Hyperedges Graph Neural Networks tested in a jet environment [4]. In these projects, I was one of the primary authors, contributing to the development and performance assessment of the detector and the algorithms.

Introduction

Despite its name, the Standard Model can be considered both a theory and a model. It aims to describe three of the four fundamental interactions currently discovered by the human kind, albeit not without difficulties:

- Strong Interactions
- Electromagnetic Interactions
- Weak Interactions

Over the past century, experiments and theoretical developments have progressed hand-in-hand, solidifying the SM's dual status. In this context, two complementary perspectives have emerged in the physics of the fundamental interactions: Quantum Field Theory, which focuses from a theoretical point of view on *fields*, and Particle Physics, which emphasizes particle-based experimentation.

The primary tools for probing the Standard Model are particle colliders, capable of achieving incredibly high energies. The pioneering result of Einstein in 1905 [5] relating energy and mass in the equation $E = mc^2$ underscores the principle that mass can be converted into energy and vice versa, explaining how it is possible to create new particles from high-energy collisions. With technological advancements since the mid-20th century, experiments have evolved from analyzing cosmic rays to actively generating particles through collisions of protons or electrons, the only known and available stable particles. At first, particle beams were directed at fixed targets, but it soon became evident that colliding beams could achieve much higher energies.

Bruno Touschek, often considered as the father of colliders, proposed the concept of an electron-positron (e^+e^-) collider in Rome during the 1960s. This innovation served as the precursor to the development of the Large Hadron Collider (LHC).

Before the linear accelerator

The Large Hadron Collider (LHC) is a proton-proton (*pp*) collider constructed by the European Organization for Nuclear Research (CERN) in Geneva between 1998 and 2008. Positioned in a 27 km underground tunnel, the LHC is the largest and most powerful particle collider to date. Since its initial run in 2009, the LHC has progressively reached higher energies and is currently (2024) operating at $\sqrt{s} = 13.6$ TeV in its third run. It occupies the same tunnel previously used by the Large Electron-Positron (LEP) collider, which collided e^+e^- from 1989 to 2000. The LHC is the final stage in a series of progressively larger accelerators that incrementally increase the energy of protons. Each LHC experiment occurs at specific points along the LHC ring, where the proton beams are collided, and data is captured by dedicated experimental setups.

Many significant milestones have been achieved at the LHC, starting with the groundbreaking discovery of the Higgs boson in 2012 by both the ATLAS and CMS experiments independently. This discovery was a monumental achievement, providing the missing piece of the puzzle in the Standard Model of particle physics and confirming the existence of the Higgs field, which is responsible for providing mass to all particles, including the Higgs boson itself.

Since the Higgs Boson's discovery, extensive research has been conducted to explore its properties in greater detail. These studies aim to validate the Higgs boson's predicted behavior and interactions as outlined by the Standard Model. Researchers have focused on measuring its mass, spin, and parity, as well as its coupling with other fundamental particles. These investigations are crucial for determining whether the Higgs boson fits within the Standard Model's framework or if it points to new physics beyond the Standard Model.

Thus the LHC has been crucial in studying the Higgs Boson's rare decay modes and production mechanisms, enhancing our understanding of electroweak symmetry breaking. The continuous data collection and analysis from successive LHC runs have allowed physicists to refine their measurements and improve the precision of these fundamental parameters. These efforts not only solidify our comprehension of the Higgs boson but also open avenues for discovering potential discrepancies that could indicate phenomena of physics beyond the Standard Model.

Thesis Content

This dissertation outlines my research and contributions as a Ph.D. student, a journey I commenced in late 2020 at the University of Rome "*La Sapienza*". Throughout my doctoral studies, I have been primarily based at CERN in Geneva, as I was awarded a scholarship as a Doctoral Student within the ATLAS Collaboration. My involvement with ATLAS has encompassed various aspects of the experiment.

It begins with a detailed theoretical foundation in Chapter 1, where I describe the underlying principles of the Standard Model, including a thorough discussion of Electroweak interactions, the Brout-Englert-Higgs mechanism, and an overview of Quantum Chromo-Dynamics, which governs strong interactions. This theoretical groundwork is essential for understanding how the Higgs boson is investigated at the LHC.

Chapter 2 includes an overview of the LHC experiment, highlighting its significance as the world's largest and most powerful particle collider. Then in the same Chapter, I provide a summarized description of the ATLAS Detector at the LHC, detailing its various subsystems and their roles in particle detection and data collection.

After the theoretical and experimental overview the following Chapters are focused on the work I personally contributed to.

The focus then shifts to Flavour Tagging in Chapter 3, where I delve into the algorithms developed for identifying heavy flavor quarks generated in the collisions and detected by the ATLAS experiment. I discuss the evolution of these algorithms and their calibration on experimental data. In the Section 3.2 a measurement of the b-tagging efficiency on data is shown, orthogonal to the standard measurement used for the calibration, that I personally performed using the Run 2 of the LHC dataset. I also performed the statistical combination with the standard b-tagging calibration used in the Collaboration.

The core of this dissertation is detailed in Chapter 4, which focuses on the experimental aspects of Higgs boson physics. The chapter begins with an overview of the Higgs boson from an experimental perspective, followed by a discussion of previous analyses involving the Higgs boson produced in association with a leptonically decaying vector boson and decaying into heavy flavor quarks, commonly referred to as $VH(bb|cc)$. This chapter presents the legacy analysis using the data from the LHC's second run. As one of the principal analysers, I played a significant role in this research, contributing extensively to the final results of the $VH(bb|cc)$ analysis. I employed a data-driven template method to estimate the QCD background in the decay channel where the vector boson is a W boson. Additionally, I explored and incorporated a previously unprobed region in the Simplified Template Cross Section (STXS) scheme, specifically at low transverse momentum of the Higgs boson when the vector boson is a W. I worked in the development and training of the Boosted Decision Trees (BDTs) to effectively discriminate between signal and background events. These machine learning models were crucial for enhancing the sensitivity of the analysis by improving the separation between the Higgs signal and various background processes. Moreover, I made substantial contributions to developing the statistical analysis strategy for the combined fit of the $VH(bb|cc)$ and the cross-check analysis of $VZ(bb|cc)$. This measurement aims to be the most precise ever performed regarding the Higgs coupling to the b-quark ($VH(bb)$) and establishes the most stringent limits on the Higgs coupling to the c-quark ($VH(cc)$) for ATLAS. Furthermore, it marks the first observation of the exclusive processes $WH(bb)$ and $VZ(cc)$.

During the last year of my Ph.D., I joined the effort in one of the Di-Higgs ATLAS analyses, which aims to probe the Higgs boson self-interaction, a critical aspect of understanding the Brout-Englert-Higgs mechanism and electroweak symmetry breaking. Chapter 5 provides an overview of the significance of this measurement, explaining why it is considered one of the most crucial objectives at the LHC following the Higgs boson discovery. The Di-Higgs analysis discussed in this dissertation focuses on events with two *b*-quarks and two photons in the final state ($H(bb)H(\gamma\gamma)$). This channel is one of the most sensitive to probe the Higgs self-coupling, which is essential for verifying the Standard Model's predictions and searching for new physics. In this chapter, I first present the state-of-the-art analysis utilizing data from the LHC's second run. Then I show my contributions for the ongoing analysis combining the Run 2 and part of the Run 3 the LHC data taking. My work has particularly focused on the impact of low transverse momentum b-tagging, which is crucial for improving the identification of *b*-jets from Higgs decays. Additionally, I explored the potential to simultaneously measure for the first time the single Higgs background in association with a Z boson decaying into *b*-quarks ($Z(bb)H(\gamma\gamma)$). This measurement would benefit from low transverse momentum *b*-tagging because of the softer spectrum of the Z boson.

In Chapter 6, I detail my research as main analyser on the potential impact of integrating a 4D tracker within the ATLAS tracker during the High-Luminosity Phase of the LHC. This upgrade could eventually occur only after the LHC's fourth run. My study highlights the benefits of an hermetic timing detector on key physics objects, such as b-tagging and vertexing, as well as its broader implications for physics analyses. Impacting the physics objects exploited by the above described analyses, timing information in the tracker could have a significant impact on enhancing the Higgs physics program, this is the case of my work as a CERN Doctoral Student.

In the final chapter, Chapter 7, I present two projects on which I collaborated outside of the ATLAS Experiment. One of these projects involved the development of the *A COnfigurable CalOrimeter simulation for AI* (COCOA). The goal of this is to provide a versatile tool for developing Machine Learning algorithms in High Energy Physics, utilizing realistic particle shower simulations. My contributions included developing the simulation framework and assessing the detector's performance. Additionally, this chapter introduces a novel approach to Global Particle Flow, employing advanced Machine Learning techniques based on Graph Neural Networks. In essence, Global Particle Flow aims to reconstruct physics objects simultaneously, streamlining the process by minimizing overlap removal between objects and leveraging redundant information from different sub-detectors. This innovative method is built upon the simulations provided by the COCOA framework. I played a key role in training the models and evaluating their performance. This new approach to Global Particle Flow is built to benefit both current and future collider experiments. This work holds broad implications for the High Energy Physics community, providing a robust framework for advancing particle reconstruction techniques.

1 | Theoretical Overview

A comprehensive theoretical exposition of the Standard Model (SM) is beyond the scope of this dissertation. Nevertheless, a concise overview of the SM and its foundational elements is essential to appreciate the motivations driving contemporary analyses at the LHC. The SM serves as both a theory and a model, encapsulating three of the four known fundamental interactions: Strong, Weak, and Electromagnetic interactions. These interactions are integral to our understanding of particle physics and are intricately woven into the fabric of the SM.

In Section 1.1, we will explore the objectives and historical evolution of the Standard Model, shedding light on its development and the pivotal discoveries that have shaped its current form. This section will provide a foundational understanding necessary to appreciate the subsequent discussions on particle interactions.

Section 1.3 delves into the realm of Quantum Chromo-Dynamics (QCD), the theory describing strong interactions. This section traces the journey from the parton model and Deep Inelastic Scattering to the sophisticated models that depict proton-proton collisions at the LHC. A thorough understanding of QCD is crucial for interpreting collision data and the behavior of quarks and gluons within protons.

Finally, Section 1.2 describes the electroweak interactions within the Standard Model, focusing on the theoretical discovery of the Higgs mechanism. This section will explain how the Higgs field imparts mass to elementary particles thorough the Spontaneous Symmetry Breaking, a cornerstone of the SM, and set the stage for discussing Higgs boson physics and its implications in modern experiments.

1.1 The Standard Model

In the '70s, weak, strong, and electromagnetic interactions had been observed for several years. Within the formalism of Quantum Field Theory (QFT) [6, 7, 8, 9, 10], it has been possible to create a predictive theory unifying the above interactions, capable of predicting new features, such as cross-sections at colliders. This theory also provides the possibility to classify all the particles discovered over the years, resulting in the Standard Model. It is the most fundamental theory in our hands and describes our understanding of elementary particles and the universe [11, 12, 13].

From a particle point of view, one can classify all particles according to their statistics. In particular, it is possible to distinguish fermions obeying the Fermi-Dirac statistic from bosons obeying the Bose-Einstein statistic. Fermions are particles with half-integer spin; all elementary fermions known have a spin of $\frac{1}{2}$. They can be separated into leptons, which participate in weak and electromagnetic interactions, and quarks, which can also interact via the strong interaction. Both quarks and leptons can be divided into three families. Bosons are particles with integer spin, and the known elementary bosons have either spin-1 (vectorial particles) or spin-0 (scalar particle). Elementary vectorial particles are known as gauge bosons, and they are distinguished according to the interaction in which they participate. The only known scalar elementary particle is the Higgs boson. All elementary particles are listed in Figure 1.1.

In the context of Quantum Field Theory, each particle can be described by a quantized field, which is the superposition of creation and annihilation operators that, when acting on the vacuum, can create or annihilate a particle. The Standard Model is based on the breakthrough concept of symmetry within the Lagrangian formalism. As well demonstrated by Nöether, for each continuous symmetry of the Lagrangian, one has a conserved quantity or a conserved current.

Studying hadron decays and conducting Deep Inelastic Scattering experiments, it became possible to develop the theory of strong interactions, leading to Quantum Chromo-Dynamics (QCD). Simultaneously, starting from nuclear β-decay and Fermi's theory in the '30s, the weak theory was developed. Important experimental discoveries include the observation of parity violation in the '50s, the Electroweak unification by Glashow, Weinberg, and Salam in the '60s, exploiting the process of Spontaneous Symmetry Breaking (SSB). This continued with the

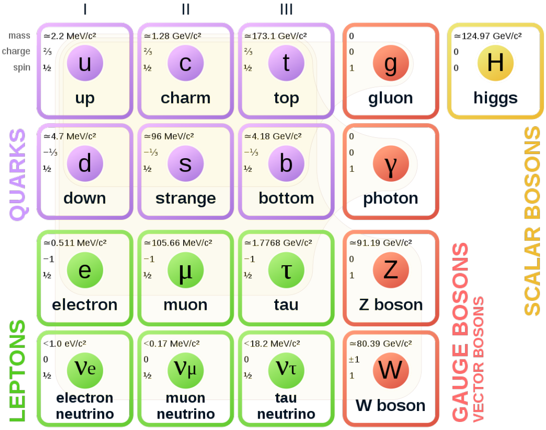


Figure 1.1. List of all elementary particles known in the Standard Model (SM). In purple are three quark families. In green are three lepton families. In red are the gauge bosons: three massive vector bosons $V = W^{\pm}, Z^0$; one massless photon γ and the 8 massless gluons g^a. In Yellow the scalar Higgs boson.

discovery of massive vector bosons in the '80s, culminating in the identification of the Higgs boson in 2012.

Nowadays the Standard Model (SM) can be described by a Lagrangian invariant under three local gauge symmetry groups [14] in the context of Spontaneous Symmetry Breaking (SSB):

$$SU(3)C \otimes SU(2)L \otimes U(1)Y \xrightarrow{SBB} SU(3)C \otimes U(1)Q \qquad (1.1)$$

The strong interactions are described by $SU(3)C$, which is the color symmetry group, while the electroweak interactions are described by $SU(2)L \otimes U(1)Y \rightarrow U(1)Q$, where, in analogy to strong isospin, $SU(2)L$ describes weak isospin, and all fermions are embedded into Left doublets and Right singlets. Finally, to include a weak neutral current, one introduces the hypercharge symmetry $U(1)_Y$. The arrow signifies that the vacuum of the theory doesn't have the same symmetry as the Lagrangian, indicating that the theory has a Spontaneously Broken Symmetry. The surviving symmetry is the known as the charge $U(1)_Q$, resulting with a single gauge boson remaining massless. This mechanism is known as the Higgs mechanism.

1.2 Electro-Weak Interactions and Higgs Mechanism

The Electro-Weak (EW) interactions and the Higgs boson searches are the main topic of this dissertation. For this reason a theoretical overview is necessary to understand the physical and historical developements of this fundamental theory.

Weak interactions at low energies are well described by the Fermi *current – current* interaction [15] which can be extended including axial currents after the discovery of parity violation. The first thing that it points out is that experimentally only left-handed helicity neutrinos and right-handed helicity antineutrinos are involved in weak interactions. In this sense as far as it is know one can use chiral states to describe leptons within the context of weak interactions. In particular it is well known that only left chiral states are selected in weak interactions. Thus the weak currents can be described by:

$$j^{\mu}_{+} = \bar{e}\gamma^{\mu}\frac{1}{2}(1-\gamma_5)\nu \qquad j^{\mu}_{-} = \bar{\nu}\gamma^{\mu}\frac{1}{2}(1-\gamma_5)e \tag{1.2}$$

The same current can be written for quark families instead of leptons. It is immediately evident that this current describes either a left-handed neutrino or a right-handed antineutrino.

The problems arose when neutral currents were observed coupling not only with left-handed particles but even with right-handed particles (such as e_R, u_R and d_R, right-handed neutrino is not considered in the SM):

$$j^{\mu}_{0} = \bar{q}\gamma^{\mu}\frac{1}{2}(c^q_V - c^q_A\gamma_5)q \tag{1.3}$$

This suggests that weak interactions can be described by a gauge theory. And for sake of simplicity only leptons are considered, quarks will be introduced for later purposes:

$$e_L = \frac{1}{2}(1-\gamma_5)e \qquad e_R = \frac{1}{2}(1+\gamma_5)e \qquad \nu_L = \frac{1}{2}(1-\gamma_5)\nu \tag{1.4}$$

The fields in any representation of the gauge group must have the same Lorentz-transformation properties, hence the representations of the gauge group divide into a left-handed doublet and a right-handed singlet:

$$L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} \qquad R = e_R \tag{1.5}$$

As mentioned before the largest possible gauge group is:

$$SU(2)_L \otimes U(1)_Y \tag{1.6}$$

The idea of electroweak theory [16, 17] is to use a mixing of third generator of $SU(2)_L$ and the generator of $U(1)_Y$ to get a neutral massless photon (γ) and a neutral massive vector boson (Z^0). Under this group the fields transform as:

$$U_{SU(2)} = e^{ig\frac{\boldsymbol{\sigma}}{2}\boldsymbol{\alpha}(x)} \qquad U_{U(1)} = e^{ig'\frac{Y}{2}\beta(x)} \tag{1.7}$$

Where σ_i are the Pauli matrices and Y is the Hypercharge operator. At this point Y can be used in linear combination with $T_3 = \sigma_3/2$ to obtain the charge operator Q with the famous Gell-Mann relation already known for $SU(3)_F$:

$$Q = T_3 + \frac{Y}{2} \tag{1.8}$$

For all the fermions the values are displayed in table 1.1.

To explain the procedure described previously one must observe that since the weak currents in eq. 1.2 can be rewritten in terms of generators of the symmetry group $SU(2)_L$ it is possible to include an extra group $U(1)_Y$ in order to obtain the neutral current.

f	Q	T_3	Y
ν_L	0	1/2	-1
e_L	-1	-1/2	-1
e_R	-1	0	-2
u_L	2/3	1/2	1/3
d_L	-1/3	-1/2	1/3
u_R	2/3	0	4/3
d_R	-1/3	0	-2/3

Table 1.1. For the first family of leptons and quarks doublet the values of charge (Q), weak isospin (T_3) and hypercharge (Y) are shown.

The reason is that the third generator of $SU(2)_L$ gives a current coupling only to left-handed particles and therefore can't represent the required neutral current. It is reasonable to include the electromagnetic current in the theory ($j^{\mu}_{em} = \bar{L}\gamma^{\mu}QL + \bar{R}\gamma^{\mu}QR$) which couples also to right, but it is not invariant under $SU(2)_L$. Thus the idea is to take two orthogonal combination of j^{μ}_{em} and j^{μ}_{0}. One combination complete the weak isospin triplet j_3^μ, while the second combination must be invariant under $SU(2)_L$ that is j_Y^μ given by the Gell-Mann relation:

$$j_{em}^\mu = j_3^\mu + \frac{1}{2} j_Y^\mu \tag{1.9}$$

To do so another symmetry group ($U(1)_Y$) had been introduced. A gauge boson field for each generator of each symmetry group appears in the theory: 3 $\mathbf{W}^\mu(x)$ for $SU(2)_L$ and a single B_μ for $U(1)_Y$. Similarly to non abelian QCD Lagrangian, it is possible to write the Electroweak Lagrangian:

$$\mathcal{L} = -\frac{1}{4}\mathbf{W}^{\mu\nu} \cdot \mathbf{W}_{\mu\nu} - \frac{1}{4} B^{\mu\nu} B_{\mu\nu} + \bar{L} i \not{D} L + \bar{R} i \not{D} R \tag{1.10}$$

In this Lagrangian the covariant derivative acts differently between left-handed and right-handed fields and from tab. 1.1 it is clear that $Y_L = -1$ and $Y_R = -2$. The covariant derivative definition is:

$$D^\mu L = (\partial^\mu + ig \frac{\boldsymbol{\sigma}}{2} \mathbf{W}^\mu - ig' \frac{1}{2} B^\mu) L \qquad D^\mu R = (\partial^\mu - ig' B^\mu) R \tag{1.11}$$

From each gauge boson it is possible to construct its tensor field and its infinitesimal gauge transformation as:

$$\begin{aligned} \mathbf{W}^{\mu\nu} &= \partial^\mu \mathbf{W}^\nu - \partial^\mu \mathbf{W}^\nu - g \mathbf{W}^\mu \wedge \mathbf{W}^\nu \\ \mathbf{W}'_\mu &= \mathbf{W}_\mu - \frac{1}{g} \partial_\mu \boldsymbol{\alpha}(x) - \boldsymbol{\alpha}(x) \wedge \mathbf{W}_\mu \\ B^{\mu\nu} &= \partial^\mu B^\nu - \partial^\mu B^\nu \\ B'_\mu &= B_\mu - \frac{1}{g'} \partial_\mu \beta(x) \end{aligned} \tag{1.12}$$

In this theory the main missing thing is a mechanism to give the mass to the gauge boson fields and to the fermions. Certainly one of the four gauge fields, or more accurately the linear combination generating the electromagnetic field A_μ coupled with j_{em}^μ, must remain massless. In this sense the assumption is that $SU(2)_L \otimes U(1)_Y$ is spontaneously broken to $U(1)_Q$ [18, 19, 20]. The details of this procedure are explained by the Higgs mechanism [7, 21]. Whatever the mechanism is the fields corresponding to charged vector bosons with mass m_W and coupled with weak currents in 1.2 are a combination of W_1^μ and W_2^μ:

$$W_\pm^\mu = \frac{1}{\sqrt{2}} (W_1^\mu \mp i W_2^\mu) \tag{1.13}$$

At the same time two electrically neutral fields of mass m_Z and zero are obtained from a rotation of W_3^μ and B^μ. The angle for this rotation is the so called Weinberg angle or θ_W. The rotation is needed because, as previously said, the combination of j_3^μ and j_Y^μ to get j_0^μ and j_{em}^μ must be orthogonal:

$$\begin{pmatrix} A^\mu \\ Z^\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B^\mu \\ W_3^\mu \end{pmatrix} \tag{1.14}$$

Finally to get a proper unification it is necessary to express the weak couplings in terms of the electromagnetic coupling e. The lepton Lagrangian at this point can be rewritten as:

$$\begin{aligned} \mathcal{L} =& \bar{L} i \not{\partial} L + \bar{R} i \not{\partial} R + \\ & - \frac{g}{\sqrt{2}} (j_+^\mu W_\mu^+ + j_-^\mu W_\mu^-) + \\ & - (g \sin\theta_W j_3^\mu + g' \cos\theta j_Y^\mu) A_\mu \\ & - (g \cos\theta_W j_3^\mu - g \sin\theta j_Y^\mu) Z_\mu \end{aligned} \tag{1.15}$$

Immediately it is clear that the Fermi coupling constant can be rewritten in terms of g and that it can be obtained from the propagator of a massive gauge boson for small transferred momenum:

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8 m_W^2} \tag{1.16}$$

Knowing that the current coupled to A_μ must be the electromagnetic current expressed in the relation 1.9 and that the coupling constant must be e:

$$g \sin\theta_W = g' \cos\theta_W = e \qquad e = \frac{g g'}{\sqrt{g'^2 + g^2}} \tag{1.17}$$

The Lagrangian can now be expressed as:

$$\begin{aligned} \mathcal{L} =& \bar{L} i \not{\partial} L + \bar{R} i \not{\partial} R + \\ & - \frac{e}{\sqrt{2}\sin\theta_W} (j_+^\mu W_\mu^+ + j_-^\mu W_\mu^-) + \\ & - e j_{em}^\mu A_\mu \\ & - \frac{e}{\sin 2\theta_W} (j_3^\mu - \sin^2\theta_W j_{em}^\mu) Z_\mu \end{aligned} \tag{1.18}$$

The amazing fact is that a gauge boson field of neutral charge is coupled asymmetrically to left-handed and right-handed fields. Comparing with the current j_0^μ in eq. 1.3 it is possible to extract the values reported in tab. 1.2 of $c_V = T_3 - 2Q\sin^2\theta_W$ and $c_A = T_3$ for all the fermions. To complete the theory some assumptions about the spontaneous symmetry breaking are necessary. As it might be noticed not only to the $W^\pm$ and Z_0 need a mass, but also fermions as well. This must holds since a mass term is not gauge invariant and with this mechanism fermions acquire a mass dynamically. The only renormalizable way is to have a scalar field coupled without derivatives to L and R. Requiring the invariance of this Yukawa term in the Lagrangian is equivalent to require the scalar to be an $SU(2)_L$ complex doublet with $Y = 1$ properly chosen:

$$\phi(x) = \begin{pmatrix} \phi^+(x) \\ \phi^0(x) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} Re\phi^+(x) + iIm\phi^+(x) \\ Re\phi^0(x) + iIm\phi^0(x) \end{pmatrix} \tag{1.19}$$

A gauge invariant term involving the scalar and the gauge fields can be added. The most general form is, including the Yukawa part:

f	c_V	c_A	c_L	c_R
ν_L	$1/2$	$1/2$	$1/2$	0
e_L	$-1/2 + 2\sin\theta_W$	-1/2	$-1/2 + \sin^2\theta_W$	$\sin^2\theta_W$
e_R	$2\sin\theta_W$	0	$\sin^2\theta_W$	$\sin^2\theta_W$
u_L	$1/2 - \frac{4}{3}\sin\theta_W$	$1/2$	$1/2 - \frac{2}{3}\sin\theta_W$	$-\frac{2}{3}\sin\theta_W$
d_L	$-1/2 + \frac{2}{3}\sin\theta_W$	-1/2	$-1/2 + \frac{2}{3}\sin\theta_W$	$\frac{1}{3}\sin\theta_W$
u_R	$-\frac{4}{3}\sin\theta_W$	0	$-\frac{2}{3}\sin\theta_W$	$-\frac{2}{3}\sin\theta_W$
d_R	$\frac{1}{3}\sin\theta_W$	0	$\frac{2}{3}\sin\theta_W$	$\frac{1}{3}\sin\theta_W$

Table 1.2. Values for the axial and vectorial fermion couplings to the Z_0. Also the values for the left and right fermion couplings are reported, where $c_L = (c_V + C_A)/2$ and $c_R = (c_V - c_A)/2$.

$$\begin{aligned} \mathcal{L} = |D_\mu\phi(x)|^2 + m^2\phi^\dagger(x)\phi(x) \\ - \lambda(\phi^\dagger(x)\phi(x))^2 - G_e[\bar{L}\phi R + \bar{R}\phi^\dagger L] \end{aligned} \tag{1.20}$$

To have a lower bound on the energy the parameter λ is required to be positive. Then the choice $m^2 > 0$ needs to break the symmetry. It means that the Lagrangian still is gauge invariant, but its vacuum is not. The potential defined is the complex version of the shape shown in figure. 1.2. The tree-approximation for the vacuum expectation value (VEV or v) is the minimal of the potential $V(\phi) = -m^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2$:

$$\langle|\phi|^2\rangle = \frac{v^2}{2} = \frac{m^2}{2\lambda} \tag{1.21}$$

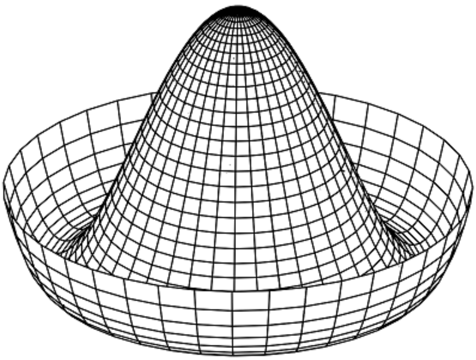


Figure 1.2. Shape of the potential for a spontaneous symmetry breaking.

From the goldstone theorem each broken generator of the symmetry, namely each generator which doesn't leave the vacuum invariant, generates a massless field. In this case there are 3 goldstone bosons. Since the Lagrangian is still invariant under $SU(2)_L \otimes U(1)_Y$ a gauge transformation can always be performed. The idea is to apply the so called unitary gauge in which $\phi^+ = 0$ and ϕ^0 is hermitian or real with a positive VEV. In some sense the goldstone boson degrees of freedom disappear in this gauge in favor of express the theory in terms of physical fields. In the unitary gauge the VEV and the field associated to the remaining scalar field (the Higgs) are:

$$\langle\phi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v\end{pmatrix} \qquad \phi = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v+H(x)\end{pmatrix} \tag{1.22}$$

Using this gauge on the kinetic term of the scalar field the mass terms for the vector bosons appear together with the Higg-Vector Boson interactions:

$$\begin{aligned}|D_\mu\phi(x)|^2 &= \frac{1}{2}|\partial_\mu\phi|^2 + \frac{1}{8}(v+H(x))^2\times\\ &\times\left[2g^2W_+^\mu W_{-\mu} + \begin{pmatrix}W_{3\mu} & B_\mu\end{pmatrix}\begin{pmatrix}g^2 & -gg'\\ -gg' & g'^2\end{pmatrix}\begin{pmatrix}W_3^\mu\\ B^\mu\end{pmatrix}\right]\end{aligned} \tag{1.23}$$

It is immediate to cross check that the diagonalization of the matrix in this expression gives the correct rotation of the fields expressed in eq. 1.14. It is evident that the symmetry is broken with a massless gauge boson, the photon A_μ. The mass of the gauge fields can be written in terms of the gauge coupling constant and the VEV obtaining:

$$m_W = \frac{vg}{2} \qquad m_Z = \frac{vg}{2\cos\theta_W} \tag{1.24}$$

From the Yukawa term there is the possibility to give a mass to the electron-like fermions and it is also noticed that the coupling of fermions to the Higgs is proportional to their mass, as well as for the gauge bosons:

$$m_e = \frac{G_e v}{\sqrt{2}} \tag{1.25}$$

In order to give the mass to the neutrino-like the charge conjugate Higgs field mus be introduced:

$$\phi_c(x) = -i\sigma_2\phi^*(x) = \begin{pmatrix} -\bar{\phi}^0(x) \\ \phi^-(x) \end{pmatrix} \xrightarrow{SBB} \frac{1}{\sqrt{2}}\begin{pmatrix} v + H(x) \\ 0 \end{pmatrix} \tag{1.26}$$

If $A \xrightarrow{SBB} B$ boh Let's note that G_e is arbitrary and, differently from the vector boson masses, the fermion masses are not predicted. It is easy to extend this method to all the families of fermions. Phenomenologically it is possible to deduce the VEV value from the Fermi constant measure in muon decays:

$$\begin{gathered} G_F = 1.1663787(6) \times 10^{-5}\ GeV^{-2} \\ v = \frac{1}{2^{1/4}G_F^{1/2}} = 246.219651(11)\ GeV \end{gathered} \tag{1.27}$$

Then it is possible to express all the constant as a function of the best measured values in the standard model such as G_F or α. A relation between the vector boson masses also occurs. Using the value of electromagnetic coupling at the electroweak scale ($\alpha(m_V) = 1/128$):

$$\begin{gathered} \rho = \frac{m_W}{\cos\theta_W m_Z} = 1 \\ m_W = \frac{ve}{2\sin\theta_W} = \frac{38.54\ GeV}{\sin\theta_W} \\ m_W = \frac{ve}{2\sin\theta_W} = \frac{77.08\ GeV}{\sin 2\theta_W} \end{gathered} \tag{1.28}$$

From purely leptonic neutral processes such as $\nu_\mu(\bar{\nu_\mu}) + e^- \rightarrow \nu_\mu(\bar{\nu_\mu}) + e^-$ it has been possible to obtain $sin^2(\theta_W) = 0.231$ which would give, without higher order corrections $m_W = 80.2\ GeV$ and $m_Z = 91.4\ GeV$.

Finally the potential can be written as the sum of mass, trilinear self-interaction and quartic self-interaction terms:

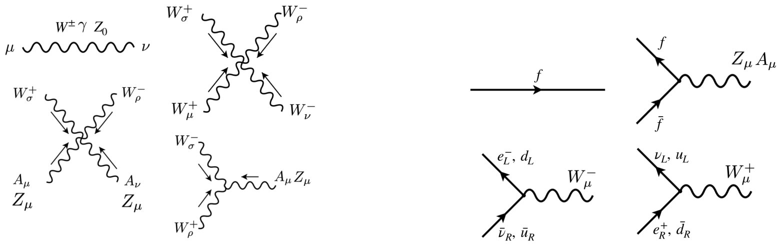
$$V(H) = \frac{1}{2}(2\lambda v^2)H^2(x) + \lambda v H^3(x) + \frac{\lambda}{4}H^4(x) \tag{1.29}$$

The Higgs mass, $M_H^2 = 2\lambda v^2$ at tree level, is not predicted in the Standard Model. Therefore the electroweak Lagrangian is summarized as in Table 1.3.

$\mathcal{L}_{kg} + \mathcal{L}_{gg} =$ $= -\frac{1}{4}\mathbf{W}^{\mu\nu} \cdot \mathbf{W}_{\mu\nu} - \frac{1}{4}B^{\mu\nu}B_{\mu\nu}$	$W^{\pm}$, Z_0 and γ kinetic terms self interactions
$\mathcal{L}_{kL} + \mathcal{L}_{Lg} = \bar{L}i\not{D}L + \bar{R}i\not{D}R$	Lepton: kinetic and gauge interaction
$\mathcal{L}_{kq} + \mathcal{L}_{qg} =$ $= \bar{Q}_L i\not{D} Q_L + \bar{u}_R i\not{D} u_R + \bar{d}_R i\not{D} d_R$	Quark: kinetic and gauge interaction
$\mathcal{L}_{kH} + \mathcal{L}_{mg} + \mathcal{L}_{Hg} = \lvert D_\mu \phi \rvert^2 - V(\phi)$	Higgs: kinematic and interactions with gauge bosons
$\mathcal{L}_{mL} + \mathcal{L}_{LH} = -G_e[\bar{L}\phi R + h.c.]$	Higgs interaction with Leptons
$\mathcal{L}_{mu} + \mathcal{L}_{uH} = -G_u[\bar{Q}_L \phi_c u_R + h.c.]$	Higgs interaction with Up-like Quarks
$\mathcal{L}_{mq} + \mathcal{L}_{qH} = -G_d[\bar{Q}_L \phi d_R + h.c.]$	Higgs interaction with Down-like Quarks

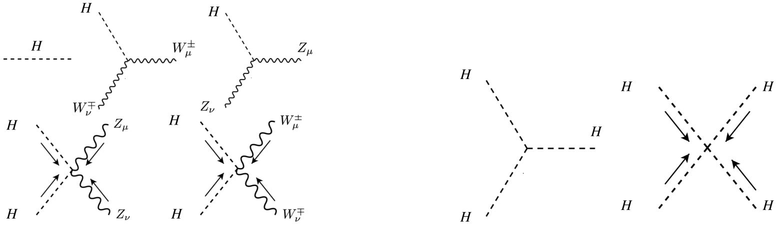
Table 1.3. Lagrangian pieces

Even before the discovery of neutral currents, the electroweak theory was extended to the interactions of hadrons with each other and with leptons. From the non conservation of the strangeness in the model of $SU(3)_F$ Cabibbo [22, 23] proposed a mixing (rotation)



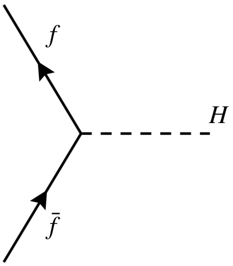
(a) Gauge boson self interaction

(b) Fermion interaction with gauge bosons



(c) Higgs interactions with gauge bosons

(d) Higgs self interaction



(e) Higgs interactions with fermions

Figure 1.3. Elementary Feynman vertices and propagator of Standar Model.

of down and strange quarks writing the quark doublet as:

$$Q_L^1 = \begin{pmatrix} u_L \\ d_L \cos\theta_C + s_L \sin\theta_C \end{pmatrix} \tag{1.30}$$

Assuming only this quark doublet and even excluding a neutral current exchange this model would lead to rates for processes like $K^0 - \bar{K}^0$

oscillations computed with the loop in fig. 1.4 much greater than observed of magnitude [24]. To avoid this difficulty Glashow, Iliopoulos and Maiani [25] proposed the existence of another up-like quark the c quark filling the second family of quarks which cancels the contribution of the u quark in the loop. The second left-handed quark doublet can be written as:

$$Q_L^2 = \begin{pmatrix} c_L \\ -d_L \sin\theta_C + s_L \cos\theta_C \end{pmatrix} \tag{1.31}$$

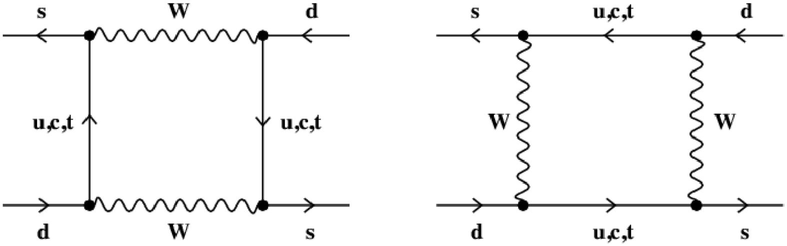


Figure 1.4. Contribution to neutral Kaons oscillations.

The discovery of the τ was the sign of the existence of a third generation. Soon after a fifth quark type (the b) was discovered and the sixth became theoretically necessary and after a long time it too was discovered. At the end the hadronic current can be expressed as:

$$j^{\mu}_{had} = \begin{pmatrix} \bar{u} & \bar{c} & \bar{t} \end{pmatrix} \gamma^{\mu} \frac{1}{2}(1-\gamma_5) V_{ckm} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.32}$$

The matrix V_{ckm} is the Cabibbo, Kobaiashi and Maskawa matrix [26]. It is extension to three generations of the quark mixing. The matrix is unitary. This is just what one would naturally expect with three quarks generations in the Higgs mechanism. The 3×3 matrix V_{ckm} appears in the fermion-gauge boson interaction in order to diagonalize the mass terms after the SSB of the quark Yukawa Lagrangian:

$$\begin{aligned} \mathcal{L}_{mq} = -\sum_{ij} G^d_{ij} [\begin{pmatrix} \bar{U}^i_L & \bar{D}^i_L \end{pmatrix} \phi d^j_R + h.c.] - \\ - \sum_{ij} G^u_{ij} [\begin{pmatrix} \bar{U}^i_L & \bar{D}^i_L \end{pmatrix} \phi_c u^j_R + h.c.] \end{aligned} \tag{1.33}$$

With only two generations (with u, d, c and s quarks) it is possible to chose the phases of the quark fields in order to have a real 2×2 matrix or rather an orthogonal matrix as made with the GIM mechanism and the Cabibbo rotation. As a consequence Charge conjugation and Parity together (CP) are simultaneously conserved . The great importance of the third quark generations is that it is not possible to choose the quark phases so that the V is real and therefore the electroweak interactions can violate CP conservation.

$$V_{ckm} \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1.34}$$

Since the third generation is weakly coupled with the first two the Cabibbo assumption works well and the CP violations are weak.

1.3 Quantum Chromo-Dynamics and Strong Interactions

As already mentioned in Section 1.1, in order to have the possibilities of studying the Higgs boson processes and all its properties described in Section 1.2 the Large Hadron Collider had been built. For this reason its fundamental to understand how protons interact at LHC to produce the interested events. These interactions are mediated by the Quantum Chromo-Dynamics and the so called strong interactions.

In this section, the strong interactions are discussed in the context of the SM [8, 11, 27, 28]. The theory of strong interactions is known as Quantum Chromo-Dynamics (QCD) and it is described by $SU(3)_C$ gauge symmetry. Each strong interacting particle carries a quantum number known as the color number. Fermions carrying color numbers are called quarks and they were postulated by Gell-Mann in 1964 [29, 30] to explain the spectrum of the strongly interacting particles known as hadrons. In this framework, mesons can be described as quark-antiquark bound states while baryons are bound states of three quarks. To make hadrons with integer charge a fractional electric charge for quarks is needed. Quarks appear in 6 different flavours placed into 3 generations: up (u), down (d), charm (c) strange (s), top (t), bottom (b). The up-like quarks have $Q = 2/3$ while the down-like quarks have $Q = -1/3$. Then the proton can be seen as a bound state of uud while the neutron as udd. The symmetry in the interchange of u and d is the strong isospin symmetry $SU(2)_I$. It can be enlarged to the approximate symmetry $SU(3)_F$ including the strange quark. All the baryon and meson multiplets fill out irreducible representations of this symmetry. To reconcile baryons as $\Delta^{++} = uuu$ with the spin-statistic theorem it has been proposed a new quantum number carried by each quark: the Color.

Thus quarks transform under the fundamental representation (3) of the $SU(3)_C$ while anti-quarks transform as the $\overline{3}$ representation. The phenomenological assumption is that no colored particle can be observed as a free state and this postulate is called confinement of color. It means that all hadrons must be invariant (singlets) of $SU(3)_C$ and the most stable combinations are quark-antiquark bound states and three quark bound states:

$$\bar{q}^i q_i \qquad \epsilon^{ijk} q_i q_j q_k \tag{1.35}$$

To test these properties $e-p$ Deep Inelastic Scattering (DIS) experiments were performed. Since the color symmetry had no other obvious physical role it was natural to identify this symmetry with a gauge group. The consequence is the appearance of a gauge field transforming with the adjoint (8) of the group. Thus this model predicted the existence of a field known as gluon which could take 8 color quantum numbers. In a non-abelian theory, it is possible to prove the asymptotic freedom [31, 32] namely the coupling constant (α_s for QCD) becomes penetrative in the high-energy regime .

The Lagrangian of QCD is the following:

$$\mathcal{L} = -\frac{1}{4} F^{\mu\nu a} F^a_{\mu\nu} + \bar{q}_i (i\not{D} - m_q)_{ij} q_j$$
$$a = 1...8 \qquad i, j = 1...3 \tag{1.36}$$

In this expression, the covariant derivative had been defined as $(D_\mu)_{ij} = \delta_{ij}\partial_\mu + ig_3 A^a_\mu \lambda^a_{ij}$ to ensure the invariance of $\mathcal{L}$ under $SU(3)_C$. The matrices λ^a_{ij} are the 8 Gell-Mann matrices that are the generators of the symmetry group. Moreover, the field tensor can be constructed by the commutator of covariant derivatives to get an invariant tensor obtaining:

$$F^a_{\mu\nu}(x) = \partial_\mu A_\nu(x)^a - \partial_\nu A_\mu(x)^a - g_3 f^{abc} A^b_\mu(x) A^c_\nu(x) \tag{1.37}$$

Thus the Lagrangian is invariant under the following local transformations of $SU(3)_C$ of fundamental representations (quarks):

$$q'_i(x) = U_{ij} = e^{i\frac{\lambda^a_{ij}}{2}\theta^a(x)} q_j(x) \tag{1.38}$$

and for adjoint representation (gluons):

$$A_\mu^{a\prime}(x) = A_\mu^a(x) - \frac{1}{g_3}\partial_\mu\theta^a(x) - f^{abc}\theta^b(x)A_\mu^c(x) \tag{1.39}$$

It is immediately evident that a non abelian group, Yang-Mills theory [33], produces a self-interacting gauge boson. In particular from the kinetic term one can observe trilinear and quartic self couplings. This feature is not possible in QED where the gauge symmetry group is the abelian $U(1)_Q$.

Starting from *ep* elastic scattering it is possible to test the inelastic region. Assuming the proton not point like, it is possible to compute the expression of the differential cross section in the Laboratory (LAB) frame. Firstly the kinematic variables of the scattering are defined. The energy of the incoming and outgoing electron are E_1 and E_3 while the angle with respect to the incoming electron is θ. Finally the the differential cross-section results to be:

$$\begin{aligned}\frac{d\sigma_{ep}}{d\Omega}|_{LAB} = & \frac{\alpha^2}{4E_1^2 sin^4(\frac{\theta}{2})}\frac{E_1}{E_3}\times \\ & \times [\frac{G_E(q^2) + \tau G_M(q^2)}{1+\tau}cos^2(\frac{\theta}{2}) + 2\tau G_M(q^2)sin^2(\frac{\theta}{2})]\end{aligned} \tag{1.40}$$

In this expression, the momentum transferred to the proton is $q^2 = (p_1 - p_3)^2 = -E_1E_3(1 - cos(\theta))$ with $\tau = -q^2/4M_P^2$. The electric and magnetic form factors ($G_E(q^2)$ and $G_M(q^2)$) can be used to understand the dimension and the structure of the proton. It is important to understand that in the case of point-like particles $G_E(q^2) = G_M(q^2) = 1$. From experiments rises the evidence of the internal structure of the proton obtaining that:

$$G_E(q^2) = \frac{G_M(q^2)}{2.79} = \frac{1}{(1 - \frac{q^2}{0.71\ GeV^2})^2} \tag{1.41}$$

Which corresponds to a radius of $r_P = 0.81\ fm$. In the case of elastic scattering $p_4^2 = M_P^2 = (q+p_2)^2$ and as a consequence $q^2 = -2M_P(E_1 - E_3)$. Only in this case, the kinematic variables are functions only of the angle θ. The expressions of $E_3(\theta)$ and $Q^2(\theta) = -q^2(\theta)$ for a given fixed E_1 are:

$$E_3(\theta) = \frac{E_1M_P}{M_P + E_1(1 - cos(\theta))} \tag{1.42}$$

$$Q^2(\theta) = \frac{2E_1^2M_P(1 - cos(\theta)}{M_P + E_1(1 - cos(\theta))} \tag{1.43}$$

Since there is only one free parameter, the cross section in case of point-like target as a function of a new variable, the fraction of energy loss by the electron defined as:

$$y = \frac{p_2 \cdot q}{p_2 \cdot p_1} = (1 - \frac{E_3}{E_1})_{LAB} = (sin^2(\frac{\theta}{2}))_{CM} \tag{1.44}$$

Increasing the energy the electrons interacts with a point-like constituent of the proton, that is a parton (either a quark or a gluon):

$$\frac{d\hat{\sigma}_{eq}}{dQ^2} = \frac{4\pi\alpha^2}{Q^4}[1 - y + \frac{1}{2}y^2]e_q^2 \tag{1.45}$$

In this expression, the possibility to have a fractional charge e_q had been taken into account. To express the inelasticity of the scattering one must assume that in the *ep* scattering the final state is no more a proton: $p_4^2 = M_X^2 > M_P^2$ where X as reported in fig. 1.5 is the set of products of an inelastic collision. Therefore a new independent variable describing the inelasticity of the collision can be built, the Bjorken variable which is defined as follows:

$$x = \frac{Q^2}{2p_2 \cdot q} \to 1 \tag{1.46}$$

At this point it is possible to introduce the Parton Distribution Functions (PDF), fundamental for the LHC physics, which are the probability to find a parton (q) with a fraction of momentum f_q of the proton. It is possible to show that in the Infinite Momentum Frame (IMF), where partons looks at rest in the transverse direction of motion during the collision, the fraction f_q corresponds to the Bjorken x. The immediate consequence is that the cross section of DIS must coincide with the scattering over all the possible point-like constituents of the proton:

$$\begin{aligned} \frac{d^2\sigma(ep)}{dxdQ^2} &= \frac{4\pi\alpha^2}{Q^4}[(1-y)\frac{F_2(x,Q^2)}{x} + y^2 F_1(x,Q^2)]e_q^2 = \\ &= \sum_q \frac{d\hat{\sigma}_{eq}}{dQ^2} f_q(x) \end{aligned} \tag{1.47}$$

The main results of this naive parton model are the Bjorken scaling and the Callan-Gross relations whose experimental signatures showed respectively that quarks are point-like and that they carry spin-$\frac{1}{2}$:

$$F_2(x,Q^2) = 2xF_1(x,Q^2) = \sum_q x f_q(x) \tag{1.48}$$

This is both a confirm of $SU(3)_F$ and $SU(3)_C$. Experimentally the fraction proton momentum carried by the up quark ($f_u = \int_0^1 x f_u(x)$) is twice the the one carried by the down quark (f_d) and at the same time something else (gluons) are missing since $f_u + f_d \approx 0.5$.

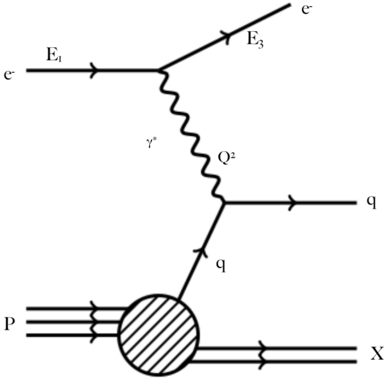


Figure 1.5. Representation of $electron - proton$ Deep Inelastic Scattering (DIS).

To simplify the notation $f_q(x) = q(x)$ for quarks and $f_g(x) = g(x)$ for gluons. Including the gluon emission in the naive parton model the scaling is violated, meaning that $q(x) \to q(x, Q^2)$. Then it is possible to construct coupled evolution equations for PDFs (Altarelli-Parisi equations [34]) which represent the non perturbative part of QCD in scattering processes. They are necessary for the LHC experiment and all the proton proton colliders.

$$\begin{aligned} \frac{dq(x,Q^2)}{d\ln Q^2} &= \int_x^1 \frac{dy}{y}[q(\frac{x}{y},Q^2)P_{qq}(\frac{x}{y}) + g(\frac{x}{y},Q^2)P_{qg}(\frac{x}{y})] \\ \frac{dg(x,Q^2)}{d\ln Q^2} &= \int_x^1 \frac{dy}{y}[g(\frac{x}{y},Q^2)P_{gg}(\frac{x}{y}) + \sum_q q(\frac{x}{y},Q^2)P_{gq}(\frac{x}{y})] \end{aligned} \tag{1.49}$$

In this expression $P_{ab}(z)$ is the splitting function, or the probability to find the parton a with fraction of proton momentum $x = zy$ generated from an initial parton b with a proton fraction momentum y. The peculiar fact is that $P_{gg}(z)$ has a non-δ distribution. This confirms the presence of self-interacting gluons.

As previously said PDFs cannot be computed perturbatively and they must be measured in scattering experiments. AP equations let

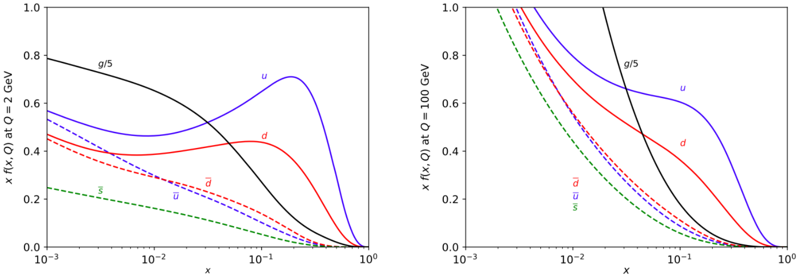


Figure 1.6. Fitted measured values for PDFs of proton in a range of $10^{-3} < x < 1$ for $Q = 2\ GeV$ (**left**) and $Q = 100\ GeV$ (**right**). The difference are explained by the Altarelli-Parisi equations. The peaks at 1/3 for u(x) and d(x) together with their proportions are the evidence of $SU(3)_F$.

us to predict the values of PDFs at higher values of Q^2 measuring them at low Q^2. At the end to give an expression for an LHC process two arbitrary scales must be included: the factorization scale (μ_f) which defines the independence of the perturbative and unperturbative part of the process and it is arbitrary to avoid Infrared (IR) divergences; the renormalization scale (μ_r) which avoid the Ultraviolet (UV) divergences at more than one loop calculations in the parton cross sections. In the end a proton-proton collision can be displayed as in fig. 1.7. As it can be seen immediately several soft and hard interaction can happen in the same collision. All these effects must be taken into account with Monte Carlo simulations, starting from the hard scattering parton collision which can be computed nowadays up to NNLO obtaining:

$$\begin{aligned}\sigma_{pp\rightarrow X} &= \\ &= \sum_{ab}\int_0^1 dx_a dx_b f_a(x_a,\mu_f^2) f_b(x_b,\mu_f^2)\hat{\sigma}_{ab\rightarrow X}(x_a,x_b,\mu_r^2,\mu_f^2) \qquad (1.50)\end{aligned}$$

Subsequently, because of asymptotic freedom, which is shown in Figure 1.8, two other phenomena are considered: Parton Showering and the Hadronization. The first represents the production of several partons from a native parton until Q^2 is high enough in each parton produced. The latter represent the hadron formation from colored partons due to color confinement of low energy partons. The set of particles produced from a native parton is called Jet. Thus in the end there is a Jet for each parton produced during the collision.

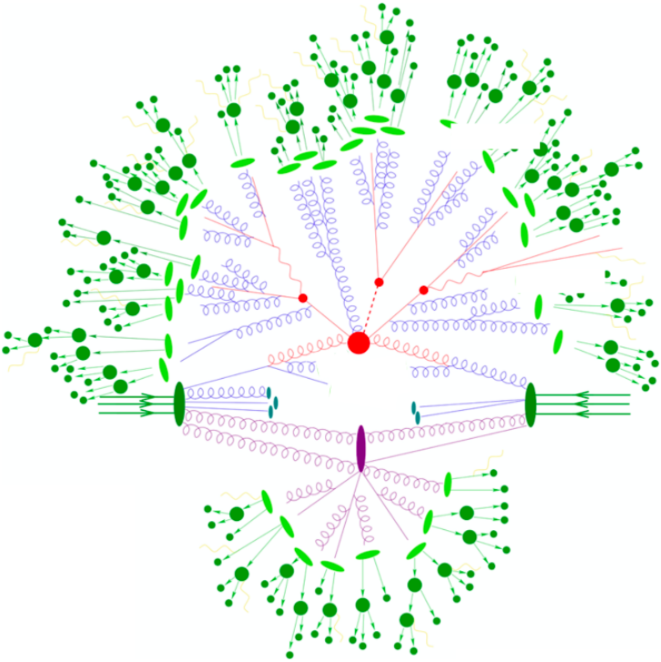


Figure 1.7. Schematic view of a *proton* − *proton* collision. In red there is the elementary parton collision that can be computed at NNNLO. In purple a secondary interaction within the same collision. PDFs are embedded into the selection of the parton. Then the blue lines represent the parton emissions. The hadronization of uncolored is represented in light green. The hadronic shower into the calorimeter detector is represented in dark green.

In a suggestive way it is possible to talk about time scales. To ensure that a parton is bound into a hadron it is necessary an approximate time of $\tau_{had} \approx \gamma R = pR^2$. In this expression γ is the Lorenz factor of the parton, p is its momentum and R is approximately the scale hadron dimension of about 1 fm. For light quarks at high energies we're talking about tens fm). During this time the parton can radiate a gluon being a quark or split in a quark-antiquark pair being a gluon it this formation time (τ_{form}) is much lower than the hadronization time. The top quark is a special case: because of his high mass ($m_t \approx 173\ GeV$) the top quark decays ($t \rightarrow W^+b$) before its hadronization, $\tau_t \approx 0.1\ fm$; as a consequence there are no hadrons composed by a top quark. The formation time can be described by the virtuality of the parton $\tau_{form} \approx p/q^2 \approx p/p_T^2$ for soft and collinear emissions. Therefore to create a new parton the condition is $\tau_{form} << \tau_{had}$ or $p_T >> 1/R$. The one can introduce the separation time or $\tau_{sep} = \tau_{form} p_T R$ representing the time for a produced parton to fall into a different hadron with respect to the native parton.

$$\tau_{form} = \frac{p}{p_T^2} \qquad \tau_{sep} = \tau_{form}(p_T R) \qquad \tau_{had} = \tau_{form}(p_T R)^2 \qquad (1.51)$$

If the three time scales are well separated with $p_T >> 1/R$ a parton shower occurs.

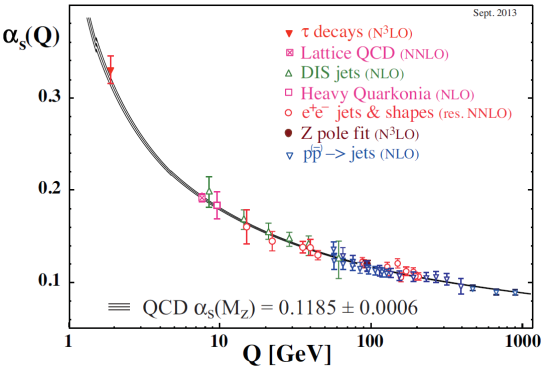


Figure 1.8. Measured values of QCD coupling constant α_s showing the asymptotic freedom.

The parton shower is due mostly to soft and collinear emissions in the perturbative region where $\alpha_s \ll 1$. This is possible because, despite the coupling is small, each emission is weighted with $\frac{\alpha_s}{2\pi} \ln^2 p$. Hard and non collinear emission do not contribute so much and as a consequence a hadron jet is most of all formed by the native parton.

2 | Experimental Overview

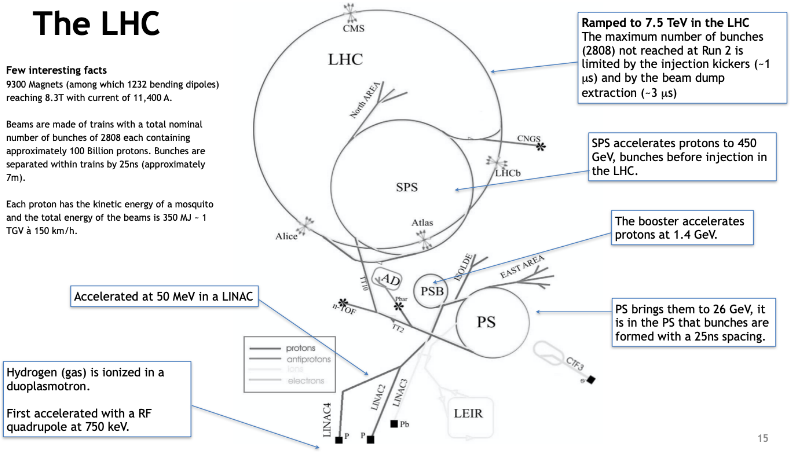
Chapter 2 provides a detailed examination of the ATLAS Detector at the Large Hadron Collider (LHC), an essential tool in the exploration of fundamental particles and their interactions. This chapter is crucial for understanding the experimental framework within which the research described in this dissertation is conducted.

In Section 2.1, we offer an overview of the Large Hadron Collider itself, the world's most powerful particle accelerator, constructed by the European Organization for Nuclear Research (CERN). This section will cover the LHC's design, its operational capabilities, and its role in advancing high-energy physics. We will discuss the LHC's infrastructure, including its 27 km ring and the series of accelerators that precede the final collision stage, elucidating how protons are accelerated to nearly the speed of light before being collided at unprecedented energies.

Furthermore, Section 2.3 will delve into the reconstruction of physics objects from raw detector data. This includes tracking and vertexing techniques, the Particle Flow algorithm, and the identification of electrons, photons, muons, jets, tau leptons, and missing energy. These reconstructed objects are critical for the analysis of collision events and the extraction of meaningful physical insights.

This chapter lays the groundwork for understanding the experimental setup and methodologies employed in the research, providing the necessary context to appreciate the complexity and capabilities of the ATLAS Detector and its pivotal role in probing the fundamental aspects of particle physics.

2.1 The Large Hadron Collider



The LHC

Few interesting facts
9300 Magnets (among which 1232 bending dipoles) reaching 8.3T with current of 11,400 A.

Beams are made of trains with a total nominal number of bunches of 2808 each containing approximately 100 Billion protons. Bunches are separated within trains by 25ns (approximately 7m).

Each proton has the kinetic energy of a mosquito and the total energy of the beams is 350 MJ ~ 1 TGV à 150 km/h.

Ramped to 7.5 TeV in the LHC
The maximum number of bunches (2808) not reached at Run 2 is limited by the injection kickers (~1 μs) and by the beam dump extraction (~3 μs)

SPS accelerates protons to 450 GeV, bunches before injection in the LHC.

The booster accelerates protons at 1.4 GeV.

PS brings them to 26 GeV, it is in the PS that bunches are formed with a 25ns spacing.

Accelerated at 50 MeV in a LINAC

Hydrogen (gas) is ionized in a duoplasmotron.

First accelerated with a RF quadrupole at 750 keV.

Figure 2.1. Schematic view of the path of protons to be injected in the Large Hadron Collider to reach an energy of $E_p = 7.5\ TeV$. The three injection rings (PSB, PS, SPS) are well visible.

The Large Hadron Collider (LHC) at Conseil Européen pour la Recherche Nucléaire (CERN) in Geneva is the world's largest and most powerful particle accelerator. It was built between 1998 and 2008 in the tunnel previously used for the Large Electron-Positron (LEP) collider. The LHC ring has a circumference of 26.7 km and it is located at a depth of between 45 and 175 m. The Collider accelerates protons (and ions, although for this thesis they are not discussed). The protons can be potentially accelerated inside the collider up to $E = 7\ TeV$, equivalent to a centre of mass energy of $\sqrt{s} = 14\ TeV$. When the LHC was built in 2010, the center of mass energy was $\sqrt{s} = 7\ TeV$ and it has increased over the years. Superconducting magnets are used to bend the protons inside the ring. A scheme of the LHC architecture is shown in Figure 2.1. At the LHC, four points of interactions of the beams are placed along the ring. For each, a main experiment is built:

- **ATLAS** (A Toroidal LHC ApparatuS)
- **CMS** (Compact Muon Solenoid)

- **ALICE** (Large Ion Collider Experiment)
- **LHCb** (Large Hadron Collider beauty)

In particular, ATLAS and CMS [35, 36] are designed to detect a wide range of signals both in the Standard Model and beyond. For this reason, they are commonly known as multi-purpose experiments. Up to now (2024), LHC has run twice and since 2022 the third Run is ongoing.

During the first Run (Run1: 2010-2012), the energy in the center of mass was $\sqrt{s} = 7 - 8\ TeV$. During the first long shutdown (LS1: 2012-2015), some of the detector systems were upgraded and the centre of mass energy was raised to $\sqrt{s} = 13\ TeV$. The second run lasted until 2018 (Run2: 2015-2018) and the collected integrated luminosity during the first two runs was $\mathcal{L} = 140 fb^{-1}$. With the end of the second long shutdown (LS2: 2018-2022), which finished on the 4th of July of 2022 (10 years after the Higgs boson discovery), the third run of LHC started (Run3: 2022-2025). During this shutdown, many upgrades were performed on LHC and the individual experiments. The centre of mass energy for Run3 is $\sqrt{s} = 13.6\ TeV$ and the instantaneous luminosity is twice that of Run2. After the third shutdown (scheduled LS3: 2026-2029), the High Luminosity era should start with an integrated luminosity expected to reach $\mathcal{L} = 3ab^{-1}$; this program is called HL-LHC.

At the moment, the procedure of injection in the Large ring works as shown in Figure 2.1. Protons must follow several phases before the collisions. In a first step, Hydrogen gas is ionized and the protons are accumulated to be accelerated up to $E_p = 50\ MeV$ energy by the Linear Accelerator 2 (LINAC2). Then they are injected into three successive circular accelerators. The first ring, the Proton Synchrotron Booster (PSB), brings the proton energy up to $E_p = 1.4\ GeV$; the second, the Proton Synchrotron (PS), up to $E_p = 26\ GeV$; and the latter, the Super Proton Synchrotron (SPS), to $E_p = 450\ GeV$. The protons are organized in bunches of $N_b = 1.1 \times 10^{11}$ protons, temporally spaced by $t = 25\ ns$. At the end, they are injected into the LHC ring where, thanks to 16 Radio-Frequency cavities (RF), the energy ramps to $E_p = 7.5\ TeV$. They work at $T = 4.5\ K$, and a potential of $V = 2MV$. The protons are confined within the ring thanks to 9300 magnets and dipoles reaching a magnetic field of $B = 8.3\ T$. Moreover, 392 quadrupoles are used to focus the beam, which converge at four points on the ring where the detectors are placed.

The synchrotron is the ultimate frontier for the collider physics [37]. The ancestor of this collider was the cyclotron which couldn't keep the orbital radius fixed with increasing energy. As a consequence all the circle must be covered by a bending magnetic field. As a consequence the synchrotron idea is to keep the radius fixed and modify the magnetic field along the trajectory. The momentum of the protons can be defined by:

$$\mathbf{p} = \beta\gamma\mathbf{v} \tag{2.1}$$

At the same time the protons are bended by the Lorentz force [38] due to magnetic dipoles along the ring.

$$\mathbf{F} = q\mathbf{v} \wedge \mathbf{B} \tag{2.2}$$

Imposing a circular orbit the Lorentz force must be equal to the centripetal force

$$\mathbf{F} = m\gamma\frac{v^2}{r^2}\mathbf{r} \tag{2.3}$$

From these qualities the famous relation between the momentum expressed in GeV (p), the magnetic field expressed in T (B) and the radius of curvature expressed in m (ρ) is obtained:

$$p[GeV] = 0.3z\rho[m]B[T] \tag{2.4}$$

In particular, to have a fixed radius for protons $p/0.3B$ must be constant; therefore, B must increase to accelerate protons. At the LHC, there are two vacuum pipes, and the protons to be collided flow into them in opposite directions. As a consequence, pairs of bending dipoles along the ring are placed. A schematic figure of the bending dipoles and quadrupoles is shown in Figure 2.2. The quadrupoles are responsible for the beam focusing.

The characteristic of an accelerating particle is the emission of radiation. In particular, the emission radiation from synchrotrons is called synchrotron radiation, and the energy loss can be estimated as:

$$\mathit{\Delta} E = \frac{E^2}{m^4\rho} \tag{2.5}$$

It is immediately clear that to reduce the energy loss, a huge ring is needed. Moreover, it is convenient to accelerate particles with a higher mass, such as protons. For example, the ratio between the energies

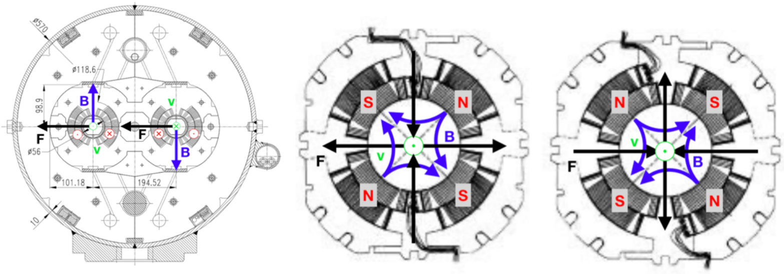


Figure 2.2. (**left**) Transverse section of magnet dipoles. The magnetic field in Blue is generated by the superconductor currents shown in Red. The proton directions flows inside the beam pipes in Green in opposite directions. The Lorentz force is shown in Black. (**right**) With the same nomenclature a quadrupole is shown. The left quadrupole focuses along the Y axis while the righ one focuses on the X axis. Subsequential quadrupoles are used to focus the beam.

lost by protons and electrons is $\approx (m_e/m_p)^4 \approx 10^{-13}$. In fact, protons lose much less energy than electrons and can be accelerated more easily. This is one of the main advantages of using hadron colliders to probe the Standard Model at the highest energies possible. One of the main disadvantages is that protons are composite particles and the events are never pure; mostly, one cannot be sure that the transverse momentum p_T in the collision is null.

Another important feature of colliders is the instantaneous luminosity ($\mathcal{L}[fb^{-1}s^{-1}]$) which enters into the cross-section formula. In particular, it is possible to express the rate ($R[s^{-1}]$) of a certain event in terms of the luminosity, containing the collider's properties, and the cross section ($\sigma[fb]$), containing the fundamental physics:

$$R = \frac{dN_{ev}}{dt} = \sigma\mathcal{L} \tag{2.6}$$

The luminosity in this expression is also called instantaneous luminosity to distinguish it from its integration over time, known as integrated luminosity. Luminosity is a parameter that can be expressed as a function of the beam parameters. In particular, assuming that the protons within the beam are collected in bunches with a Gaussian distribution, the instantaneous luminosity can be expressed as follows:

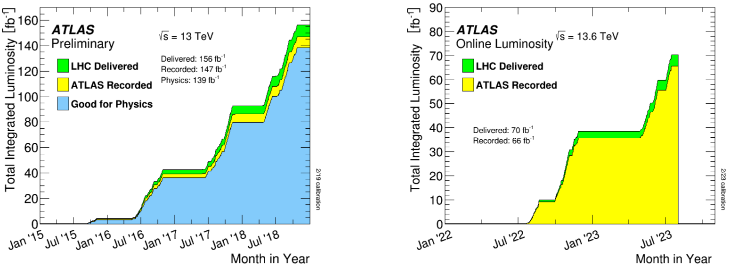
$$\mathcal{L} = \frac{N_p^2 N_b f}{4\pi\sigma_x\sigma_y}F = \frac{N_p^2 N_b f\gamma}{4\pi\epsilon_N\beta*}F \tag{2.7}$$

In this expression the nominal values (for Run2) are:

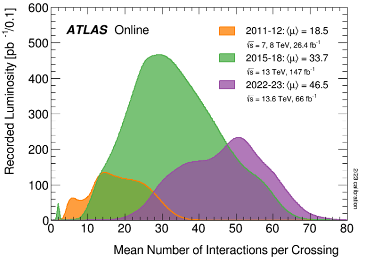
- $N_p = 1.15 \times 10^{11}$ is the number of proton in each bunch
- $N_b = 2808$ is the number of bunches
- $f = 11.245\ kHz$ is the revolution frequency
- $\sigma_{x/y} = 15.9\ \mu m$ is the transverse spatial dispersion of a beam
- $\sigma_z = 8.4\ cm$ is the longitudinal spatial dispersion of a beam
- $\beta^* = 0.55\ m$ is the focusing of the beam
- $\epsilon_N = 3.75\ \mu m\ rad$ is the normalised emittance.
- $\theta_c = 295\ \mu rad$ is the convergence angle between two beams
- $F = (1 + (\frac{\theta_C \sigma_z}{2\sigma_x})^2)^{-1/2}$ takes into account the non collinearity of the two beams
- $\mathcal{L} = 10^{34}\ fb^{-1}\ s^{-1}$

The second expression takes into account the Liouville theorem and the conservation of phase space along the ring, considering the emittance imposed at injection and defined by a distribution on the plane p_x/p_z vs x. The integrated luminosity for ATLAS over the years is shown in Figure 2.3.

Another important effect that could occur at LHC is the Pile-Up (PU), usually indicated as the average number of interactions per bunch crossing, $\langle\mu\rangle$. This phenomenon is due to multiple simultaneous interactions that can be detected by the experiment. It is possible to distinguish between in-time and out-of-time pile-ups according to their origin. The Pile-Up is said to be in-time if the events are initiated by two interactions in the same bunch; otherwise, if the events start from different bunches, they are called out-of-time pile-up. Assuming a Poissonian distribution, it is possible to give an estimation of the number of pile-ups as $\mu = \mathcal{L}\sigma_{inel}/N_b f$. This value differs depending on the beam setting, which varies in each Run of the LHC. In Run2, $\langle\mu\rangle \approx 33.7$. In Run3, it has increased to $\langle\mu\rangle \approx 33.7$ with HL-LHC.



(a) Run2 Luminosity (b) Run3 Luminosity



(c) Pile-Up Conditions

Figure 2.3. Integrated Luminosity over the years for the ATLAS experiment during Run2 (top left) and Run3 (top right). Average Pile-Up conditions during Run1, Run2 and Run3 (bottom)

2.2 The ATLAS Detector

A Toroidal LHC ApparatuS (ATLAS) is one of the four experiments along the 27 km ring of the LHC [35]. It is the largest multi-purpose particle detector ever built, measuring 25 m in diameter, 44 m in depth, and weighing 7000 t. It is followed by the Compact Muon Solenoid (CMS) detector, which is also placed along the same ring. Both the ATLAS and CMS detectors are shown in Figure 2.4.

The ATLAS detector has cylindrical symmetry and shape around the interaction point of the two beams. To achieve the highest possible coverage around the interaction point, the detector is composed as follows: There is an Inner Detector (ID), which is an inner tracking system that allows for the reconstruction of interaction vertices and the tracks of charged particles. The ID is contained within a superconducting solenoid generating a 2 T axial magnetic field, capable of bending charged particles. Outside of this tracker system, there is an Electromagnetic Calorimeter (ECAL) followed by a Hadronic Calorimeter (HCAL), which together allow for the measurement of energy deposited by electromagnetic and hadronic particles that reach it. A characteristic part of ATLAS is the final muon spectrometer (MS) composed of toroidal magnets providing a 1 T magnetic field [39] around the experiment. A schematic view of the ATLAS experiment is shown in Figure 2.5.

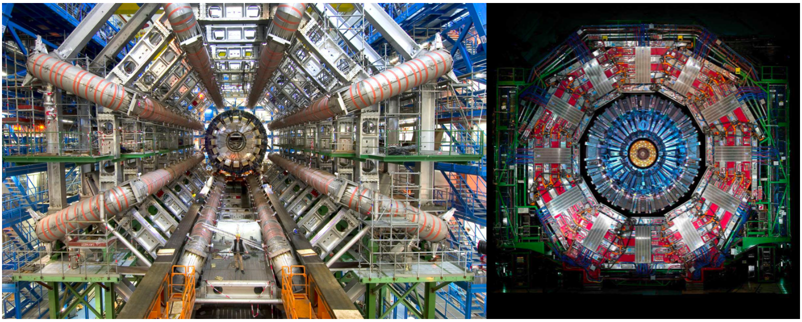


Figure 2.4. (**left**) ATLAS detector. (**right**) CMS detector.

This complex experiment has been designed to further study and investigate the Standard Model (and beyond). It is necessary for the detectors to have as fast a response as possible in order to acquire the greatest possible set of data. Moreover, each part of the detector must have enough efficiency and resolution to be able to distinguish interesting events among the collisions. To provide a brief description of the particle interactions, one can observe different behaviors for different particles. Charged particles are bent by the compact solenoid in the ID, and their momentum can be measured from their geometrical sign in the ID. The ECAL is able to measure the energy released by photons and electrons, producing electromagnetic showers within it. The HCAL, on the other hand, is able to measure the energy released by both charged and neutral strongly interacting particles such as protons, neutrons, and charged pions, generating hadronic showers within it. Muons are then able to travel many meters within the detector, leaving a sign in the ID, but are not stopped by the Calorimeters. For this reason, the MS is needed to identify them. The only particle in the Standard Model that cannot be directly measured

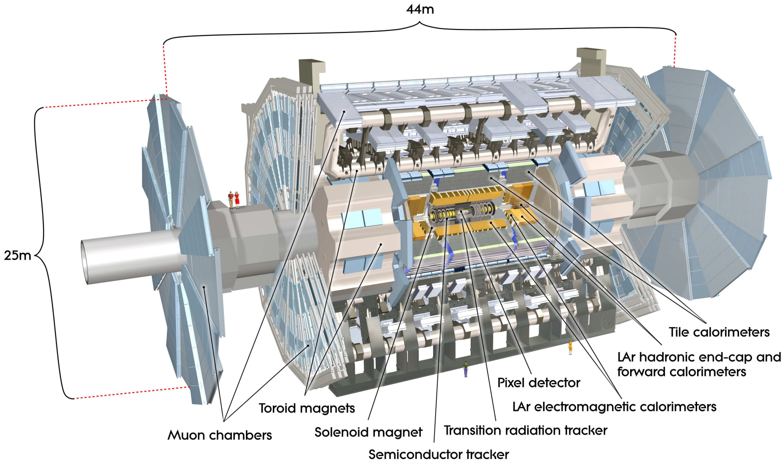


Figure 2.5. ATLAS detector and its composition.

with this detector is the neutrino, which almost doesn't interact at all with the detector. For this reason, it is possible to measure it indirectly through the missing energy in the event.

2.2.1 Coordinate System

Given the symmetry of the detector, it is useful to introduce an appropriate coordinate system. In particular, it is possible to describe the products of proton-proton collisions in terms of relativistic invariant quantities. Therefore, since the collisions are head-on, everything along the transverse plane $(x - y)$ is invariant. The most important parameters include the azimuthal angle $\phi \in (-\pi, \pi]$. The transverse momentum p_T before the collision is known to nearly vanish, and due to its conservation, it is natural to infer that the sum of the collision products' transverse momenta must also vanish.

In general, in the LAB frame, it is possible to use polar coordinates with the polar angle $\theta \in [0, \pi]$. This angle is formed with respect to the z-axis and is not Lorentz invariant. Consequently, another parameter to define is the longitudinal direction. In particular, it is possible to introduce the concept of rapidity y and its boost transformations:

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) \Rightarrow y' = y + \frac{1}{2} \ln \left(\frac{1 + \beta}{1 - \beta} \right) \qquad (2.8)$$

It becomes immediately clear that differences in rapidity are Lorentz invariant, even though rapidity itself is not.

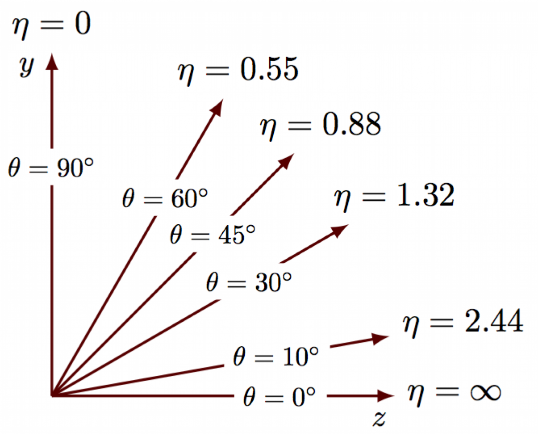


Figure 2.6. Pseudo-rapidity (η) values and corresponding polar angle (θ) values.

Evaluating the rapidity for $p \gg m$, one obtains the pseudo-rapidity η, which is itself Lorentz invariant and can be expressed as a function of the polar angle $\cos(\theta) = p_z/p$:

$$y \xrightarrow{\beta \to 1} \eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right) \tag{2.9}$$

This correspondence is shown in Figure 2.6. Another parameter, particularly important for jets, is the radius describing an angular distance that is relativistic invariant:

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} \tag{2.10}$$

2.2.2 Magnetic systems

The ATLAS experiment is equipped with four magnetic fields generated by superconductive coils. The central solenoid provides the magnetic field for the Inner Detector (ID), and three solenoids are used for the Muon Spectrometer (MS); one composes the barrel part of the spectrometer, while the last two are placed at the end-caps. A schematic of the magnets is shown in Figure 2.7.

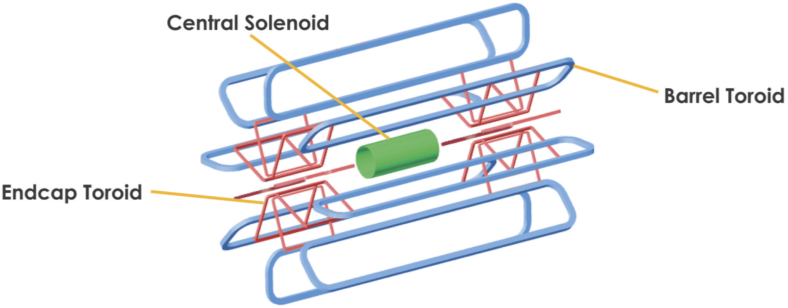


Figure 2.7. Design of the 4 magnets of the ATLAS experiment.

The central solenoid is designed with a diameter of 2 m to optimize the amount of material in front of the calorimeters, providing an axial magnetic field of 2 T. With the toroidal magnets, the length and the diameter reach about 20 m and 24 m respectively. The magnetic field varies with space from 0.1 T to 3 T.

The magnetic fields are necessary to bend charged particles and therefore to measure their momentum. As described in the previous section, the momentum depends on the radius of curvature ρ and the magnetic field:

$$p_T[GeV] = 0.3zB[T]\rho[m] \tag{2.11}$$

Thus, the transverse momentum can be measured with the measurement of the sagitta s, namely the distance between the straight direction and the bent direction. From geometrical considerations, shown in Figure 2.8, if $s \ll \rho$ then:

$$p_T[GeV] = \frac{0.3B[T]l^2[m^2]}{8s[m]} \tag{2.12}$$

2.2.3 The Inner Detector

The ATLAS Inner Detector (ID), shown in Figure 2.9, is a tracking sub-detector surrounding the interaction point. It serves as the ATLAS tracking system for Run2 and Run3. It has a diameter of 2.1 m and a length of 6.2 m, and is fully immersed in the solenoidal magnetic field. Its purpose is to measure the bending of charged particles' tracks in the magnetic field and to reconstruct the primary vertices

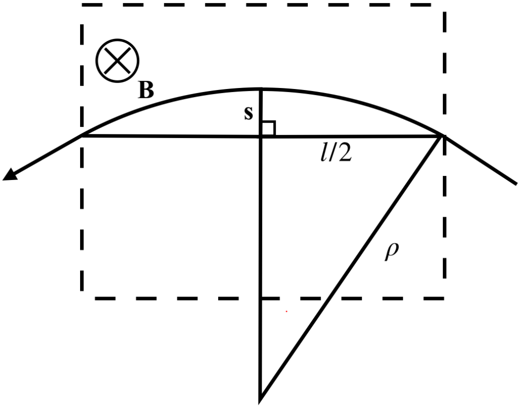


Figure 2.8. Scheme of the momentum measurement with the sagitta s of a bending particle within a magnetic field.

as well as secondary vertices from decay products. From the bent trajectory, it is then possible to reconstruct the sign of the charge of the particle crossing the ID, its relative momentum, and the extrapolated production point.

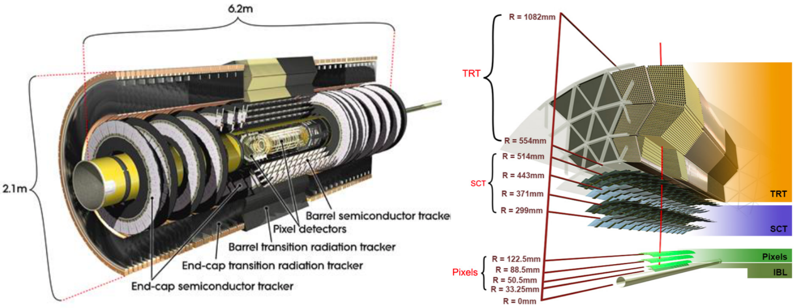


Figure 2.9. Scheme of the Inner Detector and transversal view of the four sub-systems with their spacing from the center.

The ID consists of four sub-systems, covering a total range of $|\eta| < 2.5$:

- Insertable B-Layer (IBL)

- Pixel Detector (PD)
- Semiconductor Tracker (SCT)
- Transition Radiation Tracker (TRT)

The Insertable B-Layer [40], the innermost pixel layer, was built around the beryllium beam pipe and inserted into the Pixel Detector during the first upgrade after Run1. It surrounds the beam pipe at 33 mm and enables more precise measurements near the Interaction Point (IP). It is particularly effective in distinguishing between displaced secondary vertices, i.e., decay products of B-Hadrons, and primary vertices from p-p interactions.

The Pixel Detector offers higher granularity compared to the outer detectors. It consists of 1744 silicon pixel modules arranged in three concentric barrel layers and two end-caps, each with three disks. The pixels are fragmented along η, ϕ, and z axes, with dimensions of $50 \times 400\ \mu m^2$. The last cylindrical layer is positioned 122.5 mm from the center.

The Semiconductor Tracker (SCT) spans radial distances from 299 mm to 514 mm. The detector comprises 4088 modules of silicon-strip detectors arranged in four concentric barrels and two end-caps, each with nine disks. The strips in the barrels are parallel to the solenoid field and beam axis, and have a constant pitch of 80 μm.

The Transition Radiation Tracker (TRT) is the outermost system of the ID, covering radial distances from 554 mm to 1062 mm. Unlike the inner parts, it is a gas detector. The detector consists of 298,304 proportional drift Xenon tubes, each 4 mm in diameter. The straws in the barrel region are arranged in three cylindrical layers. It is highly sensitive to distinguishing electrons from pions. Its operating principle involves transition radiation in dielectrics, which enables particle identification (PID) because the number of photons produced in the transition radiation is proportional to the mass of the crossing particle.

The ID is ultimately designed to detect particles with $p_T > 0.4\ GeV$, and it is possible to estimate the overall relative uncertainty on the momentum as follows:

$$\frac{\sigma_{p_T}}{p_T} = 0.05\%\ p_T[GeV] \oplus 1\% \tag{2.13}$$

2.2.4 The Calorimeters

Unlike the trackers, calorimeters measure the energy of a particle by absorbing it. A key feature of calorimeters is that their resolution improves with the energy of the particles. ATLAS is equipped with two types of calorimeters: the Electromagnetic Calorimeter (ECAL) and the Hadronic Calorimeter (HCAL). The layout of the ATLAS calorimeters is shown in Figure 2.10, and they are positioned around the ATLAS Inner Detector (ID) outside the solenoid magnet.

Unlike the CMS calorimeters, the ATLAS calorimeter is a sampling calorimeter, which means it consists of alternating layers of active medium and passive absorber. The overall energy resolutions for ECAL and HCAL are respectively:

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 0.7\% \qquad \frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\% \tag{2.14}$$

The critical parameters defining ECAL and HCAL are the radiation length X_0 and the interaction length λ_I, respectively.

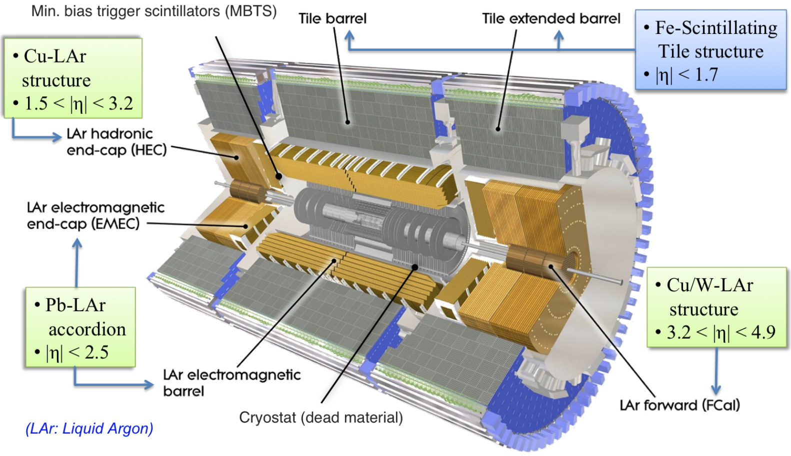


Figure 2.10. Schematic view of both Electromagnetic Calorimeter and Hadronic Calorimeter.

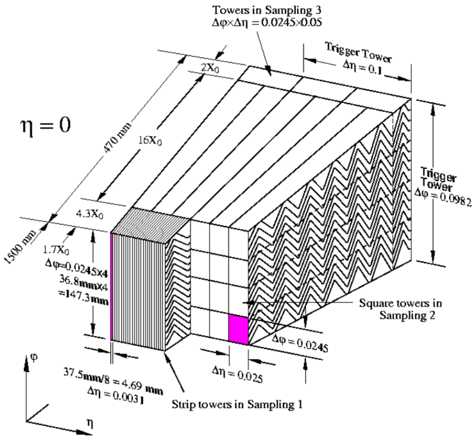


Figure 2.11. ECAL accordion geometry with LAr active material and Pb absorber.

The Electromagnetic Calorimeter

The electromagnetic calorimeter operates based on two physical processes: pair production and Bremsstrahlung. In physics, it is well known that an electron or a positron in a medium can radiate a photon (Bremsstrahlung). The characteristic length for this process is the radiation length X_0. Simultaneously, a photon with sufficient energy in a medium can split into an $e^+ - e^-$ pair. The characteristic length for this process is $\frac{9X_0}{7}$. Radiation emission occurs predominantly above the critical energy for electrons in the medium, which is defined as:

$$E_c^{solid} = \frac{610\ MeV}{Z + 1.24} \qquad E_c^{gas} = \frac{710\ MeV}{Z + 0.92} \tag{2.15}$$

Below the critical energy, electrons cease motion primarily through ionization. The sequence of pair production and Bremsstrahlung in a medium suggests that the calorimeter can be defined in terms of the number of radiation lengths it spans. Furthermore, X_0 depends on the chosen medium and follows an empirical expression:

$$X_0 = \frac{1433[g\ cm^{-1}]A}{Z(Z+1)\ln\frac{183}{Z^{1/3}}} \tag{2.16}$$

The ECAL is a sampling calorimeter consisting of alternating layers of Liquid Argon (LAr, $X_0 = 140\ mm$) as the active material and Lead (Pb, $X_0 = 5.6\ mm$) as the absorber. A simple view of this geometry is shown in Figure 2.11. The calorimeter is divided into a barrel section with $|\eta| < 1.475$ and two end-cap components with $1.375 < |\eta| < 3.2$. The total thickness of the ECAL exceeds 22 X_0 in the barrel and 24 X_0 in the end-caps.

Hadronic Calorimeter

A hadronic shower is initiated by a strongly interacting particle, such as charged pions. Neutral pions decay, with a lifetime of $\tau = 8 \times 10^{-17}\ s$, into two photons, which then produce two electromagnetic showers. Describing a hadronic shower is more complicated than describing an electromagnetic shower because other particles can be produced in nuclear reactions. The products of the hadronic shower are mostly pions, hence there could be an electromagnetic component. The energy deposited within a hadronic shower can be divided into three categories: ionizing particles (approximately 60%), neutrons (approximately 10%), and invisible energy (approximately 30%). The most important parameter of a hadronic calorimeter is the interaction length λ_I, which is the average distance a hadron must travel inside a medium before a nuclear interaction occurs. The ATLAS HCAL is composed of three parts with a total thickness at $\eta = 0$ of 11 λ_I. With this configuration, it is possible to achieve a high resolution for measuring the missing transverse energy in an event (E_T^{miss}), which is used to infer the presence of neutrinos.

The Tile Calorimeter (TC) is placed outside the ECAL. Its barrel covers the region with $|\eta| < 1.0$, with two extensions in the region of $0.8 < |\eta| < 1.7$. It is a sampling calorimeter using steel (Fe, $\lambda_I = 168\ mm$) as the absorber and scintillating tiles as the active material. Radially, the Tile Calorimeter extends from an inner radius of 2.28 m to an outer radius of 4.25 m. It is segmented in depth into three layers, approximately 1.5, 4.1, and 1.8 interaction lengths λ_I thick for the barrel, and 1.5, 2.6, and 3.3 λ_I for the extended barrel. The total thickness at $\eta = 0$ is 9.7 λ_I.

The LAr hadronic end-cap calorimeter (HEC) consists of two independent wheels per end-cap, located directly behind the end-cap electromagnetic calorimeter. To reduce the material density drop at the transition between the end-cap and the forward calorimeter, which starts at $\eta = 3.1$, the HEC extends out to $|\eta| < 3.2$, thereby overlapping with the forward calorimeter. Similarly, the HEC's η range also slightly overlaps that of the Tile Calorimeter by extending to $|\eta| > 1.5$. Each wheel is divided into two segments in depth, totaling four layers per end-cap.

The LAr forward calorimeter (FCal) is integrated into the end-cap cryostats, providing clear benefits in terms of uniformity of calorimetric coverage. The FCal is approximately 10 interaction lengths deep and consists of three modules in each end-cap: the first, made of copper, is optimized for electromagnetic measurements, while the other two, made of tungsten, predominantly measure the energy of hadronic interactions. Ultimately, it provides coverage from $3.1 < |\eta| < 4.9$.

2.2.5 The Muon Spectrometers

The muon is the most penetrating particle (excluding the neutrino) due to its mass of $m_\mu = 105.66\ MeV$. Its main characteristic is that it is a lepton and does not interact strongly. Additionally, its relatively small mass allows for a long decay time, expressed as:

$$\tau^{-1}\mu = \frac{G_F^2 m\mu^5}{192\pi^3} F\left(\frac{m_e^2}{m_\mu^2}\right) H\left(\frac{m_e^2}{m_\mu^2}, \alpha(m_\mu)\right) = (2.20 \times 10^{-6}\ s)^{-1} \tag{2.17}$$

Given that its critical energy is much higher than that of the electron:

$$E_c^{solid} = \frac{5700\ GeV}{(Z+1.47)^{0.838}} \qquad E_c^{gas} = \frac{7980\ GeV}{(Z+2.03)^{0.879}} \tag{2.18}$$

at the energies produced at LHC, the muon does not radiate Bremsstrahlung photons but primarily behaves as a Minimum Ionizing Particle (MIP). This means that the muon traverses the detectors at the minimum of the Bethe-Bloch function, losing $\frac{1}{\rho}\frac{dE}{dx} \approx 1 - 2\ \ MeV\ cm^2/g$, as depicted in Figure 2.12.

Consequently, a specialized apparatus is required to detect muons and all particles that penetrate the calorimeters. This is the purpose of the Muon Spectrometer (MS) in ATLAS. Within the range $|\eta| < 1.4$, the magnetic bending by the large barrel toroid is employed. For $1.6 < |\eta| < 2.7$, muon tracks are bent by two smaller end-cap magnets inserted at both ends of the barrel toroid. The transition region, $1.4 < |\eta| < 1.6$, utilizes magnetic fields provided by both the barrel and end-cap magnets. This configuration ensures that the field is mostly orthogonal to the muon tracks, minimizing errors due to multiple scattering.

In the barrel region, tracks are measured in chambers arranged in three cylindrical layers around the beam axis. In the transition and

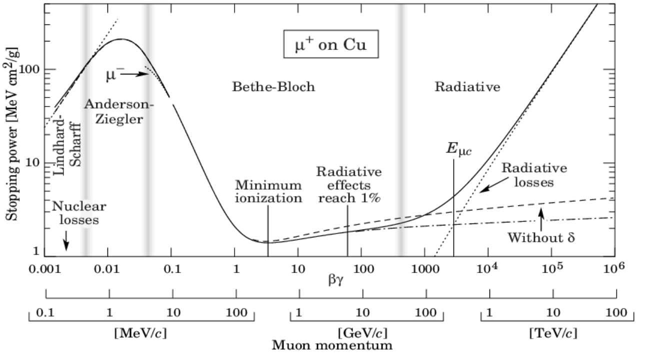


Figure 2.12. Bethe-Bloch functions for muons on Cu.

end-cap regions, the chambers are installed in planes perpendicular to the beam, also in three layers. A schematic of the system is shown in Figure 2.13.

Over most of the η-range, precision measurements are provided by Monitored Drift Tubes (MDT). These are aluminum tubes with a 3 cm diameter and lengths ranging from 1 to 6 m, filled with $Ar-CO_2$ at a 93 : 7 ratio. At the center of each tube, there is a tungsten (W) anode wire. Since the drift time of the electrons inside the tube is 700 ns, significantly greater than the 25 ns time separation between two bunches, the MDTs are not suitable as a first-level trigger system. Cathode Strip Chambers (CSC) are used in the innermost tracking layer due to their higher rate capability and time resolution. CSC covers the region of $2.0 < |\eta| < 2.7$.

The trigger system covers the pseudo-rapidity range $|\eta| < 2.4$. Resistive Plate Chambers (RPC) are used in the barrel and Thin Gap Chambers (TGC) in the end-cap regions. The trigger chambers for the muon spectrometer serve threefold purposes: they provide bunch-crossing identification, define well-established pT thresholds, and measure the muon coordinate in the direction orthogonal to that determined by the precision-tracking chambers. The resolution on the momentum depends on the muon's momentum and η. Overall, the relative uncertainty for muons at $p_T = 100\ GeV$ is $\sigma_{p_T}/p_T = 3\%$. This grows to $\sigma_{p_T}/p_T = 3\%$ for $p_T = 1\ TeV$.

During the LS2, the small wheel was replaced with a New Small Wheel (NSW) [41]. The NSW is one of the main upgrades in Run3 but is beyond the scope of this dissertation.

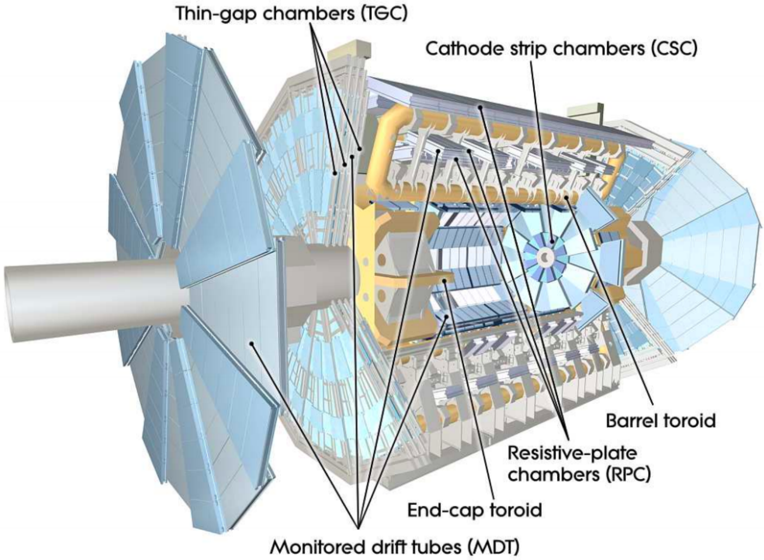


Figure 2.13. Muon spectrometer and magnetic toroid system.

2.2.6 Trigger System

During the data taking of Run2, two consecutive bunches are separated by 25 ns, providing a collision frequency of 40 MHz. Acquiring the entirety of the events would be very difficult, unnecessary, and impossible. For this reason, a multi-tier trigger system is necessary to select the interesting events for reconstruction. The trigger must make decisions quickly and use simple criteria to determine which events to store on tapes.

The trigger system from the second run is divided into two levels:

- the Level 1 (L1) Trigger, which reduces the number of events to 100 kHz with a latency time of 2.5 μs,
- the High Level Trigger (HLT), comprising the previous Level 2 (L2) and Level 3 (L3) triggers, which further reduces the number of events to $1k\ Hz$.

In this dissertation, the dataset analyzed mainly comes from the entire Run2 and from partial Run3 data taking.

2.3 Physics Objects Reconstruction

To convert the data collected by the sub-detectors of the ATLAS experiment into physics objects that can be used by the analysis teams, dedicated software named ATHENA [42] has been developed. All algorithms are centrally developed within the ATHENA framework. Each physics object is then calibrated to account for single detector effects or systematic errors during data acquisition.

The reconstruction process starts from lower-level reconstructed objects, such as hits in the Tracker or cells in the Calorimeter, to produce higher-level objects like tracks or calorimeter clusters. From these variables, it is then possible to produce physics objects like reconstructed particles. Given the complexity of this process, the algorithms used may encounter identification or isolation problems. To discriminate a reconstructed object from a similarly behaving one (background), it is also possible to define various working points to achieve a good compromise between signal efficiency and background misidentification.

2.3.1 Tracking and Vertexing

Charged particles are reconstructed from the energy deposits in the ATLAS tracking systems, the ID and the MS. This section focuses on the reconstruction in the ID [43]. During Run 2 data taking, hundreds of charged particles were produced at each BC, and their reconstruction is crucial for the physics program since many physics objects rely on the tracking system: vertexing, lepton identification, flavor tagging, etc. ATLAS tracking uses as input, hit clusters in the Pixel and SCT detectors, and drift circles in the TRT.

The two algorithms used to reconstruct a track are as follows:

- **Inside-out algorithm**: This algorithm is performed in two main steps: Seeding and Finding. Segments are created by connecting triplets of space-points, searching for patterns of hits that could belong to the same particle. This way, the algorithm, assuming some knowledge of the expected curvature of the tracks in the magnetic field and compatibility of the hit position with the track path, can efficiently reconstruct charged particles from the innermost part of the tracker outward. After identifying the track segments, these are fitted together into complete track trajectories. Track fitting uses a Combinatorial Kalman Filter

(CKF) mathematical method to estimate the most likely path of the charged particles, creating track candidates. These tracks take into account effects arising from Multiple Scattering (MS) or energy loss in the material. Once the track candidate is reconstructed, the quality of the track can be evaluated via the goodness of the fit or the number of hits associated with the track. If ambiguities arise with overlapping tracks or shared hits, an ambiguity solver is applied, ranking track candidates based on their quality. Finally, a refinement and extension to TRT are applied.

- **Outside-in algorithm**: To increase the acceptance of particles created at great distances with fewer hits in the ID, such as products from long-lived particles like heavy hadrons, photon conversions, and hadronic interactions, a complementary outside-in algorithm is used. This reconstruction starts from the TRT measurements, adding sequentially the ID hits.

Some acceptance cuts are applied to reconstructed tracks to be used in physics analyses. The tracks are required to be within the ID acceptance ($|\eta| < 2.5$), have a $p_T > 0.5\ GeV$, and pass additional quality criteria such as the number of silicon detector hits and the association with the primary vertex.

In the process of reconstructing vertices from many tracks, the focus is on identifying the primary vertex (PV) — essentially, the point where the most significant proton-proton (pp) collision within each bunch crossing occurs. This identification is accomplished through an iterative process of vertex finding and fitting. The most important quantities entering this process are the two impact parameters (IP): the longitudinal IP d_0 and the transverse IP z_0. The procedure begins with an initial guess of a vertex, known as a vertex seed. Tracks are evaluated to determine if they are likely to have originated from this proposed vertex location. This assessment relies heavily on how well the impact parameters of the measured tracks are compatible with the vertex's position. A track that matches well is considered compatible and is added to the vertex, with its influence on the vertex's proposed location being adjusted accordingly — essentially, more compatible tracks have a greater impact on determining the vertex's position. Once a track is added to a vertex, all the tracks associated with that vertex are analyzed together under the assumption they all originated from the same point. This collective assessment allows for a more precise determination of the vertex's location. However, not all tracks will be compatible with the initially proposed vertex. Those that don't fit are used to start the process again, serving as seeds for the identification of additional vertices. This cycle of evaluation, fitting, and refitting continues until no more vertices can be identified from the available tracks. For a collection of tracks to be considered a valid primary vertex candidate, it must consist of at least two tracks, each with $p_T > 0.5 GeV$. Among the identified valid candidates, the primary vertex is usually distinguished as the one with the largest sum of the squared transverse momenta of its associated tracks, but other custom methods can be implemented depending on the interested physics case. This criterion helps ensure that the selected primary vertex corresponds to the hardest collision event within the bunch crossing, given that higher momentum tracks are more likely to originate from a primary collision event.

2.3.2 Particle Flow algorithm in ATLAS

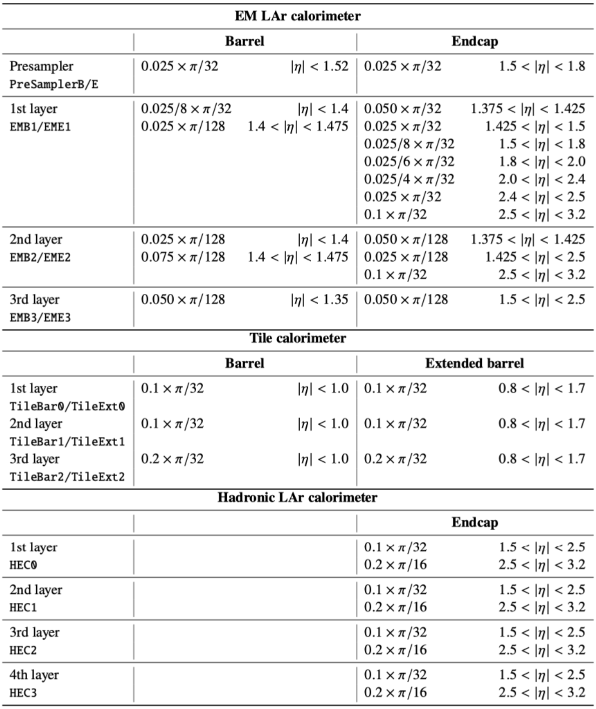
The Particle Flow algorithm [44, 45, 46, 47] is implemented in ATLAS and is designed to optimally reconstruct the energy of neutral particles [46]. The algorithm achieves this by removing the calorimeter energy deposits attributed to charged hadrons, utilizing measurements of their momenta from the Inner Detector. Ultimately, Particle Flow objects consist of the remaining calorimeter energy and tracks that are matched to the hard interaction.

The ATLAS design specifies a calorimeter energy resolution for single charged pions at the center of the detector and an inverse transverse momentum resolution for the tracker, respectively, as follows:

$$\frac{\sigma(E)}{E} = \frac{50\%}{E} \oplus 3.4\% \oplus \frac{1\%}{\sqrt{E}} \quad \sigma(1/p_T) \cdot p_T = 0.036\% \cdot p_T \oplus 1.3\% \tag{2.19}$$

where energies and momenta are in GeV. Consequently, ATLAS benefits from integrating tracking and calorimetric information into a single hadronic reconstruction step. For low-energy charged particles, the momentum resolution of the tracker is significantly better than the energy resolution of the calorimeter. Moreover, the detector's acceptance is extended to softer particles, as tracks for charged particles with a minimum transverse momentum of $p_T > 400\ MeV$ often do not surpass the noise thresholds required for Particle Flow topo-clusters, which will be explained later. The angular resolution of a single charged particle, reconstructed using the tracker, is much better than that of the calorimeter. Low-p_T charged particles originating within a hadronic jet are swept out of the jet cone by the magnetic field by the time they reach the calorimeter. Hence, using the track's azimuthal coordinate, these particles are clustered into the jet.

At high energies, however, the calorimeter resolution surpasses that of the tracker, and a combination of information from both is used for event reconstruction. In the forward direction, outside of the tracker's acceptance, only calorimeter information is utilized.



EM LAr calorimeter				
	Barrel		**Endcap**	
Presampler `PreSamplerB/E`	$0.025 \times \pi/32$	$\lvert\eta\rvert < 1.52$	$0.025 \times \pi/32$	$1.5 < \lvert\eta\rvert < 1.8$
1st layer `EMB1/EME1`	$0.025/8 \times \pi/32$	$\lvert\eta\rvert < 1.4$	$0.050 \times \pi/32$	$1.375 < \lvert\eta\rvert < 1.425$
	$0.025 \times \pi/128$	$1.4 < \lvert\eta\rvert < 1.475$	$0.025 \times \pi/32$	$1.425 < \lvert\eta\rvert < 1.5$
			$0.025/8 \times \pi/32$	$1.5 < \lvert\eta\rvert < 1.8$
			$0.025/6 \times \pi/32$	$1.8 < \lvert\eta\rvert < 2.0$
			$0.025/4 \times \pi/32$	$2.0 < \lvert\eta\rvert < 2.4$
			$0.025 \times \pi/32$	$2.4 < \lvert\eta\rvert < 2.5$
			$0.1 \times \pi/32$	$2.5 < \lvert\eta\rvert < 3.2$
2nd layer `EMB2/EME2`	$0.025 \times \pi/128$	$\lvert\eta\rvert < 1.4$	$0.050 \times \pi/128$	$1.375 < \lvert\eta\rvert < 1.425$
	$0.075 \times \pi/128$	$1.4 < \lvert\eta\rvert < 1.475$	$0.025 \times \pi/128$	$1.425 < \lvert\eta\rvert < 2.5$
			$0.1 \times \pi/32$	$2.5 < \lvert\eta\rvert < 3.2$
3rd layer `EMB3/EME3`	$0.050 \times \pi/128$	$\lvert\eta\rvert < 1.35$	$0.050 \times \pi/128$	$1.5 < \lvert\eta\rvert < 2.5$
Tile calorimeter				
	Barrel		**Extended barrel**	
1st layer `TileBar0/TileExt0`	$0.1 \times \pi/32$	$\lvert\eta\rvert < 1.0$	$0.1 \times \pi/32$	$0.8 < \lvert\eta\rvert < 1.7$
2nd layer `TileBar1/TileExt1`	$0.1 \times \pi/32$	$\lvert\eta\rvert < 1.0$	$0.1 \times \pi/32$	$0.8 < \lvert\eta\rvert < 1.7$
3rd layer `TileBar2/TileExt2`	$0.2 \times \pi/32$	$\lvert\eta\rvert < 1.0$	$0.2 \times \pi/32$	$0.8 < \lvert\eta\rvert < 1.7$
Hadronic LAr calorimeter				
			Endcap	
1st layer `HEC0`			$0.1 \times \pi/32$	$1.5 < \lvert\eta\rvert < 2.5$
			$0.2 \times \pi/16$	$2.5 < \lvert\eta\rvert < 3.2$
2nd layer `HEC1`			$0.1 \times \pi/32$	$1.5 < \lvert\eta\rvert < 2.5$
			$0.2 \times \pi/16$	$2.5 < \lvert\eta\rvert < 3.2$
3rd layer `HEC2`			$0.1 \times \pi/32$	$1.5 < \lvert\eta\rvert < 2.5$
			$0.2 \times \pi/16$	$2.5 < \lvert\eta\rvert < 3.2$
4th layer `HEC3`			$0.1 \times \pi/32$	$1.5 < \lvert\eta\rvert < 2.5$
			$0.2 \times \pi/16$	$2.5 < \lvert\eta\rvert < 3.2$

Figure 2.14. The granularity in $\Delta\eta \times \Delta\phi$ of all the different ATLAS calorimeter layers relevant to the tracking coverage of the inner detector.

As outlined in the previous section, the ID is surrounded by a 2 T magnetic field, providing track reconstruction up to $|\eta| = 2.5$. In contrast, the calorimeter provides coverage up to $|\eta| = 4.9$. Each calorimeter is segmented in the longitudinal direction into multiple layers. In the region $|\eta| < 1.8$, a pre-sampler detector is used to correct for the energy lost by electrons and photons upstream of the calorimeter. The detailed granularity, expressed as $\Delta\eta \times \Delta\phi$, of all the layers is shown in Figure 2.14.

The segmentation of the calorimeter allows for a three-dimensional reconstruction of particle showers, implemented through the topological clustering algorithm. Topo-cluster (TC) cells are grouped together based on their energy deposit and their topological position. TC cells are seeded by cells whose energy measurements E_c exceed four times the noise standard deviation σ_c: $E_c > 4\sigma_c$. TCs are then formed by adding adjacent cells whose energy exceeds two times the noise standard deviation: $E_c > 2\sigma_c$. TCs are completed by incorporating all adjacent cells to these last mentioned.

The TC algorithm is not primarily used for particle identification but rather for the separation of shower types—i.e., electromagnetic and hadronic—and for noise suppression. Moreover, the noise thresholds vary across the calorimeter layers.

In the end, the PFlow algorithm corresponds to an energy subtraction method, removing the overlap between the momentum and the calorimeter energy deposit. It can be summarized as shown in Figure 2.15.

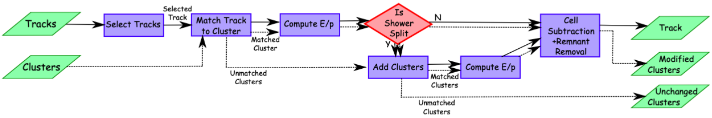


Figure 2.15. A flow chart of how the particle flow algorithm proceeds, starting with track selection and continuing until the energy associated with the selected tracks has been removed from the calorimeter. At the end, charged particles, topo-clusters which have not been modified by the algorithm, and remnants of TC which have had part of their energy removed remain.

First, well-measured tracks are selected. The algorithm then attempts to match each track to a single TC. The expected energy in the calorimeter is parametrized based on the TC position and the track momentum. The algorithm evaluates the probability that the particle energy was deposited in more than one TC and decides if it is necessary to add additional TCs for the subtraction. Finally, if the remaining energy in the TCs is consistent with the expected shower fluctuations of a single particle's signal, the TC remnants are removed.

This procedure is applied to tracks sorted in descending p_T order, first to cases where only a single TC is matched to the track, and then to the other selected tracks. This methodology is illustrated in Figure 2.16.

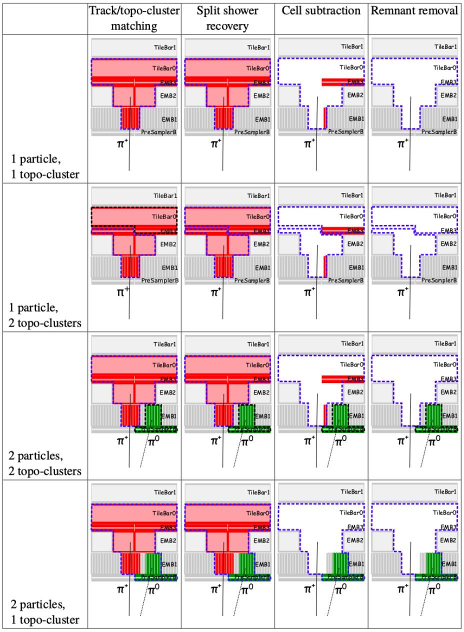


Figure 2.16. Idealised examples of how the algorithm is designed to deal with several different cases. The red cells are those which have energy from the π^+, the green cells energy from the photons from the π^0 decay, the dotted lines represent the original TC boundaries with those outlined in blue having been matched by the algorithm to the π^+, while those in black are yet to be selected. The different layers in the electromagnetic calorimeter (Presampler, EMB1, EMB2, EMB3) are indicated. In this sketch only the first two layers of the Tile calorimeter are shown (TileBar0 and TileBar1).

The samples used for the algorithm's validation include single-pion and dijet Monte Carlo (MC) samples. Charged pions dominate the charged component of the jet, which on average constitutes two-thirds of the visible jet energy. Another quarter of the jet energy is contributed by photons from neutral hadron decays, and the remainder is carried by neutral hadrons that reach the calorimeter. Because the majority of tracks are generated by charged pions, particularly at low p_T, the pion mass hypothesis is assumed for all tracks used by the particle flow algorithm to reconstruct jets.

For a given TC i, the fraction of the particle's true energy contained in the i-th TC E_i^π, relative to the total true energy deposited by the particle in all clustered cells, is defined as:

$$\varepsilon_i = \frac{E_i^\pi}{\sum_i E_i^\pi} \tag{2.20}$$

The TC with the maximum value of ε_i is designated as the leading TC: $\varepsilon_{lead} = \max_i \varepsilon_i$. The minimum number of TCs needed to capture at least 90% of the particle's true energy is denoted by n_{90} and defined by:

$$\sum_{i=0}^{n_{90}} \varepsilon_i > 90 \tag{2.21}$$

Moreover, TCs can be contaminated by several particles, hence the purity of TC i can be denoted as ρ_i and is defined by:

$$\rho_i = \frac{E_i^\pi}{E_i^{all}} \tag{2.22}$$

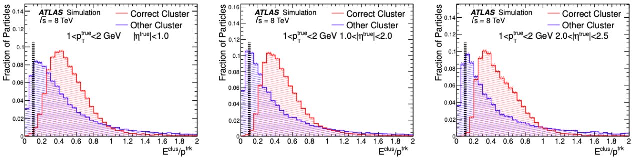
A strict selection criterion is imposed on tracks, requiring at least 9 hits in the silicon detector. Additionally, tracks must be within $|\eta| < 2.5$ and have $p_T > 0.5\ GeV$.

The algorithm attempts to match each selected track to a single topological cluster (TC) by computing the angular distance $\varDelta R$ between the extrapolated track at the second electromagnetic (EM) layer and the TC barycenter. A preliminary selection of TCs is made, requiring that for the selected i-th TC, the ratio E_i^π/p_T exceeds 10%. For TCs containing at least 90% of the true energy, the distribution of E_i^π/p_T is displayed in Figure 2.17 across different p_T and η ranges.

For TC j, which is not the nearest to the track, the distribution is peaked at less than 10%. In contrast, for the well-associated TC, the peak is at more than 50%, with the difference increasing as p_T increases.

To initiate the energy subtraction, it is necessary to have a parametrization of the average deposited energy:

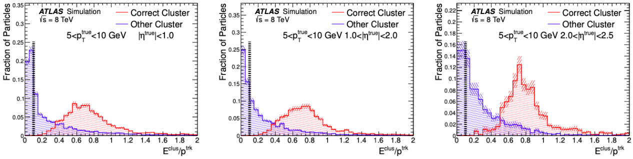
$$\langle E_{dep} \rangle = p_T \left\langle \frac{E_i^{ref}}{p_T^{ref}} \right\rangle \tag{2.23}$$



(a) $1 < p_{\mathrm{T}}^{\mathrm{true}} < 2\,\mathrm{GeV}$, $|\eta| < 1.0$

(b) $1 < p_{\mathrm{T}}^{\mathrm{true}} < 2\,\mathrm{GeV}$, $1.0 < |\eta| < 2.0$

(c) $1 < p_{\mathrm{T}}^{\mathrm{true}} < 2\,\mathrm{GeV}$, $2.0 < |\eta| < 2.5$



(d) $5 < p_{\mathrm{T}}^{\mathrm{true}} < 10\,\mathrm{GeV}$, $|\eta| < 1.0$

(e) $5 < p_{\mathrm{T}}^{\mathrm{true}} < 10\,\mathrm{GeV}$, $1.0 < |\eta| < 2.0$

(f) $5 < p_{\mathrm{T}}^{\mathrm{true}} < 10\,\mathrm{GeV}$, $2.0 < |\eta| < 2.5$

Figure 2.17. The distributions of E_i/p_T for the TC i with >90 % of the true energy of the particle and the closest other TC in ΔR. The data are taken from a dijet sample without pile-up with $20 < p_T < 500\ GeV$ and the statistical uncertainties on the number of MC simulated events are shown as a hatched band. A track is only used for energy subtraction if a TC is found inside a cone of $\Delta R = 1.64$ for which $E_i/p_T > 10$ %, as indicated by the vertical dashed line.

where "ref" indicates that the parametrization uses a single-particle sample. The expression $\left\langle \frac{E_i^{ref}}{p_T^{ref}} \right\rangle$ is a function of p_T, η, and the Layer of High Energy Density ($LHED$), which will be described later. The spread of expected energy deposition is derived from the standard deviation of the E_i^{ref}/p_T^{ref} distribution as $\sigma(E_{dep}) = p_T\ \sigma(E_i^{ref}/p_T^{ref})$. The dense electromagnetic (EM) shower has an ellipsoidal shape in the $\eta \times \phi$ plane. It is appropriate to longitudinally collocate this core to begin the energy subtraction. The $LHED$ is defined as the layer that exhibits the greatest variation in energy density relative to the variation in the number of interaction lengths in the layers. The procedure used to determine the Layer of High Energy Density (LHED) is as follows:

- First, compute the energy density per layer (l) per cell (c), expressing the volume of the cells (V_c^l) in terms of radiation lengths (X_0) as:

$$\rho_c^l = \frac{E_c^l}{V_c^l} \left[\frac{MeV}{X_0^3} \right] \tag{2.24}$$

Only cells in the TC matched with the track are considered.

- Cells are then weighted with a Gaussian centered at the extrapolated track position, with a spread of $\Delta R = 0.035$, similar to the Molière radius in an LAr calorimeter (w_c^l).
- A weighted energy density per layer is then computed as:

$$\langle \rho \rangle^l = \sum_c w_c^l \rho_c^l \tag{2.25}$$

- Finally, assuming d^l is the depth in radiation lengths of the layer l:

$$\Delta\rho^l = \frac{\langle \rho \rangle^l - \langle \rho \rangle^{l-1}}{d^l - d^{l-1}} \tag{2.26}$$

  where $d_0 = 0\ \lambda_I$ and $\langle \rho \rangle^0 = 0\ MeV/X_0^3$, and the first layer of the calorimeter has index 1.

- The Layer of High Energy Density is then determined by:

$$LHED = \max_l \Delta\rho^l \tag{2.27}$$

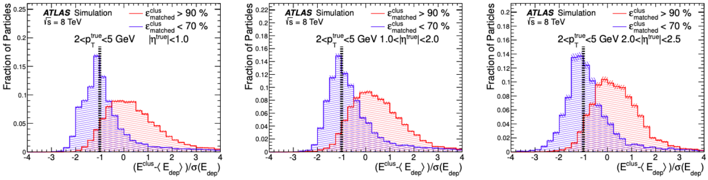
Particles do not always deposit all their energy within a single TC. The discriminant used to handle these cases is defined as the significance of the difference between the expected energy ($\langle E_{dep} \rangle$) and the energy of the matched TC (E_i):

$$S(E_i) = \frac{E_i - \langle E_{dep} \rangle}{\sigma(E_{dep})} \tag{2.28}$$

The distribution of $S(E_{dep})$ is shown in Figure 2.18 for two categories of matched TC: those with $\varepsilon_i > 90\%$ and those with $\varepsilon_i < 70\%$.

From the dashed line, it is clear that more than 90% of TCs with $\varepsilon > 90\%$ have $S(E_{dep}) > -1$. Based on this observation, a split recovery is applied when $S(E_{dep}) < -1$ for TCs within a $\Delta R = 0.2$ from the track's extrapolated position. In this way, the split recovery primarily addresses cases where $\varepsilon_i < 70\%$. Finally, the subtraction procedure can be described as follows:

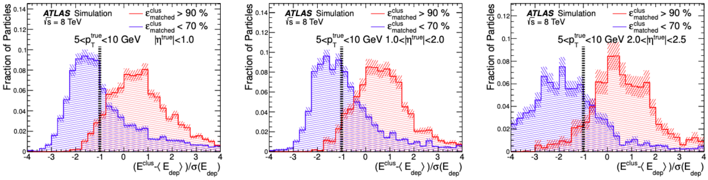
Once a set of TCs is selected, the subtraction begins. If the energy in the selected TCs, denoted as E_{TC}, is less than the expected



(a) $2 < p_{\mathrm{T}}^{\mathrm{true}} < 5\,\mathrm{GeV}$, $|\eta^{\mathrm{true}}| < 1.0$

(b) $2 < p_{\mathrm{T}}^{\mathrm{true}} < 5\,\mathrm{GeV}$, $1.0 < |\eta^{\mathrm{true}}| < 2.0$

(c) $2 < p_{\mathrm{T}}^{\mathrm{true}} < 5\,\mathrm{GeV}$, $2.0 < |\eta^{\mathrm{true}}| < 2.5$



(d) $5 < p_{\mathrm{T}}^{\mathrm{true}} < 10\,\mathrm{GeV}$, $|\eta^{\mathrm{true}}| < 1.0$

(e) $5 < p_{\mathrm{T}}^{\mathrm{true}} < 10\,\mathrm{GeV}$, $1.0 < |\eta^{\mathrm{true}}| < 2.0$

(f) $5 < p_{\mathrm{T}}^{\mathrm{true}} < 10\,\mathrm{GeV}$, $2.0 < |\eta^{\mathrm{true}}| < 2.5$

Figure 2.18. The significance of the difference between the energy of the matched TC and the expected deposited energy $\langle E_{dep} \rangle$ and that of the matched topo-cluster, for $\pi^{\pm}$ when < 70 % and > 90 % of the true deposited energy in TC is contained in the matched TC for different p_T and $|\eta|$ ranges. The vertical line indicates the value below which additional TC are matched to the track for cell subtraction. Subfigures (a)–(f) indicate that a single cluster is considered $(93, 95, 95, 94, 95, 91)$ % of the time when $\varepsilon > 90$ %; while matched additional TC are considered $(49, 39, 46, 56, 52, 60)$ % of the time when $\varepsilon < 70$ %. The data are matched taken from a dijet sample without pile-up with $20 < p_T < 500\ GeV$ and the statistical uncertainties on the number of MC simulated events are shown as a hatched band.

energy $\langle E_{dep} \rangle$, then all the energy is removed. Otherwise, a cell-by-cell subtraction must be performed. Starting from the extrapolated track position in the *LHED*, a parametrized shower model is used to extrapolate the most likely energy density profile in each layer. This profile is determined from a pure $\pi^{\pm}$ sample and depends on p_T, η, and *LHED*. Then, equally spaced rings in $\eta - \phi$ space are created around the extrapolated track, ensuring that at least one cell is contained in each ring. Subsequently, the average energy density of each ring is computed, and the rings are ranked in descending order of energy density.

The subtraction begins from the ring of high energy density within the *LHED*. If the energy in the ring is less than the energy needed to reach $\langle E_{dep} \rangle$, then each cell in the ring is scaled down, and the process is halted. An example of this subtraction procedure is shown in Figure 2.19.

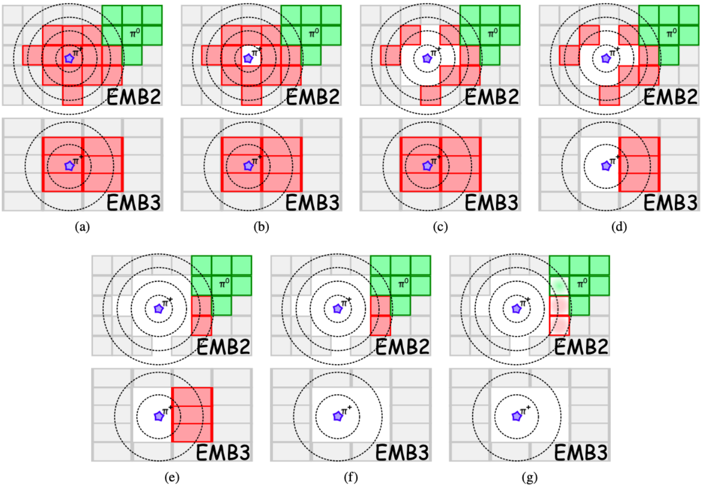
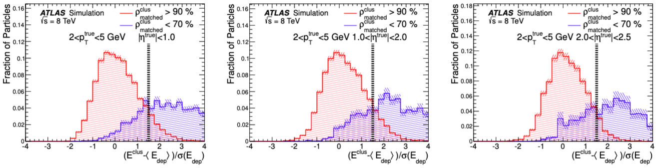


Figure 2.19. An idealised example of how the cell-by-cell subtraction works. Cells in two adjacent calorimeter layers (EMB2 and EMB3) are shown in grey if they are not in clusters, red if they belong to a π^+ cluster and in green if contributed by a π^0 meson. Rings are placed around the extrapolated track (represented by a star) and then the cells in these are removed ring by ring starting with the centre of the shower, (a), where the expected energy density is highest and moving outwards, and between layers. This sequence of ring subtraction is shown in subfigures (a) through (g). The final ring contains more energy than the expected energy, hence this is only partially subtracted (g), indicated by a lighter shading.

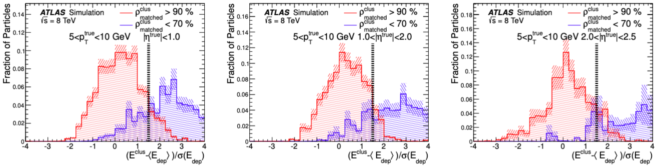
If the remaining energy in the TC is less than $1.5\sigma(E_{dep})$, it is assumed that the TC was created by a single particle, and the residual energy is considered as a shower fluctuation. Figure 2.20 illustrates how this criterion effectively separates multiple-particle from single-particle TCs. After the final step, the set of selected tracks and residual TCs should ideally represent the reconstructed event with no double counting of charged and neutral particles, respectively.



(a) $2 < p_T^{true} < 5$ GeV, $|\eta^{true}| < 1.0$

(b) $2 < p_T^{true} < 5$ GeV, $1.0 < |\eta^{true}| < 2.0$

(c) $2 < p_T^{true} < 5$ GeV, $2.0 < |\eta^{true}| < 2.5$



(d) $5 < p_T^{true} < 10$ GeV, $|\eta^{true}| < 1.0$

(e) $5 < p_T^{true} < 10$ GeV, $1.0 < |\eta^{true}| < 2.0$

(f) $5 < p_T^{true} < 10$ GeV, $2.0 < |\eta^{true}| < 2.5$

Figure 2.20. The significance of the difference between the energy of the matched TC and the expected deposited energy $\langle E_{dep} \rangle$ for $\pi^{\pm}$ with either < 70 % or > 90 % of the total true energy in the matched TC originating from the $\pi^{\pm}$ for different p_T and $|\eta|$ ranges. The vertical line indicates the value below which the remnant TC is removed, as it is assumed that in this case no other particles contribute to the TC. Subfigures (a)–(f) indicate that when $\rho > 90$ % the remnant is successfully removed $(91, 89, 94, 89, 91, 88)$ % matched of the time; while when $\rho < 70$ % the remnant is retained $(81, 80, 76, 84, 83, 91)$ % of the time. The data are matched taken from a dijet sample without pile-up with $20 < p_T < 500\ GeV$ and the statistical uncertainties on the number of MC simulated events are shown as a hatched band.

2.3.3 Electrons

Electron reconstruction is a critical process in ATLAS. This process involves identifying and characterizing electrons produced in particle collisions. Given the fundamental role electrons play in a wide range of physical phenomena, accurate electron reconstruction is crucial.

The reconstruction of electrons begins with the detection of signals in the electromagnetic calorimeter. When an electron (or a positron) enters the calorimeter, it initiates a shower of secondary particles. This shower produces a characteristic spatial and energy distribution pattern within the calorimeter, which can be used to identify the incoming particle as an electron and to measure its energy.

However, electron reconstruction is not limited to calorimeter data alone. It also involves the tracking system in the central region with

$|\eta| < 2.47$. The combination of tracking and calorimeter data provides a comprehensive picture of each electron's properties by estimating the energy from the calorimeter and the directional information from the tracker.

A visual sketch of the electron is shown in Figure 2.21.

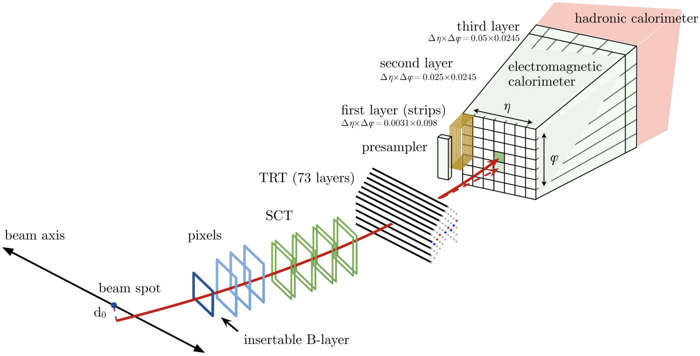


Figure 2.21. Sketch of an electron crossing the ATLAS sub-systems. The electron path is drawn with a solid red line, whereas the dashed red line corresponds to a Bremsstrahlung photon. Figure taken from Ref. [48]

A crucial aspect of electron reconstruction is distinguishing genuine electron signals from those of other particles and from various forms of background noise. For instance, photons can create similar signals in the calorimeter but do not leave tracks in the tracking detector. Sophisticated algorithms and pattern recognition techniques are employed to differentiate electrons from photons and other background signals. These algorithms consider factors such as the shape of the energy deposit in the calorimeter, the presence or absence of a corresponding track, and the consistency of the track with the primary collision vertex. In addition to identifying individual electrons, electron reconstruction algorithms must also address complications such as bremsstrahlung radiation. As electrons pass through material, they can emit bremsstrahlung photons, which may then convert back into electron-positron pairs, complicating the reconstruction process. Advanced techniques are used to correct for these effects and to accurately reconstruct the original electron's properties.

The process of electron reconstruction relies on calibrated topo-clusters. Tracks compatible with these topo-clusters are then fitted using a Gaussian Sum Filter method [49], specifically designed to accommodate the nonlinear effects of bremsstrahlung radiation during electron track fitting. The topo-clusters with associated tracks are then used to construct a dynamically-sized super-cluster [50]. This expansive super-cluster is designed to incorporate additional photon clusters, which may emanate from the bremsstrahlung phenomenon associated with the initiating electron.

In this context, several sources of background can be found, including jets enriched with electromagnetic components, electrons from photon conversions, and those originating from the semi-leptonic decay of heavy-flavour hadrons. Despite being real electrons, those arising from photon conversions and heavy-flavour hadron decays are treated as background in the context of electrons generated during the primary collision event. A sophisticated identification algorithm is deployed to separate background electrons from the "interesting" ones. This algorithm uses a likelihood discriminant where the processes $Z \rightarrow ee$ and $J/\Psi \rightarrow ee$ are considered as signal and di-jet events as background. Operational thresholds are established at three levels—Loose, Medium, and Tight—to balance minimizing signal loss against maximizing background rejection.

To enhance the selectivity of electron candidates, isolation criteria can be applied based on nearby tracks or energy deposits. Various isolation criteria are applied depending on a predefined distance and the p_T of the electron candidate.

The initial estimation of the electron's energy is derived from the super-cluster, and further refinements are applied from Monte Carlo-derived corrections. All remaining differences are considered via dedicated calibration on data, such as the Z boson decaying into electrons. The overall efficiency, as the product of reconstruction, identification, and isolation, is then determined using a tag-and-probe method on $Z \rightarrow ee$ events. This approach applies Scale Factors (SF) on the order of 1% as a function of E_T and η to account for the mentioned differences between data and MC in terms of electron efficiency reconstruction.

2.3.4 Photons

Photon reconstruction primarily relies on the electromagnetic calorimeter, which is designed to detect particles that interact electromagnetically, such as electrons and photons [50]. When a photon enters the calorimeter, it triggers a cascade of interactions leading to an electromagnetic shower. This shower, comprising numerous secondary particles, generates a distinctive pattern of energy deposits across the calorimeter cells. These patterns are crucial for identifying the incident particle as a photon and for accurately measuring its energy.

Unlike electrons, photons do not leave tracks in the inner detector due to their lack of charge. Therefore, photon identification relies solely on the analysis of calorimeter data, without the evidence provided by the tracking system. The primary challenge in photon reconstruction is distinguishing photons from electrons (or other electromagnetic showers initiated by charged particles) and mitigating background noise from other sources.

Analysis of Shower Shape: Algorithms dedicated to analyzing the energy distribution within the calorimeter are key in separating photons from electrons and other background signals. The shape of the electromagnetic shower, particularly its compactness and symmetry, serves as a distinguishing factor. Photons typically exhibit a tighter shower profile than electrons, which may show effects of bremsstrahlung. [51]

Isolation Criteria: Isolation measures further refine photon identification by evaluating extraneous energy deposits surrounding the photon candidate. Photons emerging from high-energy particle decays or directly from the hard scatter are likely to be more isolated compared to those emanating from jet fragmentation. [52]

Photon Conversions: Identifying photons that convert to electron-positron pairs before reaching the calorimeter enhances the photon identification process. The ATLAS detector is equipped to track these conversions, providing additional evidence of the photon's presence and aiding in the precise measurement of its energy.

Calibrations and Corrections to the Calorimeter: Initial estimates of photon energy, based on calorimeter data, are adjusted to the electromagnetic scale. These estimates are refined through considerations of material effects and calibrations. Corrections are also refined using Monte Carlo simulations and control data samples (e.g., $Z \to ee\gamma$ events where an electron emits a photon) to fine-tune the energy scale and resolution. [53]

A multi-variable analysis, often incorporating machine learning, is utilized to optimize the separation between signal photons and background. Operational thresholds (Loose, Medium, Tight) are designed to balance selection efficiency with purity.

2.3.5 Muons

Muon reconstruction is another essential component of the ATLAS experiment [54], and a dedicated Muon Spectrometer (MS) has been built to measure it. Muons, the heavier cousins of electrons, penetrate much further through matter, making their detection and reconstruction a unique challenge. At LHC energies, muons behave like minimum ionizing particles in the calorimeter and can be measured via their track signature in the Inner Detector (ID) and MS.

In the ATLAS experiment, muons are reconstructed through various methods, each depending on the signature they leave in the various sub-detectors. The five primary types of reconstructed muons are:

- *Stand-Alone Muons* (SA): These muons are identified solely based on information from the muon spectrometer, without any input from the inner tracking system. Stand-alone muons are particularly useful for identifying high-momentum muons that may not leave significant tracks in the inner detector due to their penetrating nature.

- *Combined Muons* (CB): Combined muons are reconstructed by integrating tracks from both the muon spectrometer and the inner tracking system. This method provides the most precise measurement of the muon's momentum, as it utilizes the full trajectory information across the detector. The combination process involves fitting tracks from both systems to form a single, optimized trajectory.

- *Segment Tagged Muons* (ST): These muons are primarily identified through the inner tracking system but are tagged as muons by matching a track segment in the muon spectrometer to the extrapolated track from the inner detector. Segment tagged muons allow for the identification of muons that only leave partial information in the muon spectrometer.

- *Calorimeter Tagged Muons* (CT): Calorimeter tagged muons are identified when a track in the inner detector is associated with an energy deposit in the calorimeter that is consistent with a muon's interaction. This type is useful for identifying low-momentum muons that may not penetrate through to the muon spectrometer.

- *Extrapolated Muons* (ME): Extrapolated muons are identified when muon candidates in the inner tracking system are extrapolated to the muon spectrometer, but without a requirement for specific matched hits in the spectrometer. This method can help in identifying muons in regions of the spectrometer with reduced coverage or efficiency.

Each of these reconstruction methods plays a crucial role in ATLAS's ability to accurately identify muons, allowing for a wide range of muon momenta and interaction types to be analyzed.

Distinguishing genuine muon signals from background sources is a critical aspect of muon reconstruction. Backgrounds can include hadronic punch-through or decays in flight of other particles mimicking muon signals. Advanced algorithms and pattern recognition techniques are employed to differentiate true muons from these backgrounds. These algorithms assess factors such as the consistency of the track across different detector layers, the energy deposit profile, and alignment with the primary collision vertex.

As with electrons, isolation criteria also play a crucial role in muon selection, helping to distinguish prompt muons originating from the primary interaction from those produced in secondary decays. By evaluating the absence of significant energy deposits or other particle tracks in the vicinity of a muon candidate, muons involved in clean signatures of new physics phenomena can be effectively isolated.

The initial momentum measurement of muons is derived from the curvature of their path in the magnetic field, with subsequent adjustments based on detailed simulation models and calibration against known processes, such as the decay of Z bosons to muon pairs. The overall efficiency of muon reconstruction, identification, and isolation is quantitatively assessed using data-driven methods, notably the tag-and-probe technique applied to $Z \to \mu\mu$ events. Scale factors are determined to correct for discrepancies between simulation and real data in terms of muon reconstruction efficiency, typically on the order of a few percent, depending on p_T and η.

2.3.6 Jets

In Section 1.3, jets and their formation through parton showering and hadronization processes were discussed. As mentioned, Monte Carlo (MC) simulations are necessary to estimate background and Standard Model predictions to observe discrepancies with theory. At the LHC, the effort required is much greater because jets are collimated sprays of particles produced abundantly in high-energy particle collisions. Jets have played a central role in the discovery and property measurements of many fundamental particles, such as the gluon (g) and the top quark (t-quark). As highlighted in the previous section, they were also indispensable in studying Higgs boson couplings to the bottom quarks (b-quarks). When studying the dynamics of quark and gluon scattering, it became necessary to perform quantitative analyses and calculations that go beyond event shapes. For these analyses to be possible, a deterministic set of rules on how particles are combined into jets from an experimental viewpoint had to be defined. An intuitive definition of a jet algorithm involves summing the momenta of all particles within a fixed-size cone. However, naive cone algorithms are not infrared and collinear (IRC) safe — meaning they must be insensitive to arbitrarily low-energy particles and collinear splittings. IRC safety is a useful theoretical requirement for making calculations in perturbative QCD and serves as a convenient language for describing experimental robustness against noise and detector granularity.

The standard jet reconstruction algorithm used in ATLAS is the anti-k_t algorithm, which is IRC safe. The hadrons produced in the jet can be measured in the Inner Detector (ID) and the calorimeter. The anti-k_t algorithm [55] is favored over the k_t [56] and Cambridge/Aachen [57] algorithms for jet identification. It introduces distances d_{ij} between entities (particles or pseudo-jets) i and j, and d_{iB} between entity i and the beam (B). The clustering process is inclusive and proceeds by identifying the smallest of these distances, d_{ij}^{min}:

- If $d_{ij}^{min} < d_{iB}$, then entities i and j are merged.
- If $d_{ij}^{min} > d_{iB}$, then i is classified as a jet and is removed from the list of entities.

This process is iterated until the list of entities is empty. The main distinction of the anti-k_t algorithm from the k_t algorithm lies in the definition of the distances:

$$\begin{aligned} d_{ij} &= \min(k_{t_i}^{2p}, k_{t_j}^{2p})\frac{\Delta_{ij}^2}{R^2} \\ d_{iB} &= k_{t_i}^{2p} \end{aligned} \tag{2.29}$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, and k_{t_i}, y_i, and ϕ_i are, respectively, the transverse momentum, rapidity, and azimuth of particle i. In addition to the usual radius parameter R, an extra parameter p is added to govern the relative power of the energy scale versus the geometrical scale (Δ_{ij}).

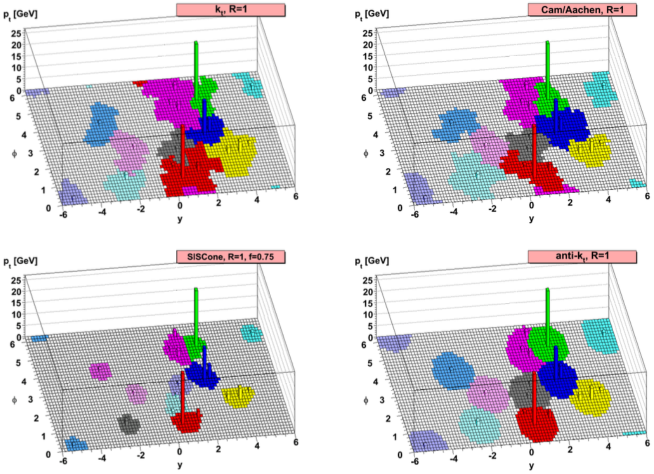


Figure 2.22. A sample parton-level event, clustered with four different jets algorithms, illustrating the "active" catchment areas of the resulting hard jets. **top left** k_t algorithm with $R = 1$ parameter. **top right** Cambridge/Aachen algorithm with $R = 1$ parameter. **bottom right** anti-k_t algorithm with $R = 1$ parameter.

This parameter determines the algorithm used. In particular, for $p = 1$, one recovers the k_t algorithm. It can generally be shown that for $p > 0$, the behavior of the jet algorithm with respect to soft radiation is similar to that observed for the k_t algorithm, because the ordering between particles matters, and for finite Δ, this is maintained for all positive values of p. Negative values of p might initially seem pathological, but they are not. The behavior with respect to soft radiation will be similar for all $p < 0$, so here only $p = -1$ is considered, referring to it as the anti-k_t jet-clustering algorithm. Finally, for $p = 0$, one recovers the inclusive Cambridge/Aachen algorithm, which weights the distances without taking into account the momentum of the entities. The behavior of different algorithms is shown in Figure 2.22.

Consider an event with a few hard particles (high k_t) and many soft particles (low k_t); the distance d_{1i} between a hard particle 1 and

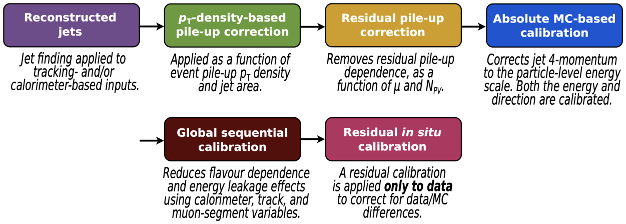


Figure 2.23. Stages of jet energy scale calibrations. Each one is applied to the four-momentum of the jet. From [58]

a soft particle i is determined exclusively by k_{t_1} and $\varDelta_{1i}$. Generally, the distance between two soft particles d_{ij} will be much greater. If a hard particle has no hard neighbors within a distance $2R$, then it will simply accumulate all the soft particles within a circle of radius R, resulting in a perfectly conical jet.

If another hard particle 2 is present such that $R < \varDelta_{12} < 2R$, then there will be two hard jets. It is not possible for both to be perfectly conical. If $k_{t_1} \gg k_{t_2}$, then jet 1 will be conical, and jet 2 will be partly conical. However, if $k_{t_1} = k_{t_2}$, neither jet will be conical, and the overlapping part will simply be divided by a straight line equally between the two. Similarly, one can determine what happens when $\varDelta_{12} < R$. Here, particles 1 and 2 will cluster to form a single jet, but the jet shape will strongly depend on the relation between their momenta. If $k_{t_1} \gg k_{t_2}$, then it will be a perfect cone around jet 1. If $k_{t_1} \simeq k_{t_2}$, the shape will instead be more complex, being the union of cones (radius $< R$) around each hard particle plus a cone (of radius R) centered on the final jet.

In this thesis, 3 types of jets are considered: small-R jets (SR), large-R jets (LR), and Variable Radius jets (VR).

Small-R Jets

In this thesis I refer to Small-R jets for those jets built with Particle Flow objects described in section 2.3.2 and using anti-k_T with R=0.4. The calibration scheme used is described in [58] and summarized in Figure 2.23. The calibration process involves initial simulation-based corrections, followed by in situ techniques to adjust for data and simulation differences and to measure jet resolution.

The calibration aim is to measure and correct the jet energy scale

(JES) and resolution (JER).

- *Smulation-based Jet Calibrations*: Jets are initially calibrated using simulation-based corrections. This involves pile-up corrections to remove the excess energy contributions from additional proton-proton interactions and absolute jet energy scale (JES) and pseudorapidity (η) calibrations. These corrections account for non-compensating calorimeter responses, energy losses in inactive material, out-of-cone effects, and biases in the jet η reconstruction.

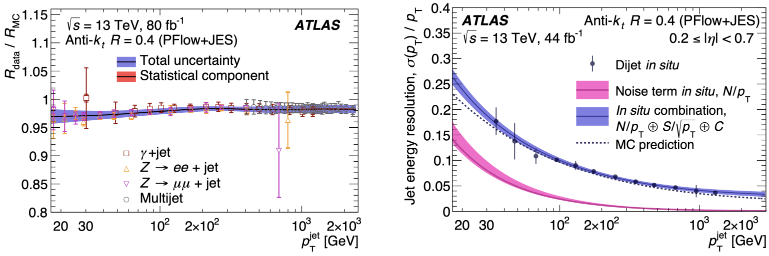
- *Global Sequential Calibration (GSC)*: After the initial corrections, the GSC applies a series of multiplicative corrections based on global jet observables (such as the longitudinal structure of energy depositions within the calorimeters, tracking information, and activity in the muon chambers) to reduce the effects of fluctuations in the jet energy measurement and improve the jet resolution. This step aims to minimize the dependence of the reconstructed jet response on the jet's internal structure and the fluctuating nature of its formation, thereby improving the precision of jet energy measurements.

- *In Situ Calibrations*: The final step involves in situ techniques that correct for differences between data and simulation and measure the jet resolution. These techniques use well-calibrated reference objects (like photons and Z bosons) and systems of jets to calibrate the jet energy scale directly from data, ensuring that the jet energy measurements in the ATLAS experiment accurately reflect the true jet energies. This step is crucial for correcting any residual discrepancies between the simulated calibrations and the actual data observed in the experiment.

After the JES calibration the Jet Energy Resolution is measured in di-jet events relying on scalar balance between the two jets in the events. The ratio of the Jet Response between data and the nominal Monte Carlo as a function of p_T is shown in Figure 2.24(a) while the Jet Energy Resolution as a function of p_T is shown in Figure 2.24(b).

Since the JER is better on Monte Carlo than on data a residual Jet energy smearing is applied to simulations.

Finally a Jet Vertex Tagger (JVT) [59] is applied to Jets with $p_T < 120GeV$ and $|\eta| < 2.5$ assigning a score to the jet of being associated to the Hard Scatter event or to a Pile-Up event.



(a) (b)

Figure 2.24. (a) Ratio of the PFlow+JES jet response in data to that in the nominal MC event generators as a function of jet p_T for Z+jet, γ+jet, and multijet in situ calibrations. (b) The relative jet energy resolution as a function of p_T for fully calibrated PFlow+JES jets.

Large-R Jets

In this thesis I refer to Large-R jets for those jets built with Particle Flow objects using anti-k_T with $R = 1.0$. These objects are commonly used to reconstruct hadronically decaying massive particles such as Z/W, Higgs bosons or Top quark at high transverse momentum where the final state particles are collimated and can be reconstructed in a single large radius jet. The Large-R jet constituents are Topological Calorimeter-Cell Clusters that are calibrated with a Local Cell Weighting (LCW) described in [60]. This calibration aims at calibrating the calorimeter signals cluster-by-cluster to compensate signal losses due to clustering. All the jet object constituents are forced to point at the Hard Scatter vertex and they are assumed to be massless.

These jets are then *groomed* to minimise the contamination due to Pile-Up which can be significant due to the large radius used. Then a *trimming* [61] and the implementation for this thesis follows ad described in [62] and in Figure 2.25 :

- First the initial constituents of the already formed Large-R jet are re-clustered into smaller sub-jets using a k_t algorithm with $R = 0.2$.

- If such a sub-jet carries less than 5% of the Large-R jet it gets removed.

- The remaining constituents after this trimming procedure will form the groomed Large-R jet.

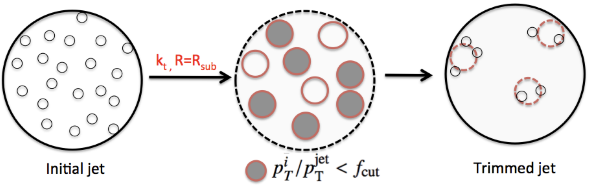
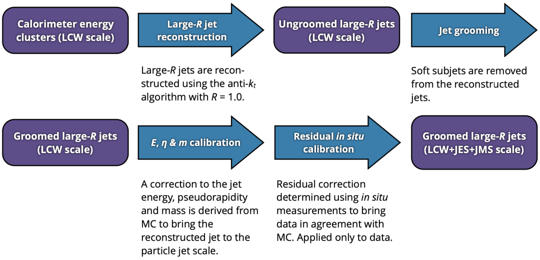


Figure 2.25. Diagram depicting the jet trimming procedure from [63].

An summary overview of the reconstruction and calibration is then described in [64] and shown in Figure 2.26.



Large-*R* jets are reconstructed using the anti-k_t algorithm with $R = 1.0$.

Soft subjets are removed from the reconstructed jets.

A correction to the jet energy, pseudorapidity and mass is derived from MC to bring the reconstructed jet to the particle jet scale.

Residual correction determined using *in situ* measurements to bring data in agreement with MC. Applied only to data.

Figure 2.26. Overview of the Large-R jet reconstruction and calibration procedure described in this paper. The calorimeter energy clusters from which jets are reconstructed have already been adjusted to point at the event's primary hard-scatter vertex. [64]

The calibration procedure is important to reconstruct precisely the mass of the resonance that is studied with the Large-R jet. First the JES is corrected to particle-level using truth matched jet with same trimming procedure applied. Then a pseudorapidity correction is applied. Since the jet mass is sensitive to wide-anglesd soft radiation the Jet Mass Scale (JMS) is calibrated separately using *di-jets* events similarly to JES calibration. Then similarly to Small-R jets a residual *in-situ* calibration is applied to JES and JMS using the balance methods for JES and the R_{trk} procedure [64] as well as a Forward Folding Method [65] for JMS.

The mass of the Large-R Jet is then given by the combined jet mass

[66] combining track and calorimeter based measurements linearly as:

$$m_{comb} = w_{calo} \times m_{calo} + w_{track} \times m_{trk} \times \frac{p_T^{calo}}{p_T^{trk}} \tag{2.30}$$

The combined Jet Mass resolution is shown in Figure 2.27 with w_{calo} and w_{trk} choosed to sum up to 1 and to minimize the m_{comb} variance.

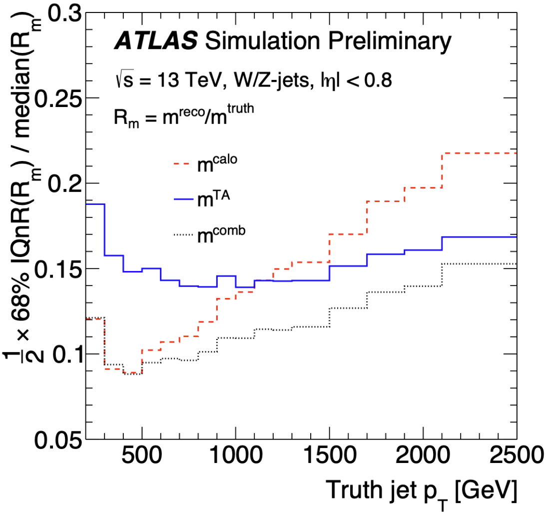


Figure 2.27. Fractional large-R jet mass resolution for the calorimeter, the track- assisted and the combined jet mass as a function of the simulated jet pT. [66]

Variable Radius Jets

Variable Radius (VR) Jets [67] are reconstructed using anti-k_t algorithm with a radius parameter that depends 2.28 recursively on the p_T of the jet itself as:

$$R = \frac{30GeV}{p_T}; \quad R_{min} = 0.02; \quad R_{max} = 0.4 \tag{2.31}$$

VR Jets constituents are tracks and these jets are used in this thesis for Flavour Tagging purposes in boosted environment. These

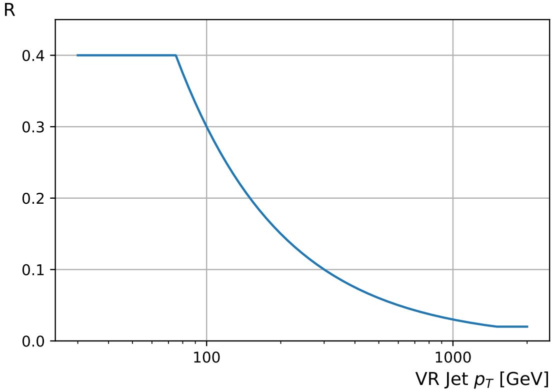


Figure 2.28. anti-k_t radius of VR track-jets as a function of the jet p_T

values are optimized for the $H \to b\bar{b}$ at high transverse momentum where the two b-quarks hadronize collimated to each other.

As the tracks entering the jet algorithm have the momentum already calibrated, no further calibration is applied to track jets. The energy scale of such jets is about 30% less than the calorimeter jets because of the missing neutral component.

2.3.7 Tau Leptons

In the ATLAS experiment at the Large Hadron Collider (LHC), the reconstruction and identification of tau leptons are crucial for a wide range of physics analyses. Tau leptons are particularly interesting because they are the heaviest of the three generations of leptons, and they can decay into lighter leptons (electrons or muons) or into hadrons (primarily pions). This decay versatility makes their identification complex. Taus are primarily identified in two decay modes:

- Leptonic decays, where a tau decays into an electron or a muon and neutrinos. These decays leverage the established electron and muon identification and reconstruction techniques within ATLAS but require additional discrimination to separate them from direct electron or muon production processes.

- Hadronic decays, where a tau decays into hadrons (most commonly pions) and a neutrino [68]. The reconstruction of hadronically decaying taus (often denoted as $\tau_{had}\tau_{had}$) is more challenging due to the need to distinguish them from the large background of quark and gluon jets. ATLAS uses sophisticated algorithms based on Recurrent Neural Networks (RNN) [69] that combine tracking information from the inner detector with calorimeter energy deposits to identify the characteristic signatures of tau decays, such as the number of charged tracks (one or three, corresponding to the number of charged pions) and the specific energy deposition patterns in the calorimeter.

2.3.8 Missing Energy

The last fundamental aspect of the ATLAS physics object reconstruction is the transverse missing energy (E_T^{miss}) [70]. This observable is crucial to probe events with particles that can not be detected directly such as neutrinos or unknown particles that do not interact with the detector materials.

The conservation of momentum in transverse plane implies that the vector sum of the transverse momenta (p_T) of all particles produced in a proton-proton collision should be zero. However, some particles, like neutrinos, escape detection, leading to an imbalance in the measured transverse momentum. This imbalance is quantified as E_T^{miss}, representing the magnitude and direction of the missing transverse momentum.

The ATLAS detector is designed to measure the momenta of charged particles, the energy of electromagnetic and hadronic showers, and to identify muons with high precision. However, particles that do not interact with the detector components, or interact only weakly, contribute to E_T^{miss}. The computation of E_T^{miss} therefore involves combining the momenta and energy measurements of all the already-calibrated final state objects adding a track-based soft term quantifying the energy not associated to any object.

$$\vec{E}_T^{miss} = -\sum_{i \in obj} \vec{E}_T^{i} - \vec{E}_T^{soft} \tag{2.32}$$

A dedicated Overlap Removal is then applied to avoid double counting.

3 | Flavour Tagging

Flavour Tagging (FTag) is a crucial method used to identify the flavour of the parton from which a jet originates. It is indispensable in the analysis of jets originating from heavy-flavour quarks and is important for both Standard Model measurements and searches for physics Beyond the Standard Model (BSM). This method focuses on tagging jets from the hadronization of b-quarks (b-tagging) and c-quarks (c-tagging). The essential role of b-tagging is highlighted in this thesis through its application in the $VH \to b\bar{b}$ analysis, which forms the core of this thesis. Additionally, it is significant in Di-Higgs analyses, where one of the main channels, $HH \to b\bar{b}\gamma\gamma$, is depicted later. Furthermore, the significance of b-tagging extends to top-quark physics, since the top quark decays almost exclusively to $t \to Wb$.

In this chapter, sophisticated algorithms employed for FTag in ATLAS are depicted, with a particular focus on the evolution of these techniques through the different runs of the LHC. Initially, the chapter reviews the DL1r algorithm, which has been instrumental during Run2. Then the next-generation algorithms are described, GN1 and GN2, which are based on graph neural networks, developed for Run 3, they can also be used for a reprocessing of Run 2 data.

In an alternative approach to calibration, this thesis also describes a different method for calibrating b-tagging efficiency. This method was applied to VR track jets using $t\bar{t}$ semi-leptonic events and was subsequently combined statistically with the mainstream calibration.

Throughout the thesis, jets are classified based on the flavour of the originating particle, labeled as b-, c-, light- (comprising u, d, s quarks, or gluons) or tau-jets. This truth labelling is derived from Monte Carlo simulations, where the presence of b-hadron, c-hadron, or tau leptons within a cone of $\Delta R = 0.3$ around the jet axis and with a $p_T > 5$GeV is used for truth-level labelling of the jet. Once the corresponding particle is detected, the jet is labeled accordingly; otherwise, it is designated as a light-flavour jet. Moreover, any jet, whether from simulated or actual data, that meets the criteria of a tagging algorithm's operating point is tagged correspondingly as b-tagged, c-tagged, or untagged.

This chapter aims to provide a comprehensive understanding of the current and future state of flavour tagging algorithms and calibrations, setting the stage for the detailed analyses and measurements that follow and that employ these methods.

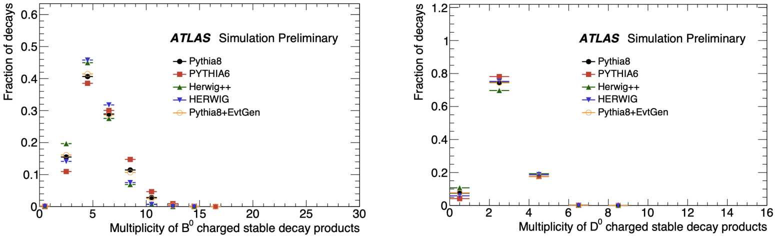
3.1 Algorithms

Heavy flavour quarks that hadronize into the detector have specific properties that can be exploited by "*ad hoc*" algorithms to effectively discriminate between b-jets, c-jets, and light-flavour jets. These properties include the presence of a heavy-flavour (HF) hadron within the jet that has longer lifetimes, higher mass, and specific decay signatures compared to lighter quarks and gluons. Consequently, heavy-flavour hadrons can travel measurable distances from the primary interaction point before decaying, leading to secondary vertices and tracks with significant impact parameters. At the core of flavour tagging, especially for b-jets and c-jets, are algorithms that integrate these characteristics to effectively separate them from light-flavour jets. This section explores the innovative computational approaches and machine learning techniques that have enhanced the accuracy and efficiency of flavour tagging, focusing particularly on the advancements through the various LHC runs.

As seen in 1.3, a quark hadronizes into a hadron due to confinement. The energy scale of this process occurs at large α_s, where perturbation theory does not hold. For this reason, fragmentation functions are used to describe the probability of a given hadron being created from the emitted quark. These fragmentation functions depend on the flavour of the quark and on the p_T of the generated hadron [71].

The average lifetime of B-hadrons and D-hadrons is $\mathcal{O}(1ps)$. This means that depending on the transverse momentum, the hadrons can travel a few millimeters in the transverse plane before decaying, providing a clear signature of a secondary vertex. Moreover, heavy flavour hadrons have a large mass, of the order of a few GeV, and can decay into multiple charged particles, as seen in Figure 3.1, carrying the largest fraction of the jet momentum.

To further distinguish b-quarks from c-quarks it is possible to exploit the peculiarity of b-jets. Since the $|V_{cb}| >> |V_{ub}|$ most of b-hadrons will decay into c-hadrons exhibiting a tertiary vertex in the decay chain within the jet.



(a) (b)

Figure 3.1. Multiplicity of charged decay products from a B^0 decay (a) and from a D^0 decay (b) as predicted from different MC generators from Ref. {ATL-PHYS-PUB-2014-008

A schematic example of Heavy Flavour jets and light-flavour jets is shown in Figure 3.2.

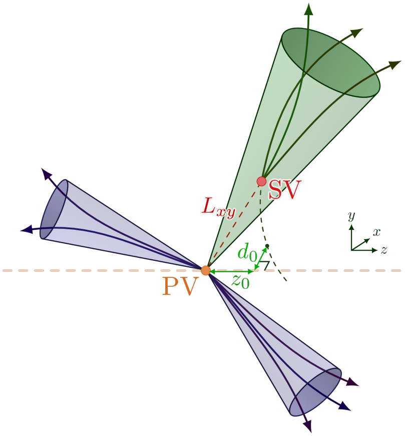


Figure 3.2. Schematic description of a Heavy-Flavour jet with tracks from the Heavy-Flavour hadron decay compared to light flavour jets.

With the advancement of machine learning techniques, the evolution of the Flavour Tagging algorithm has been extraordinary. Multidimensional correlation of track parameters is a suitable task for advanced machine learning approaches that can best identify the flavour of jets originating from heavy-flavour quarks.

In the following, the metric used to assess the discrimination power of an algorithm will be the efficiency or accuracy in identifying a given flavour jet compared to the mistag efficiency or inaccuracy in identifying a different flavour jet. In particular, the inverse of the mistag will be defined as the rejection power. For this reason, a common figure of metric will be the light/c-jet rejection as a function of the b-jet efficiency for b-tagging or light/b-jet rejection as a function of the c-jet efficiency for c-tagging. In Figure 3.3, the b-tagging evolution over time of different FTag algorithms is shown. It is clear that with the same efficiency ($\epsilon = 70\%$), the rejection power on light- and c-jets is about 4 times larger with the state-of-the-art algorithm (GN2) compared to early Run2 algorithms.

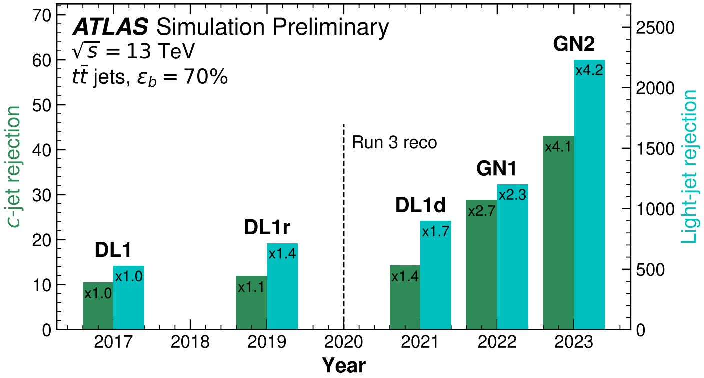


Figure 3.3. ATLAS FTag algorithm progression with time in terms of performance from Ref. [72].

Before the advent of Graph Neural Network-based single-stage approaches such as GN1 and GN2, the FTag algorithms were based on a *two*-stage approach where *low*-level taggers were combined into *high*-level taggers, returning a score prediction of the jet label. A schematic view of such algorithms (DL1, DL1r, and DL1d, ordered by time and performance) is shown in Figure 3.4.

Before describing the various algorithms, it's important to mention that all these algorithms are evaluated on Monte Carlo simulations and, for this reason, a dedicated calibration to data is necessary. Moreover,

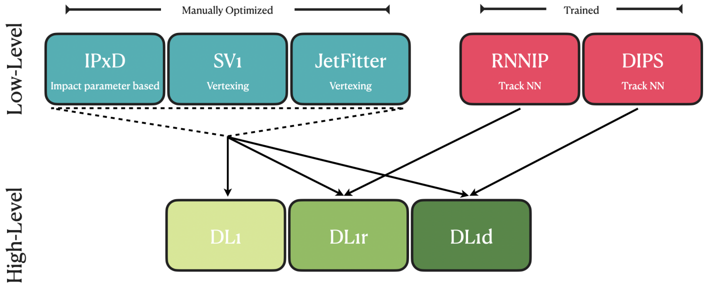


Figure 3.4. Schematic view of *DL*1, *DL*1*r*, *DL*1*d* algorithms based on *two*-stage approaches

since different Monte Carlo generators can lead to different results, it is also possible to provide MC-to-MC Scale Factors as a source of uncertainty to account for such differences.

Monte Carlo Samples. For the training and evaluation of the ATLAS Run2 FTag algorithms [73], a hybrid sample is created composed by SM $t\bar{t}$ events and high-mass $Z' \to q\bar{q}$ events.

The $t\bar{t}$ sample Matrix Element calculation is performed using Powheg Box v2 [74, 75] interfaced with Pythia 8.2 [76] for the parton showering simulation. Moreover, $t\bar{t}$ events with at least one leptonically decaying W-boson are considered. The Z' sample is generated entirely with Pythia 8.2, modifying the cross-section of the Hard-Scattering process to obtain a very broad width of the resonance and the branching fraction was set to one-third for each $b\bar{b}$, $c\bar{c}$, and light quark pairs.

Finally, as the p_T and η of the jets are inputs to the FTag algorithms for both the DL1 series and the GNN based algorithms, a resampling procedure is applied to ensure that the jet kinematics are equally distributed among each flavour category for p_T and η.

The p_T resampled distribution of the *hybrid* sample is shown in Figure 3.5.

Object Selection. FTag relies mainly on charge-particle tracks, vertices and hadronic jets. Only tracks with $p_T > 500\ MeV$ within the acceptance of the ATLAS ID ($|\eta| < 2.5$) are used in jet flavour tagging. These tracks have to further satisfy "*Loose*" selection criteria to assess a good quality and reject fake and poorly measured tracks.

Computing the Primary Vertex (PV) is also particularly important for Flavour Tagging since it defines the reference point from which

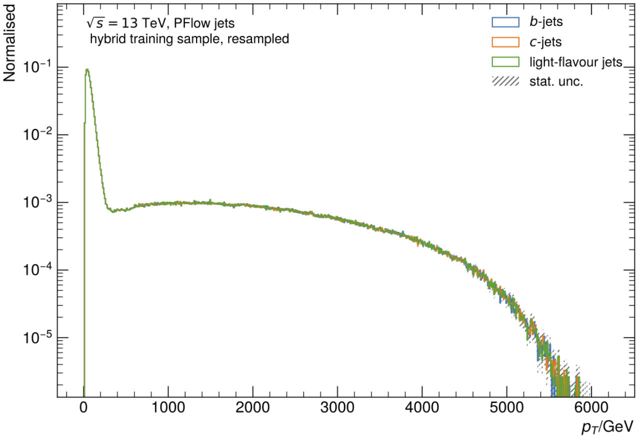


Figure 3.5. Resampled p_T distribution of the *hybrid* sample formed by $t\bar{t}$ and Z' events.

displacements of secondary vertices and tracks are computed. The resolution on the reconstructed PV is about $\mathcal{O}(10\mu m)$ in the longitudinal direction and $\mathcal{O}(1\mu m)$ transversal one [77]. The PV is chosen with the standard method of the highest sum of squared transverse momenta of contributing tracks: $\max_{vtx}\sum_{trk\in vtx}(p_T^{trk})^2$

All the reconstructed jets considered are built with Particle Flow Objects usign anti-k_t algorithm with radius parameter $R = 0.4$. Jets are required to be within the ID acceptance ($|\eta| < 2.5$) and with a $p_T > 20$ GeV. To reduce the contamination from Pile-Up jets the Jet Vertex Tagging (JVT) algorithm is applied to jets with $p_T < 60$ GeV and $|\eta| < 2.4$.

3.1.1 The DL1 Series

As shown in Figure 3.4 DL1 series of taggers are based on a *two*-stage approach. First *low*-level taggers are applied to reconstruct the characteristic features of b-jets.

The first kind of *low*-level tagger is based on the large impact parameter (IP) of the tracks originated from heavy flavour hadron decays. The IP based taggers are IPxD [78], manually optimized, RNNIP [79], a Recurrent neural network used for DL1r tagger and DIPS, based on Deep Sets Neural Networks [80].

The second type of *low*-level tagger is based on the Secondary

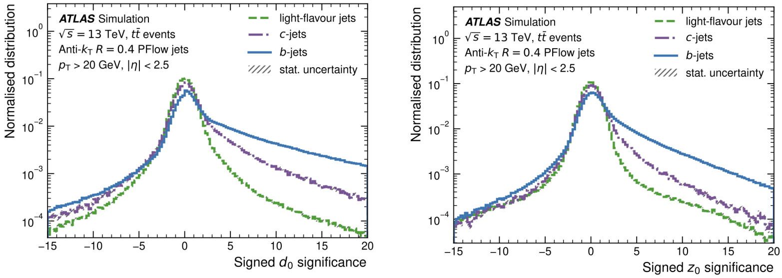
Vertex finding. The algorithm used are SV1 [81], attempting to reconstruct an inclusive secondary vertex and JetFitter [82] aiming at fully reconstructing the full *PV-to-b-to-c* decay chain.

IP based algorithms

As already mentioned, the IP-based algorithms leverage the significant impact parameters (IPs) that tracks from the decay of heavy flavour hadrons (both c and b-hadrons) often exhibit due to their substantial lifetimes. This results in b-/c-hadrons that travel measurable distances from the primary collision point before decaying. The primary metrics utilized by these algorithms include the IP itself and its significance, which compares the IP to its uncertainty.

$$S_{d_0} = \frac{d_0}{\sigma_{d_0}}; \qquad S_{z_0} = \frac{z_0}{\sigma_{z_0}} \tag{3.1}$$

We will commonly refer to the signed impact parameter where the sign is essentially given by the scalar product of the IP vector with the jet axis direction. The significance enhances the discrimination power between jets containing heavy flavour hadrons and the others. In Figure 3.6 this track level discrimination is shown comparing the transverse and longitudinal impact parameters for different jet flavours.



(a) (b)

Figure 3.6. The (a) transverse and (b) longitudinal signed impact parameter significance of tracks for *b*-jets, *c*-jets and light-flavour jets in $t\bar{t}$ events from Ref. [73].

IPxD The IPxD algorithms represent a class of IP-based taggers that utilize detailed histograms of IP significance to assign likelihood scores to each track within a jet (p_b^{trk}, p_c^{trk}, p_l^{trk}). IP3D embeds IP2D since it exploits the correlation between transversal and longitudinal impact parameter while the latter is based only on transversal dimension. The methodology behind IPxD involves segmenting the IP significance into bins (this is performed in both d_0 and z_0 for IP3D), each corresponding to a particular likelihood for the track to belong to a given type of jet. From this it is possible to compute a jet-level Log Likelihood Ratio discriminant as the sum of the logarithms of the per-track probability ratios for each jet-flavour hypothesis:

$$LLR_{b-l} = \sum_{trk} log \frac{p_b^{trk}(S_{d_0/z_0})}{p_l^{trk}(S_{d_0/z_0})} \tag{3.2}$$

This can be computed for each combination of two of the three hypothesis: b vs. l, b vs. c, c vs. l.

RNNIP RNNIP enhances traditional IP-based tagging algorithms by integrating a recurrent neural network (RNN) architecture, specifically designed to handle sequential data with inherent correlations. Unlike static models that treat input features independently, RNNIP capitalizes on the sequential nature of tracks within a jet, recognizing patterns and dependencies that unfold over the sequence of tracks.

The core component of RNNIP is its use of Long Short-Term Memory (LSTM) cells, a type of RNN architecture optimized to avoid the vanishing gradient problem, allowing it to learn long-range dependencies. Each LSTM cell processes individual track data sequentially, updating its internal state based on the IP significance, track kinematics, and possibly other relevant track attributes such as the number of hits, the track's angular relation to the jet axis ΔR or the track relative momentum with respect to the jet, defined as p_T^{frac}. This process not only assesses individual track features but also their dynamic interaction with the features of preceding tracks within the jet. Tracks are sorted by descending S_{d_0} since RNNs do require a specific order. A schematic drawing of the RNNIP architecture is shown in Figure 3.7. Each track is originally described as a vector of features that is then converted thorough the RNN block into a 50-dimension vector input to a fully connected Neural Network that in the end provides the RNNIP tagging scores. These continuous scores represent the probability that the input jet originates from by a b-/c- or light-flavour quark and for this reason a Soft Max node is used to ensure that: $p_b + p_c + p_l = 1$.

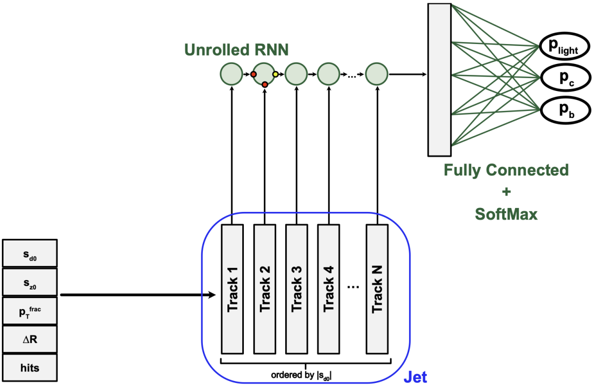


Figure 3.7. Schematic drawing of the RNNIP neural network architecture. The signed transverse and longitudinal impact parameters input variables correspond to the lifetime-correlated signed transverse and longitudinal impact parameter significances, while p_T^{frac} and ΔR represent the fraction of transverse momentum carried by the track relative to the jet and the angular distance between the track and the jet axis, respectively.

The probabilities can then be combined together to create a *b*-/*c*-tagging discriminant function. This approach will be generalised later to all the algorithms outputting separate jet-flavour probabilities.

$$D^{b/c}_{RNNIP} = \ln\left(\frac{p_{b/c}}{f_{c/b}p_{c/b} + (1 - f_{c/b})p_l}\right) \tag{3.3}$$

In general $f_{b/c}$ represents the prior *b-vs-c* jet flavour composition of the sample and can be tuned to change the relative importance of *b*-/*c*- vs light-flavour jet rejection. The *b*-tagging score evaluated on $t\bar{t}$ events with $pT \in [20, 250]$ GeV is shown in Figure 3.8.

DIPS The DIPS algorithm represents a further advancement by applying the Deep Sets theory [83], which is inherently permutation invariant and thus does not require sorting or sequential processing of tracks. This approach aligns more closely with the physical nature of

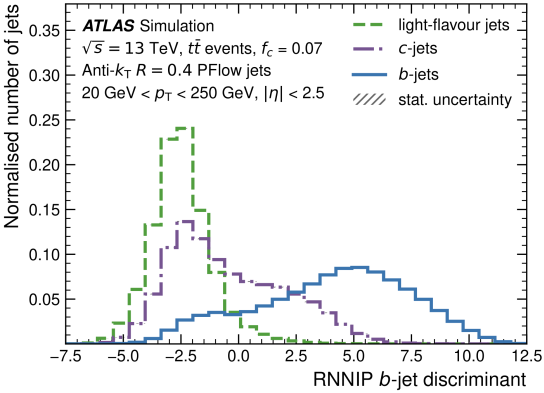


Figure 3.8. Distributions of the RNNIP *b*-btagging discriminant D^b_{RNNIP} of the RNNIP *b*-tagging algorithm for *b*-jets, *c*-jets and light-flavour jets in the baseline $t\bar{t}$ simulated events from Ref. [73]

jets, where the order of tracks is not intrinsically meaningful. DIPS processes all tracks simultaneously through a shared neural network, shown in Figure 3.9, aggregating their information before making a final classification. This method is particularly effective at handling the varied decay patterns and lifetimes of both *b*- and *c*-hadrons.

Secondary Vertexing based algorithms

Vertexing algorithms tackle the problem from a different point of view. These algorithms exploit the properties of decay vertices built out of multiple tracks that are displaced from the primary proton-proton collision point due to the relatively long lifetimes of Heavy Flavour hadrons, rather than the single displaced tracks.

SV1 SV1 is an algorithm that constructs a secondary vertex by fitting tracks that are significantly displaced from the primary vertex. This displacement is indicative of the decay of a Heavy-Flavour hadron. The algorithm uses a combination of vertex decay properties to determine the likelihood of the jet containing such a hadron. It assigns a weight based on the significance of the displacement, the mass of the reconstructed vertex, the number of tracks contributing to the vertex, and the ratio of the energy at the secondary vertex to the total jet energy.

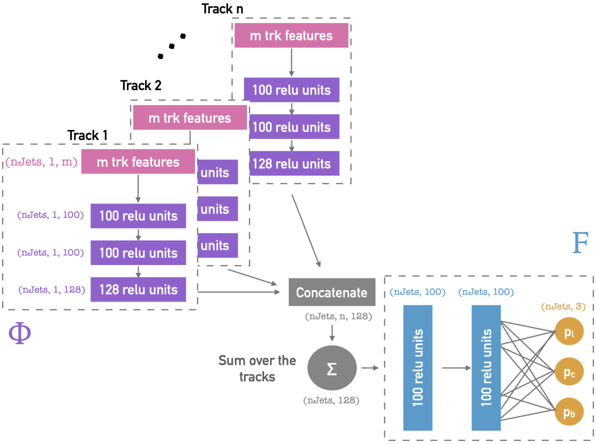


Figure 3.9. Architecture for the DIPS algorithm. The number of hidden units in the different neural network layers correspond to the final optimized architecture from Ref. [80]

SV1 runs iteratively computing the track-to-vertex association score with a χ^2 test at each iteration. The track with largest χ^2 contribution is removed and the vertex is fitted again. This process continues until the χ^2 is acceptable for a vertex and the invariant mass of the vertex is less than 6 GeV.

To evaluate the SV1 performance a simple feed forward network is used with SV1 outputs and p_T and η of the jet.

JetFitter Jet Fitter extends the secondary vertex approach by attempting to reconstruct the entire decay chain, not just one inclusive decay vertex. This algorithm integrates multiple vertices in a jet into a common decay topology, making it effective in identifying jets with complex decay structures.

High-level taggers

To maximise the flavour-tagging performance for Run 2, the output quantities of the low-level algorithms are combined using deep-learning classifiers, based on fully connected multi-layer feed-forward neural networks (NN), forming the *DL1* series.

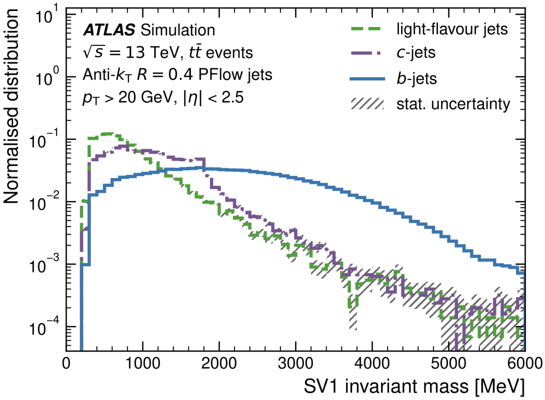


Figure 3.10. Invariant mass of secondary vertices reconstructed by the SV1 algorithm for b-jets, c-jets and light-flavour jets in the baseline $t\bar{t}$ simulated events from Ref. [73].

The hybrid sample used contains 70% of the jets from the $t\bar{t}$ sample and the remaining 30% from the Z' sample. As shown in Figure 3.4, DL1 exploits all the *low*-level tagger (IP2D, IP3D, SV1, JetFitter) outputs as input. The improvements from DL1r and DL1d is the extra input of RNNIP and DIPS score respectively. The architecture is shown in Figure 3.11.

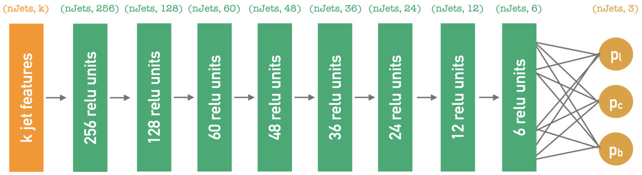


Figure 3.11. Schematic drawing of the DL1r neural network architecture. A similar architecture is used for DL1d from Ref. [73].

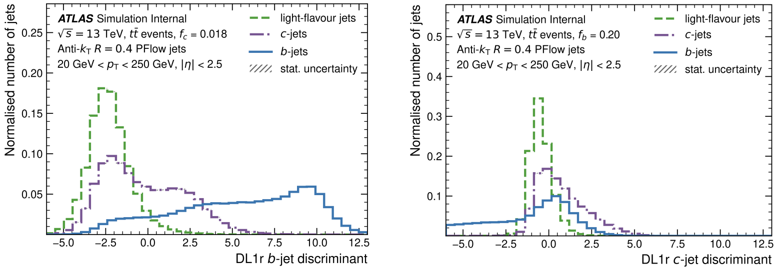
Such algorithms exploit the correlation among the *low*-level taggers to improve the discrimination between different flavour jets. Each algorithm returns a jet level probability for each flavour kind: p_b, p_c and p_l.

Similarly to what it has been shown for RNNIP it is possible to build a b-/c-tagging score:

$$D_{b/c} = \ln\left(\frac{p_{b/c}}{f_{c/b}p_{c/b} + (1 - f_{c/b})p_l}\right) \tag{3.4}$$

As already mentioned $f_{c/b}$ are tunable to adjust the relative rejection of the two backgrounds jets. In particular, f_c is tuned for b-tagging to adjust the relative background rejection between c-jets and light-jets. For DL1r the optimized values chosen are $f_c = 0.018$ and $f_b = 0.2$.

The b-/c-tagging discriminant scores are shown in Figure 3.12 for DL1r



(a) (b)

Figure 3.12. Distributions of the outputs of the DL1r algorithm (a) b-tagging (b) c-tagging score for b-jets, c-jets and light-flavour jets from Ref. [73].

Performances

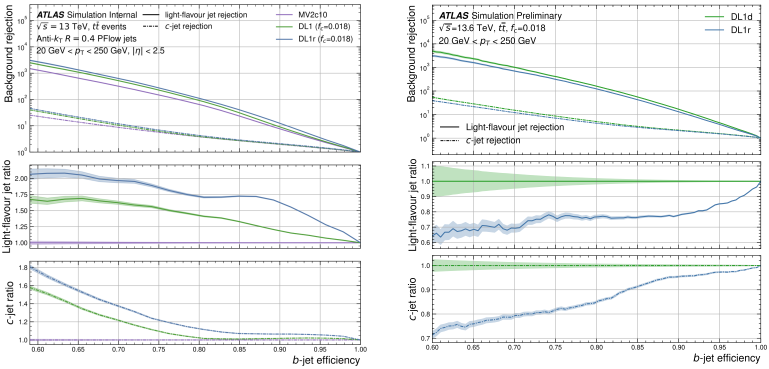
Flavour tagging performances are evaluated with efficiencies of correctly identify the correct flavour and the probability of mistakenly identifying a background jet (mis-tag rate). In particular the inverse of the mis-tag rate is defined as the rejection: $rej = 1/\epsilon$

Starting from the $D_{b/c}$ distributions it is possible to build Receiver Operating Characteristics (ROC) curves expressed as backgrounds rejection as a function of the tagging efficiency. These curves can be computed by scanning the $D_{b/c}$ distribution. In case of b-tagging for each D_b value it is possible to compute the tagging efficiency and the background rejection as:

$$\epsilon_b = \frac{N_{b-jet}(D_b > D*)}{N_{b-jet}} \qquad \epsilon_{c/l-jet} = \frac{N_{c/l-jet}(D_b > D*)}{N_{c/l-jet}} \tag{3.5}$$

The same reasoning is applicable to c-tagging just by swapping b with c. It is clear from this formula how to map the $D_{b/c}$ discriminants into the mentioned ROC curves.

In Figure 3.13(a) the performance in Run2 of the b-tagging DL1 and DL1r algorithms is shown compared to the Run1 baseline algorithm based on a Boosted Decision Tree (MV2) [78]. It can be seen that the DL1 series outperforms the MV2 series improving the light-jet rejection up to a factor of 2 and the c-jet rejection up to 80%. From Figure 3.13(b) with a Run3 $t\bar{t}$ sample, it can be seen that DL1d further improves over DL1r up to a further 30% on both c-/light-jet rejections.



(a) (b)

Figure 3.13. (a) The light-flavour and c-jet rejections as a function of b-tagging efficiency for different high-level taggers including MV2, DL1 and DL1r, The lower two panels show the ratio of the light-flavour jet rejection and the c-jet rejection of the algorithms to MV2c10. The statistical uncertainties on the rejection are calculated using binomial uncertainties and are indicated as coloured bands. Run2 $t\bar{t}$ samples from Ref. [73]. (b) The light-flavour jet (solid line) and c-jet (dashed line) rejection for the latest DL1r and DL1d algorithms. Run3 $t\bar{t}$ samples from Ref. [84].

Many ATLAS analyses of the Run2 LHC use the so called "*fixed-cut operating points*" defined by the inclusive efficiency obtained by a fixed cut on the $D_{b/c}$ score. For this reason it is also interesting to evaluate the performances for a given *working point* (WP). In particular it is interesting to observe the efficiency as a function of the transverse momentum of the jet, as shown in Figure 3.14 for the inclusive b-tagging efficiency of 77% (77% WP).

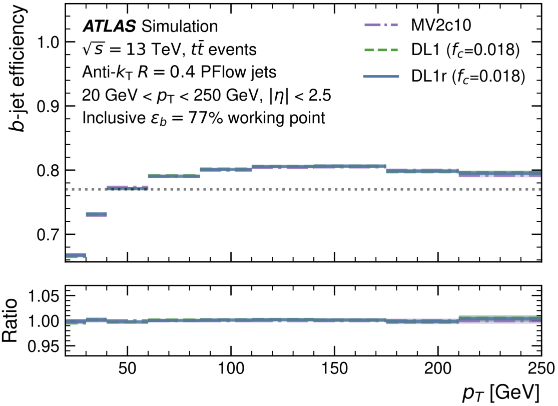


Figure 3.14. b-jet efficiency as a function of the p_T of the jet for the 77% WP from Ref. [73].

All the algorithms exhibit a consistent pattern, indicating an inherent decline at low transverse momentum. This trend at low at low p_T is mostly caused by a drop in resolution due to Multiple Scattering in the tracker and a lower pion reconstruction efficiency due to hadronic interactions. Additionally, the tracks produced at low p_T are less energetic ($p_T < 400$ GeV), which may lead to incomplete reconstruction in the Inner Detector. This relationship will be important for later discussions in this dissertation, particularly regarding the limitations imposed by inefficiencies at low p_T in the analyses.

It is also interesting to show the performances on c-tagging as displayed in 3.15

3.1.2 GNN Revolution

With the advent of Graph Neural Networks (GNN) it was clear that they would be the natural evolution of Flavour Tagging algorithms. GNNs are suitable to perform the tagging in one go. The state-of-the-art tagger for ATLAS is GN2, the evolution of GN1 [85]. The GN2 tagger is based on Transformer architecture [86]. Because of their permutation invariance GN1 and GN2 are perfect to deal with tracks associated to a jet to probe the structure of the jet and its flavour. The GNN networks have been refined with time and also parallel tasks

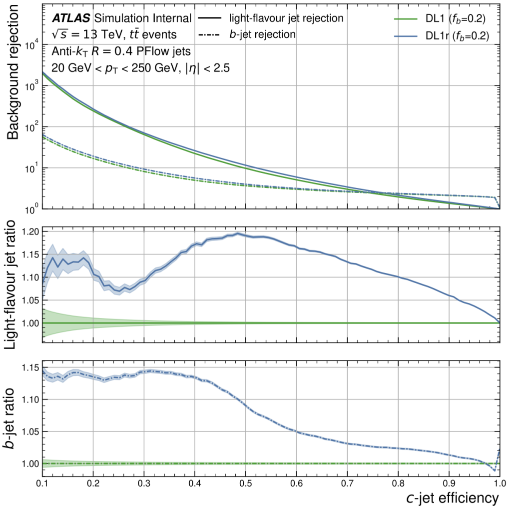


Figure 3.15. The light-flavour and b-jet rejections as a function of c-jet efficiency for DL1 and DL1r. The lower two panels show the ratio of the light-flavour jet rejection and the b-jet rejection of DL1r to DL1. The statistical uncertainties on the rejection are calculated using binomial uncertainties and are indicated as coloured bands from Ref. [73].

are included to drive the network towards better performances guided by the physics knowledge.

In Figure 3.16 the difference between the old-fashioned algorithms and the GNN based approach is shown. Tracks associated to the jets are the nodes of the graph and the GNN is trained in a supervised way to extract the jet flavour prediction, the origin label of the track and a simplified vertexing.

The GN1 and GN2 architecture

Given a jet a graph is built out of the tracks associated to it. Each track is a node with a vectorial representation of its track information. GN1 input is thus composed by 2 variables for the jet (p_T and η after the resampling) and the set of 21 variables associated to each track listed in Table 3.1. The maximum amount of tracks (nodes) allowed is 40. In those few cases with more than 40 tracks associated to the jet, the tracks with the lowest transverse impact parameter ($S(d_0)$)

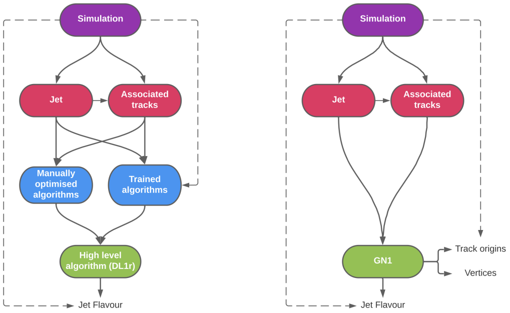


Figure 3.16. Comparison of the existing flavour tagging scheme (left) and GN1 (right). The existing approach utilises low-level algorithms (shown in blue), the outputs of which are fed into a high-level algorithm (DL1r). Instead of being used to guide the design of the manually optimised algorithms, additional truth information from the simulation is now being used as auxiliary training targets for GN1. The solid lines represent reconstructed information, whereas the dashed lines represent truth information from Ref. [85].

are discarded. The main inputs to the network are the signed Impact Parameter significances, the remaining track parameters with their uncertainty and the angular distance with respect to the jet axis.

In parallel to the jet flavour classification, two auxiliary tasks are defined for the network driven from physics reasons. The presence of auxiliary tasks improves the performances. The first auxiliary task is a node classification aiming at predict the label or origin of each track. Each track is labelled in exclusive categories as listed in Table 3.2. The second auxiliary task is needed to find non-primary vertices. It is an edge prediction that aims to find track-pair vertex compatibility. This task removes the need for a dedicated secondary vertexing algorithm.

The GN1 architecture is shown in Figure 3.17. Tracks are first concatenated with jet inputs and then they are fed into the GNN. Each node will have 21 track features associated and 2 jet features associated and can be represented as a vector of 23 dimension: t_i. Each node is first mapped into a 64 dimension vector with a per-track initialization network composed by 3 hidden layers similar to a Deep Set model [83].

The graph is then built from the output of the per-track initialization. It is a fully connected graph where each node is connected to the other nodes. At each iteration, the node representation is updated

Jet Input	**Description**
p_T, η	jet kinematic
Track Input	**Description**
$(q/p,\ d_0,\ z_0 sin\theta)$	track parameters
$\mathit{\Delta}\eta$, $\mathit{\Delta}\phi$	angles relative to jet
$(\sigma_{q/p},\ \sigma_\theta,\ \sigma_\phi)$	uncertainty on track parameters
$S(d_0)$, $S(z_0)$	signed Impact Parameter significances
Pix, SCT, IBL	Number of hits, shared hits, holes
leptonID	label for electron or muon (only GN1Lep)

Table 3.1. Input features to the GN1 model. Basic jet kinematics, along with information about the reconstructed track parameters and constituent hits are used. The track leptonID is an additional input to the GN1 Lep model.

Truth Label	**Description**
Pile-Up	Not from the primary interaction
Fake	Created from the hits of multiple particles
Primary	Does not originate from any secondary decay
FromB, FromBC	From the decay of a *b*-hadron
FromC	From the decay of a *c*-hadron
OtherSecondary	From other secondary interactions and decays

Table 3.2. Truth origins which are used to categorise the physics process that led to the production of a track.

aggregating the features of the neighbouring nodes. Each node is first fed into a fully connected layer **W** to update the representation of each node. The updated nodes are then used to compute the edge score for each node pair: $v(t_i, t_j)$.

$$v(t_i, t_j) = \mathbf{a}^{\perp}\theta\left[\mathbf{W}t_i \oplus \mathbf{W}t_j\right] \tag{3.6}$$

In this expression $\oplus$ means that the two vectors are concatenated into a 128 dimension vector, θ is a scalar activation function and **a** a fully connected layer. These scores can be used to compute attention weights:

$$a_{ij} = \text{softmax}[v(t_i, t_j)] \tag{3.7}$$

Afterwards, the node representation is computed as the weighted sum over each updated node and then passed via an activation function

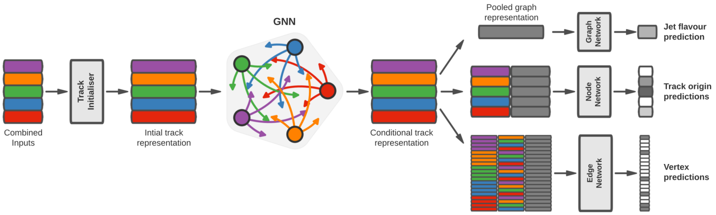


Figure 3.17. The network architecture of GN1. Inputs are fed into a per-track initialisation network, which outputs an initial latent representation of each track. These representations are then used to populate the node features of a fully connected graph network. After the graph network, the resulting node representations are used to predict the jet flavour, the track origins, and the track-pair vertex compatibility from Ref. [85]

σ:

$$t_i^{'} = \sigma\left[\sum_j a_{ij}\mathbf{W}t_j\right] \tag{3.8}$$

There are $N_{head} = 2$ heads mechanism executed in parallel that are averaged together for the GN1 case.

$$t_i^{'} = \frac{1}{N_{heads}}\sum_k t_i^{',k} \tag{3.9}$$

This sequence of mathematical operations defines a graph network layer. Three of such layers are sequentially stacked together. The output of these graph network layers is used to build a global representation of the jet. In particular a weighted sum is computed among all the nodes using the attention weights learned during the training. Three separated fully connected network are then used for the separate network tasks as shown in Table 3.3. The global classification is used to predict the flavour of the jet and it has three outputs: p_b, p_c and p_l.

The training is performed minimizing the total Loss function defined as a weighted sum of the three tasks Losses:

$$L_{tot} = L_{jet} + \alpha L_{vtx} + \beta L_{trk} \tag{3.10}$$

where $\alpha = 1.5$ and $\beta = 0.5$.

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

Table 3.3. A summary of GN1's different classification networks used for the different training objectives. The hidden layers column contains a list specifying the number of neurons in each layer.

GN2 architecture is similar to GN1 with a different attention mechanism similar to transformers and few minor differences. After the per-track initialization the initialized graph is composed by $N_{head} = 8$ multi-head attention blocks. The edge representation uses the scalar dot products.

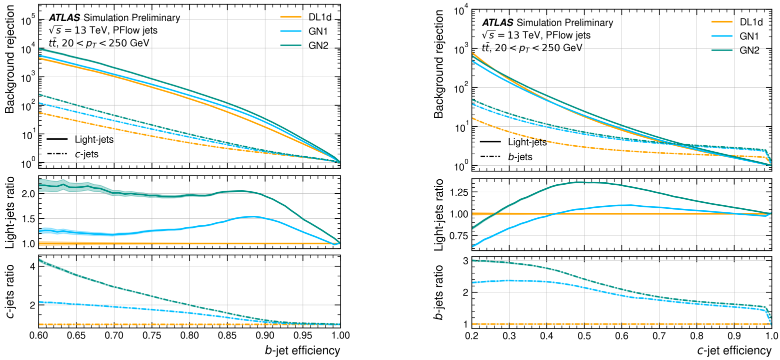
$$v(t_i, t_j) = \frac{\mathbf{W}t_i \cdot \mathbf{W}t_j}{\sqrt{s}} \tag{3.11}$$

where s indicates the dimension of the vector $\mathbf{W}$. Then similarly to GN1 a_{ij} is computed and $t_i^{'}$ is updated. GN2 differently from GN1 has 8 attention blocks instead of 2 and the number of graph layers is 6 instead of 3. Moreover stochastic dropout layers are included to reduce the overfitting. Finally another important difference in GN2 from its predecessor GN1 is that it is trained with $\mathcal{O}(200M)$ jets, about 5 times the size of the sample used for GN1.

The performances can be shown in terms of ROC curves comparing GN2 to the best performing *two*-stages algorithm DL1d [72]. From the jet flavour prediction a $D_{b/c}$ score can be built as for the DL1 series to plot the ROC curve. from Ref. 3.18, it is clear that GN2 improvements are significant. In particular on $t\bar{t}$ events for the b-tagging it is possible to improve up to a factor of 2 in light-jet rejection and a factor of 4 in c-jet rejection as compared to DL1d. For the c-tagging instead the improvement is up to 30% in light-jet rejection and up to a factor of 3 in b-jet rejection.

3.2 Calibration

All the Flavour Tagging algorithms are trained and evaluated on Monte Carlo samples. The efficiencies and mistagging rates of these algorithms need to be measured on real data to correct the source of errors on Monte Carlos due to mis-modelling. We refer to these



(a) (b)

Figure 3.18. (a) The c-jet and light-jet rejections as a function of the b-jet tagging efficiency. (b) The b-jet and light-jet rejections as a function of the c-jet tagging efficiency. Jet from a $t\bar{t}$ sample with $20 < p_T < 250$ GeV. The ratio with respect to the performance of the DL1d algorithm is shown in the bottom panels. A value of $f_c = 0.018$ is used in the calculation of the tagging discriminant for DL1d, $f_c = 0.05$ is used for GN1, and $f_c = 0.1$ is used for GN2. Binomial error bands are denoted by the shaded regions. All taggers are evaluated using the ATLAS software stackfrom Ref. [72]

measurements as calibrations since these measurements are used *a posteriori* to correct the Monte Carlo simulations.

Calibrations are performed for each jet collection used in the analysis, PFlow- and VR Track-jets for Run2, and for each fixed-cut Working Point (60%, 70%, 77% and 85% are the standard WP for b-tagging in Run2). The calibration is performed for the b-tagging efficiency and each flavour mis-tagging efficiency. The result of the calibration is expressed in terms of data/MC Scale Factors (SF) as a function of some jet kinematic variable such as p_T.

$$SF_b = \frac{\epsilon_b^{data}}{\epsilon_b^{MC}} \tag{3.12}$$

The calibration is usually performed selecting a phase space enriched in b-/c- or light-jets depending on the calibration. The standard method for the b-jet efficiency calibration uses di-leptonically decaying $t\bar{t}$ events [87, 88, 89] since top quark is know to decay almost exclusively to a W vector boson and a b-quark: $t \to bW$.

In this section I will show an alternative calibration of b-tagging efficiency in semi-leptonically decaying $t\bar{t}$ events of VR Track-jets for DL1r algorithm and Run2 dataset, which was carried out in the context of this thesis work. The methodology is similar to what has been done for PFlow-jets in [90].

3.2.1 Semi-leptonic calibration of b-efficiency

In this section I present the measurement of the b-tagging efficiency in semi-leptonically decaying $t\bar{t}$ events. A schematic diagram of the process is represented in figure 3.19 where one top-quark decays leptonically via $t \to bW(\to l\nu)$ while the second top decays hadronically as $t \to bW(qq^{'})$. This measurement allows to measure, and thus calibrate, the b-tagging efficiency up to an higher p_T range because of the higher Branching Ratio of the hadronic decay of the W with respect to the leptonic channel:

$$BR(W \to e/\mu + \nu_{e/\mu}) = 21.3\% \qquad BR(W \to qq^{'}) = 67.4\% \tag{3.13}$$

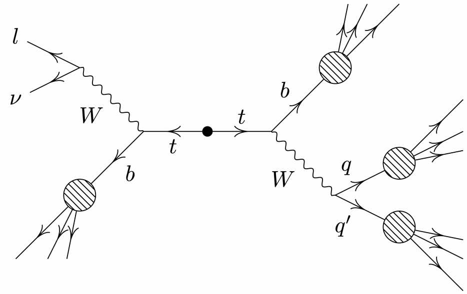


Figure 3.19. Schematic diagram of a semi-leptonic $t\bar{t}$ event at leading order.

The methodology used differs from the standard one [87] used in di-leptonic $t\bar{t}$ events as here a *Tag and Probe* strategy is applied. In particular a jet is tagged associated to the hadronic top (t_{had}) and the efficiency is then measured on the b-jet associated to the leptonic top (t_{lep}).

This measurement is performed using the entire Run2 (2015-2018) dataset at $\sqrt{s} = 13\ TeV$ collected by ATLAS corresponding to an integrated luminosity of $\mathcal{L} = 139\ fb^{-1}$.

The events selected in the analysis are mostly from $t\bar{t}$ production. A smaller contribution is given by W boson production in association with jets, W+jets, or a top-quark production (Wt single-top). Other minor backgrounds are $t\bar{t}$ associated with a vector boson, single-top production in the s- or t-channels and diboson events (WW, WZ, ZZ). The nominal $t\bar{t}$ sample is generated at matrix element level with POWHEGBOX v2 and the parton showering is processed with PYTHIA 8.2. Two main samples are used as source for Monte Carlo modelling uncertainty: one sample uses MADGRAPH5 aMC@NLO v2.6 as alternative Matrix Element generation and same Parton Showering as the nominal sample, while the second alternative sample has the Parton Shower processed with HERWIG 7.04 and same Matrix Element as the nominal sample.

Event reconstruction

A single electron or muon trigger is used to target the charged lepton in the event. Electron candidates are required to pass *tight*-ID criteria and to satisfy $p_T^e > 28$ GeV. Muons are instead reconstructed requiring to pass the *medium*-ID criteria and to have a $p_T^{\mu} > 28$ GeV. An additional track-based isolation criterion is applied to leptons to reduce the background from non-prompt leptons:

$$I_R = \frac{1}{p_T^{(e(\mu))}} \sum_{\Delta R(trk,e(\mu))<\Delta R_e(\mu)} p_T^{trk} < 0.06 \tag{3.14}$$

$$\Delta R_{e(\mu)} = \min\left(0.2(0.3), \frac{10\ \text{GeV}}{p_T^{e(\mu)}}\right) \tag{3.15}$$

For the electron an extra isolation is imposed at calorimeter level requiring that the transverse energy of the topo-cluster in $\Delta R < 0.2$ sums up to less than 6% of the electron transverse energy.

The leptons are also required to be matched to the primary vertex with $|z_0 sin\theta| < 0.5\ mm$ and $S(d_0) < 5$ for electrons and $S(d_0) < 3$ for muons.

Variable Radius Track-jets are then constructed using anti-k_t algorithm. The VR Track-jets are calibrated and corrected for Pile-Up effects. JVT is also applied to jets with $p_T < 60$ GeV to reduce the jets non arising from primary pp collisions. The track-jets are required to have $p_T > 5$ GeV, since they are lacking the reconstruction of the neutral component from the calorimeter information.

Large-R jets are reconstructed from topo-clusters with LCW scheme using anti-k_t algorithm with Radius parameter $R = 1.0$. Large-R jets with $p_T > 100$ GeV and $|\eta| < 2.0$ are considered.

An overlap removal procedure is finally applied to avoid double countings in the objects. Electron candidates within $\Delta R = 0.01$ from a muon candidate are removed. Afterwards the closest jet within $\Delta R = 0.2$ from an electron is discarded. If then the electron is within $\Delta R = 0.4$ of the axis of any other jet, the jet is kept and the electron is removed. Muons satisfying $\Delta R(\mu, jet) < 0.04 + 10$ GeV$/p_T^{\mu}$ are removed if the jet has at least three tracks from the primary vertex otherwise the jet is removed and the muon is kept.

Event Selection

Once the reconstruction is applied a dedicated event selection is applied to build a region of the phase space with high purity in b-jets where we can ensure a good measurement of the b-tagging efficiency.

In order to enhance the purity of the semi-leptonic $t\bar{t}$ sample, each event must contain one lepton with $p_T^l > 70$ GeV. The missing transverse momentum must satisfy $E_T^{miss} > 70$ GeV, which helps in identifying events where significant momentum is carried by the neutrino from the leptonic W. At least one Large-R jet with $p_T > 300$ GeV with a mass $m > 100$ GeV must be present. If more than one such jet exists, the one with the highest p_T is chosen as the t_{had} candidate. Additionally, the event should contain at least two VR Track jets with $p_T > 5$ GeV. At least one of these jets must satisfy a b-tagging selection with a Working Point of $\epsilon_b = 77\%$ using DL1r algorithm. This b-tagged jet must match geometrically the Large-R jet candidate with $\Delta R(b - jet, t_{had}) < 1.0$. In the following this jet will be considered as the *Tag* jet in the *Tag and Probe* method.

To select the *probe* jet the closest to the lepton is chosen and the efficiency will be measured later on this jet. This probe jet is required to be separate on the transverse plane from the t_{had} with $|\Delta\phi(jet, t_{had})| > 1.0$ ensuring a spatial separation between the two top-quarks in the transverse plane.

Calibration strategy

Simulated Monte Carlo samples are separated into signal and backgrounds after the event selection is applied. The signal events are $t\bar{t}$ events with the *probe* jet that is truth b-labelled. The parameter of interest of this analysis is the ϵ_b, which is the probability to tag as a b-jet a truth b-labelled jet. The backgrounds are either non-$t\bar{t}$ events or $t\bar{t}$ events with the *probe*-jet which is not b-labelled. This analysis is performed as a function of the transverse momentum of the *probe*-jet: p_T^{probe}.

The inclusive distribution of the *probe*-jet is shown in Figure 3.20. The standard di-leptonic calibration of VR Track-jets is commonly performed in the range $10 < p_T < 250$ GeV. This calibration allows to extend the calibration up to $p_T = 500$ GeV. In this region the standard calibration lacks information and simulation-based extrapolation methods are normally used to estimate the calibration Scale Factors in this high transverse momentum phase space, at the cost of significant increase of the uncertainty. This calibration instead allows to have a better and precise estimate of the efficiency at high transverse momentum.

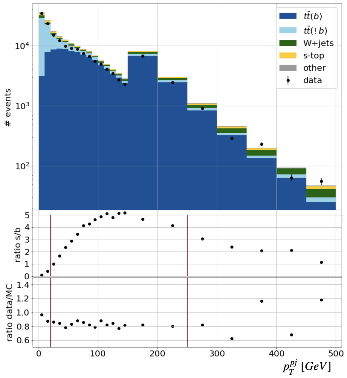


Figure 3.20. Inclusive distribution of the transverse momentum of the *probe*-jet used to measure the b-tagging efficiency. The ratio panels show the signal over background ratio and the data to Monte Carlo ratio before the fit is performed.

The semi-leptonic calibration presented is performed in the range $60 < p_T < 500$ GeV. The bin chosen are arbitrarily selected to be compatible with the di-leptonic calibration for later combination adding an extra bin at high transverse momentum and allowing a finer binning in $100 < p_T < 250$ GeV. A schematic view of the binning of the two calibration is shown in Figure 3.21.

Relevant variables used in the analysis are shown in Figure 3.22.

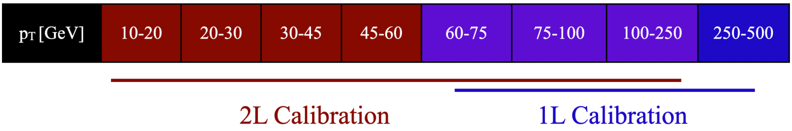


Figure 3.21. Schematic p_T binning of the di-leptonic and semi-leptonic $t\bar{t}$ calibrations of VR Track-jets.

In particular in Figure 3.22(a) the invariant mass of the lepton and the *probe*-jet (m_{lj}) is plotted. This variable, that describes the t_{lep} allows to discriminate the resonant top from the non-resonant backgrounds. The bulk of this distribution mainly contains signal $t\bar{t}$ events while the tail at large m_{lj} is dominated by non-signal $t\bar{t}$ events an W+jets events. Monte Carlo simulations describe reasonably well the data. An expected overall offset is observed due to a known $t\bar{t}$ mis-modelling at high transverse momentum of the top-quark.

As already mentioned the b-tagging efficiency ϵ_b is measured in bins of p_T. For each p_T bin signal and control regions are built. A signal-enriched region is defined requiring $20 < m_{lj} < 130$ GeV to extract ϵ_b. This signal region is further splitted based on the b-tagging result on the *probe*-jet. If the jet is tagged we denote the region as SR^{pass} otherwise SR^{fail}.

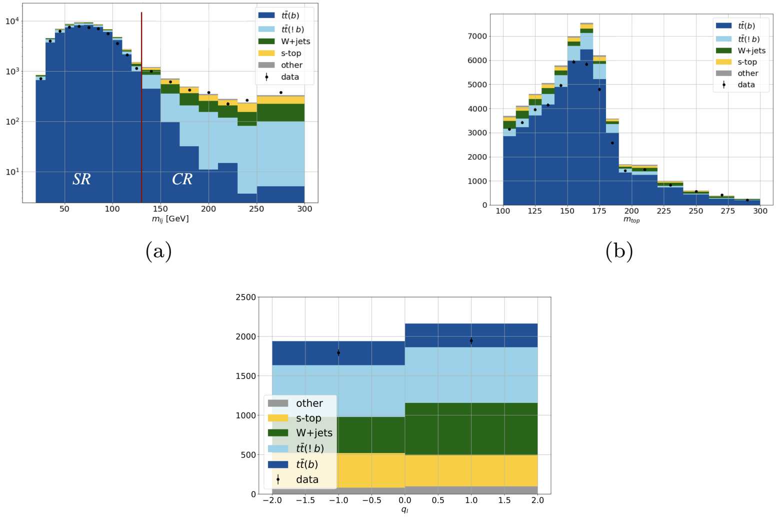
	SR^{pass}	SR^{fail}	CR^{+}	CR^{-}
Selection	$20 < m_{lj} < 130$ GeV		$m_{lj} > 130$ GeV	
Lepton charge	inclusive		+	-
Probe-jet b-tagged?	yes	no	inclusive	
Variable	m_{lj}	m_{lj}	$m_{t_{had}}$	$m_{t_{had}}$

Table 3.4. Scheme of the event categorization.

A background enriched region is defined by $m_{lj} > 130$ GeV. This region is further splitted depending on the charge of the lepton q^{lep} (CR^{+}, CR^{-}) to control the W+jets background exploiting the charge asymmetry as shown in Figure 3.22(c).

In the signal-enriched regions the variable fitted is m_{lj} while in the control-enriched regions the variable fitted is $m_{t_{had}}$. A schematic view of the regions is shown in Table 3.4. For each p_T bin a different binning is used for the fitted variables

For each p_T bin a likelihood $\mathcal{L}(\epsilon_b, \vec{\mu}; p_T)$ is constructed as a product



(a)

(b)

(c)

Figure 3.22. (a) Invariant mass of the lepton and *probe*-jet system (m_{lj}). The red vertical line defines the separation between signal and control regions. (b) Mass of the t_{had} reconstructed from the Large-R jet. (c) Charge distribution of the lepton. This charge asymmetry in *pp* collisions is exploited to control the W+jets background where this effect is caused by the charge asymmetry in proton valence quark composition.

of Poisson over all the analysis bins. In this extent, ϵ_b is the efficiency on data and it is the Parameter of Interest (POI) left floating in the fit. $\vec{\mu} = (NF_{t\bar{t}}, NF_{W+j})$ are the set of two floating normalization factors of the main backgrounds. The systematic treatment will be described later. The efficiency on Monte Carlo is estimated as the ratio of number of signal $t\bar{t}$ events in SR^{pass} over the total number of signal $t\bar{t}$ events in both SR^{pass} and SR^{fail}.

$$\epsilon_b^{MC} = \frac{N_{t\bar{t}(b)}^{SR^{pass}}}{N_{t\bar{t}(b)}^{SR}} \tag{3.16}$$

The Monte Carlo efficiency is then used to scale the two signal regions as:

$$SR^{pass} \rightarrow \epsilon_b \frac{SR^{pass}}{\epsilon_b^{MC}} \qquad SR^{fail} \rightarrow (1-\epsilon_b)\frac{SR^{fail}}{1-\epsilon_b^{MC}} \tag{3.17}$$

In the end the Likelihood can be written as follows:

$$\begin{aligned}
\mathcal{L}(\epsilon_b, \vec{\mu}) = \prod_{i\in[SR,CR]} & \frac{e^{-\lambda_i^R}(\lambda_i^R)^n}{n!} \\
\lambda_i^{SR^{pass}} &= NF_{t\bar{t}}[\epsilon_b s_{t\bar{t}(b)} + s_{t\bar{t}(!b)}] + NF_{W+j} b_{W+j} + b_{others} \\
\lambda_{SR^{fail}} &= NF_{t\bar{t}}[(1-\epsilon_b) s_{t\bar{t}(b)} + s_{t\bar{t}(!b)}] + NF_{W+j} b_{W+j} + b_{others} \\
\lambda_{CR^{+/-}} &= NF_{t\bar{t}} s_{t\bar{t}} + NF_{W+j} b_{W+j} + b_{others}
\end{aligned} \tag{3.18}$$

It worth to mention that ϵ_b will impact only $t\bar{t}(b)$ since we are not evaluating also the mis-tag rate of non b-tagged jets. Finally the fit is performed maximising the Likelihood separately for each p_T and for the four standard Fixed Cut b-tagging Working Points (60%, 70%, 77% and 85%).

An example of post-fit distributions for the 77% WP and for the $75 < p_T < 100$ GeV bin is shown in Figure3.23.

Systematic Uncertainties The uncertainties are classified into two primary categories: modelling uncertainties and detector-related uncertainties.

Process modelling uncertainties account for variations in simulation settings and are assessed by comparing nominal event samples with alternative samples. This category includes the impacts of different parton shower and hadronization models. These systematic uncertainties affect the b-tagging efficiency by approximately $1 - 2.3\%$. The uncertainties related to ISR and FSR range from about 0.1% to 0.8%. The uncertainty related to the single-top modelling is dominated by two main sources. The first refers to the diagram-subtraction scheme (DS) compared with the nominal diagram-removal scheme (DR) resulting in an uncertainty of $1.3 - 2.4\%$. A second uncertainty takes into account a normalization effect varying the nominal normalization of ±20% resulting in an uncertainty of less than 0.5%. For the other processes the modelling impact is found to be negligible.

Detector related uncertainties are associated to the physics object reconstruction. Lepton uncertainties results in an effect of less than 0.3%. The VR Track-jet are considered to be negligible and thus are not considered. The uncertainties associated to the Large-R jet reconstruction are less than 0.7%. All the other uncertainties are found to be negligible.

The effect of each systematic uncertainty is evaluated performing a Likelihood fit for each systematic variation and considering the relative difference with respect to the nominal fit:

$$\delta^{syst.}\epsilon_b/\epsilon_b = \frac{\epsilon_b^{syst.} - \epsilon_b^{nom.}}{\epsilon_b^{nom.}} \tag{3.19}$$

The breakdown of the uncertainties associated to the ϵ_b is shown in Table 3.5 for the 77% b-tagging Working Point.

p_T [GeV]	**60-75**	**75-100**	**100-150**	**150-200**	**200-250**	**250-500**
Total unc.	3.7%	2.9%	2.2%	2.6%	4.2%	5.0%
Stat unc.	1.1%	0.9%	0.9%	1.4%	2.6%	3.8%
Syst. unc.	3.5%	2.8%	2.0%	2.2%	3.3%	3.2%
$t\bar{t}$ modelling	3.2%	2.2%	1.2%	1.1%	2.8%	2.0%
$t\bar{t}$ ME (Powheg vs aMC@NLO)	2.5%	1.3%	0.8%	0.9%	1.7%	1.4%
$t\bar{t}$ PS (Pythia8 vs Herwig7)	1.7%	1.5%	0.9%	0.2%	2.1%	0.6%
$t\bar{t}$ ISR var.	0.4%	0.4%	<0.1%	0.5%	0.4%	0.8%
$t\bar{t}$ FSR var.	0.8%	0.4%	0.2%	0.1%	0.3%	0.7%
$t\bar{t}$ scale var.	0.8%	0.4%	<0.1%	0.3%	0.4%	0.7%
single-top modelling	1.4%	1.7%	1.5%	1.9%	1.7%	2.5%
Wt DR vs DS	1.3%	1.6%	1.5%	1.8%	1.6%	2.4%
Normalization	0.3%	0.4%	0.4%	0.4%	0.3%	0.5%
V+jets scale variations	0.2%	0.2%	0.2%	0.2%	0.4%	0.4%
Large-R jet uncertainties	0.3%	0.2%	0.1%	0.2%	0.7%	0.2%
Large-R JMS/JMR	0.2%	0.1%	<0.1%	0.2%	0.7%	0.2%
Large-R JES (Modelling)	<0.1%	<0.1%	<0.1%	<0.1%	0.1%	<0.1%
Large-R JES (Statistics)	<0.1%	<0.1%	<0.1%	<0.1%	<0.1%	<0.1%
Large-R JES (Mistag)	<0.1%	<0.1%	<0.1%	<0.1%	<0.1%	<0.1%
Large-R JES (Detector)	<0.1%	-	-	<0.1%	<0.1%	<0.1%
Flavour tagging uncertainties	<0.1%	<0.1%	<0.1%	<0.1%	0.1%	<0.1%
light-mistag	<0.1%	<0.1%	<0.1%	<0.1%	<0.1%	<0.1%
c-mistag	<0.1%	<0.1%	<0.1%	<0.1%	<0.1%	<0.1%
FTag extrapolations	-	-	<0.1%	<0.1%	<0.1%	-
Lepton uncertainties	0.1%	<0.1%	<0.1%	<0.1%	0.2%	0.3%
p_T^{μ} calibration	<0.1%	<0.1%	<0.1%	<0.1%	<0.1%	0.1%
ϵ_{el} corrections	<0.1%	<0.1%	<0.1%	<0.1%	<0.1%	0.1%
ϵ_{μ} corrections	0.1%	<0.1%	<0.1%	<0.1%	0.1%	0.3%
E_{el} calibration	<0.1%	-	<0.1%	<0.1%	<0.1%	<0.1%
Others	0.1%	<0.1%	0.1%	0.1%	0.2%	0.3%
pileup	<0.1%	<0.1%	<0.1%	<0.1%	<0.1%	<0.1%
JVT	<0.1%	<0.1%	<0.1%	<0.1%	<0.1%	<0.1%
flavour comp./resp.	<0.1%	<0.1%	<0.1%	<0.1%	0.1%	<0.1%
η inter-calibration	<0.1%	<0.1%	<0.1%	<0.1%	<0.1%	<0.1%
High p_T	-	-	-	-	-	-

Table 3.5. Summary of uncertainties contribution impacting the measurement of the b-tagging efficiency ϵ_b for each p_T bin. The results are for the 77% WP. Each category is the result of the sum in quadrature of the individual sub-components.

The results are then shown for each b-tagging working point in Figure 3.24 and 3.25 where the measured b-tagging efficiency ϵ_b and the Scale Factors ($\epsilon_b/\epsilon_b^{MC}$) as a function are respectively shown. This analysis leads to a precision on VR Track-jets b-tagging efficiency up to 5% in a transverse momentum phase space ($250 < p_T < 500$ GeV) where the standard extrapolation methods usually have an uncertainty of more than 13%, improving thus the calibration at high p_T of almost a factor of 3.

3.2.2 Statistical combination of different calibrations

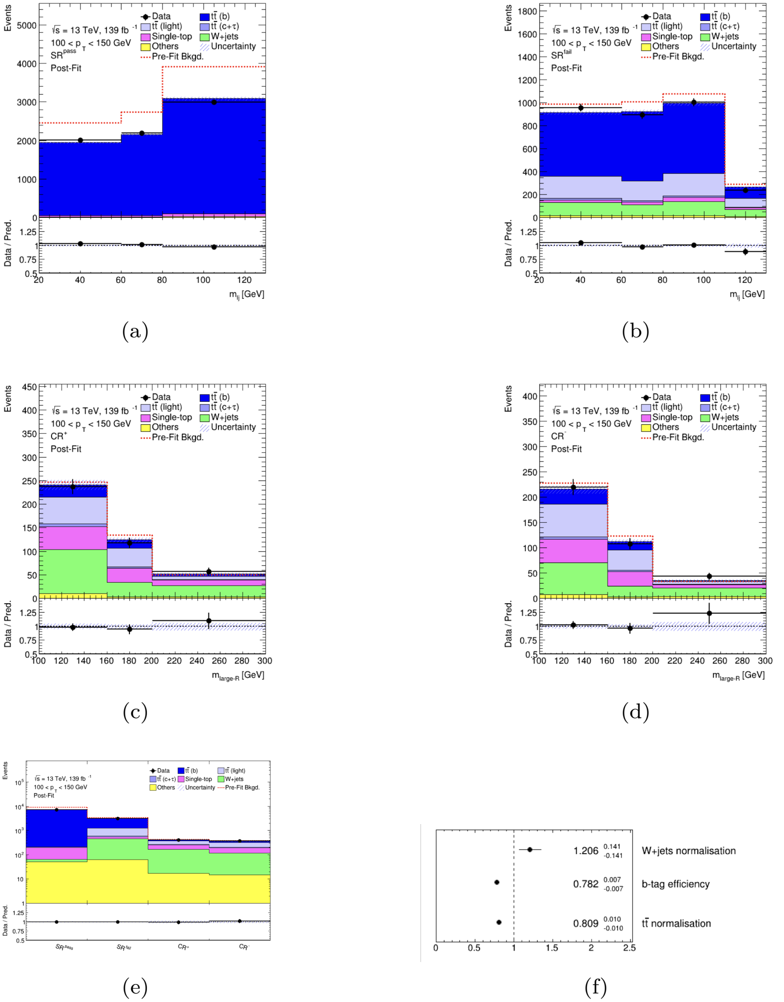
In this brief section the statistical combination of different and orthogonal b-tagging calibration is shown. In particular the standard di-leptonic $t\bar{t}$ calibration can be combined with the semi-leptonic calibration described in the previous section allowing to have the best measurement of the b-tagging efficiency on the widest range as possible. In the following I will describe the general approach for the combination and its application to two different kind of jets: PFlow jets and VR Track-jets.

To perform the combination a Likelihood fit is performed. Some assumptions are made in this methods. Firstly, each b-tagging measurement is considered with a Gaussian Likelihood. Each systematic is associated with a Gaussian prior and is considered to contribute linearly to the combined value, thus second order or higher contributions are neglected. Finally each systematic is considered as 100% correlated in p_T and if it contributes to more than one measurement it is also considered as 100% correlated between the measurements. It is thus possible to write the Likelihood function as the product in each p_T bin and each b-tagging calibration of the Gaussians centered at the measured Scale Factor ($\hat{SF}$) with a width given by the statistical uncertainty on the measurement $\delta SF^{stat.}$. The likelihood is then summarised in the Equation 3.20 where each systematic is contributing to the mean linearly with a fitted with parameter $\theta^{syst.}$ with a Gaussian prior centered in 0 and sigma 1.

$$
\begin{aligned}
\mathcal{L} = \prod_{p_T}^{bins} \prod_{i=1L,2L}^{calib} \mathcal{G}\left[SF(p_T)|\mu,\sigma\right] \prod_{s}^{syst} \mathcal{G}[\theta_{is}|0,1] \\
\mu = \hat{SF}_i(p_T)\left(1 + \sum_{s}^{syst} \frac{\delta SF_{is}}{\hat{SF}_i}(p_T)\theta_{is}\right) \\
\sigma = \delta SF_i^{stat}(p_T)
\end{aligned}
\tag{3.20}
$$

The result of the combination is shown in Figure 3.27 and 3.28 for the PFlow jets and VR Track-jets respectively. The p_T binning used for the PFlow combination is shown in Figure 3.26 while for the VR Track-jets it is shown in Figure 3.21.

From Figures 3.27 3.28 it is possible to observe how the combination not only allows to calibrate jets to a higher transverse momentum but it also improves the uncertainty in those p_T bins where the calibration is performed by both methods.

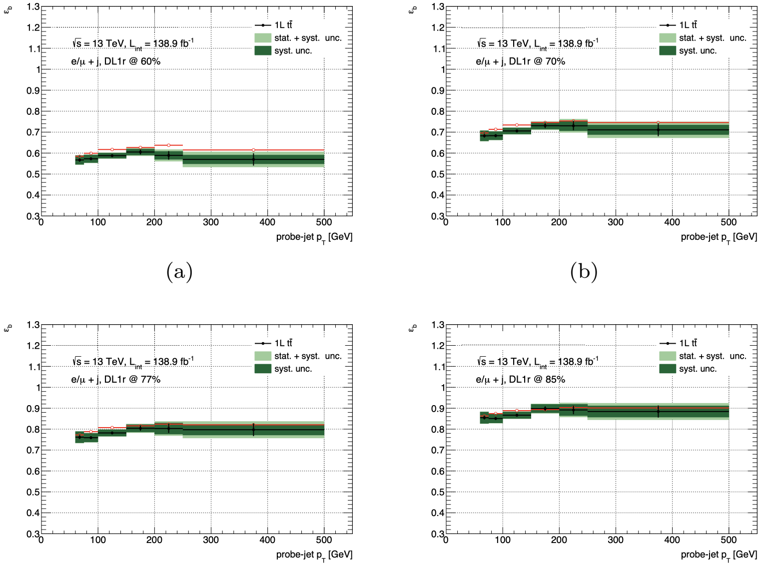


(a) (b)

(c) (d)

(e) (f)

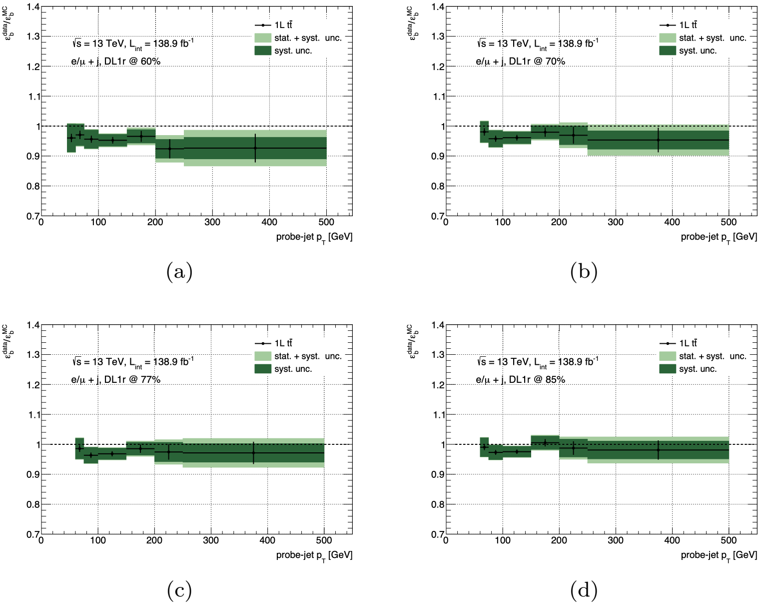
Figure 3.23. Post-fit distributions of the analysis regions for the 77% b-tagging WP in the bin $100 < p_T < 150$ GeV. (a) Signal Region passing the b-tagging : SR^{pass}. (b) Signal Region failing the b-tagging : SR^{fail}. (c) Control Region with positive lepton charge: CR^{+} (d) Control Region with negative lepton charge: CR^{-} (e) Inclusive bin distribution for each region. The pre-fit distribution is shown with a dashed red line. (f) Floating parameters in the fit: Normalization factors for the signal ($t\bar{t}$) and the main background (W+jets) and the b-tagging efficiency ϵ_b.



(a) (b)

(c) (d)

Figure 3.24. Measurement of the *b*-tagging efficiency as a function of the p_T of the *probe*-jet for the four standard Working Points (a) 60%, (b) 70%, (c) 77% and (d) 85%. The dark-green error band is associated to the overall systematic uncertainty while the light-green shows the systematic uncertainty summed in quadrature with the statistical error displayed by the black error band. The red reference line describes the reference Monte Carlo *b*-tagging efficiency ϵ_b.



(a) (b)

(c) (d)

Figure 3.25. Measurement of the b-tagging Scale Factors as a function of the p_T of the *probe*-jet for the four standard Working Points (a) 60%, (b) 70%, (c) 77% and (d) 85%. The dark-green error band is associated to the overall systematic uncertainty while the light-green shows the systematic uncertainty summed in quadrature with the statistical error displayed by the black error band.

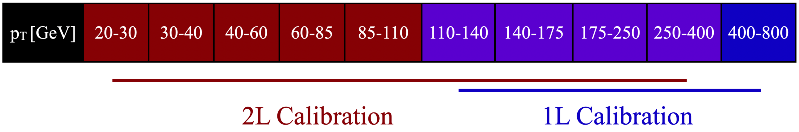
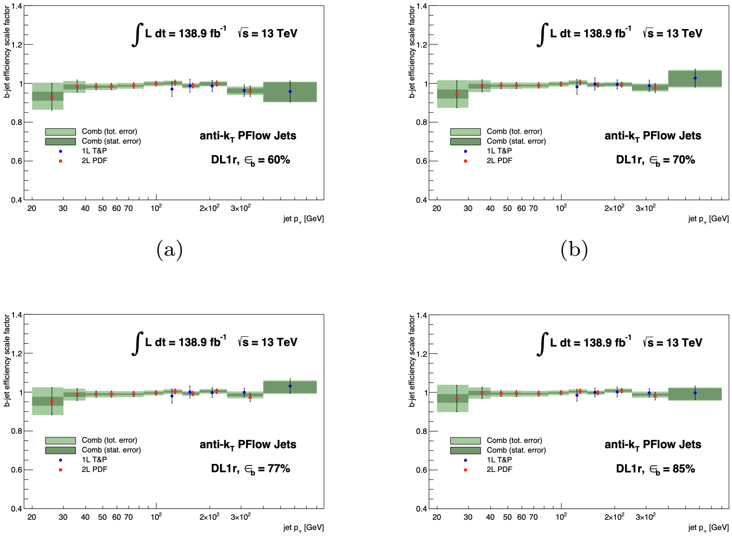


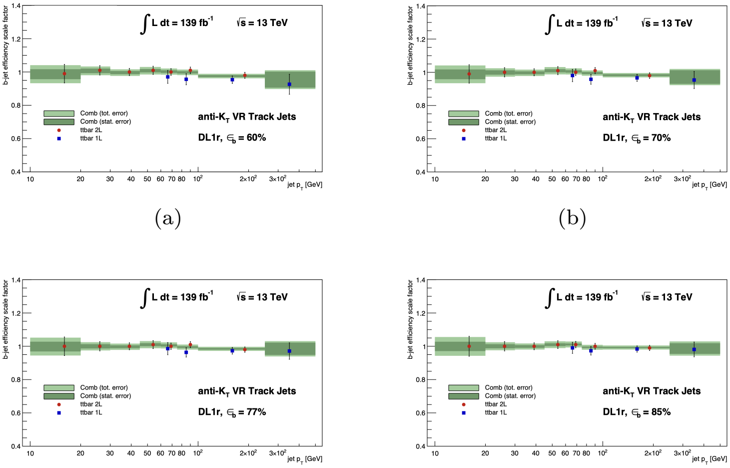
Figure 3.26. Schematic p_T binning of the di-leptonic and semi-leptonic $t\bar{t}$ calibrations of PFlow jets.



(a) (b)

(c) (d)

Figure 3.27. Combination of b-tagging Scale Factors as a function of p_T of the jet. The combination is performed on the PFlow jets and the calibration considered are the di-leptonic and semi-leptonic calibration on $t\bar{t}$ events. The Combination is performed separately for each b-tagging Working Point: (a) 60%, (b) 70%, (c) 77% and (d) 85%.



(a) (b)

(c) (d)

Figure 3.28. Combination of b-tagging Scale Factors as a function of p_T of the jet. The combination is performed on the VR Track-jets and the calibration considered are the di-leptonic and semi-leptonic calibration on $t\bar{t}$ events. The Combination is performed separately for each b-tagging Working Point: (a) 60%, (b) 70%, (c) 77% and (d) 85%.

4 | The Higgs-strahlung: A Legacy Analysis

In Section 1.2, the electroweak spontaneous symmetry breaking mechanism, which generates the masses of the Standard Model particles, was described. This mechanism introduces a new scalar massive particle: the Higgs boson. It is well known that in 2012, the Higgs boson was discovered with a mass of approximately 125 GeV by the ATLAS Ref. [91] and CMS Ref. [92] collaborations, a value that was not predicted beforehand.

This chapter begins with a description of Higgs physics from an experimental point of view (Section 4.1), highlighting the main results of the ATLAS collaboration. Following this, the physics analysis of the Higgs production in association with a leptonically decaying vector boson, with the Higgs decaying into b- or c-quarks, is briefly discussed (Section 4.2): $VH(bb|cc)$, where V represents either W or Z.

The primary focus of this chapter, and the dissertation as a whole, is the Run 2 Legacy analysis performed to simultaneously extract the Higgs coupling to b- and c-quarks (Section 4.3). This measurement leverages advanced techniques to improve upon previous results, resulting in the most precise measurement of the Yukawa coupling to the b-quark to date. It also places the most stringent upper limits on the signal strength of the Yukawa coupling to the c-quark. Additionally, this measurement aims to provide the first observation of the $WH(bb)$ Higgs process and the $VZ(cc)$ diboson process by the ATLAS collaboration.

4.1 The Higgs Boson

The presence of a Higgs boson in a physics event proceeds through two different steps: the production processes and the decay modes.

The cross sections for the production of a Standard Model Higgs boson as a function of $\sqrt{s}$, for pp collisions, are summarized in Figure 4.1 (left) with relative theoretical uncertainties Ref. [93, 94].

The production of the Higgs boson at the LHC can occur through several channels. The main channels, in terms of cross-section, are gluon fusion (ggF), weak-boson fusion (VBF), Higgs-strahlung (VH), and associated production with a pair of quarks (mostly $t\bar{t}$). In Figure 4.2, the main Feynman diagrams contributing to these processes are shown. After the measurement of the Higgs mass, it has been possible to infer its production and decay rates. They were found to be compatible with the SM predictions. The Higgs couples to the mass of the particles through the Vacuum Expectation Value (VEV, $v \approx 246$ GeV).

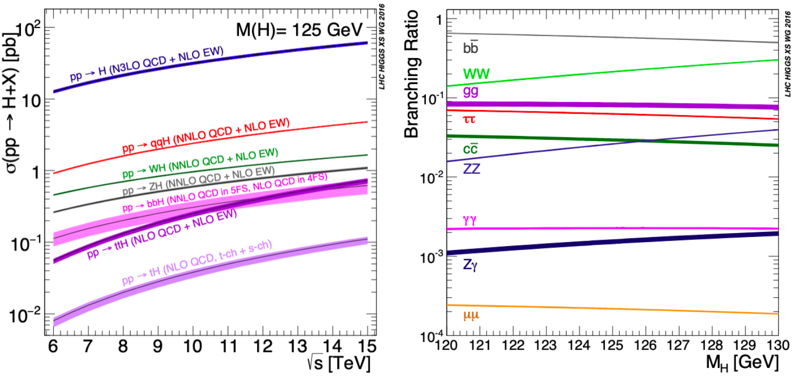


Figure 4.1. (**left**) The SM Higgs boson production cross sections as a function of the center of mass energy, $\sqrt{s}$, for pp collisions. The VBF process is indicated here as qqH. The theoretical uncertainties are indicated as bands. (**right**) The branching ratios for the main decays of the SM Higgs boson near $m_H = 125$ GeV. From Ref. [93]

In particular, the SM Higgs boson couplings to vector bosons (V) and fermions (f), as well as the Higgs boson self-coupling (this will be discussed in detail in Section 5), at tree level are:

$$\begin{aligned} g_{HVV} &= \frac{2m_V^2}{v}\delta_V \quad g_{HHVV} = \frac{2m_V^2}{v^2}\delta_V \quad \delta_Z = \frac{1}{2};\ \delta_W = 1 \\ g_{Hf\bar{f}} &= \frac{\sqrt{2}m_f}{v} \quad g_{HHH} = \frac{m_H^2}{2v} \quad g_{HHHH} = \frac{m_H^2}{2v^2} \end{aligned} \tag{4.1}$$

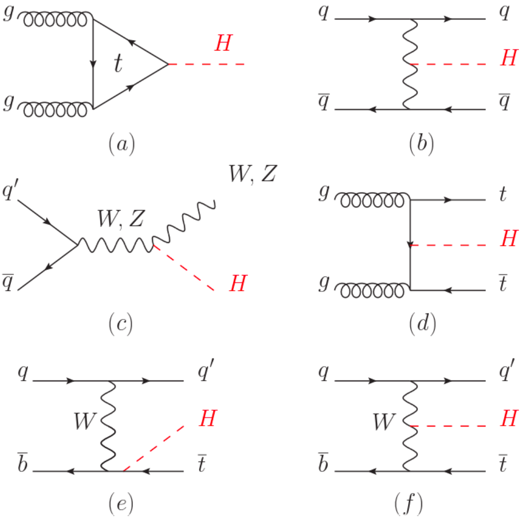
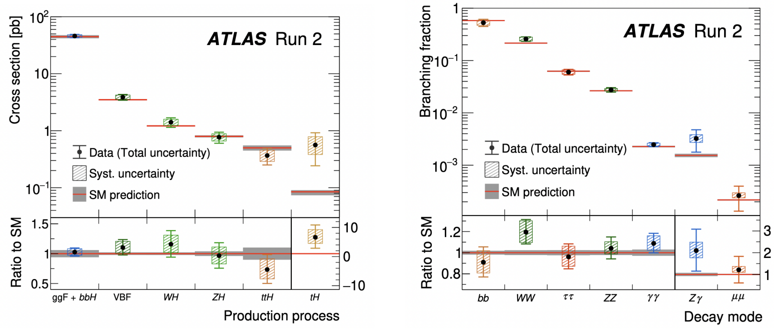


Figure 4.2. Main Leading Order Feynman diagrams contributing to the Higgs production in (a) gluon fusion, (b) Vector-boson fusion, (c) Higgs-strahlung (or associated production with a gauge boson), (d) associated production with a pair of top (or bottom) quarks, (e-f) production in association with a single top quark. with top quarks. From Ref. [93]

At high-energy hadron colliders, the Higgs boson production mechanism with the largest cross section is the gluon-fusion process (ggF) mediated by the exchange of a virtual, heavy top quark. The SM Higgs production mode with the second-largest cross section at the LHC is vector boson fusion (VBF). The scattered quarks give rise to two hard jets in the forward and backward regions of the detector. The peculiar characteristics of VBF processes can be exploited to distinguish them from overwhelming QCD backgrounds. After the application of specific selection cuts, the VBF channel provides a clean environment, not only for Higgs searches but also for the determination of Higgs boson couplings at the LHC.

After ggF and VBF, the most important Higgs production mechanism at the LHC, and relevant for this dissertation, is the associated production with W or Z vector bosons (Higgs-strahlung). As neither the Higgs boson nor the vector bosons are stable particles, their decays also have to be taken into account. The leptonic decay of the associated vector boson provides a clear signature in the event at the LHC, drastically reducing the QCD background and allowing probing of the Higgs decays despite the lower cross-section.

Because of the high mass of the top-quark ($m_t \simeq 173$ GeV) and the couplings shown in Eq. 4.1, the Higgs cannot decay into a $t\bar{t}$ pair. Consequently, the most probable decay is into the heaviest accessible quark, the b-quark with a mass of $m_b \simeq 4.2$ GeV: $H \rightarrow b\bar{b}$. All the branching ratios are shown in Figure 4.1 (right), in particular, $BR(H \rightarrow b\bar{b}) = 58.4\%$, assuming a SM Higgs boson with a mass of 125 GeV.



(a) (b)

Figure 4.3. Observed and predicted Higgs boson production cross-sections and branching fractions. (a) Cross sections for different Higgs boson production processes are measured assuming SM values for the decay branching fractions. (b) Branching fractions for different Higgs boson decay modes are measured assuming SM values for the production cross sections. The lower panels show the ratios of the measured values to their SM predictions. The vertical bar on each point denotes the 68% confidence interval. The p-value for compatibility of the measurement and the SM prediction is 65% for (a) and 56% for (b). [95]

A summary of the measurements by the ATLAS collaboration of the Higgs boson production and decays with the Run 2 dataset is shown in Figure 4.3 and reported in Ref. [95], published ten years after the discovery of the Higgs boson. All measurement results are compatible with the Standard Model (SM) predictions.

The Higgs boson production rates are probed by the likelihood fit to observed signal yields described earlier. As the production cross section σ_i and the branching fraction B_f for a specific production process i and decay mode f cannot be measured separately without further assumptions, the observed signal yield for a given process is expressed in terms of a single signal-strength modifier:

$$\mu_{i,f} = \frac{\sigma_i B_f}{\sigma_i^{SM} B_f^{SM}} \tag{4.2}$$

where the superscript ‘SM’ denotes the corresponding SM prediction.

To determine the value of a particular Higgs boson coupling strength, a simultaneous fit of many individual production times branching fraction measurements is required. The coupling fit presented in Ref. [95] is performed within the k-framework Ref. [96] with a set of parameters k that affect the Higgs boson coupling strengths without altering any kinematic distributions of a given process.

Within this framework, the cross section times the branching fraction for an individual measurement is parameterized in terms of the multiplicative coupling strength modifiers k. A coupling strength modifier k_p for a production or decay process via the coupling to a given particle p is defined as $k_p^2 = \sigma_p/\sigma_p^{SM}$ or $k_p^2 = \mathit{\Gamma}_p/\mathit{\Gamma}_p^{SM}$, respectively, where $\mathit{\Gamma}$ is the partial decay width into a pair of particles p. The parameterization takes into account that the total decay width depends on all decay modes included in the present measurements, as well as currently undetected or invisible, direct or indirect decays predicted by the SM (such as those to gluons, light quarks, or neutrinos) and hypothetical decays into non-SM particles.

Let’s consider a class of models where the coupling strength modifiers for W, Z, t, b, τ, μ, and c are treated independently. All modifiers are assumed to be positive. It is assumed that only SM particles contribute to the loop-induced processes, and modifications of the fermion and vector-boson couplings are propagated through the loop calculations. Invisible or undetected non-SM Higgs boson decays are not considered. These models allow us to test the predicted scaling of the couplings of the Higgs boson to the SM particles as a function of their mass using the reduced coupling strength modifiers.

Figure 4.4 shows the results for two scenarios: one with the coupling to c-quarks constrained by $k_c = k_t$ in order to cope with the low sensitivity to this coupling, and the other with k_c left as a free parameter in the fit. All measured coupling strength modifiers are found to be compatible with their SM prediction.

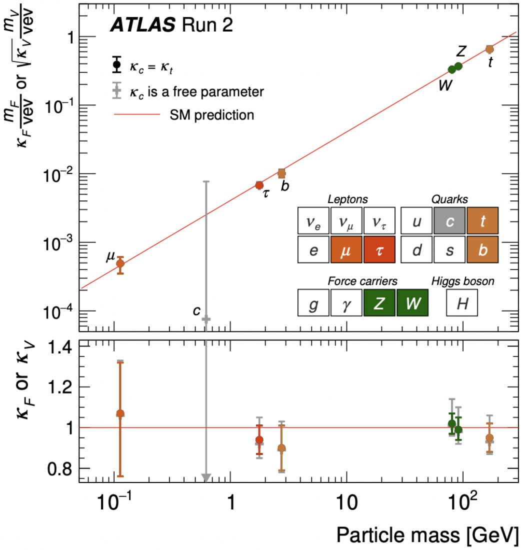


Figure 4.4. Reduced Higgs boson coupling strength modifiers and their uncertainties. They are defined as $k_F m_F / v$ for fermions ($F = t$, b, τ, μ) and $\sqrt{k_V} m_V / v$ for vector bosons as a function of their masses m_F and m_V. Two fit scenarios with $k_c = k_t$ (coloured circle markers), or k_c left free-floating in the fit (grey cross markers) are shown. Loop-induced processes are assumed to have the SM structure, and Higgs boson decays to non-SM particles are not allowed. The vertical bar on each point denotes the 68% confidence interval. The p-values for compatibility of the combined measurement and the SM prediction are 56% and 65% for the respective scenarios. The lower panel shows the values of the coupling strength modifiers. The grey arrow points in the direction of the best-fit value and the corresponding grey uncertainty bar extends beyond the lower panel range. Ref. [95]

It is relevant to observe that there is no experimental evidence of the Higgs boson coupling to the first generation of fermions or the second generation of quarks. Before the LHC was built, it was thought to be impossible to measure the Higgs coupling to c-quarks at the LHC due to the branching fraction being about 20 times smaller than that for b-quarks ($B_c(H \rightarrow c\bar{c}) = 2.89\%$ in Ref. [97]). However, recent studies have shown that with the improvement obtained by novel techniques, this may no longer be the case, and evidence could potentially be obtained by the end of the LHC's operation.

For this reason, measuring and analyzing the coupling of the Higgs to heavy flavor quarks plays a pivotal role in our understanding of the Standard Model.

4.2 The Run2 VH Analyses

As mentioned, the Higgs decay to c-quarks shares a similar topology with the decay to b-quarks. The crucial differences that make measuring the coupling to the second generation of quarks challenging are the approximately 20 times lower branching ratio and the complexity of Flavor Tagging, described in Chapter 3, for tagging c-quarks. The benchmark analyses for the $H \rightarrow b\bar{b}$ Ref. [98] and $H \rightarrow c\bar{c}$ Ref. [99] used the full Run 2 dataset. These analyses will be briefly described here, to provide a benchmark for the "*legacy*" analysis that is the main topic of this thesis work.

As already mentioned, the $H \rightarrow b\bar{b}$ and $H \rightarrow c\bar{c}$ decays are challenging to measure in the ggF and VBF production modes due to the large background from multi-jet events. Therefore, the most sensitive production channel is the associated production of a Higgs boson with a Vector boson W or Z, briefly referred to as VH. The leptonic decay of the Vector bosons significantly suppresses the multi-jet background.

The $H \rightarrow b\bar{b}$ decay mode was first observed by ATLAS and CMS Ref. [100, 101], and since then, the focus has shifted towards precision measurements. ATLAS has published three additional results using the full Run 2 dataset of 139 fb^{-1}. The first [102], similar to the observation analysis, focuses on Higgs bosons produced at low transverse momentum with the two b-quarks resolved into two distinct jets. The second [103] targets the Higgs boson when it is highly boosted and reconstructed with a single large-radius jet. The third [98], described in Section 4.2, combines the two approaches with a careful study of the overlap region.

The $H \rightarrow c\bar{c}$ decay mode is much harder to measure, and direct searches have set upper limits on the cross-section times the branching fraction of the process [99]. In Section 4.2, the latest measurement of this process is shown, setting an upper limit of 26 times the signal strength predicted by the Standard Model for the $VH(cc)$ process. The CMS collaboration measured an expected (observed) upper limit of 7.6 (14.4) times the Standard Model prediction Ref. [104].

The $H \to b\bar{b}$ Analysis

The $H \to b\bar{b}$ analysis Ref. [98] has been performed with the full Run 2 dataset, as a function of the vector boson transverse momentum, p_T^V, which is related to the transverse momentum of the Higgs boson, in kinematic fiducial volumes as defined within the Simplified Template Cross-Section (STXS) framework [105].

Two orthogonal regimes are defined based on p_T^V. At low p_T^V, the two b-quarks will generate a pair of jets that can be reconstructed separately. In this region, the Higgs candidate is reconstructed with two distinct b-tagged jets using the anti-k_t algorithm with a radius parameter of 0.4. This regime will be referred to as *resolved* in the following discussion.

At higher p_T^V, the b-quarks are more collimated due to the Higgs boost:

$$\Delta R(b_1, b_2) \simeq \frac{2m_H}{p_T^H} \tag{4.3}$$

Thus, the Higgs candidate is reconstructed with a single large-radius jet ($R = 1.0$) with substructure containing two VR Track-jets. This regime will be referred to as *boosted* in the following discussion. The switch between the two regimes occurs at $p_T^V = 400$ GeV, ensuring consistency with the STXS scheme. In the end, the signal is extracted in STXS $p_T^{V,t}$, the truth transverse momentum of the vector boson, categories as shown in Figure 4.5.

It is possible to extract the signal strength for the WH and ZH processes (μ_{WH}^{bb}, μ_{ZH}^{bb}), validating the analysis on the diboson process VZ (μ_{VZ}^{bb}).

The combination yields an inclusive signal strength for the $VH(bb)$ signal of:

$$\mu_{VH}^{bb} = 1.00^{+0.12}_{-0.11}(\text{stat.})^{+0.14}_{-0.13}(\text{syst.})$$

corresponding to an expected (observed) significance of 6.3 (6.4) standard deviations over the background-only hypothesis.

$$\mu_{WH}^{bb} = 1.03^{+0.19}_{-0.19}(\text{stat.})^{+0.21}_{-0.19}(\text{syst.}) \tag{4.4}$$

$$\mu_{ZH}^{bb} = 0.97^{+0.17}_{-0.17}(\text{stat.})^{+0.18}_{-0.15}(\text{syst.}) \tag{4.5}$$

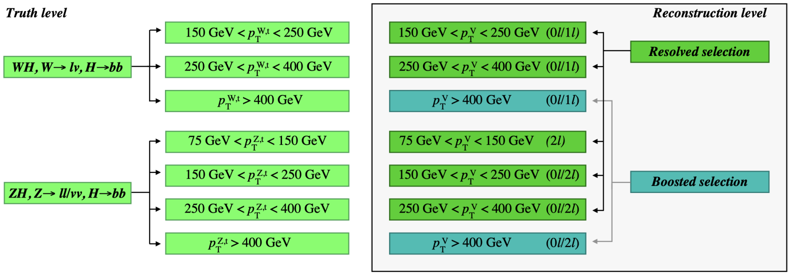


Figure 4.5. Sketch of the ZH and WH STXS p_T^V, t categories for which cross-sections are extracted in the combination and the corresponding categories of reconstructed p_T^V from Ref. [105].

resulting in 3.9 (4.1) and 5.0 (4.6) expected (observed) standard deviations from the background-only hypothesis for the WH and ZH production modes, respectively.

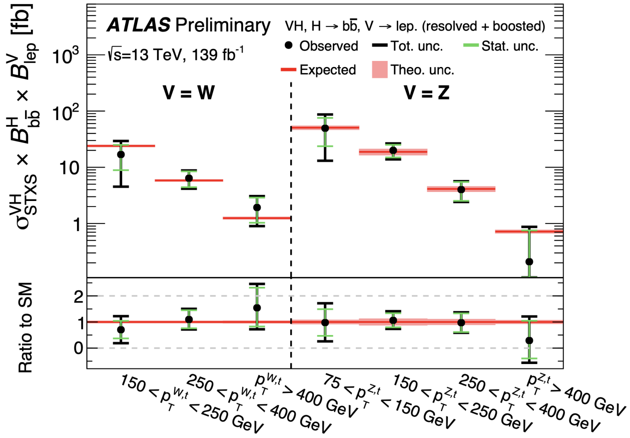


Figure 4.6. Measured VH cross-sections times the $H \to b\bar{b}$ branching fraction in the STXS scheme Ref. [98].

The $H \to c\bar{c}$ Analysis

Given the similar topology, the $H \to c\bar{c}$ analysis Ref. [99] has been performed with a similar strategy to the $H \to b\bar{b}$ analysis. Due to the low branching fraction of the $H \to c\bar{c}$ process, about 2.89%, the coupling to the second generation of quarks has yet to be observed at the LHC. This analysis uses the *resolved* scheme for the entire p_T^V spectrum, and a binned profiled likelihood fit to the m_{cc}, the invariant mass of the two c-tagged jets, distribution is performed. The fitted signal strength is:

$$\mu_{VH}^{cc} = -9 \pm 10(\text{stat.}) \pm 12(\text{syst.})$$

For the μ_{VH}^{cc} signal strength, an upper limit of $31(26)^{+12}_{-8}$ is expected (observed) at the 95% Confidence Level. The limits for each lepton channel and the combinations are summarized in Figure 4.7.

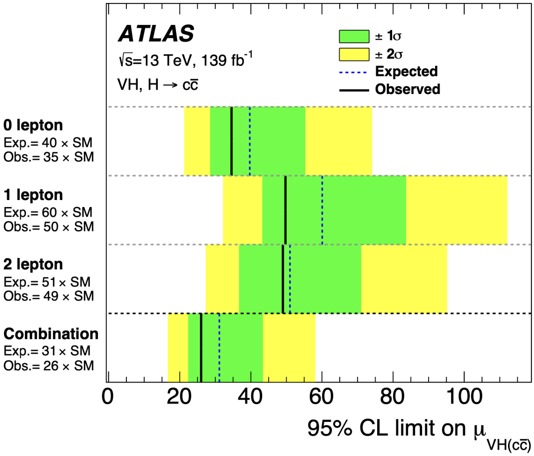


Figure 4.7. The observed and expected 95% CL upper limits on the cross-section times the branching fraction normalized to its SM prediction in each lepton channel and for the combined fit. The single-channel limits are obtained using a 5-POI fit, in which each channel has separated $VH(cc)$ parameters of interest from Ref. [99].

This analysis is also validated by studying the diboson production $VZ(c\bar{c})$ and $VW(cq)$, with expected (observed) significances of 2.2 (2.6) and 4.6 (3.8), respectively.

4.3 VH(bb|cc) Legacy Analysis

Since Ref. [98, 99], many improvements in terms of techniques have been achieved by the ATLAS collaboration. For this reason, a re-analysis of the Run 2 data has been performed. The goal is to extract the best sensitivity to the $H \to b\bar{b}$ and $H \to c\bar{c}$ decays through a combined analysis using an improved version of the b-tagging algorithm, implementing a Multi-Variate Analysis (MVA) in all the signal-enhanced regions of the analysis, and correlating the systematic uncertainties among the regions to maximize the power of the data to extract the Parameters Of Interest (POI).

Due to the similar topology of the $H \to b\bar{b}$ and $H \to c\bar{c}$ decays, there is a significant benefit in performing the analysis simultaneously. Additionally, the backgrounds of the two analyses can be treated coherently.

The analysis strategy is schematized in Figure 4.8 across the three different analysis regimes: $H \to b\bar{b}$ *resolved*, $H \to b\bar{b}$ *boosted*, and $H \to c\bar{c}$.

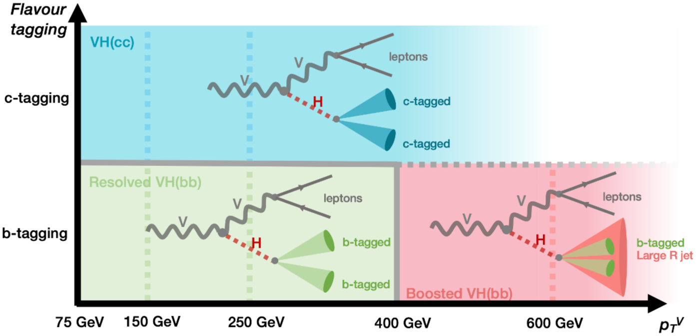


Figure 4.8. Illustration of the analysis regimes considered in the $VH(b\bar{b}|c\bar{c})$ analysis. The $H \to b\bar{b}$ and $H \to c\bar{c}$ regimes are separated through the requirement of b- and c-tags, and the boosted and resolved regimes are separated with a cut at $p_T^V = 400$ GeV requirements are not fulfilled. Note that jets are considered for c-tagging only if b-tagging requirements are not fulfilled.

The analysis targets $VH(c\bar{c})$ and $VH(b\bar{b})$ as signal processes. Additionally, a cross-check analysis is considered with the Standard Model diboson processes $VZ(c\bar{c})$ and $VZ(b\bar{b})$ as signal processes.

The main backgrounds in both $H \to b\bar{b}$ and $H \to c\bar{c}$ cases come from V+jets and Top quark processes (mainly $t\bar{t}$). The V+jets background is split into three different components: V+hf (bb, cc), V+mf (bc, bl, cl), and V+lf. The top quark background is also split by di-jet flavor, with Top(bb)[1] treated separately as the main background for the $H \to b\bar{b}$ case. Minor backgrounds include single-top processes, multi-jet, and non-signal diboson processes such as WW.

This analysis strongly relies on Flavor Tagging to target the $H \to b\bar{b}$ and $H \to c\bar{c}$ final states. The DL1r algorithm, described in Section 3.1.1, is used for this purpose. A Pseudo-Continuous Flavor Tagging scheme is implemented to achieve an exclusive categorization in terms of Flavor Tagging for each jet. A two-dimensional plane of b-tagging vs c-tagging scores divides the plane into five orthogonal regions (two b-tagging , one c-tagging tight, one c-tagging loose, and a non-tagged region), as shown in Figure 4.9.

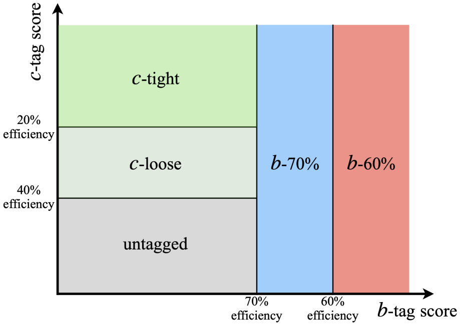


Figure 4.9. DL1r Pseudo Continuous Flavour Tagging algorithm.

The analysis is split into three channels depending on the leptonic decay of the vector boson: 0-*lepton* for $Z \to \nu\bar{\nu}$, 1-*lepton* for $W \to l\nu_l$, and 2-*lepton* for $Z \to l^+l^-$. Each channel is further divided into different regions that enter the statistical analysis, as shown in the scheme in Figure 4.10. Three event displays are shown in Figure 4.12 for each lepton channel.

In the $H \to b\bar{b}$ *resolved* regime, exactly two b-tagged jets are required, and the regions are defined in bins of p_T^V : low-p_T^V $[75, 150]$ GeV (1-/2-*lepton* only), medium-p_T^V $[150, 250]$ GeV, and high-p_T^V $[250, 400]$ GeV. Events are further separated into 2 and 3 jets categories; in the 0-*lepton* channel, an exactly 4-jets category is added, while in the 2-*lepton* channel, 4 or more jets are also included. A control region based on the angular distance ($\Delta R(b1, b2)$) of the two Higgs candidate jets is created. The control region with large $\Delta R(b1, b2)$ is denoted as CRHigh.

---

[1] We refer here and the future to Top as $t\bar{t}$ $+Wt$

For the 1-*lepton* channel, a control region with low $\Delta R(b1, b2)$ is added, denoted as CRLow.

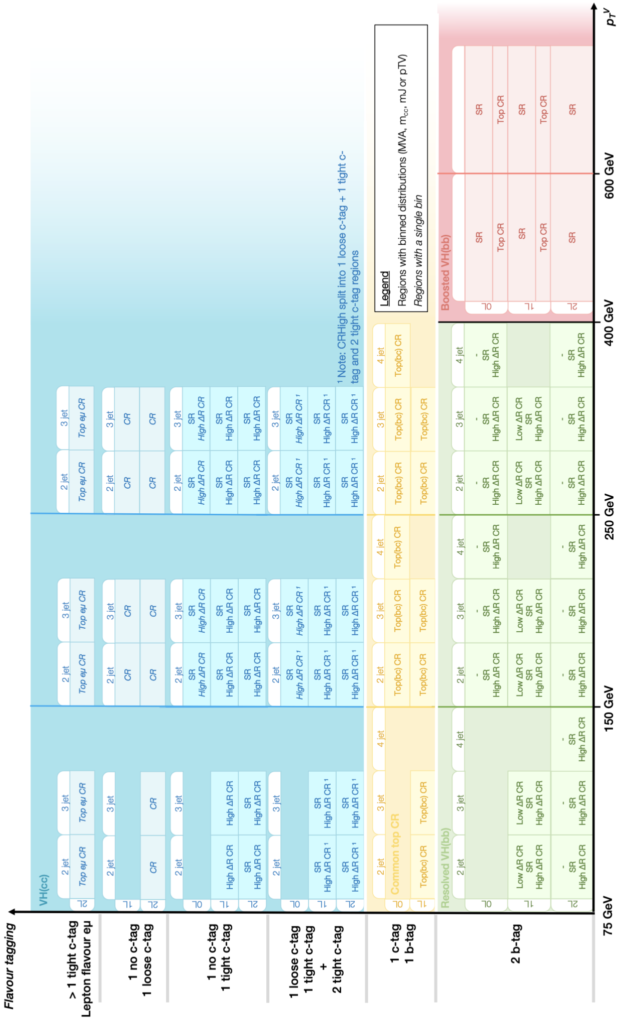


Figure 4.10. Illustration of the analysis regions considered in the analysis. In regions marked in indicated in italicized font, only the total yields are used in the fit. In all other regions either the BDT or di-jet mass distribution is considered.

For the $H \rightarrow c\bar{c}$ regime, a similar categorization is used with the following p_T^V regions: low-p_T^V $[75, 150]$ GeV (1-/2-*lepton* only), medium-p_T^V $[150, 250]$ GeV, and high-p_T^V $[250, \infty]$ GeV. Signal regions are required to have at least one tightly c-tagged jet. Similar to $H \rightarrow b\bar{b}$, each region has a CRHigh. An additional region is created to control the V+lf background, requiring a loosely c-tagged jet and a non-tagged jet. Top-enriched regions are shared between $H \rightarrow b\bar{b}$ and $H \rightarrow c\bar{c}$, requiring a b-tagged and a c-tagged jet for the 0-/1-*lepton* channels. In the 2-*lepton* channel, the top control region is built using different flavor requirements for the leptons.

The regions separating $H \rightarrow b\bar{b}$ and $H \rightarrow c\bar{c}$ cases in the resolved regimes based on the tagging requirements passed by the jets composing the Higgs candidate are shown in Figure 4.11.

In the *boosted* $H \rightarrow b\bar{b}$ regime, events are divided into two p_T^V regions: $[400, 600]$ GeV and $[600, \infty]$ GeV. The signal region requires that the two leading sub-jets of the large-radius jet be b-tagged. In the 0-/1-*lepton* channels, events with a b-tagged jet not matched with the large-radius jet will form a dedicated $t\bar{t}$ control region. In this *boosted* regime has a small contamination from $H \rightarrow c\bar{c}$ because no *boosted* regime is built for the latter.

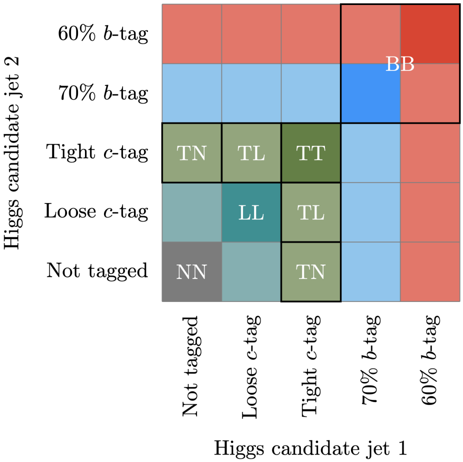


Figure 4.11. Tagging regions separating $H \rightarrow b\bar{b}$ and $H \rightarrow c\bar{c}$ in the *resolved* regime based on the tag requirement passed by the jets used for the Higgs candidate.

Compared with Ref. [98], a Multi-Variate Analysis (MVA) is performed to enhance the sensitivity of the measurement in both the *boosted* $H \to b\bar{b}$ and $H \to c\bar{c}$ regimes by training a Boosted Decision Tree (BDT) in each region of the analysis.

Different kinds of systematic uncertainties are considered in this analysis. Firstly, *experimental* uncertainties are considered for all the physics objects used in the analysis. Dedicated calibrations are then performed to provide the Flavour Tagging-related uncertainties. Secondly, *modelling* uncertainties are derived as two-point systematics by comparing the nominal Monte Carlo samples with alternative ones, QCD scale variations, and Parton Distribution Function variations.

Generally, three different kinds of uncertainties can be defined:

- **Normalization uncertainties**: Normalizations that can be extracted from data in pure regions or set to Monte Carlo predictions with an overall uncertainty derived from different generator comparisons.

- **Acceptance ratios**: Uncertainties on acceptance effects between different regions in the analysis with the same normalization factor. These uncertainties are derived by computing the double ratio of yields in two regions between two generators. The sum in quadrature of all the variations is considered.

- **Shape uncertainties**: Uncertainties on the BDT shape are derived using a multivariate approach (namely CARL), which trains a Neural Network to differentiate between the nominal and the alternative sample.

4.3.1 Samples

The dataset used for this analysis is the full Run 2 collected by the ATLAS detector between 2015 and 2018, corresponding to 140 fb^{-1} at $\sqrt{s} = 13$ TeV.

All the signal and background processes are simulated using Monte Carlo generators, except for the multi-jet background in the 1-*lepton* channel and the $t\bar{t}$ background in the 2-*lepton* channel, which are estimated using data-driven methods. The estimate of the multi-jet background is described in detail in Section 4.3.7.

The summary of the Monte Carlo samples used for the backgrounds is shown in Table 4.1.

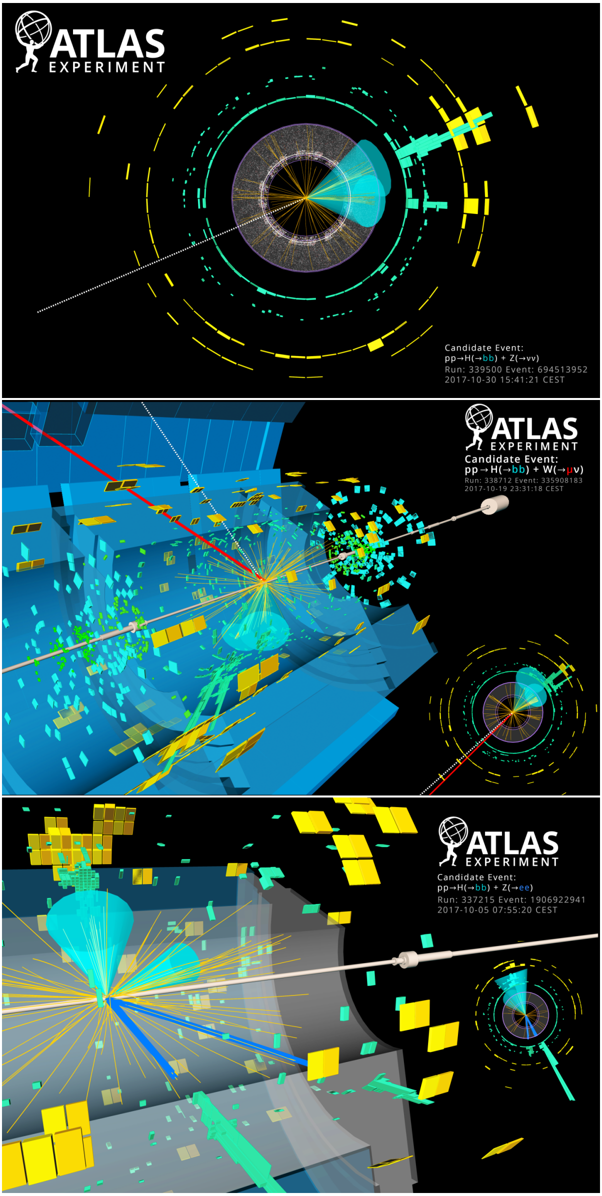


Figure 4.12. Event displays for 0-/1-/2-*lepton* channels respectively.

Process	Matrix Element	PDF Set (ME)	Parton Shower	σ order
$qq \to VH$	POWHEG-BOX v2 [75, 106] + GoSam [109] + MiNLO [110, 111]	NNPDF3.0NLO [107]	Pythia-8.245 [108]	NNLO(QCD)+ NLO(EW)
$gg \to ZH$	POWHEG-BOX v2	NNPDF3.0NLO	Pythia-8.245	NNLO(QCD)+ NLL(EW)
Z+ jets	Sherpa 2.2.11	NNPDF3.0NLO	Sherpa 2.2.11	NNLO
$t\bar{t}$	Powheg-Box v2	NNPDF3.0NLO	Pythia-8.230	NNLO+NNLL
single-top (s/t)	Powheg-Box v2	NNPDF3.0NLO	Pythia-8.230	NLO
single-top (Wt)	Powheg-Box v2	NNPDF3.0NLO	Pythia-8.230	Approx. NNLO
$qq \to VV$	Sherpa 2.2.11	NNPDF3.0NLO	Sherpa 2.2.11	NLO
$gg \to VV$	Sherpa 2.2.2	NNPDF3.0NLO	Sherpa 2.2.2	NLO

Table 4.1. Nominal Monte Carlo samples used in $VH(b\bar{b}|c\bar{c})$ analysis. The PDF sets in the table are the ones used for the matrix element.

4.3.2 Object and Event Selection

In this section, the object reconstruction and event selection procedures are described.

Object Reconstruction

Leptons used in this analysis are reconstructed as described in Sections 2.3.3 and 2.3.5. The selection criteria are detailed in Table 4.2 for electrons and Table 4.3 for muons. Unlike electrons and muons, τ-leptons are reconstructed hadronically and identified with Machine Learning techniques based on Recurrent Neural Networks, as described in Section 2.3.7. The selection criteria for τ-leptons are detailed in Table 4.4.

Electron Selection	p_T	η	ID	d_0^{sig} w.r.t. BL	$\|\Delta z_0 \sin\theta\|$	Isolation
VH-Loose	> 7 GeV	$\|\eta\| < 2.47$	LH Loose	< 5	< 0.5 mm	Loose_VarRad
ZH-Signal	> 27 GeV	Same as VH-Loose				
WH-Signal	Same as ZH-Signal		LH Tight	Same as ZH-Signal		HighPtCaloOnly

Table 4.2. Electron selection requirements

Muon Selection	p_T	η	ID	d_0^{sig} w.r.t. BL	$\|\Delta z_0 \sin\theta\|$	Isolation
VH-Loose	> 7 GeV	$\|\eta\| < 2.7$	Loose quality	< 3	< 0.5 mm	Loose_VarRad
ZH-Signal	> 27 GeV	$\|\eta\| < 2.5$	Same as VH-Loose			
WH-Signal	> 25 GeV when $p_T^V > 150$ GeV > 27 GeV when $p_T^V < 150$ GeV	$\|\eta\| < 2.5$	Medium quality	< 3	< 0.5 mm	HighPtTrackOnly

Table 4.3. Muon selection requirements

Three different anti-k_t jet collections have been used in this analysis for the various regimes:

Lepton	p_T	η	n_{trk}	ID
tau	> 20 GeV	$\lvert\eta\rvert < 2.5$	1 or 3 tracks	Loose

Table 4.4. Hadronic tau selection requirements

- **Small-R Jets**: As described in Section 2.3.6, these jets are built with a radius parameter of $R = 0.4$ using calorimeter information.
- **Large-R Jets**: As described in Section 2.3.6, these jets are built with a radius parameter of $R = 1.0$ using calorimeter information.
- **Variable-R Track-Jets**: As described in Section 2.3.6, these jets are built with a variable radius parameter between $0.02 < R < 0.4$ using track information.

The jet selection criteria are summarized in Table 4.5.

Jet Category	Tight Jet Cleaning	p_T	η	JVT/fJVT
Signal Jet	true	> 20 GeV	< 2.5	JVT > 0.5 for $\lvert\eta\rvert < 2.4$, $p_T < 60$ GeV jets
Forward Jet	true	> 30 GeV	$2.5 < \lvert\eta\rvert < 4.5$	fJVT < 0.5 for $p_T < 120$ GeV jets
Large-R jets	-	$p_T > 250$ GeV	$\lvert\eta\rvert < 2.0$	-
Track jets	-	$p_T > 10$ GeV	$\lvert\eta\rvert < 2.5$	-

Table 4.5. Jet selection requirements

Finally, the missing transverse momentum (MET or E_T^{miss}) is used as a sign of neutrinos for the 0-/1-*lepton* channels with the standard ATLAS procedure described in Section 2.3.8.

Event Selection

All the events use lepton triggers to efficiently select the leptonically decaying vector boson. The 0-*lepton* channel uses MET triggers with online thresholds of 70 GeV for the data collected during 2015, 90 and 110 GeV for data in 2016, and 110 GeV for the 2017/18 data. In the 1-*lepton* channel, muonic events are triggered with the same triggers as for the 0-*lepton* for p_T^V above 150 GeV, while for low-p_T^V single-muon triggers are used. In the leptonic case, single-electron triggers are required. In the 2-*lepton* channel, triggers are similar to the 1-*lepton* channel, changing the threshold for the muon case to 250 GeV.

Building the V and H Candidates

The Vector Boson In the 0-*lepton* channel, the p_T^V is the missing energy, requiring $MET > 150$ GeV where the trigger efficiency is above 90%. The scalar sum of the jets (S_T) is required to be larger than 120(150) GeV for the 2(3)-jet events, removing mis-modelling in the simulation of the jet activity.

A special case is considered when a τ_{had} is reconstructed in the 0-*lepton* channel from the W decay. These events require a further selection on the transverse mass of the W boson (m_T^W[2]). Imposing $m_T^W > 10$ GeV reduces fake τ contamination. These events are considered as in the 1-*lepton* channel, and in this case, the p_T^V is built as the vector sum of the MET and the transverse momentum of the τ: $\vec{p}_T^W = \vec{E}_T^{miss} + \vec{p}_T^{\tau}$.

As for the τ case, in the 1-*lepton* channel the p_T^V is the vector sum of the MET and the lepton transverse momentum: $\vec{p}_T^W = \vec{E}_T^{miss} + \vec{p}_T^{l}$. A $p_T^W > 75$ GeV is required.

For the 2-*lepton* channel, the two leptons are required to have the same flavor. In the case of the muon sub-channel, the two leptons are also required to have opposite charge. The invariant mass of the di-leptonic system is required to be compatible with the Z boson resonance: $81 < m(ll) < 101$ GeV in $H \to b\bar{b}$ *resolved* or $H \to c\bar{c}$, and $66 < m(ll) < 116$ GeV in $H \to b\bar{b}$ *boosted*. The p_T^V in this case is the vector sum of the two lepton systems: $\vec{p}_T^Z = \vec{p}_T^{l1} + \vec{p}_T^{l2}$.

The Higgs Boson In the $H \to b\bar{b}$ *resolved* or $H \to c\bar{c}$ regime, the Higgs candidate is built by exploiting the PCFT for each jet in the event. In these regimes, only Small-R jets are considered. To avoid overlaps, only events without any *b*-tagged jet and with at least one *c*-tagged jet (both loose or tight) are considered for the $H \to c\bar{c}$ regime. For the $H \to b\bar{b}$ *resolved* regime, exactly 2 *b*-tagged jets are required with a further veto on tight *c*-tagged jets for the 0-/1-*lepton* channels. In the 0-/1-*lepton* channels, events containing both *b*- and tight *c*-tagged jets are considered for the Top(bc) Control Regions. With these selections, the Higgs candidate is built using the two *b*-tagged jets in the $H \to b\bar{b}$ *resolved* regime. For the $H \to c\bar{c}$ case, jets are sorted as tight > loose > no *c*-tag. If two jets have the same *c*-tagging information, the one with the highest p_T is prioritized. The first two jets with this sorting are considered as the Higgs candidate. For the

---

[2] $m_T^W = \sqrt{2p_T^l E_T^{miss}(1 - \cos\Delta\phi(l, E_T^{miss}))}$

events entering the Top(bc) Control Regions, the Higgs candidate is built using the b-tagged and tight c-tagged jets, both selected with the highest p_T within their tagging category. In the 2-*lepton* channel, no tight c-tagging veto is applied in the Signal Regions, so these kinds of events will enter the Signal Regions.

After building the Higgs candidate, the two jets are sorted by p_T: the leading jet is denoted as j_1 while the sub-leading jet is j_2. The leading jet candidate is required to have $p_T^{j_1} > 45$ GeV while the sub-leading jets must have $p_T^{j_2} > 20$ GeV. Due to a known mis-modelling of soft collinear gluon splitting in V+jets, the invariant mass of the Higgs candidate is required to be $m(j_1, j_2) > 50$ GeV.

For the $H \to b\bar{b}$ *boosted* regime, track-jets are required to pass the 85% Working Point. The Higgs candidate is considered from the highest p_T Large-R jet. The Higgs candidate is required to have at least two track-jets associated with the Large-R jet. Three track-jets are considered for the tagging requirement. The signal region is required to have exactly two b-tagged sub-jets. The mass of the Large-R jet is required to be more than 50 GeV.

Further selection requirements are applied to suppress multi-jet and other backgrounds. A summary of all the criteria is reported in Table 4.6 for the $H \to b\bar{b}$ *resolved* and $H \to c\bar{c}$ regimes and in Table 4.7 for the $H \to b\bar{b}$ *boosted* regime.

4.3.3 Event Categorization

As shown in Figure 4.8, events are categorized into p_T^V bins and number of jet categories.

Signal regions are defined from the Flavour Tagging categorization as shown in Figure 4.11 by the solid black line. In the $H \to b\bar{b}$ *resolved* regime, the signal region is built with exactly two b-tagged jets. In the $H \to c\bar{c}$ regime, the signal regions are built with at least one tight c-tagged jet: TT with exactly two tight c-tagged jets, TL with a tight and a loose c-tagged jet, and TN with a tight c-tagged and a non-tagged jet.

To better constrain the various backgrounds of the analysis, several Control Regions have been built.

A High $\Delta R(j_1, j_2)$ Control Region (CRHigh) is built with a continuous cut on ΔR as a function of the p_T^V to constrain V+jets and $t\bar{t}$ normalization and shape. An example for the 2-jet category and the 2-*lepton* channel is shown in Figure 4.13 for the signal distribution.

Analysis regime	*$H \to b\bar{b}$ resolved*	$H \to c\bar{c}$
	Common Selections	
Jets	≥ 2 signal jets	
Candidate jets tagging	2 b-tags	≥ 1 c-tag [a]
Leading Higgs candidate jet p_T	> 45 GeV	
Sub-leading Higgs candidate jet p_T	> 20 GeV	
m_{bb} or m_{cc}	> 50 GeV (before correction)	
ΔR(jet1, jet2)	Upper cut $\Delta R < \pi$	
	0 Lepton	
Trigger	E_T^{miss} triggers	
Jets	≤ 4 jets	≤ 3 jets
Additional jets tagging	no *c*-tag	no *b*-tag
Leptons	0 VH-loose lepton	
E_T^{miss}	> 150 GeV	
S_T	> 120 (2 jets), > 150 GeV (3p jets)	
m_T^W	> 10 GeV (for events with at least one hadronic τ)	
$\lvert \min \Delta\phi(E_T^{miss}, \text{jet}) \rvert$	$> 20°$ (2 jets), $> 30°$ (3 jets)	
$\lvert \Delta\phi(E_T^{miss}, H) \rvert$	$> 120°$	
$\lvert \Delta\phi(\text{jet1, jet2}) \rvert$	$< 140°$	
	1 Lepton	
Trigger	e channel: single electron trigger μ channel: single muon trigger ($p_T^V < 150$ GeV) E_T^{miss} triggers (above)	
Jets	≤ 3 jets	
Additional jets tagging	no *c*-tag	no *b*-tag
hadronic τ	no hadronic τ	
Leptons	1 WH-signal lepton > 1 VH-loose lepton veto	
E_T^{miss}	> 30 GeV (e channel)	
S_T	> 120 (2 jets), > 150 GeV (3 jets)	
m_T^W	> 20 GeV ($75 < p_T^V < 150$ GeV only)	
	2 Lepton	
Trigger	e channel: single electron trigger μ channel: single muon trigger ($p_T^V < 250$ GeV) E_T^{miss} triggers (above)	
Additional jets tagging	-	no *b*-tag
Leptons	2 VH-loose leptons (≥ 1 ZH-signal lepton) Same flavour, opposite-charge for $\mu\mu$	
m_{ll}	$81 < m_{ll} < 101$ GeV	

Table 4.6. Summary of the event selection in the 0-, 1- and 2-lepton analyses.

Selection	0 lepton channel	1 lepton channel		2 lepton channel	
		e sub-channel	μ sub-channel	e sub-channel	μ sub-channel
Trigger	E_T^{miss}	Single electron	E_T^{miss}	Single electron	E_T^{miss}
Leptons	0 VH-*loose* lepton	1 WH-*signal* lepton no second VH-*loose* lepton no hadronic τ		$\geq$ 1 ZH-*signal* lepton 2 VH-*loose* leptons	
E_T^{miss}	$>$ 250 GeV	$>$ 50 GeV	-	-	
p_T^V	$>$ 400 GeV				
Large-R jet	at least one large-R jet $p_T >$ 250 GeV, $\lvert\eta\rvert < 2$				
Track-Jets	at least two track-jets $p_T >$ 10 GeV, $\lvert\eta\rvert < 2.5$ matched to the leading large-R jet				
b-jets	exactly two of the leading three track-jets matched to the leading large-R must be b-tagged				
m_J	$>$ 50 GeV				
min $\Delta\phi(E_T^{miss}, \text{jets})$	$> 30°$	-		-	
$\Delta\phi(E_T^{miss}, H_{\text{cand}})$	$> 120°$	-		-	
m_{ll}	-	-		66 GeV $< m_{ll} <$ 116 GeV	
Lepton flavor	-	-		two lepton same flavour	
Lepton charge	-	-		opposite sign muons	

Table 4.7. Event Selection for the three channels of the boosted $H \to b\bar{b}$ analysis.

This function is built to contain 95% (85%) of the signal below the ΔR cut for the 2(3)-jet category. The TT and LT Signal Regions are then merged together into a named XT region while the respective CRHigh are kept separate. In the 1-*lepton* channel, a Low $\Delta R(j_1, j_2)$ Control Region (CRLow) is also built to constrain the W+hf background. This region is defined to keep 90% of the diboson signal in the Signal Region.

The already defined Top(bc) Control Region with a b-tag and a tight c-tag jet is used to better constrain the Top backgrounds in the 0-*lepton* and 1-*lepton* channels. In the 2-*lepton* channel, a high-purity $t\bar{t}$ control region is built with different flavour leptons.

Finally, a V+lf region is also created in the 1-*lepton* and 2-*lepton* channels using a LN tag region with a loose c-tag jet and a non-tag jet.

A summary of the event categorization for all the control regions is reported in Figure 4.14.

For the $H \to b\bar{b}$ *boosted* regime, the Signal Region is defined by a Large-R jet composed of exactly two b-tagged track-jets matched to the Large-R jet. A top control region is built with additional b-tagged track jets outside the Higgs candidate. In the 0-*lepton* and 1-*lepton*

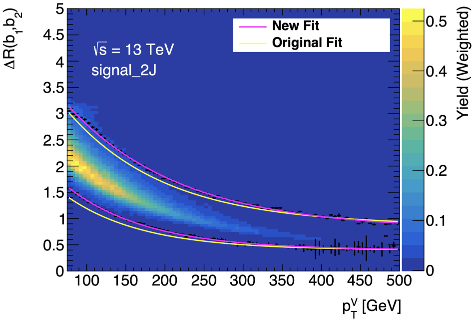


Figure 4.13. Signal distribution of ΔR vs p_T^V in the 1-*lepton* channel and the 2 jet category. The pink lines shows the function delimiting the High/Low ΔR control region. In yellow the reference line used in the previous iteration of the analysis.

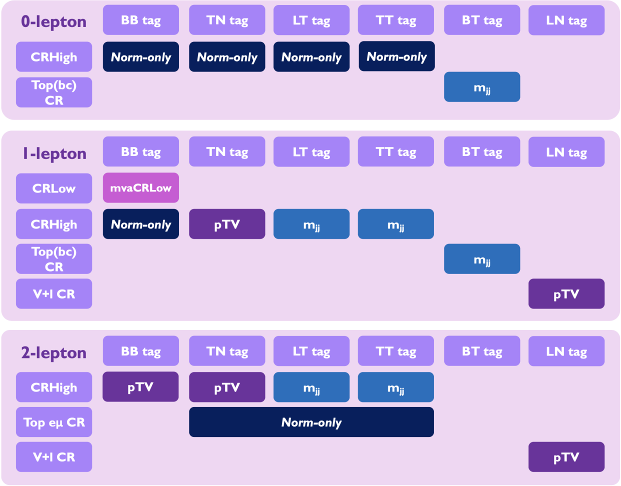
channels, the region can be further split into *High Purity* (HP) and *Low Purity* (LP) based on the number of calo Small-R jets outside the Large-R jets. This information, which was used in the previous *boosted*-only analysis to create two separate Control Regions, is now used in the BDT training but keeps HP and LP merged in the analysis.

4.3.4 Jet Corrections

A series of jet corrections is applied to improve the Higgs candidate momentum resolution.

- **Muon In-Jet Correction**: This correction aims to correct the momentum of the b-/c-tag jet with the muonic decay of the heavy flavour hadron that is not captured by the calorimeter. If a muon is found in a tagged jet with $\Delta R(\mu, j) < \min(0.4, 0.04 + 10\text{GeV}/p_T^{\mu})$, the muon 4-momentum is then added to the jet after subtracting the calorimeter energy deposit.

- **Final State Radiation Recovery**: Applied to the 2-*lepton* channel to events with 3 or more jets. This correction aims to improve the $m(j_1, j_2)$ resolution and peak. Continuous cuts are applied on the sum of ΔR between the third or fourth jets with the two Higgs candidate jets. The closest jet below the cut is considered as an FSR-recovered jet.



0-lepton	BB tag	TN tag	LT tag	TT tag	BT tag	LN tag
CRHigh	Norm-only	Norm-only	Norm-only	Norm-only		
Top(bc) CR					m_{jj}	

1-lepton	BB tag	TN tag	LT tag	TT tag	BT tag	LN tag
CRLow	mvaCRLow					
CRHigh	Norm-only	pTV	m_{jj}	m_{jj}		
Top(bc) CR					m_{jj}	
V+l CR						pTV

2-lepton	BB tag	TN tag	LT tag	TT tag	BT tag	LN tag
CRHigh	pTV	pTV	m_{jj}	m_{jj}		
Top eμ CR		Norm-only				
V+l CR						pTV

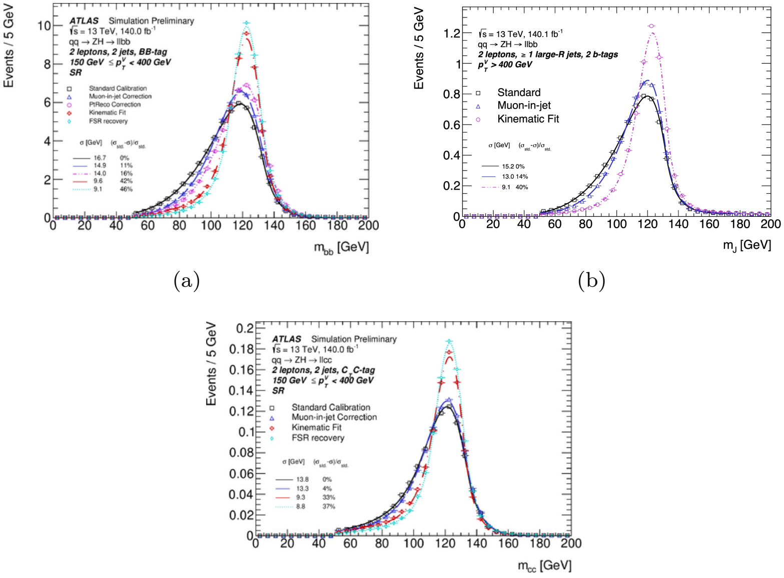
Figure 4.14. Scheme of the event categorization with the variable used in the fit in each control region.

- **PtReco Correction**: This correction aims to correct the missing energy in semi-leptonic decays of heavy hadrons and out-of-cone effects of b-jets. It is applied only to the $H \to b\bar{b}$ *resolved* regime. It is derived from the simulation as the ratio of p_T between the jets after the Muon In-Jet correction and the truth jet p_T on b-tag jets.
- **Kinematic Fit**: This correction is applied only in the 2-*lepton* channel where the event is expected to be reconstructed entirely. A kinematic fit is performed to improve the $m(j_1, j_2)$ resolution following the methods used in Run 1 Ref. [112, 113].

A summary of the jet corrections is shown on signal samples in Figure 4.15 for the 2-*lepton* channel. Overall, the mass resolution can be improved by up to 40%.

4.3.5 Truth Tagging strategy

The default Flavour Tagging method Ref. [114] is defined as Direct Tagging, where a jet passes or fails the tag requirement based on the score of DL1r. The efficiencies and mis-tag rates for various flavour jets are shown in Table 4.8.



(a) (b)

(c)

Figure 4.15. Tagged energy correction performance in different analysis schemes in the 2-lepton channels. Events in different nJet or p_T^V categories are shown inclusively in different analysis schemes. The events are taken from signal MC samples.

With tagging efficiency and rejection rate limits, only part of the events satisfy the requested tagging conditions while many other events are discarded from the analysis. For this reason, many regions in the analysis lack Monte Carlo statistical sensitivity, leading to a large Monte Carlo statistical uncertainty. An alternative approach, known as *truth tagging*, uses all the events by weighting each event with the efficiency of being selected. This event weight is essentially the product of the tagging efficiencies of the tagged jets times the inefficiencies of the non-tagged jets. A traditional way of implementing *truth tagging* is via 2D efficiency maps in p_T and η. Tagging efficiency can have non-trivial dependencies on other variables, but a higher-dimensional map would cause a lack of statistics in each tag bin. For this reason, Graph Neural Networks have been used Ref. [115] to exploit this multi-dimensional parametrization of the tagging efficiency. The input

PCFT bin name	**b-jets**	**c-jets**	**light-jets**	**τ-jets**
Untagged	14.5%	47.4%	92.4%	60.3%
Loose c-tag	11.5%	20.5%	6.5%	18.5%
Tight c-tag	4.8%	24.2%	0.9%	19.5%
70% b-tag	11.2%	5.2%	0.1%	1.7%
60% b-tag	58%	2.6%	0.05%	0.5%

Table 4.8. Jet efficiencies for different PCFT bins.

to the network is composed of a graph where a node represents a jet. The input features are the kinematics of the jets and their constituents. A schematic view of the GNN-based Truth Tagging methods is shown in Figure 4.16.

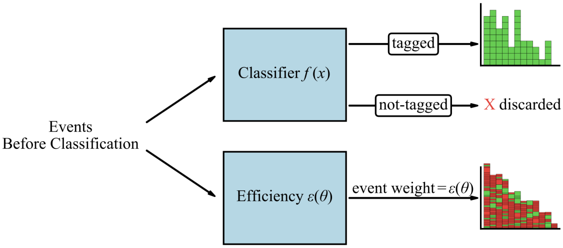


Figure 4.16. GNN based *Truth Tagging* scheme. Ref. [115]

GNN-based *Truth Tagging* methods can provide smoother distributions of the tagging efficiency compared to the standard Direct Tagging while maintaining a reasonable closure with the Direct Tagging distribution. Since the efficiency on b-jets is quite high (ca. 70%), the impact of *Truth Tagging* should be quite limited. For this reason, this method is applied to $H \to b\bar{b}$ resolved, only to non-b-jets: *Hybrid Tagging.*

The final strategy is thus decided as follows:

- *$H \to b\bar{b}$ resolved*: GNN-based *Hybrid Tagging* applied to V+jets and single-top s-/t-channels. All the other samples use *Direct Tagging.*
- *$H \to b\bar{b}$ boosted*: GNN-based *Truth Tagging* applied to all samples and jet flavors except for signal and diboson.
- $H \to c\bar{c}$: GNN-based *Hybrid Tagging* applied to a randomized 8% of the V+jets and single-top s-/t-channel samples in the TT

tagging category. The remaining 92% uses *Direct Tagging* and is categorized in the remaining TL, TN, LN, BT categories. The other samples use *Direct Tagging*.

4.3.6 A Multi-Variate Analysis

		$H \to b\bar{b}$ *resolved* $H \to c\bar{c}$			$H \to b\bar{b}$ *boosted*		
Variable	**Name**	0-lepton	1-lepton	2-lepton	0-lepton	1-lepton	2-lepton
$m(j_1, j_2)$	mBB / mJ	×	×	×	×	×	×
$m(j_1, j_2, j_3)$	mBBJ	×	×	×			
$p_T^{j_1}$ or p_T^{trk}	pTB1 / pTBTrkJ1	×	×	×	×	×	×
$p_T^{j_2}$	pTB2 / pTBTrkJ2	×	×	×	×	×	×
$p_T^{j_3}$	pTBTrkJ3				×	×	×
$\sum_{i>2} p_T^i$	sumPtAddJets	×	×	×			
$\Delta R(j_1, j_2)$	dRBB / deltaRTrkJbTrkJ	×	×	×	×	×	×
$\|\Delta\eta(j_1, j_2)\|$	dEtaBB	×					
$\mathrm{bin}_{DL}(j_1)$	bin_btagB1 / bin_bTagBTrkJ1	×	×	×	×	×	×
$\mathrm{bin}_{DL}(j_2)$	bin_btagB2 / bin_bTagBTrkJ2	×	×	×	×	×	×
P_T^V	pTV	×	×	×	×	×	×
E_T^{miss}	MET	×	×		×	×	
$E_T^{\mathrm{miss}}/\sqrt{S_T}$	METSig			×			
$\|\Delta y(\vec{V}, H_{\mathrm{cand}})\|$	dYVBB / deltaYVJ		×	×		×	×
$\|\Delta\phi(\vec{V}, H_{\mathrm{cand}})\|$	dPhiVBB / absdeltaPhiVJ	×	×	×	×	×	×
$\min[\Delta\phi(\vec{\ell}, j_1 \text{ or } j_2)]$	dPhiLbmin		×				
m_{eff}	MEff	×					
m_T^W	mTW		×				
m_{top}	Mtop		×				
m_{ll}	mLL			×			
$\cos\theta(\vec{\ell}, \vec{Z})$	cosThetaLep			×			×
$\left(p_T^{\ell} - E_T^{\mathrm{miss}}\right)/p_T^W$	lepPtBalance					×	
p_T^{ℓ}	pTL					×	
$N_{\mathrm{trackjetsinJ}}$	NAdditionalCaloJets				×	×	×
$N(\mathrm{add.smallR-jets})$	NMatchedTrackJetLeadFatJet				×	×	×
Colour	Colour				×	×	×
$\min(\Delta R(b, j))$	minDRBJets	×	×				

Table 4.9. MVA variables used for the 0-, 1- and 2-lepton channels in the resolved and boosted topologies for the VH($\to b\bar{b}$) and VH($\to c\bar{c}$) analyses. The VH($\to c\bar{c}$) analysis only probes the resolved topology. The Higgs candidate system, composed of the 2 small-R jets ($b\bar{b}$ or $c\bar{c}$) for the resolved regime or of the leading large-R jet (J) of the event for the boosted regime, is sometimes denoted H_{cand}.

One of the other main changes with the previous analysis is the use of Multi-Variate Analysis discriminant in all the Signal Regions of the analysis. In the previous $H \to b\bar{b}$ *boosted* and $H \to c\bar{c}$ analyses, the sensitivity was extracted by fitting the invariant mass of the Higgs candidate. In this case, Boosted Decision Trees (BDT) have been used to discriminate signal from backgrounds in all regions. For each region, two separate sets of training have been performed: one for the VH signal and one for the VZ diboson cross-check.

In the $H \to b\bar{b}$ *resolved* and $H \to c\bar{c}$ analyses, the training is performed separately in the XT, NT, and BB regions. In the 0-*lepton* channel, the training is performed inclusively for $p_T^V > 150$ GeV and then a training is performed for each jet category (2, 3, or 4 jets). In the 1-*lepton* channel, the training is instead split at $p_T^V = 150$ GeV and between 2 and 3 jet categories. Similarly, for the 2-*lepton* channel, with the only difference that the split in jet categories is between 2 and $\geq$3 jets.

In the $H \to b\bar{b}$ *boosted* regime, a training is performed in each lepton channel.

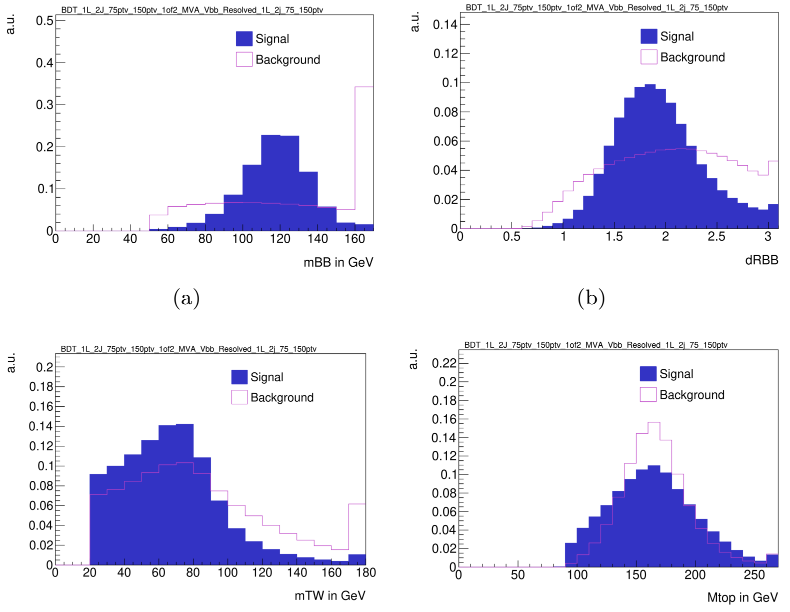
The summary of the input variables to the BDT is shown in Table 4.9. A few variables used as input to the BDT are shown in Figure 4.17 for the 1-*lepton* channel and 2 jet category.

The performance and over-training are checked by comparing the evaluation of the performance on the training and evaluation samples. In Figure 4.18(a), the ROC curve with its Area Under the Curve (AUC) is shown for the 1-*lepton* channel 2 jet category, while the over-training check is shown in Figure 4.18(b).

A similar BDT is trained in the CRLow of the 1-*lepton* channel in the $H \to b\bar{b}$ *resolved* regime. This network aims to separate the $W + bb$ component from the Top in this low ΔR Control Region to better control the $W + bb$ background.

Output transformation Since the output of the BDT is designed to maximize the separation of the signal and background populations, the optimal sensitivity of the MVA is not necessarily achieved with the default binning, which is equidistant and very fine. To achieve optimal binning, the following factors should be taken into account: low BDT output scores are populated by background events and almost entirely depleted of signal events, high BDT output scores have a good signal-to-background ratio, and the statistical uncertainty of each bin.

Therefore, a transformation of the BDT output distribution, designed to optimize the final sensitivity and reduce the number of bins,



(a) (b)

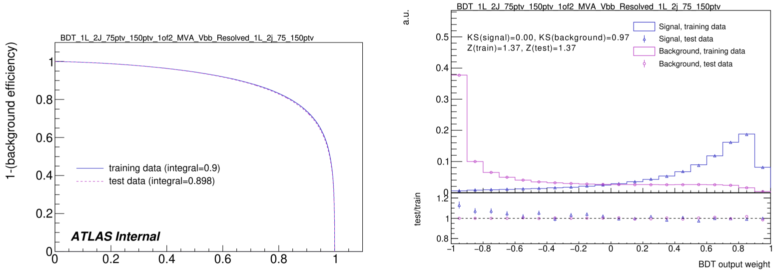
(c) (d)

Figure 4.17. Few input distribution to the BDT for the 1-*lepton* channel case and 2 jet category.

is implemented. This transformation is called Transformation D, and the formula used is as follows:

$$Z = z_s \frac{n_s}{N_s} + z_b \frac{n_b}{N_b} \tag{4.6}$$

In this equation, N_s and N_b are the total number of signal and background events, respectively, while n_s and n_b are the number of events in a given bin. The default 500 bins are then merged from high to lower BDT scores, requiring n_s and n_b to satisfy $Z > 1$. In this sense, z_s and z_b are tunable parameters optimized in each analysis region. The total number of bins varies between 4 to 15 in all the regions. An example of the BDT score distribution before and after transformation can be seen in Figure 4.19.



(a) (b)

Figure 4.18. Overtraining checks for the 2-jet category and in the $p_T^V \in [75, 150]$ GeV region for BDT trainings performed with BDT output distributions for signal and background obtained for the training (line) and test (dots) samples. The bottom panel shows the ratio of the test over training BDT output distributions for signal and background.

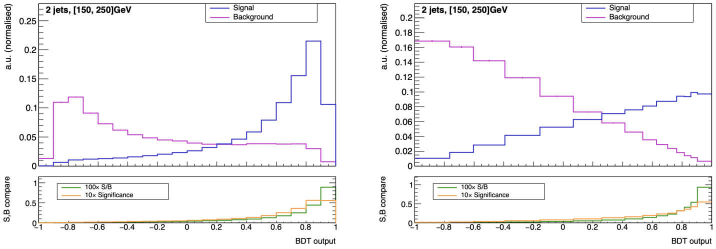


Figure 4.19. Example of comparison between before and after the transformation of the BDT output score.

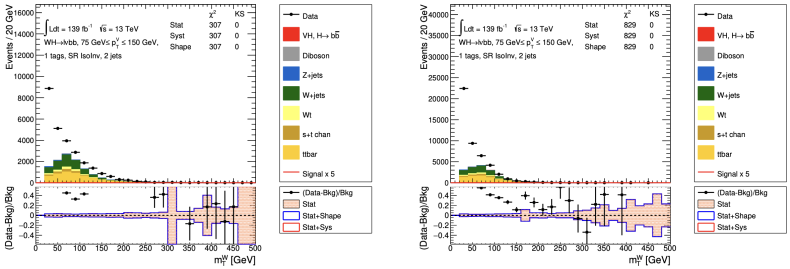
4.3.7 Multijet estimation

In the 1-lepton channel, the contribution of the multijet background is small but still significant compared to the other backgrounds, particularly at low values of p_T^V. To minimize the effect of multijet contamination, isolation criteria are employed separately for the electron and muon sub-channels. For the electron sub-channel, the isolation is based on calorimeter contamination using a variable denoted as *TopoEtCone20*, while for the muon sub-channel, isolation is based on a track base using a variable denoted as *PtCone20*.

The multijet background contribution varies across different p_T^V regions. In the low-p_T^V region, it is more significant compared to the two other regions (medium-/high-p_T^V). However, in the boosted regime, the multijet contribution is found to be negligible when compared to other Monte Carlo backgrounds. The multijet background is thus estimated separately in the low-p_T^V and inclusively in the medium-/high-p_T^V regions. In the low-p_T^V region, a cut on the transverse mass of the W-candidate, $m_T^W > 20$ GeV, is applied to reduce the multijet background. Additionally, in the electron sub-channel, a cut in $E_T^{miss} > 30$ GeV is used to further reduce the multijet contamination in all the p_T^V regions.

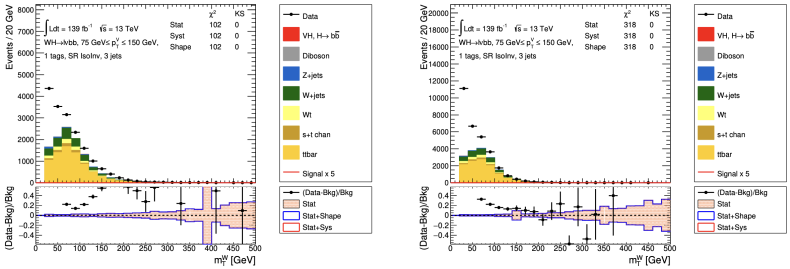
The remaining multijet background cannot be accurately estimated through simulation due to practical limitations in generating sufficient events for the analysis phase space, as well as the inherent challenge in modeling fake leptons. Therefore, a data-driven approach has been employed to estimate the impact of this background, utilizing a template method.

Template Method



(a) 2-jet Electrons

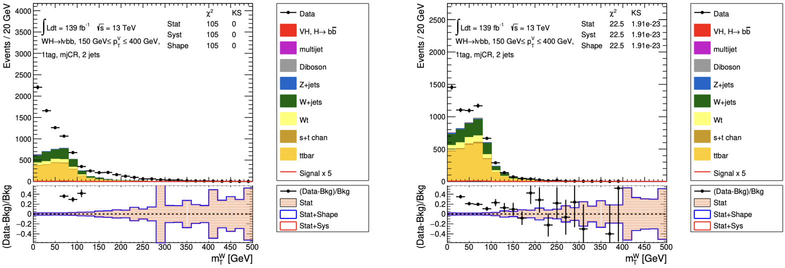
(b) 2-jet Muons



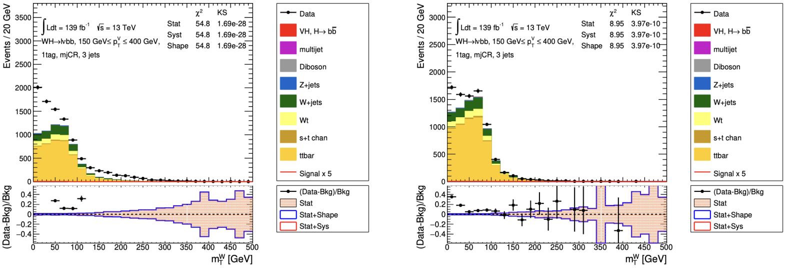
(c) 3-jet Electrons

(d) 3-jet Muons

Figure 4.20. low-p_T^V region templates



(a) 2-jet Electrons (b) 2-jet Muons



(c) 3-jet Electrons (d) 3-jet Muons

Figure 4.21. medium+high-p_T^V region templates

Under the assumption that the QCD background is small compared with the other MC samples, the template method is used to extract the shape of the multijet background in the Signal Region from an orthogonal Control Region enriched with multijet events ("mjCR"). The "mjCR" must be as similar as possible to the "SR" in terms of multijet composition in order to extrapolate it from one region to the other. The multijet contribution is then obtained by fitting some distribution in the "SR" where the fit involves a multijet template extracted from the "mjCR" and the other MC templates in the "SR". In the end, the fit will return the Normalization Factor (NF) for the multijet and for each component left floating in the fit. In order to be efficient, the shape of the multijet template must be different from the other MC shapes.

Because different isolation criteria are used for electrons and muons, and also because the sources of fake electrons and fake muons are different, independent template fits are performed in each of the electron and muon sub-channels separately. Independent template fits are also performed in the low-p_T^V and the medium+high-p_T^V regions. Finally, independent template fits are performed in the 2-jet and 3-jet categories. Overall, there are 8 independent regions where the fit is performed: electron/muon, low-/medium+high-p_T^V, 2-/3-jet.

In addition to this categorization, the analysis splits events into three exclusive regions: CRLow, SR, and CRHigh. The same categorization holds for the multijet estimation, where the two Control Regions (CR) are treated as a single bin to avoid shape effects from them.

For each signal region, a multijet-enriched control region, "mjCR", is obtained by requiring the inversion of the tight isolation cut required on the lepton, as shown in Table 4.10. In the SR, the tight isolation cut for the electron requires $TopoEtCone20 < \max(1.5\%p_T^{el}, 3.5)$ GeV, while the tight muon isolation requires $PtCone20 < 1.25$ GeV. However, for the muon isolation, a soft track isolation is maintained also in the "mjCR" with `PtCone20`< 4 GeV. The statistics are still limited in "mjCR" when 2 b-tags are required. In order to reduce the impact of statistical fluctuations in the multijet template, events with a single *b*-tag jet are used.

	Isolation, 2 b-tags	Inverted Isolation, 1 b-tag
Electron	$TopoEtCone20 < max(1.5\%p_T^{el}, 3.5)\ GeV$	$TopoEtCone20 > max(1.5\%p_T^{el}, 3.5)\ GeV$
Muon	$PtCone20 < 1.25\ GeV$	$1.25 < PtCone20 < 4\ GeV$

Table 4.10. Summary of differences in lepton isolation in "SR" (Isolation 2 b-tag) and in "mjCR" (Inverted isolation 1 b-tag)

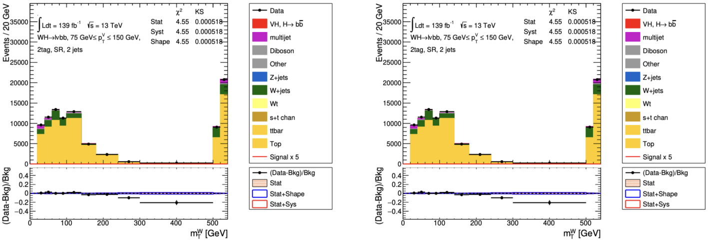
A multijet template can be constructed for any variable from the data in the "mjCR" as the difference between data and the non-multijet background obtained from simulation. The already mentioned transverse mass of the W-candidate, m_T^W, is found to be the variable offering the best discrimination between multijet and non-multijet backgrounds, and therefore it will be chosen for the nominal template to be fitted. In this analysis, the main backgrounds for the 1-lepton channel arise from pair and single production of top quarks, and from W+jets production. Smaller backgrounds are Z+jets and diboson production. For the multijet estimation purpose, all these non-multijet backgrounds are denoted as electroweak backgrounds in the following.

The plots in Figures 4.20 and 4.21 show the m_T^W distribution for data and electroweak backgrounds in the "mjCR" for the low-/medium-p_T^V regions, respectively, where the multijet template is extracted. In each region, the distribution for 2-/3-jets categories and El/Mu sub-channels is shown.

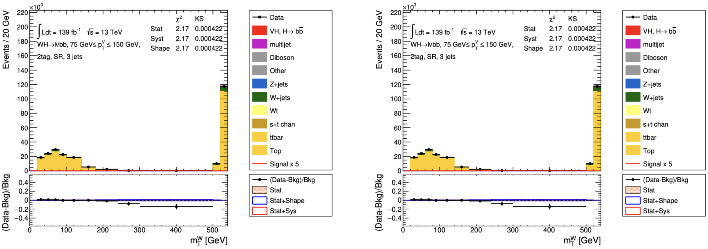
In each of the eight "SR," a fit to the m_T^W distribution is performed to simultaneously extract the normalization factor (NF) for the multijet and the electroweak backgrounds. As already mentioned, the CRHigh and CRLow are included in the fit as single bin distributions.

Practical Implementation

The $t\bar{t}$ and the W+jets processes are dominant in the "SR" where the fit is performed, and their normalization can have a large impact on the multijet estimate. Being smaller than the other backgrounds, the diboson and the Z+jets samples are kept at their nominal value. The single-top is merged with the $t\bar{t}$ background, with the result collectively called the Top background.



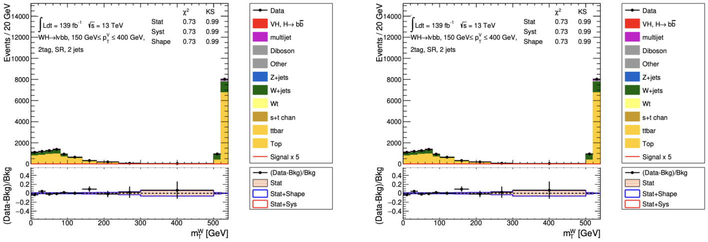
(a) 2-jet Electrons (b) 2-jet Muons



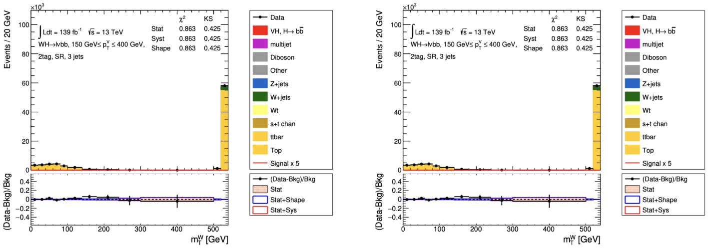
(c) 3-jet Electrons (d) 3-jet Muons

Figure 4.22. low-p_T^V post fit in "SR"

The events with a single leptonic W decay exhibit a characteristic peak in terms of m_T^W, while this is not the case for the multijet background, thus explaining the reason for the choice of this variable. The discrimination between Top and W+jets is driven by the different composition in CRHigh and CRLow for these backgrounds.



(a) 2-jet Electrons (b) 2-jet Muons



(c) 3-jet Electrons (d) 3-jet Muons

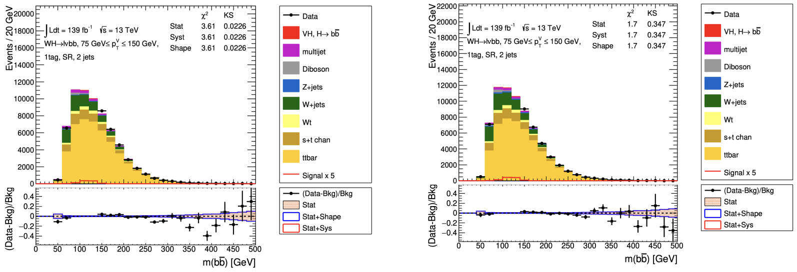
Figure 4.23. medium+high-p_T^V post fit in "SR"

Technically, the multijet fit is implemented as a template fit to data in each of the eight regions, with the CRLow and CRHigh included as two bins at the end of the m_T^W distribution in SR. The binning of the m_T^W distribution is optimized to have roughly constant MC statistical uncertainty. The multijet, Top, and W+jets have their templates and a floating normalization factor in the fit.

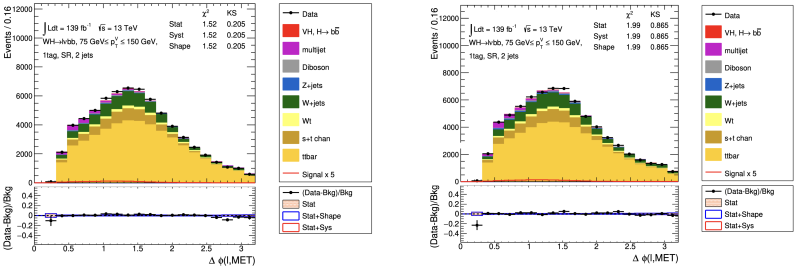
	Electrons		Muons	
NF	Top ($t\bar{t}$+s-top)	W+jets	Top ($t\bar{t}$+s-top)	W+jets
low-p_T^V 2-jet	0.991 ± 0.011	1.83 ± 0.10	1.000 ± 0.010	1.57 ± 0.08
low-p_T^V 3-jet	0.988 ± 0.007	2.02 ± 0.12	0.969 ± 0.006	1.84 ± 0.09
med+high-p_T^V 2-jet	0.994 ± 0.02	1.12 ± 0.07	0.976 ± 0.02	1.14 ± 0.05
med+high-p_T^V 3-jet	0.947 ± 0.01	1.25 ± 0.08	0.929 ± 0.001	1.31 ± 0.01

Table 4.11. Normalization factors obtained for Top and W+jets MC in the SR fit of m_T^W.

The normalization factors for Top and W+jets are presented in Table 4.11 and are affected by anti-correlation. Postfit plots are shown in Figures 4.22 and 4.23 for the low-/medium+high-p_T^V regions, respectively. Additionally, in Figure 4.24, the post-fit plots for other variables such as m_{bb} and $\Delta\phi(l, \text{MET})$ are shown for the 2-jet category for Electron and Muon sub-channels, respectively.



(a) m_{bb} Electron (b) m_{bb} Muon



(c) $\Delta\phi(l, MET)$ Electron (d) $\Delta\phi(l, MET)$ Muon

Figure 4.24. m_{bb} and $\Delta\phi$ post-fit distribution for 2-jet Electrons and Muons in "SR".

Systematic uncertainty

In both Electron and Muon sub-channels, a set of systematic uncertainties are considered. Two sources of uncertainty are taken into account, affecting the shape and the normalization of the MVA distribution of the multijet used for the main analysis. The systematic uncertainties that impact the shape are described first; they are also considered for the normalization.

The following sources of shape uncertainty are considered:

- An evaluation of the uncertainty introduced by the extrapolation of the multijet template from the full inverted isolation region to the signal region is considered. Instead, a reduced inverted-isolation region is defined, with additional isolation cuts applied to the inverted isolation region defined in Table 4.10. In the Electron sub-channel, this is done by additionally requiring $TopoEtCone20 < 12$ GeV, and in the Muon sub-channel $PtCone20 < 2.9$ GeV.

- For the nominal multijet template shape, the subtraction of electroweak MC from data in the "mjCR" is evaluated without applying any normalization factor (data – MC); The impact of using the normalization factors is thus evaluated with this systematic uncertainty. Therefore, the template shape is re-evaluated with the application of the nominal normalization factors before the subtraction (data – MC*NF).

In both of the cases above, the difference between the nominal and alternative shapes is taken as a systematic uncertainty. The shape systematic uncertainty is then passed to the final fit while the relative normalization uncertainty is cumulated independently with the following normalization uncertainties.

	low-p_T^V mj (%)	medium+high-p_T^V mj (%)
2jet El	$2.6^{+0.3}_{-0.7}$	$1.5^{+1.4}_{-1.4}$
3jet El	$0.2^{+0.7}_{-0.2}$	$0.6^{+0.3}_{-0.6}$
2jet Mu	$2.3^{+0.3}_{-0.8}$	$0.5^{+1.2}_{-0.5}$
3jet Mu	$1.2^{+0.2}_{-0.4}$	$0.0^{+0.0}_{-0.0}$

Table 4.12. multijet fractions relative to the total simulation, along with their associated uncertainty in the 2-jet and 3-jet and for **low-/medium-**p_T^V regions.

The impact of a systematic variation on the multijet normalization is indirectly driven by changes to the m_T^W template distributions and the relative yields in the "SR." The individual contributions to the normalization uncertainty are added in quadrature to give the overall normalization uncertainty for each "SR." In addition to the sources considered for the shape uncertainties, the following ones are considered exclusively for the normalization uncertainty:

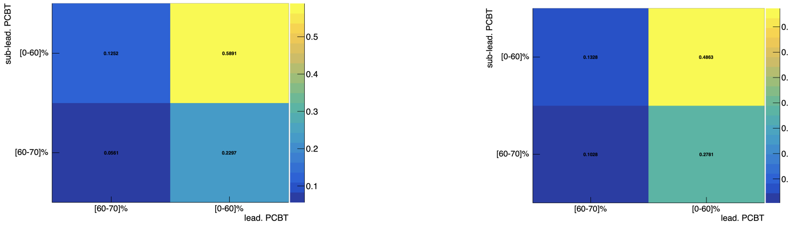
- In the low-p_T^V region only, including the $m_T^W < 20$ GeV events in the template fit enhances the multijet yields.

- In the medium+high-p_T^V region and Electron sub-channel only, including the $MET < 30$ GeV events in the template fit.

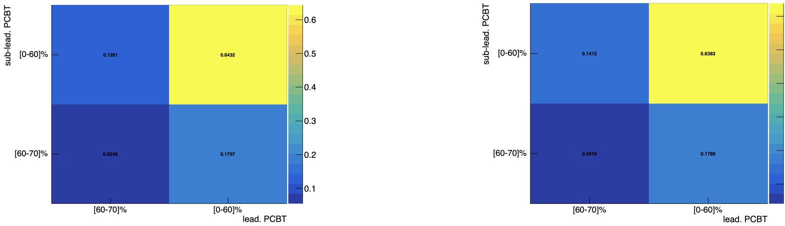
- Fitting using an alternative variable for the fit rather than m_T^W. In particular, the variable selected is $\Delta\phi(l, \mathrm{MET})$, which is correlated with m_T^W and shows good discrimination between the multijet and the electroweak template shapes.

The combination of all the 5 normalization uncertainties (2 coming from the shape systematics) gives rise to the fraction of multijet compared to all the samples, presented in Table 4.12 for **low-/medium-p_T^V** regions respectively.

b-tagging Score of Multijet Events



(a) 2 jet Electron (b) 3 jet Electron



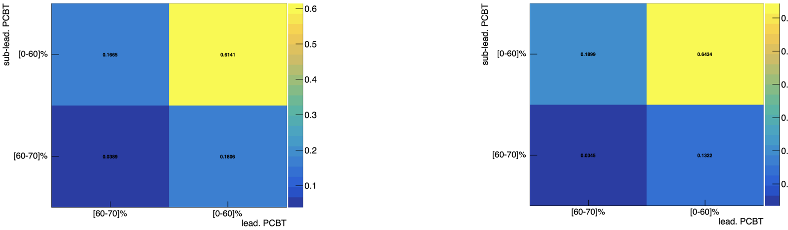
(c) 2 jet Muon (d) 3 jet Muon

Figure 4.25. PCBT scores in "pcbtCR" low-p_T^V

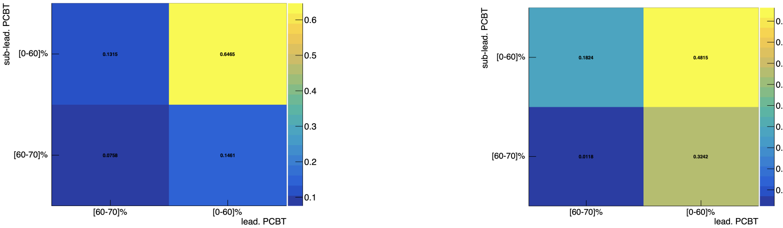
The MVA used to perform the fit in the main analysis takes as input the Pseudo Continuous B-Tagging score (PCBT) of the two b-tagged jets. Since the multijet events are estimated as the subtraction of electroweak from the data in the "mjCR," which is in the 1-tag region with the inverted isolation requirement on the lepton, a solution had to be found to assign the PCBT scores to the multijet events.

For this purpose, an emulation of the PCBT scores distribution of the 2 b-tagged jets for multijet from an orthogonal region to the "SR" and the "mjCR" is needed. This region, called "pcbtCR," is built by requiring 2 b-tags and also the inversion of the isolation criteria on the leptons. From this region, 2-by-2 maps of leading vs subleading jet PCBT scores have been created. These distributions are 2x2 because each tagged jet has efficiency lower than 70% and can thus lie in the tight PCBT bin ($\epsilon_b \in [0, 60]$) or loose PCBT bin ($\epsilon_b \in [60, 70]$).

In this region, for each bin in the 2-by-2 maps, the same electroweak Monte Carlo to data subtraction is applied, obtaining, after the normalization, the multijet PCBT scores template maps shown in Figures 4.25 and 4.26.



(a) 2 jet Electron (b) 3 jet Electron



(c) 2 jet Muon (d) 3 jet Muon

Figure 4.26. PCBT scores in ”pcbtCR” medium+high-p_T^V

These 2-by-2 normalized distributions provide the joint probabilities for leading and subleading jets of 2 b-tag multijet events to have PCBT scores above or below the 60% efficiency threshold. They are used to assign PCBT scores to both the leading and subleading jets in each data or electroweak background event from the "mjCR" according to the above probabilities obtained from "pcbtCR".

In Figure 4.27, a summary sketch of the multijet estimation is described with all the regions involved.

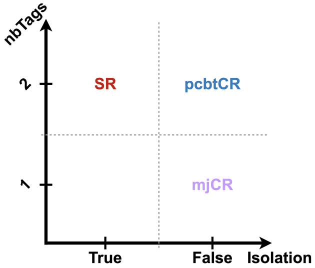


Figure 4.27. Schematic view of the multijet estimation procedure.

A similar procedure is applied to $H \to b\bar{b}$ *boosted*. In $H \to b\bar{b}$ *boosted*, we found that the multijet contamination is negligible. The same procedure is applied as for the resolved regime. The multijet shape is extracted from an orthogonal region separately for the electron and muon sub-channels, requiring exactly 0-btag VR track jets inside the larger jet candidate and the standard isolation inversion on the lepton as for the resolved regime. Then the shape is fitted in the standard isolated 2-btag signal region, including a single bin distribution for the Control Region.

With this configuration, the fitted multijet yields are observed to be less than 2×10^{-5} events for both Electron and Muon sub-channels.

4.3.8 Signal and Background modelling

As already mentioned, the modeling uncertainties are divided into three macro categories: normalization, acceptance, and shape uncertainties.

Since many uncertainties are derived from alternative samples or variations as two-point systematics, the list of the alternative samples is shown in Table 4.13.

Signal

The VH signal is divided into three different final states depending on the leptonic decay of the vector boson: $Z \to \nu\bar{\nu}$, $W \to l\nu_l$, and $Z \to l^+l^-$. Here the modeling scheme for the various categories is discussed.

Sample	Nominal Generator	Alternative Generators	Systematic Effects
VH	Powheg + Pythia 8	Powheg + Herwig 7	μ_R, μ_F, ISR/FSR, PDF
Diboson	Sherpa 2.2.11	Powheg + Pythia 8, Sherpa 2.2.1	μ_R, μ_F, PDF, electroweak corrections
$t\bar{t}$ and single top	Powheg + Pythia 8	Powheg + Herwig 7, MadGraph 5_aMC@NLO + Pythia 8	Additional ISR/FSR, DS/DR (single-top Wt)
V+jets	Sherpa 2.2.11	MadGraph FxFx Sherpa 2.2.11	μ_R and μ_F, PDF, electroweak corrections

Table 4.13. Summary of samples in the analysis, their nominal and alternative generators and systematic effects used to assess signal and background modelling uncertainties.

One of the main results of this analysis is the fiducial measurements in STXS bins as shown in the scheme in Figure 4.28. It provides 8 POI for the ZH, splitted in number of additional Jets, and 5 POI for WH processes, both divided in $p_T^{V,t}$ bins.

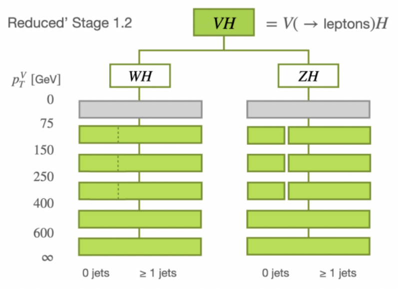


Figure 4.28. STXS scheme used in this analysis.

The most important uncertainties are the QCD scale variations coming from the factorization and renormalization scale variation (μ_F and μ_R). Then a set of uncertainties related to the PDF+α_s are also considered. Electroweak corrections are also taken into account from NNLO not considered contributions. Uncertainty on the Branching Ratio is considered with about 1.61% on the $H \to b\bar{b}$ and $^{+5.53\%}_{-1.99\%}$ on the $H \to c\bar{c}$.

Background

In the following, a set of modeling uncertainties is described for each background process.

W+jets This background is one of the main backgrounds in the 1-lepton channel. As already mentioned, this background can be divided depending on the flavor composition of the jets. W+hf defines events with heavy-flavored jets, either W+bb or W+cc. W+mf defines events with mixed-flavored jets (W+bc, W+bl, W+cl, W+bτ, and W+cτ). W+lf defines all the other processes with light-flavor jets or τ.

For the $H \to b\bar{b}$ *resolved* and $H \to c\bar{c}$ cases, W+hf, W+mf, and W+lf normalization factors are directly estimated in the fit via the Control Regions with floating parameters. These parameters are decorrelated in p_T^V and the number of jets.

A set of acceptance uncertainties is considered. These kinds of uncertainties are estimated as a double ratio between different Monte Carlo samples and different regions. In particular, these are used to extrapolate the uncertainty from a high-purity phase space to a low-purity phase space. These are mostly applied in the 0-lepton and 1-lepton channels where the W+jets background is mostly present. Lepton channel extrapolations from 1-lepton to 0-lepton range from 3% to 30% among the various W+jets flavor compositions. Given the complexity of the flavor composition, acceptance uncertainties are applied to W+cc from W+bb (Wcc/Wbb ratio) ranging from 4% to 20% in the $H \to b\bar{b}$ depending on the p_T^V and the number of jet categories and an overall 25% on the $H \to c\bar{c}$. Other flavor ratios acceptance uncertainties are applied: W+bc for the Wbc/Wcl ratio ranging from 14% to 24%; W+bl for the Wbl/Wcl ratio ranging from 12% to 29%. W+c(b)τ for Wc(b)tau/Wcl ratio of about 20%. W+l(τ)τ for Wl(tau)tau/Wl ratio of 9%.

Then a set of extrapolation uncertainties is considered between SR and CR. These uncertainties range from 3% to 65%.

Finally, an extra acceptance uncertainty is applied in the 4-jet category in the 0-lepton channel ranging from 12% to 20%.

Also, four kinds of shape uncertainties are considered:

- CARL: Firstly, shape uncertainties are estimated with a Calibrated Likelihood Ratio Estimator (CARL) algorithm that is a reweighting technique based on a Neural Network that is trained to distinguish between the nominal and alternative Monte Carlo generators. For W+jets, CARL evaluates the differences between SHERPA 2.2.11 and MADGRAPH FXFX samples.

- SHERPA 2.2.1: The nominal sample seems to disagree with data in the p_T^V shape more than the SHERPA 2.2.1. SHERPA 2.2.1

uncertainties are derived to reweight Sherpa 2.2.11 to Sherpa 2.2.1 p_T^V distribution.

- QCD Scale Variations: Similarly to signal samples, μ_F and μ_R variations are considered as shape uncertainties.
- Electroweak variations: This is a minor shape uncertainty. It's a multiplicative correction added as a shape uncertainty.

For the $H \to b\bar{b}$ *boosted* regime, only W+hf is floated while W+mf and W+lf have 36% and 38% uncertainties, respectively.

A series of acceptance uncertainties are applied on flavor composition, ranging from 9% to 15%, on lepton channel uncertainty, about 20%, on p_T^V extrapolation, around 3%, and SR-CR acceptances, ranging from 16% to 27%.

The same shape uncertainties are considered as for the $H \to b\bar{b}$ resolved case.

Z+jets The Z+jets background mainly contributes to the 0-lepton and 2-lepton channels. This background, similarly to the W+jets, is divided into different flavor compositions. Z+hf is defined for the heavy-flavored jets for Z+bb and Z+cc. Z+mf is for mixed-flavor jets composed of Z+bc, Z+bl, and Z+cl. All the remaining will compose the Z+lf.

Similarly to W+jets in $H \to b\bar{b}$ *resolved* and $H \to c\bar{c}$, a normalization factor is defined for Z+hf, Z+mf, and Z+lf for each p_T^V bin and number of jet categories.

In the Z+jets case, the lepton acceptance uncertainty is extrapolated from 2-lepton to 0-lepton with an uncertainty ranging from 2% to 10%.

Flavour composition uncertainties are then estimated for Zbb/Zcc ratio, Zbc/Zcl ratio, and Zbl/Zcl ratio with uncertainties ranging from 4% to 12%.

Extrapolation between SR and CR is also considered, ranging from 5% to 30%.

The same shape uncertainties as for the W+jets are used for the Z+jets.

For the $H \to b\bar{b}$ *boosted* regime, Z+hf is left floating while Z+mf and Z+lf have 35% uncertainty.

Similarly to the W+jets case, flavor acceptance uncertainties are applied ranging from 6% to 9%. A 3% lepton channel uncertainty is also considered. The p_T^V extrapolation is considered with 15%

uncertainty and the SR-CR acceptance uncertainty ranges from 15% to 25% for different flavor components.

The shape uncertainties considered are the same as in the $H \to b\bar{b}$ *resolved* case.

Top The main backgrounds related to top quark production in this analysis are $t\bar{t}$ and single-top production. Single-top production can be further divided into three samples: single-top s-channel, t-channel, and Wt. This background is mainly estimated for the 0-lepton and 1-lepton channels, while for the 2-lepton channel, a data-driven method is used. In the 2-lepton channel, an extrapolation of 0.8% in the $H \to b\bar{b}$ *resolved* case is applied. For the $H \to c\bar{c}$ regime and the $H \to b\bar{b}$ *boosted* regime, the Top background is a minor component, hence normalization uncertainties are applied, ranging from 4.4% to 25%.

In the rest of the paragraph, the Top background uncertainties are associated with the 0-lepton and 1-lepton channels. $t\bar{t}$ and single-top Wt are treated as correlated into a Top process for $H \to b\bar{b}$ *resolved* and $H \to c\bar{c}$ regimes to minimize the sensitivity to their interference.

Two separate normalizations are considered, one for Top(bb) and Top(bq/qq) derived respectively in CRHigh and Top(bc) Control Region. These normalization factors are divided into p_T^V bins and the number of jet categories.

Top background samples can be divided into 3 components based on flavor: Top(bb), Top(bq) for Top(bc), and Top(bl); Top(qq) merging Top(cc), Top(cl), and Top(ll). Acceptance ratios are applied between bl and bc components and between qq and bc with uncertainties of 5% and 10%, respectively. Acceptance uncertainties are applied also between $t\bar{t}$ and single-top Wt processes, ranging from 12% to 48%.

Lepton channel acceptance uncertainties are also applied from 1-lepton to 0-lepton with uncertainties ranging from 2% to 8%.

A series of acceptance uncertainties between SR and CR are applied, ranging from 2% to 10%.

Similar acceptance uncertainties are applied to s-/t-channels, which are minor Top backgrounds.

For Top processes, the following shape uncertainties are considered:

- CARL: It is used to consider the two-point systematic between the nominal POWHEG + PYTHIA 8 with the two alternative samples, one for the Matrix Element variation (MADGRAPH5 AMC@NLO) and one for the Parton Showering (HERWIG 7). For the single-top Wt sample, the only shape uncertainties estimated with CARL are for the Diagram Removal sample.

- single-top Wt DR-DS: To estimate the interference between single-top Wt and $t\bar{t}$, an additional uncertainty is applied between Diagram Removal (nominal) and the alternative Diagram Subtraction. This uncertainty is considered as both shape and normalization.
- ISR and FSR: Additional shape uncertainties are considered for Initial/Final State Radiations using alternative samples.

In the $H \to b\bar{b}$ *boosted* regime, a normalization for $t\bar{t}$ is considered, and channel and SR-CR acceptance ratios are considered, ranging from 6% to 20%.

For the single-top processes, a 25% normalization uncertainty is considered for the Wt channel, while 10% and 4.6% normalization uncertainties are considered for the s-/t-channels, respectively.

Lepton channels, SR-CR, and p_T^V acceptance ratios are also taken into account, ranging from 20% to 40%.

The same considerations hold for the shape uncertainties as for the $H \to b\bar{b}$ *resolved* scenario.

Diboson This resonant background has a final state similar to the signal one, with the Higgs boson substituted with a Z boson. Diboson backgrounds consist of WW, WZ, and ZZ samples.

Main acceptance uncertainties are computed from generator differences, QCD scale variations, electroweak corrections, α_s variations, and PDF variations.

Normalization uncertainties ranging from 16% to 30% are applied to diboson samples in the $H \to b\bar{b}$ *resolved* and $H \to c\bar{c}$ regimes, while ranging from 17% to 27% in the $H \to b\bar{b}$ *boosted* regime.

Similarly to the V+jets background samples, lepton, p_T^V, number of jets, and SR-CR acceptances are considered, ranging from 3% to 40%.

The same kinds of uncertainties are considered for the diboson background as for the V+jets: CARL to reweight SHERPA 2.2.11 to SHERPA 2.2.1, QCD scale variations, PDF shape uncertainties, and electroweak corrections.

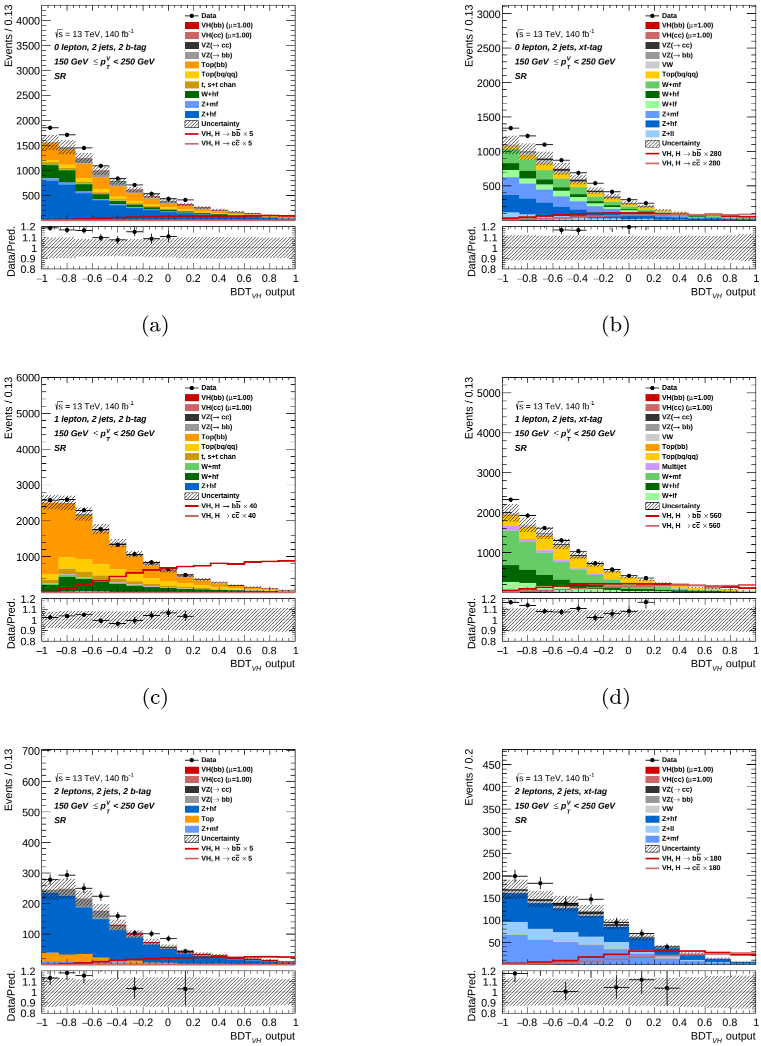
A summary of the modeling uncertainties for each sample and regime can be displayed in Table 4.14.

Uncertainties	*$H \to b\bar{b}$ resolved, $H \to c\bar{c}$*	*$H \to b\bar{b}$ boosted*
VH		
VH normalisation/acceptance		
$H \to b\bar{b}$ BR	1.61%	
$H \to c\bar{c}$ BR	$^{+5.53\%}_{-1.99\%}$	
Diboson		
WW/ZZ/WZ normalisation	16%/17%/19%	16%/17%/27%
ggVV normalisation	30%	30%
Lepton channel acceptance	2-23%	7%
Njet acceptance	10%-30%	-
pTV acceptance	3%-16%	8%-40%
SR/CR acceptance	6%-16%	-
STXS like binning acceptance		
Z+jets		
Z + hf normalisation	Floating	Floating
Z + mf normalisation	Floating	35%
Z + lf normalisation	Floating	35%
Z + hf flavour composition ratio	8%-12%	6%-9%
Z + mf flavour composition ratios	4%-10%	6%-9%
Z + lf flavour composition ratios	5%-30%	-
topCR-SR extrapolation ratio	2%-10%	15%-25%
2- to 0-lepton acceptance ratio	2%-10%	3%
pTV extrapolation	-	15%
W+jets		
W + hf normalisation	Floating	Floating
W + mf normalisation	Floating	36%
W + lf normalisation	Floating	38%
W + hf flavour composition ratios	4%-25%	11%
W + mf flavour composition ratios	14%-29%	9%-15%
W + lf flavour composition ratios	9%	-
High/Low-AR CR-SR extrapolation ratios	2%-63%	-
topCR-SR extrapolation ratio	- 16%-27%	
1- to 0-lepton acceptance ratio	3%-30%	20%
pTV extrapolation	-	3%
Njet extrapolation	12%-20%	-
Top (0- and 1-lepton resolved)		
Top(bb) normalisation	Floating	-
Top(bq/qq) normalisation	Floating	-
Flavour acceptance ratios	5%-10%	-
1- to 0-lepton acceptance ratio	2%-8%	-
High/Low-AR CR-SR extrapolation ratios	2%-10%	-
Wt / tt ratio	12%-48%	-
Top (2-lepton resolved)		
Normalisation	Floating ($H \to c\bar{c}$), 0.08% ($H \to b\bar{b}$)	-
Single-top s/t-channel (0- and 1-lepton resolved)		
Normalisation	4-10%	-
High/Low-AR CR-SR extrapolation ratios	3%-17%	-
pTV extrapolation ratios	7%-15%	-
Njet acceptance	15%	-
1- to 0-lepton acceptance ratio	6%	-
tt and single-top (boosted)		
tt normalisation	-	Floating
single-top normalisation	-	4.6%-25%
1- to 0-lepton acceptance ratio	-	40%
topCR-SR acceptance ratio	-	20%
single-top pTV extrapolation	-	20%
Multi-jet (1-lepton)		
Normalisation	20%-100%	-

Table 4.14. Summary of uncertainties in the resolved and boosted $H \to b\bar{b}$ and $H \to c\bar{c}$ analyses.

4.3.9 Data/MC SR Pre-fit comparison

In this section, few example of shape and normalization comparison between Data and Monte Carlo Post-fit. In Figure 4.29 Pre-fits for the SR in the medium-p_T^V bin and 2-jet category for the BB and XT tag regions.



(a) (b)

(c) (d)

(e) (f)

Figure 4.29. SR Pre-fit distribution in the medium-p_T^V ($150 < p_T^V < 250$ GeV) for the 2-jet category and BB region (left) and XT region (right). (a-b) 0-*lepton* channel. (c-d) 1-*lepton* channel. (e-f) 2-*lepton* channel.

4.3.10 Statistical framework

The fit is performed over all analysis regions, comprising 168 regions in total: 59 Signal Regions and 109 Control Regions. The primary objective of the fit is to determine whether the collected data reveals a significant Higgs signal.

This involves testing two hypotheses: the Null hypothesis (H_0), which assumes only background events are present, and the alternative hypothesis (H_1), which assumes both background and signal events are present. By conducting this test, we can either reject H_0 or find evidence supporting H_1. The test employs an estimator called the p-value, which quantifies the probability that the observed data is compatible with the Null hypothesis. A smaller p-value indicates a greater deviation from H_0.

In particle physics, a "5σ" threshold is used to claim a discovery or observation. This means that the observed p-value is more than 5σ away from the center of the Gaussian p-value distribution representing the Null hypothesis. Consequently, the probability that the observed data are compatible with H_0 is less than 2.7×10^{-7}, effectively ruling out the Null hypothesis. When the signal is too small, as in the $H \to c\bar{c}$ case, the Null hypothesis assumes the presence of the signal, and the test is performed to set upper limits on the production rate of the signal, with a threshold set to achieve a 95% Confidence Interval.

The fit is performed by maximizing the likelihood function, defined as follows. Since this is a counting experiment, the Poissonian distribution is suitable for expressing the probability of observing a certain number of events. Thus, a Poissonian term is included in the likelihood:

$$\mathcal{L}(\mu) = \prod_{i \in \text{bins}} \text{Pois}(N_i | \mu s_i + b_i) \tag{4.7}$$

Here, N_i is the number of measured data events in the i-th bin, while s_i and b_i are the expected signal and background events in the same bin according to the simulation. μ is the signal strength and the Parameter Of Interest (POI) that defines the ratio with the Standard Model of the $\sigma \times BR$.

Systematic uncertainties are incorporated into the fit via Nuisance Parameters (NP) $\vec{\theta}$, which affect the signal or the backgrounds: $s_i, b_i \to s_i(\vec{\theta}), b_i(\vec{\theta})$. These enter the likelihood fit, normalized to 1 and centered to 0, as Gaussian penalty terms:

$$\mathcal{L}(\vec{\theta}) = \prod_{\theta\in\vec{\theta}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\theta^2}{2}\right) \tag{4.8}$$

These represent auxiliary measurements. Deviation of the best fit values from 0 is defined as *pull*, while a decrease in variance is defined as *constraint*. Generally, a *constraint* on a NP means that there is sufficient sensitivity from the data to measure that auxiliary NP.

It is possible to have NP without any prior knowledge ($\vec{\tau}$); these NP are referred to as *free-floating* NP and are usually the Normalization Factors of the main backgrounds. The pre-fit values of such NP are set to 1.

Uncertainties related to the Monte Carlo statistics are defined via γ parameters. These parameters affect the background samples as $b_i(\vec{\theta}) \to \gamma_i b_i(\vec{\theta})$. These terms enter the fit as:

$$\mathcal{L}(\vec{\gamma}) = \prod_{i\in\text{bins}} \text{Gaus}(\beta_i|\gamma_i\beta_i, \sqrt{\gamma_i\beta_i}) \tag{4.9}$$

Where $\beta_i = 1/\sigma^2_{rel}$. Overall, the likelihood function is defined as $\mathcal{L}(\vec{\mu}, \vec{\theta}, \vec{\gamma}, \vec{\tau})$. The best fit value is obtained by maximizing the likelihood or minimizing $-\ln\mathcal{L}$. This procedure is implemented via MINUIT [116] implemented via RooFit framework [117] and RooStats toolkit [118].

The likelihood is then used to test the background-only hypothesis ($\mu = 0$ for both $H \to b\bar{b}$ and $H \to c\bar{c}$) using the profile likelihood ratio:

$$\lambda(\mu = 0) = \frac{\mathcal{L}(\mu = 0)}{\mathcal{L}(\hat{\vec{\mu}})} \tag{4.10}$$

Values of the likelihood ratio close to 1 indicate good agreement with the background-only hypothesis, while values close to 0 indicate deviations from the background-only hypothesis.

A statistical test q_μ can be built as the negative log of the likelihood ratio to avoid numerical instabilities from multiplying small numbers:

$$q_0 = -2\ln\lambda(0); \hat{\mu} \geq 0 \qquad q_\mu = 0; \hat{\mu} < 0 \tag{4.11}$$

It is useful to build a p-value to quantify the compatibility of the result with the background-only hypothesis:

$$p_\mu = \int_{q_{\mu,obs}}^{\infty} f(q_0|0)dq_0 \tag{4.12}$$

where $q_{\mu,obs}$ is the observed value in data. From the p-value, the significance Z can be computed as the inverse Gaussian cumulative distribution (Φ^{-1}) of the p-value:

$$Z = \Phi^{-1}(1 - p) \tag{4.13}$$

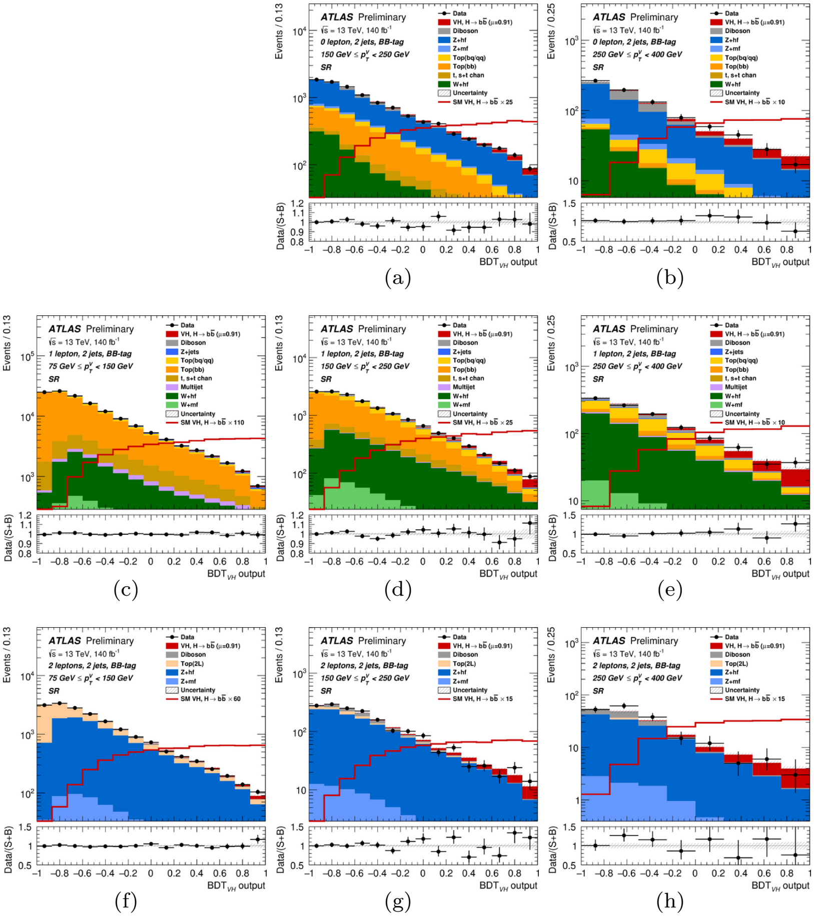
In the $H \to c\bar{c}$ case, an upper limit Confidence Level (CL) of 95% is computed on the signal strength. For the upper limit, a modified frequentist CL_s method is used Ref. [119, 120].

Given the likelihood ratio between $\mu = 1$ (signal+background hypothesis) and $\mu = 0$ (background-only hypothesis), a 95% CL can be inferred by the value of the signal strength μ returning a p-value of 0.05.

To evaluate the expected sensitivity of the analysis, an *Asimov* fit is performed, where the *Asimov* dataset is built as the sum of all the simulated signal and background samples.

4.3.11 Data/MC SR Post-fit comparison

In this section, few example of shape and normalization comparison between Data and Monte Carlo Post-fit. In Figure 4.30 and Figure 4.31 Post-fits for the SR in the 2-jet category for the BB and XT tag regions respectively.

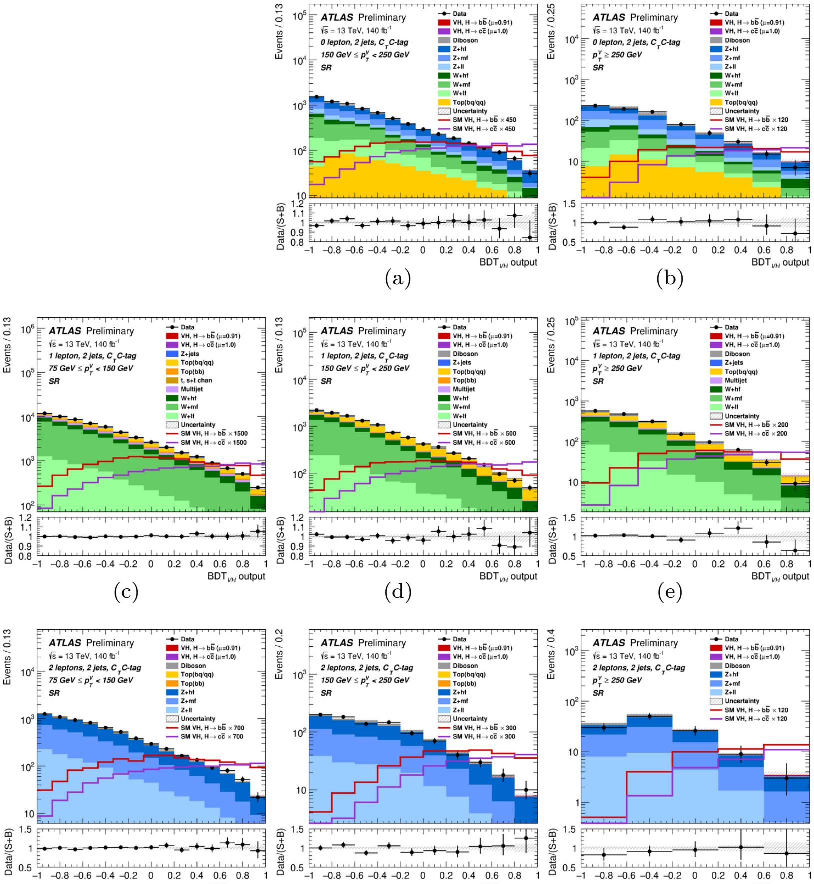


(a) (b)

(c) (d) (e)

(f) (g) (h)

Figure 4.30. The BDT$_{\mathrm{VH}}$ post-fit distributions in the 75 GeV< $p_{\mathrm{T}}^{\mathrm{V}}$ <150 GeV (left), 150 GeV< $p_{\mathrm{T}}^{\mathrm{V}}$ <250 GeV (middle) and 250 GeV< $p_{\mathrm{T}}^{\mathrm{V}}$ <400 GeV (right) signal regions of the Hbb category in the 0-lepton (top), 1-lepton (middle) and 2-lepton (bottom) channel for events with 2 jets. The background contributions after the VH fit are shown as filled histograms. The Higgs boson signal VH,H→bb is shown as a filled histogram on top of the fitted backgrounds normalised to the signal yield extracted from data ($\mu = 0.91$), and unstacked as an unfilled histogram, scaled by a value reported in the legend for better visualisation.



(a) (b)

(c) (d) (e)

(f) (g) (h)

Figure 4.31. The BDT$_{\text{VH}}$ post-fit distributions in the 75 GeV< p_{T}^{V} <150 GeV (left), 150 GeV< p_{T}^{V} <250 GeV (middle) and p_{T}^{V} >600 GeV (right) signal regions of the Hcc XT category in the 0-lepton (top), 1-lepton (middle) and 2-lepton (bottom) channel for events with 2 jets. The background contributions after the VH fit are shown as filled histograms. The VH,H→ canti-c signal and the contribution from VH,H→bb are shown unstacked as unfilled histograms, scaled by the factor indicated in the legend.

4.3.12 Results

VH Results

For a Higgs boson of mass 125 GeV, when all the channels and regions are combined, the fitted values for the VH signal strength are:

$$\hat{\mu}_{VH}^{bb} = 0.91^{+0.16}_{-0.14} = 0.91^{+0.10}_{-0.10}(\text{stat.})^{+0.12}_{-0.11}(\text{syst.}) \tag{4.14}$$

$$\hat{\mu}_{VH}^{cc} = 1.0^{+5.4}_{-5.2} = 1.0^{+4.0}_{-3.9}(\text{stat.})^{+3.6}_{-3.5}(\text{syst.}). \tag{4.15}$$

Both values show good agreement with the SM and their correlation is 5%. The background only hypothesis for the $H \rightarrow b\bar{b}$ case is rejected with 8.0(7.40)σ. This result shows the large improvement with respect to the previous result of expected 6.3σ with the new techniques implemented in this Legacy analysis.

For the $H \rightarrow c\bar{c}$ case, the resulting observed upper limit on the signal strength with 95% CL is 11.3 with an expectation of 10.4. This results, which is compatible with respect to the 7.6(14.4) expected(observed) by the CMS collaboration, also shows the great improvement with the latest result from the ATLAS collaboration of 31(26).

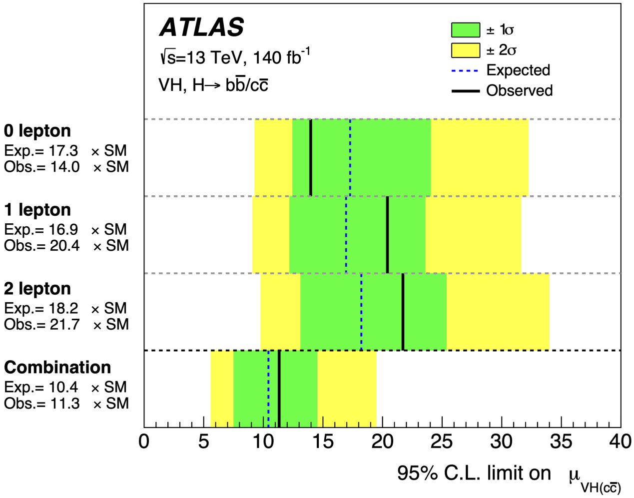


Figure 4.32. 95% CL expected and observed limits on the $H \rightarrow c\bar{c}$ signal strength for each lepton channel and for the combination.

For the $H \rightarrow c\bar{c}$ the signal strength measured, μ_{VH}^{cc}(Leg. Run2) =

1 ± 4(stat.), is compatible, within 1.1σ, with the previous Run 2 result of $\mu^{cc}_{VH}(\text{Run2}) = -9 \pm 10(\text{stat.})$, under the conservative assumption of full statistical correlation. If the systematic uncertainties are also considered the compatibility increases depending on the degree of correlation between the uncertainties in the two analyses. Assuming only statistical correlation between the two analysis, the two measures are compatible within 0.65σ.

It is also possible to show the results when measuring the signal strengths separately for each lepton channel. Table 4.15 shows the best estimation of the $H \to b\bar{b}$ signal strength in each leptonic channel. Figure 4.32 shows instead the expected and observed limits on the $H \to c\bar{c}$ signal strength when fitting a POI for each lepton channel. This result shows how all the lepton channels have similar sensitivities.

	$\mu^{\hat{bb}}_{VH}$	σ Exp.	σ Obs.	$\mu^{\hat{cc}}_{VH}$
0-*lepton*	0.84 ± 0.24	5.71	3.96	-5.68 ± 8.59
1-*lepton*	0.95 ± 0.20	5.57	5.35	3.79 ± 8.8
2-*lepton*	0.92 ± 0.28	4.50	4.13	4.34 ± 8.89
Combination	0.91 ± 0.15	8.0	7.4	0.98 ± 5.30

Table 4.15. Best estimate of the signal strength μ^{bb}_{VH} for each lepton channel and relative Expected and Observed standard deviation from the background only hypothesis and of the signal strength for $\mu^{\hat{cc}}_{VH}$ for each lepton channel.

It is also possible to show the results of the combined fit when measuring the signal strength separately for the WH and ZH production processes. The fitted values for the $H \to b\bar{b}$ signal strengths are:

$$\hat{\mu}^{bb}_{WH} = 0.95^{+0.21}_{-0.19} \tag{4.16}$$

$$\hat{\mu}^{bb}_{ZH} = 0.87^{+0.23}_{-0.20} \tag{4.17}$$

In the WH and ZH production modes the background only hypotheses are rejected with an expected(observed) significance of 5.46(5.26) and 5.67(4.95) respectively. This result provides the first observation of $WH, H \to b\bar{b}$ process.

The main result of this analysis is the STXS measurement with the 13 Parameters Of Interest for the $H \to b\bar{b}$ process described in Figure 4.28. In Figure 4.33 the best estimate of the signal strength for each POI in the fit.

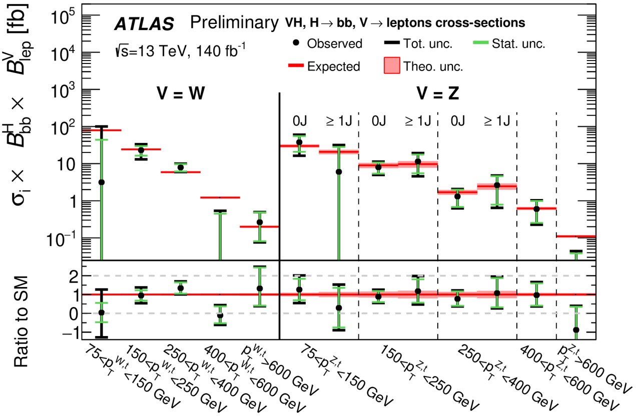
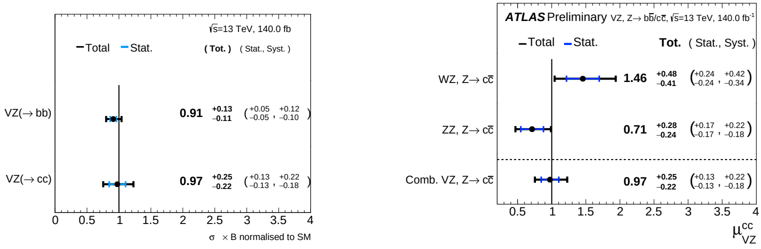


Figure 4.33. Best estimate of the signal strength for each of the 13 POI. It provides the signal strength in each of STXS stage 1.2 bin.

VZ Cross-check Results



(a) (b)

Figure 4.34. (a) Signal Strength for the Diboson VZ process splitted between $H \to b\bar{b}$ and $H \to c\bar{c}$ decay channels. (b) Signal Strength for the Diboson $VZ, H \to c\bar{c}$ process splitted between WZ and ZZ.

It is possible to validate the analysis on diboson production. The signal strength of the processes are:

$$\hat{\mu}^{bb}_{VZ} = 0.91^{+0.13}_{-0.11} \tag{4.18}$$

$$\hat{\mu}^{cc}_{VZ} = 0.97^{+0.25}_{-0.22} \tag{4.19}$$

This result is in perfect agreement with the standard model prediction. The background only (i.e. when the diboson is considered as signal) hypothesis is rejected for both the $H \to b\bar{b}$ and $H \to c\bar{c}$ with an expected(observed) significance of 15.0(13.8) and 5.3(5.2) standard deviation for the two cases respectively. This results in the first observation of the $VZ, Z \to c\bar{c}$ process by the ATLAS detector. It is also possible to decorrelate between the WZ and ZZ processes and the following signal strengths are found:

$$\hat{\mu}^{bb}_{WZ} = 1.00 \pm 0.22 \tag{4.20}$$

$$\hat{\mu}^{bb}_{ZZ} = 0.81 \pm 0.15 \tag{4.21}$$

$$\hat{\mu}^{cc}_{WZ} = 1.46 \pm 0.45 \tag{4.22}$$

$$\hat{\mu}^{cc}_{ZZ} = 0.71 \pm 0.26 \tag{4.23}$$

For the $H \to b\bar{b}$ analysis the background only hypothesis is rejected with an expected(observed) significance of 6.47(6.41) and 11.86(9.93) standard deviations for the WZ and ZZ respectively, while for the $H \to c\bar{c}$ case the background hypothesis is rejected with 2.7(3.9) and 4.2(3.1) standard deviations respectively. These results are shown in Figure 4.34 reporting the first evidence of both $WZ, Z \to c\bar{c}$ and $ZZ, Z \to c\bar{c}$.

Results validation

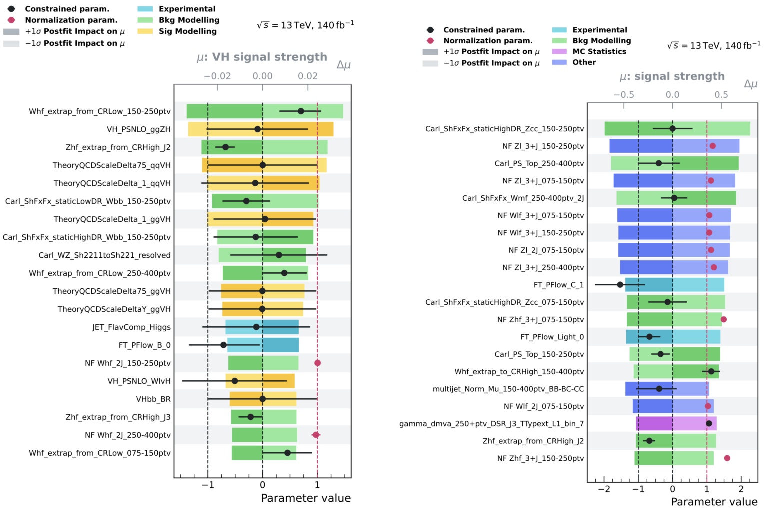
The effects of the uncertainties on the VH, WH and ZH for $H \to b\bar{b}$ and VH for $H \to c\bar{c}$ are displayed in Figure 4.16.

The contribution of each uncertainty is estimated as the difference in quadrature of uncertainty on the signal strength on the nominal fit (σ_μ) and the same uncertainty estimated with all but the interested NP are fixed at their nominal value during an alternative fit ($\sigma_{\mu,i}$) obtaining:

$$\sigma_i = \sqrt{\sigma^2_\mu - \sigma^2_{\mu,i}} \tag{4.24}$$

For both $H \to b\bar{b}$ and $H \to c\bar{c}$ statistical and systematic contributions are similar. The main source of systematic uncertainty is given by the modelling of the Monte Carlo.

The ranking of the impact of the most impacting 20 Nuisance Parameters is shown in Figure 4.35. Each impact is evaluated by the variation of the signal strength as a function of $\pm 1\sigma$ variation on the relative NP.



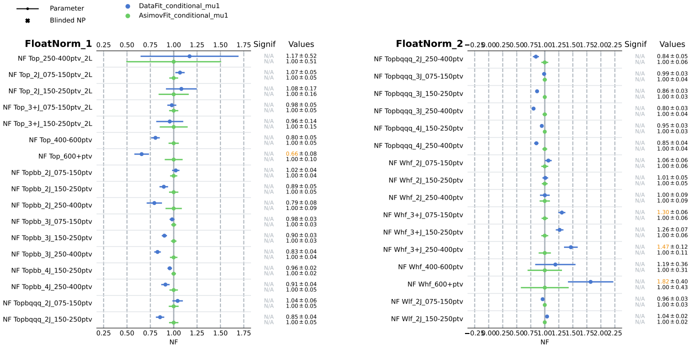
(a) (b)

Figure 4.35. Importance ranking of the impact of individual NP on the signal strengths for the $H \to b\bar{b}$ (a) and $H \to c\bar{c}$ (b) processes.

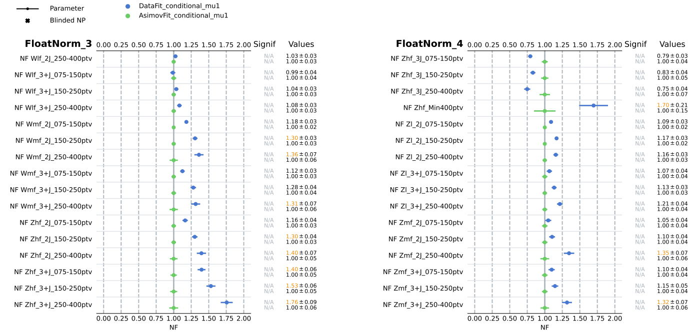
It is possible to show the pull of the Nuisance Parameters to show the auxiliary measurements performed by this analysis. The Normalization Factors are shown in Figure 4.36 where it is possible to observe the clear mis-modelling of the V+jets Sherpa 2.2.11 Monte Carlo as a function of p_T^V . In Figure 4.37 the pull plot relative to CARL systematic uncertainties for the background modelling is reported for the three main backgrounds: Top, W+jets and Z+jets.

POI	$\hat{\mu}^{bb}_{VH}$	$\hat{\mu}^{bb}_{WH}$	$\hat{\mu}^{bb}_{ZH}$	$\hat{\mu}^{cc}_{VH}$
Best estimate	$0.91^{+0.16}_{-0.14}$	$0.95^{+0.21}_{-0.19}$	$0.87^{+0.23}_{-0.20}$	$1.0^{+5.4}_{-5.2}$
Impact on error				
Total	0.151	0.200	0.220	5.29
Statistical	0.097	0.139	0.151	3.94
Systematic	0.116	0.160	0.160	3.53
Statistical uncertainties				
Data statistical	0.089	0.129	0.137	3.70
$t\bar{t}$ $e\mu$ control region	0.009	0.002	0.020	0.06
Background floating normalisations	0.034	0.053	0.040	1.23
Other VH floating normalisation	0.007	0.013	0.007	0.24
Simulation samples size	0.023	0.034	0.030	1.61
Experimental uncertainties				
Jets	0.028	0.039	0.025	1.00
E_T^{miss}	0.009	0.005	0.018	0.24
Leptons	0.004	0.003	0.008	0.23
b-tagging: b-jets	0.020	0.016	0.023	0.30
b-tagging: c-jets	0.013	0.020	0.010	0.73
b-tagging: light-flavour jets	0.006	0.010	0.004	0.67
Pile-up	0.009	0.017	0.003	0.24
Luminosity	0.006	0.007	0.006	0.08
Theoretical and modelling uncertainties				
Signal	0.073	0.066	0.112	0.56
Z + jets	0.039	0.018	0.079	1.76
W + jets	0.055	0.087	0.027	1.41
$t\bar{t}$ and Wt	0.018	0.033	0.018	1.03
Single top quark (s-, t-ch.)	0.010	0.019	0.003	0.15
Diboson	0.032	0.040	0.048	0.51
Multi-jet	0.006	0.011	0.005	0.57

Table 4.16. Breakdown of the statistical and systematic contribution to the signal strength of the VH processes.

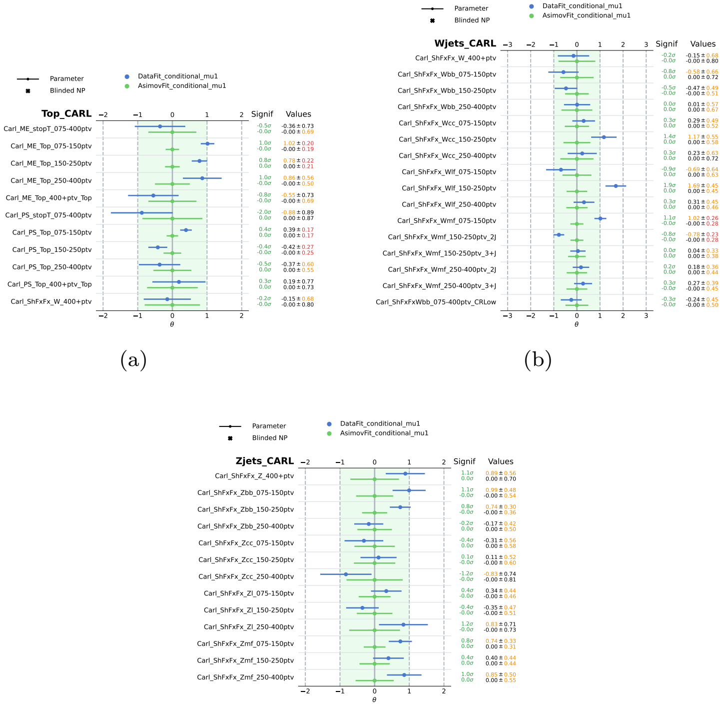


(a)



(b)

Figure 4.36. Post-Fit Normalization Factors for all the backgrounds.



(a) (b)

(c)

Figure 4.37. Post-Fit Modelling systematic uncertainties derived with CARL for the three main backgrounds Top (a), W+jets (b) and Z+jets (c).

5 | A Di-Higgs Analysis

The Spontaneous Symmetry Breaking described in Section 1.2 represents the simplest model and mechanism to provide mass to the particles in the Standard Model. Studying the Higgs self-coupling is fundamental for understanding the Standard Model and exploring Beyond the Standard Model physics. It offers insights into the stability of the Electroweak vacuum and helps determine if the Higgs is the scalar particle predicted by the minimal dynamical symmetry breaking of the Standard Model.

Section 5.1 provides a brief introduction to the Higgs self-coupling and its significance within the Standard Model. Section 5.2 summarizes the state-of-the-art Di-Higgs analysis conducted using the Run 2 data from ATLAS, focusing on events with two *b*-quarks and two photons in the final state which have been conducted in the past by the Collaboration. This channel is one of the most sensitive to the Higgs self-coupling.

Finally, Section 5.3 presents the work which was performed as part of this thesis on the ongoing analysis combining data from Run 2 and partial data from Run 3. This section introduces new ideas and improvements aimed at enhancing the sensitivity of obtained by the Run 2 analysis by enhanced statistics and improved techniques.

5.1 The Higgs Self Coupling

As described in Section 1.2, the Higgs potential is included in the Standard Model Lagrangian as:

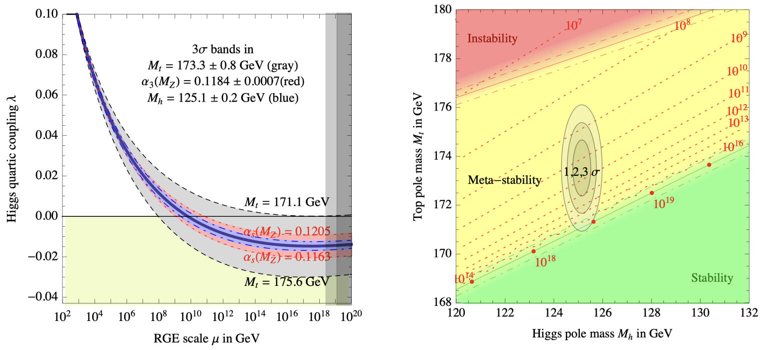
$$V(\phi) = -\mu^2|\phi|^2 + \lambda|\phi|^4 \tag{5.1}$$

When expanding around the vacuum expectation value (VEV v), the potential becomes:

$$V(\phi(x)) \overset{\phi\to v+H(x)}{\longrightarrow} \frac{1}{2}m_H^2 H(x)^2 + \lambda v H(x)^3 + \frac{1}{4}\lambda H(x)^4 \tag{5.2}$$

At tree level, the measurement of the VEV and the Higgs mass gives us the value of the self-coupling $\lambda = \frac{m_H^2}{2v^2} = 0.129$.

Measuring the Higgs self-coupling has significant implications for our understanding of the Higgs potential shape. In particular, the stability of the Electroweak vacuum is ensured by the positivity of λ up to the Planck energy scale ($\Lambda_{\text{Planck}} = 10^{19}$ GeV).

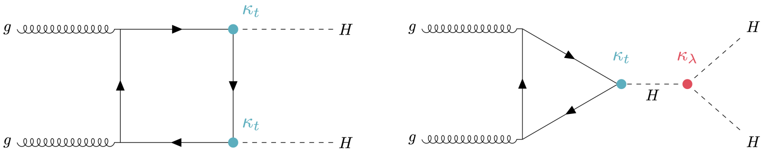


(a) (b)

Figure 5.1. (a) Re-normalization Group evolution of λ for different Higgs and Top masses. (b) Two dimensional diagram of the Electroweak stability as a function of the experimental values of the Higgs and Top masses.

Solving the Re-normalization Group Equations (RGE) it is possible to infer the *running* of the self-coupling, namely the self-coupling as a function of the energy scale. This value is determined essentially by the Higgs mass and the *running* of the Yukawa coupling of the Top quark, which is related to the Top quark mass. From the measured values of the Higgs and Top masses it is possible to establish, solving the coupled RGE, that we live in a meta-stable universe [121, 122].

It means that the lambda coupling is positive at the Electroweak scale but after *running* to the Plank scale, it becomes close to zero as it is shown in Figure 5.1(a). This is shown in Figure 5.1(b) where the near-criticality region is shown as a function of the Higgs and Top masses. This means that if we assume that only the SM exists up to the Plank scale these experimental values prove that the universe has a phase transition with a time scale of at least $\mathcal{O}(10^{65})$ yr [123].



(a) Box Diagram

(b) Triangle Diagram

Figure 5.2. Leading Order ggF diagrams for the Di-Higgs production.

If the SM only is assumed up to the Plank scale, the Higgs and Top quark masses are enough to establish the stability of the Electroweak vacuum. For this reason, a direct measure of the λ can provide insights on the presence of physics Beyond the SM. Moreover the value of λ controls the dynamic of a potential Electroweak phase transition and a measure of it may explain the Baryogenesis. In particular it is notably that the presence of a Baryon asymmetry could be explained by a first-order transition [124].

For this reason, it is crucial to have a measurement of the Higgs self-coupling λ. The best way for a direct measurement is the Higgs boson pair production (Di-Higgs or HH). Figure 5.2 shows the two Leading Order Feynman diagrams for the gluon-gluon fusion (ggF HH) production. Figure 5.2(a) , with the *box* diagram, is proportional to the Yukawa coupling of the top quark, while Figure 5.2(b) , containing the *triangular* diagram, is proportional to the Higgs self-interaction.

In the context of the κ-framework, the vertices are highlighted with κ_t and $\kappa_\lambda = \lambda/\lambda^{SM}$. In the Standard Model case, $\kappa_t = \kappa_\lambda = 1$, and significant deviations from these values indicate that Beyond Standard Model physics may be involved. The two diagrams interfere destructively, as shown in Figure 5.3, leading to an overall cross-section at NNLO of $\sigma^{SM}_{ggF}(m_H = 125\text{GeV}) = 31.05^{+6\%}{-23\%}(\text{scale} + mt) \pm 3\%(\text{PDF} + \alpha_s)$ fb.

In Figure 5.4 it is shown how different values of κ_λ impact the Di-Higgs signal m_{HH} distribution. From the right plot is evident how higher κ_λ values would have a larger impact at low m_{HH} for the ggF HH production. A similar behavior is observed for the VBF HH production.

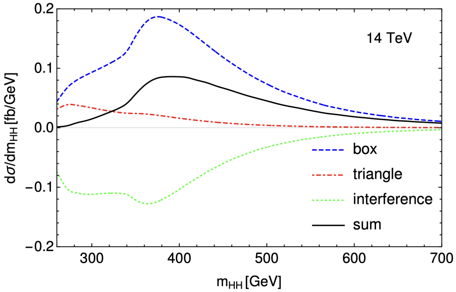


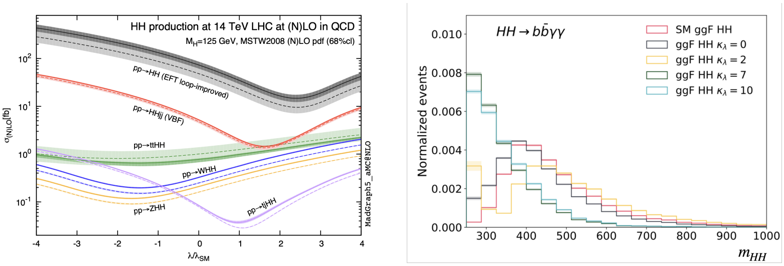
Figure 5.3. Distribution of the Di-Higgs invariant mass system for the two diagrams and their interference. [125]

The contribution of the two diagrams can be described as a function of the Di-Higgs system invariant mass (m_{HH}). At low values of m_{HH} the triangle diagram is dominant and if λ is larger than its SM value the soft m_{HH} spectrum may be enhanced.

At $m_{HH} > 350$ GeV, when the two top quarks can be produced on-shell, the box diagram becomes more relevant. In this phase space, it is easier to test and probe the signal strength of the Standard Model Di-Higgs process (μ_{HH}). In particular, it is possible to test the Di-Higgs against the background-only hypothesis.

In the analysis considered in this dissertation, the second most relevant Di-Higgs production process is also considered: the Vector Boson Fusion production, which has a cross-section computed at N^3LO of $\sigma^{SM}_{VBF}(m_H = 125\text{GeV}) = 1.726^{+0.03\%}_{-0.04\%}(\text{scale} + mt) \pm 2.1\%(\text{PDF} + \alpha_s)$ fb. This production channel is sensitive, apart from the self-coupling, to the $VVHH$ and VVH interactions, which are described in the context of the κ-framework with κ_{2V} and κ_V.

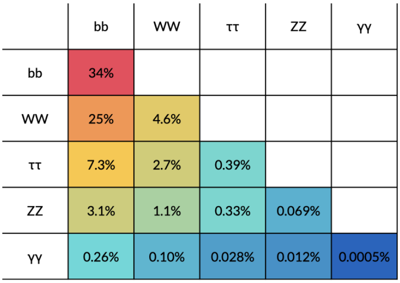
Several final states can be produced with the production of two Higgs bosons. Since the main decay channel of a single Higgs boson is to b-quarks with about 60% of the branching ratio, requiring at least two b-quarks in the final state enhances the overall branching ratio



(a) (b)

Figure 5.4. (a) Cross-section as a function of the κ_λ for the different production channels of Di-Higgs processes.[126] (b) ggF Di-Higgs production m_{HH} distribution for different values of κ_λ.

of the process. Figure 5.5 shows a summary of all the main decay channels with all the combinations of the two Higgs bosons.



	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 5.5. Di-Higgs Branching Ratios.

It is possible to parameterize the ggF signal as a function of κ_λ and κ_t and the two diagrams (T: Triangular, B: Box):

$$\sigma(pp \to HH) = |\mathcal{A}(\kappa_t, \kappa_\lambda)|^2 = |\kappa_t^2 B + \kappa_t \kappa_\lambda T|^2 = \tag{5.3}$$

$$= \kappa_t^4 \left[|B|^2 + \frac{\kappa_\lambda}{\kappa_t}(B * T + BT*) + \frac{\kappa_\lambda^2}{\kappa_t^2} T^2 \right] \tag{5.4}$$

From this expression and with some calculations, it is possible to obtain an analytic formula for the cross-section as a function of κ_λ [127].

5.2 Legacy Run2 H(bb)H($\gamma\gamma$)

In this section, one of the main Di-Higgs analyses performed with the Run 2 dataset using the ATLAS experiment is described [128]. In particular this analysis targets the $HH \rightarrow b\bar{b}\gamma\gamma$ decay channel. This final state benefits from the large $H \rightarrow b\bar{b}$ branching ratio (about 60%), and the low $H \rightarrow \gamma\gamma$ branching ratio (around 0.23%) is compensated by the high efficiency in trigger and reconstruction for photons as well as their excellent energy resolution.

This analysis requires two photons and two b-jets in the final states as candidates for the two produced Higgs bosons. The analysis relies on di-photon triggers with nominal p_T thresholds of 35 GeV and 25 GeV for the leading and sub-leading photon candidates, respectively. The two-photon system is required to have a mass ($m_{\gamma\gamma}$) between 105 and 160 GeV and transverse momenta above 35% and 25% of $m_{\gamma\gamma}$.

Jets are required to be reconstructed from PFlow objects using the anti-k_t algorithm with $R = 0.4$. Jets are required to have $p_T > 25$ GeV, and for those jets with $p_T < 60$ GeV and $|\eta| < 2.4$, the standard Jet Vertex Tagger (JVT) algorithm is applied to match them with the primary vertex. To identify the Higgs candidate decaying as $\mathrm{H} \rightarrow \mathrm{b\bar{b}}$, events are required to have exactly two b-tagged jets using the DL1r algorithm with a nominal efficiency of 77%. Non-tagged jets are ranked based on the Pseudo Continuous BTagging (PCBT) score and then by p_T. Events with leptons in the final state or with more than five central jets with $p_T > 10$ GeV are discarded to suppress $\mathrm{t\bar{t}}$ production and $t\bar{t}H(\rightarrow \gamma\gamma)$.

With these requirements, the signal efficiency of this pre-selection is around 13(9)% for SM ggF(VBF) HH production.

Events are further classified based on the modified m_{HH}:

$$m^*_{bb\gamma\gamma} = m_{bb\gamma\gamma} - (m_{bb} - 125 \text{ GeV}) - (m_{\gamma\gamma} - 125 \text{ GeV}) \tag{5.5}$$

This variable improves the four-object mass resolution, particularly for resonant signal particles decaying into a pair of Higgs bosons, due to detector resolution effects canceling out. The effect is shown in Figure 5.6.

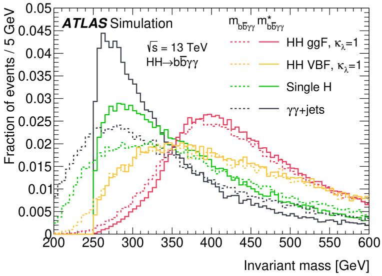


Figure 5.6. Reconstructed four-body mass for the SM HH and single H production processes and for the $\gamma\gamma$+jets background. Dashed lines represent the distribution of $m_{bb\gamma\gamma}$ while solid lines represent the distribution of $m^{*}_{bb\gamma\gamma}$. Distributions are normalized to unit area from Ref. [129].

At low values of m_{HH} ($m^{*}_{bb\gamma\gamma} < 350$ GeV), the analysis is sensitive to large variations in κ_{λ}, while for high values of m_{HH} ($m^{*}_{bb\gamma\gamma} > 350$ GeV), the analysis is sensitive to the SM HH production.

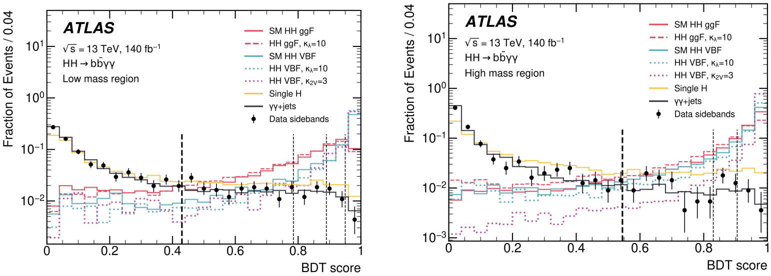
In each of the two categories, indicated as *low mass* and *high mass*, a Boosted Decision Tree (BDT) is trained to distinguish the signals from the various backgrounds of the analysis.

The backgrounds of this analysis can be divided into two categories: the continuous background and the single Higgs resonant backgrounds. Resonant backgrounds include single Higgs bosons decaying to two photons produced by ggF, VBF, in association with a W or Z, with a $t\bar{t}$ or $b\bar{b}$ pair, or single top. For the continuous background, events are denominated as $\gamma\gamma$+jets or $t\bar{t}\gamma\gamma$. All the signal and background samples are produced using Monte Carlo events.

In the high mass region, the signal target samples used are both SM HH events and κ_{λ} variations for VBF production, while in the low mass region, the signal target samples correspond to non-SM HH production with $\kappa_{\lambda} = 10$, 5.6, and the VBF κ_{λ} variations.

The BDT distribution is agnostic to the $m_{\gamma\gamma}$ distribution, which is the final variable used to fit the final parameter of interest (POI). The BDT score distribution is shown in Figure 5.7.

After training, three categories in the high mass region and four categories in the low mass region are defined based on the BDT discriminants. These values are chosen to optimize the overall final significance. The optimization is performed by requiring at least 9 events in the *sidebands* ($m_{\gamma\gamma} \notin [120, 130]$ GeV) in order to have enough statistics to constrain the continuous background from data.



(a) (b)

Figure 5.7. BDT score distribution for simulated signal and background samples for (a) low mass and (b) high mass regions from Ref. [128].

The signal is extracted with an unbinned fit to the di-photon mass ($m_{\gamma\gamma}$) distribution in each category. The signal and background functions are modeled independently with analytical functions.

The signal and resonant background events are modeled by a double-sided Crystal Ball function [130], where the shape parameters are obtained by fitting the simulated SM samples and then fixed in the final fit. The continuous backgrounds are modeled with exponential functions, whose normalization and shape are inferred from data. Each exponential function has two degrees of freedom.

Moreover, the *spurious signal* [131] is defined as the maximum absolute value of the bias on the fitted signal yield in multiple signal+background fits to the background-only template, performed by varying the mass between 123 and 127 GeV in 0.5 GeV intervals. In each fit, the impact of the spurious signal is smaller than 20% of the statistical uncertainty plus twice the statistical uncertainty of the spurious signal itself.

This analysis is statistically dominated, with systematic uncertainty impacting less than 5% on the final result. More details on systematic uncertainties can be found in [128].

The results are derived using a simpler and similar procedure to the one described for the analysis in Chapter 4. A likelihood function is constructed as a function of the parameter of interest (POI, α) and the auxiliary measurements (θ). The likelihood is the product of the

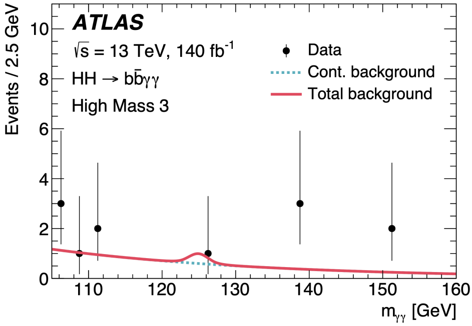
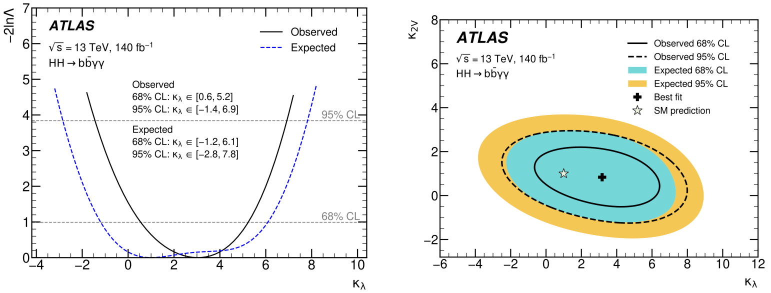


Figure 5.8. Fit on the $m_{\gamma\gamma}$ distribution in the Highest BDT bin for the high mass region. from Ref. [128].

seven independent likelihoods in each BDT category bin and m_{HH} region. Then, a profile-likelihood-ratio test is used to extract the 68% and 95% confidence level (CL) intervals for κ_λ. Upper limits on the Di-Higgs signal strength (μ_{HH}) are also derived. The fit result of a background-only fit is shown in Figure 5.8. No significant excess has been observed, setting a 95% CL upper limit to 4.0 on μ_{HH}. This result must be compared with the expected values of 5.0(6.4) under the background-only hypothesis ($\mu_{HH} = 0(1)$).



(a)

(b)

Figure 5.9. (a) Observed and expected value of the negarive log-likelihood ratio as a function of κ_λ when the other coupling modifiers are fixed to the SM predictions. (b) Likelihood contours at 68% and 95% in the ($\kappa_\lambda, \kappa_{2V}$) plane when all the other modifiers are set to their SM predictions. [128]

From the negative logarithm of the likelihood as a function of the κ values, it is possible to set limits on the κ_λ and κ_{2V} modifiers. For the self-coupling, the expected (observed) limits are $-1.4 < \kappa_\lambda < 6.9$ ($-2.8 < \kappa_\lambda < 7.8$) and $-0.5 < \kappa_{2V} < 2.7$ ($-1.1 < \kappa_{2V} < 3.3$) at 95% confidence level (CL). The likelihood scan for κ_λ and the two-dimensional constraint in the $(\kappa_\lambda, \kappa_{2V})$ plane are shown in Figure 5.9.

5.3 Run2 + partial Run3 Analysis and new techniques

In this Section, preliminary studies for the combination of the Run 2 with a partial Run 3 dataset are shown. The Run 3 is ongoing and this analysis targets to develop new techniques and methods in view of the full Run 2 + Run 3 combination.

One of the main innovations for this analysis will be the improvement coming from the jet Flavour Tagger. This analysis as many of the analyses involving b-quark jets is now starting to use the new advanced GNN based Tagger: GN2, described in Section 3.1.2.

As shown in Figure 3.3 GN2 can improve about a factor of 4 with respect to DL1r in light-jet rejection keeping the same b-tagging efficiency. This reflects automatically in the analysis into a better rejection of the continuous background composed by $\gamma\gamma$+jets where the jets are mis-tagged by the algorithm while keeping a similar nominal b-tagging efficiency. The FTag improvement can be used with a different approach. It is possible to use a looser Working Point with the new tagger that improves the b-tagging efficiency while keeping a similar light-jet rejection compared to DL1r. This would guarantee an increased signal efficiency at the cost of an increased $\gamma\gamma$+bb continuous background.

In the legacy analysis where DL1r with a 77% nominal Working Point has been used the signal efficiency of requiring exactly two b-jets was around 39% which is less compared to the nominal expected $77\%^2 \simeq 59\%$. This effect is due to the fact that even if the average nominal efficiency is 77%, this efficiency is not flat as a function of the p_T of the jet as it can be seen from Figure 3.14. This effect convoluted with the signal p_T distributions of the jets, shown in Figure 5.10, which peaking at lower values leads to a lower efficiency.

A similar behaviour in p_T is expected from GN2. To understand

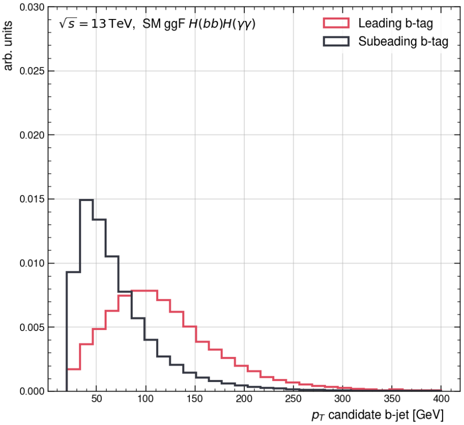


Figure 5.10. Distribution of the p_T of the leading and sub-leading Higgs candidate jets in the ggF HH sample. The distributions peak at p_T values where the b-tagging efficiency is lower than the nominal 77%.

the impact of b-tagging on the signal efficiency two different cases have been studied, keeping in mind that the same study must be performed on the backgrounds to assess the overall impact on the analysis.

A first tested approach is to use a looser Working Point with GN2 that has the similar light-jet rejection as DL1r: 82% Working Point. This approach would enhance the signal efficiency but it would also increase the $\gamma\gamma + bb$ continuous background while keeping the $\gamma\gamma + jj$ approximately unchanged. The second approach would be to use a custom b-tagging Working Point flat in p_T. This can be obtained by a p_T dependent cut on the b-tagging score D_b discussed in the Equation 3.5. To emulate and simplify such a Working Point a Pseudo-Flat approach has been used. A p_T dependent Working Point has been applied. For $p_T > 90$ GeV the b-tagging at 77% Working Point is used while the looser 82% Working Point is applied for jets with $p_T < 90$ GeV.

The results on the signal efficiency for GN2 are shown in Table 5.1.

As mentioned, one of the resonant background for the Di-Higgs analysis is the single Higgs ZH production with the Z boson decaying to b-quarks. This process $(Z(bb)H(\gamma\gamma))$ has not been observed and with new techniques and the Run 3 dataset it could be extracted in a

	HH	
GN2	**Signal Efficiency**	**Variation w.r.t. nominal**
Default 77% WP	14.0%	–
Pseudo-Flat	15.1%	8%
82% WP	15.8%	13%

Table 5.1. Signal Efficiency and variation with respect to nominal for $HH \to b\bar{b}\gamma\gamma$ process.

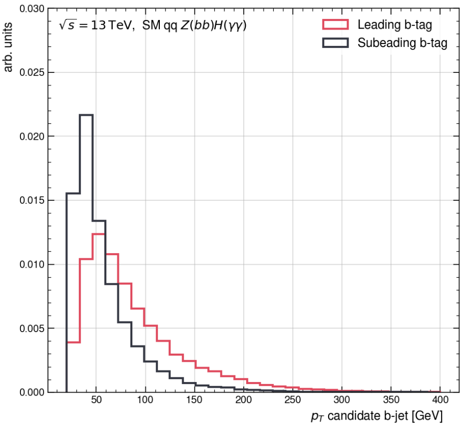


Figure 5.11. Distribution of the p_T of the leading and sub-leading Higgs candidate jets in the ZH sample. The distributions peak at p_T values where the b-tagging efficiency is lower than the nominal 77%.

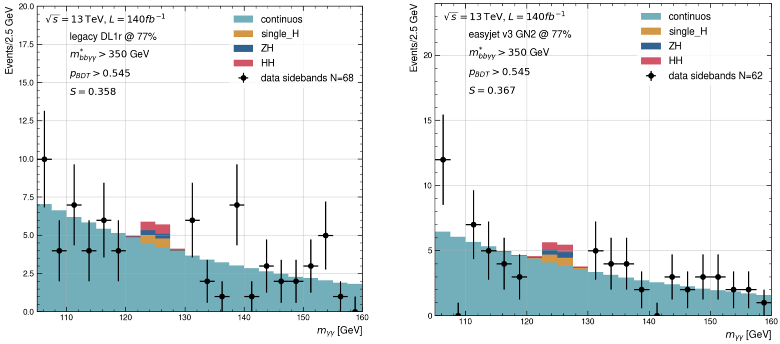
similar phase-space as the Di-Higgs one. As shown in Figure 5.11 the Z has a softer spectrum. For this reason it is possible to apply similar b-tagging studies to the ZH case. The results are shown in Table 5.2 where a larger improvement is obtained by raising the b-tagging efficiency with Fixed-Cut Working Point or with the Pseudo-Flat Working Point.

At the time of this dissertation, the analysis is still under development. Nevertheless, some preliminary results can be provided. It is important to highlight that many samples are yet to be produced, especially for the Run 3 dataset. Currently, the available continuous background lacks sufficient statistics, so a data-driven approach is used to extract the significance.

As previously mentioned, it is well known that the high-mass region is the most sensitive to Standard Model Di-Higgs production. For

	ZH	
GN2	**Signal Efficiency**	**variation w.r.t. nominal**
Default 77% WP	33.3%	–
Pseudo-Flat	38.4%	15%
82% WP	38.0%	14%

Table 5.2. Signal Efficiency and variation with respect to nominal for $ZH \to b\bar{b}\gamma\gamma$ process.



(a) (b)

Figure 5.12. $m_{\gamma\gamma}$ distribution in the high-mass and high BDT score ($p_{BDT} > 0.545$) for the Di-Higgs, Single-Higgs and ZH Monte Carlo samples stacked on top of the *ad-hoc* continuous background extracted from a fit to the Data sidebands. Comparison between (a) legacy distribution and (b) Run 2 reprocessing with GN2 tagger.

this study, the new Run 2 samples are evaluated using the same BDT as in the legacy Run 2 analysis. It is then possible to build a Signal Region by applying cuts on these two variables, requiring $m^*_{bb\gamma\gamma} > 350$ GeV and $p_{BDT} > 0.545$. Since the continuous background has no Monte Carlo sample available with enough statistics, it is possible to exploit the $m_{\gamma\gamma}$ data side-bands: $m_{\gamma\gamma} \notin [120, 130]$ GeV. A two-parameter exponential is fitted in the $m_{\gamma\gamma} \in [105, 120] \cup [130, 160]$ GeV region. This function is then used to sample the one-dimensional $m_{\gamma\gamma}$ continuous distribution with sufficient statistics to ensure that statistical fluctuations do not impact the final result.

From this, the binned significance of the signal compared to the background-only hypothesis can be extracted as:

$$S_i = \sqrt{-2((s_i + b_i)\ln(1 + s_i/b_i) - s_i)} \tag{5.6}$$

This formula is derived from a likelihood ratio Asimov test, assuming that the number of observed events equals the expected number. In the expression, i represents the i-th $m_{\gamma\gamma}$ bin, while s_i and b_i are the numbers of signal and background events in that bin, respectively.

Using this expression, the overall statistical-only significance obtained in the high mass and high BDT score regions is $S = 0.367$. Considering all the applied approximations, this can be compared to the significance obtained with the legacy Run 2 samples using the same selection criteria, which resulted in a significance of $S = 0.358$. This improvement of 2.5% is driven by the reduction of fake b-jets in the lower BDT score bins. Since the b-tagging efficiency of GN2 is the same as the legacy value of DL1r, no improvement is expected in the high BDT bins ($p_{BDT} > 0.905$), where the background is negligible and the signal is not increased. Therefore, it is necessary to use a b-tagging Working Point with a larger efficiency to enhance the signal efficiency while keeping the background at the same level as the DL1r tagger.

Considering the enhanced luminosity ($L_{Run3} = 58.6$ fb^{-1}) collected with the partial Run 3 dataset, it is possible to approximately extrapolate the significance to the combined Run 2 and partial Run 3 data by scaling it by a factor of $\sqrt{1 + L_{Run3}/L_{Run2}} \simeq 1.19$, leading to an overall improvement of 20% in addition to any potential improvement obtained with the GN2 tagger.

6 | Four Dimensional Tracking for High-Lumi LHC

In this chapter a first investigation of the usage of precise timing information in the new generation silicon trackers is presented. The ATLAS full-silicon Inner Tracker (ITk) detector will be installed for the HL-LHC Run4, the main focus of this Chapter is on a potential upgrade of the ITk beyond Run4. The current Inner Tracking system will be upgraded during the Long Shutdown 3 (2026-2029) to cope with the extreme conditions of the High-Luminosity phase of the LHC. ATLAS will also be installing a High-Granularity Timing Detector in the forward region to mitigate the Pile-Up effect. Due to the hard radiation in the proximity of the Interaction Point the innermost layers of the ITk are designed to be replaced after about 2000 fb^{-1} collected data. This represents a unique opportunity to bring in technological innovation and expand the physics potential of HL-LHC by including fast-timing through 4-dimensional (4D) tracking in the ATLAS barrel region.

While HGTD will provide unique handles to improve the reconstruction of physics objects in the forward region, its capability is limited by its reduced η acceptance. There are also compelling physics reasons to consider fast-timing in the central region. In particular, barrel timing information can significantly improve the identification of b-jets, enhancing the prospects to observe Di-Higgs.

The presented work has been recently published as an ATLAS Pub-Note [1].

6.1 Impact of 4D Tracking in ATLAS Beyond Run 4

Picosecond timing information is a ground-breaking new feature to mitigate the impact of Pile-Up and enhance the reconstruction of events at high luminosity hadron colliders. Timing, being uncorrelated with spatial information, allows to disentangle Pile-Up interactions that are very close in space but well-separated in time. This is relevant for the High-Luminosity phase of the LHC (HL-LHC), in which Pile-Up scenarios of about 200 simultaneous interactions are expected [132] and it becomes even more important for future hadron collider concepts, such as the Future Circular Collider (FCC-hh), where the Pile-Up is foreseen to be almost one order of magnitude larger compared to HL-LHC. Both the ATLAS and CMS experiments will incorporate dedicated precision timing detector layers for the HL-LHC upgrade. ATLAS will install a High Granularity Timing Detector (HGTD) in the forward region [133] while CMS will install a MIP Timing detector surrounding hermetically its tracker [134]. The next generation of silicon detectors for charged-particle tracking in such environments is expected to feature 4-dimensional (4D) trackers capable of measuring simultaneously spatial and temporal coordinates [135].

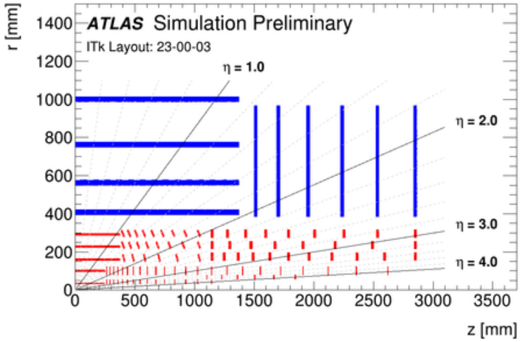


Figure 6.1. The schematic depiction of the ITk Layout 23-00-03 [136] Only one quadrant and only active detector elements are shown. The active elements of the Strip detector are shown in blue, and those of the Pixel detector are shown in red.

The ATLAS Inner Detector for the HL-LHC Run 4 will be upgraded to a full-silicon Inner Tracker (ITk) [137, 138] which has excellent spatial resolution but does not have fast-timing capabilities. The ITk layout for Run4 is displayed in Figure 6.1.

A High-Granularity Timing Detector (HGTD) [139, 133] will be installed in the forward pseudo-rapidity region, providing track timing information at the $30 - 50\ ps$ level.

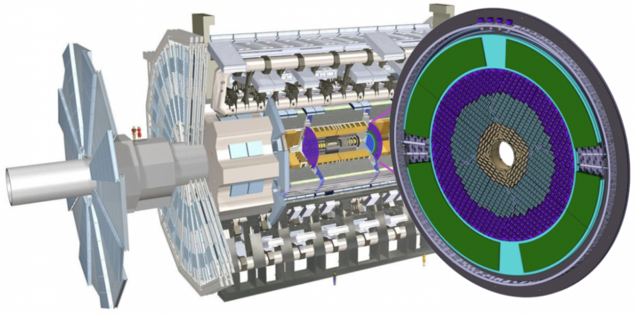


Figure 6.2. Position of the HGTD within the ATLAS Detector. The HGTD acceptance is defined as the surface covered by the HGTD between a radius of 120 mm and 640 mm at a position of z = $\pm 3.5m$ along the beam-line, on both sides of the detector. [139]

While HGTD will help to mitigate the effects of Pile-Up and will lead to gains in specific physics benchmarks such as Vector Boson Fusion (VBF) processes, its coverage is restricted to the forward region, limiting its overall impact to the ATLAS physics programme. In contrast to CMS, whose MIP Timing Detector [134] will provide hermetic timing coverage up to $|\eta| < 3.0$, HGTD only covers the region $2.4 < |\eta| < 4.0$. This limited η acceptance has a non-trivial effect even in the reconstruction of forward jets, typical of VBF signatures. The reason is that the times of forward tracks within jets need to be compared with the time of the hard-scatter primary vertex (t_{HS}) reference time for placing timing selection cuts. The identification of t_{HS} becomes challenging when only forward tracks are available. This means that the ability of HGTD to tag Pile-Up forward jets depends on how many forward tracks originate in the hard-scatter process. A recent study, which can be found in [133], shows that the forward Pile-Up jet rejection with HGTD can improve by more than a factor of 2 with the knowledge of the precise time of the Hard-Scatter as it is shown in Figure 6.3.

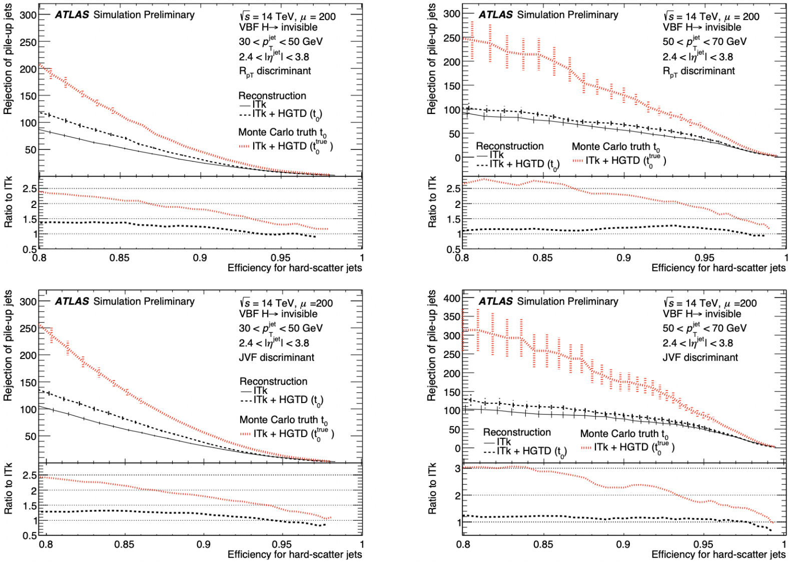


Figure 6.3. The rejection of Pile-Up jets as a function of the efficiency for hard-scatter jets with $30 < p_T < 50$ GeV (left) and $50 < p_T < 70$GeV (right) using the R_{pT} (top) and JVF discriminants (bottom) in VBF $H \to inv.$ events with an average of 200 Pile-Up collisions per bunch crossing. The solid lines are obtained without using the time information. Black dashed (red fine-dashed) lines are obtained using time information and the reconstructed (true) t_{HS} of the event. The error-bands reflect the statistical uncertainty of the samples. Target working points use 85% of efficiency for Hard Scatter jets from Ref. [133].

The two innermost pixel layers of the ITk are anticipated to be replaced after a certain number of years of data taking beyond Run 4 due to the high radiation dose received, offering the opportunity to extend fast-timing information to the ATLAS central pseudo-rapidity region, for hermetic timing coverage.

There are three key opportunities that hermetic 4D tracking could enable:

- improved vertex t_{HS} for forward Pile-Up jet suppression complementing and extending HGTD performance, and for the identification of long-lived particle (LLP) decays with large lifetime $c\tau$.

- improved reconstruction of physics objects in the central region, particularly b-tagging , long-lived particle decays with small $c\tau$, and missing transverse energy.
- improved track and vertex reconstruction (resolution, purity, efficiency, and CPU consumption).

The studies presented in this dissertation aims at quantifying the impact of hermetic 4D tracking coverage in ATLAS for the first time, with emphasis on vertex reconstruction, flavour tagging and subsequential impact on the physics analyses. Low-level investigations on how to improve track reconstruction, resolution, purity, and efficiency, and to reduce CPU consumption by filtering out-of-time hits, are not considered in the following and will be taken into account in longer-term studies.

The ITk detector shown in Figure 6.1 is an all-silicon tracking detector made of two sub-systems: a Pixel Detector [138] surrounded by a Strip Detector [137]. The Strip Detector has four strip double-module layers in the barrel region and six end-cap disks, covering the pseudo-rapidity range up to $|\eta| < 2.7$. The Pixel Detector consists of five flat barrel layers and multiple inclined or vertical ring-shaped end-cap disks, extending the coverage up to $|\eta| < 4.0$. During the last years, the layout of the ITk has undergone several updates compared to the one reported in [138]. The details of the latest version can be found in [140].

The replaceable layers of the Pixel Detector are the innermost ones, located at 34.0 mm and 99.0 mm in the barrel and 33.2 mm, 58.7 mm and 80.0 mm in the end-caps. In terms of radiation tolerance, the innermost part the ITk is designed to collect up to 2000 fb^{-1} of data before being replaced. This corresponds (with a safety factor of 1.5) to a radiation dose of about 10 MGy [138], one order of magniture more than the current ATLAS Inner Detector. In turn, this requirement implies that a key feature for any new 4D tracking technology used for replacing the pixel modules will be radiation hardness [141].

All the following studies are based on Monte Carlo simulations estimated for the HL-LHC. The generation of simulated samples includes in-time Pile-Up pp interactions happening in the same bunch crossing as well as the effect on the detector response due to interactions from bunch crossings before or after the one containing the hard interaction.

The sample used is a $t\bar{t}$ sample produced at the center of mass energy of $\sqrt{s} = 14\ TeV$ to evaluate the performance of vertex reconstruction and b-tagging at an average Pile-Up $\langle\mu\rangle = 200$. This sample was generated using PowhegBox interfaced with Pythia for the Parton Showering. The EvtGen 1.2.0 program [142] was used for the simulation of the b- and c-hadron decays.

6.1.1 The Pile-Up Challenge

The increasing level of Pile-Up through the subsequent runs of the LHC, leading to an average Pile-Up of 200 during HL-LHC, constitutes one of the biggest computational and performance challenges of LHC physics. From a computational standpoint, the large number of simultaneous collisions makes the information to process for each event much larger and complex. For example, the space-point (or hit) combinatorics to reconstruct the trajectory of charged-particles increases enormously. To set the scale, for each event, $\mathcal{O}(10^4)$ charged-particles will be produced during HL-LHC within the detector acceptance, leading to $\mathcal{O}(10^5)$ hits. With the combinatorics scaling roughly as the square of the number of charged-particle tracks, any improvements in the association of hits to the correct track and, in turn, of the tracks to the correct vertex or jet can have a large and broad impact in the LHC programme.

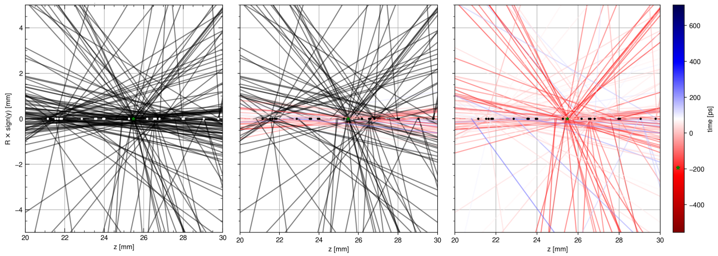
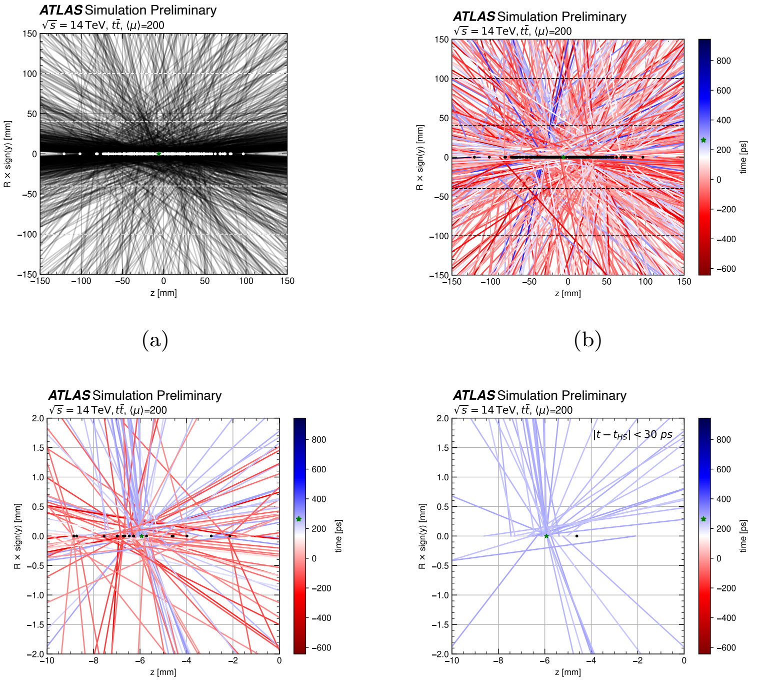


Figure 6.4. Longitudinal view of a simulated $t\bar{t}$ event with $\langle\mu\rangle = 200$ with (left) no track-timing, (center) with track timing in the HGTD pseudorapidity coverage $2.4 < |\eta| < 4.0$ and (right) full hermetic track-timing information coverage.

From a performance point of view, Pile-Up contaminates reconstructed physics objects, such as vertices or jets arising from the hard-scatter interaction, reducing the sensitivity of all physics analyses. While the computational aspects are left for future studies, this work discusses the impact of timing information on ATLAS performance and physics, focusing on the availability of track timing in the central pseudorapidity region.

Assuming the time as a colour associated to for each track it is possible to visualise an event in the three cases where there is no timing detector, with HGTD only and with a full hermetic coverage. This aspect is qualitatively shown in Figure 6.4 where it clarifies why HGTD only is not enough to determine the t_{HS} .



(a) (b)

(c) (d)

Figure 6.5. Longitudinal view of a simulated $t\bar{t}$ event with $\langle\mu\rangle = 200$ with (1(a)) no track-timing compared to (1(b),1(c),1(d)) with track-timing information. The truth hard-scatter is indicated with a green star, while the truth Pile-Up vertices are displayed as solid circles. The reconstructed tracks displayed associated with the truth vertices are displayed. [1]

The association of tracks to vertices becomes challenging in two main situations. The first situation is when the typical separation between Pile-Up vertices becomes comparable or smaller than the longitudinal impact parameter resolution. This is the case for tracks beyond $|\eta| > 2.4$ and this fact motivated the addition of the HGTD to the ATLAS Experiment.

The second situation is the identification of displaced tracks from heavy flavour decays in the central pseudo-rapidity region. In this case, while the longitudinal impact parameter resolution of prompt tracks, z_0 , is significantly smaller than in the forward 0 region, tracks from b-hadron decays can have large impact parameters that are comparable or larger than the separation between Pile-Up vertices. Beyond these two main cases, even the association of tracks to primary vertices can be challenging in the presence of close-by Pile-Up vertices.

As an example, Figure 6.5(a) shows the longitudinal view of a simulated $t\bar{t}$ event with in the ATLAS Inner Detector in absence of timing information. When track-timing information is available, as depicted in Figure 6.5(b) with the usage of different colours to displays tracks belonging to different interactions, the association of tracks to vertices becomes more unambiguous. The concept can be better visualised in Figure 6.5(c) where a zoom on the z axis is performed. In Figure 6.5(d) track-timing information is explicitly used to apply a compatibility time cut between the tracks and the hard-scatter vertex $|t_{trk} - t_{HS}| < 30\ ps$ in order to clean up the environment, which shows very promising potential in the improvement of the track association to the hard-scatter vertex.

6.1.2 Track-time assignment procedure

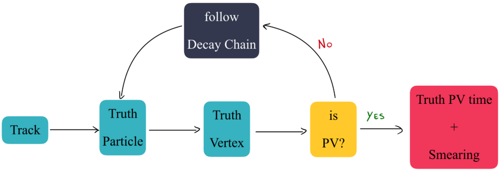


Figure 6.6. Flow chart of the track-time assignment.

The very first step in the evaluation of the impact of 4D tracking on higher level objects is the assignment of time to tracks. As a preliminary procedure, the track-time is assigned from truth level information. Starting from a reconstructed track, the time of the corresponding truth-level information primary vertex is retrieved and the time of that vertex, which is relative to the simulated bunch crossing, is assigned to the track after applying a Gaussian smearing. Several widths are investigated for the Gaussian smearing, to emulate the detector response in terms of time resolution, namely 30 ps, 60 ps and 90 ps. For the small fraction of tracks, of the order of a few percents, which do not have truth information due to simulation and reconstruction effects, a random time is assigned based on the beam-spot time profile. This procedure is schematized in Figure 6.6.

6.1.3 Impact on Object Reconstruction

The addition of precision timing information in the central tracker can have a broad impact on the ability of the ATLAS detector to reduce Pile-Up contamination and improve the reconstruction of physics objects. This section focuses only on two key areas as examples of the potential for improvements: the precise determination of the t_{HS} for all events and the identification of b-jets.

The knowledge of the t_{HS} is a crucial element for the full exploitation of HGTD in the forward region. The limited acceptance of the HGTD makes this determination challenging in events with relatively low forward jet activity. One example are VBF events containing only one jet within the HGTD acceptance. In this case, HGTD can assign a time to the jet, but still needs a reference hard-scatter vertex time from the more central jet in order to be able to either reject or accept this jet. While improvements on HGTD-based algorithms to reconstruct t_{HS} are expected, the availability of the time information for every track in there event will significantly boost the HGTD potential by providing precise t_{HS} containing a single forward jet.

The reconstruction of t_{HS}, in conjunction with the time-stamp of every track, can also provide means to reduce Pile-Up contamination in the reconstruction of the hard-scatter vertex itself. One known limitation of the current vertex reconstruction algorithms under HL-LHC conditions is vertex merging due to primary vertices occurring very close in space. This effect leads to a region of between 0.5 mm and 1 mm where it is not possible to find reliably near-by vertices. Recent proposals for improved vertex reconstruction using machine learning, such as [143] promise to improve this challenge, but are still far from solving it. Track and vertex time information is a very clear way to address this challenge.

Another application of t_{HS} is to improve the calorimeter time resolution, which is limited by the knowledge of the hard-scatter vertex time. This improvement can be exploited to expand the reach to search for long lived particles decaying to displaced photons or jets.

The availability of time information for tracks in the barrel region will have a major impact in the identification of jet flavours, i.e. flavour tagging. The long lifetime of B- and D-hadrons requires the selection of tracks with large impact parameters for the identification of heavy-flavour-quark-intiated jets. The need for large selection windows around the longitudinal impact parameters implies more Pile-Up contamination that can create fake secondary vertices. Therefore, while the HGTD physics case was made on the basis of impact parameter detector resolution being comparable or larger than the expected average distance between primary vertices, the case for barrel track time relies on the physics of B- and D-hadrons decays leading to similarly large impact parameters relative to the vertex separation.

In a second stage a description of the impact of 4D tracking on the ability to identify *b*-jets is studied.

While this Chapter focuses on these two important applications with broad physics impact, the potential of a 4D tracker extends well beyond. For example, pixel hit time information can provide new ways to improve track seeding and reconstruction. Even if only one tracker layer is instrumented with pico-second timing, it might be possible to reject track seeds that are incompatible to each other in time and drastically reduce CPU time. One particularly interesting direction this could enable is the possibility to lower the minimum track transverse momentum threshold of 0.9 GeV in the central region. These investigations are left for future studies.

Determination of t_{HS}

Primary vertex reconstruction in ATLAS is performed using the Adaptive Multi Vertex Finder (AMVF) [77], based solely on spatial information. The AMVF algorithm assigns tracks to multiple vertices, incorporating a weight, w_{trk}, which represents the likelihood of a track belonging to a specific vertex. For this study, only reconstructed vertices which are not purely made up of Pile-Up tracks are considered. The Pile-Up only Hard Scatter vertices are about $\mathcal{O}(1\%)$. The t_{HS} is then computed as the weighted average of the track times:

$$t_{HS}^{reco} = \frac{\sum_{trk} w_{trk} t_{trk}}{\sum_{trk} w_{trk}} \tag{6.1}$$

In the following three cases in the calculation of t_{HS} are considered:

- *All Tracks*, i.e. all the tracks contributing to the vertex;
- *HS only*, i.e. only the tracks that have been produced from the HS vertex;
- *Clustering*, i.e. implementing a time based clustering algorithm.

The *HS only* case makes use of truth matching and is meant to quantify the ideal performance assuming a perfect track-to-vertex association. The clustering algorithm utilized in this task is the Density-based spatial clustering of applications with noise (DBSCAN) [144]. DBSCAN is specifically designed to identify clustered points amidst noise, making it highly suitable for the current application. In this context, it serves as a simplified version of 4D vertexing, where a 1D temporal clustering is applied in addition to the spatially reconstructed vertex obtained from AMVF.

The calculation of t_{HS} is repeated for each of the track resolution scenarios (30 ps, 60 ps, 90 ps).

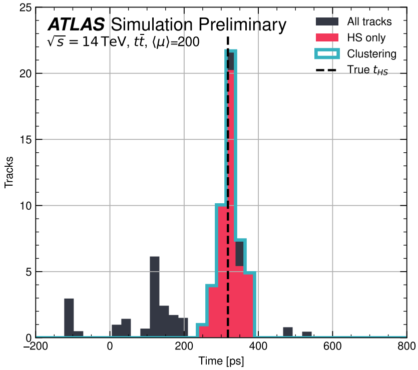


Figure 6.7. Distribution of the track time in the 30 ps track resolution scenario for a single event, for the tracks that belong to a reconstructed vertex. Three cases are shown: *All Tracks*, i.e. all the tracks contributing to the vertex, *HS only*, i.e. only the tracks that have been produced from the HS vertex and *Clustering*, i.e. tracks that belong to a re-clustered vertex. [1]

Figure 6.7 shows an example of a distribution of track time for a reconstructed vertex in the three cases described above, when the track time resolution is emulated to be 30 ps. In the *All tracks* case, one can notice spurious track clusters compared to the *HS only* case. These tracks will affect the time resolution of the vertex if they are included in the computation. The clustering procedure instead removes those spurious contributions, allowing to determine the t_{HS} with a resolution close to that of the *HS only* case.

After determining the t_{HS} , it is possible to estimate its resolution on the HS vertex time with the RMS of the distribution $\varDelta t = t_{HS}^{truth} - t_{HS}^{reco}$. This is a rough estimation because of the non Gaussian tails where the Pile-Up contributes the most.

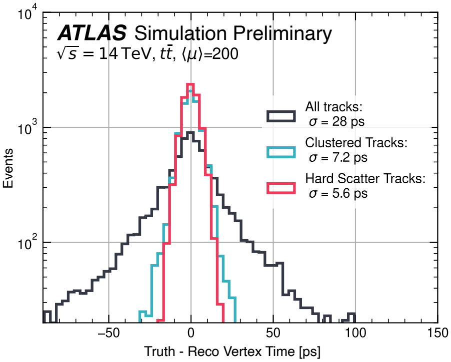
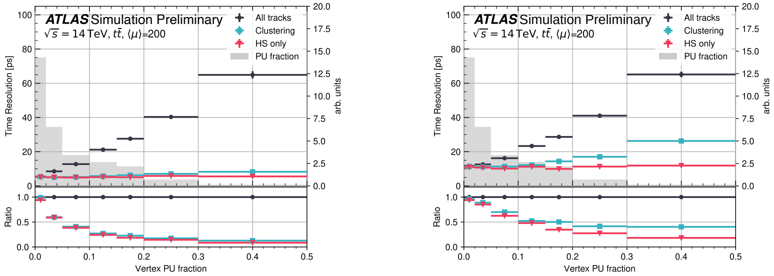


Figure 6.8. Distribution of the t_{HS} resolution defined as the width of a Gaussian distribution fitted on the difference between the truth and the reconstructed vertex time. The 30 ps track resolution scenario is used here. The resolution obtained when using all tracks is compared with that obtained with an *a posteriori* clustering technique as well as with that obtained by using only HS tracks. [1]

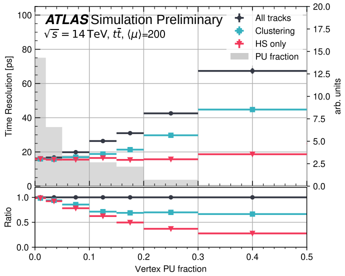
Figure 6.8 shows the improvement of the t_{HS} resolution when introducing the clustering algorithm. When the resolution on the track-time is 30 ps, the t_{HS} resolution would be 28 ps if all tracks were used while it improves to about 7.2 ps when re-clustering the tracks. This is very close to the 5.6 ps obtained when considering only the tracks from the hard scatter, demonstrating that while there is still space for improvement which might arise from the usage of a simultaneous treatment of the space and time coordinates in the vertex finding and fitting procedures, a very good resolution is already achieved.

To better understand the Pile-Up impact on the determination of the t_{HS} resolution, the resolution is studied as a function of the fraction of Pile-Up tracks in the vertex (PU fraction = #PU $trk_{\in vtx}/\#trk_{\in vtx}$), as shown in Figure 6.9



(a) 30 ps resolution

(b) 60 ps resolution



(c) 90 ps resolution

Figure 6.9. Resolution of the t_{HS} in three track resolution scenarios (30 ps, 60 ps and 90 ps respectively for Figure 6.9(a), Figure 6.9(b) and Figure 6.9(c)) shown as a function of the PU fraction per vertex. The resolution obtained when using all tracks is compared with that obtained with an *a posteriori* clustering technique as well as with that obtained by using only HS tracks. The distribution of the PU fraction is also overlaid. [1]

The PU fraction has a smoothly falling distribution and it is binned to obtain almost constant statistics per bin. In each bin, the difference between the truth and the reconstructed vertex time distribution follows exactly a Gaussian function, whose width is used to estimate the resolution.

While the majority of the vertices have low Pile-Up contamination arising from the good performance of the tracking system, a non negligible number of vertices has a sizeable Pile-Up contamination. The ability of removing the Pile-Up contamination depends strongly on the track-time resolution. At 30 ps shown in Figure 6.9(a), the clustering algorithm leads to almost the ideal resolution across the entire PU fraction spectrum. When increasing the track resolution, the improvement in the t_{HS} resolution decreases as shown in Figure 6.9(b) for the 60 ps case and 6.9(c) for the 90 ps case.

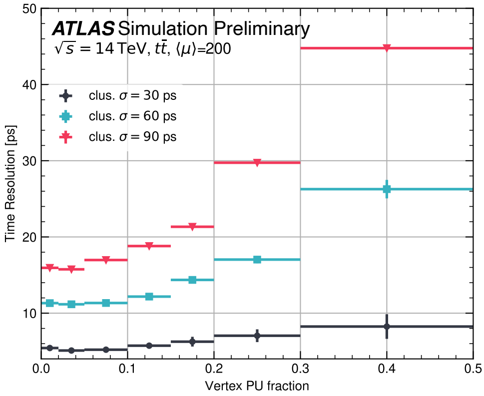


Figure 6.10. Comparison of the t_{HS} resolution obtained with an *a posteriori* temporal 1D clustering technique in three track resolution scenarios 30 ps, 60 ps and 90 ps shown as a function of the PU fraction per vertex. [1]

Figure 6.10 is a comparison of the t_{HS} resolution obtained with the clustering algorithm for the three track resolution cases, as a function of the PU fraction.

Along with the PU fraction, it is possible to study the t_{HS} resolution as a function of two additional variables: the well-established average spatial Pile-Up density $\langle\rho_z\rangle$ and the newly introduced average temporal Pile-Up density $\langle\rho_t\rangle$. In the following formulas σ_z and σ_t are the spatial and temporal standard deviation of the beam-spot.

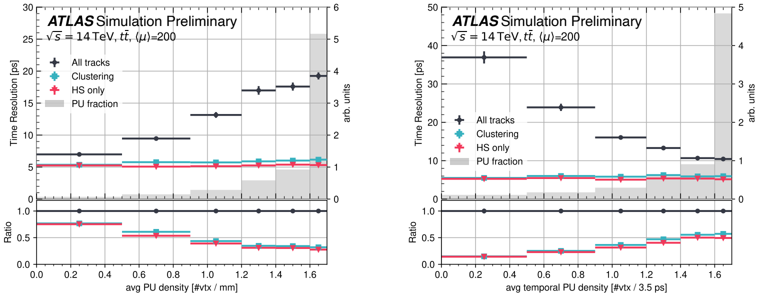
The average spatial Pile-Up density is defined as:

$$\langle\rho\rangle(z_{HS}) = \frac{\langle\mu\rangle}{\sqrt{2\pi}\sigma_z}\exp\left(-\frac{z^2_{\mathrm{HS}}}{2\sigma^2_z}\right) \tag{6.2}$$

The average temporal Pile-Up density is defined as:

$$\langle\rho\rangle(t_{HS}) = \frac{\langle\mu\rangle}{\sqrt{2\pi}\sigma_t}\exp\left(-\frac{t^2_{\mathrm{HS}}}{2\sigma^2_t}\right) \tag{6.3}$$

While these variables provide only averaged information and are not sensitive to the local Pile-Up contamination, they are complementary to the PU fraction defined earlier as they do not depend on the performance of the reconstruction algorithms and, as such, will remain stable should the tracking and vertexing algorithms change.



(a) t_{HS} resolution vs $\langle\rho\rangle(z_{HS})$ (b) t_{HS} resolution vs $\langle\rho\rangle(t_{HS})$

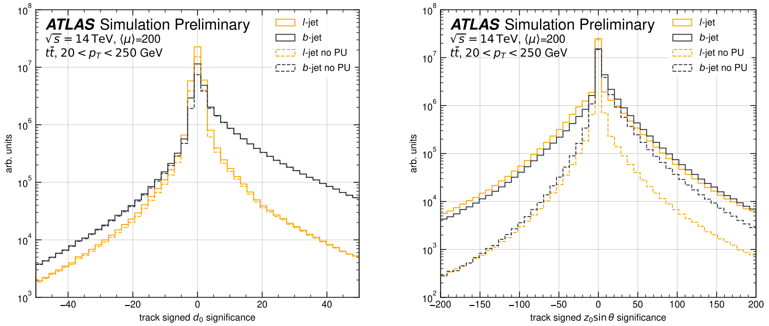
Figure 6.11. Resolution of the t_{HS} in the 30 ps track resolution scenario as a function of two different variables: Figure 6.11(a) as a function of the average spatial Pile-Up density $\langle\rho\rangle(z_{HS})$ and Figure 6.11(b) as a function of the average temporal Pile-Up $\langle\rho\rangle(t_{HS})$ density. The resolution obtained when using all tracks is compared with that obtained with an *a posteriori* clustering technique as well as with that obtained by using only HS tracks. The distribution of spatial and temporal Pile-Up density is also overlaid in the corresponding figures. [1]

The t_{HS} resolution as a function of spatial and temporal $\langle\rho\rangle$ is shown in Figure 6.11 for the 30 ps track-time resolution scenario. The distribution of spatial and temporal average Pile-Up density is also overlaid in the corresponding figures. On average, in the HL-LHC scenario there will be about one vertex every 0.7 mm or every 2.3 ps. In both distributions, there is a flattening of the resolution when the Pile-Up tracks are removed by means of the clustering algorithm. The difference between the two is in the trend of the resolution achieved when all tracks are used. The resolution is worse at high average spatial Pile-Up densities due to the large presence of vertices nearby in the z-coordinate and the increasing amount of mis-associated tracks as a function of $\langle\rho\rangle(z_{HS})$. Instead, the t_{HS} resolution appears to be worse at small average temporal Pile-Up densities. This is because the vertexing algorithm uses only spatial information, making tracks from vertices close in space but very far in time still compatible with the same vertex. The number of mis-associated tracks is roughly flat as a function of $\langle\rho\rangle(t_{HS})$, as expected since time is unused in the current vertexing algorithms. Tracks that belong to vertices that are far in time populate the low $\langle\rho\rangle(t_{HS})$ bins. The contamination from such spurious tracks will have a larger impact on the average vertex time compared to what happens at high $\langle\rho\rangle(t_{HS})$ where the tracks belong to vertices close by in time, therefore not impacting the average vertex time as much as they do in the low $\langle\rho\rangle(t_{HS})$ regime.

Flavour Tagging

As widely discussed in the previous chapters, Flavour Tagging is a crucial technique employed in particle physics experiments to identify the flavour of the initial state parton from which a jet originates. As described in Section 3.1, through innovative experimental techniques and sophisticated algorithms, we are able to successfully tag these particles with high efficiency and accuracy. The state-of-the-art flavour tagging algorithm in ATLAS used in this section is GN1 [85] described in Section 3.1.2 that in the meantime has been superseded by GN2. In this work it is demonstrated that, beside the already high performances that this architecture can reach with the new ITk detector [145], timing information allows to improve the performances considerably. As mentioned in Section 3.18, GN1 uses a one stage Graph Neural Network that takes as input jet and track level information. The output of the algorithm is a set of scores for each jet, which represents the probability that the jet be tagged as a $b-$, $c-$ or $l-$quark-initiated jet. The focus of this study is on light-jets vs b-jets discrimination, but the conclusions are expected to extend also to c-jets similarly.

One of the key features utilised in b-tagging algorithms is the large life time of a B-hadron in the jet, $\tau_B \sim 1.5$ ps. This results



(a) d_0 significance

(b) z_0 significance

Figure 6.12. Track 6.12(a) transverse and 6.12(b) longitudinal impact parameter significance with and without PU for b-jet and light-jets. [1]

in the b-jet having the peculiar signature of a displaced vertex with respect to the beam-line. The tracks arising from this displaced vertex have larger signed longitudinal (z_0) and transverse (d_0) impact parameters, defined with respect to the primary vertex, compared to tracks originating from light-quark-initiated jets. The observables used to capture such features are the impact parameter significances, i.e. the ratio between the impact parameter and their estimated error σ:

- $S(d_0) = d_0 \ /\sigma_{d_0}$ for the transverse impact parameter significance;
- $S(z_0) = z_0 \ /\sigma_{z_0}$ for the longitudinal impact parameter significance.

Being GN1 a track-based algorithm, maximizing the discrimination power of such variables is crucial to better identify light- and b-jets. In the high Pile-Up environment foreseen during HL-LHC, flavour tagging becomes more challenging because pile up tracks mis-associated to hard-scatter jets worsen the discrimination power in the tails of such distributions. Figure 6.12 shows the track impact parameter significance with and without PU for b-jet and light-jets. It becomes evident that eliminating PU contamination can enhance the differentiation among flavours.

Additionally, it is anticipated that the primary influence will be on the longitudinal impact parameter due to the expected minimal displacement of PU vertices in the transverse plane.

To evaluate the impact of track timing on GN1, a temporal track-level variable that will be used as an additional input to GN1 is defined. This new variable is called track-time significance $S(t_{trk})$ and it is defined as follows:

$$S(t_{trk}) = \frac{|t_{trk} - t_{HS}|}{\sigma_t} \tag{6.4}$$

In this equation, σ_t is the track time resolution injected by smearing the truth level information, as explained in Section 6.1.2. The t_{HS} is derived from truth level information with a Gaussian smearing applied based on the vertex studies presented earlier.

Figure 6.13 shows that the distribution of this variable comprises two main components: the tracks originating from the hard scatter, indicated as HS, which will populate the core of the distribution, and the track originating from Pile-Up, indicated as PU, which will have a higher significance.

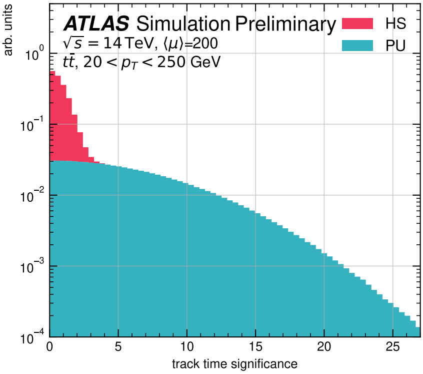


Figure 6.13. Distribution of the track-time significance for tracks that have been truth-matched to the hard-scatter (HS) or to a Pile-Up (PU) vertex. [1]

From now on, the enhanced architecture of GN1 which uses $S(t_{trk})$ is defined as GNT, GN1 with Time information. The other input variables are identical to those defined in [145].

The statistics of the $t\bar{t}$ sample used in this analysis is divided into three parts: a training (28×10^5 jets), a validation (7×10^5 jets) and a test (26×10^5 jets) set. The total statistics of the sample used in this study is equivalent to that used in previous ITk publications [145], but it is limited compared to the results presented in the Run 3 studies, therefore changes in efficiency and rejection can be expected in the future if larger data-sets will be used.

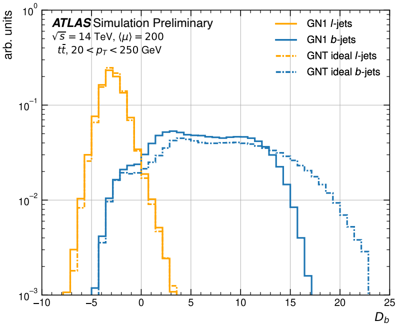


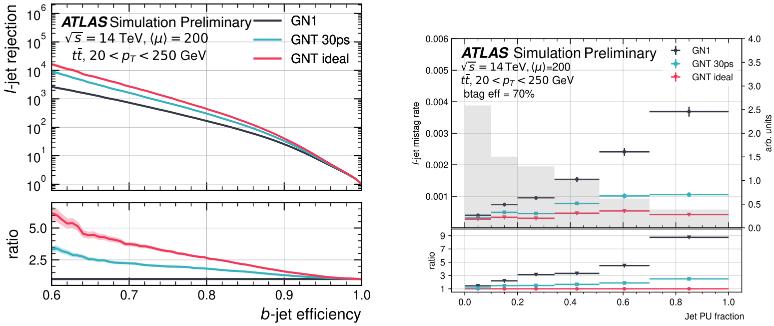
Figure 6.14. Distribution of the b-tagging score D_b for light- and b-jets in a comparison between GN1 vs GNT. [1]

The outputs of the networks, i.e. the three scores (p_b, p_c, p_l) introduced earlier and representing the probabilities for the jet to be tagged with a certain flavour, can be combined to create a global discriminant D_b indicating the probability that a jet is a b-jet defined as follows:

$$D_b = \log \frac{p_b}{f_c \cdot p_c + (1 - f_c) \cdot p_l} \tag{6.5}$$

In this formula f_c is a parameter related to the fraction of c-jets, that can be tuned. For GN1, it is set to $f_c = 0.05$. For GNT, this parameter can be optimized as well. Since the c-jets performances are not the focus of this study, f_c is tuned to have an almost constant c-jets rejection at a given b-jet efficiency working point, $\epsilon_b = 77\%$. A working point is defined by the b-tagging efficiency obtained as the integral of the D_b distribution above a certain cut on the b-jet curve. For each b-tagging efficiency, the corresponding rate of light-jets that are mis-tagged as b-jets can be calculated as the integral of the light-jet curve above the same cut.

Figure 6.14 shows the D_b discriminant for light- and b-jets in a comparison between GN1 vs GNT with no smearing on track time (GNT ideal). It can be noticed that in general D_b takes large values for b-jets and smaller values for light-jets, as expected from its definition, but it is apparent that the discrimination between the two flavours improves when introducing timing information.

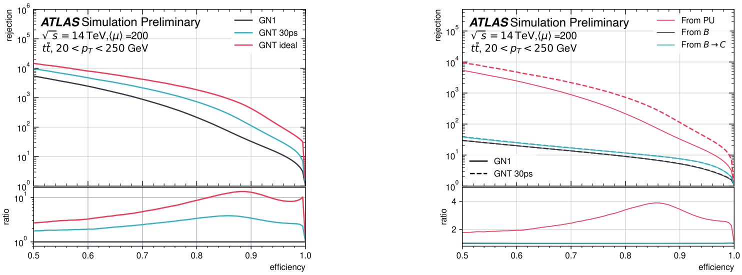


(a) (b)

Figure 6.15. (a) Background rejection as a function of the b-tagging efficiency for different setups of the GNN algorithm: with perfect timing (GNT), with timing information in the 30 ps track resolution scenario (GNT30) and without timing (GN1 std). (b) The light-jet mistag rate as a function of the PU fraction for a fixed $\epsilon_b = 70\%$ b-tagging working point. Different setups of the GNN algorithm are shown: with perfect timing (GNT), with timing information in the 30 ps track resolution scenario (GNT30) and without timing (GN1 std). The distribution of the PU fraction is also overlaid. [1]

A more quantitative assessment of the performance is obtained by studying the light-jet rejection as a function of the b-tagging efficiency. Figure 6.15(a) shows a comparison between three setups of the b-tagging algorithm: with track time and without smearing (GNT ideal), with track time information in the 30 ps track resolution scenario (GNT 30ps) and without timing (GN1). A large improvement arises from the usage of $S(t_{trk})$. For example, for a 77% b-tagging working point, the light-jet rejection improves by a factor of about 3.5 in the ideal case scenario and by a factor of more than 2 in the 30 ps track resolution scenario.

This improvement is analyzed through several figures of merit. For example, Figure 6.15(b), shows the light-jet mistag rate for a fixed $\epsilon_b = 70\%$ b-tagging working point, as a function of the jet PU fraction (jet PU fraction = $\#\text{PU}\ trk_{\in jet}/\#trk_{\in jet}$) defined similarly as in Sect. 6.1.3, but this time by considering the fraction of PU tracks in a jet rather than in a vertex. Different setups of the GNN algorithm are shown: with perfect timing (GNT), with timing information in the 30 ps track resolution scenario (GNT30) and without timing (GN1 std). It can be observed that the light-jet mistag rate as a function of the PU fraction in the jet gets flattened in GNT, with improvements of almost one order of magnitude in the highest PU fraction bins.



(a) PU track classification

(b) PU, FromB and FromBC track classification

Figure 6.16. Performance of the track classification task in GN1 and GNT. In 6.16(a) rejection of non-PU tracks as a function of the efficiency of identifying PU tracks and in 6.16(b) rejection of non-PU or non-FromB or non-FromBC tracks as a function of the efficiency of categorising correctly the tracks within the corresponding category. [1]

One of the auxiliary tasks of the GNN algorithm is the prediction of the origin of each track within the jet. The track is labelled with one of the exclusive categories defined in Ref. [85], after analyzing the particle interaction that led to its formation. These categories include tracks from Pile-Up, indicated as *PU tracks*, as well as tracks from the decay of a B-hadron, indicated as *FromB tracks*, and tracks from the decay of a C-hadron which itself is from the decay of a B-hadron, indicated as *FromBC tracks*. By examining the differences in the labels that tracks are assigned when using GN1 or GNT, as shown in Figure 6.16, we can gain insights into the source of the observed effects. Figure 6.16(a) illustrates that, as expected, track timing information helps mostly in discriminating PU tracks from non-PU tracks. The rejection of non-PU tracks could be improved up to a factor of 10 in the ideal case with no smearing emulating the time resolution, and up to a factor of 4 in the case of 30ps track-time smearing. The classification task in Figure 6.16(b) shows also that while the main improvement arises from a much accurate classification PU tracks, the classification of tracks from heavy flavour hadrons (FromB or FromBC) is mainly unchanged when using track-timing information.

To further understand the requirement on a future 4D tracker, the impact of several track time resolution scenarios on b-tagging efficiency and light-jet rejection is also studied. The results are reported in Figure 6.17. Even in the pessimistic 90 ps track resolution case, for a 77% b-tagging working point, the light-jet rejection improves by a factor of about 1.5.

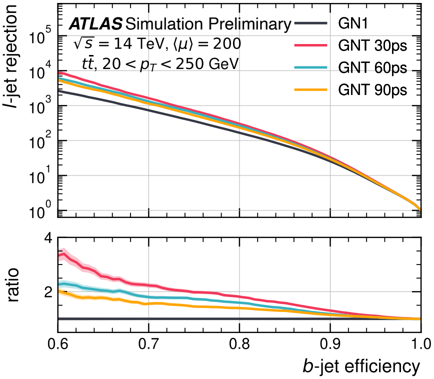


Figure 6.17. Background rejection as a function of the b-tagging efficiency for different track resolution scenarios: 30 (GNT 30ps), 60 (GNT 60ps) and 90 (GNT 90ps), in a comparison with the default algorithm without track timing (GN1). [1]

Two additional effects which are expected to impact the performance of 4D tracking have also been emulated in the case of 30ps track time resolution, i.e. the impact of missing or mis-associated hits-on-track. To study missing hits-on-track, it is assumed that the timing information in the ITk will come from the second innermost pixel layer. If the track analysed has no hits in the second innermost layer, t_{trk} is extracted to emulate the HS profile but its error σ_t is assumed to be 180 ps, i.e. as large as the temporal spread of the beam-spot. This corresponds to saying that no clear discrimination between HS and PU tracks can be made on the basis of timing information. The impact is shown in Figure 6.18. For example, for a 77% b-tagging efficiency working point, a small degradation of about 5% is present in the light-jet rejection when comparing the results with the default

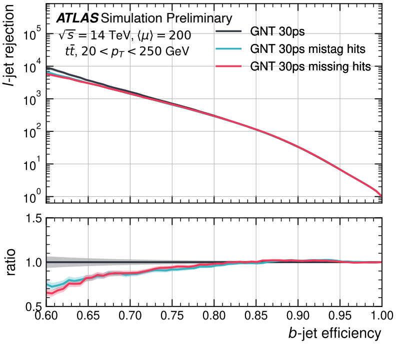


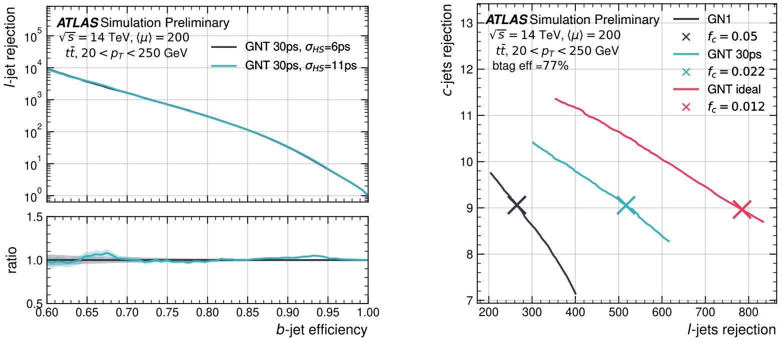
Figure 6.18. Background rejection as a function of the b-tagging efficiency for different setups of the GNT algorithm in the 30 ps track-time resolution scenario: with no assumption on mis-associated or missing hits (GNT 30 ps), emulating tracks which have received a timestamp from mis-associated hits (GNT 30 ps mistag hits), and emulating tracks with missing hits in the second ITk pixel layer (GNT 30 ps missing hits). [1]

performance of GNT at 30 ps. Assessing the impact of mis-associated hits-on-track requires a more complex procedure. While studying these effects in detail is beyond the scope of this proof-of-concept and will require access to fully simulated hits, a simple emulation is carried out in this note to understand at which scale the performance might be influenced. The so-called Truth Match Probability value, or P_{match}, is used. The formula for P_{match} is given below:

$$P_{match} = \frac{2N^{pix}_{common} + N^{strip}_{common}}{2N^{pix}_{track} + N^{strip}_{track}} \tag{6.6}$$

where the denominator contains the number of pixel and strip detector hits attributed to the track, while the numerator contains the number of pixel and strip hits common to both the track and the truth particle. Each track is assigned a P_{match} between 0 and 1. A score of 1 indicates that each hit composing a track was attributed correctly. While P_{match} is not perfect for this study because it includes information from the strip detectors which are not replaceable in HL-LHC, it still provides a good proxy for the number of mis-associated hits and their correlation with other track parameters. By using P_{match} in these studies it is assumed that mis-associated hits are mis-associated in both time and space. The excellent performance of the ITk is confirmed by the large amount of tracks that have a P_{match} score of 1 or very close to 1. In fact, only about 1% of the tracks have a $P_{match} < 0.8$. It is used this cut value and assign a *wrong* (i.e. taken randomly from a PU vertex) time stamp to the tracks which have $P_{match} < 0.8$. The effect of this emulated mis-tagged hits is presented in Figure 6.18. The degradation in performance is comparable with that of the missing hits-on-track case.

Several studies were performed to characterise the flavour tagging results obtained when using track-time information.

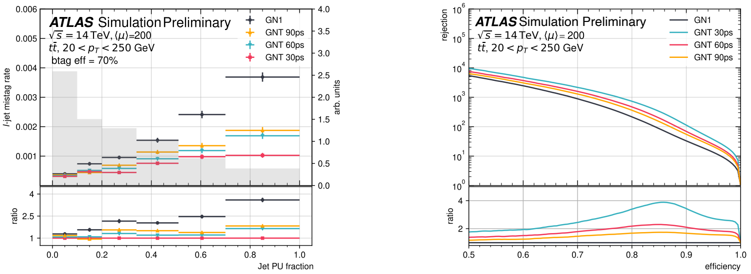


(a) (b)

Figure 6.19. (a) Light-jet rejection as a function of the b-tagging efficiency for two setups of the GNT algorithm: with a track-time resolution of 30 ps and either 6 ps or 11 ps t_{HS} resolution, representing the two extreme values of t_{HS} resolution obtained as a function of the vertex PU fraction shown in Figure 6.10. (b) Scan for the determination of f_c, where GN1 is shown along with GNT ideal and GNT 30ps. [1]

For example, the t_{HS} resolution has a slight remaining dependence on the vertex PU fraction after we apply the *a posteriori* clustering procedure. Figure 6.19(a) shows the impact of different resolutions on the t_{HS} for a given track-time resolution. In particular, in the 30 ps track-time resolution case, the two extreme values of t_{HS} resolution obtained as a function of the vertex PU fraction are either 6 ps or 11 ps (see Figure 6.10). The impact on *b*-tagging performance is negligible.

One of the parameters that can be tuned in the GNN *b*-tagging algorithm is f_c, a tunable parameter to adjust the relative importance of p_c, p_l in the discriminant. Figure 6.19(b) shows a scan for the determination of f_c, where GN1 is shown along with GNT (the ideal case when perfect time resolution is assumed on both tracks and HS vertex) and GNT 30ps (the case with 30ps resolution on track-time and 6ps resolution on t_{HS}). These resolution values are justified from Figure 6.8.

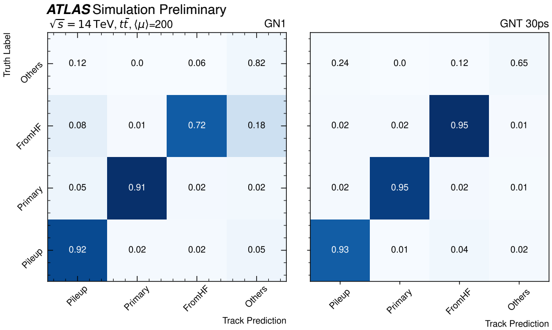


(a) (b)

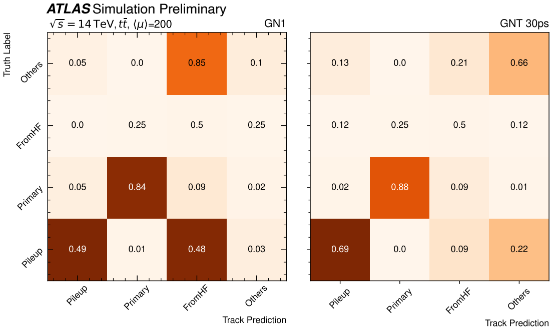
Figure 6.20. (a) The light-jet mistag rate as a function of the jet PU fraction for a fixed $\epsilon_b = 70\%$ b-tagging working point. Different setups of the GNN algorithm are shown: without timing (GN1) and with timing information in the 30 ps, 60 ps and 90 ps track-time resolution scenario (GNT 30ps, GNT 60ps, GNT 90ps). The distribution of the jet PU fraction is also overlaid. (b) Performance of the PU track classification task in GN1 and GNT. The rejection of non-PU tracks as a function of the efficiency of identifying PU tracks is shown for different setups of the GNN algorithm: without timing (GN1) and with timing information in the 30 ps, 60 ps and 90 ps track resolution scenario (GNT 30ps, GNT 60ps, GNT 90ps). [1]

Figure 6.20(a) shows the l-jet mistag rate as a function of the jet PU fraction for a fixed $\epsilon_b = 70\%$ b-tagging working point in different track-time resolution scenarios, while Figure 6.20(b) shows the PU track classification as a result of one of the auxiliary tasks of the GNN b-tagging algorithm. As expected, the l-jet mistag rate and the confusion between non-PU and PU tracks increases with the worsening of the track-time resolution.

Figure 6.21 shows the classification of the tracks as predicted by the GN1 and GNT 30ps algorithms. In particular, 6.21(a) for b-jets in events in which the b-jets migrate from not being "loose" tagged at $\epsilon_b = 85$ % to being "tight" tagged at $\epsilon_b = 60$ %; 6.21(b) for l-jets that are tight tagged as b-jet at $\epsilon_b = 60$ % without time. The matrices are more diagonal in the GNT 30ps case, indicating a superior accuracy of the track classification in the time-assisted algorithm.



(a) Track classification for b-jets



(b) Track classification for *light*-jets

Figure 6.21. Track classification in GN1 and GNT 30ps: 6.21(a) for b-jets in events in which the b-jets migrate from not being "loose" tagged at $\epsilon_b = 85$ % to being "tight" tagged at $\epsilon_b = 60$ %; 6.21(b) for l-jets that are tight tagged as b-jet at $\epsilon_b = 60$ % without time. [1]

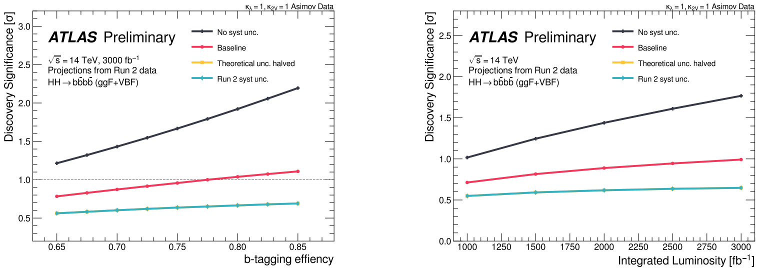
6.1.4 Impact on Higgs boson pairs analysis

The impact of 4D tracking on object reconstruction will largely benefit several areas of the HL-LHC Physics programme. In the following, an example related the searches for Higgs boson pair (HH) production is reported.

As widely discussed in Chapter 5, measuring the Higgs boson self-coupling λ is one of the key point of the LHC and its High Luminosity phase.

Several searches for Higgs pair production have been performed by the ATLAS [146, 147, 148] and CMS Collaboration [149, 150, 151, 152] with the existing data-sets. The current results obtained by combining the three leading final states $HH \rightarrow b\bar{b}b\bar{b}$, $HH \rightarrow b\bar{b}\tau\tau$ and $HH \rightarrow \gamma\gamma b\bar{b}$ [153, 154] show that to observe a HH signal with the Run 2 data, it would have to be about 3 times as large as what the SM predicts. These results have been extrapolated to the HL-LHC data-set by both ATLAS and CMS [155, 156, 157, 158, 154] and show that an evidence for SM-like HH production is feasible by combining the same three final states mentioned above. Additional improvements may arise from including other final states.

Nevertheless, more data or better algorithms are needed to increase the precision with which we will extract λ and draw more quantitative conclusions about the Higgs potential. As HH is a very rare process, most of the analyses look for a signal in a final state in which at least one of the two Higgs bosons decays into its most abundant mode, i.e. in a pair of b-quarks. Therefore improving b-tagging is one of the most powerful ways to improve HH searches.



(a) Discovery significance as a function of the b-tagging efficiency.

(b) Discovery significance as a function of the integrated luminosity.

Figure 6.22. The discovery significance extrapolated from the Run 2 analysis to HL-LHC for the $HH \rightarrow b\bar{b}b\bar{b}$ process in 6.22(a) evaluated for a range of b-tagging efficiency values from 65% to 85% assuming the light-jet rejection rate as the 77% working point used in Run 2 and in 6.22(b) evaluated at various integrated luminosity values ranging from 1000 fb^{-1} to 3000 fb^{-1}. The yellow line overlaps the cyan one in the plots. Details can be found in Ref. [158].

The ATLAS $HH \to b\bar{b}b\bar{b}$ Run 4 extrapolation study [158] considered the impact of potential improvements in b-tagging efficiency in the analysis sensitivity. Figure 6.22 shows the $HH \to b\bar{b}b\bar{b}$ discovery significance. In particular, Figure 6.22(a) shows the discovery significance as a function of b-tagging efficiency values ranging from 65% to 85%, assuming as light-jet rejection rate that of the 77% working point used in the corresponding Run 2 analysis. In the scenario of statistical only errors (No syst. unc.), a 5% increase in b-tagging efficiency (from 77% to 82%) corresponds to an increase in the $HH \to b\bar{b}b\bar{b}$ discovery significance of about 0.3 σ. As shown in Figure 6.22(b), this corresponds to the sensitivity enhancement of more than 500 fb^{-1} of data. Figure 6.15(a) and 6.17 show that this improvement in b-tagging efficiency could be within reach with 4D tracking.

Depending on the point in time at which the innermost pixel layers of the ITk could be replaced with a 4D tracking system, the improvement might decrease (i.e. less time to exploit the new technology), but it must be noted that similar level of improvements will be brought in each one of the HH searches with at least one $H \to b\bar{b}$ decay, making 4D tracking a very powerful way to enhance the reach of HH searches.

The Di-Higgs application is one of the main physics cases where the timing could improve and push the cutting-edge physics. Nevertheless, many other physics analyses would benefit from the timing information in the tracking system. Some examples are reported in [1], focusing on Long Lived Particle searches or VBF Higgs production with $H \to inv.$ decay mode.

6.1.5 Future perspectives

This timing detector investigation with the ATLAS detector is a preliminary study of the feasibility of such a detector.

Track timing information can play a pivotal role not only at the LHC but also in future collider experiments where new challenges will arise from different physical and technological conditions. For this reason perspectives studies are being conducting with detector-agnostic simulations. The A Common Tracking Software (ACTS) [159] tool interfaced with the Open Data Detector (ODD) [160] is used to simulate events.

Within this context, two main projects are being conducted. Although a detailed description is beyond the scope of this dissertation, a brief overview of the two primary studies using the ACTS toolkit is provided here.

The first project involves a comprehensive implementation of 4-Dimensional vertexing, encompassing both finding and fitting techniques. These methods are generally outlined in Section 2.3.1.

The second project investigates the performance of Flavour Tagging on a general detector using the DIPS algorithm, described in Section 3.1.1

7 | ML applications to Global Particle Flow

Artificial Intelligence (AI) emerged in the 1950s with the goal of emulating human learning using machines. In Machine Learning (ML), human-provided input data and predicted answers are used to train algorithms to learn rules. Ideally, an ML algorithm identifies a transformation that converts data into a more useful representation for a given task. Over the years, advancements in computing power have enabled the development of more complex ML structures, known as Neural Networks [161, 162]. ML algorithms are particularly well-suited for tasks in High Energy Physics experiments, where they can efficiently exploit multidimensional correlations between various input variables. This has been demonstrated in analyses where Boosted Decision Trees are used to distinguish between signal and background processes.

In this chapter, we explore another application of ML: detector reconstruction. These algorithms can be used to process low-level detector data, such as calorimeter cells or inner detector tracks, to generate higher-level particle information. We present a Graph Neural Network (GNN)-based algorithm for simultaneous particle reconstruction (Global Particle Flow) in dense environments. Section 7.1 introduces a Configurable Calorimeter simulation to generate the events used in subsequent sections for Global Particle Flow. In Section 7.2, we compare a Hyperedge GNN architecture to other GNN architectures, highlighting the potential impact of these new algorithms. These studies are broadly applicable and demonstrate a strong use case for current and future collider experiments.

7.1 COCOA: A COnfigurable CalOrimeter simulation for AI

In this Section A configurable calorimeter simulation for AI (CoCoA) applications is presented [3], based on the GEANT4 toolkit and interfaced with the PYTHIA event generator. This open-source project is aimed to support the development of machine learning algorithms in high energy physics that rely on realistic particle shower descriptions, such as reconstruction, fast simulation, and low-level analysis. Specifications such as the granularity and material of its nearly hermetic geometry are user-configurable. The tool is supplemented with simple event processing including topological clustering, jet algorithms, and a nearest-neighbors graph construction. Formatting is also provided to visualise events using the Phoenix event display software.

Algorithms incorporating machine learning (ML) methods are a new paradigm in reconstruction, calibration, identification, simulation and analysis of High Energy Physics (HEP) experimental data. In recent years, various ML architectures have been deployed to optimize low-level tasks such as clustering, reconstruction, fast simulation, pileup suppression and object identification [163, 164]. For example, ML-based fast calorimeter simulation relies on accurate target data to train a fast conditional generative model $p(D|T)$, where T denotes the true set of stable final state particles produced in the collision and D is the set of resulting detector hits. In particle reconstruction, on the other hand, the inverse process $D \to T$ is modelled by predicting a set of particles $R(D)$ to approximate T as accurately as possible. The development of such algorithms requires a realistic, highly-granular simulation of particle detector response going beyond parameterized detector models frequently used in studies of particle physics phenomenology such as DELPHES [165]. In particular due to the complexity of particle showers in calorimeters, a detailed, microscopic simulation of interactions between particles and detector material is needed in order to develop low-level ML algorithms exploiting such features.

Recent research efforts to study calorimeter shower properties using ML [166, 167, 2] made use of the GEANT4 [168] simulation toolkit for simple detector geometries. However, an open source detector simulation with a realistic cylindrical geometry and hermetic coverage is yet to be adopted by the HEP community for ML studies and beyond.

Aiming to bridge this gap, the COnfigurable Calorimeter simulatiOn for Ai (CoCoA) was developed [169], which uses GEANT4 [168] to implement detailed shower simulation for particles in a full-coverage, highly-segmented sensitive volume comparable to that of multipurpose detectors at the LHC. CoCoA offers research teams a common, curated benchmark tool for the benefits of straightforward detector configuration and comparability of research results. The program source code [170] is linked together with a technical documentation on the project website[1].

The emphasis of this software package is on realistic calorimeter simulation. No realistic digitization and electronic readouts are implemented and energy loss due to these processes are neglected in this package. For the same reason, simplified tracking is included in CoCoA to model particle deflection in a magnetic field and energy depositions upstream of the calorimeter. A sophisticated open source toolkit suitable for tracking studies based on silicon hits is provided by [171].

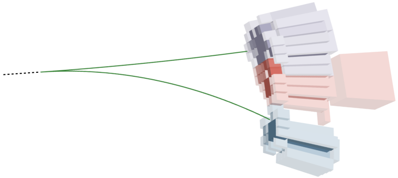


Figure 7.1. Visualization of a photon (dashed line) with energy 50 GeV converting to two electrons (green lines) producing three distinct clusters in the CoCoA central electromagnetic calorimeter. The cluster shown in red contains an additional cell in the first layer of the hadronic calorimeter due to a noise fluctuation. Cells are shown with an opacity proportional to energy over noise ratio divided by 4.6, the threshold for topoclustering seeds.

Usability for ML-based studies is a core motivation in the design of the CoCoA code. Datasets generated by CoCoA have featured in two recent applications of ML to particle reconstruction and fast simulation [172, 4]. To this end, the main parameters of the calorimeters are largely configurable, including their material, granularity, depth and the amount of readout noise. Similarly the inclusion of material interactions in the tracking region is optional.

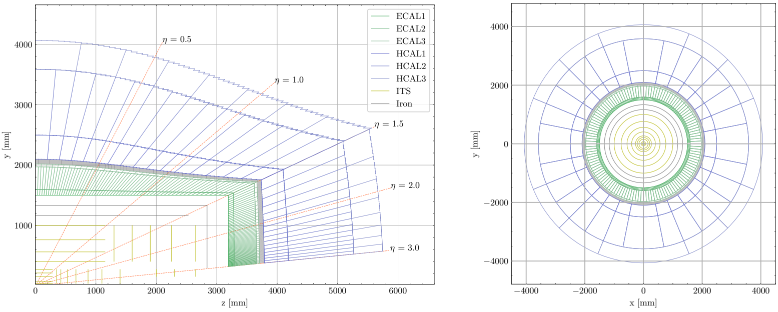
---

[1] `https://cocoa-hep.readthedocs.io/en/latest/index.html`

For comparisons with benchmark reconstruction approaches, output data from CoCoA are conveniently interfaced to standard topological clustering and jet clustering algorithms. The output includes a record of energy contributions to each cell by truth particles for supervising cell-level predictions and edge lists for connecting cells and tracks in a graph to support geometric deep learning models. Finally, the default geometry has been formatted for rendering in the Phoenix event display software [173], along with a script to export event output files for visualization. An example is shown in Figure 7.1.

The sophisticated CoCoA calorimeter simulation and its data post-processing provides users easy access to datasets suitable to train models for current collider experiments or for more general algorithms development and benchmarking. In addition, the open-source nature of the package and its visualization support have the potential for use cases in education and science communication in HEP.

7.1.1 Detector Design



(a) (b)

Figure 7.2. Positive quadrant scheme of CoCoA. Right-handed orthogonal coordinate system x-y-z, where z-axis is the principal axis of the detector and a constant z refers to a circular cross-section of the detector. (a) yz-projection showing the CoCoA ITS, subsequent iron layers, calorimeter system in the barrel and end-cap region, overlaid on lines marking constant pseudorapidity η. (b) xy-projection shows the barrel region of the same subsystems at $z = 0$.

The major components of CoCoA are an inner tracking system (ITS) surrounded by an electromagnetic calorimeter (ECAL) and

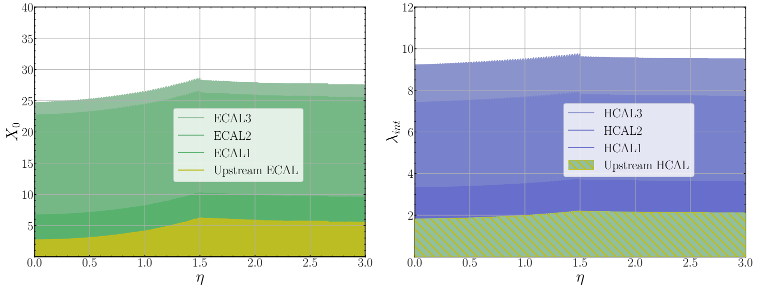


Figure 7.3. (left) Cumulative amounts of material, in units of radiation length X_0 and as a function of η, in front of and within the electromagnetic calorimeter system. (right) Cumulative amounts of material, in units of interaction length λ_{int} and as a function of η, in front of and within the hadronic calorimeter system.

finally a hadronic calorimeter (HCAL). These subsystems are arranged concentrically and are symmetric in azimuthal angle $-\pi < \phi \leq \pi$ as shown in Figure 7.2. No muon spectrometer is considered in this design and muons are reconstructed as tracks with the ITS. The goal of this design is to accurately model the relevant outputs of a multipurpose detector at the LHC while being simplified by the exclusion of detailed components like readout electronics, cabling, and support structures. The detector design is largely configurable, with its default parameter values chosen to achieve response characteristics comparable to that of the current ATLAS detector. Following is a detailed description of each subsystem.

The ITS consists of hollow cylinders in the central detector part and disks at both of its ends, each of which are centered around the beamline. Each of these components consists of a silicon layer of 150 μm thickness in case of the disks and the five innermost cylinders and 320 μm in case of the 4 outermost cylinders. Each silicon layer is accompanied by an iron layer of 350 μm thickness in order to provide a simulation for support material. The ITS only serves the purpose of simulating the interaction of particles with matter upstream of the calorimeter. The resulting detector hits are not used for tracking purposes. The default value of the magnetic flux density present in the ITS amounts to 3.8 T. Finally, two layers of iron totalling 4.4 cm in depth are added to represent support or cryostat material in front of the calorimeter.

The inner surface of the calorimeter system is a cylinder with a radius of 150 cm and a length of 6387.8 mm immediately enclosing the iron layers and the ITS. The calorimeters are separated into a central barrel region covering the pseudorapidity range $|\eta| < 1.5$ and two end-cap regions extending the coverage up to $\eta = 3$ by default. Both the ECAL and the HCAL are divided into 3 concentric layers, with each layer being further segmented into cells with edges of constant η and ϕ.

Layer	Depth	Segmentation ($\eta \times \phi$)	Std. dev. noise per cell per event [MeV]
ECAL 1	$4\,X_0$	256×256	13
ECAL 2	$16\,X_0$	256×256	34
ECAL 3	$2\,X_0$	128×128	41
HCAL 1	$1.5\,\lambda_{\mathrm{int}}$	64×64	75
HCAL 2	$4.1\,\lambda_{\mathrm{int}}$	64×64	50
HCAL 3	$1.8\,\lambda_{\mathrm{int}}$	32×32	25

Table 7.1. Calorimeter default design values regarding layer depths in terms of radiation lengths X_0 (ECAL) and hadronic interaction lengths λ_{int} (HCAL), granularity and energy noise levels.

The cell granularity for each layer is configurable by setting the number of equal divisions in η and (separately) ϕ. The depth of the cells in every layer is designed to be nearly constant in η to ensure that the fraction of a particle's energy deposited in each layer does not depend on the incident angle. This design, leading to layer shapes of the form $1 /\cosh\eta$, provides a uniform calorimeter thickness as a function of pseudo-rapidity. CoCoA will thus have a more uniform response than a pure circular cylindrical shape.

The CoCoA calorimeter material is a compound using an equivalent molecule approximation, mixing an absorber and an active material with a constant proportion. Both the materials and their proportion can be configured for the ECAL and the HCAL individually. By default, the ECAL is made of a mixture of lead and liquid argon, corresponding to the ATLAS ECAL materials. The volume proportion amounts to $1 : 3.83$, resulting in a radiation length of $X_0 = 2.5$ cm. The ECAL and HCAL are separated by an iron layer with a default thickness of 80 mm. The HCAL is made of a mixture of iron and polyvinyl toluene plastic material with a volume proportion of

1.1 : 1.0, resulting in a nuclear interaction length of $\lambda_{int} = 26.6$ cm. The integrated radiation and interaction length measured from the interaction point (IP) at the center to the end of the HCAL is shown as a function of the pseudorapidity in Fig 7.3.

While this calorimeter design represents a homogeneous detector, a spread in the resolution of reconstructed energies in accordance with a sampling calorimeter design is emulated by means of configurable sampling fraction parameters for the ECAL and the HCAL individually. In lieu of a complete simulation of active and passive material, the sampling is emulated by accounting only for a fraction of the GEANT4 energy deposition steps for all particles in the calorimeter showers. The steps to be removed are chosen randomly. The sum of the total deposited energy by those steps is computed and the total energy released is estimated by inverse scaling of the total deposited energy by the corresponding fraction.

Noise, as for example from electronics, is simulated by the addition of random amounts of energy following a Gaussian distribution centered around zero. The noise is independently added to each cell. Negative energies are allowed as is typically the case as a result from the subtraction of pedestals. If such downward fluctuations are significant in size, those negative energy cells can be clustered into topoclusters. The implementation of pile-up collision events from additional proton interactions is left for future development. The default choices of materials and smearing parameters provided in Tab. 7.1 are chosen in order to approximate single-particle responses of the ATLAS calorimeter system [174, 175].

7.1.2 Data Processing

Every event is processed according to the workflow presented in Figure 7.4. First, primary particles are generated at the IP by means of the PYTHIA8 Monte Carlo event generator [176]. A broad range of primary physics processes is available to the user, ranging from the generation of single particles as well as single jets up to more complicated final states with large multiplicities of jets and leptons in the final state.

The set of final state, stable particles is stored in the output file and passed on to the detector simulation described in the previous Section, where the propagation of these particles and their interactions with the detector material is simulated in GEANT4 [168]. The model of hadronic interactions is chosen in accordance with the ATLAS and

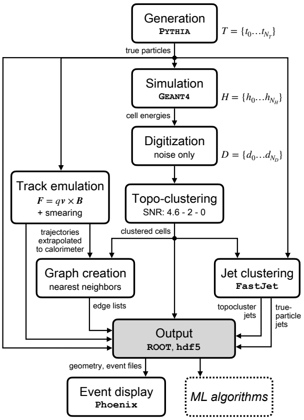
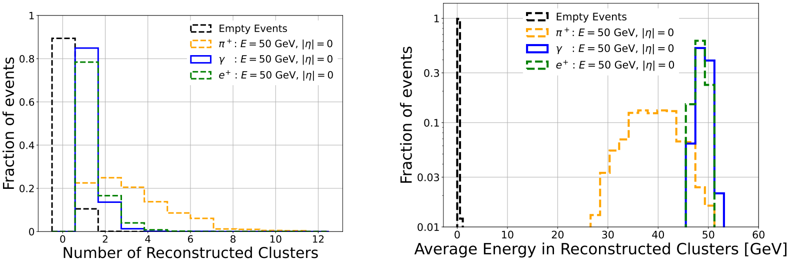


Figure 7.4. CoCoA workflow. Primary particles generated with the PYTHIA library are introduced to CoCoA. Their interactions with the detector material is simulated by means of the GEANT4 toolkit. Calorimeter cells identified by a topological clustering algorithm are stored in the output `ROOT` file together with true particles, emulated tracks, and particle trajectories extrapolated from the IP through the calorimeter according to the equations of motion. A nearest-neighbors-based graph is constructed and stored via edge lists connecting source and destination nodes amongst the output cells and tracks. Jets made of true particles as well as topoclusters are stored in the output file as well. Events in the output file can be parsed for visualization in PHOENIX.

CMS detector simulations. The sum of the energies deposited in each calorimeter cell is stored. Electronic noise is simulated by the addition of random energy offsets to each cell for which Tab. 7.1 provides the default values of standard deviation for each layer.

For the purpose of particle reconstruction, the origin of energy deposits in each cell is stored via a list of parent particle indices which contributed energy into the cell and a list of weights recording their relative contribution to the total cell energy. Cells which received their dominant contribution from electronic noise are assigned an index of -1.

In order to limit the number of calorimeter cells stored in the output file to a reasonable level, low-energy cells dominated by noise contributions are suppressed using a topological clustering algorithm [60]. Only cells contained in the resulting “topoclusters” are stored



(a) (b)

Figure 7.5. Number (a) and average energy (b) of reconstructed clusters in CoCoA for events with a single charged pion, electron, or photon shot at $\eta = 0$. Results are also shown for clusters reconstructed in empty events due to electronic noise. The default topological calorimeter cell clustering settings are used.

in the output file. Topoclusters are seeded by single cells which are required to contain a deposited energy well above the noise level, where the threshold of this signal-to-noise ratio (SNR) is 4.6 for CoCoA by default, while a value of 4.0 is used for the ATLAS experiment. This difference is chosen in order to achieve a better agreement between ATLAS and CoCoA in terms of the topocluster multiplicity distribution for single charged and neutral pions as well as pure noise events (Figure 7.5). Starting with the seeding cells, all neighbouring cells are added to the cluster if their SNR is above another threshold, where the default value is set to 2. Finally, all further neighbouring cells above a third threshold are added, which by default is set to 0. Cells with negative energy can be included, based on their absolute value, or excluded entirely (default configuration). Topocluster candidates containing multiple local maxima in ECAL cell energy each surpassing 400 MeV are split into separate topoclusters.

In order to support particle reconstruction studies which include high energy primary photons, electron-positron pairs from photon conversions taking place in the ITS upstream its two outermost iron layers are stored in the CoCoA output file as well. Tracks emanating from photon conversions and also primary electron tracks are used to construct groups of topoclusters denoted as *superclusters* associated with electron and photon showers. The superclustering procedure in CoCoA follows the criteria described in [177], designed to improve electron energy reconstruction by incorporating nearby energy deposits from bremsstrahlung. It also includes criteria for grouping multiple clusters that are related by a pair of nearby tracks to a photon conversion vertex, thus improving reconstructed photon energy. In the photon conversion shown in Figure 7.1, for example, the CoCoA output contains a supercluster which combines the three topoclusters shown. Due to the simplified tracking, the criteria on number of track hits are not applied. The CoCoA implementation does not focus on electron and photon identification; rather, superclusters are only formed using tracks linked to primary or conversion electrons.

	ECAL layer			HCAL layer		
	1	2	3	1	2	3
k (c-c) inter-layer	1	2	2	2	2	1
k (c-c) same layer	8	8	8	6	6	6
k (t-c)	4	4	4	3	3	3
$\Delta R^{\max}$ (c-c)	0.05	0.07	0.14	0.30	0.30	0.60
$\Delta R^{\max}$ (t-c)	0.15	0.15	0.40	1.10	1.10	2.00

Table 7.2. Default k (number of nearest-neighbors) and maximum ΔR separation used to define edges in the fixed graph creation. Edges between cells are denoted "c-c" while edges between tracks and cells are denoted "t-c".

While the event simulation based on the GEANT4 [168] toolkit determines particle trajectories according to their equations of motion and their interactions with detector material, CoCoA implements a particle tracking based only on the equations of motion for the benefit of downstream tasks. These tracks are extrapolated to the entry surface as well as each layer of the calorimeter and the resulting η and ϕ coordinates are stored in the output file.

For user convenience, an interface to the FastJet [178] library is provided that clusters primary particles as well as topological calorimeter cell clusters into jets. The user can choose the specific jet clustering algorithm accordingly, with the anti-k_{T} algorithm set as default.

For each event in the output data a fixed heterogeneous graph containing cells and tracks is provided by means of two lists storing the indices of source and destination nodes for each edge. The edges are created based on k nearest neighbors in angular distance with k being user-configurable per calorimeter layer and edge type. Three edge types are defined: track-to-cell, cell-to-cell inter-layer, and cell-to-cell across neighboring calorimeter layers (tracks are not directly connected). The user can configure for each of these types both how many edges to construct in a ΔR-ordered neighborhood and also with a maximum ΔR (where $\Delta R^2 = \Delta\eta^2 + \Delta\phi^2$). The default values are given in Tab. 7.2.

The final output file produced by CoCoA stores an array of features for each event which are associated with the following sets: cells that participated in topoclusters, tracks, topoclusters, truth particles and decay record, graph edges, and jets. The output file format is `ROOT` but can be converted to `hdf5` format using a script provided in the repository.

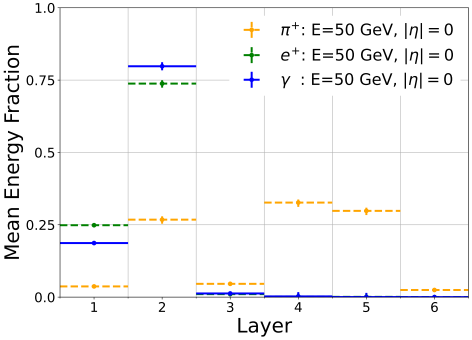


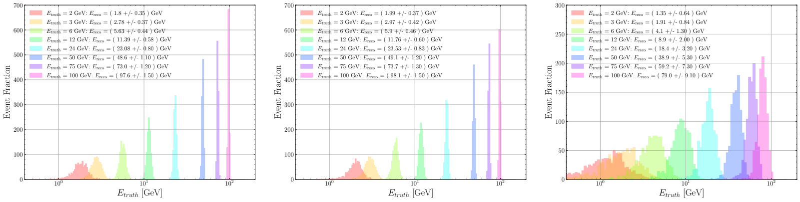
Figure 7.6. Energy deposited by electrons, photons and charged pions for each calorimeter layer. The electron and photon showers are limited to the electromagnetic calorimeter (layers 1 to 3) while the charged pion showers reach deep into the hadronic calorimeter (layers 4 to 6)

.

7.1.3 Detector Performance

In the following, the performance of CoCoA is investigated by means of single particles which are generated at the IP. For each particle type and momentum under investigation, the event generation is repeated in order to gather a statistically significant amount of events.

The correct reconstruction of particle energies is demonstrated in Figure 7.5, which compares the distributions of multiplicities (Figure 7.5(a)) and energy sums (Figure 7.5(b)) of topoclusters for charged pions, photons, electrons and events containing only noise contributions, denoted as empty events. In most of the empty events, the cell energies do not pass the noise threshold of the clustering algorithm. For those events in which this threshold is passed, the average cluster energy sum amounts to 36 MeV in line with the low noise levels provided in Tab. 7.1. The photons and electrons mostly result in one cluster, while their energy is reconstructed with only a small variation. In comparison, the charged pion events result in larger variations of the cluster multiplicity and energy sum distributions due to the higher degree of variations in deposited energies for the hadronic showers. The average cluster energy sum is below the initial charged pion energy due to the involved nuclear interactions of the shower particles with the detector material, which are not counted as detectable energy. A hadronic calibration procedure is not performed within CoCoA but left for downstream tasks.



(a) (b) (c)

Figure 7.7. Reconstruction energies for central electrons (a), photons (b) and charged pions (c). The average reconstructed energies follow the initial particle energies while the energy resolution improves as the initial particle energy increases.

Patterns of energy depositions across the calorimeter are demonstrated in Figure 7.6 in terms of fractions of deposited energy per calorimeter layer for electrons, photons and charged pions. As a consequence of the material budget presented above in Figure 7.3, the electrons and photons deposit most of their energy in the electromagnetic calorimeter, in particular in the second calorimeter layer, while the charged pions reach the hadronic calorimeter layers where they deposit most of their energy, in line with energy deposition patterns at collider-detector experiments.

Figure 7.7 shows distributions of the reconstructed energies for central electrons, photons and charged pions with different initial energies. The energy resolution provided by the calorimeter response improves as the initial particle energy increases. It is larger for charged pions compared to electrons and photons, as expected because of the existence of large sampling fluctuations for hadronic showers compared to electromagnetic showers.

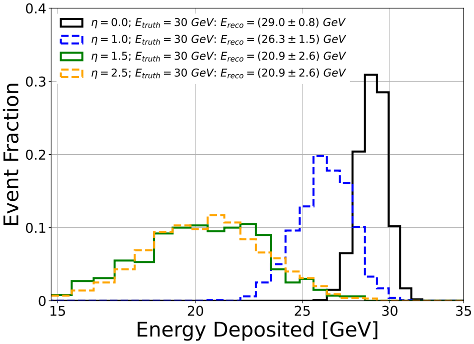


Figure 7.8. Reconstructed energies of electrons with an energy of 30 GeV for different directions η. As the initial electron momentum is directed closer to the beamline, the difference between the reconstructed energy and the intial particle energy increases due to the iron traversed by the electron upstream the calorimeter (dead material).

Figure 7.8 shows the reconstructed energies of a single electron emitted at different initial η. The average reconstructed energy is always lower than the initial particle energies, with the difference growing with the particle η. This is due to the energy depositions in the iron contained in the ITS upstream the calorimeter, in accordance with the material map presented in Figure 7.3.

Figure 7.9 quantifies the energy resolution as a function of true particle energy, comparing electrons with charged pions. For each particle type, the relative energy resolution depending on the particle energy is fitted using least-squares to the following common form of the resolution function:

$$\frac{\sigma(E_{\text{reco}})}{E_{\text{truth}}} = \frac{a}{\sqrt{E_{\text{truth}}}} \oplus \frac{b}{E_{\text{truth}}} \oplus c \tag{7.1}$$

where the best-fit parameters are provided within the figure. Here a, b and c refers to the stochastic, electronic noise and constant terms, respectively. The larger fitted coefficient of the sampling term for hadronic shower compared to electromagnetic is related to the larger

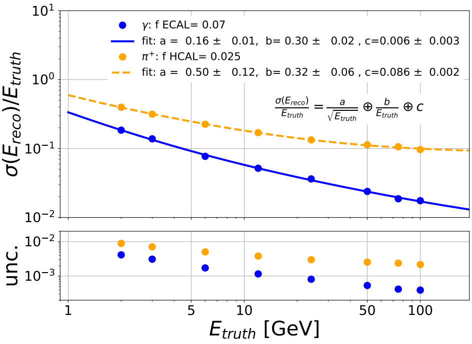


Figure 7.9. The relative energy resolution $\sigma(E_{\mathrm{reco}})/E_{\mathrm{truth}}$ is plotted as a function of E_{truth} for eight different truth energy values and fitted with the relative-resolution function, for photon and pion, respectively. The average sampling fraction f for ECAL and HCAL are shown in the legend.

value of sampling fraction f configured for the ECAL and HCAL separately (0.07 and 0.025, respectively).

The values of the parameters, appearing in Equation 7.1, are individually evaluated for photon as $a = (\,0.16{\pm}0.01\,)\,\mathrm{GeV}$, $b = (\,0.30\pm 0.02\,)\sqrt{\mathrm{GeV}}$ and $c = 0.006 \pm 0.003$. The same numbers for charged pions are found to be $a = (\,0.50 \pm 0.12\,)\,\mathrm{GeV}$, $b = (\,0.32 \pm 0.06\,)\sqrt{\mathrm{GeV}}$ and $c = 0.086 \pm 0.002$. The noise term is compatible with the input noise values, the sampling term is as expected from the sampling emulation.

The performance of the simulated detector has been so far probed using single particles. To illustrate the detector performance in a more realistic event environment, proton proton collision producing an on-shell W boson, decaying to an electron and neutrino, i.e. $pp \to W \to e + \nu$ were simulated. The electron is reconstructed using the superclustering algorithm described in Sec. 7.1.2, and its energy is calibrated in order to compensate for the loss due to scattering in the ITS and iron layers upstream the ECAL. The missing transverse momentum (MET) is calculated from the rescaled clusters, as the opposite the vector sum over visible transverse momenta in the whole event. Finally the transverse W mass, m_T^W, is computed from the reconstructed W four-momentum and compared with the corresponding truth level distribution in Figure 7.10.

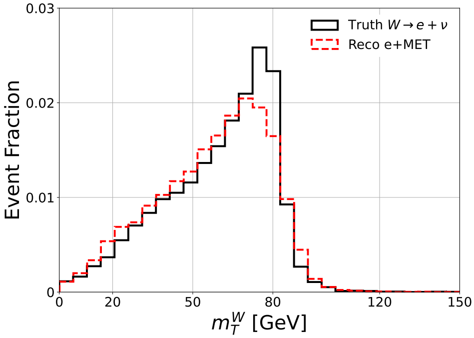
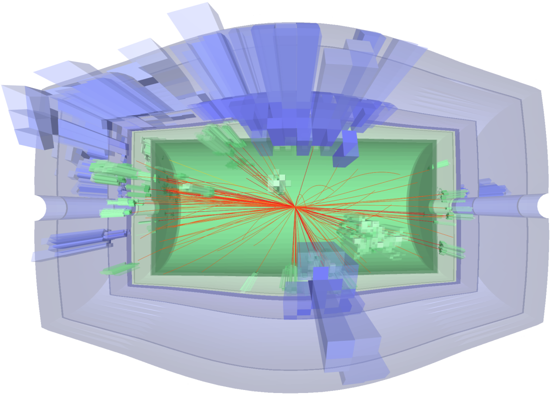


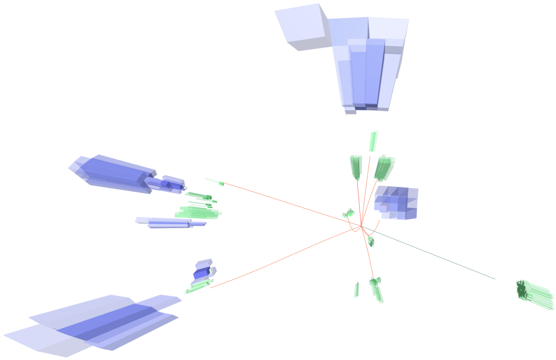
Figure 7.10. The transverse W mass m_T^W distribution is plotted for leptonically decaying W events. The black curve shows the truth distribution whereas the red curve is obtained from the vector sum of reconstructed lepton momentum and the MET in the event. The peak location of the two distributions are well aligned, demonstrating that the event-level reconstructed MET is trustworthy within the CoCoA framework.

7.1.4 Event Display

Visualization of detector geometry and examples of hits for individual events is important for communicating results, and interpreting downstream tasks such as reconstruction and event selection. The default geometry of the CoCoA detector was ported into the open-source framework Phoenix, chosen for its versatility and user support. An example event display is shown in Figure 7.11. The growing interest in ML approaches to low-level analysis tasks such as event or jet reconstruction in a realistic detector underscores the importance of leveraging the rich feature space of calorimeter showers for improving these tasks. Providing an open, configurable, and realistic calorimeter simulation, CoCoA will facilitate the development of such algorithms and ultimately expand the physics reach of current and next-generation collider experiments. The thorough treatment of particle interactions in Geant4 and the full-coverage, highly-granular design of CoCoA calorimeter system enable an accurate representation of the complex data environment present in the ATLAS and CMS experiments at the LHC. To quantify this resemblance, an investigation of the single-particle response characteristics, in terms of topological clustering performance and energy resolution for electromagnetic and hadronic



(a)



(b)

Figure 7.11. Phoenix event displays configured using the CoCoA detector geometry, showing the charged particle tracks and calorimeter hits generated by (a) $pp \to t\bar{t}$ and (b) $pp \to W \to e\nu$ events simulated with Pythia8. In (a), a cutaway of the CoCoA calorimeter volumes is shown along with the clustered cells, while in (b) only the cells are shown. The electron from the W decay in (b) is indicated by a green line. Both displays are shown in perspective view, such that nearer objects appear larger. Different shades of green and blue represent the different layers of ECAL and HCAL, respectively, while cell opacity is determined by cell signal-to-noise ratio.

showers, has been carried out. Finally, additional aides including data post-processing, event visualization, and documentation for CoCoA has been provided to further encourage use.

In future, it is foreseen the inclusion of pileup, direct interface to MadGraph [179] output file and include python binding for CoCoA.

7.2 Global Particle Flow with GNNs

The task of reconstructing particles from low-level detector response data to predict the set of final state particles in collision events represents a set-to-set prediction task requiring the use of multiple features and their correlations in the input data. Three separate set-to-set neural network architectures are presented aiming to reconstruct particles in events containing a single jet in a fully-simulated calorimeter [4]. Performance is evaluated in terms of particle reconstruction quality, properties regression, and jet-level metrics. The results demonstrate that such a high-dimensional end-to-end approach succeeds in surpassing basic parametric approaches in disentangling individual neutral particles inside of jets and optimizing the use of complementary detector information.

Testing theories in high energy physics rely on the ability to reconstruct high energy particle collision events using information recorded by particle detectors. General-purpose detectors enable this primarily through two sources of information: charged particle trajectories (tracks) measured in an inner tracking region and energy deposited by particle showers in a surrounding array of calorimeter cells.

Currently, experiments at the CERN Large Hadron Collider (LHC) employ parameterized particle-flow algorithms, which combine track and calorimeter information in a complementary way while avoiding double counting.

The performance of particle-flow algorithms is limited to an extent by detector design specifications, such as the precision and size of the inner tracking system, the magnetic field strength in the tracking volume, the granularity of the calorimeters, and their energy resolution. However, a number of intrinsic factors complicate the task of particle reconstruction in the LHC environment: the busy and often collimated signatures resulting from proton collisions, the presence of multiple simultaneous scattering events (pileup), and finally, the extensive and irregular array of sensitive elements required for granularity and angular coverage.

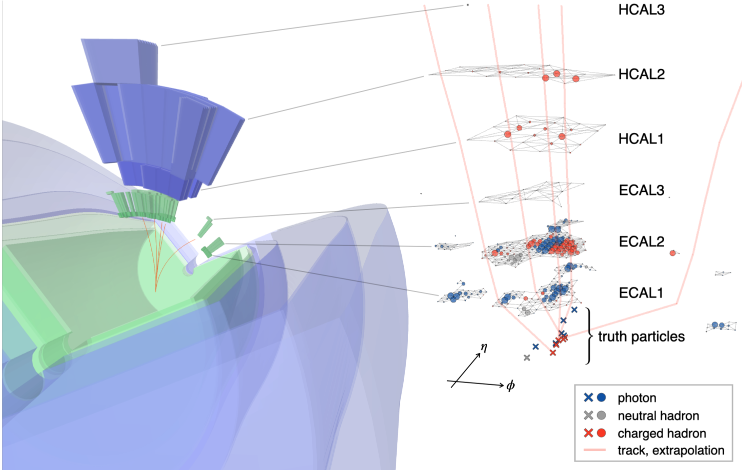


Figure 7.12. A depiction of a single-jet event from the test dataset in both the COCOA calorimeter layers (left) and as an input graph in $\eta - \phi$ space (right). On the left, the actual geometry of the calorimeter cells is shown, while on the right, they are represented by spheres with sizes proportional to their energy divided by noise threshold (up to a maximum value). Lines represent tracks and their projected locations in η and ϕ in each calorimeter layer. Connections between calorimeter cells are the edges formed during graph construction (inter-layer edges and track-cell edges are not shown). The markers at the bottom right indicate the $\eta - \phi$ coordinates of the truth particles.

There are two main approaches to particle-flow algorithms. The approach used by the ATLAS collaboration [180], described in Section 2.3.2, involves subtracting the expected shower profile for each track in an event from the calorimeter deposits to infer the energy contributed by nearby neutral particles. The CMS collaboration, on the other hand, employs a *global* particle-flow algorithm where final state particles of different types are reconstructed simultaneously [181]. Global particle-flow algorithms allow a high physics analysis flexibility and eliminate the need for overlap-removal algorithms while better exploiting the strengths of each sub-detector system.

Here a different approach the global particle-flow paradigm is shown using machine learning (ML) models operating on graph data. As in other applications to particle physics, ML brings the advantage of replacing parameterized cuts (for example, in energy subtraction schemes) with fully differentiable decision boundaries in the full space of relevant features in data. The expressiveness of ML models also opens new possibilities, such as reconstructing individual neutral particles inside of jets. Similarly, the choice to represent input data as graphs is motivated by several advantages: graphs more naturally capture the spatial correlations encoded in irregular detector geometry and also are well-suited for the sparsity and variable cardinality of the input set. Graph neural networks (GNN) have therefore emerged as an architecture of choice in recent particle reconstruction models, as they have in other particle physics tasks [182].

In a collision event, the true set of particles T upstream to the detector sensitive volume gives rise to a set of detector-level hits D. So the input set comprising the detector record is sampled from $p(D|T)$. Then global particle-flow reconstruction is the set-to-set task where the input set of detector-level hits D is transformed into a typically much smaller output set R comprising N_R predicted particles. The predictions of a successful reconstruction algorithm $R(D)$ will correctly model the cardinality N_T of T and the properties (class, momentum, and angular coordinates) of its members. Several ML approaches have been proposed in the literature to predict $R(D)$.

In [183] the object condensation (OC) approach was proposed, which clusters nodes or pixels in latent space to form candidate objects, in this case, particles. Recently, OC has been used to predict clusters in CMS data [184, 185], where the authors focused on reconstruction efficiency and energy regression of showers from single particles embedded in pileup. Here an OC with modifications is presented as explained in Section 7.2.2 for the purpose of establishing a performance baseline for an ML-based particle reconstruction.

The reduction in size from input to output set is handled in the MLPF [186] approach by assigning input nodes to particle classes in the output set or else to a dedicated *neglect class*. This approach was also recently successfully tested using CMS data [187, 188], where the model predictions were trained to match the output candidates from a standard particle-flow algorithm. For predicting true particles, MLPF is limited to cases where one or more clusters can be associated to each particle. It would therefore be required to define a fractional target definition in order to efficiently reconstruct particles that do not contribute a dominant fraction of energy in any single cluster (for example, a significant percentage of low-p_{T} photons).

Two new GNN-based algorithms are here proposed comparing their performance alongside a modified OC implementation as a baseline and a parameterized particle flow algorithm. Rather than full proton-proton collision events, the focus is on events comprising a single jet, which represent the typical domain over which inter-particle correlations are expected to play a significant role in reconstruction. The dataset incorporates full GEANT4 [189] treatment of particle showers in a nearly-hermetic calorimeter simulation [3]. An example of a simulated single jet event is shown in Figure 7.12. In the true particle-flow paradigm, this approach is built around the idea of combining low-level features from calorimeter showers with the complementary information provided by tracks.

It is found that a novel application of recurrent hypergraph learning leads to the most accurate results and preserves a high degree of interpretability. This is achieved thanks to a physics-inspired approach which allows the network to exploit the relationships between properties of the target particles and their energy deposits in the detector.

7.2.1 Dataset

Detector simulation

Unlike the full detector models used to simulate experiments at the LHC, publicly-available codes such as DELPHES [190] do not model particle interactions with sufficient complexity to enable training a network with the full calorimeter signature available at real detectors. This motivated the development of the Configurable Calorimeter simulatiOn for AI (COCOA) package [3], described in Section 7.1, which is used to generate the datasets in this work.

The geometry used is the default described for the COCOA calorimeter detector. The geometric coverage of the COCOA calorimeter is split into a barrel ($0.0 < |\eta| < 1.5$) and two identical endcaps ($1.5 < |\eta| < 3.0$) regions. The endcap region is situated in a hermetic way such that there is no void in the transition region. In depth, the calorimeter has a total of six concentric layers: the first three layers comprising an electromagnetic calorimeter (ECAL) and the next three a hadronic calorimeter (HCAL). The calorimeters have uniform segmentation in η and ϕ enabling high spatial resolution, as listed in Tab. 7.1. The geometric depth of the cells is modulated as $1 / \cosh\eta$ in order to achieve a constant effective interaction depth with increasing η.

The inner region of COCOA is immersed in a uniform axial magnetic field of 3.8T that extends until a radius of 150 cm, where four 1.1 cm layers of iron immediately precede the ECAL. The ECAL is modeled as a homogeneous calorimeter by mixing lead and liquid argon, corresponding to ATLAS calorimeter materials, in volume proportion 1.2 : 4.6 leading to a radiation length of $X_0 = 2.5$ cm. For the HCAL, iron is used as the absorber material, and polyvinyl toluene plastic material as the scintillating active material. These are mixed with a volume proportion 1.1 : 1.0, yielding a nuclear interaction length of $\lambda_{int} = 26.6$ cm. The simulated energy deposits in each layer are smeared to reproduce the expected sampling energy resolution. For the used dataset, the hadronic sampling term is 10%. The effect of pileup and electronic noise is mimicked using normal distributions centered at zero with widths varying according to the layer. The choice of material and smearing parameters is tuned to reproduce the ATLAS calorimeter system's single-particle response.

The effect of tracking is emulated by smearing truth charged particles with a resolution $\frac{\sigma(p)}{p} = a \times p$ with $a = 10^{-5}/\text{GeV}$. The smearing of the track direction is neglected as it is expected to have a subdominant effect in the problem of interest.

Charged particles produced from hadrons decaying-in-flight above a transverse radius $R > 75$ mm (250 mm) in the barrel (endcap) have no tracks associated to them. To focus on the reconstruction of particles as they appear at the calorimeter, the dataset simulates photon conversions only at the stage of the iron layer prior to the calorimeters, while material interactions within the tracker are emulated solely by the track q/p smearing.

Dataset generation

Event generation, followed by parton shower and hadronization is performed with PYTHIA8 [176] with a single initial state quark or gluon particle. The parton initial energy is sampled in the range 10 GeV $-$ 200 GeV, and angular coordinates are distributed uniformly in the range $\eta \in [-2.5, 2.5]$, $\phi \in [-\pi, \pi]$. Final state particles are interfaced with GEANT4 to simulate their interaction with material, both showering in the calorimeter and scattering and e.g. photon conversions in the iron layer preceding it. Additional pileup collisions were not simulated. The targets of the machine learning algorithms are final state stable particles with transverse momentum above 1 GeV, which reach the calorimeter.

In each event, a standard clustering algorithm is used to group calorimeter cells into "topoclusters" based on their proximity and deposited energy, following the algorithm described in Ref. [191] with minor modifications. First, an energy over noise ratio of $(\frac{E}{\sigma}) > 4.6$ is used to identify cluster seeds. For each seed, a two-stage search is performed in its vicinity to group neighboring cells with nonzero energy. The first search collects neighboring cells with an energy-to-noise threshold ratio above 2, and the second search further extends the clusters with cells that have energy above 0. Finally, the algorithm applies a set of rules to merge topoclusters sharing seed cells and split topoclusters formed by particles in close proximity.

A record is kept of contributing particles and their energy contribution to each cell. Electronic noise is simulated in the calorimeter cells at realistic levels and dominates a fraction of the clustered cell. A small fraction of topoclusters, therefore, consist purely of cells where noise was the dominant contributor. One or more such topoclusters are present in 23% of the training events.

In summary, the data used for ML comprise the following object collections: cells which belong to a topocluster, all tracks that reached the calorimeter, and the set of particles which entered the calorimeter. An identical configuration is used to generate the dataset of 50000 events for training and the independent dataset of 30000 events for testing. In addition, a "gluon jet" dataset containing 30000 events is generated by replacing the single incident quark by a gluon with the same initial energy and angular distributions. The quark and gluon jet datasets are provided in the link[2]. The results obtained with this gluon jet sample are discussed in Section 7.2.3. Figure 7.13 summarizes the number of various entities stored in both the single jet and gluon jet test datasets.

Fiducial particle definitions for reconstruction targets

In a collision event, not all particles produced can be reconstructed in the detector. When defining target truth particles it is important to account only for those that can be detected, *i.e.* those that are within the detector acceptance and have sufficiently high transverse momentum to be reconstructed. Beyond these simple criteria, particles produced in the collision can later decay or interact with the detector and convert, radiate or interact and produce other particles.

---

[2] Training and test datasets: `https://doi.org/10.5281/zenodo.7699681`

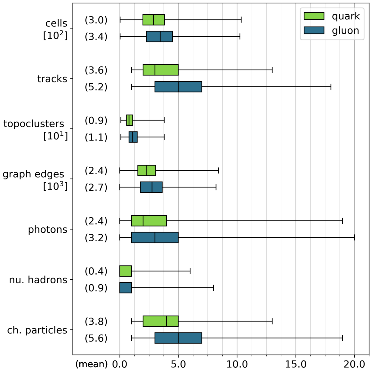


Figure 7.13. Composition of the quark and gluon test samples in terms of the cardinality of different sets of entities contained. The mean value in each case is written in parentheses while the range and quartiles of the distributions over events are shown in the box plot.

The specific definition of the particles that are targets for the reconstruction algorithm, referred to as fiducial particles, is important to remove ambiguities during training and in assessing the performance of reconstruction. To qualify as fiducial particles, truth stable particles must have the following properties:

- $p_{\mathrm{T}} > 1$ GeV
- be produced before the first calorimeter layer
- release a nonzero amount of energy in the calorimeter

Additional consideration would be needed to achieve a more realistic environment where bremsstrahlung, pair production, and the presence of soft particles in general might result in highly collimated topologies, above the spatial reconstruction capabilities of the detector. In this work, the absence of pileup and the absence of material in the inner tracking region justify the use of the three fiducial criteria described above.

Input graph

A fixed heterogeneous graph is built out of each event by connecting calorimeter cells and tracks based on their proximity. Each cell

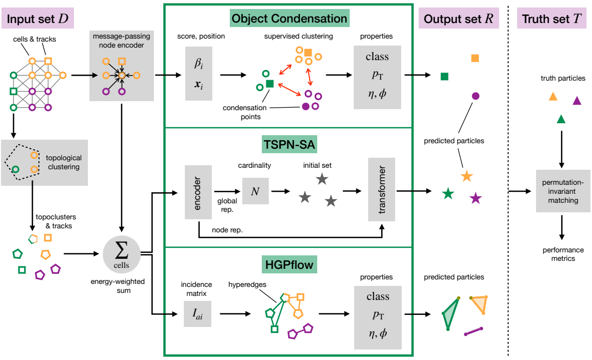


Figure 7.14. Comparison of the different ways in which the three ML reconstruction algorithms map the input set of nodes in the form of a graph to the output set of predicted particles, to be compared with the set of truth particles. Colors indicate distinct particles and the nodes for which they are the dominant contributor.

is connected to the k nearest cells in the same calorimeter layer, where $k = 8$ in the ECAL and $k = 6$ in the HCAL. Additionally, each cell is connected to the single nearest cell in its immediately adjacent layer(s). A cell in layer l can only receive incoming edges from other cells if they are separated in $\varDelta R$ by less than $\{d^{c-c}_{\max}\}_l = \{0.05, 0.07, 0.14, 0.30, 0.30, 0.60\}$ for the six calorimeter layers. A set of indices and weights is assigned per cell listing the true particles which contributed and their relative contribution to the total cell energy. An index of -1 is given to energy contributions from noise.

Tracks are likewise connected to cells based on closest separation in $\varDelta R$ between the cell and the projected $\eta - \phi$ coordinate of the track in the corresponding calorimeter layer. A track is connected to a maximum number of $k = 4$ cells in each ECAL layer and $k = 3$ cells in each HCAL layer. For track-cell edges, a larger maximum $\varDelta R$ is allowed: $\{d^{t-c}_{\max}\}_l = \{0.15, 0.15, 0.40, 1.10, 1.10, 2.00\}$. A depiction of the graph connectivity for tracks and cells is shown in Fig. 7.12.

Topoclusters are represented in the input graph by a separate set of nodes with edges connecting each to the set of cells belonging to the topocluster. The angular coordinates of a topocluster are taken at its energy barycenter.

7.2.2 Particle reconstruction algorithms

Parameterized particle-flow algorithm

To compare the performance of the ML algorithms, a traditional parameterised particle-flow (PPflow) algorithm [180] is implemented, following the prescription in Section 2.3.2. The algorithm aims at subtracting the energy deposited in the calorimeter from charged particles associated to tracks. To this end, shower templates are derived from single π^+ samples and parameterized as a function of the track p_{T} and the layer where the first nuclear interaction takes place. The energy subtraction is performed in concentric rings of radius equal to a single cell pitch built from the extrapolated track position in each calorimeter layer. The ring's energy is progressively subtracted from the topoclusters until the expected total energy determined in the single π^+ template is reached. The remaining energy in the topoclusters after this subtraction is considered as originating from neutral particles. The PPflow algorithm does not aim at reconstructing the single particles composing the jets, but rather it is designed to estimate the overall neutral energy component for each topocluster.

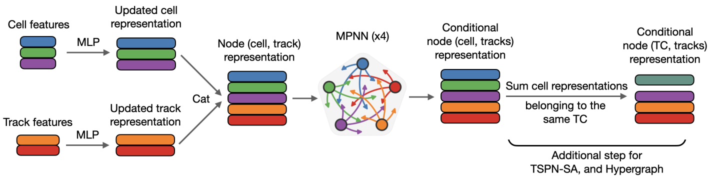


Figure 7.15. The encoding model used to derive a learned node representation.

Common description for the ML algorithms

Three ML-based particle reconstruction models for the set-to-set prediction $R(D)$ are described: object condensation (OC) as an existing ML baseline, transformer set prediction network with slot attention (TSPN-SA), and a hypergraph architecture (HGPflow). Descriptions of each algorithm is given in Section 7.2.2. Section 7.2.3 compares their relative particle-level performance and a comparison to the PPflow baseline for jet reconstruction.

There are commonalities to all three algorithms. Predicted particles in each case are inferred from the node features (skip connections) concatenated with a node representation vector from a common encoder network. Tracks are treated similarly in each case: the charged particle cardinality in an event is set by the number of tracks, and the charged particles' η and ϕ are determined directly from the tracks without regression. While the OC algorithm takes calorimeter cells as input nodes, the other two algorithms use topoclusters instead, to reduce the dimensionality of the input. Fig. 7.14 illustrates the core differences between the ways each algorithm maps the set of detector-level nodes D to the set of predicted particles R.

The choice to use calorimeter cells compared to using the coarser topoclusters can be compared in terms of an *injective condition*: the degree to which the energy deposit in a node can be mapped back to a single parent particle. In the case of cells, although contributions from more than one parent particle are present in general, the injective condition is more valid than in the case of topoclusters. Since the injective condition is an assumption of the OC algorithm (i.e. in the definition of the entries of I_{ki} in Eq. 7.2), this motivates the choice of cells as input nodes.

Having contributions to a node from more than one parent particle can be learned in the TSPN-SA architecture in an unsupervised way via node-particle attention. The HGPflow architecture, on the other hand, is fully equipped to disentangle multiple-particle contributions to a node thanks to supervised learning of the incidence matrix. In both cases, computing gradients for predictions on edges becomes significantly more expensive for cell-level inputs compared to topocluster-level inputs, which was the main motivation for choosing the latter.

The loss associated with predicted particle properties is computed similarly in each algorithm. Particle class is trained using a categorical cross-entropy term between the predicted and the target class. A mean squared error loss term is used to regress continuous properties η_i and $p_{\mathrm{T},i}$. The ϕ prediction is trained using $1 - \cos\left(\phi^{\text{pred}} - \phi^{\text{targ}}\right)$.

The total number of trainable parameters in the neural network blocks of the OC, TSPN-SA, and HGPflow algorithms is compared in Tab. 7.3 including the node encoding network in each. An estimate of their computational performance is also shown. For each of the three algorithms, hyperparameter optimization scans have not been performed, except on the threshold cuts used during inference for OC and HGPflow. The code for the algorithms is provided in the github

Algorithm	# Parameters Total	(Node enc.)	Speed* [ms/event]	Memory [MiB/event]
OC	1.8M	(0.2M)	249	1480
TSPN-SA	1.5M	(0.2M)	465	1448
HGPflow	1.8M	(0.2M)	257	1394

*algorithms not optimized for execution time.

Table 7.3. Comparison of the three particle reconstruction algorithms in terms of model size and computational resources. The number of trainable parameters belonging to the node encoding model is shown alongside the total. The time per event is averaged over 100 single jet events evaluated sequentially, and the memory is estimated as the peak memory over the same. Results are obtained on the same GPU (NVIDIA TITAN RTX).

page[3].

Graph nodes encoding

Each event is represented as a heterogeneous graph comprising track, cell, and topocluster nodes connected by edges as defined in Section 7.2.1. The embedding model described in the following is shared among the different network architectures. Fig. 7.15 illustrates the network components of the encoding model: input feature vectors associated with track and cell nodes are passed through separate networks to embed them in a common representation space of dimension 100. The cell features input to the embedding are (energy, position, ϕ, η, layer). Similarly, the track input features are the track parameters (q/p, θ, ϕ, d_0, z_0)[4] and the extrapolated η-ϕ coordinates of the track at each calorimeter layer. The latter are important features because a charged particle after exiting the magnetic field travels in a straight line which no longer points back to the origin (assuming no material interactions). In addition to their hidden representation, these nodes are also given an additional binary feature which flags whether they originate from cells or tracks.

---

[3] Code for algorithms: `https://github.com/nilotpal09/hg-tspn-pflow`

[4] The track impact parameters d_0 and z_0 measure the distance of closest approach of the track to the beam line in the transverse and longitudinal directions, respectively.

The node encodings are then updated to incorporate the graph relational structure via 4 successive blocks of message passing along edges. In each block, a dedicated network is used with the following three inputs concatenated: current node representation, sum of representations from neighboring nodes, and a graph-level global representation (the mean of all current node representations). Following the message passing blocks, topocluster representations are computed by the energy-weighted mean of the cell representation vectors belonging to the topocluster.

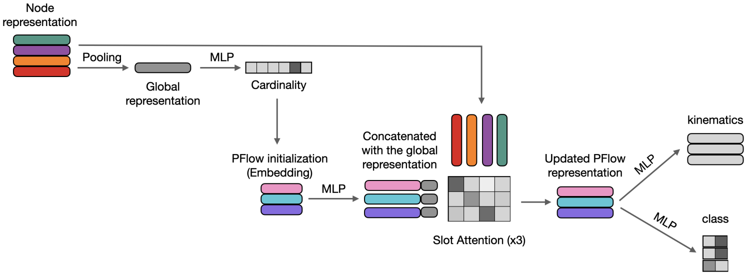


Figure 7.16. The TSPN-SA architecture. The cardinality of the set of output particles is predicted from the global representation, while their properties are predicted from representation vectors resulting from successive slot attention blocks.

Modified object condensation (OC)

The OC algorithm was proposed in [183] for tasks of segmenting a set of input nodes into a set of target objects and prediction of their properties, which it does simultaneously. In the particle reconstruction case, the input set comprises tracks and calorimeter cells and the output set of objects are the progenitor particles (“parents”) with their classes and properties. The set-to-set procedure for OC is shown in the top row of Fig. 7.14. The implementation follows the original OC approach with some modifications which are stated in the following description.

The OC algorithm is based on clustering nodes according to their parents in a learned few-dimensional space x. The clustering is supervised by adding to the loss potentials defined on this space: a repulsive potential $\hat{V}(x) \propto \max(2 - \Delta x, 0)$ between nodes that belong to different parent particles, and an attractive one $\check{V}(x) \propto \Delta x^2$ for nodes having a common parent. The goal is that after training the resulting clusters of nodes will correspond to the set of parent particles.

However, calculating the sum of N^2 pairwise potential terms during training becomes expensive for problems of even moderate N. This is addressed by designating a single representative node for each parent particle, called a *condensation point*, to impose the potentials on all other nodes during training. A separate network is trained to predict a score, $\beta \in [0, 1]$, with a target value of 1 for condensation points and 0 otherwise. An increasing function $q(\beta) = \operatorname{arctanh}^2\beta + q_{\min}$ (with $q_{\min}$ a hyperparameter) is used in analogy to charge in the loss term responsible for the clustering potentials:

$$L_V = \frac{1}{N}\sum_{i=1}^{N} q(\beta_i)\sum_{k=1}^{K}\left[I_{ki}\breve{V}_{c,k}(x_i) + (1 - I_{ki})\hat{V}_{c,k}(x_i)\right] \tag{7.2}$$

where I_{ki} is an $N \times K$ matrix determining whether particle k is the parent of node i.

For each node i, which is a cell, the properties loss is of the same form as discussed earlier, where target class is either photon, neutral hadron or charged particle. Similar to L_V above, the particle property loss is also weighted by $q(\beta)$ such that nodes with the highest β receive the most supervision during training. These nodes are ultimately selected for the output set during inference by requiring their predicted $\beta > t_\beta$ and that they be separated in the clustering space by $\Delta x > t_d$, where t_β and t_d are two threshold hyperparameters.

Compared to the original OC model, this implementation has two modifications connected to the condensation score β. The condensation points defined during training do not have a physical meaning and are learned in an unsupervised way. In this approach, the following physics-oriented definition is used:

$$\mathrm{CP}_k^T = \begin{cases} \text{track} \in k, & \text{if } k \text{ is charged particle} \\ \operatorname{argmax}_z(\text{cells} \in k), & \text{if } k \text{ is neutral particle} \end{cases} \tag{7.3}$$

where z is the energy over noise threshold ratio for each cell. This definition removes the need to identify a representative node for charged particles, assuming a 1-1 mapping to tracks in the event. For neutral particles, on the other hand, the β prediction is fully supervised and can be interpreted as the likelihood that a cell has maximal z in a given shower. Since this cell also serves as an approximate location of the shower center, the η and ϕ for neutral particles are regressed via a learned offset to the cell η-ϕ coordinates. During inference, condensation points passing the t_b and t_d thresholds are further required to be classified as either photon or neutral hadron, whereas cells classified as charged particles are discarded (since this role is fulfilled by tracks).

A second modification compared to the original OC approach is that instead of a $\sim (1-\beta)$ regression-type loss computed on condensation points only, the β prediction is trained using a binary cross-entropy loss evaluated for all nodes. The reasoning behind a classification-type loss is to directly penalize the network for predicting large β for nodes which are not condensation points, i.e. false positives. In an ablation study, each modification was seen to bring substantial improvement at essentially no additional model complexity.

Besides the two modifications above, the OC implementation differs from that of [185] in a few regards. Firstly, they propose an upgrade to the original OC algorithm where particles are represented in the clustering space not only by a singular condensation point but by the β-weighted average over a learned distance scale. Finally, the proposed model has significantly more parameters (Tab. 7.3), with GravNet blocks [192] replaced by the node encoding model. The choice of network block sizes has not been optimized for computational efficiency and allocates a large proportion of its parameters to the node prediction networks compared to the message-passing networks.

Transformer Set Prediction Network with Slot Attention

The Transformer Set Prediction Network (TSPN) was initially developed for the permutation-invariant encoding and decoding of variable-size sets of feature vectors [193]. The utility of this for set-to-set problems in particle physics is clear: a model is needed that predicts an output set of entities (i.e. particles) based on an input set of different entities (i.e. calorimeter clusters, tracks), where both sets typically have different cardinality. The model is divided into two networks: the first for predicting neutral particles and the second for predicting charged particles (discussed later).

As shown in the top part of Fig. 7.16, the first architecture starts with a set encoder network whose output is used to predict the number of neutral particles, N^{pred} This prediction is trained using a categorical cross-entropy loss over 25 classes, which is an upper bound on the number of neutral particles per event. The cardinality prediction is used during test time. During training, the truth cardinality is enumerated to form a set of numbers that are passed through embedding layers to instantiate the initial set of random vectors. These vectors combined with the global representation vector function as queries for a series of 3 slot-attention (SA) blocks [194]. The SA blocks are not part of the original TSPN proposal but were found to be fundamental for performance in this application (abbreviated TSPN-SA). Each block contains 3 iterations where the particle hidden representation is updated using the attention-weighted representations of the topoclusters and tracks in the event. Finally, the updated hidden representations are inputs to two dedicated neural network blocks aimed at predicting the kinematics (p_{T},η,ϕ) and the class of the particle candidates.

The neutral particle properties loss function for the TSPN-SA algorithm is defined as the sum of a categorical cross-entropy term for class t_i and a mean-squared error (MSE) term for the continuous properties $\boldsymbol{p}_i = \{\eta_i, \phi_i, p_{\mathrm{T},i}\}$ computed for each particle candidate. The target particles are defined by matching to the set of predicted particles using the Hungarian assignment algorithm [195], with the loss itself being the distance metric.

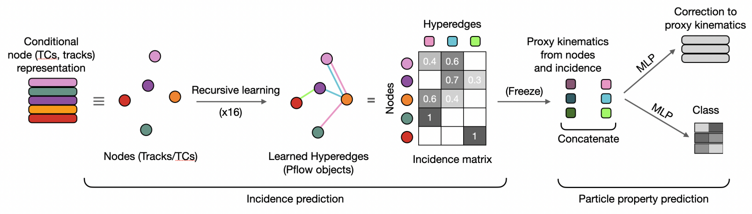


Figure 7.17. The two stages of learning in the HGPflow algorithm. The objective of the first stage is to predict the fractional entries in the energy-weighted incidence matrix, where columns correspond to hyperedges (i.e. particles). This is done by accumulating the loss over a sequence of recursive updates. In the second stage, the incidence matrix is frozen and the network minimizes losses for particle property predictions, defined relative to proxy quantities.

The second part of the TSPN-SA algorithm, for predicting charged particles, makes explicit use of the prior knowledge originating from track-particle objects. Each track is promoted to a particle, such that the cardinality of the output set is fixed by the number of tracks. Similar to the neutral architecture, SA blocks are used to update the hidden representation of the charged particle candidates. Unlike the model for neutral particles, the only predicted quantity for charged particles is their transverse momentum in order to improve over the track resolution. For charged particles, the track index is used to match with the corresponding target particle. As in the OC algorithm, the η and ϕ of the charged particles are taken directly from their representative tracks. Finally, the total loss for the TSPN-SA algorithm is the sum of neutral particle cardinality loss, the neutral particle properties loss, and the charged particle p_{T} MSE loss. These loss terms are minimized simultaneously during training.

HGPflow: particles as hyperedges

A hypergraph is a generalization of a graph where hyperedges can each connect one, two, or multiple nodes (Fig. 7.17). While connectivity in a graph of N nodes is described by an $N \times N$ adjacency matrix, a hypergraph containing K hyperedges is described by an incidence matrix $I^{(N\times K)}$. In the context of particle reconstruction, calorimeter deposits and tracks can be represented as nodes in a hypergraph, while each particle is represented by a hyperedge connecting the set of nodes to which it contributed. Here the *HGPflow* is described: an algorithm that treats particle reconstruction as a task of learning hyperedges and their properties. There are two objectives in the training of HGPflow:

1. predict the incidence matrix defining the hyperedges
2. predict the hyperedge (i.e. particle) properties

The first objective is similar to the task of separating overlapping charged and neutral showers which was the focus of [196]. In this first stage, the HGPflow network predicts $(N+1)\times K$ entries comprising a zero-padded incidence matrix and an additional row of binary values that indicates whether the particle corresponding to a given column exists or not. Since the number of particles per event varies, the number of columns K is set to an upper bound on the number of particles estimated from the training set (in this case $K = 30$). To express a non-injective map from particles to nodes, a target incidence matrix is defined having fractional rather than binary-valued entries. The entry relating node i to particle a is the following:

$$[I]_{ia} = \frac{E_{ia}}{\sum\limits_{\text{particles } b} E_{ib}} = \frac{E_{ia}}{E_i} \tag{7.4}$$

where E_{ia} is the amount of energy that particle a contributes to the total energy E_i of node i. For nodes which are tracks, incidence entries are simply 1, whereas for topoclusters they compute the fraction of the topocluster's energy that came from a given particle. An example of target and predicted incidence matrix entries are shown in Fig. 7.18 for one event.

Predicted rows in the incidence matrix are normalized using Softmax (i.e. sum over all hyperedges for a given node is 1) before being compared to the target via Kullback–Leibler divergence loss.

$$L_{\text{inc}} = \sum_{a} \text{KL}_i \left(I_{ia}^{\text{targ}}, \text{Softmax}_i(I_{ia}^{\text{pred}}) \right) \tag{7.5}$$

The predicted entries of the indicator row are passed through a sigmoid function and compared to the (binary) target entries using a binary cross entropy loss function. Predicted columns are rearranged using the Hungarian algorithm to minimize the loss.

The incidence matrix prediction network is trained using the recurrent strategy proposed by [197], described briefly hereafter. The loss in Eq. 7.5 is calculated for a sequence of 16 refinement blocks each comprising an updated prediction of the incidence matrix followed by an update of node representation vectors V, and hyperedge representation vectors E. The iteration $t \to t+1$ is performed with the following three successive steps:

$$I_{ia}^{t+1} = \phi_I \left(v_i^t, e_a^t, I_{ia}^t \right) \tag{7.6}$$

$$V^{t+1} = \phi_V \left(\left\{ v_i^t, \rho_{E\to V}(i,t), v^0 \middle| i = 1 \dots n \right\} \right) \tag{7.7}$$

$$E^{t+1} = \phi_E \left(\left\{ e_a^t, \rho_{V\to E}(a,t) \quad \middle| a = 1 \dots k \right\} \right) \tag{7.8}$$

where $\rho_{E\to V}(i,t) = \sum_a I_{ia}^{t+1} e_a^t$ and $\rho_{V\to E}(a,t) = \sum_i I_{ia}^{t+1} v_i^t$ are aggregations of node (v) and hyperedge (e) representation vectors weighted by the updated incidence matrix. The updates are performed at each step using the same networks ϕ_I, ϕ_V, and ϕ_E, where the latter two networks are DeepSets models [198].

To reduce computational cost, not every iteration of the backward pass is included in the gradient step. Two sequences of 4 adjacent iterations are randomly selected out of the 16 for which the incidence loss is computed and added to the loss from the prediction at the end of the sequence.

The second training objective of the HGPflow network (Fig. 7.17c) is to predict particle properties for each hyperedge. The corresponding loss function contains classification and regression terms evaluated by matching predicted and target particles using the Hungarian algorithm.

Particles corresponding to hyperedges where the predicted indicator was below the threshold are matched to dummy targets and weighted by zero in the loss. Classification between photons and neutral hadrons is performed for hyperedges which do not contain a track and are thus identified as neutral particles. The regression task benefits from a unique advantage enabled by learning the incidence matrix (Eq. 7.4): particle kinematics can be approximated as weighted sums and averages over the input features of the topoclusters contained in the hyperedge. Proxy quantities (denoted ˆ) for energy and angular coordinates can be computed as:

$$\hat{E}_a = \sum_{\text{nodes } i} E_i I_{ia} \ , \qquad \{\hat{\eta_a}, \hat{\phi_a}\} = \sum_{\text{nodes } i} \{\eta_i, \phi_i\} \tilde{I}_{ia} \tag{7.9}$$

where a dual incidence matrix $\tilde{I}$, normalized over node instead of particle indices, can be defined:

$$\tilde{I}_{ia} = \frac{E_{ia}}{\sum\limits_{\text{nodes } j} E_{ja}} = \frac{E_{ia}}{E_a} = \frac{E_i \cdot I_{ia}}{\sum\limits_{\text{nodes } j} (E_j \cdot I_{ja})} \tag{7.10}$$

The property prediction networks in HGPflow are therefore given the simpler objective of learning corrections to the approximate values from Eq. 7.9. The loss terms used for the property predictions follow the earlier description.

Therefore, neutral particle kinematics $(p_\mathrm{T}, \eta, \phi)$ are regressed by predicting an offset to the proxy values in Eq. 7.9. For charged particles, an offset is likewise predicted for the p_T measured from the associated track. The properties loss is computed by matching predicted and target particles using the Hungarian algorithm [195].

7.2.3 Performance of particle reconstruction in jets

One of the most challenging tasks of global particle flow algorithms is the reconstruction of particles in dense environments, in particular jets. In this Section, the performance of the ML reconstruction algorithms will be assessed by quantifying the similarity between the set of predicted and set of target particles. The following four types of metrics are meant to evaluate the cardinality, class, and properties predictions:

- Efficiency and fake rate

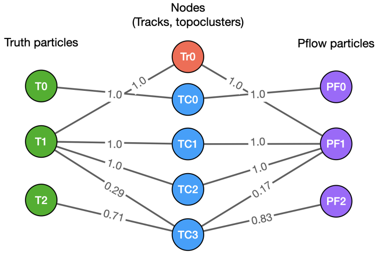


Figure 7.18. Schematic representation of the truth and predicted incidence matrix in HGPflow for one event. The left part of the diagram shows the three truth particles in the event. One of them has a track (Tr) associated to it. The three particles deposit their energy into four topoclusters (TC). The links represent the fractional energy originated by a given particle in a given topocluster or track. The right part of the diagram shows the predicted values of the incidence matrix for each reconstructed particle.

- Classification purity
- Particle angular and momentum resolution
- Jet-level quantities

The efficiency and fake rate are defined as follows:

$$\epsilon \equiv \frac{N(\text{matched pred})}{N(\text{targ})} \quad , \quad f \equiv \frac{N(\text{unmatched pred})}{N(\text{pred})} \tag{7.11}$$

The quality of the regression tasks is evaluated from distributions of their residuals, defined as $(y_{\text{targ}} - y_{\text{pred}})/y_{\text{targ}}$ for particle property $y \in \{p_{\text{T}}, \eta, \phi\}$.

Particle matching

Predicted and target particles are matched using the Hungarian algorithm to find the pairings which minimize the distance between their properties, defined by the following metric:

$$d_{\text{match}} = \sqrt{c_{p_T} (\Delta p_{\text{T}} / p_{\text{T}}^{\text{truth}})^2 + c_{\Delta R} \Delta R^2} \tag{7.12}$$

where Δ denotes the difference between a predicted and target property, and $\Delta R^2 = \Delta\eta^2 + \Delta\phi^2$. Matching is performed separately for neutral particles and charged particles since the latter are distinguished by the presence of a track. The coefficients c_{p_T} and $c_{\Delta R}$ are set to 1 and 5 for neutral particles while for charged, matching is based only on ΔR (i.e. $c_{p_\mathrm{T}} = 0$). Prioritizing spatial matching helps decouple reconstruction efficiency from classification accuracy, which in particular will dominate at low-p_T because of the similarity of photon and neutral hadron signatures.

In each event, when the cardinality of the predicted set of particles is larger than that of the target, the non-matched predictions are labeled as “fake” particles. Conversely, inefficiency arises when not enough neutral or charged particles were predicted in order to match every target.

Charged particle performance

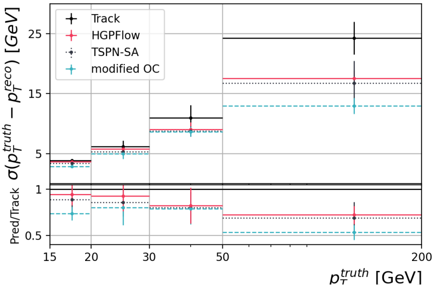


Figure 7.19. Resolution of charged particle-flow candidates and tracks as a function of the associated particle transverse momentum. At high p_T the particle-flow candidates show improved resolutions over the tracks.

Charged particles include electrons, muons, and charged hadrons. In jets charged pions produced during hadronization account for around 90% of all charged particles. Leptons such as electrons and muons are present in less than 3% of the jets. Electrons are produced from photon conversions and hadrons decaying in flight while muons are mostly produced by the latter mechanism. Given the large class imbalance and the fact that no dedicated studies have been performed to improve the classification of electrons inside jets, the three classes are grouped together and characterized as a single class of charged particles.

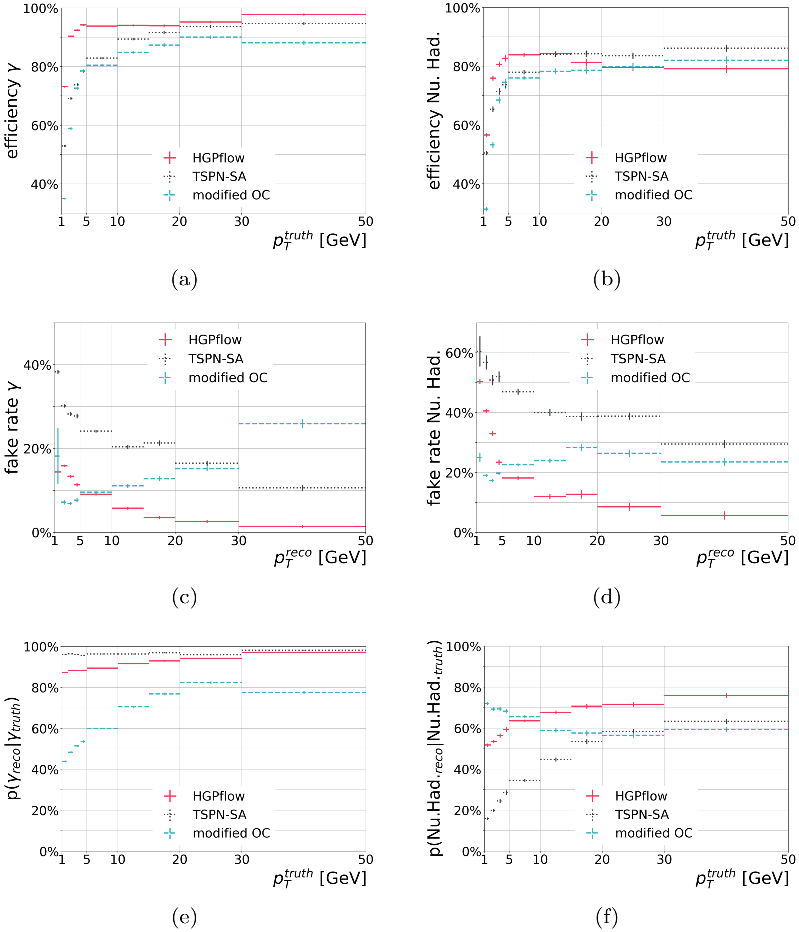
Tracking efficiency presents an upper bound on the efficiency of charged particle reconstruction (see Section 7.2.1). Since fake tracks are not emulated in the data, charged particle fake rates are neglected in this study. In any case, the rate of fake tracks at 1 GeV is typically at the percent level for the ATLAS and CMS experiments, which is expected to have a small impact.

In cases where the track belonging to a charged particle is not reconstructed, the target particle is relabelled to avoid confusing the network during training. Charged hadrons without a track are relabelled as neutral hadrons, and electrons without a track as photons. Photon pairs from neutral pion decays prior to the calorimeter are treated as two distinct target particles.

A key characteristic of charged particle reconstruction is the resolution of p_{T} with respect to the true value. It is well known that at low transverse momentum tracks provide the best momentum estimate over the calorimeter resolutions. An opposite trend appears at high energies where the calorimeter systems provide the most accurate energy measurement. Fig. 7.19 shows the resolutions of charged particles reconstructed with the three ML approaches and compared to the track resolutions for charged particles with $p_{\mathrm{T}} > 15$ GeV. Below this value, the particle p_{T} regression is replaced with simply the track p_{T} since an improvement is not expected. An increasing improvement at high p_{T} is observed for all the ML algorithms demonstrating that indeed the complementarity between the calorimeter and tracking measurements has been learned during training.

Photon and neutral hadron performance

The presence of photons inside jets is mainly due to decays of neutral pions and to a lesser extent from bremsstrahlung processes. Long-lived neutral hadrons on the other hand trace back to the shower of the initial partons. Disentangling these two components is not a trivial task and is detector-dependent – for COCOA it is observed that 70% of neutral hadrons below 5 GeV release all their energy in the ECAL layers, making it difficult to distinguish them from photons. This fraction steeply decreases with the energy of the initial hadrons to approach percent levels at around 20 GeV. Neutral particle reconstruction is further complicated inside the collimated environment because of the frequent overlaps between neutral showers. For this reason, efficiency and fake rate plots are computed by considering photon and neutral hadron predictions inclusively, without requiring a match between



(a) (b)

(c) (d)

(e) (f)

Figure 7.20. Top: efficiency of matching predicted neutral particles to truth photons (a) and neutral hadrons (b) in a jet as a function of the associated truth particle p_{T}. Middle: fake rate, i.e. probability that predicted photons (c) and neutral hadrons (d) are not matched to any truth neutral particle, as a function of the predicted particle p_{T}. Bottom: the probability that the predicted neutral particles which are matched to truth photons (e) and neutral hadrons (f) are assigned the correct class. For each curve, the misclassification probability is simply the difference of the curve from 1.

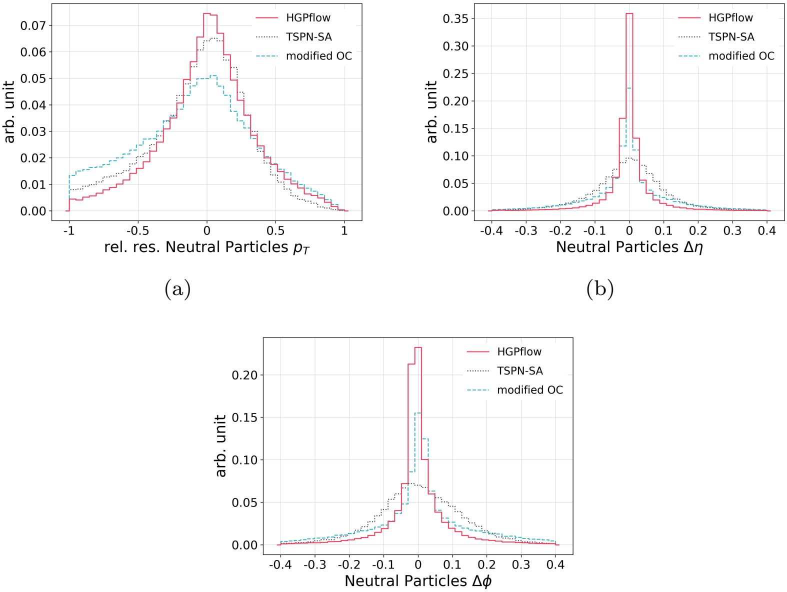
predicted and target class.

Efficiencies, fake rates, and class prediction accuracy for photons and neutral hadrons are shown in Fig. 7.20. The efficiency of reconstructing photons with $p_{\mathrm{T}} > 2$ GeV is above 90% for HGPflow, rising to 98% for photons above $p_{\mathrm{T}} > 30$ GeV. For TSPN-SA and OC, the efficiency for photons reaches 95% and 90%, respectively. Neutral hadrons above 5 GeV are reconstructed with efficiencies ranging from 76% to 86% for the three algorithms. Fig. 7.20(c) shows that the rate of producing unmatched photon predictions drops from 16% (30%) at a predicted p_{T} of 2 GeV to 1.4% (11%) above 30 GeV for HGPflow (TSPN-SA). For neutral hadrons, the fake rate (Fig. 7.20(d)) is a factor of 2-4 times larger across the full p_{T} range for HGPflow, and 1.6-2.8 times larger for TSPN-SA.

For the TSPN-SA and HGPflow algorithms, reconstructing neutral particles at low-p_{T} is challenging because a large fraction of the target particles does not contribute a leading amount of energy to any topocluster in the event (33% of photons and 25% of neutral hadrons). In HGPflow, this is compensated in a supervised manner by learning subdominant contributions to topoclusters as fractional entries in the predicted incidence matrix. This limitation could be further overcome by using cell-level input nodes such as for the OC algorithm.

Efficiency and fake rate plots are complemented by studying the probability of misclassification between photons and neutral hadrons. In Fig. 7.20(e) and Fig. 7.20(f), both HGPflow and TSPN-SA algorithms exhibit high accuracy of classification for predictions matched to photons, and for neutral hadrons an accuracy that rises with p_{T}: from 51% to 76% for HGPflow and from 16% to 63% for TSPN-SA. A lower accuracy is expected when considering that due to the class imbalance between photons and neutral hadrons of 5.8 : 1 (inclusive), a random classifier would have an accuracy of roughly 15% for neutral hadrons. Moreover, the class imbalance is p_{T}-dependent, with the proportion of photons dropping off faster than neutral hadrons.

The OC algorithm behaves similarly to the others for reconstruction efficiency of neutral particles, albeit with lower performance in most bins. For the fake rate, and in particular, for photons, OC exhibits an increase with p_{T}. The classification accuracy for OC is also lower for photons and shows a different trend in p_{T} for neutral hadrons compared to HGPflow and TSPN-SA. These differences are related to the fact that in OC neutral particle predictions correspond to a subset of calorimeter cells passing the selection defined by the t_b and t_d threshold cuts. This introduces a sensitivity of the neutral particle cardinality to the number of cells per particle, which grows as a function of particle p_{T}. The trend in fake rate appears to reflect this. Furthermore, towards high p_{T} a growing majority of cells belong to showers of charged particles, which makes the classification task more challenging in the OC approach (which involves 3 classes, unlike HGPflow and TSPN-SA). Introducing p_{T}-dependent weights on the β and x prediction tasks could help counter this effect.



(a)

(b)

(c)

Figure 7.21. Distributions of relative residuals between predicted and true p_{T} (a), η (b), and ϕ (c) of reconstructed neutral particles.

Finally, relative residuals are used to quantify the ability to correctly predict the p_{T}, η, and ϕ of the reconstructed neutral particles, shown in Figs. 7.21(a), 7.21(b), and 7.21(c), respectively. The HGPflow algorithm shows the best performance at estimating accurately both angular variables and momentum. It is interesting to note that the TSPN-SA algorithm has the worst performance for the angular variables. This is related to the usage of topoclusters in a less supervised way compared to HGPflow, where it is known which topoclusters contributed to the formation of a particle. The OC algorithm instead uses the more granular calorimeter cells directly showing similar performance to the HGPflow algorithm for angular regression.

The models also perform differently for neutral particle p_{T} regression, shown in Fig. 7.21(a). The OC model has a tendency to overestimate the neutral particle p_{T}. A similar trend is less pronounced in the TSPN-SA regression, while the distributions of predictions from HGPflow exhibit the least skew, in addition to the smallest mean and variance of the three.

Jet-level performance

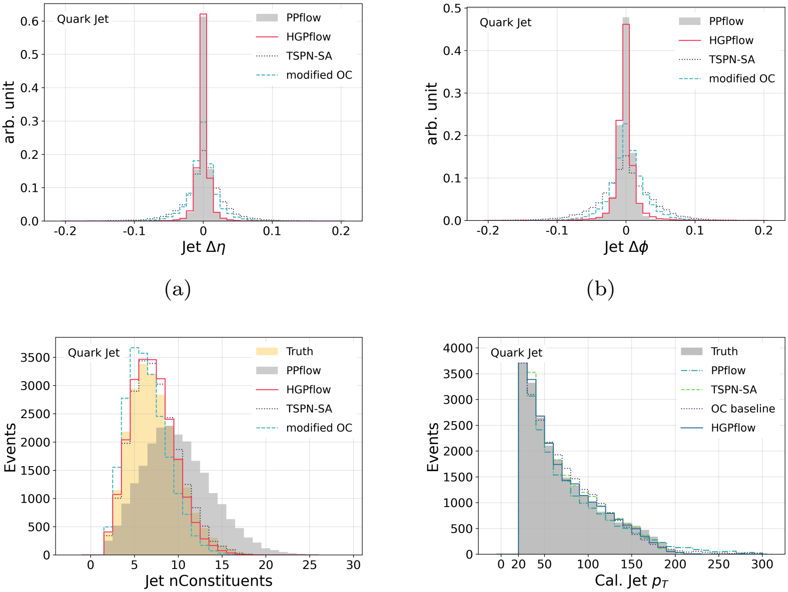
The ability to efficiently reconstruct jets and correctly predict their properties is a priority for experiments at the LHC. Jet performance depends on the overall efficiency, fake-rates and kinematic regression of the constituent particles, therefore being an important test for the ML reconstruction algorithms.

Following evaluation of the networks, jets are built using the anti-k$_t$ algorithm [55] with a radius parameter of 0.4 and a minimum number of 2 constituents. Three sets of jets with differing input constituents are defined:

- **Truth jets**: jets built using the set of target particles
- **ML jets**: jets built using the sets of particles predicted by the OC, TSPN-SA, and HGPflow algorithms
- **PPflow jets**: jets built using tracks and topoclusters with the charged energy subtraction procedure of a parameterized particle-flow algorithm (see Section 7.2.2)

The number of constituents is shown in Fig. 7.22(c) for each algorithm and compared to that of the true jet. The ML algorithms, by accounting for neutral particles in the jet, are able to model this reasonably well. On the other hand, the PPflow distribution overestimates the truth distribution as expected, since its constituents are tracks and topoclusters rather than particles.

In order to further optimize the jet-p_{T} and provide a more quantitative figure of the jet resolutions, a simplistic calibration is applied. First jet p_{T} residual distributions are computed in different p_{T} bins. For each, a dedicated scale factor is computed. A functional fit is performed and the corresponding scale factor is applied to reconstructed



(a) (b)

(c) (d)

Figure 7.22. Jet-level performance metrics shown for the three algorithms and the PPflow reference in comparison to true jets. Angular residuals between reconstructed and true jets (a and b), the number of jet-constituents in (c), and the calibrated p_T relative residuals in (d).

jets based on their p_T. This procedure is applied separately to each reconstruction algorithm. Relative residuals are shown in Fig. 7.22. As observed for neutral particles, HGPflow shows the best performance at the jet-level. In terms of jet angular observables, HGPflow is comparable to the traditional PPflow approach while for jet p_T resolution it shows a 24% improvement relative to PPflow. The TSPN-SA jet p_T resolution is better than OC, while for angular observables OC performs slightly better.

To help visualize the jet reconstruction task, Fig. 7.23 displays an event from the test dataset showing predictions from the trained HGPflow algorithm. In this example, each of the four neutral particles at truth level can be matched to a predicted particle with the correct class and an η-ϕ prediction consistent within the cell granularity. The

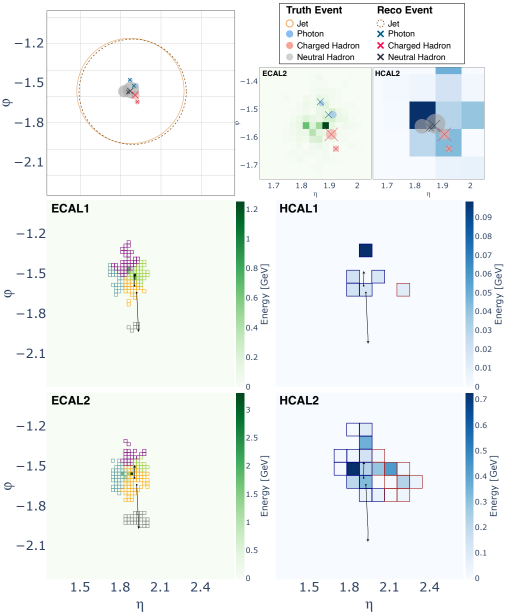
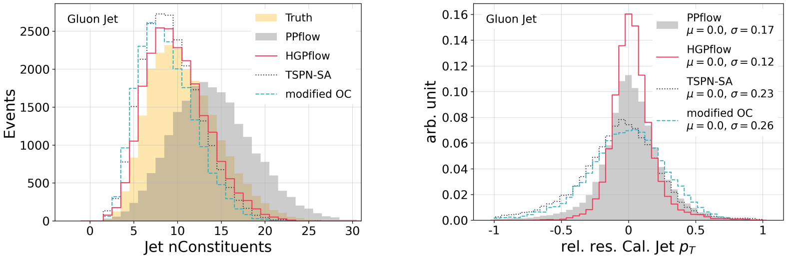


Figure 7.23. Event display of a single jet event from the test dataset. In the top left panel, the $\eta - \phi$ coordinates of truth particles with momentum above 1 GeV are shown as circles along with the set of predicted particles from the HGPflow algorithm shown as crosses. The set of particles and their p_T at truth level are as follows: two photons in blue (1.8, 3.0 GeV), a pair of neutral K_L^0 mesons in grey (12.3, 22.3 GeV), and two charged pions in red (2.2, 6.5 GeV). The circles of $R = 0.4$ represent anti-k$_t$ jets built from the truth (solid) and predicted (dashed) particle sets that nearly overlap in η-ϕ and have p_T agreement within 35%. In the top right panels, the truth and predicted particles are shown overlaid on a zoomed region of the ECAL2 and HCAL2 layers. In the bottom panels, the detector-level information serving as input to the reconstruction algorithms is shown for each of the first two layers of both ECAL and HCAL in the same $\eta - \phi$ plane. Cells that have the same border color belong to the same topocluster. Green and blue fill is used to indicate the energy of cells in the ECAL and HCAL layers, respectively. The arrows describe the tracks for charged particles from the interaction point with the arrowheads indicating the angular coordinate extrapolated at the given layer.

calorimeter panels illustrate the tight arrangement of topoclusters used as input nodes for the HGPflow prediction

Performance on gluon jets



(a) (b)

Figure 7.24. Distributions of the number of constituents per jet (a), and distributions of relative residuals between predicted and true p_{T} (b) for the three ML particle reconstruction algorithms evaluated on the sample of gluon jets.

To study the ability of the particle reconstruction algorithms to generalize beyond the training data to a new physics process, the trained models is evaluated on the dataset of single gluon jet events described in Section 7.2.1. The difference arising at the parton shower for gluon-initiated jets reflects itself in the dataset feature distributions, for example, the larger multiplicity of cells, tracks, and particles shown in Fig. 7.13. Since the appropriate upper bound on the number of particles was determined based on the training dataset, a single gluon jet found to contain > 30 particles was excluded from the dataset in order to evaluate HGPflow. The results for the three ML algorithms are shown along with the PPflow comparison in Fig. 7.24. Overall, the algorithms demonstrate an ability to generalize: the number of predicted constituents is shifted slightly lower with respect to truth compared to the quark jet case (Fig. 7.22(c)), while the jet relative p_{T} residual distributions are comparable to Fig. 7.22(d). The rank of the algorithms in terms of performance remains the same as before, with HGPflow again boasting narrower jet p_{T} resolution than the PPflow comparison.

7.2.4 Perspectives

Compared to the OC and TSPN-SA algorithms and the PPflow benchmark, HGPflow shows the best performance in terms of jet momentum resolution, which was not directly a training objective. This traces back to superior modeling of neutral particle momentum shown in Fig. 7.21(a). Unlike the other ML models, which must learn implicitly that a given energy deposit in the calorimeter cannot be associated with more than one parent particle, HGPflow benefits from being structured around the concept of energy conservation. Successfully predicting an incidence matrix defined via Eq. 7.4 and the hyperedge indicator row entails knowing the energy contributions a given topocluster received from all particles (Fig. 7.18). Furthermore, the normalization ensures that energy attributed to a given particle candidate is not counted again in assignments to other particles. Since both the hyperedge representation and the proxy for neutral particle energy (Eq. 7.9) are weighted by entries of the incidence matrix, the property predictions which stem from these inputs inherit a bias towards energy conservation.

The hypergraph approach allows common elements of both the OC and TSPN-SA approaches to be handled in a more clear formalism. In the OC potential loss (Eq. 7.2), a binary-valued incidence matrix I_{ik} functions as a lookup table determining whether a node is repelled or attracted to the representation of a particle (i.e. condensation point). The clustering of nodes according to parent particle can thus be thought of as an indirect way of learning I_{ik}, limited by the extent to which the injective condition applies. Likewise, the TSPN-SA algorithm is built around an attention matrix between particle candidates k and nodes i from the input set which resembles an incidence matrix, although it is normalized along columns rather than rows. The attention weights also have a latent rather than physical meaning and are learned in an unsupervised way. On the other hand, HGPflow not only explicitly predicts the incidence matrix, which is the key to unraveling overlapping particle showers, but expresses it in the physical basis of energy contributions with the advantage mentioned previously.

It can be anticipated that the structure of HGPflow can be extended in at least three ways. First, the input set granularity has been set without tuning to that of topoclusters from a standard ATLAS-like algorithm. This granularity can be increased to further enable the segmentation of overlapping energy deposits from nearby particles.

Second, in the trainings, the two objectives of incidence and properties prediction have been carried out nearly independently. However, a more powerful representation learning scheme could lead to a model which learns these two objectives in a synergistic way, allowing the incidence prediction to be informed by the properties prediction and vice versa. Finally, while Tab. 7.3 indicates an acceptable inference time of HGPflow, more optimization is needed to reduce its training duration. This could be achieved by hyperparameter optimization in the recurrent training configuration and by exploring alternatives to the Hungarian matching, which the authors of [197] identified as a computational bottleneck.

Similarly, besides the modifications proposed in [184], it is reasonable to suspect that the OC algorithm can be substantially improved in future work. While the TSPN-SA and HGPflow algorithms both have neural network layers for information exchange following the node encoding (i.e. successive attention and incidence weighted-updates), the node predictions of the modified OC algorithm are likely limited by a comparatively narrow receptive field. One way to increase the receptive field in OC is to add additional message passing blocks in the node encoder model, although a limited study of this option did not lead to conclusive improvement. Another possibility is introducing an attention mechanism. The distance between nodes i and j in the latent clustering space entails a term $a_{ij} = x_i \cdot x_j$ with the form of attention, e.g. in the attractive potential, $\breve{V} \propto \mathit{\Delta} x^2 \ni -2a_{ij}$. Therefore the clustering mechanism in OC seems a natural point to introduce transformer blocks for enhanced information exchange, at the cost of computing additional gradients for the set of edges.

Several opportunities emerge for future investigation on new datasets. The performance reported in this work for $R = 0.4$ quark and gluon jets suggests the application to substructure reconstruction in large-R jets from boosted boson decays. The goal of studying particle reconstruction on a single jet dataset was to focus on the local system of overlapping particle showers which represents the kernel of the problem at the full-event level. For this reason, it is reasonable that reconstructing full events could proceed by mapping the same trained model onto spatial partitions of the detector hits D defined by topological and jet clustering algorithms, for instance. Given an effective scheme to deal with potential overlap, each partition could be treated as an approximately isolated set of input nodes, graph edges, and attention or incidence matrix weights. In this case the resource requirements reported in Tab. 7.3 would scale linearly with the number of partitions. Ignoring overlap, an upper bound on the number of $R = 0.4$ partitions of a full event could be estimated as $2\pi \cdot 6/0.4 \simeq 10^2$.

Studying the robustness of the models in the presence of pileup will also be an important follow-up task (see [185] for existing work in this direction). Likewise, the impact of interactions with material upstream of the calorimeter needs to be thoroughly addressed in a future dataset. In this case, electron pair production from early conversions and photons from bremsstrahlung will require a thoughtful definition of target particles to ensure that they can be feasibly distinguished during training. For example, photon conversions could be treated as a separate class, and electrons could be defined as targets depending on the quality of their associated track (if present at all).

For charged particles, the algorithms learned to exploit the complementary information provided by calorimeter activity to improve on the measured track momentum. The efficiency of reconstructing photons and neutral hadrons reached 90% and 80%, respectively, for $p_\mathrm{T} > 10$ GeV. The neutral particle fake rates were more variable for each algorithm, with the best performance being 10% and lower for $p_\mathrm{T} > 10$ GeV.

Jets formed from the predicted particles were compared to those from the true particles and also a parameterized particle flow baseline. HGPflow showed the best performance and surpassed the baseline in terms of both angular and momentum resolution of the jet. This can be explained from the fact that the hypergraph formalism is structured around energy conservation, which also makes its predictions more interpretable from a physics point of view. The suitability of the hypergraph formalism for the set-to-set task of particle reconstruction is yet to be fully leveraged. By demonstrating the potential of ML algorithms to disentangle the jet dense environment, these findings motivate the application to full collision events in future work.

8 | Conclusions

The Standard Model stands as the most precise theory available for understanding the origin and mechanics of the universe. Despite its successes, it leaves several intriguing questions unanswered, making it essential to push the boundaries of High Energy Physics experiments.

One of the most compelling questions is about the nature of the Higgs boson. Since its discovery in 2012, it has been crucial to determine whether the observed particle conforms to the predictions of the Standard Model or exhibits any deviations. So far, no significant discrepancies have been detected, but many production and decay channels remain unexplored.

To probe and study the Higgs boson, significant technological and scientific efforts have been undertaken. The Large Hadron Collider (LHC) was constructed, featuring two multi-purpose experiments, ATLAS and CMS, dedicated to this and many other aspects.

This dissertation discusses the Higgs boson production in association with a leptonically decaying Vector boson using the ATLAS Detector at the LHC. The Higgs boson is studied in its decays to heavy flavor quarks (b or c). The leptonic decay of the Vector boson is leveraged to enhance the signature's clarity, making this production channel more suitable compared to the gluon-gluon fusion process, which is obscured by QCD background. Given that the branching ratios for the two decay channels are approximately 60% and 3% respectively, one of the major challenges is entrusted to Flavor Tagging algorithms, which discriminate between the different quark flavors produced in the Higgs decay.

This analysis was conducted using the complete Run 2 dataset, implementing new Flavor Tagging algorithms and Multivariate Analysis techniques to simultaneously extract the Higgs coupling to b and c quarks.

The primary results of this analysis include the first observation of the single process $WH, H \to b\bar{b}$, the first observation of the process $VZ, Z \to c\bar{c}$ and setting the best ATLAS limits on the process $VH, H \to c\bar{c}$.

This measurement is fundamentally important as it was previously thought impossible to observe the Yukawa coupling of the c quark during the LHC era. However, recent studies and advanced technologies have shown that this might be achievable.

Related to Higgs physics and its implications on the Standard Model, this dissertation also describes one of the expected milestones for the HL-LHC: the measurement of the Higgs self-coupling. This parameter is directly probed through Di-Higgs analyses. One of the best channels, and the focus of this dissertation, is that with two Higgs bosons decay into 2 b-quarks and 2 photons respectively. Despite the low branching fraction of this process, the purity in the photon system allows for high sensitivity. This dissertation outlines the potential impact of new studies on the sensitivity of this process, particularly how innovative b-tagging techniques could significantly enhance this measurement.

The ability to achieve these significant milestones is closely linked to the experiment's performance. Specifically, the described analyses heavily rely on Flavor Tagging algorithms aimed at identifying the parton flavor produced by the Higgs boson. Improving these algorithms' performance translates to greater sensitivity in the involved measurements.

This dissertation demonstrates, for the first time, the potential impact of a 4-Dimensional tracker in ATLAS for the HL-LHC, highlighting improvements in vertexing and b-tagging performance. Additionally, it shows how these performance enhancements lead to better physics analysis outcomes. The 4D trackers are likely to be installed in next-generation collider experiments, making these studies relevant not only for the LHC but also on a broader scale.

This dissertation illustrates also how innovative machine learning approaches based on Graph Neural Networks (GNN) can improve collider physics reconstruction by exploiting multidimensional correlations among low-level event variables. A detector-agnostic simulation for machine learning studies was developed, followed by a Hypergraph GNN-based approach to Global Particle Flow, demonstrating improvements over classical algorithms. These techniques could significantly maximize the reach of the LHC experiments and future colliders.

High energy physics represents the frontier of human knowledge, pushing the limits of the imaginable in every aspect. This work showcases how human collaboration enables achieving great results, both individually and collectively.

Bibliography

[1] ATLAS Collaboration. *Investigating the impact of 4D Tracking in ATLAS Beyond Run 4*. Tech. rep. Geneva: CERN, 2023. URL: `https://cds.cern.ch/record/2870326`.

[2] F. A. Di Bello et al. “Towards a computer vision particle flow”. In: *Eur. Phys. J. C.* 81.2 (2021). ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-021-08897-0`. URL: `http://dx.doi.org/10.1140/epjc/s10052-021-08897-0`.

[3] F. A. Di Bello et al. “Configurable calorimeter simulation for AI applications”. In: *preprint: arXiv:2303.02101* (2023).

[4] F. A. Di Bello et al. “Reconstructing particles in jets using set transformer and hypergraph prediction networks”. In: (Dec. 2022). arXiv:2212.01328. arXiv: `2212.01328 [hep-ex]`.

[5] Albert Einstein. “On the electrodynamics of moving bodies”. In: *Annalen Phys.* 17 (1905), pp. 891–921. DOI: `10.1002/andp.200590006`.

[6] Steven Weinberg. *The Quantum theory of fields. Vol. 1: Foundations*. Cambridge University Press, June 2005. ISBN: 978-0-521-67053-1.

[7] Steven Weinberg. *The quantum theory of fields. Vol. 2: Modern applications*. Cambridge University Press, Aug. 2013. ISBN: 978-1-139-63247-8.

[8] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to quantum field theory*. Reading, USA: Addison-Wesley, 1995. ISBN: 978-0-201-50397-5.

[9] Sidney Coleman. *Aspects of Symmetry: Selected Erice Lectures*. Cambridge, U.K.: Cambridge University Press, 1985. ISBN: 978-0-521-31827-3. DOI: `10.1017/CBO9780511565045`.

[10] C. Itzykson and J.B. Zuber. *Quantum Field Theory*. International Series In Pure and Applied Physics. New York: McGraw-Hill, 1980. ISBN: 978-0-486-44568-7.

[11] Francis Halzen and Alan D Martin. *Quark & Leptons: An Introductory Course In Modern Particle Physics*. John Wiley & Sons, 2008.

[12] Yorikiyo Nagashima. *Elementary particle physics: Quantum Field Theories and Particles*. Vol. 1. Wiley Online & Sons, 2010.

[13] Yorikiyo Nagashima. *Elementary particle physics: Foundations of the standard model*. Vol. 2. Wiley Online & Sons, 2013.

[14] Howard Georgi. *Lie algebras in particle physics*. 2nd ed. Vol. 54. Reading, MA: Perseus Books, 1999.

[15] E. Fermi. "An attempt of a theory of beta radiation. 1." In: *Z. Phys.* 88 (1934), pp. 161–177. DOI: `10.1007/BF01351864`.

[16] S.L. Glashow. "Partial Symmetries of Weak Interactions". In: *Nucl. Phys.* 22 (1961), pp. 579–588. DOI: `10.1016/0029-5582(61)90469-2`.

[17] Abdus Salam and John Clive Ward. "Electromagnetic and weak interactions". In: *Phys. Lett.* 13 (1964), pp. 168–171. DOI: `10.1016/0031-9163(64)90711-5`.

[18] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (1964). Ed. by J.C. Taylor, pp. 321–323. DOI: `10.1103/PhysRevLett.13.321`.

[19] Peter W. Higgs. "Broken symmetries, massless particles and gauge fields". In: *Phys. Lett.* 12 (1964), pp. 132–133. DOI: `10.1016/0031-9163(64)91136-9`.

[20] Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (1964). Ed. by J.C. Taylor, pp. 508–509. DOI: `10.1103/PhysRevLett.13.508`.

[21] Steven Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266. DOI: `10.1103/PhysRevLett.19.1264`.

[22] Nicola Cabibbo. "Unitary symmetry and leptonic decays". In: *Physical Review Letters* 10.12 (1963), p. 531.

[23] Murray Gell-Mann and M Levy. "The axial vector current in beta decay". In: *Nuovo Cim.* 16 (1960), p. 705. DOI: `10.1007/BF02859738`.

[24] Gustavo C. Branco, Luis Lavoura, and Joao P. Silva. *CP Violation.* Vol. 103. Oxford Science Publication, 1999.

[25] S.L. Glashow, J. Iliopoulos, and L. Maiani. “Weak Interactions with Lepton-Hadron Symmetry”. In: *Phys. Rev. D* 2 (1970), pp. 1285–1292. DOI: `10.1103/PhysRevD.2.1285`.

[26] Makoto Kobayashi and Toshihide Maskawa. “CP Violation in the Renormalizable Theory of Weak Interaction”. In: *Prog. Theor. Phys.* 49 (1973), pp. 652–657. DOI: `10.1143/PTP.49.652`.

[27] Taizo Muta. *Foundations of Quantum Chromodynamics: An Introduction to Perturbative Methods in Gauge Theories, (3rd ed.)* 3rd. Vol. 78. World scientific Lecture Notes in Physics. Hackensack, N.J.: World Scientific, 2010. ISBN: 978-981-279-353-9.

[28] R.N. Cahn and G. Goldhaber. *The experimental foundations of particle physics.* Cambridge: Cambridge Univ. Press, 2009. ISBN: 978-0-521-52147-5. DOI: `10.1017/CBO9780511609923`.

[29] Murray Gell-Mann. “A Schematic Model of Baryons and Mesons”. In: *Phys. Lett.* 8 (1964), pp. 214–215. DOI: `10.1016/S0031-9163(64)92001-3`.

[30] G. Zweig. “An SU(3) model for strong interaction symmetry and its breaking. Version 1”. In: *CERN Document Server* (Jan. 1964).

[31] David J. Gross and Frank Wilczek. “Ultraviolet Behavior of Nonabelian Gauge Theories”. In: *Phys. Rev. Lett.* 30 (1973). Ed. by J.C. Taylor, pp. 1343–1346. DOI: `10.1103/PhysRevLett.30.1343`.

[32] H.David Politzer. “Reliable Perturbative Results for Strong Interactions?” In: *Phys. Rev. Lett.* 30 (1973). Ed. by J.C. Taylor, pp. 1346–1349. DOI: `10.1103/PhysRevLett.30.1346`.

[33] Chen-Ning Yang and Robert L. Mills. “Conservation of Isotopic Spin and Isotopic Gauge Invariance”. In: *Phys. Rev.* 96 (1954). Ed. by Jong-Ping Hsu and D. Fine, pp. 191–195. DOI: `10.1103/PhysRev.96.191`.

[34] Guido Altarelli and G. Parisi. “Asymptotic Freedom in Parton Language”. In: *Nucl. Phys. B* 126 (1977), pp. 298–318. DOI: `10.1016/0550-3213(77)90384-4`.

[35] G. Aad et al. “The ATLAS Experiment at the CERN Large Hadron Collider”. In: *JINST* 3 (2008), S08003. DOI: `10.1088/1748-0221/3/08/S08003`.

[36] S. Chatrchyan et al. “The CMS Experiment at the CERN LHC”. In: *JINST* 3 (2008), S08004. DOI: `10.1088/1748-0221/3/08/S08004`.

[37] Helmut Wiedemann. *Particle Accelerator Physics*. Graduate Texts in Physics. Berlin, Germany: Springer, 2015. ISBN: 978-3-319-18316-9. DOI: `10.1007/978-3-319-18317-6`.

[38] John David Jackson. *Classical Electrodynamics*. Wiley, 1998. ISBN: 978-0-471-30932-1.

[39] Claus Grupen and Boris Schwartz. *Particle detectors*. Vol. 26. Cambridge, UK: Cambridge Univ. Pr., 2008.

[40] M Capeans et al. *ATLAS Insertable B-Layer Technical Design Report*. Tech. rep. 2010. URL: `https://cds.cern.ch/record/1291633`.

[41] Bernd Stelzer. *The New Small Wheel Upgrade Project of the ATLAS Experiment*. Tech. rep. Geneva: CERN, 2016. DOI: `10.1016/j.nuclphysbps.2015.09.182`. URL: `https://cds.cern.ch/record/1958265`.

[42] *ATLAS Computing: technical design report*. Technical design report. ATLAS. Geneva: CERN, 2005. URL: `https://cds.cern.ch/record/837738`.

[43] “Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2”. In: *The European Physical Journal C* 77.10 (Oct. 2017). ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-017-5225-7`. URL: `http://dx.doi.org/10.1140/epjc/s10052-017-5225-7`.

[44] H. J. Behrend et al. “An Analysis of the Charged and Neutral Energy Flow in e^+e^- Hadronic Annihilation at 34-GeV, and a Determination of the QCD Effective Coupling Constant”. In: *Phys. Lett.* 113B (1982), pp. 427–432. DOI: `10.1016/0370-2693(82)90778-X`.

[45] D. Buskulic et al. “Performance of the ALEPH detector at LEP”. In: *Nucl. Instrum. Meth.* A360 (1995), pp. 481–506. DOI: `10.1016/0168-9002(95)00138-7`.

[46] Morad Aaboud et al. "Jet reconstruction and performance using particle flow with the ATLAS Detector". In: *The European Physical Journal C* 77.7 (2017), p. 466.

[47] CMS Collaboration. "Particle-flow reconstruction and global event description with the CMS detector". In: *JINST* 12.10 (2017), P10003. DOI: `10.1088/1748-0221/12/10/P10003`. arXiv: `1706.04965 [physics.ins-det]`.

[48] "Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13$ TeV". In: *The European Physical Journal C* 79.8 (Aug. 2019). ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-019-7140-6`. URL: `http://dx.doi.org/10.1140/epjc/s10052-019-7140-6`.

[49] *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*. Tech. rep. Geneva: CERN, 2012. URL: `https://cds.cern.ch/record/1449796`.

[50] *Electron and photon reconstruction and performance in ATLAS using a dynamical, topological cell clustering-based approach*. Tech. rep. Geneva: CERN, 2017. URL: `https://cds.cern.ch/record/2298955`.

[51] Morad Aaboud et al. "Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run 2 data collected in 2015 and 2016". In: *Eur. Phys. J. C* 79.3 (2019), p. 205. DOI: `10.1140/epjc/s10052-019-6650-6`. arXiv: `1810.05087 [hep-ex]`.

[52] Georges Aad et al. "Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data". In: *JINST* 14.12 (2019), P12006. DOI: `10.1088/1748-0221/14/12/P12006`. arXiv: `1908.00005 [hep-ex]`.

[53] Morad Aaboud et al. "Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data". In: *JINST* 14.03 (2019), P03017. DOI: `10.1088/1748-0221/14/03/P03017`. arXiv: `1812.03848 [hep-ex]`.

[54] Georges Aad et al. "Muon reconstruction and identification efficiency in ATLAS using the full Run 2 *pp* collision data set at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J. C* 81.7 (2021), p. 578. DOI: `10.1140/epjc/s10052-021-09233-2`. arXiv: `2012.00578 [hep-ex]`.

[55] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The anti-k_t jet clustering algorithm". In: *JHEP* 04 (2008), p. 063. DOI: `10.1088/1126-6708/2008/04/063`. arXiv: `0802.1189 [hep-ph]`.

[56] S. Catani et al. "Longitudinally invariant K_t clustering algorithms for hadron hadron collisions". In: *Nucl. Phys. B* 406 (1993), pp. 187–224. DOI: `10.1016/0550-3213(93)90166-M`.

[57] Yuri L. Dokshitzer et al. "Better jet clustering algorithms". In: *JHEP* 08 (1997), p. 001. DOI: `10.1088/1126-6708/1997/08/001`. arXiv: `hep-ph/9707323`.

[58] "Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$TeV with the ATLAS detector". In: *The European Physical Journal C* 81.8 (Aug. 2021). ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-021-09402-3`. URL: `http://dx.doi.org/10.1140/epjc/s10052-021-09402-3`.

[59] "Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$TeV using the ATLAS detector". In: *The European Physical Journal C* 76.11 (Oct. 2016). ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-016-4395-z`. URL: `http://dx.doi.org/10.1140/epjc/s10052-016-4395-z`.

[60] "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". In: *The European Physical Journal C* 77.7 (July 2017). ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-017-5004-5`. URL: `http://dx.doi.org/10.1140/epjc/s10052-017-5004-5`.

[61] David Krohn, Jesse Thaler, and Lian-Tao Wang. "Jet trimming". In: *Journal of High Energy Physics* 2010.2 (Feb. 2010). ISSN: 1029-8479. DOI: `10.1007/jhep02(2010)084`. URL: `http://dx.doi.org/10.1007/JHEP02(2010)084`.

[62] "Identification of boosted, hadronically decaying W bosons and comparisons with ATLAS data taken at $\sqrt{s} = 8$ TeV". In: *The European Physical Journal C* 76.3 (Mar. 2016). ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-016-3978-z`. URL: `http://dx.doi.org/10.1140/epjc/s10052-016-3978-z`.

[63] "Performance of jet substructure techniques for large-R jets in proton-proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector". In: *Journal of High Energy Physics* 2013.9 (Sept. 2013). ISSN: 1029-8479. DOI: `10.1007/jhep09(2013)076`. URL: `http://dx.doi.org/10.1007/JHEP09(2013)076`.

[64] “In situ calibration of large-radius jet energy and mass in 13 TeV proton–proton collisions with the ATLAS detector”. In: *The European Physical Journal C* 79.2 (Feb. 2019). ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-019-6632-8`. URL: `http://dx.doi.org/10.1140/epjc/s10052-019-6632-8`.

[65] “Optimisation of large-radius jet reconstruction for the ATLAS detector in 13TeV proton–proton collisions”. In: *The European Physical Journal C* 81.4 (Apr. 2021). ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-021-09054-3`. URL: `http://dx.doi.org/10.1140/epjc/s10052-021-09054-3`.

[66] *Jet mass reconstruction with the ATLAS Detector in early Run 2 data.* Tech. rep. Geneva: CERN, 2016. URL: `https://cds.cern.ch/record/2200211`.

[67] David Krohn, Jesse Thaler, and Lian-Tao Wang. “Jets with variable R”. In: *Journal of High Energy Physics* 2009.06 (June 2009), pp. 059–059. ISSN: 1029-8479. DOI: `10.1088/1126-6708/2009/06/059`. URL: `http://dx.doi.org/10.1088/1126-6708/2009/06/059`.

[68] *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC.* Tech. rep. Geneva: CERN, 2015. URL: `https://cds.cern.ch/record/2064383`.

[69] *Identification of hadronic tau lepton decays using neural networks in the ATLAS experiment.* Tech. rep. Geneva: CERN, 2019. URL: `https://cds.cern.ch/record/2688062`.

[70] Morad Aaboud et al. “Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV”. In: *Eur. Phys. J. C* 78.11 (2018), p. 903. DOI: `10.1140/epjc/s10052-018-6288-9`. arXiv: `1802.08168 [hep-ex]`.

[71] Yasmine Sara Amhis et al. “Averages of b-hadron, c-hadron, and τ-lepton properties as of 2018”. In: *Eur. Phys. J. C* 81.3 (2021), p. 226. DOI: `10.1140/epjc/s10052-020-8156-7`. arXiv: `1909.12524 [hep-ex]`.

[72] *Jet Flavour Tagging With GN1 and DL1d. Generator dependence, Run 2 and Run 3 data agreement studies.* Tech. rep. Geneva: CERN, 2023. URL: `https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2023-01/`.

[73] Georges Aad et al. “ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset”. In: *Eur. Phys. J. C* 83.7 (2023), p. 681. DOI: `10.1140/epjc/s10052-023-11699-1`. arXiv: `2211.16345 [physics.data-an]`.

[74] Stefano Frixione, Paolo Nason, and Giovanni Ridolfi. “A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction”. In: *JHEP* 09 (2007), p. 126. DOI: `10.1088/1126-6708/2007/09/126`. arXiv: `0707.3088 [hep-ph]`.

[75] Stefano Frixione, Paolo Nason, and Carlo Oleari. “Matching NLO QCD computations with Parton Shower simulations: the POWHEG method”. In: *JHEP* 11 (2007), p. 070. DOI: `10.1088/1126-6708/2007/11/070`. arXiv: `0709.2092 [hep-ph]`.

[76] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. “A Brief Introduction to PYTHIA 8.1”. In: *Comput. Phys. Commun.* 178 (2008), pp. 852–867. DOI: `10.1016/j.cpc.2008.01.036`. arXiv: `0710.3820 [hep-ph]`.

[77] *Development of ATLAS Primary Vertex Reconstruction for LHC Run 3*. Tech. rep. Geneva: CERN, 2019. URL: `https://cds.cern.ch/record/2670380`.

[78] *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run.* Tech. rep. Geneva: CERN, 2017. URL: `https://cds.cern.ch/record/2273281`.

[79] *Identification of Jets Containing b-Hadrons with Recurrent Neural Networks at the ATLAS Experiment.* Tech. rep. Geneva: CERN, 2017. URL: `https://cds.cern.ch/record/2255226`.

[80] *Deep Sets based Neural Networks for Impact Parameter Flavour Tagging in ATLAS*. Tech. rep. Geneva: CERN, 2020. URL: `https://cds.cern.ch/record/2718948`.

[81] *Secondary vertex finding for jet flavour identification with the ATLAS detector*. Tech. rep. Geneva: CERN, 2017. URL: `https://cds.cern.ch/record/2270366`.

[82] *Topological b-hadron decay reconstruction and identification of b-jets with the JetFitter package in the ATLAS experiment at the LHC.* Tech. rep. Geneva: CERN, 2018. URL: `https://cds.cern.ch/record/2645405`.

[83] Manzil Zaheer et al. *Deep Sets*. 2018. arXiv: `1703.06114 [cs.LG]`.

[84] *Performance of the Run 3 ATLAS b-tagging algorithms*. Tech. rep. Geneva: CERN, 2018. URL: `https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2022-004/`.

[85] *Graph Neural Network Jet Flavour Tagging with the ATLAS Detector*. Tech. rep. Geneva: CERN, 2022. URL: `https://cds.cern.ch/record/2811135`.

[86] Ashish Vaswani et al. "Attention is all you need". In: *Advances in neural information processing systems* 30 (2017).

[87] Georges Aad et al. "ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J. C* 79.11 (2019), p. 970. DOI: `10.1140/epjc/s10052-019-7450-8`. arXiv: `1907.05120 [hep-ex]`.

[88] Georges Aad et al. "Measurement of the c-jet mistagging efficiency in $t\bar{t}$ events using pp collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS detector". In: *Eur. Phys. J. C* 82.1 (2022), p. 95. DOI: `10.1140/epjc/s10052-021-09843-w`. arXiv: `2109.10627 [hep-ex]`.

[89] *Calibration of light-flavour b-jet mistagging rates using ATLAS proton-proton collision data at $\sqrt{s} = 13$ TeV*. Tech. rep. Geneva: CERN, 2018. URL: `https://cds.cern.ch/record/2314418`.

[90] *Measurement of the b-jet identification efficiency for high transverse momentum jets in $t\bar{t}$ events in the lepton + jets channel with the ATLAS detector using Run 2 data*. Tech. rep. Geneva: CERN, 2021. URL: `https://cds.cern.ch/record/2753734`.

[91] Georges Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Lett. B* 716 (2012), pp. 1–29. DOI: `10.1016/j.physletb.2012.08.020`. arXiv: `1207.7214 [hep-ex]`.

[92] Serguei Chatrchyan et al. "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC". In: *Phys. Lett. B* 716 (2012), pp. 30–61. DOI: `10.1016/j.physletb.2012.08.021`. arXiv: `1207.7235 [hep-ex]`.

[93] M. Tanabashi et al. "Review of Particle Physics". In: *Phys. Rev. D* 98.3 (2018), p. 030001. DOI: `10.1103/PhysRevD.98.030001`.

[94] M Carena et al. “Status of Higgs boson physics”. In: *Review of Particle Physics, Chin. Phys., Chin. Phys. C* 40 (2016), p. 100001.

[95] Georges Aad et al. “A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery”. In: *Nature* 607.7917 (2022). [Erratum: Nature 612, E24 (2022)], pp. 52–59. DOI: `10.1038/s41586-022-04893-w`. arXiv: `2207.00092 [hep-ex]`.

[96] J R Andersen et al. “Handbook of LHC Higgs Cross Sections: 3. Higgs Properties”. In: (July 2013). Ed. by S Heinemeyer et al. DOI: `10.5170/CERN-2013-004`. arXiv: `1307.1347 [hep-ph]`.

[97] A. Djouadi, J. Kalinowski, and M. Spira. “HDECAY: A Program for Higgs boson decays in the standard model and its supersymmetric extension”. In: *Comput. Phys. Commun.* 108 (1998), pp. 56–74. DOI: `10.1016/S0010-4655(97)00123-9`. arXiv: `hep-ph/9704448`.

[98] *Combination of measurements of Higgs boson production in association with a W or Z boson in the $b\bar{b}$ decay channel with the ATLAS experiment at $\sqrt{s} = 13$ TeV.* Tech. rep. Geneva: CERN, 2021. URL: `https://cds.cern.ch/record/2782535`.

[99] Georges Aad et al. “Direct constraint on the Higgs-charm coupling from a search for Higgs boson decays into charm quarks with the ATLAS detector”. In: *Eur. Phys. J. C* 82 (2022), p. 717. DOI: `10.1140/epjc/s10052-022-10588-3`. arXiv: `2201.11428 [hep-ex]`.

[100] Morad Aaboud et al. “Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector”. In: *Phys. Lett. B* 786 (2018), pp. 59–86. DOI: `10.1016/j.physletb.2018.09.013`. arXiv: `1808.08238 [hep-ex]`.

[101] A. M. Sirunyan et al. “Observation of Higgs boson decay to bottom quarks”. In: *Phys. Rev. Lett.* 121.12 (2018), p. 121801. DOI: `10.1103/PhysRevLett.121.121801`. arXiv: `1808.08242 [hep-ex]`.

[102] Georges Aad et al. “Measurements of WH and ZH production in the $H \to b\bar{b}$ decay channel in pp collisions at 13 TeV with the ATLAS detector”. In: *Eur. Phys. J. C* 81.2 (2021), p. 178. DOI: `10.1140/epjc/s10052-020-08677-2`. arXiv: `2007.02873 [hep-ex]`.

[103] Georges Aad et al. “Measurement of the associated production of a Higgs boson decaying into b-quarks with a vector boson at high transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Phys. Lett. B* 816 (2021), p. 136204. DOI: `10.1016/j.physletb.2021.136204`. arXiv: `2008.02508 [hep-ex]`.

[104] Armen Tumasyan et al. “Search for Higgs Boson Decay to a Charm Quark-Antiquark Pair in Proton-Proton Collisions at s=13 TeV”. In: *Phys. Rev. Lett.* 131.6 (2023), p. 061801. DOI: `10.1103/PhysRevLett.131.061801`. arXiv: `2205.05550 [hep-ex]`.

[105] Morad Aaboud et al. “Measurement of VH, H $\to$ b$\overline{\text{b}}$ production as a function of the vector-boson transverse momentum in 13 TeV pp collisions with the ATLAS detector”. In: *JHEP* 05 (2019), p. 141. DOI: `10.1007/JHEP05(2019)141`. arXiv: `1903.04618 [hep-ex]`.

[106] Simone Alioli et al. “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”. In: *JHEP* 06 (2010), p. 043. DOI: `10.1007/JHEP06(2010)043`. arXiv: `1002.2581 [hep-ph]`.

[107] Richard D. Ball et al. “Parton distributions for the LHC Run II”. In: *JHEP* 04 (2015), p. 040. DOI: `10.1007/JHEP04(2015)040`. arXiv: `1410.8849 [hep-ph]`.

[108] Torbjörn Sjöstrand et al. “An introduction to PYTHIA 8.2”. In: *Comput. Phys. Commun.* 191 (2015), pp. 159–177. DOI: `10.1016/j.cpc.2015.01.024`. arXiv: `1410.3012 [hep-ph]`.

[109] G. Cullen et al. “Automated One-Loop Calculations with GoSam”. In: *PoS* RADCOR2011 (2011), p. 013. DOI: `10.22323/1.145.0013`. arXiv: `1201.2782 [hep-ph]`.

[110] Keith Hamilton, Paolo Nason, and Giulia Zanderighi. “MINLO: Multi-Scale Improved NLO”. In: *JHEP* 10 (2012), p. 155. DOI: `10.1007/JHEP10(2012)155`. arXiv: `1206.3572 [hep-ph]`.

[111] Gionata Luisoni et al. “$HW^{\pm}$/HZ + 0 and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO”. In: *JHEP* 10 (2013), p. 083. DOI: `10.1007/JHEP10(2013)083`. arXiv: `1306.2542 [hep-ph]`.

[112] F Ahmadov et al. *Supporting Document for the Search for the bb decay of the Standard Model Higgs boson in associated (W/Z)H production with the ATLAS detector*. Tech. rep. Geneva: CERN, 2014. URL: `https://cds.cern.ch/record/1645654`.

[113] G Aad et al. *Invariant Mass Studies for the $H \to bb$ Measurements for LHCP*. Tech. rep. Geneva: CERN, 2013. URL: `https://cds.cern.ch/record/1541965`.

[114] Andrea Coccaro et al. *Flavor Tagging Efficiency Parametrisations with Graph Neural Networks*. Tech. rep. Geneva: CERN, 2022. URL: `https://cds.cern.ch/record/2810447`.

[115] Francesco Armando Di Bello et al. "Efficiency Parameterization with Neural Networks". In: *Comput. Softw. Big Sci.* 5.1 (2021), p. 14. DOI: `10.1007/s41781-021-00059-x`. arXiv: `2004.02665 [hep-ex]`.

[116] F. James and M. Roos. "Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations". In: *Comput. Phys. Commun.* 10 (1975), pp. 343–367. DOI: `10.1016/0010-4655(75)90039-9`.

[117] Wouter Verkerke and David P. Kirkby. "The RooFit toolkit for data modeling". In: *eConf* C0303241 (2003). Ed. by L. Lyons and Muge Karagoz, MOLT007. arXiv: `physics/0306116`.

[118] Gregory Schott. "RooStats for Searches". In: *PHYSTAT 2011*. Geneva: CERN, 2011, pp. 199–208. DOI: `10.5170/CERN-2011-006.199`. arXiv: `1203.1547 [physics.data-an]`.

[119] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *Eur. Phys. J. C* 71 (2011). [Erratum: Eur.Phys.J.C 73, 2501 (2013)], p. 1554. DOI: `10.1140/epjc/s10052-011-1554-0`. arXiv: `1007.1727 [physics.data-an]`.

[120] Alexander L. Read. "Presentation of search results: The CL_s technique". In: *J. Phys. G* 28 (2002). Ed. by M. R. Whalley and L. Lyons, pp. 2693–2704. DOI: `10.1088/0954-3899/28/10/313`.

[121] Giuseppe Degrassi et al. "Higgs mass and vacuum stability in the Standard Model at NNLO". In: *JHEP* 08 (2012), p. 098. DOI: `10.1007/JHEP08(2012)098`. arXiv: `1205.6497 [hep-ph]`.

[122] Dario Buttazzo et al. "Investigating the near-criticality of the Higgs boson". In: *JHEP* 12 (2013), p. 089. DOI: `10.1007/JHEP12(2013)089`. arXiv: `1307.3536 [hep-ph]`.

[123] Anders Andreassen, William Frost, and Matthew D. Schwartz. “Scale Invariant Instantons and the Complete Lifetime of the Standard Model”. In: *Phys. Rev. D* 97.5 (2018), p. 056006. DOI: `10.1103/PhysRevD.97.056006`. arXiv: `1707.08124 [hep-ph]`.

[124] Manuel Reichert et al. “Probing baryogenesis through the Higgs boson self-coupling”. In: *Phys. Rev. D* 97.7 (2018), p. 075008. DOI: `10.1103/PhysRevD.97.075008`. arXiv: `1711.00019 [hep-ph]`.

[125] J. Alison et al. “Higgs boson potential at colliders: Status and perspectives”. In: *Rev. Phys.* 5 (2020). Ed. by Biagio Di Micco et al., p. 100045. DOI: `10.1016/j.revip.2020.100045`. arXiv: `1910.00012 [hep-ph]`.

[126] R. Frederix et al. “Higgs pair production at the LHC with NLO and parton-shower effects”. In: *Phys. Lett. B* 732 (2014), pp. 142–149. DOI: `10.1016/j.physletb.2014.03.026`. arXiv: `1401.7340 [hep-ph]`.

[127] *Validation of signal Monte Carlo event generation in searches for Higgs boson pairs with the ATLAS detector.* Tech. rep. Geneva: CERN, 2019. URL: `https://cds.cern.ch/record/2665057`.

[128] Georges Aad et al. “Studies of new Higgs boson interactions through nonresonant HH production in the $b\bar{b}\gamma\gamma$ final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *JHEP* 01 (2024), p. 066. DOI: `10.1007/JHEP01(2024)066`. arXiv: `2310.12301 [hep-ex]`.

[129] Georges Aad et al. “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”. In: *Phys. Rev. D* 106.5 (2022), p. 052001. DOI: `10.1103/PhysRevD.106.052001`. arXiv: `2112.11876 [hep-ex]`.

[130] M. Oreglia. “A Study of the Reactions $\psi' \to \gamma\gamma\psi$”. Other thesis. Dec. 1980.

[131] *Recommendations for the Modeling of Smooth Backgrounds.* Tech. rep. Geneva: CERN, 2020. URL: `https://cds.cern.ch/record/2743717`.

[132] Brian Petersen et al. *Expected performance of the ATLAS detector under different High-Luminosity LHC conditions*. Tech. rep. Geneva: CERN, 2021. URL: `https://cds.cern.ch/record/2765851`.

[133] *Tagging and suppression of pile-up jets in the forward region using timing information with the ATLAS detector at $\sqrt{s}$ = 14 TeV at HL-LHC*. Tech. rep. Geneva: CERN, 2022. URL: `https://cds.cern.ch/record/2834329`.

[134] Collaboration CMS. *A MIP Timing Detector for the CMS Phase-2 Upgrade*. Tech. rep. Geneva: CERN, 2019. URL: `https://cds.cern.ch/record/2667167`.

[135] Doug Berry et al. "4-Dimensional Trackers". In: *Snowmass 2021*. Mar. 2022. arXiv: `2203.13900 [physics.ins-det]`.

[136] *Expected Size and Readout Rate of the Data of the ATLAS ITk Pixel Detector*. Tech. rep. Geneva: CERN, 2022. URL: `https://cds.cern.ch/record/2800852`.

[137] *Technical Design Report for the ATLAS Inner Tracker Strip Detector*. Tech. rep. Geneva: CERN, 2017. URL: `https://cds.cern.ch/record/2257755`.

[138] *Technical Design Report for the ATLAS Inner Tracker Pixel Detector*. Tech. rep. Geneva: CERN, 2017. DOI: `10.17181/CERN.FOZZ.ZP3Q`. URL: `https://cds.cern.ch/record/2285585`.

[139] *Technical Design Report: A High-Granularity Timing Detector for the ATLAS Phase-II Upgrade*. Tech. rep. Geneva: CERN, 2020. URL: `https://cds.cern.ch/record/2719855`.

[140] *Expected tracking and related performance with the updated ATLAS Inner Tracker layout at the High-Luminosity LHC*. Tech. rep. Geneva: CERN, 2021. URL: `https://cds.cern.ch/record/2776651`.

[141] D. Braga et al. "Electronics for Fast Timing". In: *Snowmass 2021*. Mar. 2022. arXiv: `2204.00149 [physics.ins-det]`.

[142] David J. Lange. "The EvtGen particle decay simulation package". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 462.1 (2001). BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines, pp. 152–155. ISSN: 0168-9002. DOI: `https://doi.org/10.1016/S0168-`

9002(01)00089-4. URL: `https://www.sciencedirect.com/science/article/pii/S0168900201000894`.

[143] V. Kostyukhin et al. “Improving primary-vertex reconstruction with a minimum-cost lifted multicut graph partitioning algorithm”. In: *Journal of Instrumentation* 18.07 (2023), P07013. DOI: `10.1088/1748-0221/18/07/P07013`. URL: `https://dx.doi.org/10.1088/1748-0221/18/07/P07013`.

[144] Martin Ester et al. “A density-based algorithm for discovering clusters in large spatial databases with noise.” In: *kdd*. Vol. 96. 34. 1996, pp. 226–231.

[145] *Neural Network Jet Flavour Tagging with the Upgraded ATLAS Inner Tracker Detector at the High-Luminosity LHC*. Tech. rep. Geneva: CERN, 2022. URL: `https://cds.cern.ch/record/2839913`.

[146] ATLAS Collaboration. “Search for Higgs boson pair production in the two bottom quarks plus two photons final state in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector”. In: *Phys. Rev. D* 106 (2021), p. 052001. DOI: `10.1103/PhysRevD.106.052001`. arXiv: `2112.11876 [hep-ex]`.

[147] 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”. In: (2022). arXiv: `2209.10910 [hep-ex]`.

[148] ATLAS Collaboration. “Search for nonresonant pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *CERN-EP-2022-235* (2023). arXiv: `2301.03212 [hep-ex]`.

[149] CMS Collaboration. “Search for nonresonant Higgs boson pair production in final states with two bottom quarks and two photons in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$”. In: *JHEP* 03 (2021), p. 257. DOI: `10.1007/JHEP03(2021)257`. arXiv: `2011.12373 [hep-ex]`.

[150] CMS Collaboration. “Search for nonresonant Higgs boson pair production in final state with two bottom quarks and two tau leptons in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$”. In: (2022). arXiv: `2206.09401 [hep-ex]`.

[151] CMS Collaboration. “Search for Higgs Boson Pair Production in the Four b Quark Final State in Proton–Proton Collisions at $\sqrt{s} = 13\,\text{TeV}$”. In: *Phys. Rev. Lett.* 129 (2022), p. 081802. DOI: `10.1103/PhysRevLett.129.081802`. arXiv: `2202.09617 [hep-ex]`.

[152] CMS Collaboration. “Search for nonresonant pair production of highly energetic Higgs bosons decaying to bottom quarks”. In: (2022). arXiv: `2205.06667 [hep-ex]`.

[153] ATLAS Collaboration. “Constraining the Higgs boson self-coupling from single- and double-Higgs production with the ATLAS detector using pp collisions at $\sqrt{s} = 13\,\text{TeV}$”. In: (2022). arXiv: `2211.01216 [hep-ex]`.

[154] CMS Collaboration. “A portrait of the Higgs boson by the CMS experiment ten years after the discovery”. In: *Nature* 607 (2022), pp. 60–68. DOI: `10.1038/s41586-022-04892-x`. arXiv: `2207.00043 [hep-ex]`.

[155] ATLAS Collaboration. *Projected sensitivity of Higgs boson pair production in the $bb\tau\tau$ final state using proton–proton collisions at HL-LHC with the ATLAS detector*. ATL-PHYS-PUB-2021-044. 2021. URL: `https://cds.cern.ch/record/2798448`.

[156] ATLAS Collaboration. *Measurement prospects of Higgs boson pair production in the $b\bar{b}\gamma\gamma$ final state with the ATLAS experiment at the HL-LHC*. ATL-PHYS-PUB-2022-001. 2022. URL: `https://cds.cern.ch/record/2799146`.

[157] ATLAS Collaboration. *Projected sensitivity of Higgs boson pair production combining the $b\bar{b}\gamma\gamma$ and $b\bar{b}\tau^+\tau^-$ final states with the ATLAS detector at the HL-LHC*. ATL-PHYS-PUB-2022-005. 2022. URL: `https://cds.cern.ch/record/2802127`.

[158] ATLAS Collaboration. *HL-LHC prospects for the measurement of Higgs boson pair production in the $b\bar{b}b\bar{b}$ final state and combination with the $b\bar{b}\gamma\gamma$ and $b\bar{b}\tau^+\tau^-$ final states at the ATLAS experiment*. ATL-PHYS-PUB-2022-053. 2022. URL: `https://cds.cern.ch/record/2841244`.

[159] Xiaocong Ai et al. “A Common Tracking Software Project”. In: *Comput. Softw. Big Sci.* 6.1 (2022), p. 8. DOI: `10.1007/s41781-021-00078-8`. arXiv: `2106.13593 [physics.ins-det]`.

[160] Corentin Allaire et al. *OpenDataDetector*. Version v2. Apr. 2022. DOI: `10.5281/zenodo.6445359`. URL: `https://doi.org/10.5281/zenodo.6445359`.

[161] François Chollet. *Deep Learning with Python*. Manning, Nov. 2017. ISBN: 9781617294433.

[162] Michael A Nielsen. *Neural networks and deep learning*. Vol. 2018. Determination press San Francisco, CA, USA: 2015.

[163] D. Guest, K. Cranmer, and D. Whiteson. "Deep Learning and its Application to LHC Physics". In: *Ann. Rev. Nucl. Part. Sci.* 68 (2018), pp. 161–181. DOI: `10.1146/annurev-nucl-101917-021019`. arXiv: `1806.11484 [hep-ex]`.

[164] HEP ML Community. *A Living Review of Machine Learning for Particle Physics*. `https://iml-wg.github.io/HEPML-LivingReview/`. URL: `https://iml-wg.github.io/HEPML-LivingReview/`.

[165] A. Mertens. "New features in Delphes 3". In: *J. Phys. Conf. Ser.* 608.1 (2015). Ed. by L. Fiala, M. Lokajicek, and N. Tumova, p. 012045. DOI: `10.1088/1742-6596/608/1/012045`.

[166] M. Paganini, L. de Oliveira, and B. Nachman. "CaloGAN: Simulating 3D high energy particle showers in multilayer electromagnetic calorimeters with generative adversarial networks". In: *Phys. Rev. D* 97 (1 2018), p. 014021. DOI: `10.1103/PhysRevD.97.014021`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.97.014021`.

[167] S. R. Qasim et al. "Learning representations of irregular particle-detector geometry with distance-weighted graph networks". In: *Eur. Phys. J. C.* 79.7 (2019). ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-019-7113-9`. URL: `http://dx.doi.org/10.1140/epjc/s10052-019-7113-9`.

[168] J. Allison et al. "Recent developments in Geant4". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 835 (2016), pp. 186–225. ISSN: 0168-9002. DOI: `https://doi.org/10.1016/j.nima.2016.06.125`. URL: `http://www.sciencedirect.com/science/article/pii/S0168900216306957`.

[169] Anton Charkin-Gorbulin et al. "Configurable calorimeter simulation for AI applications". In: *Mach. Learn. Sci. Tech.* 4.3 (2023), p. 035042. DOI: `10.1088/2632-2153/acf186`. arXiv: `2303.02101 [hep-ex]`.

[170] Sanmay Ganguly et al. *cocoa-hep/cocoa-hep: 0.1.1*. Version 0.1.1. `https://doi.org/10.5281/zenodo.7700475`. Mar. 2023. DOI: `10.5281/zenodo.7700475`. URL: `https://doi.org/10.5281/zenodo.7700475`.

[171] A. Corentin et al. *Open Data Detector*. 2022. DOI: `https://doi.org/10.5281/zenodo.4674401`.

[172] F. A. Di Bello et al. "Conditional Generative Modelling of Reconstructed Particles at Collider Experiments". In: (Nov. 2022). arXiv:2211.06406. arXiv: `2211.06406 [hep-ex]`.

[173] E. Moyse et al. "The Phoenix event display framework". In: *EPJ Web Conf.* 251 (2021), p. 01007. DOI: `10.1051/epjconf/202125101007`. URL: `https://doi.org/10.1051/epjconf/202125101007`.

[174] ATLAS Collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST* 3 (2008), S08003. DOI: `10.1088/1748-0221/3/08/S08003`.

[175] The ATLAS collaboration. "Measurement of the energy response of the ATLAS calorimeter to charged pions from $W^{\pm} \to \tau^{\pm}(\to \pi^{\pm}\nu_{\tau})\nu_{\tau}$ events in Run 2 data". In: *Eur. Phys. J. C* 82.3 (), p. 223. DOI: `10.1140/epjc/s10052-022-10117-2`. arXiv: `2108.09043 [hep-ex]`.

[176] C. Bierlich et al. "A comprehensive guide to the physics and usage of PYTHIA 8.3". In: *SciPost Phys. Codebases* (2022), p. 8. DOI: `10.21468/SciPostPhysCodeb.8`. URL: `https://scipost.org/10.21468/SciPostPhysCodeb.8`.

[177] The ATLAS collaboration. "Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data". In: *J. Instrum.* 14.12 (Dec. 2019), P12006–P12006.

[178] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "FastJet user manual". In: *Eur. Phys. J. C.* 72.3 (2012). DOI: `10.1140/epjc/s10052-012-1896-2`. URL: `https://doi.org/10.1140%2Fepjc%2Fs10052-012-1896-2`.

[179] J. Alwall et al. “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”. In: *JHEP* 07 (2014), p. 079. DOI: `10.1007/JHEP07(2014)079`. arXiv: `1405.0301 [hep-ph]`.

[180] ATLAS collaboration. “Jet reconstruction and performance using particle flow with the ATLAS Detector”. In: *Eur. Phys. J. C* 77.7 (2017), pp. 1–47. DOI: `10.1140/epjc/s10052-017-5031-2`. arXiv: `1703.10485 [hep-ex]`.

[181] A. M. Sirunyan, CMS collaboration, et al. “Particle-flow reconstruction and global event description with the CMS detector”. In: *JINST* 12.10 (2017), P10003.

[182] J. Shlomi, P. Battaglia, and J. R. Vlimant. “Graph neural networks in particle physics”. In: *Mach. Learn. Sci. Technol* 2.2 (2020), p. 021001. ISSN: 2632-2153. DOI: `10.1088/2632-2153/abbf9a`. URL: `http://dx.doi.org/10.1088/2632-2153/abbf9a`.

[183] J. Kieseler. “Object condensation: one-stage grid-free multi-object reconstruction in physics detectors, graph, and image data”. In: *Eur. Phys. J. C* 80.9 (2020), pp. 1–12. ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-020-08461-2`. URL: `http://dx.doi.org/10.1140/epjc/s10052-020-08461-2`.

[184] S. R. Qasim et al. “Multi-particle reconstruction in the High Granularity Calorimeter using object condensation and graph neural networks”. In: *EPJ Web Conf.* 251 (2021), p. 03072.

[185] S. R. Qasim et al. “End-to-end multi-particle reconstruction in high occupancy imaging calorimeters with graph neural networks”. In: *Eur. Phys. J. C* 82.8 (2022), p. 753. DOI: `https://doi.org/10.1140/epjc/s10052-022-10665-7`.

[186] J. Pata et al. “MLPF: Efficient machine-learned particle-flow reconstruction using graph neural networks”. In: *Eur. Phys. J. C* 81.5 (2021), pp. 1–14.

[187] J. Pata et al. “Machine Learning for Particle Flow Reconstruction at CMS”. In: *J. Phys. Conf. Ser.* 2438.1 (2022), p. 012100. DOI: `10.1088/1742-6596/2438/1/012100`. URL: `https://dx.doi.org/10.1088/1742-6596/2438/1/012100`.

[188] F. Mokhtar et al. "Progress towards an improved particle flow algorithm at CMS with machine learning". In: *ACAT 2022*. Mar. 2023. arXiv: `2303.17657`.

[189] S. Agostinelli et al. "GEANT4: A simulation toolkit". In: *Nucl. Instrum. Meth.* 506.3 (2003), pp. 250–303. DOI: `10.1016/S0168-9002(03)01368-8`.

[190] J. de Favereau et al. "DELPHES 3: a modular framework for fast simulation of a generic collider experiment". In: *JHEP* 2014.2 (2014), pp. 1–26. DOI: `10.1007/jhep02(2014)057`. URL: `http://dx.doi.org/10.1007/JHEP02(2014)057`.

[191] ATLAS collaboration. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". In: *Eur. Phys. J. C* 77.7 (2017), pp. 1–73. DOI: `10.1140/epjc/s10052-017-5004-5`. arXiv: `1603.02934 [hep-ex]`.

[192] S. R. Qasim et al. "Learning representations of irregular particle-detector geometry with distance-weighted graph networks". In: *Eur. Phys. J. C* 79.7 (2019), p. 608. DOI: `10.1140/epjc/s10052-019-7113-9`. arXiv: `1902.07987 [physics.data-an]`.

[193] A. R. Kosiorek, H. Kim, and D. J. Rezende. "Conditional Set Generation with Transformers". In: *preprint: arXiv:2006.16841* (2020). arXiv: `2006.16841 [cs.CV]`.

[194] F. Locatello et al. "Object-centric learning with slot attention". In: *Adv. Neural Inf. Process. Sys.* 33 (2020), pp. 11525–11538.

[195] H. W. Kuhn. "The Hungarian method for the assignment problem". In: *Naval research logistics quarterly* 2.1-2 (1955), pp. 83–97.

[196] Francesco Armando Di Bello et al. "Towards a Computer Vision Particle Flow". In: (Mar. 2020). arXiv: `2003.08863 [hep-ex]`.

[197] D. W. Zhang, G. J. Burghouts, and C. G. M. Snoek. "Recurrently Predicting Hypergraphs". In: *preprint: arXiv:2106.13919* (2021).

[198] M. Zaheer et al. "Deep Sets". In: *Adv. Neural Inf. Process. Sys.* 30 (2017).