Searches for CP-Violation in $H \to \tau\tau$ Decay with the ATLAS Detector

Jianhuan Xiang

A Thesis Submitted to
The Hong Kong University of Science and Technology
in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy
in Physics

December, 2022, Hong Kong

Authorization

I hereby declare that I am the sole author of the thesis.

I authorize the Hong Kong University of Science and Technology to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize the Hong Kong University of Science and Technology to reproduce the thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

Jianhuan Xiang

16, December, 2022

Searches for CP-Violation in $H \to \tau\tau$ Decay with the ATLAS Detector

by

Jianhuan Xiang

This is to certify that I have examined the above PhD thesis
and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by
the thesis examination committee have been made.

Supervisor: Prof. Dr. Kirill Prokofiev

Department Head: Prof. Dr. Jiannong Wang

16, December, 2022

摘要

本论文介绍了对希格斯玻色子与 τ-轻子之间的 Yukawa 耦合的 CP 特性的测量。该测量使用了在 2015-2018 年间从大型强子对撞机中的 ATLAS 探测器中采集的总计 139 fb^{-1} 的质子对撞数据。这些质子对撞的质心能量为 $\sqrt{s} = 13$ TeV。该研究用以测量 CP 特性的可观测量由 τ-轻子的衰变产物中可探测的部分构建。希格斯玻色子与 τ-轻子之间违反了 CP 对称性的相互作用由一个 CP 混合角 ϕ_τ 表示。该混合角 ϕ_τ 在 68% 置信度下的期望值为 $0^\circ \pm 28^\circ$，在 95.5% 置信度下的期望值为 $0^{\circ\,+75^\circ}_{\;-70^\circ}$。观测到的混合角 ϕ_τ 在 68% 置信度下的期望值为 $9^\circ \pm 16^\circ$，在 95.5% 置信度下的期望值为 $9^\circ \pm 34^\circ$。此结果可以在 3.4 个标准差的区间内排除纯 CP-odd 的假设。本实验的观测符合标准模型的预期。

此外，本文还介绍了有关区分探测器中由堆积信号造成的粒子射流信号的研究。Run-2 时期在 ATLAS 中使用的区分探测器中由堆积信号造成的粒子射流信号的算法基于 K-近邻算法。本研究中，基于此算法的模型使用在 2015-2018 年间在 ATLAS 探测器采集的数据重新训练。此外，一个新的基于神经网络的模型被开发用以区分由堆积信号造成的粒子射流信号。新模型的背景信号抑制率在所有给定的工作环境中都得到了提升。

本文还介绍了一个用神经网络训练的对 CP 特性敏感的可观测量。在一个简化了的数据模型中，该可观测量相比此前设计的可观测量以更高的确信度排除纯 CP-odd 的假设。这项研究会为未来对 CP 特性的精确测量做出贡献。

Abstract

This thesis presents a measurement of the CP-properties of the Yukawa coupling between the Higgs boson and $\tau-$lepton. The measurement uses the proton–proton collision data collected from 2015 to 2018 with the ATLAS detector at the Large Hadron Collider. Totally 139 fb^{-1} proton–proton collision data is collected at a center-of-mass energy of $\sqrt{s} = 13$ TeV. This study investigates the CP-properties with CP-sensitive observables defined by the visible decay products of $\tau-$leptons. CP-violating interactions between the Higgs boson and $\tau-$lepton are described by the CP-mixing angle ϕ_τ. The expected value of ϕ_τ according to the Standard Model is $0^\circ \pm 28^\circ$ at 68% confidence level, and $0^{\circ\,+75^\circ}_{\;\;-70^\circ}$ at 95.5% confidence level obtained from a simulated dataset. The observed value of ϕ_τ is $9^\circ \pm 16^\circ$ at 68% confidence level, and $9^\circ \pm 34^\circ$ at 95.5% confidence level. The pure CP-odd hypothesis is disfavoured at 3.4 standard deviations. The observation is consistent with the Standard Model expectations.

This thesis also presents studies of pile-up jet tagging. The pile-up jet tagging algorithm is based on the K-Nearest Neighbor method. It is trained with the data collected from 2015 to 2018 with the ATLAS detector. A new pile-up jet tagging algorithm based on neural network is developed. It provides improvement to the background rejection at all working points.

In this thesis, a new CP-sensitive observable is developed using neural network approach. The trained observable excludes the pure CP-odd hypothesis at higher confidence level in a simplified sample compared with the previously designed observable. These studies will contribute to future measurements of the CP-properties.

Acknowledgments

I had the pleasure of working with many remarkable people during the past years. At the end of my PhD study, there are many people I would like to thank.

First of all, I would like to thank my supervisor Kirill Prokofiev, for his support throughout my studies. Thanks him for guiding and encouraging me in the research, and the help during the thesis writing. It is my honor to work in his group during my PhD study.

I am grateful to the postdoc researchers Ki Lie, Hok Chuen Cheng, and Antonio De Maria. They conducted the measurements of CP-properties at the ATLAS detector. I learned a lot from them about experimental particle physics and the analysis. And they are always helpful and took the time to assist me in all matters. I would like to particularly thank Ki Lie, for his helps during my travel to CERN and his help in my thesis writing. I am also thankful to Matthew Klein and Louis Portales, who supervised my qualification task at CERN. I learned a lot from them about the software and data structure at CERN.

I would like to thank Elzbieta Richter-Was and her group for their participation in my research of CP-sensitive observables. They provide several valuable suggestions in my study and they kindly provide the Monte Carlo simulated samples used in their previous research. These samples help improve my machine learning model.

I would also like to thank Tianyi Yang, Yuchen Yan, Lingfeng Li and Sijun Xu. I am inspired from the conversation with them. I also learned new knowledge and latest scientific discovery from them. I also receive many reference articles from them during my thesis writing.

I would like to thank Manqi Ruan and Young-Kee Kim for supervising my research at CEPC during my Master's degree. I would also like to thank professor Tao Liu. His particle physics course is the start of my particle physics researches. He also provides many suggestions during my PhD study.

It is difficult to summarize all the people who contribute. I would like to thank all the professors and colleague I met during my study. I would like to thank the HKUST, the University of Chicago and the Copenhagen University for their valuable courses. I would like to thank all the student I talked with in the University. I would like to thank people at the CERN and the CEPC. My research is based the experiments done with the effort from everyone.

Finally, I would like to express my deepest gratitude to my family for their continuous love, help, and support.

Hope everyone stay healthy and achieve their goal in the future.

Contents

Abstract in Chinese **3**

Abstract **5**

Acknowledgements **7**

Contents **9**

Introduction **11**

Chapter 1. Theoretical background **13**

1.1 The Standard Model and the Higgs boson 13

1.2 The Higgs mechanism 14

1.3 The baryon asymmetry of the Universe 16

1.4 The CP-violation 18

1.5 Higgs boson production and decay in the proton-proton collision 19

Chapter 2. Experimental setup **23**

2.1 The Large Hadron Collider 23

2.2 The ATLAS detector 26

2.3 Data reconstruction 36

Chapter 3. Jet-Vertex-Tagger Retraining **43**

3.1 Hard-scattering jet identification 43

3.2 Jet-Vertex-Tagger retraining 47

Chapter 4. CP-violation in Higgs boson decays to τ pairs **57**
4.1 Decays of $\tau-$lepton 57
4.2 Methods to measure the CP-mixing angle 60
4.3 Investigation of 3-prong CP observables 68

Chapter 5. Measurement of CP properties in the $H \to \tau\tau$ coupling **73**
5.1 Experimental data and Monte Carlo simulation 73
5.2 Event selection 75
5.3 Background estimation 79
5.4 Expected event number 80
5.5 Systematic uncertainties 80
5.6 Maximum-likelihood fit 83
5.7 Results 86

Chapter 6. Further Studies **95**
6.1 Investigation of CP-sensitive 3p0n-1p1n observables 95
6.2 Main types of machine learning techniques 96
6.3 Searching for 3p0n-1p1n observable with reinforcement learning 97
6.4 Results 102

Chapter 7. Summary **105**

Appendices **108**
A 3-prong observables 108

References **115**

Introduction

The ATLAS and CMS collaborations at the Large Hardon Collider (LHC) discovered a particle consistent with the Standard Model (SM) Higgs boson in 2012 [1]. Several studies have been performed to investigate the properties of this Higgs boson. Decay of the Higgs boson to a pair of $\tau-$leptons has been observed by the CMS and ATLAS experiments in 2018 and 2019 respectively [2, 3]. This decay allows measurements of the charge conjugation and parity (CP) properties of the Yukawa coupling of the Higgs boson to the $\tau-$lepton. The SM Higgs boson is predicted to be purely CP-even. Studies performed by ATLAS and CMS experiments on the interaction between the Higgs boson and gauge bosons have not shown any deviations from the SM predictions so far. A study on the interaction between Higgs boson and gauge bosons published in 2015 has excluded the pure CP-odd structure at a 99.98% confidence level [4]. Nevertheless, the presence of a small CP-odd admixture has not been yet excluded. If CP-odd contributions to the Higgs boson interactions exist, they may appear at the tree level in the Higgs boson-fermion interactions. CP-odd contributions can only appear at the loop level in the Higgs boson interaction with the SM gauge bosons [5]. Measurements of coupling between the Higgs boson and $\tau-$lepton might provide evidence for CP-mixing in the Higgs sector. Discovery of CP-mixing will indicate the existence of new physics beyond the SM. It may also contribute to the explanation of observed asymmetry between matter and anti-matter in our Universe [6].

In this thesis, a measurement of the CP property in the interaction between the Higgs boson and $\tau-$lepton is presented. The analysis is based on the observables constructed from the measured parameters of the $\tau-$lepton decay products. The proton-proton collision data recorded by the ATLAS detector from 2015 to 2018 are used. The proton beams collided at a center-of-mass energy of $\sqrt{s} = 13$ TeV.

This thesis also presents a development of the pile-up jet tagging algorithm based on neural network. The pile-up jet tagging algorithm is part of the reconstruction, which takes detector signals as input and reconstructs the kinematic information of the particles. This new algorithm improves background rejection and processing

speed. It has been adopted as the new default jet-vertex-tagging algorithm in the ATLAS Run-3 data analysis which has started in 2022.

This thesis also includes the investigation of a new CP-sensitive observable for the $H \to \tau\tau$ decay. It demonstrates the potential of constructing new observables with machine learning techniques. The new observables may improve the sensitivity of measurements in future analysis. The presented methodology can also be applied to other analyses in particle physics.

Chapter 1

Theoretical background

1.1 The Standard Model and the Higgs boson

The Standard Model (SM) is a quantum field theory that describes the elementary particles and the fundamental interactions. It predicted the existence of several elementary particles: top, bottom and charm quarks, W and Z bosons, gluons, Higgs boson and others. All these particles have been discovered in experiments [7, 8, 9, 10, 11, 12]. The Higgs boson was discovered in 2012, by the experiments conducted at the Large Hadron Collider [12]. Figure 1.1 shows the elementary particles of the Standard Model and their quantum numbers. The quarks u, d, c, s, t, b and the leptons e, ν_e, μ, ν_μ, τ, ν_τ are fermions. The bosons γ, $W^{\pm}/Z$ and g carry electromagnetic force, the weak force and the strong force, respectively.

The SM predicts the Higgs boson to be a neutral particle with spin 0, and it has no strong force interaction. The existence of the Higgs boson is predicted by the Standard Model before its discovery, as a consequence of the Higgs mechanism [13]. It is predicted to couple with itself, fermions, W and Z bosons. The existence of the Higgs boson solves the problem that the fermions and W and Z bosons are not allowed to have mass term. The description of how these particles acquire mass from the interaction with the Higgs boson is in Section 1.2.

The Standard Model is a gauge-invariant quantum field theory (QFT) with gauge symmetry $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$ [14]. Quantum chromodynamics (QCD) is a QFT with symmetry group $SU(3)_c$, that describes the interaction between quarks and gluons. The $SU(2)_L \otimes U(1)_Y$ symmetry describes the electromagnetic and weak interactions, also named the Electroweak interaction (EW).

The interactions in the Standard Model are described by the SM Lagrangian. Each term of the Lagrangian is invariant under the corresponding gauge transformation. The terms of electroweak interaction satisfy the $SU(2)_L \otimes U(1)_Y$ gauge

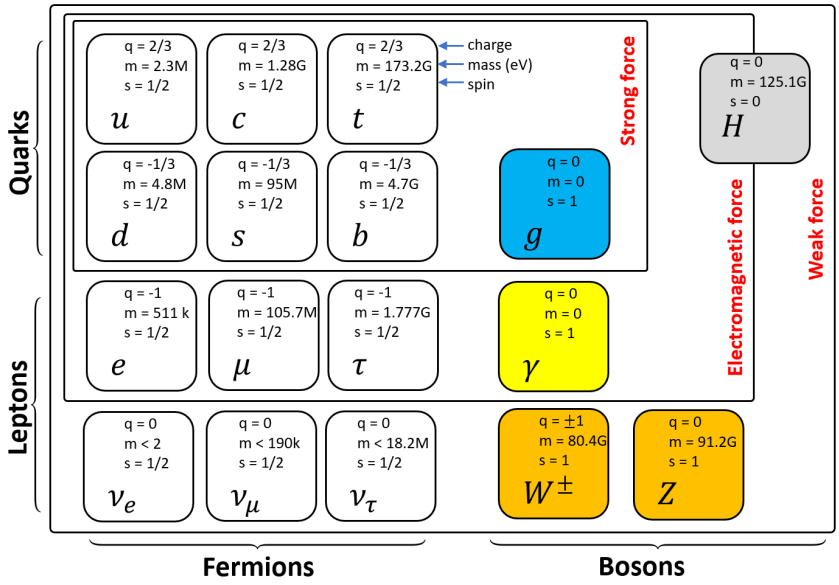


Figure 1.1: Elementary particles in the Standard Model. The mass, charge and spin are labeled in the corner.

symmetry, where Y represents the weak hypercharge of the $U(1)_Y$ symmetry. The fields associated with $SU(2)$ symmetry $W_i^\mu (i = 1, 2, 3)$ only couple to left-handed spinor fields, leading to a subscript L in the $SU(2)$ symmetry. The Lagrangian of electroweak interaction does not contain the mass term of gauge bosons $\propto m_V^2 V^\mu V_\mu$. Such a mass term is not invariant under the $SU(2)_L \otimes U(1)_Y$ transformation. The mass term of fermions $\propto m_f \bar{\psi}\psi$ is also not allowed, since the left-handed particles transform differently compared to the right-handed particles [15]. The gauge bosons $W^\pm$ bosons, Z boson and fermions acquire mass from the Higgs mechanism [13].

1.2 The Higgs mechanism

The Higgs mechanism was first introduced in the 1960s [13]. It explained the masses of gauge bosons, and successfully predicted the existence of the Higgs boson. In 2013, the Nobel Prize in physics is awarded to François Englert and Peter W. Higgs for the theoretical discovery of the Higgs mechanism, confirmed by the discovery of the Higgs boson performed at the Large Hadron Collider [12].

Consider an isospin doublet of complex scalar fields,

$$H = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}, \tag{1.1}$$

a Higgs potential $V(H)$ is written as [15]:

$$V(H) = |H| = \mu^2 H^\dagger H + \lambda (H^\dagger H)^2. \tag{1.2}$$

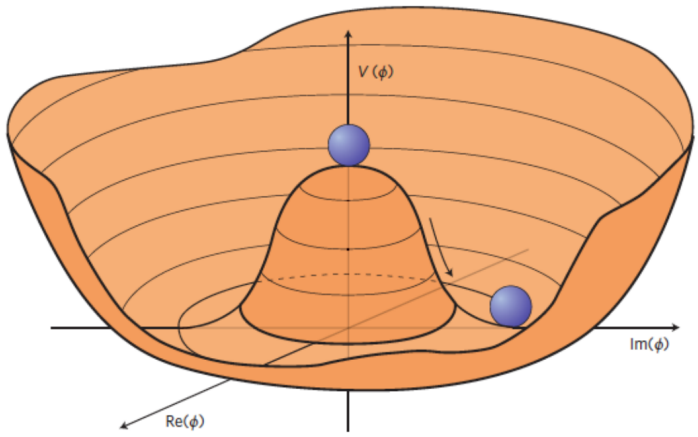


Figure 1.2: Higgs potential with degenerate ground states when $\mu^2 < 0$ [16].

The parameters μ^2 and λ describe the shape of the potential. The Higgs potential has degenerate ground states when $\lambda > 0$ and $\mu^2 < 0$. Figure 1.2 shows the graphical representation of the Higgs potential in this case. The potential $V(H)$ gets a minimum at $|H| = \sqrt{H^\dagger H}$. The vacuum expectation value (VEV) of the doublet H is:

$$\langle H \rangle = |H| = \sqrt{\frac{-\mu^2}{2\lambda}} = \frac{\upsilon}{\sqrt{2}}, \tag{1.3}$$

where υ is the vacuum expectation value of the Higgs potential.

By choosing a non-zero value for the Higgs potential in the vacuum state, the electroweak symmetry is broken spontaneously. A concise way is to choose the $\phi^+ = 0$ and make ϕ^0 real. This choice is called the unitary gauge. Then the vacuum expectation value is represented by:

$$\langle 0|H|0\rangle = \begin{pmatrix} 0 \\ \frac{\upsilon}{\sqrt{2}} \end{pmatrix}. \tag{1.4}$$

With the vacuum expectation value $\upsilon \approx 246$ GeV [17], the scalar doublet H can be parameterized as:

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \upsilon + h \end{pmatrix}. \tag{1.5}$$

As a result, the following Lagrangian terms represent the mass of the gauge bosons [15]:

$$\mathcal{L}_{mass} = \frac{1}{2}\left(\frac{\upsilon g}{2}\right)^2 (W_\mu^- W^{-\mu} + W_\mu^+ W^{+\mu}) + \frac{1}{2}\left(\frac{\upsilon\sqrt{g^2+g'^2}}{2}\right)^2 Z_\mu Z^\mu, \qquad (1.6)$$

where g and g' are coupling constants related to electroweak interaction.

The $W\pm$ bosons gain mass of:

$$m_{W^\pm} = \frac{\upsilon g}{2},$$

and the Z boson gains a mass of:

$$m_Z = \frac{\upsilon\sqrt{g^2+g'^2}}{2}.$$

The latest measurement of the mass of the $W^\pm$ bosons and the Z boson are $m_{W^\pm} = 80.397 \pm 0.012$ GeV, $m_Z = 91.1876 \pm 0.0021$ GeV [17], while the SM predicts $m_{W^\pm} = 80.357 \pm 0.004_{input} \pm 0.004_{theory}$ GeV [18] and $m_Z = 91.1875 \pm 0.0021$ GeV [19]. The fermion masses are granted by the interaction between the Higgs field with the fermions fields, which is called the Yukawa interaction. The fermion mass can be expressed by:

$$m_f = \frac{\lambda_{yf}\upsilon}{\sqrt{2}} \quad ,$$

where the λ_{yf} represents the Yukawa coupling for the fermion [15].

The Higgs boson also gains mass through the Higgs mechanism [15]. The mass of the Higgs boson is expressed by:

$$m_{Higgs} = \sqrt{2\lambda\upsilon^2}.$$

Photon and gluons do not interact with the Higgs field at tree level. These particles do not acquire mass through the Higgs mechanism.

1.3 The baryon asymmetry of the Universe

Although the Standard Model explains the elementary particles and interactions well, there are still observed phenomena that it cannot explain. One of them is baryon asymmetry.

Baryons are heavy subatomic particles that are made up of three quarks. Baryon number is a quantum number assigned to quarks and hadrons. A quark has a baryon number of $\frac{1}{3}$, and an anti-quark has a baryon number of $-\frac{1}{3}$. A proton or neutron

has a baryon number of 1, corresponding to the sum of the baryon numbers of three quarks. Baryon asymmetry η is defined as:

$$\eta = \frac{N_B - N_{\bar{B}}}{N_\gamma}. \tag{1.7}$$

It describes the difference between the number of baryons N_B and antibaryons $N_{\bar{B}}$ in the observed Universe. The N_γ is the number of photons in the cosmic background radiation [20].

It is a natural assumption that the early Universe was neutral, and the same amount of matter and anti-matter was produced in the Big Bang. This assumption is challenged by the observation that much less anti-matter than matter is presently found in Nature. The observed Universe contains negligible anti-matter compared with matter. The Baryon asymmetry of the Universe was measured with the cosmic microwave background (CMB). The Wilkinson Microwave Anisotropy Probe (WMAP) team measured the CMB radiation and provide an estimated $\eta \sim \mathcal{O}(10^{-10})$ [21, 22]. Even though it is a small number, the Standard model predicts a much lower value $\eta \sim \mathcal{O}(10^{-20})$ [23].

One of the possible ways to generate the matter-antimatter asymmetry is through baryogenesis [24]. Baryogenesis is a hypothesis that describes the physical process to produce matter-antimatter asymmetry in the early universe. In 1967s Andrei Sakharov proposed three necessary conditions for baryogenesis to occur [25].

1. Process that violates the baryonic number exists.

2. Both the charge conjugation (C) symmetry and the combination of charge conjugation and parity symmetries (CP) are violated in the baryogenesis process.

3. The process violating the baryon asymmetry must happen under conditions outside the thermal equilibrium.

Assuming the Universe is initially baryon symmetric, the condition 1 is required to reach the observed Universe with the baryon asymmetry. Otherwise, the baryon number is conserved.

The condition 2 requires both C and CP symmetry to be violated. The charge conjugation (C), parity (P), and time-reversal (T) symmetry are three fundamental symmetries. The charge symmetry transforms a particle into its corresponding anti-particle. Parity is the inversion of all spatial coordinates. Time reversal reverses the direction of the time axis. When C is conserved, the same number of particle and anti-particle is created or annihilated. When CP symmetry is not violated, the same number of left-handed particles and right-handed anti-particles is created or annihilated.

Within thermal equilibrium, a process that changes the baryon number has the same rate as its inverse process that changes it back. The condition 3 is required by baryogenesis, otherwise the overall rate of baryon number changes is canceled out [26].

This thesis studies CP-violation in the Higgs fermionic decay, using data obtained with the Large Hadron Collider.

1.4 The CP-violation

CP-transformation is an operation for studing the invariance under the combination of charge conjugation (C) and parity (P) transformation. A process with CP-symmetry conserves the laws of physics when the particles are replaced by their anti-particles and the spatial coordinates are inverted. CP-violation happens when the laws of physics change under the CP-transformation.

Parity symmetry breaking was observed in the experiment in the early 1950s [27]. One of the experiments demonstrating parity violation is the $\tau - \theta$ puzzle, where the same mass τ, θ mesons decay into three and two pions, respectively. The τ, θ mesons are the former names of the neutral Kaon (K) meson. In 1956, T. D. Lee and C. Yang proposed that the $\tau - \theta$ puzzle could be explained by parity violation in the weak interactions [27]. After that, an assumption of higher-level CP-symmetry was made.

The evidence for CP-violation was discovered in 1964, in the neutral K-meson decay [28]. Further experiments showed more evidence in the B meson decays and Λ_b baryon decays [29, 30]. In 2012, The LHCb Collaboration reported the discovery of CP-violation in D^0 meson decay [31].

The mechanism explaining the CP-violation discovered in the K, B and D mesons decay was implemented in the Standard Model. In 1973, Makoto Kobayashi and Toshihide Maskawa introduce the Cabibbo-Kobayashi-Maskawa (CKM) matrix that describes the transformation between the mass eigenstate of quarks q and the weak eigenstates q' [32]. The existence of complex terms of CKM matrix leads to CP-violation, which explains all the found sources of CP-violation [33].

In the Standard Model, the Higgs boson is purely CP-even and no CP-violation is expected. An observation of a mixed CP-even and CP-odd state would indicate the existence of a new source of CP-violation in the Higgs sector. Such sources are possible with potential new physics beyond the Standard Model. Experiments on the bosonic decay of the Higgs boson have excluded the pure pseudoscalar state of the Higgs boson [34, 35]. The CP-even coupling between the Higgs boson and

gauge bosons is in the tree-level, which is the lowest-order interaction process. The potential CP-odd coupling of the Higgs boson to gauge bosons may only occur at higher-order levels, leading to the measurement suppressed by the mass scale of new physics [36].

In contrast, both CP-even and CP-odd coupling between the Higgs boson and fermions can be present at the tree-level [5]. The fermionic couplings can be studied in both processes of the Higgs boson production and the Higgs boson decay. The measurements on the CP-violation in the Higgs production process have been performed by the ATLAS and CMS Collaborations at the LHC, excluding the pure CP-odd Yukawa coupling at 3.9σ [37] and 3.2σ [38], respectively.

1.5 Higgs boson production and decay in the proton-proton collision

The Large Hadron Collider (LHC) is the largest particle accelerator in the world. It is designed to produce proton-proton collision at a maximum center-of-mass energy of 14 TeV. During the first operational run in the LHC, which is usually called the Run-1, the accelerated proton beams were collided, with a center-of-mass energy around 7 TeV in 2011 and 8 TeV in 2012. The second operational run (Run-2) was started in 2015 with upgraded detectors. During the Run-2 experiments of LHC in 2015-2018, the proton beams collided at a center-of-mass around 13 TeV [39].

In 2012, the ATLAS [12] and CMS [40] collaborations announced the discovery of a new particle compatible with the predicted Standard Model Higgs boson. The analysis by ATLAS experiment was performed using the 4.8 fb^{-1} data collected at $\sqrt{s} = 7$ TeV and 5.8 fb^{-1} data collected at $\sqrt{s} = 8$ TeV. This analysis combined the measurements in the $H \to ZZ^*$, $H \to \gamma\gamma$, $H \to WW^*$, $H \to b\bar{b}$ and $H \to \tau^+\tau^-$ decay channels for $\sqrt{s} = 7$ TeV data, and in the $H \to ZZ^*$, $H \to \gamma\gamma$ and $H \to WW^*$ channels for $\sqrt{s} = 8$ TeV data.

The latest measured value of the mass of the Higgs boson published by Particle Data Group (PDG) is $m_H = 125.10 \pm 0.14$ GeV [17], based on the experiments at the LHC data until 2022.

The expected cross sections of the Higgs boson production depend on the center-of-mass energy of protons collision ($\sqrt{s}$). Figure 1.3 shows the Standard Model Higgs boson production cross sections as a function of the center-of-mass energy.

This thesis studies the CP-violation in the Higgs fermionic decay in the Run-2 experimental setup. At $\sqrt{s} = 13$ TeV and $m_H = 125.10$ GeV, we expect the dominant production process of the Higgs to be gluon fusion (ggF). Other channels

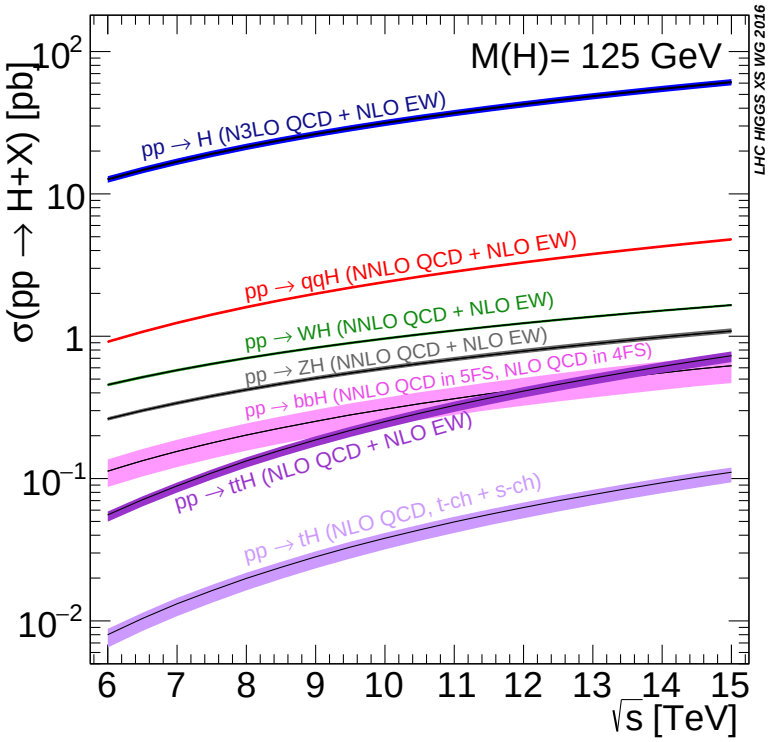


Figure 1.3: Theoretical predicted SM Higgs boson production cross sections in $p-p$ collisions as functions of centre-of-mass energy$\sqrt{s}$ [41].

$\sqrt{s}(TeV)$	ggF	VBF	WH	ZH	$t\bar{t}H$	total
13	$48.5^{+4.6\%}_{-6.7\%}$	$3.78^{+2.2\%}_{-2.2\%}$	$1.37^{+2.6\%}_{-2.6\%}$	$0.88^{+4.1\%}_{-3.5\%}$	$0.50^{+6.8\%}_{-9.9\%}$	55.1
14	$54.7^{+4.6\%}_{-6.7\%}$	$4.28^{+2.2\%}_{-2.2\%}$	$1.51^{+1.9\%}_{-2.0\%}$	$0.99^{+4.1\%}_{-3.7\%}$	$0.60^{+6.9\%}_{-9.8\%}$	62.1

Table 1.1: Theoretical predicted Higgs boson production cross sections for a Higgs boson with $m_H = 125.10 \pm 0.14$ GeV at different center of mass energy ($\sqrt{s}$) [41].

such as vector boson fusion (VBF), W/Z-associated Higgs strahlung (WH/ZH), and top quark pair associated production ($t\bar{t}H$) also contribute [17]. The Feynman diagrams corresponding to the dominant Higgs production at the LHC are shown in figure 1.4. Table 1.1 lists the theoretically predicted cross sections for these production modes of the Higgs boson. The ggF is calculated at the next-next-next-to-leading-order (N3LO) level. The VBF, WH, ZH processes are calculated at the next-next-to-leading-order (NNLO) level. The ttH is calculated at the next-to-leading-order (NLO) level.

The expected branching ratios of the Higgs boson decays are functions of the mass of the Higgs boson. Figure 1.5 illustrates the relation between the branching ratio of each decay mode and the mass of the Higgs boson.

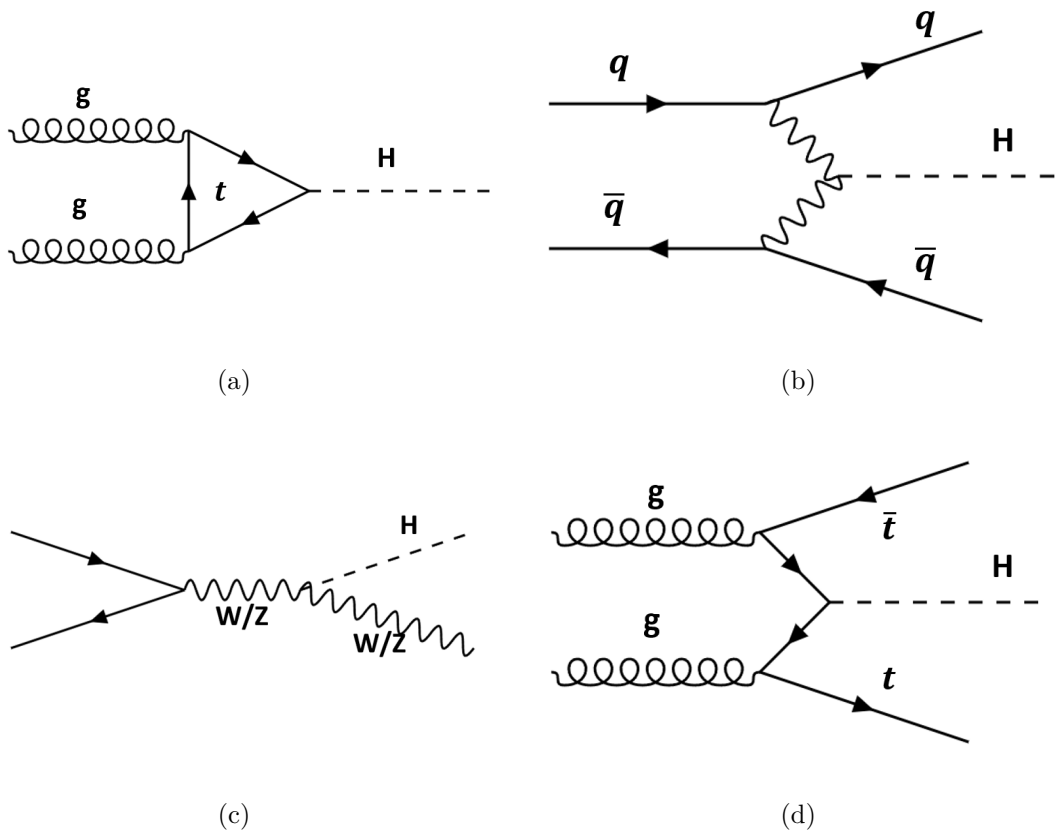


Figure 1.4: Main production channels of the Higgs boson in the $p-p$ collision at $\sqrt{s} = 13$ TeV. (a) gluon fusion (ggF); (b) vector boson fusion (VBF); (c) W/Z-associated Higgs strahlung (WH/ZH); (d) top quark pair associated production ($t\bar{t}H$).

Decay mode	branching ratio
$b\bar{b}$	$57.4^{+3.23\%}_{-3.29\%}\%$
$\tau^-\tau^+$	$6.29^{+5.71\%}_{-5.63\%}\%$
$c\bar{c}$	$2.89^{+12.21\%}_{-12.18\%}\%$
$\mu^-\mu^+$	$0.0218^{+5.95\%}_{-5.86\%}\%$

Table 1.2: Expected branching ratio for the main fermionic decay modes of the Standard Model Higgs boson [17, 41].

For a Standard Model Higgs boson with $m_H = 125.10 \pm 0.14$ GeV, the dominant fermionic decay modes are shown in the table 1.2. The $b\bar{b}$ decay mode has the highest branching ratio in the fermions. The CP-analysis in this mode is difficult due to the heavy QCD background in the LHC. The analysis of the LHC Run-2 data showed evidence of the $b\bar{b}$ decay mode [42] and $\tau^-\tau^+$ decay mode [43] in 2018. The $\mu^-\mu^+$ decay mode has a much lower branching ratio than the $\tau^-\tau^+$ decay mode [17]. The events containing the process of Higgs decaying into $\mu^-\mu^+$ pair are rare in

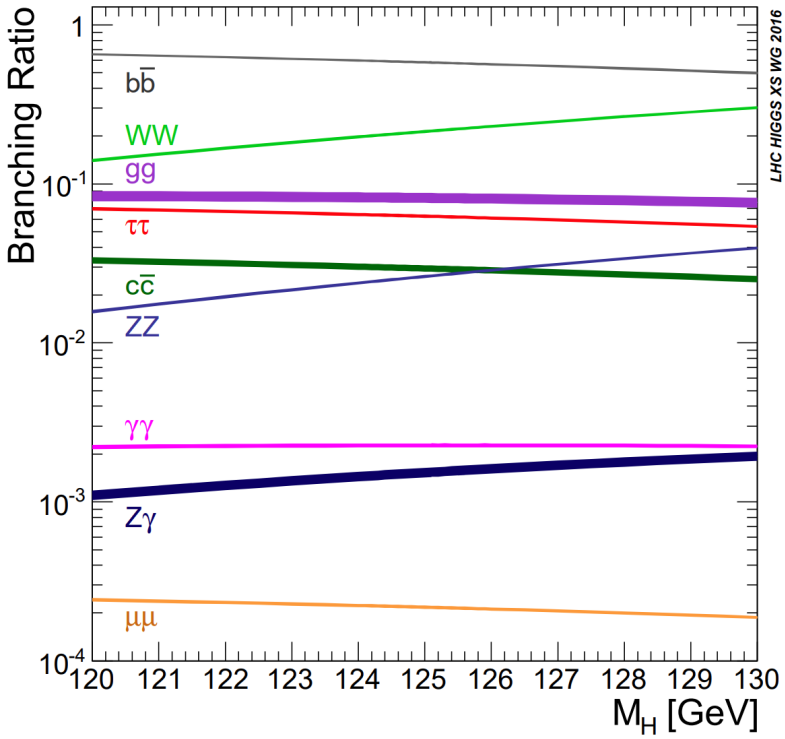


Figure 1.5: Expected branching ratios for the main decay modes in different Higgs boson mass (M_H) [17].

the data collected during the Run-2. Both the ATLAS [44] and CMS experiments [45] have found the evidence of $H \to \mu\mu$ decay, but the sensitivity has not reached the level of discovery so far.

Chapter 2

Experimental setup

2.1 The Large Hadron Collider

At the time of writing this thesis, the Large Hadron Collider (LHC) is the largest and most powerful particle accelerator. The accelerator complex at CERN is a succession of machines that accelerate particles to increasingly higher energies. Each machine boosts the energy of a beam of particles before injecting it into the next machine in the sequence. The LHC is the last element in this chain. The designed maximum collision energy of LHC is 14 TeV [46, 47].

Linear accelerator 4 (Linac4) is the source of proton beams for the CERN accelerator complex. It accelerates negative hydrogen ions (H_2^-, consisting of a hydrogen atom with an additional electron) to 160 MeV to prepare them to enter the Proton Synchrotron Booster (PSB). The ions are stripped of their two electrons during injection from Linac4 into the PSB, leaving only protons. These are accelerated to 2 GeV for injection into the Proton Synchrotron (PS), which pushes the beam up to 26 GeV. Protons are then sent to the Super Proton Synchrotron (SPS), where they are accelerated up to 450 GeV [47]. After that, the protons are transferred to the two beam pipes of the LHC. The beam in one pipe circulates clockwise while the beam in the other pipe circulates anticlockwise. It takes 4 minutes and 20 seconds to fill each LHC ring, and 20 minutes for the protons to reach the energy of 6.5 TeV. Beams circulate for many hours inside the LHC beam pipes under normal operating conditions. The single beam lifetime typically starts at around 20 hours. The lifetime of the beam increases to over 30 hours after 6 hours of stable beams [48]. The two beams are brought into collision inside four detectors –ALICE, ATLAS, CMS, and LHCb [47]. Figure 2.1 shows the schematic view of the CERN Accelerator complex.

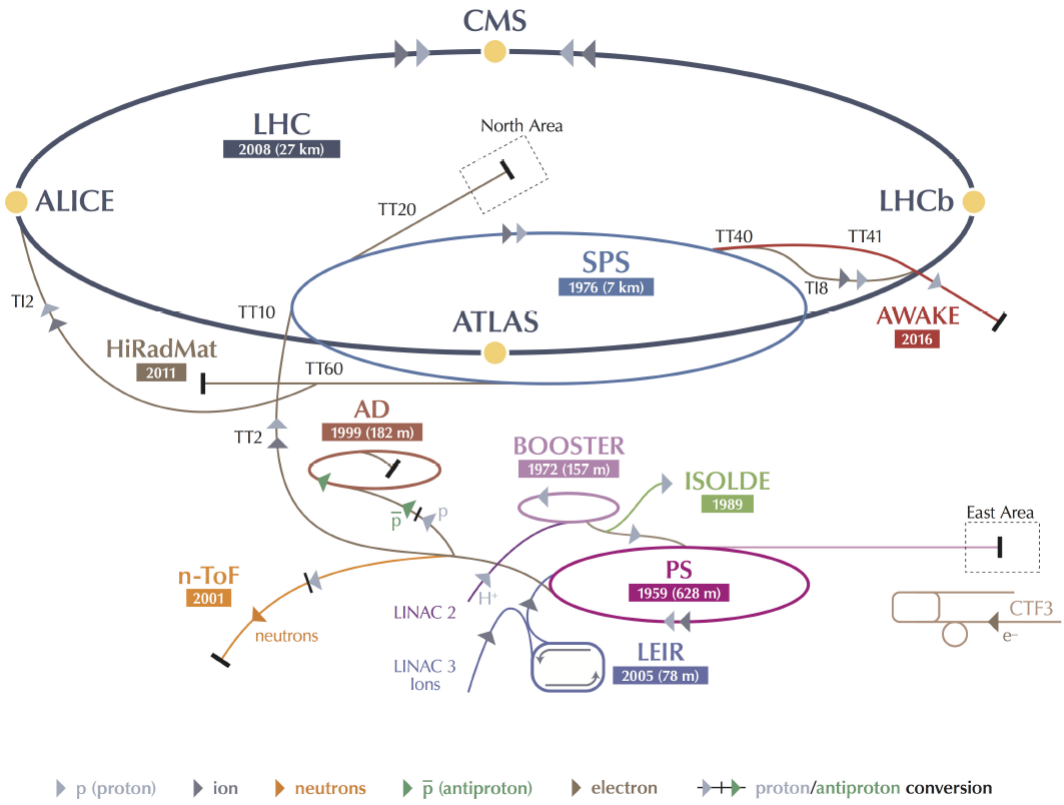


Figure 2.1: Schematic view of the CERN Accelerator complex [39].

Protons are not the only particles accelerated in the LHC. Lead ions for the LHC start from a source of vaporized lead and enter Linac3 before being collected and accelerated in the Low Energy Ion Ring (LEIR). They then follow the same route to maximum energy as the protons [49].

Two LHC operational runs have been finished. After each operational run, the machine was undergoing a maintenance period for about two years.

The number of events generated in the LHC collisions is given by:

$$N = \sigma_{event} \int L dt, \tag{2.1}$$

where the σ_{event} is the cross section of the interaction, L is the instantaneous luminosity. The cross section is expressed with unit barns where 1 barn = 10^{-24} cm^2. The integrated luminosity is measured in units b^{-1}. The integrated luminosity is often simplified as the luminosity. The instantaneous luminosity L can be expressed as:

$$L = \frac{f n_b N_p^2 \gamma}{4\pi\epsilon\beta^*} F, \tag{2.2}$$

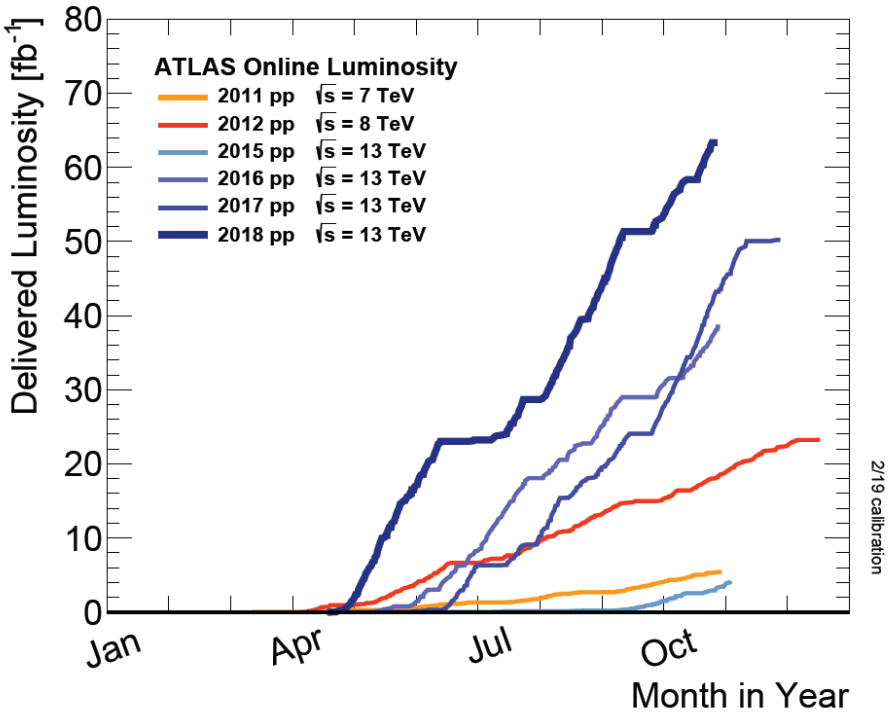


Figure 2.2: Delivered luminosity for each year of Run-1 and Run-2 [39].

where f is the revolution frequency [50]. The N_p is the typical number of protons per bunch, n_b is the number of bunches per beam, γ is the relativistic factor, ϵ is the normalized transverse beam emittance, β^* is the beta function at the collision point, and F is a geometric luminosity reduction factor due to the crossing angle at the interaction point. For head-on collision, the value of factor F is 1 and it becomes smaller when the crossing angle increases. The product $\epsilon\beta^*$ is directly related to the beam size at the interaction point. In order to achieve a high luminosity, more high population bunches need to collide at high frequency with the beam size as small as possible [49].

Figure 2.2 shows the integrated luminosity delivered to the ATLAS detect in each year during Run-1 and Run-2. There was more luminosity delivered to the ATLAS detector during 2015-2018 compared with the Run-1 period from 2011-2012 [39].

There are multiple radio frequency (RF) cavities inside the LHC to constrain the proton into bunches. The RF accelerates the protons and distributes bunches of protons that are spaced apart from each other. Each bunch contains around 10^{11} protons. These bunches are separated from each other by 25 ns. There are about 2808 bunches loaded in the LHC during operation [39].

The protons are accelerated in opposite directions in the LHC and collide inside the detectors. Just prior to the collision, the bunch is squeezed by focusing magnets

to maximize the collision rate. There are four main experiments and detectors along LHC. A Toroidal LHC ApparatuS (ATLAS) [39], Compact Muon Solenoid (CMS) [51], Large Hadron Collider beauty (LHCb) [52], and A Large Ion Collider Experiment (ALICE) [53].

ATLAS and CMS are general-purpose detectors. They are designed for wide-ranged experiments and their results can be used for cross-verification, such as testing the Standard Model and searching for new physics beyond the Standard Model. The LHCb focuses on precision measurements of CP-violation and rare decays of b-hadrons. ALICE aims to study physics of strongly interacting matter in the collisions of heavy nuclei.

2.2 The ATLAS detector

The ATLAS detector is one of the two major general-purpose detectors at the LHC. It has a cylindrical shape surrounding the LHC beam pipe. Figure 2.3 shows the structure of the ATLAS detector. The detector has a length of 44 meters and a height of 25 meters. The beams collide at the center of the ATLAS detector at a frequency of 40 MHz.

Figure 2.4 shows the luminosity-weighted distribution of the mean number of interactions per crossing for the proton collisions in ATLAS Run-2. There are five curves in the plot. The area under each curve represents the integrated luminosity recorded in that period. The total recorded integrated luminosity during 2015-2018 is 146.9 fb^{-1}. The average number of interactions is 33.7 in the whole Run-2 period. There are pile-up effects in the recorded data. The pile-up is caused by several $p-p$ collisions recorded as one event. There are in-time pile-up and out-of-time pile-up. The in-time pile-up is the additional p-p collisions occurring in the same bunch-crossing as the collision of interest. The out-of-time pile-up is due to the electronics delay of the detectors. The unfinished detector response is reconstructed as out-of-time pile-up collisions.

ATLAS is composed of a series of modular sub-detectors. The major sub-modules are the Inner Detector, the Calorimeters and the Muon Spectrometer. Each of them records certain information of the passing particles. The structure of ATLAS can be divided into two parts, the barrel and the end-caps. The detectors in the barrel region locate in a cylinder along the beam line. The detectors in the end-cap region are arranged perpendicular to the beam line to measure the particles in low angle.

The coordinate system used to describe the ATLAS detector is summarised as follows [54]. The nominal interaction point is defined as the origin of the coordinate

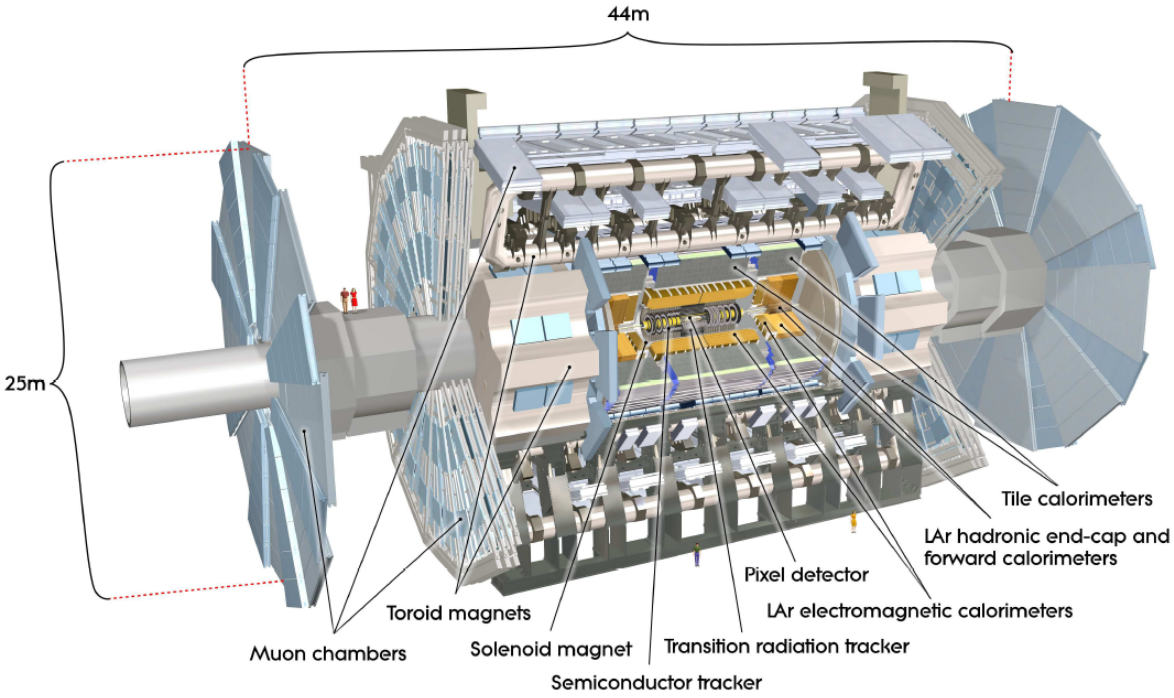


Figure 2.3: The ATLAS detector [54].

system, while the beam direction defines the z-axis and the x-y plane is transverse to the beam direction. The positive x-axis is defined as pointing from the interaction point to the centre of the LHC ring and the positive y-axis is defined pointing upwards. The side-A of the detector is defined as that with positive z and side-C is that with negative z. The azimuthal angle ϕ is measured as usual around the beam axis, and the polar angle θ is the angle from the beam axis. The pseudorapidity is defined as:

$$\eta = -ln\left(tan\left(\frac{\theta}{2}\right)\right). \tag{2.3}$$

The transverse momentum p_T, the transverse energy E_T, and the missing transverse energy E_T^{miss} are defined in the x-y plane. The distance ΔR in the pseudorapidity-azimuthal angle space is defined as:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \tag{2.4}$$

Magnet system

In the ALTAS detector, the trajectories of charged particles are bent by magnetic field to measure their momentum and charge. The magnetic field is generated by

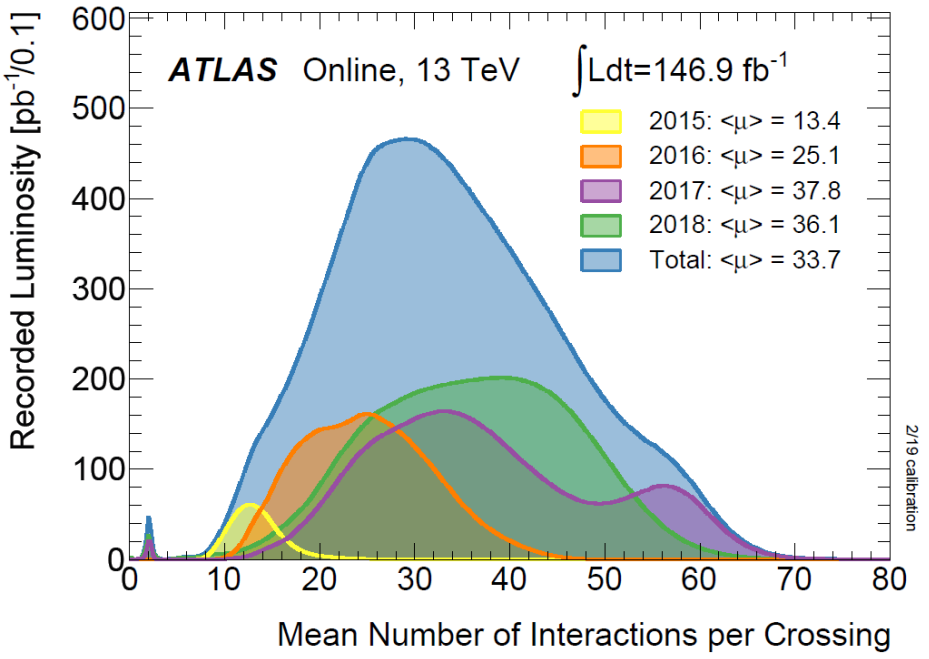


Figure 2.4: Distributions of the number of interactions per bunch-crossing [39].

superconducting magnet system. There are two main sections of the magnet system: central solenoid magnet and toroid magnet.

The ATLAS solenoid surrounds the inner detector at the core of the experiment. This powerful magnet is 5.6 m long, 2.56 m in diameter and weighs over 5 tonnes. It provides a 2 T magnetic field in just 4.5 cm thickness.

The ATLAS toroids use a series of eight coils to provide a magnetic field of up to 3.5 T, used to measure the momentum of muons. There are three toroid magnets in ATLAS: two at the ends of the experiment, and one massive toroid surrounding the centre of the experiment. At 25.3 m in length, the central toroid is the largest toroidal magnet ever constructed. It is unique in particle physics and an iconic element of ATLAS. It uses over 56 km of superconducting wire and weighs about 830 tonnes. The end-cap toroids extend the magnetic field to particles leaving the detector close to the beam pipe. Each end-cap is 10.7 m in diameter and weighs 240 tonnes [54].

The Inner Detector

The Inner Detector is composed of three subdetectors: the Pixel Detector, the Semiconductor Tracker (SCT), and the Transition Radiation Tracker (TRT). The

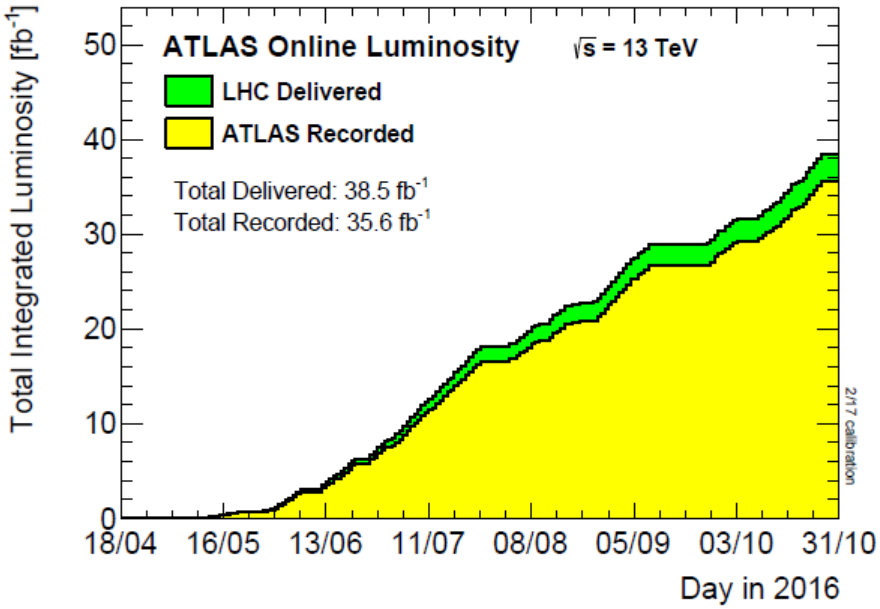


Figure 2.5: Integrated luminosity Recorded by ATLAS detector during 2016. Around 92% of delivered luminosity are detected [39].

Inner Detector is 2.1 meters in diameter and 6.2 meters in length. The layout is presented in figure 2.7. The Pixel Detector is closest to the beam. The SCT surrounds the Pixel Detector and TRT surrounds the SCT. The main function of the Inner Detector is to measure the trajectory of charged particles. As charged particles traverse the detector, they leave behind a track of small energy deposits (or hits) in each sub-detector, allowing the reconstruction of the trajectory of the particle. The Pixel and SCT detectors cover the region $|\eta| < 2.5$ and TRT cover the region $|\eta| < 2.0$. Figure 2.8 shows the structure of the Inner Detector. The Pixel Detector covers the radius from 33.25 mm to 122.5 mm. The SCT covers a radius of 299 mm to 514 mm, It is surrounded by the TRT with 528 mm in thickness.

The Pixel Detector consists of 1744 silicon pixel modules. These modules are arranged in four concentric barrel layers and three disks at each end-caps. The end-caps detector is placed at each end of a barrel-shaped detector to provide the most complete coverage in detecting particles. Each module has 47232 pixels, covering an area of 16.4×60.8 mm ^{2} and having a thickness of 250 μm. The signal in a module is read out by 16 radiation-hard front-end chips bump-bonded to the sensor. There are approximately 80.4 million readout channels in the Pixel Detector of ATLAS. Hits in a pixel are read out if the signal exceeds a tunable threshold. The pulse height is measured using the Time-over-Threshold (ToT) technique. The ToT technique records a signal pulse by measuring the pulse width when the electric pulse has a

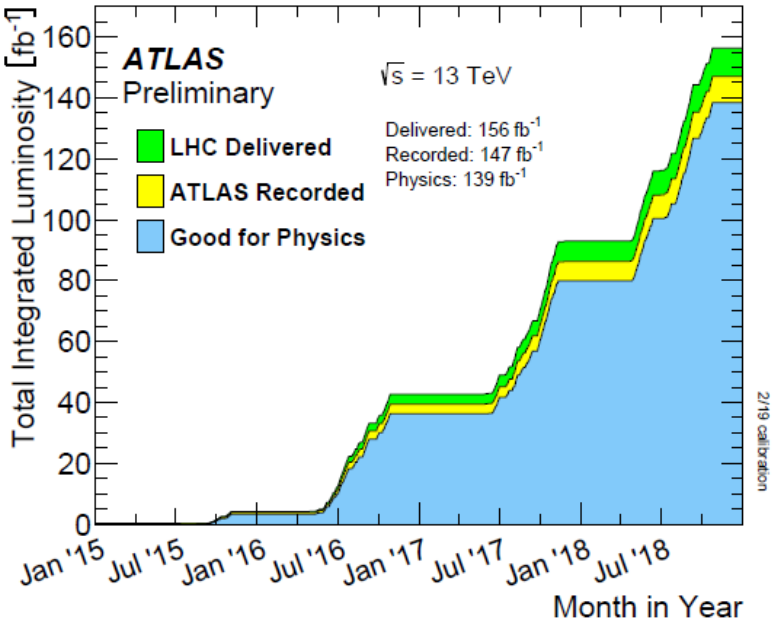


Figure 2.6: The total integrated luminosity during Run-2 [39].

voltage higher than the given threshold. The spatial resolution of the pixel detector is 10 μm in the azimuthal direction and 60 μm along the beam direction [56, 57].

The SCT consists of four coaxial cylindrical layers in the barrel region and nine disks on each endcap. All layers can read out a position in two dimensions. The SCT contains 4088 modules of silicon microstrip detectors. The microstrip is an electromagnetic detector constructed with a flat conductor suspended over a ground plane. The conductor and ground plane are separated by a dielectric material. Each layer or disk provides measurements from two strips arranged at an angle, to construct a space-points. Each of these silicon microstrips has a constant pitch of 80 μm and a thickness of 285 μm. There are around 6.3 million readout channels in the SCT. It gives a resolution of 17 μm in the azimuthal direction and 580 μm along the beam direction [58].

The TRT consists of thin-walled proportional drift straws [59]. Each proportional drift straw is a tube with a diameter of 4 mm, filled with a gas mixture of $Xe(70\%)$, $CO_2(27\%)$ and $O_2(3\%)$. There is a gold-plated tungsten anode wire of 31 μm diameter at the center of each proportional drift straw. When charged particles pass through the tube, the gas is ionized. The ions are drifted to the center wire and send an electric signal as a readout. There are 52544 straws of 1440 mm in length in the barrel region and 122880 straws with length of 370 mm in the end-caps, giving approximately 351000 readout channels and providing 130 μm spatial resolution in

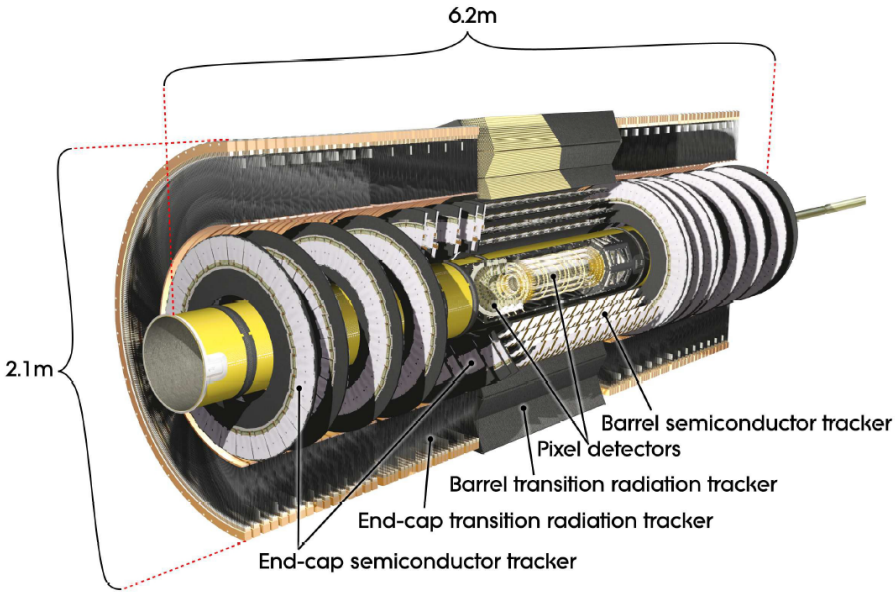


Figure 2.7: The ATLAS Inner Detector [54].

the azimuthal direction. The spatial layout of TRT is designed such that a charged particle with $|\eta| < 2.0$ with $p_T > 0.5$ GeV passes through at least 30 straws. These continuous readouts show a track trace in the detector [59].

Calorimeters

Surrounding the Inner Detector are the Calorimeters, which provide energy measurements of showering particles [54]. The ATLAS calorimeter system is illustrated in figure 2.9. There are two major sub-detectors, Electromagnetic Calorimeter (EMCal) and Hadronic Calorimeter (HCal).

The ATLAS Electromagnetic Calorimeter is a high-granularity sampling calorimeter. It uses liquid argon (LAr) as the active material and lead plates as the passive absorber. The EMCal submodules are arranged in an accordion-like geometry to avoid azimuthal cracks. The EMCal in the barrel region covers the range $|\eta| < 1.475$ and the end-cap region covers $1.375 < |\eta| < 3.2$. There is a crack region $1.375 < |\eta| < 1.5$ where the measurement is affected by additional material needed to cool the inner detector. This region is usually excluded from the analysis.

The calorimeter barrel is divided into three layers with different granularities. Layer 1 is segmented into strips of $\Delta\eta \times \Delta\phi = 0.0031 \times 0.098$, which helps separating γ/π and measuring the electromagnetic shower direction. The second layer takes 70% in the thickness of the EMCal. It collects most of the shower energy at a granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.0245$. The third layer coarsely segmented in

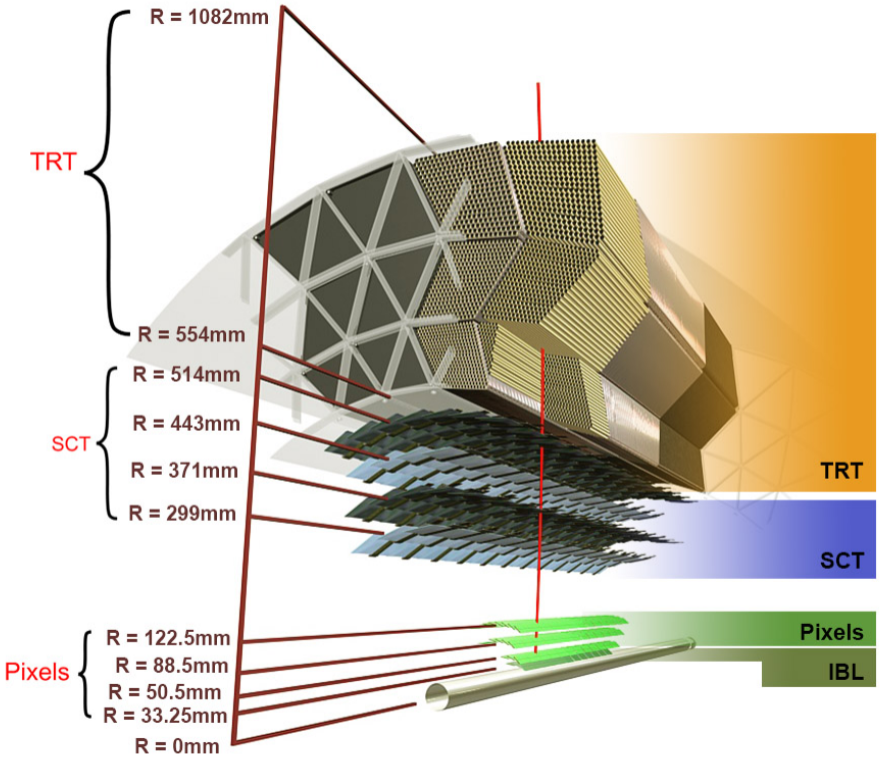


Figure 2.8: The three complementary subdetectors in the barrel region of the ATLAS Inner Detector [55].

η, with a granularity of$\Delta\eta \times \Delta\phi = 0.0031 \times 0.098$. It collects the shower tail and provides extra depth. The layout of layers in the EM barrel calorimeter is shown in figure 2.10.

The total thickness of EMCal is 22 radiation lengths in the barrel and 24 radiation lengths in the end-caps. The radiation length X_0 is defined as the mean path length of a high-energy electron losing all but $1/e$ of its energy by bremsstrahlung in the material [60]. The energy resolution of the EMCal was measured using a test-beam. After noise subtraction, the energy resolution is (E in GeV):

$$\frac{\sigma(E)}{E} = \frac{10\%}{\sqrt{E}} \oplus 0.7\%. \tag{2.5}$$

The hadronic jets pass through the EMCal and reach the Hadronic Calorimeter placed outside of EMCal. The HCal is a scintillating Tile Calorimeter (TileCal), which uses scintillating tiles as the active material and steel as the absorber [61]. The HCal covers the range $|\eta| < 1.7$. The HCal is divided azimuthally into 64 wedges of the size $\Delta\phi \approx 0.1$ [61]. The scintillator tiles are arranged normal to the radial direction. The wavelength shifting fibers at each edge of the scintillating

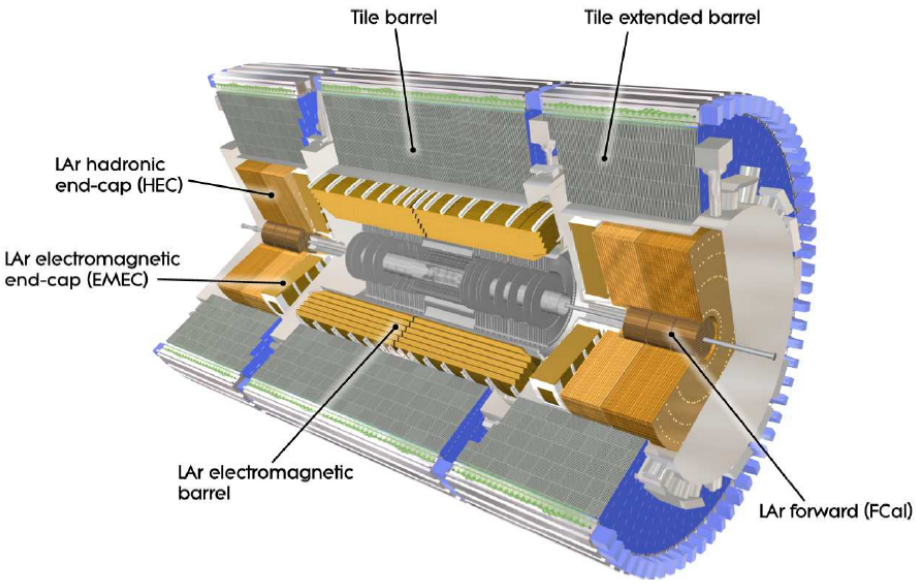


Figure 2.9: The ATLAS Calorimeters [54].

tiles collect the photon and generate amplified readout by photomultiplier tubes (PMTs). These fibers define the 3D geometry of a cell. The cell dimension is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in the first and second layers. and $\Delta\eta \times \Delta\phi = 0.2 \times 0.1$ in the third layer. The granularity of the HCal is lower than the EMCal, but the hadronic shower spread larger than the electromagnetic showers, making the accuracy acceptable. The HCal system also contains hadronic end-cap calorimeters (HEC). The HEC are copper/liquid-argon calorimeters. The HEC cover a range of $1.5 < |\eta| < 3.2$ [54].

In addition to the end-cap calorimeters, the Forward Calorimeter (Fcal) provides coverage over $3.1 < |\eta| < 4.9$. It is composed of electromagnetic modules and hadronic modules. The electromagnetic modules use LAr as active material and copper as absorber. The hadronic modules also use LAr as active material but use tungsten as absorber. The Fcal is designed to measure electromagnetic and hadronic showers [54].

The Muon Spectrometer

The muon spectrometer (MS) is the outermost ATLAS subdetector. The muons have a much higher penetration depth than other particles. The muon spectrometer measures the momentum of muon and helps reconstruct the trajectory. It provides a standalone measurement of muons within $|\eta| < 2.7$. It consists of four types of detectors: Monitored Drift Tubes (MDTs), Cathode Strip Chambers (CSCs),

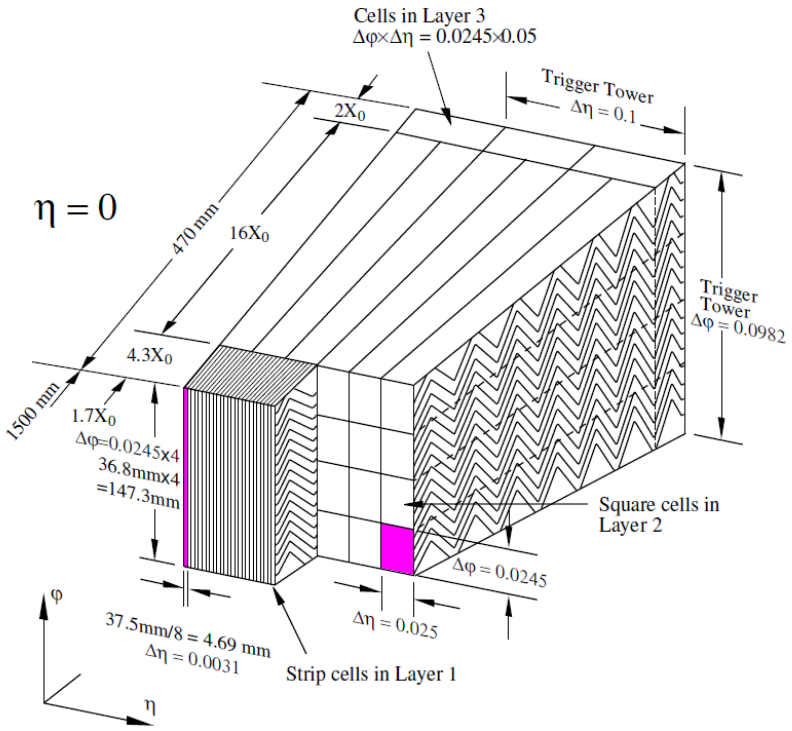


Figure 2.10: Three layers of ATLAS barrel Electromagnetic Calorimeter [54]. The X_0 is the radiation length of electron in the material.

Resistive Plate Chambers (RPCs), and Thin Gap Chambers (TGCs) [62]. Figure 2.11 shows the layout of ATLAS muon spectrometer [54].

The precision momentum measurement is performed by the MDT. The MDT cover the region $|\eta| < 2.7$, except for the innermost layer, where it covers the range $|\eta| < 2$. The basic element of the MDT is a tube with a diameter of 29.97 mm. The tube is made of aluminum and is filled with Ar/CO_2 gas (93%/7%) at 3 bar. The MDT achieves an average resolution of 35 μm for a muon passing through a MDT chamber [54].

The hardware limit of MDT to receive signals is 150 Hz/cm^2. This limit is exceeded in the end-cap region of the detector. In the forward region within $2 < |\eta| < 2.7$, the MDT is replaced by CSC, due to its higher capability. The CSC can be safely operated at a counting rate of 1000 Hz/cm^2, which is sufficient for muon detection in the $2 < |\eta| < 2.7$ region. The tubes of the CSC are filled with Ar/CO_2 gas (80%/20%). The CSC reaches a resolution of 40 μm in the radial direction [54].

The MDT and CSC are used for reconstructing the trajectories of muons. The RPC and TGC are used for triggering. The RPC detectors locate at the barrel region $|\eta| < 1.05$. The RPC is a gaseous parallel electrode-plate detector. It is composed of two plate parallel to each other at a distance of 2 mm. It is filled with

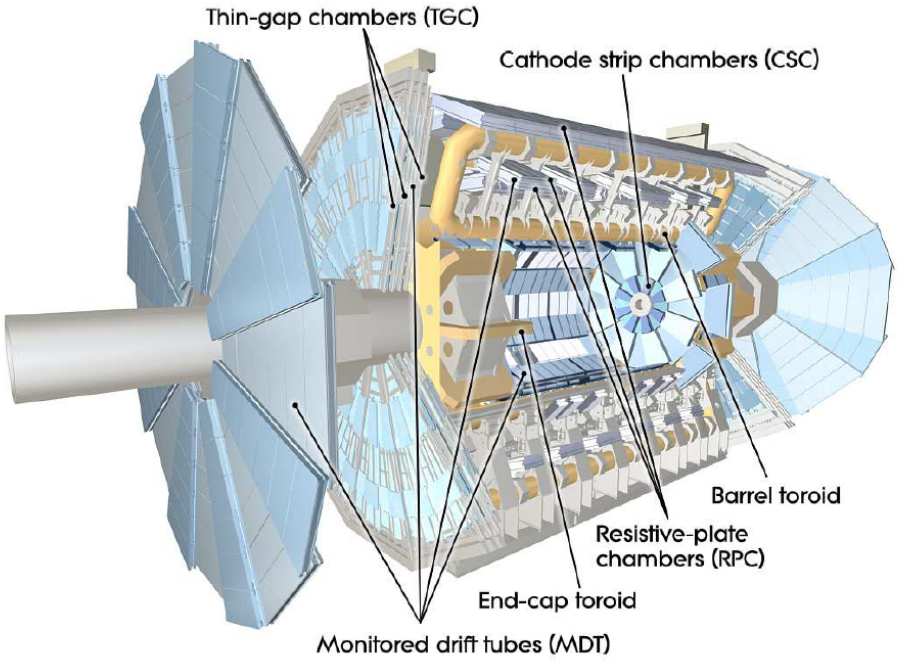


Figure 2.11: The ATLAS Muon Spectrometer [54].

a gas mixture of $C_2H_2F_4$/Iso-C_4H_{10}/SF_6 (94.7%/5%/0.3%). The TGC detectors are located at the end-cap region $1.05 < |\eta| < 2.4$. They have a time resolution of 1.5 ns. The TGC is composed of chambers filled with a gas mixture of CO_2/n-pentane gas (55%/45%). It provides a time resolution of 4 ns. With the signal collected by the PRC and TGC, the trigger system delivers a muon candidate within tens of nanoseconds after the passage of a particle [54].

Triggers

The trigger system of ATLAS is designed to select collisions to be recorded. The trigger system is formed by a hardware-based first-level trigger (L1) and a software-based high-level trigger (HLT). The HLT contains a software-based second-level trigger (L2) and event filter [54].

The L1 selects events based on the presence of particles with large transverse momentum, such as muons, electrons, photons, jets, and hadronic decaying τ. It also selects events with large total transverse energy or large missing transverse energy. The maximum acceptance rate of the L1 trigger is 75 kHz and possibly to be upgraded to 100 kHz in the future. Decision of the L1 trigger is made within 2.5 μs after a bunch-crossing occurs.

The L2 trigger is seeded by regions-of-interest that has been labeled possible trigger objects by the L1 trigger. It further limits the data rate to 3.5 kHz with the information of coordinates, energy, and types. On average, an event is processed around 40 ms.

The event filter is an offline trigger. It selects events with a maximum rate of 200 Hz and stores the selected events in an offline storage [63].

The L1 and L2 use the information from the calorimeters and the muon chambers for their decision. The overall HLT algorithm also uses the information from the inner detector to refine the selection. With better energy deposition and track reconstruction, it significantly improves the identification of particles.

2.3 Data reconstruction

The interaction between the produced particles and the detector is recorded as electronic signals, with a combination of position, time, and energy deposit. The computing algorithms are designed to reconstruct the physics information from the recorded detector data.

Tracks and primary vertices

In particle physics, a track is the reconstructed trajectory of a charged particle in the detector. The vertex is the position where an interaction happens. The reconstruction of tracks and vertices is the basis of object identification. For charged particles, the track information is reconstructed with the inner detector. The muon reconstruction also uses the muon spectrometer. There are three stages for reconstructing a track with the inner detector: pre-processing, track-finding, and post-processing stage [64]. The pre-processing stage uses the raw data from the Pixel and SCT detectors. The detected information is converted to positions in the lab frame.

The track-finding stage involves a set of different tracking strategies [65]. These strategies are optimized for different cases. By default, the track-finding algorithm finds a combination of space points as a track seed. These space points are associated to the detected signals in the pixel layers and the first SCT layer. Then the algorithms attempt to extend these seeds in the SCT region to get a track candidate. A set of fitting algorithms are applied to reject the outlier candidates and mis-identified tracks. These selected tracks are associated to the detected information in the path. Then the found tracks are refitted with full information of all detectors in the inner detector to further improve its accuracy [54].

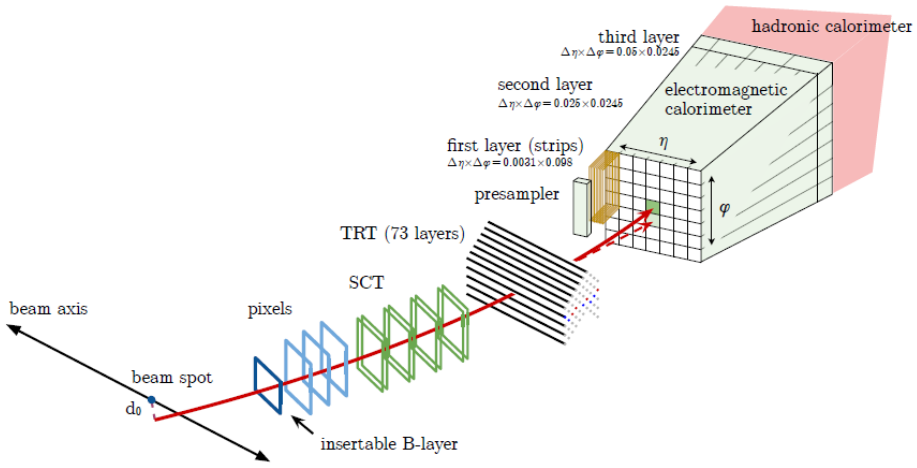


Figure 2.12: Illustration of the reconstruction of electron [64].

The primary vertex is defined as the point where $p-p$ collision occurred. Primary vertices are reconstructed from a collection of reconstructed tracks. These tracks must satisfy the following requirements. Each of these tracks must has $p_T > 400$ MeV and $|\eta| < 2.5$. There must be at least 9 recorded detected interactions in the SCT if the track is in $|\eta| < 1.65$, or 11 detected interactions if the track is in $1.65 < |\eta| < 2.5$. There must be at least 1 detected interaction in the first two layers of the Pixel detector. The track has at most 1 shared detected interaction with the other reconstructed tracks in the Pixel detector, and has at most 2 shared detected interaction with the other tracks in the SCT. The track should not has a hole in the Pixel detector and at most 1 hole in the SCT, where the hole represents a place on a detector surface that is expected passed by but not observed [66].

These selected tracks are used as input of vertex finding and vertex fitting algorithms to get primary vertices. It is an iterative procedure. In each iteration, it finds the best fit value of a vertex position that is associated to at least two tracks. The less compatible tracks are down-weighted. Then the vertex position is recalculated. After one vertex is determined, the tracks incompatible with the vertex are used as input of vertex finding algorithm again, until no additional vertex can be found [66].

Electrons and photons

The electron and photon have similar behavior in the EMCal. The electron reconstruction uses the reconstructed track and the clusters in the calorimeter. A cluster represents a set of detected interaction positions that close to each other, since a particle always leaves signals in multiple calorimeter elements.

Figure 2.12 demonstrates the detection of an electron. The electron passes through the pixels, SCT, TRT, presampler and finally reaches the calorimeter. The red dash line shows a photon produced during the interaction of this electron with the detector [64]. It also reaches the calorimeter and leaves an energy deposit near the electron [67].

The electron track identification is done in two steps. The first step is pattern recognition. It attempts to use the π hypothesis to fit the energy loss of the candidate. If it is not likely to be a π, an electron hypothesis is fitted to the signal where up to 30% energy loss due to bremsstrahlung at each intersection is expected. Then the Gaussian Sum Filter algorithm is applied to the electron candidate to take non-linear bremsstrahlung into consideration [68].

Muons

Muons are reconstructed from the tracks detected by the inner detector and the muon spectrometer. When hits are detected in the muon chamber, the searching algorithm will try to match them with the other related signals. If it happens in the MDT, it will try to search each layer to get a segment, which represents a set of connected signal positions. If it is found in the CSC, the segment is reconstructed in the $\eta - \phi$ plane. And the RPC and TGC can measure the vector orthogonal to the bend plane. Two segments or a high-quality segment can build a track. There are four types of reconstructed muon tracks: combined muons, segment-tagged muons, calorimeter-tagged muons, and extrapolated muons [69].

The combined muons use both the inner detector and the muon spectrometer information. It requires a global fit to match information from different sources. Some hits may be removed or added to improve the fit. Most muon candidates are reconstructed starting from the outermost hit and matching the hits of the inner detector tracks. Segment-tagged muons are those reconstructed with at least one segment of the track in the MDT or the CSC chambers additional to the MS information. This method is used when the track only leaves information in one layer of the MS chambers after the trigger. Calorimeter-tagged muons use the energy deposit in the calorimeter. Muons reconstructed in this way have the lowest purity among the four types, but have high recovery acceptance when $|\eta| < 0.1$. The extrapolated muons are those with only MS information. This method can recover the muons in the $2.5 < |\eta| < 2.7$ range, which has exceeded the range of the inner detector [69].

Some muon background candidates come from pion and kaon decays. To suppress this kind of background, quality requirements are applied. These requirements are

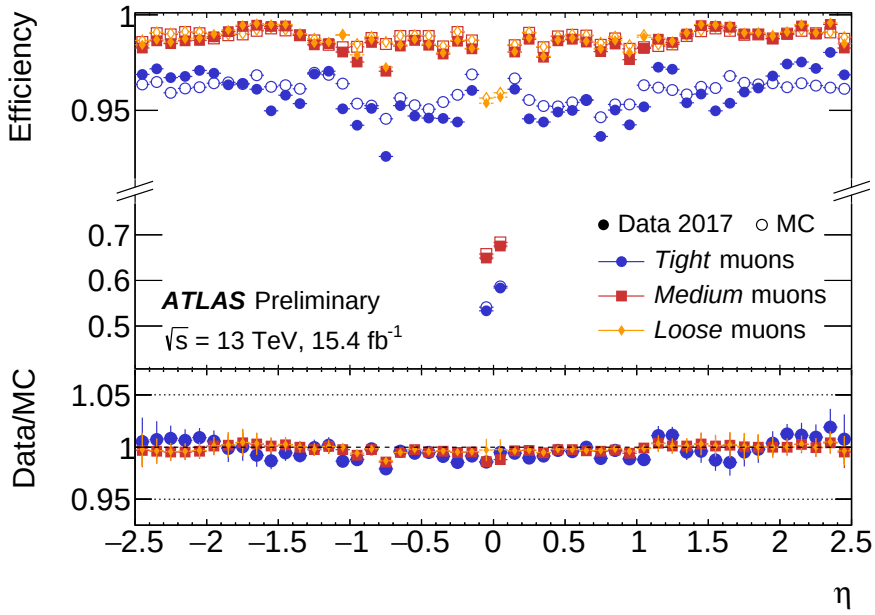


Figure 2.13: Muon reconstruction efficiencies for the Loose/Medium/Tight identification algorithms [70].

based on the number of hits in the ID and MS. Figure 2.13 shows the efficiency of muon reconstruction in Loose/Medium/Tight identification categories.

Jets

Jets are collimated sprays of hadrons produced in the hadronization process. Jets are reconstructed with the calorimeter readout and the particle trajectories detected with the inner detector. The calorimeter cells are used as input and expanded from a calorimeter seed cell with a growing volume algorithm. The topologically connected calorimeter cells are called topo-clusters. Each topo-cluster is calibrated with the energy deposit of electromagnetically interacting particles [71].

The calibration of reconstruction consists of several steps. First, the hard-scattering vertex is selected as the primary vertex with the largest $\sum p_T^2$, summed over all the tracks associated with it. A correction to the jets is applied according to the choice of hard-scattering vertex. Then the effect of pile-up is removed. After that, the four-momentum of jets are corrected by simulation-based calibration. Improvements are made by taking calorimeter readout, muon segment, track variables, and global sequential calibration into account. The jet energy is calibrated with the real data [72].

The pile-up jets are removed with an algorithm called Jet-Vertex-Tagger (JVT). JVT is a multivariate algorithm trained with the ATLAS Monte Carlo simulated

samples. It is applied to jets within a certain range of transverse momentum (p_T) and pseudorapidity (η). The jets within $20 < p_T < 30$ GeV, $|\eta| < 2.4$ range are given JVT scores between 0 and 1. The JVT score close to 1 represents a high possibility of being a hard-scattering jet. The jets with $p_T > 50$ GeV are assumed to be hard-scattering jet [73].

Hadronic τ decays

The hadronically decayed $\tau-$lepton candidates are seeded by jets. These jets are reconstructed using an anti-k_t algorithm [74, 75, 76]. The clusters of calorimeter cells are used as inputs to the anti-k_t algorithm. The $\tau-$lepton candidates seeded by jets are additionally required to have $p_T > 10$ GeV and $|\eta| < 2.5$ [77, 76].

The $\tau-$lepton vertex is selected according to the fraction of momentum from the tracks associated with a jet. The tracks are required to have $p_T > 1$ GeV. The shortest distance from the track to the $\tau-$lepton candidates vertex in the transverse plane $d_0 < 1$ mm. The shortest distance in the longitudinal plane, $|\Delta z_0 sin(\theta)| < 1.5$ mm, where Δz_0 is the closest approach along the longitudinal axis, and θ is the polar angle of the track [77].

The direction of the $\tau-$lepton candidate is calculated using the vector sum of the clusters of calorimeter cells within a cone limited by the angular difference $\Delta R < 0.2$ to the direction of the seed jet, originated from the $\tau-$lepton vertex. The energy of the $\tau-$lepton candidate is calculated from the energy deposition measured in these clusters of calorimeter cells. A calibration algorithm is applied to correct the energy deposition to the average value of the energy carried by the $\tau-$lepton candidate [77].

The reconstructed $\tau-$lepton candidates contain other particles and jet-like signatures. A recurrent neural networks (RNN) discriminant is developed to separate the $\tau-$lepton from jets initiated by quarks or gluons [76]. And a Boosted Decision Tree (BDT) discriminant is developed to separate misidentified hadronic $\tau-$lepton decay due to electrons [77]. The $\tau-$lepton selected by both discriminants is required to have 1 or 3 associated tracks, and have $p_T > 20$ GeV and $|\eta| < 2.47$ (excluding the region $1.37 < |\eta| < 1.52$).

Missing transverse energy

Not all the particles are visible in the detector. Some particles such as neutrinos hardly leave any signal in the detector. In this case, the corresponding transverse momentum will not be recorded. According to momentum conservation, the missing

transverse energy and momentum can be measured by summing up the transverse momentum of all visible parts.

Apart from the neutrino, some other visible particles can also have missing momenta, such as the electron, muon, photon, jets. The detectors also record energy deposits in calorimeters with no object associated to. These signals also contribute to the missing momenta [78].

Chapter 3

Jet-Vertex-Tagger Retraining

This chapter introduces the Jet-vertex-Tagger (JVT) algorithm, the multivariate algorithm designed to distinguish the hard-scattering jets from the pile-up jets. The author of this thesis developed new JVT models and retrained these models with the ATLAS Run-2 data and simulated samples, and expanded the applicable working range from $20 < p_T < 50$ GeV, $|\eta| < 2.4$ to $20 < p_T < 110$ GeV, $|\eta| < 2.5$. The new model with the best score improves the background rejection rate in three working points corresponding to the signal efficiency of 0.97, 0.92, and 0.85. These studies have not been applied to the recent ATLAS Run-2 analysis of measuring the CP-property in the $H \to \tau\tau$ decay. One of the new algorithms based on neural network has been adopted as the new default JVT algorithm in the current ATLAS Run-3 data analysis, which started in 2022. It may inspire future improvements of jet reconstruction.

3.1 Hard-scattering jet identification

At the LHC, multiple collisions happen when the beam bunches cross. Most of the events contain no collision of interest. Some events contain one or more collisions of interest. In ATLAS, the primary vertices are defined as the reconstructed positions of the $p - p$ collisions. In each event, the hard-scattering vertex is defined as the primary vertex with the largest summed p_T^2 of all trajectories it associated to. All the other reconstructed primary vertices are defined as pile-up vertices. The hard-scattering jets are defined as those associated with the hard-scattering vertex. The pile-up jets are those not associated with the hard-scattering vertex, no matter whether they are caused by in-time pile-up or out-of-time pile-up. The pile-up jets are rejected by a multivariable algorithm called Jet-Vertex-Tagger (JVT). The JVT

developed for ATLAS Run-1 data is based on the KNN algorithm with two variables $JVF/corrJVF$ and R_{p_T} as inputs.

Figure 3.1 shows an example of $p-p$ collision in the ATLAS detector. It is a reconstructed $Z \to \mu\mu$ event from the ATLAS Run-1 data, where the centre-of-mass energy is $\sqrt{s} = 8$ TeV. The two thick yellow lines represent the two muon trajectories from the Z boson generated from the hard-scattering collision. All other lines represent trajectories from the pile-up collisions [73].

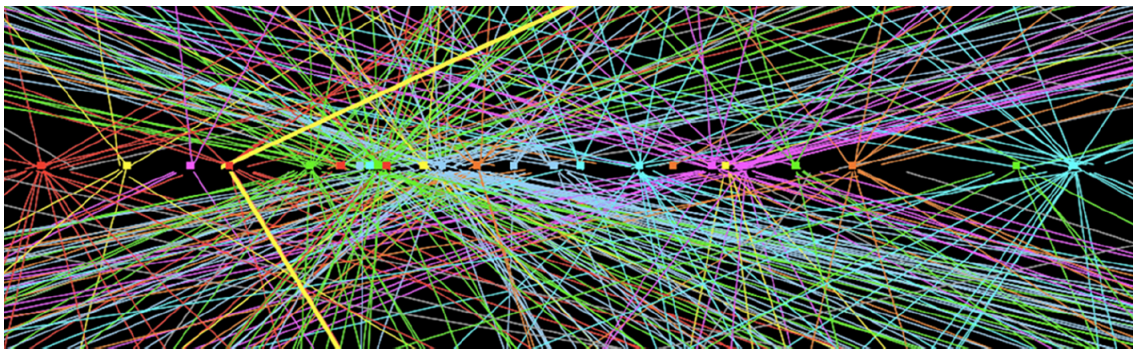


Figure 3.1: A $Z \to \mu\mu$ event from the ATLAS Run-1 data. The lines represent the trajectory of detected particles. The dots represent the collision points [73].

Jet Vertex Tagger

The JVT is trained with Run-1 data and Monte Carlo samples. The local neighborhood is defined dynamically as the 100 nearest neighbors around the test point using a Euclidean metric in the R_{p_T}-$corrJVF$ space. The two observables R_{p_T} and $corrJVF$ are calculated from the detected vertex and track information [79]:

$$R_{p_T} = \frac{\sum_k p_T^{trk_k}(PV_0)}{p_T^{jet}}, \tag{3.1}$$

$$corrJVF = \frac{\sum_k p_T^{trk_k}(PV_0)}{\sum_l p_T^{trk_l}(PV_0) + \frac{\sum_{n\geq 1}\sum_l p_T^{trk_l}(PV_n)}{C_w N_{trk}^{PU}}}. \tag{3.2}$$

The observable R_{p_T} is defined as the ratio between the p_T sum of all tracks from the primary vertex to the p_T of jet. The $p_T^{trk_k}(PV_0)$ is the transverse momentum of each track originated from the hard-scattering primary vertex. The p_T^{jet} is the transverse momentum of the tested jet. Figure 3.2 shows the distribution of R_{p_T} for the pileup and hard-scatter jets with $20 < p_T < 30$ GeV [79].

The observable $corrJVF$ is calculated using all the tracks related to the tested jet. The N_{trk}^{PU} is the number of tracks originating from the pile-up vertex. When

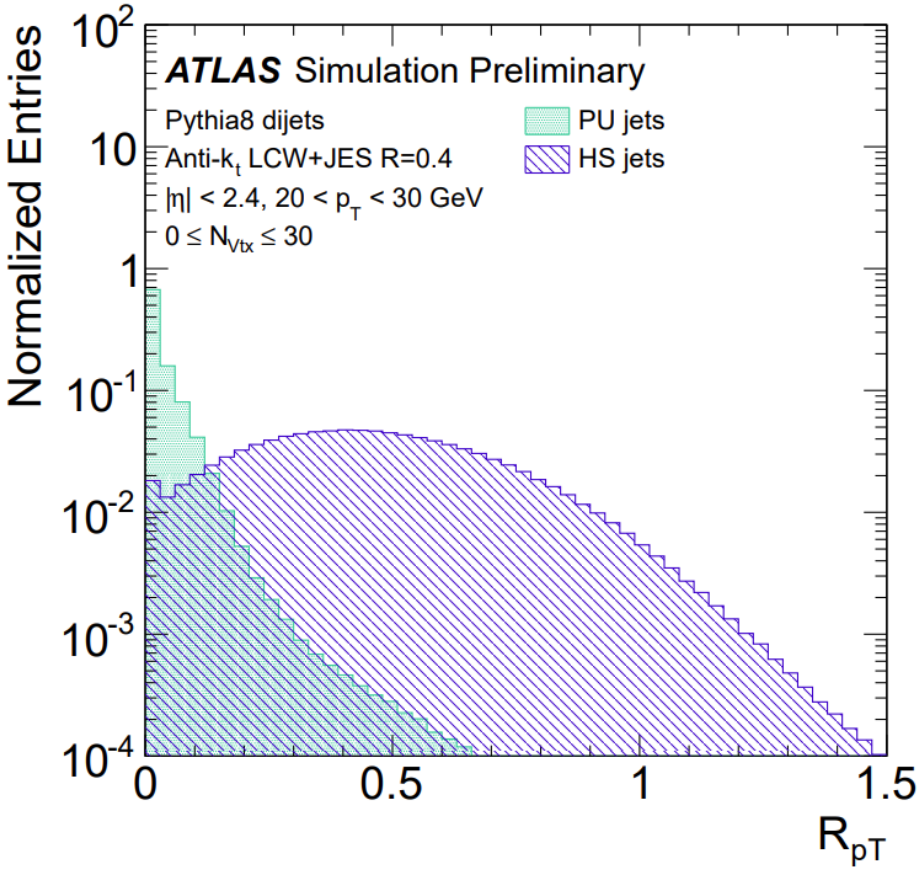


Figure 3.2: R_{p_T} distribution for pileup and hard-scatter jets with $20 < p_T < 30$ GeV in ATLAS Run-1 Monte Carlo simulated sample [79].

the number of primary vertex $n \geq 1$, the $p_T^{trk_i}(PV_n)$ is the transverse momentum of each track originated from the pile-up vertex. The corrJVF is the ratio between the p_T sum of all the tracks from the primary vertex to the p_T of all the tracks associated with the tested jet, with those from the pile-up vertex reweighted. There is a constant $C_w = 0.01$ in the weight of the pile-up tracks. The value of C_w is manually chosen [79]. Figure 3.3 shows the distribution of $corrJVF$ for the pile-up and hard-scattering jets with $20 < p_T < 30$ GeV [79].

The K-nearest-neighbors (KNN) algorithm is a non-parametric classification method [80]. Figure 3.4 shows a demonstration of the KNN algorithm. There are 7 red and 7 blue dots in a 2-D plane of R_{p_T}- $corrJVF$. The red dots represent the signal, which is the hard-scattering jets. The blue dots represent the background, which is the pile-up jets. The value of k is set to 6. The green dot is the jet to be tested. In the nearest 6 dots, 4 of them are signals and 2 of them are backgrounds. When

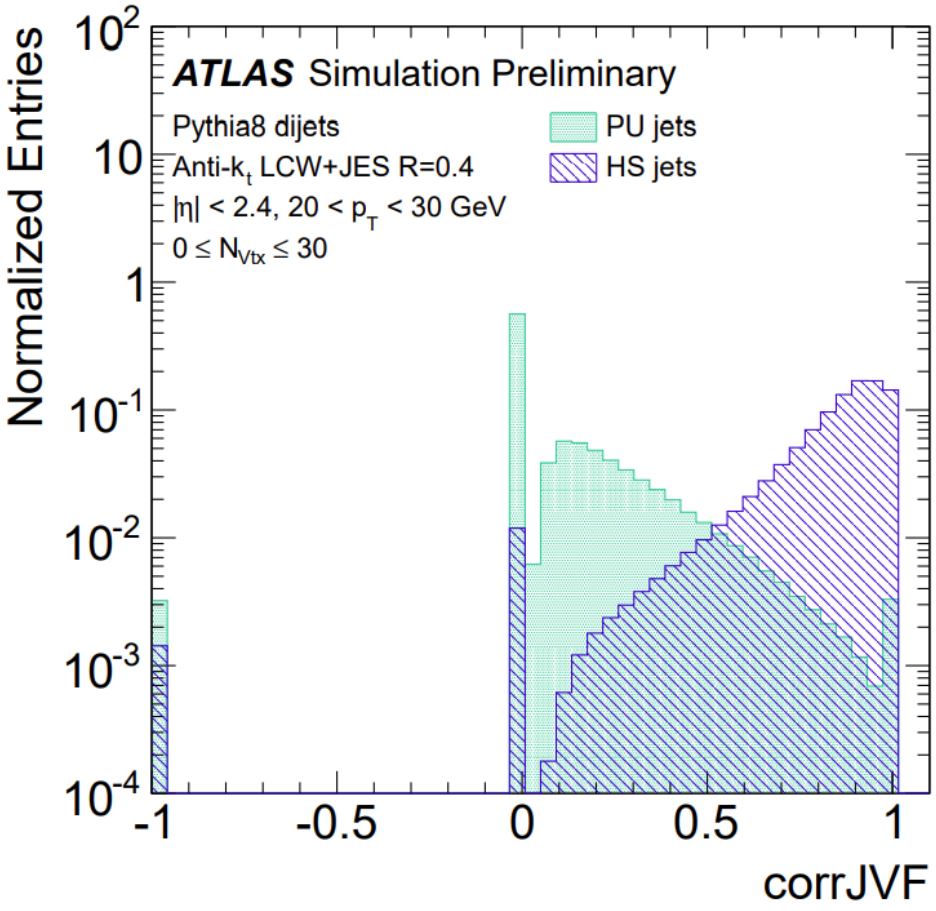


Figure 3.3: CorrJVF distribution for pileup and hard-scatter jets in ATLAS Run-1 Monte Carlo simulated sample [79].

both the weights of signal and background are 1, The KNN score is 0.67. If the working point is selected to be 0.50, the jet is identified as a hard-scattering jet. If the working point is selected to be 0.70, the jet is identified as a pile-up jet.

Figure 3.5 shows the JVT distribution for hard-scattering and pileup jets with $20 < p_T < 30$ GeV in the Run-1 MC sample [79]. The hard-scattering jets have JVT score close to 1 and pileup jets are close to 0. Working point is assigned for certain analysis depending on the required signal efficiency. The jets with JVT scores larger than the working point are identified as hard-scattering jets, otherwise pileup jets. A working point closer to 1 rejects more pileup jets but reduces the signal efficiency. A lower cut retains more signal events but more background is left.

The JVT algorithm was updated for Run-2 analysis with the data collected in 2015 [81], where the energy scale was re-calibrated. The Run-2 JVT still used the

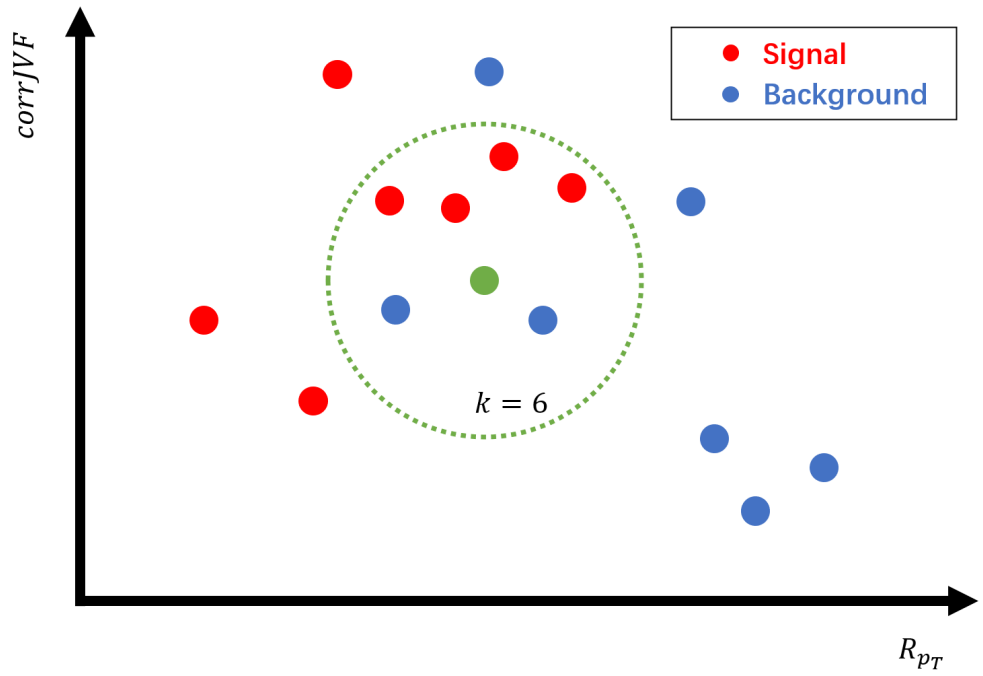


Figure 3.4: Demonstration of the KNN method. The green dot is a tested jet.

KNN method in R_{p_T}-$corrJVF$ space.

3.2 Jet-Vertex-Tagger retraining

The ATLAS Run-2 was performed under a center of mass energy larger than Run-1 and the detectors were upgraded. The Run-2 JVT was required to identify the hard-scattering jet in a larger range. The Run-1 JVT was designed to work in the $20 < p_T < 50\ GeV, |\eta| < 2.4$ range. The Run-2 analysis required the JVT to work in a range of $20 < p_T < 110$ GeV and $|\eta| < 2.5$. It was retrained with Run-2 Monte Carlo simulated sample and validated with the Run-2 data collected in 2018.

There are two Monte Carlo (MC) simulated samples used in the Jet-Vertex-Tagger retraining. Both of the MC samples are generated with Powheg [82] and Pythia8 [83]. The MC sample used in the training is a di-jet sample generated with the parton distribution function set NNPDF23LO [84], which contains enough hard-scattering and pileup jets in each event. MC sample of $Z \rightarrow \mu\mu$ events is generated with the parton distribution function set CTEQ6L1 [85]. The $Z \rightarrow \mu\mu$ sample is used to test a set of selection cuts, which select a hard-scattering jet enriched region in Run-2 data collected in 2018. The selected events are used to validate the trained JVT models.

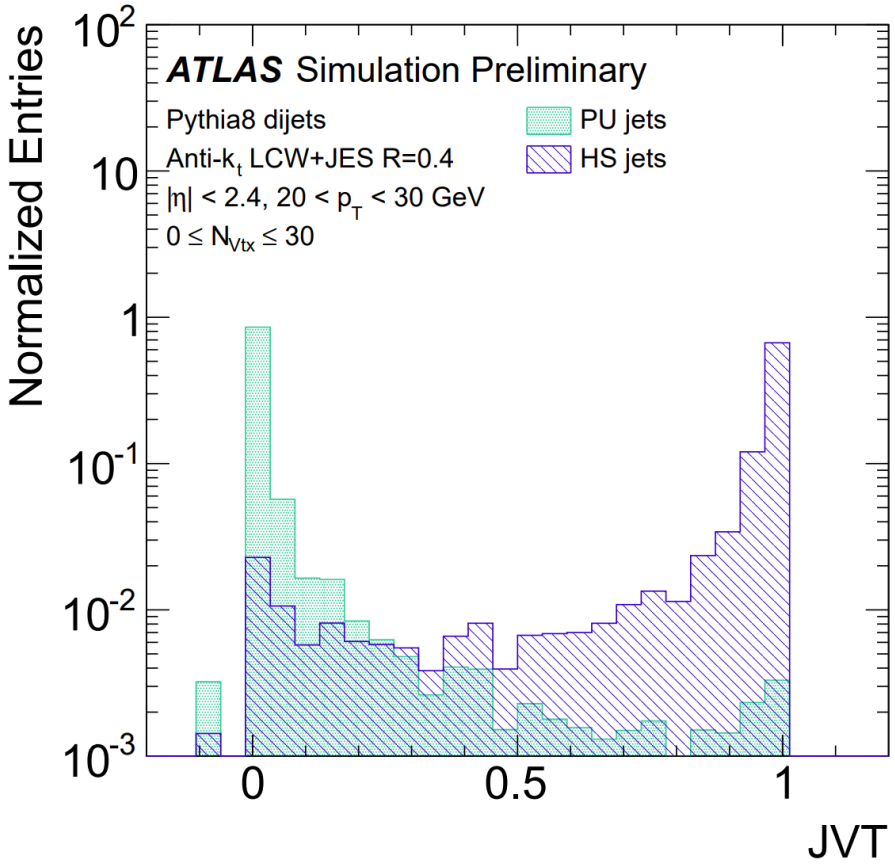


Figure 3.5: Distribution of JVT for pileup and hard-scatter jets with $20 < p_T < 30$ GeV in ATLAS Run-1 Monte Carlo simulated sample [79].

The MC samples are reconstructed with the ATLAS reconstruction packages, assigning a JVT score for all reconstructed jets in each event. This JVT score is used as a baseline of JVT retraining. The first model used in the JVT retraining was the KNN method in R_{p_T}-$corrJVF$ space. This model trained the JVT with 30000 hard-scattering jets as signal and 30000 pileup jets as background. Both signal and background weights were set to 1. This model is named KNN-V2 due to the two input variables R_{p_T} and $corrJVF$.

A new JVT model was developed using the KNN method, with four input variables. In additional to R_{p_T} and $corrJVF$, the new methods also use the transverse momentum p_T and pseudorapidity η. This method is called KNN-V4. Figures 3.6, 3.7, 3.8 and 3.9 shows the distributions of the four variables: $corrJVF$, R_{p_T}, η, and p_T, respectively.

The current JVT trained with the KNN method requires the storage of trained samples. When a sample is tested, the KNN method calculates the distance between the sample to all the points in the storage. The complexity increases with the

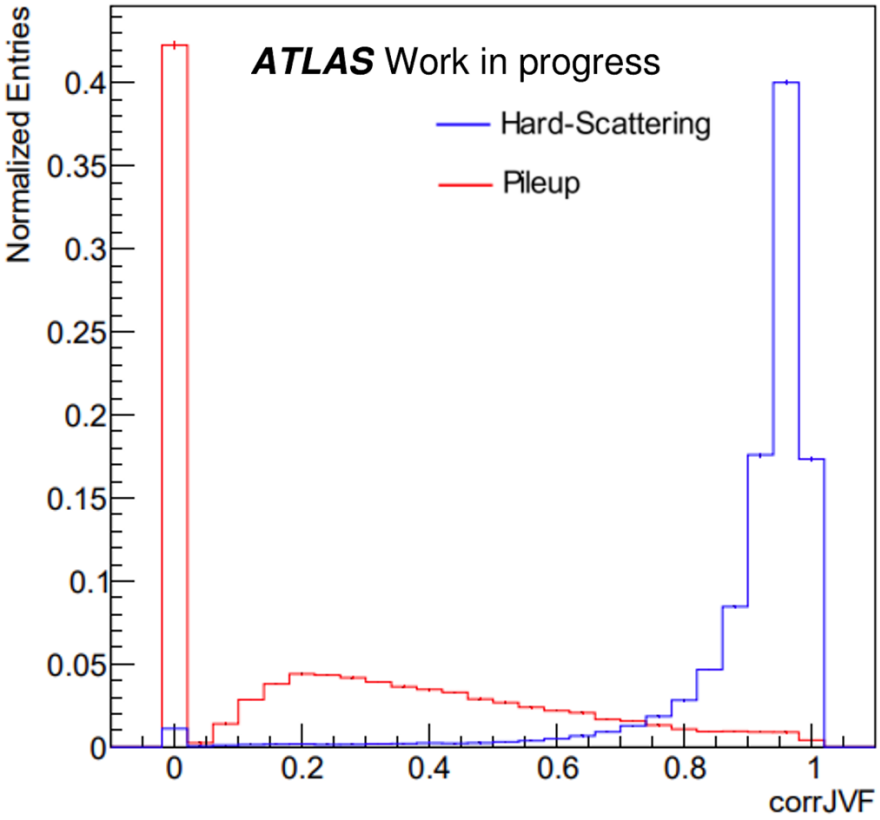


Figure 3.6: Normalized distributions of the variable $corrJVF$. The blue histogram is the distribution in the hard-scattering jets. The red histogram is the distribution in the pile-up jets.

size of the training dataset. The JVT is a common tool in the ATLAS analysis. The calculation is done for every dataset collected by ATLAS. An algorithm with constant complexity will improve the calculation speed in the Run-3 and future analysis.

Besides the KNN method, two new models based on neural network were designed. The KNN method requires the storage of trained samples. When a sample is tested, the KNN method calculates the distance between the sample to all the points in the storage. The complexity increases with the size of the training dataset. The neural network algorithms have a complexity independent on the size of the input dataset. In a neural network, the input variables of each event are converted to a vector. The vector passes through a calculation and returns a value or a vector as output. The calculation contains changeable parameters. After each calculation, a loss is calculated with the output of calculation. The parameters in the calculation is updated with the loss. The calculation expression, the loss and the way to update the parameter depend on how the neural network is designed. In the Run-2 JVT training, the neural network algorithms are applied.

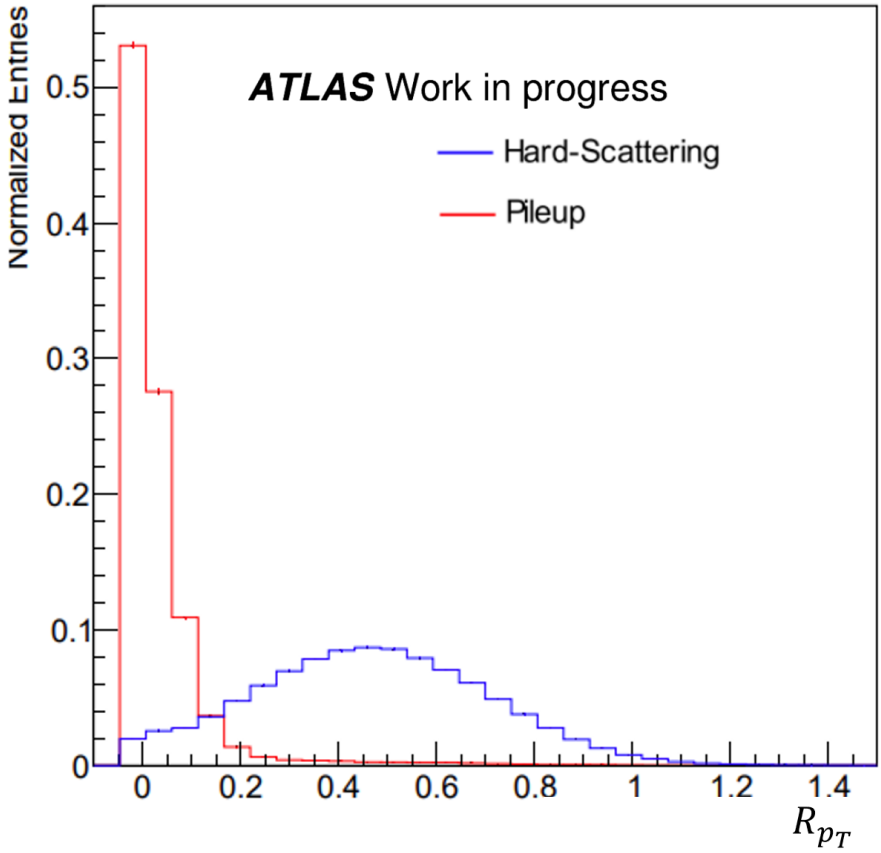


Figure 3.7: Normalized distributions of the variable R_{p_T}. The blue histogram is the distribution in the hard-scattering jets. The red histogram is the distribution in the pile-up jets.

The input of JVT is the R_{p_T}, $corrJVF$, p_T and η. It forms a dimension-4 vector. The calculation is designed as a combination of three matrices, three vectors and activation functions. Each matrix with size (n, m) correspond to a vector with dimension n. The vector is called a bias vector. The input vector with dimension-m is multiplied with an $n \times m$ matrix and generates an output vector with dimension-n. The output vector is added with the bias vector with the same dimension. Then each element in the result of the sum is put in an activation function. The activation function of the JVT training is selected to be $tanh(x)$. The activation function provides non-linear transformation and limits the value of the output to a fixed range. The processes described above are repeated in each layer of a neural network. There are totally three layers in the JVT training. The three matrices have sizes of $(14, 4)$, $(9, 14)$ and $(1, 9)$. There are also three corresponding bias vectors with dimensions 14, 9, and 1. All the values in the three matrices and bias vectors are changeable. These values are updated after each cycle, based on the output of the network.

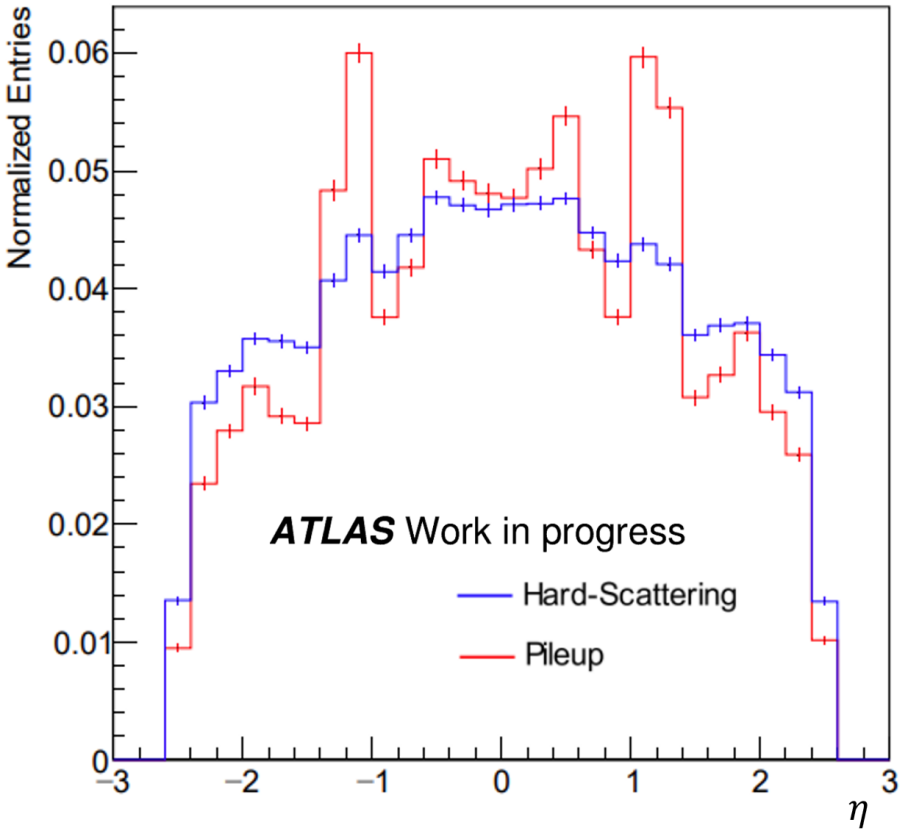


Figure 3.8: Normalized distributions of the variable η. The blue histogram is the distribution in the hard-scattering jets. The red histogram is the distribution in the pile-up jets.

The final output of the last layer returns a number. The $tanh()$ transform each element in the input vector to a value between -1 and 1. The result is linearly rescaled by the function $y = \frac{x+1}{2}$, where x is the output of the neural network and y is the output of this algorithm. The rescaling makes the output consistent with the previous JVT range from 0 to 1. The JVT is trained with supervised learning. Each jet in the training dataset is assigned a number. For hard-scattering jet, it is set to one. For pile-up jet, it is set to zero. The square of the difference between the returned value and the assigned number is used as the loss. A lower loss represents a better result. The gradient descent method is used to update the changeable values in matrices and vectors. The method calculates the gradient and updates the parameter with a negative gradient with a scale. The scale is called the learning rate. In this training it is set to 0.02. If the result of a model converges to a value, the loss reaches a local minimum with the gradient descent.

In the neural network designed for the JVT model, all the parameters in each matrix and bias vector are updated with the gradient descent method. This network is also called the multilayer perceptron (MLP) algorithm. The model trained with

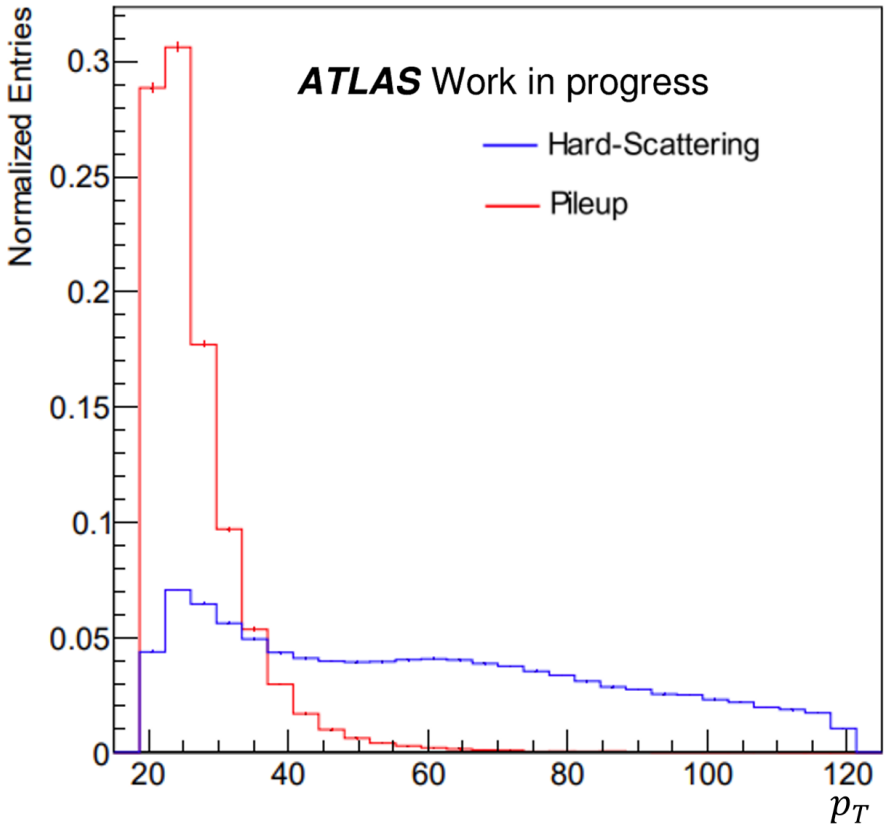


Figure 3.9: The normalized distributions of the variable p_T. The blue histogram is the distribution in the hard-scattering jets. The red histogram is the distribution in the pile-up jets.

R_{p_T}, $corrJVF$, p_T and η is named MLP-V4. And the one trained with R_{p_T} and $corrJVF$ is called MLP-V2. The MLP-V4 and MLP-V2 are trained using the same amount of signal and background samples with the KNN methods. These MLP algorithms take fewer storage resources than the JVT algorithm based on the KNN method. For the KNN-V4 method, the model requires a storage space for 240000 numbers, while the MLP-V4 model only requires a storage space for 215 numbers. During the calculation, the KNN methods require memory of the calculated distance. The MLP methods only require a memory no larger than the size of matrices. The MLP-V4 method has advantages in the storage and memory cost compared with the KNN methods. The storage and memory cost of KNN increases with the size of training dataset, while the MLP methods do not. The time cost in calculating the JVT with the KNN methods linearly increases with the size of the training dataset [86], while the neural network methods have a constant time cost when calculating the JVT score. The neural network method is promising in dealing with a large amount of training data.

The performance of a supervised distinguishing algorithm can be quantified by

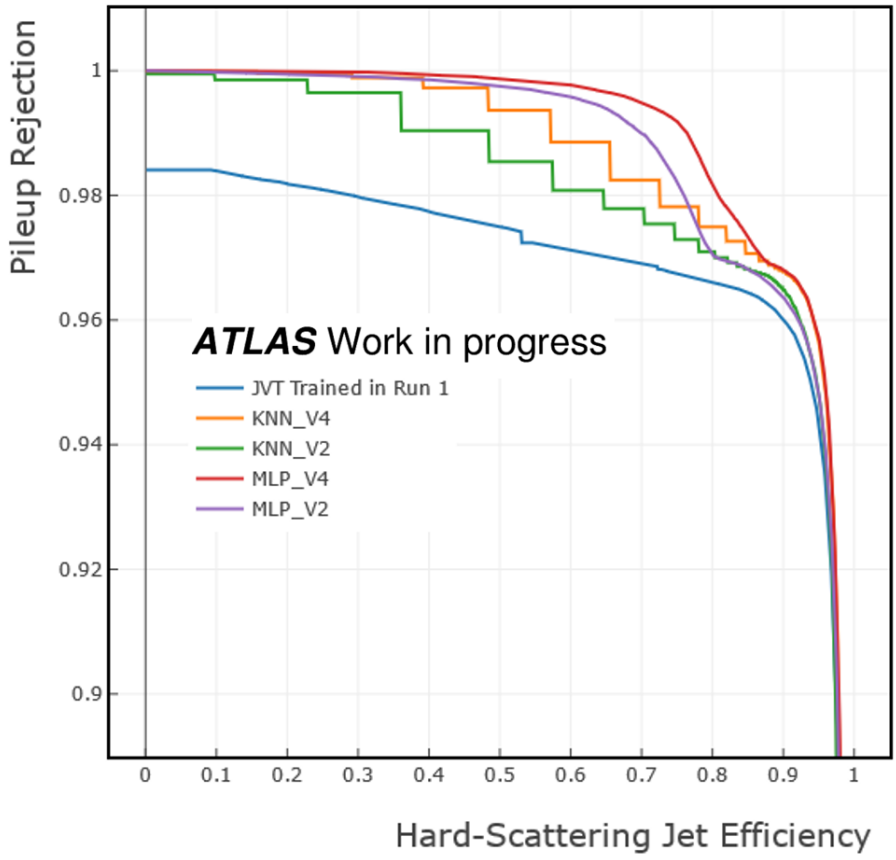


Figure 3.10: Receiver operating characteristic curve of all tested methods. The blue line is the JVT score provided by reconstruction packages. The others are the models trained in ATLAS Run-2 sample. The curve in red is for MLP-V4 method. It has the best pile-up rejection at almost all hard-scattering efficiency.

the Receiver Operating Characteristic (ROC) curve. It is a curve on the 2-D plane of the signal efficiency and background rejection rate. When the signal efficiency increases, the selection is looser, leading to a lower background rejection rate, and vice versa. The area under the ROC curve is called AUC, which is a score for judging the model. The AUC is 1 when it is an ideal distinguishing model, and it is around 0.5 when the model separates the two categories without any bias.

Among all the tested methods, the MLP-V4 shows the best performance, which provides a pile-up rejection of 0.957 at a hard-scattering efficiency of 0.95. The performance of all tested methods is shown in figure 3.10. The curve in red is for the MLP-V4 method. It has the best pile-up rejection in almost the whole range of the hard-scattering efficiency.

Selection cuts
$p_{TZ} > 30$ GeV
$\lvert p_{TZ} - p_{Tjet}\rvert < 5$ GeV
$\lvert\eta_Z + \eta_{jet}\rvert < 0.3$
$\lvert\Delta\phi(Z, jet)\rvert > 2.8$

Table 3.1: Selection cuts of Hard-scattering jet enriched region in the ATLAS Run-2 Monte Carlo simulated sample. The leading jet is the jet with the largest transverse momentum among the jets in an event. In the selection cut, the "jet" represents the leading jet in the event.

Jet Vertex Tagger validation

The Retrained JVT is validated within a hard-scattering enriched region. It uses the $Z \to \mu\mu + jets$ events from Run-2 data collected in 2018. The hard-scattering enriched region in $Z \to \mu\mu + jets$ sample is selected with the cut listed in table 3.1. Only the jets with the largest transverse momentum in an event are counted. The Z bosons are reconstructed with two μ in the final state. The transverse momentum of the Z boson is required to be larger than 30 GeV. The difference between the transverse momentum of the Z boson and the leading jet is required to be smaller than 5 GeV. The η and ϕ cuts guarantee that the Z and the leading jet are in the back-to-back position. This selection cuts correctly select the hard-scattering leading jet in 89.1% of the events in the Monte Carlo simulated $Z \to \mu\mu + jets$ testing sample. This method was used in the Run-1 JVT training, the selection cuts are adjusted to optimize the proportion of the hard-scattering leading jet [79].

The same selection cuts are applied to the ATLAS Run-2 $Z \to \mu\mu + jets$ data, in order to estimate the performance of JVT in handling the hard-scattering jets. Figure 3.11 and figure 3.12 show the normalized JVT distribution in the hard-scattering enriched data, Monte Carlo simulated hard-scattering jets, and pileup-jet samples. Both figures use the ATLAS Run-2 samples. Figure 3.11 uses the JVT score provided by the reconstruction packages [79]. Figure 3.12 is generated for the MLP-V4 method. In both figures, a reasonable agreement is observed between the distribution of JVT in the hard-scattering enriched data and the one in the Monte Carlo simulated hard-scattering jets. This result validates the effectiveness of the JVT in real data.

Working point of Jet-Vertex-Tagger

Three working points are defined in ATLAS analysis, named "loose", "medium" and "tight". These three working points correspond to the signal efficiency 97%,

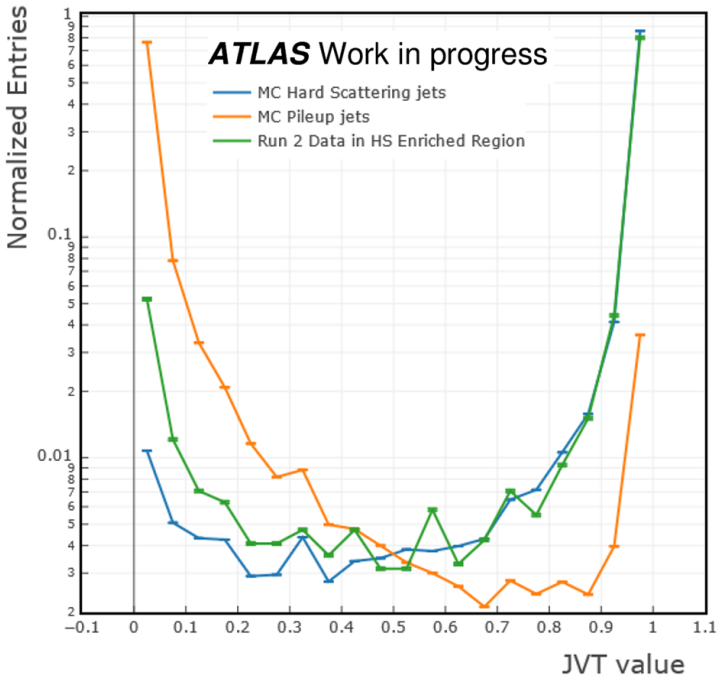


Figure 3.11: Normalized JVT distribution in hard-scattering enriched data, Monte Carlo simulated hard-scattering jet and pileup jet samples. The JVT is calculated using the JVT trained in ATLAS Run-1 samples. The histogram is drawn with log scale in y-axis.

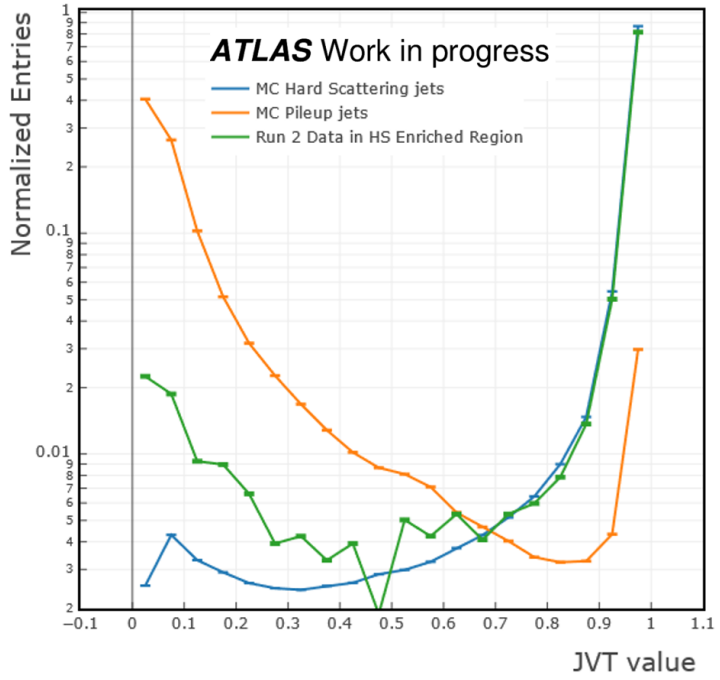


Figure 3.12: Normalized JVT distribution in hard-scattering enriched data, Monte Carlo simulated hard-scattering jet and pileup jet samples. The JVT is calculated with MLP_V4 method trained in ATLAS Run-2 samples. The histogram is drawn with log scale in y-axis.

Model	JVT loose SE = 0.97		JVT medium SE = 0.92		JVT tight SE = 0.85	
	WP	BR	WP	BR	WP	BR
JVT	0.295	0.915	0.828	0.956	0.955	0.964
KNN-V2	0.170	0.916	0.640	0.961	0.870	0.968
MLP-V4	0.525	0.931	0.904	0.966	0.953	0.973

Table 3.2: Work point of JVT in loose, medium and tight cut. The "SE" represent the signal efficiency. The "WP" represent the working point. The "BR" represent the background rejection. The model JVT represent the one used in the reconstruction packages. The KNN-V2 is the model use the input variables R_{p_T} and $corrJVF$. The MLP-V4 model is the one with best performance at all the three signal efficiencies, which is a neural network model with four input variables: R_{p_T}, $corrJVF$, p_T and η.

92% and 85%, respectively. Table 3.2 shows the working point values together with the corresponding pileup-jet rejection rates. Three JVT models are listed: the one currently used, the KNN-V2, and the MLP-V4. The MLP-V4 has slightly better background rejection at all three working points.

Chapter 4

CP-violation in Higgs boson decays to τ pairs

4.1 Decays of $\tau-$lepton

There is no CP-violation expected in the interaction between the Higgs boson and the other particles in the Standard Model. The SM Higgs boson is predicted to be purely CP-even. The Run-1 data excluded the pure CP-odd state at 99.98% confidence level [4]. CP-mixing is still possible in the interaction between the Higgs boson and the other particles.

One way to probe the CP nature of the Higgs boson coupling is to measure the spin correlations of the $\tau-$lepton pair from the decay of the Higgs boson. The $\tau-$lepton is the unique choice in the measurement. The branching ratio of Higgs decaying to $\mu-$lepton pair or electron-positron pair is much lower than decaying into $\tau-$lepton pair. The method of measuring the spin correlations between the lepton pair used in this thesis also requires the leptons decay in the tracking volume of detector, which is not applicable for the electrons and muons.

The $\tau-$leptons decay both leptonically and hadronically. Table 4.1 lists the fraction of the dominant decay processes of the $\tau-$lepton. Both leptonic decay and hadronic decay channels are used in the analysis. In the leptonic decay channels, the $\tau-$lepton decays into an electron or a muon, with two undetectable neutrinos. The hadronic decay of τ has one neutrino in the final state, as well as charged and neutral pions: $\pi^{\pm}$ and π^{0}. These decay channels are labeled with alphanumeric codes. The leptonic decay channels are labeled with ℓ. The hadronic decay channels are named as XpYn, where "X" is the number of charged pions and the "Y" is the number of neutral pions. For example, in the $H \to \tau\tau$ process, when one τ decays

Decay mode	Branching fraction	Notation
$\mu^-\bar{\nu}_\mu\nu_\tau$	17.4 %	ℓ
$e^-\bar{\nu}_e\nu_\tau$	17.8 %	
$h^-\nu_\tau$	11.5 %	1p0n
$(\pi^-\nu_\tau)$	(10.8 %)	
$h^-\pi^0\nu_\tau$	25.9 %	1p1n
$(\pi^-\pi^0\nu_\tau)$	(25.5 %)	
$h^- \geq 2\pi^0\nu_\tau$	10.8 %	1pXn
$(\pi^-2\pi^0\nu_\tau)$	(9.3 %)	
$3h^\pm\nu_\tau$	9.8 %	3p0n
$(3\pi^\pm\nu_\tau)$	(9.3 %)	

Table 4.1: Branching fractions of dominant leptonic and hadronic τ decay modes [17] and their notations used in the analysis. "$h^\pm$" includes $\pi^\pm$ and $K^\pm$. The hadronic decay is dominated by decays to $\pi^\pm$, which are shown in parenthesis.

into the leptonic channel and the other decay into one charged pion and a neutrino, it is called a lepton-hadron channel, or lephad. It is labeled as $\ell - 1p0n$.

Some BSM Models predicts CP-violation in the Yukawa interaction. The CP-violation can either be caused by CP-mixing angle or Yukawa coupling magnitude [87]. This study focuses on the measurement of CP-mixing angle. Some other BSM Models, such as Two Higgs Doublet Model (THDM), predicts CP-violation caused by Yukawa coupling magnitude. The THDM requires the Higgs field to be composed of two complex scalar field. Each complex scalar field acquires nonzero vaccum-expectation-value v_1 and v_2, where $v = \sqrt{v_1^2 + v_2^2} = 246$ GeV. Both scalar fields interacts with $\tau-$lepton pair, with different coupling constant. The CP-violation arises due to the combination of two Yukawa interactions [88]. Further analysis on the CP-violation caused by Yukawa coupling magnitude may be carried on if evidence of two Higgs scalar fields is detected.

In a model of CP-violation caused by CP-mixing angle, a general model-independent effective Yukawa interaction between the Higgs boson and $\tau-$lepton pair can be parametrized as [89]:

$$\mathcal{L}_{h\tau\tau} = -\frac{m_\tau}{v}\kappa_\tau(\cos\phi_\tau\bar{\tau}\tau + \sin\phi_\tau\bar{\tau}i\gamma_5\tau)h \quad , \tag{4.1}$$

where the vacuum expectation value of the Higgs field $v = 246$ GeV, $\kappa_\tau > 0$ is the reduced Yukawa coupling strength, and $\phi_\tau \in [-\pi/2, \pi/2]$ is the "CP-mixing" angle that parametrizes the relative contributions of the CP-even and CP-odd components to the $H\tau\tau$ coupling. The Standard Model hypothesis predicts a $\phi_\tau = 0$, corresponding to pure CP-even. The pure CP-odd hypothesis corresponds to $\phi_\tau = \pi/2$. Any other value of ϕ_τ corresponds to CP-mixing state.

The information of the scalar-pseudoscalar mixing angle ϕ_τ is encoded in the transverse spin correlation in the $H \to \tau\tau$ decay, which is independent of the Higgs production. The differential decay width in the Higgs rest frame can be written as [89, 90]:

$$d\Gamma_{h\tau\tau} \propto 1 - s_z^- s_z^+ + \cos(2\phi_\tau)(s_\perp^- \cdot s_\perp^+) + \sin(2\phi_\tau)[(s_\perp^- \times s_\perp^+) \cdot \hat{k}^-], \qquad (4.2)$$

where $\hat{k}^-$ is the normalised τ^- spatial momentum in the Higgs boson rest frame, $\hat{s}^\pm$ are the unit spin vectors of $\tau^\pm$ in their respective τ rest frames, and $s_\perp^\pm(s_z^\pm)$ are the transverse (longitudinal) components of $\hat{s}^\pm$ with respect to $\hat{k}^-$. The information of ϕ_τ can be obtained from the correlation between the transverse components of the spin of $\tau-$lepton.

Equation 4.2 shows that the longitudinal components of the spin correlation are insensitive to the CP-mixing angle, and that the sensitivity lies only in the transverse spin correlation of the di-τ system. The τ spin correlation is observable by measuring the angular distributions of the τ decay products. In the $\tau^- \to \pi^-\nu_\tau$ decay, the π^- momentum points preferably in the same direction as the τ spin in the τ^- rest frame. In the $\tau^+ \to \pi^+\bar{\nu}_\tau$ decay the π^+ momentum points preferably in the opposite direction of $\tau-$lepton spin in the τ^+ rest frame. The spin correlation can thus be measured by considering the angular correlations of the τ decay products.

This can be done by defining a "signed" acoplanarity angle $\varphi^*_{\mathrm{CP}} \in [0, 2\pi]$ as the angle between the τ decay planes spanned by the τ decay products in the zero-momentum frame (ZMF) of the visible di-τ decays. Figure 4.2 is an illustration of the observable φ^*_{CP}. All the four vectors in the calculation are boosted to the visible di-τ ZMF. The visible di-τ ZMF is the ZMF of all the detected particles in the final state of $H \to \tau\tau$. It is an approximation of the Higgs rest frame due to the absence of the reconstruction of neutrinos in the $\tau-$lepton decay products.

Figure 4.1 illustrate the Lorentz transformation from the lab frame to the di-τ ZMF. In a $H \to \tau\tau$ event, the Higgs boson may has an initial speed to a random direction in the lab frame. The di-τ ZMF is a frame moving at a constant relative speed to the lab frame. By boosting all the 4-vectors to the di-τ ZMF with Lorentz transformation, the analysis can be done in a frame with a rest Higgs boson.

If the $\tau-$leptons are from a scalar Higgs boson, the φ^*_{CP} peaks at π, while in the case of a pseudo-scalar Higgs boson, the φ^*_{CP} to peak at 0 and 2π. Figure 4.3 shows the normalized distribution of φ^*_{CP} in the CP-even and CP-odd $H \to \tau\tau$ decay and in the $Z \to \tau\tau$ process. The black long-dash dotted line shows a CP-mixing model with $\phi_\tau = -\frac{\pi}{4}$. Equation 4.3 shows the relationship between the ϕ_τ and the decay width for $H \to \tau\tau$ decay [89]:

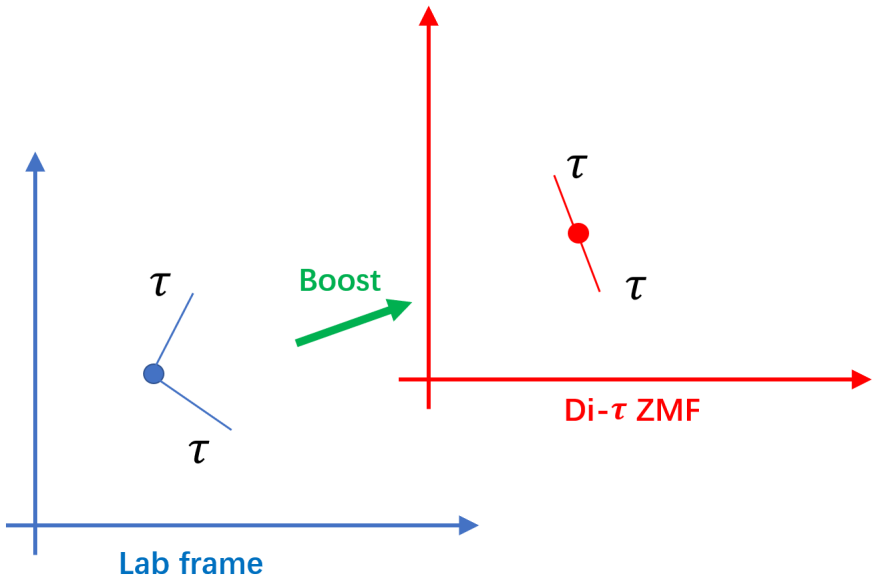


Figure 4.1: Illustration of the Lorentz transformation from the lab frame to the di-τ ZMF.

$$d\Gamma_{H\to\tau^+\tau^-} \approx 1 - b(E_+)b(E_-)\frac{\pi^2}{16}\cos(\varphi^*_{\mathrm{CP}} - 2\phi_\tau), \tag{4.3}$$

where the functions $b(E_\pm)$ express the spin analyzing power of a given decay mode, in which for direct decays $\tau^\pm \to \pi^\pm\nu_\tau$ and $\tau^\pm \to a^{L,T,\pm}\nu_\tau$ it equals $+1$. For other decays such as $\tau^\pm \to \rho^\pm\nu_\tau$, it is generally energy dependent.

An important feature of the φ^*_{CP} observable is that it follows a flat and uniform distribution for background processes, so that the presence of background does not bias the sensitivity to ϕ_τ. For example, the differential cross section of the $Z \to \tau\tau$ process in the $\tau^+\tau^-$ ZMF has no dependence of φ^*_{CP} in LO and NLO QCD. The red flat line in figure 4.3 represent the ϕ^* distribution of the $Z \to \tau\tau$ process [90]. In the experiment, the detector effects from reconstruction and analysis selection can introduce shape distortion [89, 90]. These effects may affect the experimental sensitivity of the CP-mixing angle. The methods used in this analysis to build the φ^*_{CP} observables with respect to the $\tau\tau$ decay modes are based on [89].

4.2 Methods to measure the CP-mixing angle

Impact parameter method

The impact parameter (IP) method can be applied to any decay modes that have a non-vanishing impact parameter. It is particularly useful in the decays where there is only one visible particle in the τ decay, such as the hadronic decay $\tau^\pm \to \pi^\pm\nu_\tau$

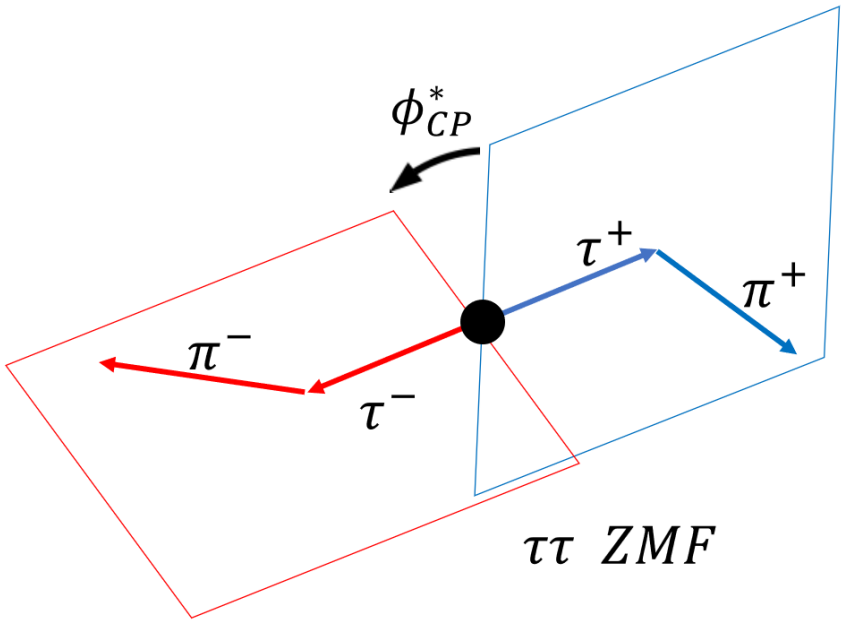


Figure 4.2: Illustration of a $H \to \tau^+\tau^- \to \pi^+\pi^- + 2\nu$ decay in the $\tau\tau$ Zero Momentum Frame. The decay planes are spanned by the momenta of the τ leptons and the pions. The angle acoplanarity angle $\varphi^*_{\mathrm{CP}} \in [0, 2\pi]$ is sensitive to the CP-mixing angle.

and leptonic decays $\tau^\pm \to \ell^\pm \nu_l \nu_\tau$. In these decays, the τ decay plane is formed from the spatial vector $\mathbf{q}^\pm$ of the charged particle ($\pi^\pm$, $\ell^\pm$) and the 3-dimensional impact parameter vector $\mathbf{n}^\pm$ of the charged particle [89, 91, 92, 93, 94].

Figure 4.4 shows how the impact parameter is defined ideally. The primary vertex (PV) is where the Higgs boson is produced. $\mathbf{q}^-$ denotes the visible charged particle in the final state. The position of the visible particle is called the secondary vertex. The impact parameter, which is labeled as $\mathbf{n}^\pm$, is defined as the vector that starts from the PV and points to the closest point to the track of the visible particle.

Experimentally the reconstruction precision of the secondary vertex of τ decay is poor. In this analysis, the impact parameter is constructed using the position parameters d_0 and z_0 of the charged particle. The d_0 is defined as the distance between the closest position of the reconstructed track to the beam line. The z_0 is the longitudinal value of that position. The impact parameter is calculated by assuming the secondary vertex located at ($d_0 cos\phi$, $d_0 sin\phi$, z_0), where ϕ is the azimuthal angle of the visible component.

The 4-momenta of the charged particle and the impact parameter are initially measured and defined in the laboratory frame. The impact parameter is expanded to a 4-vector $(0, x_{IP}, y_{IP}, z_{IP})$, so that it can be transformed in Lorentz transformation. They are boosted to the visible di-τ ZMF [89].

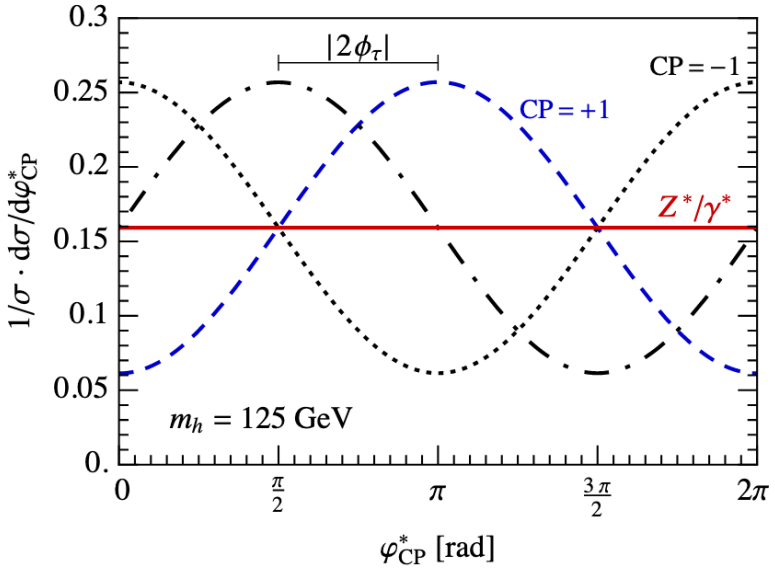


Figure 4.3: Normalized φ^*_{CP} distribution for $pp \to H/Z^*/\gamma^* \to \tau^+\tau^-$ production in the $\tau^+\tau^- \to \pi^+\pi^- + 2\nu$ decay [90]. The distribution for a CP-even (CP-odd) Higgs boson is shown by the blue dashed line (black dotted line). For a Higgs boson of CP mixture with $\phi_\tau = -\frac{\pi}{4}$ the distribution is given by the black long-dash dotted line. The distribution for the $Z \to \tau\tau$ process is shown by the solid red line, which is one of the main background of the φ^*_{CP} measurements.

The boosted impact parameter vector is labeled as $\widehat{\mathbf{n}}^{*\pm}$. It is used to calculate the angle between the two decay planes. Each of the decay planes is formed by the impact parameter and its corresponding charged particle $\widehat{\mathbf{q}}^{*\pm}$. The angle φ^* is expressed as:

$$\varphi^* = \arccos(\widehat{\mathbf{n}}^{*+}_{\perp} \cdot \widehat{\mathbf{n}}^{*-}_{\perp}), \tag{4.4}$$

where $\widehat{\mathbf{n}}^{*\pm}_{\perp}$ represents the component of the impact parameter that is verticle to the direction of the charged particle $\mathbf{q}^{\pm}$.

The CP-sensitive observable φ^*_{CP} ($0 \leq \varphi^*_{\mathrm{CP}} \leq 2\pi$) is defined by:

$$\varphi^*_{\mathrm{CP}} = \begin{cases} \varphi^* & \text{if } \mathcal{O}^*_{\mathrm{CP}} \geq 0; \\ 2\pi - \varphi^* & \text{if } \mathcal{O}^*_{\mathrm{CP}} < 0\,, \end{cases} \tag{4.5}$$

where the $\mathcal{O}^*_{CP}$ is defined as:

$$\mathcal{O}^*_{CP} = \widehat{\mathbf{q}}^{*-} \cdot (\widehat{\mathbf{n}}^{*+}_{\perp} \times \widehat{\mathbf{n}}^{*-}_{\perp}). \tag{4.6}$$

Figure 4.5 illustrates the φ^*_{CP} construction used in the IP method.

The performance of the IP-method is related to the impact parameter vector size and its uncertainty. A value called d_0^{sig} is defined as the transverse track impact

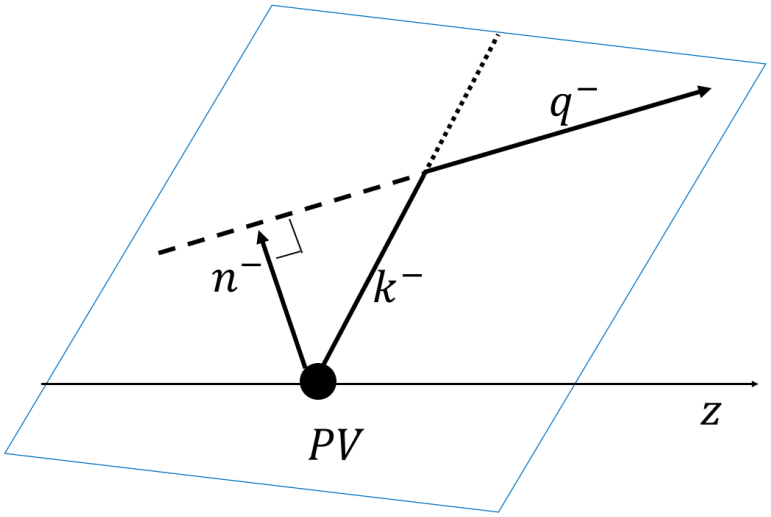


Figure 4.4: Definition of the impact parameter vector $\mathbf{n}^{\pm}$ in the plane of the decay $\tau^{\pm} \to \pi^{\pm}\nu_{\tau}$ or $\tau^{\pm} \to \ell^{\pm}\nu_l\nu_{\tau}$ in the laboratory frame [91]. Here, $\mathbf{q}^-$ is the measured charged particle momentum, PV is the τ production vertex, and $\mathbf{k}^-$ is the 3-momentum of the τ^-.

parameter d_0 divided by its uncertainty:

$$d_0^{sig} = \frac{d_0}{\sigma_{d_0}}. \tag{4.7}$$

The IP-method works better in the τ decay products with higher d_0^{sig}. In order to maximize the CP sensitivity of φ^*_{CP}, the events are divided into two regions called High and Low. In IP-method, The High and Low regions are separated by a cut on the $|d_0^{sig}|$.

ρ method

In the case of indirect decays $\tau^{\pm} \to \rho^{\pm}\nu$, with $\rho^{\pm} \to \pi^{\pm}\pi^0$, the τ decay plane can be formed from the spatial vectors of the charged components and neutral components. The 4-momentum vectors $\mathbf{q}^{*\pm}$ denotes the charged components. The $\mathbf{q}^{*0\pm}$ represent the neutral components from $\rho^{\pm}$, respectively. All these vectors are boosted to the visible di-τ ZMF. Using the normalized spatial vectors $\widehat{\mathbf{q}}^{*\pm}$ and $\widehat{\mathbf{q}}^{*\mathbf{0}\pm}$, the angle φ^* is expressed as:

$$\varphi^* = \arccos(\widehat{\mathbf{q}}^{*0+}_{\perp} \cdot \widehat{\mathbf{q}}^{*0-}_{\perp}), \tag{4.8}$$

where the $\widehat{\mathbf{q}}^{*0\pm}_{\perp}$ represent the verticle component of the neutral pion. The variable $\mathcal{O}^*_{CP}$ is defined as:

$$\mathcal{O}^*_{\mathrm{CP}} = \widehat{\mathbf{q}}^{*-} \cdot (\widehat{\mathbf{q}}^{*0+}_{\perp} \times \widehat{\mathbf{q}}^{*0-}_{\perp}). \tag{4.9}$$

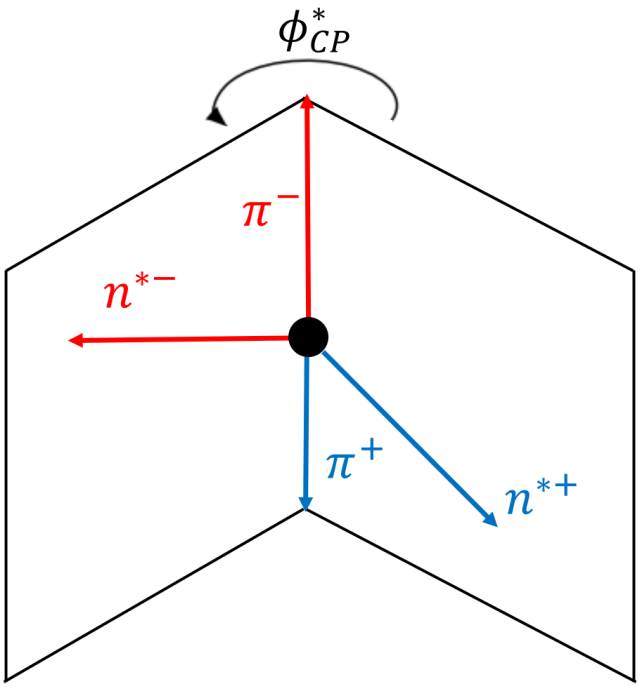


Figure 4.5: Definitions of 3-vectors and φ^*_{CP} in the visible di-τ ZMF used in the impact parameter method [89].

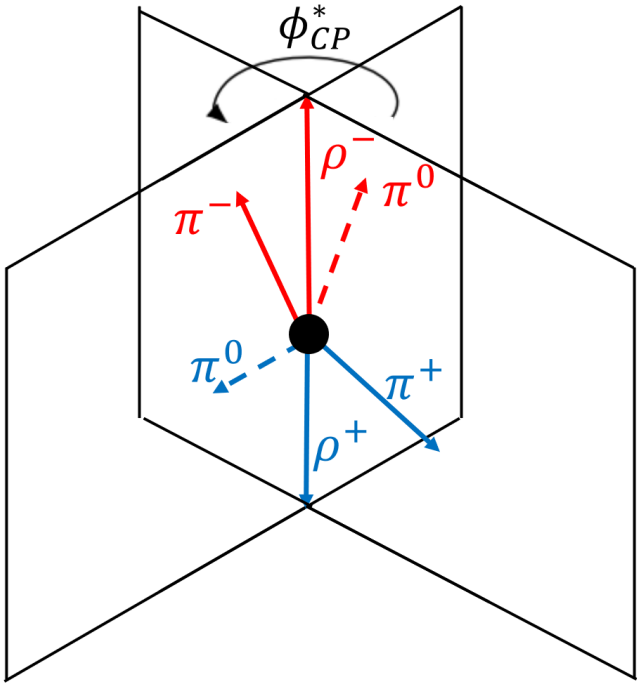


Figure 4.6: Definitions of 3-vectors and ϕ^* in the visible di-τ ZMF used in the ρ method [89].

The angle $\varphi^{*\prime}$ is defined as:

$$\varphi^{*\prime} = \begin{cases} \varphi^{*} & \text{if } \mathcal{O}^{*}_{\mathrm{CP}} \geq 0; \\ 2\pi - \varphi^{*} & \text{if } \mathcal{O}^{*}_{\mathrm{CP}} < 0. \end{cases} \tag{4.10}$$

The distribution of $\varphi^{*\prime}$ is not sensitive to ϕ_τ. The sensitivity appears when the $\varphi^{*\prime}$ is corrected using the variables y_- and y_+. The variables y_- and y_+ are defined as:

$$y_- = \frac{E_{\pi^-} - E_{\pi^0}}{E_{\pi^-} + E_{\pi^0}}, \qquad y_+ = \frac{E_{\pi^+} - E_{\pi^0}}{E_{\pi^+} + E_{\pi^0}}. \tag{4.11}$$

The $E_{\pi^{\pm,0}}$ are the energies of the pions measured in the lab frame. The variable y_- is calculated for the decay products of the τ^-. The y_+ corresponds to the decay products of the τ^-. The $\varphi^{*}_{\mathrm{CP}}$ is shifted by a π when the product $y_+ y_- < 0$:

$$\varphi^{*}_{\mathrm{CP}} = \begin{cases} \varphi^{*\prime} & \text{if } y_+ y_- \geq 0; \\ \varphi^{*\prime} + \pi & \text{if } y_+ y_- < 0. \end{cases} \tag{4.12}$$

Figure 4.6 illustrates the definition of the 3-vectors and $\varphi^{*}_{\mathrm{CP}}$ in the visible di-τ ZMF used in the ρ method.

In the analysis, the events with the larger absolute values of the upsilon product $y_+ y_-$ are observed to have higher sensitivity to ϕ_τ. The value of $|y_+ y_-|$ is used to select events into different categories. This method is applied to 1p1n-1p1n and 1p1n-1pXn decay channels. In the case of 1pXn decay, the vector of neutral component is defined as the sum of the 4-momenta of all neutral pions.

IP-ρ method

In the case of two τ leptons decaying into different decay channels, where one has neutral components and the other does not, a combined method can be applied. The combined method is called IP-ρ method. For the decay $\tau^- \to \pi^- \nu_\tau$, the τ decay plane is spanned by the normalised spatial vector $\widehat{\mathbf{q}}^{*-}$ and the transverse component of the impact parameter vector, $\widehat{\mathbf{n}}^{*-}_{\perp}$. For the $\rho^+ \to \pi^+ \pi^0$, the decay plane is calculated as in the ρ-method. All the 4-vectors are boosted to the visible di-τ ZMF. The angle φ^* and the correlation variable $\mathcal{O}^{*}_{CP}$ are defined by:

$$\varphi^{*} = \arccos(\widehat{\mathbf{q}}^{*0+}_{\perp} \cdot \widehat{\mathbf{n}}^{*-}_{\perp}), \qquad \mathcal{O}^{*}_{\mathrm{CP}} = \widehat{\mathbf{q}}^{*-} \cdot (\widehat{\mathbf{q}}^{*+}_{\perp} \times \widehat{\mathbf{n}}^{*-}_{\perp}), \tag{4.13}$$

that can be combined to:

$$\varphi^{*\prime} = \begin{cases} \varphi^{*\prime} & \text{if } \mathcal{O}^{*}_{CP} \geq 0 \ ; \\ 2\pi - \varphi^{*\prime} & \text{if } \mathcal{O}^{*}_{CP} < 0 \ . \end{cases} \tag{4.14}$$

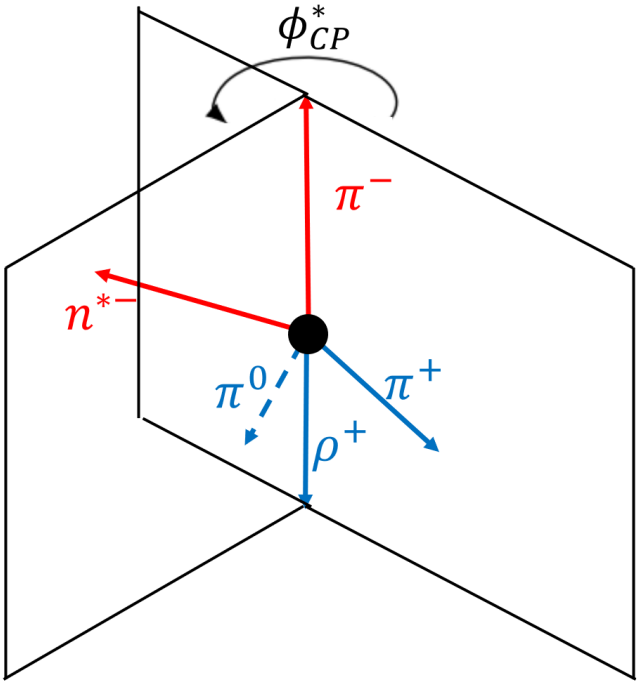


Figure 4.7: Definitions of 3-vectors and φ^*_{CP} in the visible di-τ ZMF used in the IP-ρ method [89].

Using the sign of y_+, defined in Equation 4.11 and computed in the laboratory frame, the observable φ^*_{CP} is defined as:

$$\varphi^*_{\mathrm{CP}} = \begin{cases} \varphi^{*\prime} & \text{if } y_+ \geq 0 \ ; \\ \varphi^{*\prime} + \pi & \text{if } y_+ < 0 \ . \end{cases} \tag{4.15}$$

Figure 4.7 illustrates the definition of the 3-vectors and φ^*_{CP} in the visible di-τ ZMF used in the IP-ρ method.

a_1-method for 3-prong decays

The a_1-method is used in the case of 3π decays with $\tau \to a_1\nu_\tau$, where $a_1^\pm \to \rho^0\pi^\pm$ and $\rho^0 \to \pi^+\pi^-$. The a_1-method is an extension of the ρ-method to 3-prong decays. This method selects a pair of charged π's from the 3-prong decay to form an intermediate ρ_0 particle. Together with the remaining charged π, a φ^*_{CP} observable can be constructed as done in the ρ-method. The charged and neutral components $\pi^\pm$ and π^0 determine respectively the spatial vectors; and here in the a_1-method, ρ^0 is taken as the neutral components.

Some ambiguities exist in this method as compared to the ρ-method. First, the choice of $\pi^\pm$ pair to form the intermediate ρ^0 meson is somewhat arbitrary, attributed

Method	Channel	Charged component	Neutral component
ρ	1p1n	$\pi^{\pm}$	π^0
a_1	3p0n	$\pi^{\pm}$ (the highest pT track)	$\pi^{\pm}\pi^{\pm}$ pair (the remaining tracks)

Table 4.2: Components used as spatial vectors in φ^*_{CP} construction in ρ and a_1 methods, respectively, in the corresponding decay modes.

to the broad resonance peak of ρ^0, especially at the reconstructed objects level. Second, the variable y used in the ρ-method needs to be re-defined to account for the mass effects. In the 1p1n case, the mass terms are negligible due to the small difference of $\pi^{\pm}$ and π^0 masses and are therefore dropped from the original definition in practice. In the a_1 case, the mass difference between $\pi^{\pm}$ and the duplet is no longer negligible. The associated mass terms in the upsilons have a more significant effect to the CP-sensitivity of the constructed φ^*_{CP}. The variable y_{a1} is defined as [95]:

$$y^{\pm}_{a_1} = \frac{E_{\rho^0} - E_{\pi^{\pm}}}{E_{\rho^0} + E_{\pi^{\pm}}} - \frac{m^2_{a_1} - m^2_{\pi^{\pm}} + m^2_{\rho^0}}{2m^2_{a_1}}.$$

Two charged π tracks with the lowest p_T are chosen to form a duplet, without requiring opposite charges. The duplet is treated as a particle with 4-momenta equals to the sum of 4-momenta of the two selected πs. The duplet is taken to replace the neutral component in the ρ method, while the highest-p_T charged π is taken as the charged component. The choice exhibits a close-to-optimal CP-sensitivity in the constructed φ^*_{CP} observable among other choices using all three tracks from the a_1 cascade decays.

Another possibility, explored in Ref [95], is to use only a pair of tracks forming the intermediate ρ^0. This method requires a pair of pions with opposite charges. The pion with the same charge to the τ it decays from is chosen as the charged component and the oppositely charged one replaces the neutral component of the ρ-method. This method provides a similar CP-sensitivity to the currently chosen method. Besides the methods metioned above, attempts are made on constructing CP-sensitivity observable for the $\tau \to a_1\nu_\tau$ decay with neural network.

In this analysis, only the 3p0n decay are considered. Only the 1p1n-3p0n and 3p0n-1p1n (ℓ-3p0n) events are considered in the $\tau_{had}\tau_{had}$ ($\tau_{lep}\tau_{had}$) channel. The 3p decay with more than one π^0 (3pXn) and 3p-3p events have much lower statistics and exhibit very small CP-sensitivity in the corresponding φ^*_{CP} observables, as shown in [95]. The 3p0n High and Low regions are defined with $|y_{a_1}|$ and another variable d_0^{sig} or $|y|$, depending on the decay mode of the other $\tau-$lepton decay.

4.3 Investigation of 3-prong CP observables

The 3-prong observable is defined as an extension to the ρ-method. The IP-method requires an impact parameter calculated from the primary vertex position, transverse position of the secondary vertex, and the direction of $\pi^{\pm}$. In the 3-prong decay of $\tau-$lepton, there are three $\pi^{\pm}$ detected. Each of these $\pi^{\pm}$ has an impact parameter pointing to a different direction. The neutral component can hardly be defined from these impact parameters in the 3-prong decay.

The a_1-method is described in section 4.2. The track with the highest p_T among the three tracks is selected as the charged component. The sum of the rest two tracks is used as the "neutral" component. In the 1-prong ρ-method, the "neutral" component is always neutral. In 3-prong decay, the decay plane is formed by a $\pi^{\pm}$ and the other component $\hat{k}$, which is selected manually. In some selection strategies, $\hat{k}$ may not be neutral. To distinguish it from a real neutral component in the 1-prong channels, it is named the "fork component" in the rest of this section.

Besides the a_1-method, there are many other methods to define the CP-sensitive observable for 3-prong decay. Two types of observables are used in the investigation of the 3-prong CP observables. For the type-1 observables, one $\pi^{\pm}$ is selected as the charged component, and the other two $\pi^{\pm}$ are selected as the fork component. The decay plane is defined by the charged component and the fork component. For the type-2 observables, one of the $\pi^{\pm}$ with the same charge to the $\tau-$lepton is selected as the charged component, and the other two $\pi^{\pm}$ are selected as the fork component. The decay plane is defined by the two opposite charged $\pi^{\pm}$ in the fork component.

The observables of type 1 do not require neutral fork components. There are three types of charge requirements of the fork component: no-charge-requirement, requiring the opposite charge $\pi^{\pm}$, and requiring the same charge $\pi^{\pm}$. The a_1-method is an example of no charge requirement. For the observables with no-charge-requirement or those requiring opposite charge, the fork components are selected according to the p_T of tracks or the invariant mass of the fork component. The invariant mass of the fork component is compared with the $m_{\rho} = 775$ MeV. For each $H \to \tau\tau$ event, there is only one choice of selecting the same charge $\pi^{\pm}$ as the fork component.

The observables constructed using a ρ-method may be shifted by π depending on the sign of a variable y. Two choice of y are used for each tested observable: $y^{\pm}_{a_1}$, and $y^{\pm}_{\rho-like}$ defined as follows:

$$y^{\pm}_{a_1} = \frac{E_{\rho^0} - E_{\pi^{\pm}}}{E_{\rho^0} + E_{\pi^{\pm}}} - \frac{m^2_{a_1} - m^2_{\pi^{\pm}} + m^2_{\rho^0}}{2m^2_{a_1}},$$

fork component	y	D_{KL} ($\times$10000)
Two $\pi^{\pm}$ with the lowest p_T	y_{a1}	69.35
Two $\pi^{\pm}$ with invariant mass farther from 775 MeV	$y_{\rho-like}$	58.16
Two same charged $\pi^{\pm}$	$y_{\rho-like}$	40.25
Opposite charged $\pi^{\pm}$ with lower invariant mass	$y_{\rho-like}$	28.72
Opposite charged $\pi^{\pm}$ with invariant mass farther from 775 MeV	$y_{\rho-like}$	24.66
Two $\pi^{\pm}$ with the largest p_T	y_{a1}	19.92
Two same charged $\pi^{\pm}$	y_{a1}	19.26
Two $\pi^{\pm}$ with the invariant mass farthest from 775 MeV	$y_{\rho-like}$	15.85
Opposite charged $\pi^{\pm}$ with invariant mass farther from 775 MeV	y_{a1}	14.95
Two $\pi^{\pm}$ with the lowest p_T	$y_{\rho-like}$	9.563

Table 4.3: Top 10 tested type-1 observables, ranking with the KL-divergence between CP-even and CP-odd distribution of the proposed observable.

$$y^{\pm}_{\rho-like} = \frac{E_{\pi^{\pm}} - E_{\rho^0}}{E_{\pi^{\pm}} + E_{\rho^0}}. \tag{4.16}$$

One of the ways to compare the effectiveness of these observables is to calculate the Kullback–Leibler (KL) divergence [96] between the normalized distribution of observables that reweighted to the CP-even and CP-odd states. The KL divergence is a method that quantifies how close a probability distribution is to another. The following equation shows how KL divergence is calculated in two discrete distributions p and q, where p_i and q_i are the corresponding probabilities in the i-th bin:

$$D_{KL}(p||q) = \sum_i p_i log\left(\frac{p_i}{q_i}\right). \tag{4.17}$$

The KL divergence $D_{KL}(p||q)$ is a non-negative number. The KL divergence is equal to zero only when $p(x) = q(x)$. A lower KL divergence value represents a smaller difference between the two distributions. All the tested observables are measured in the visible di-τ ZMF.

Table 4.3 lists the top 10 tested observables, ranking with the KL divergence between the reweighted CP-even and CP-odd distributions. The distributions of the observables are estimated using the Monte Carlo simulated sample. The plots of the distributions of these observables are shown in appendix A.

Among all the observables constructed using type 1 selection strategy, the observable selecting two $\pi^{\pm}$ with the lowest p_T and corrected by y_{a1} has the highest D_{KL} score. This observable is the one used in the a_1-method. Figure 4.8 shows the distribution of the observable constructed with the a_1-method.

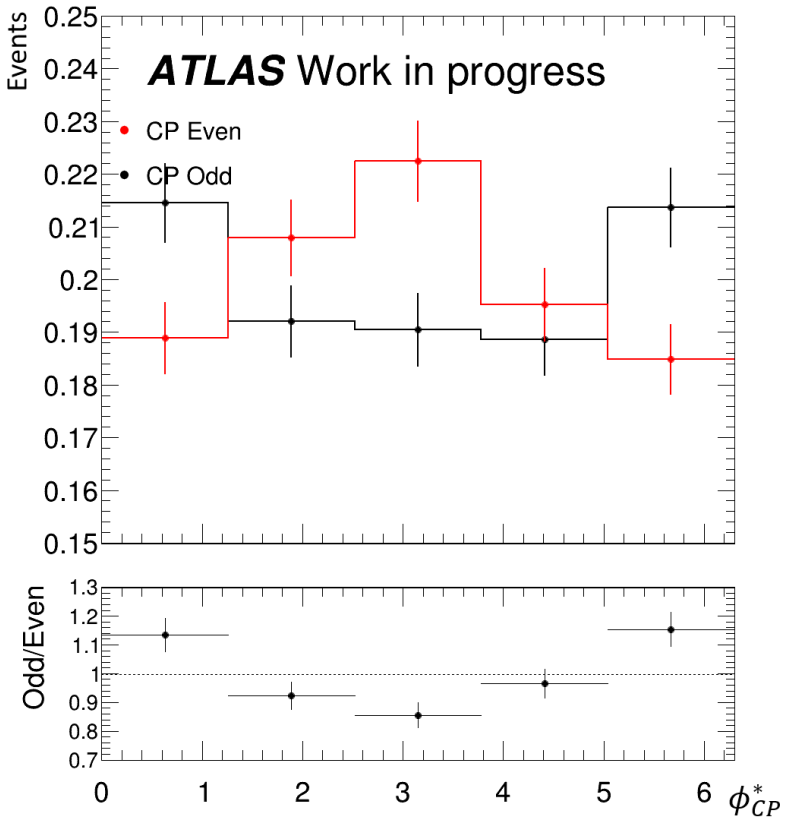


Figure 4.8: CP-sensitive observable candidate using type-1 method. Two $\pi^{\pm}$ with the lowest p_T are selected to form the fork component.

The observable using the same charges $\pi^{\pm}$ to form the fork component also performs well. Figure 4.9 shows the distribution of the observable constructed with this method.

The type-2 method is an alternative method of calculating the CP-sensitive observable. Instead of using the fork component and the other $\pi^{\pm}$ to construct the decay plane, type-2 methods select the fork component and construct the decay plane with the two $\pi^{\pm}$s in the fork component. This method requires two $\pi^{\pm}$s with opposite charges, otherwise, the orientation of the decay plane cannot be determined.

The calculation of type-2 observables also uses the ρ-method. The neutral component is replaced by the $\pi^{\pm}$ with the opposite charge to the τ it decays from. The variable $y_{\rho\rho}$ that shifts the observable is defined as:

$$y^{\pm}_{\rho\rho} = \frac{E_{\pi^{\pm}} - E_{\pi^{\mp}}}{E_{\pi^{\pm}} + E_{\pi^{\mp}}}.$$

There are 6 observables constructed with type-2 method. Table 4.4 lists these

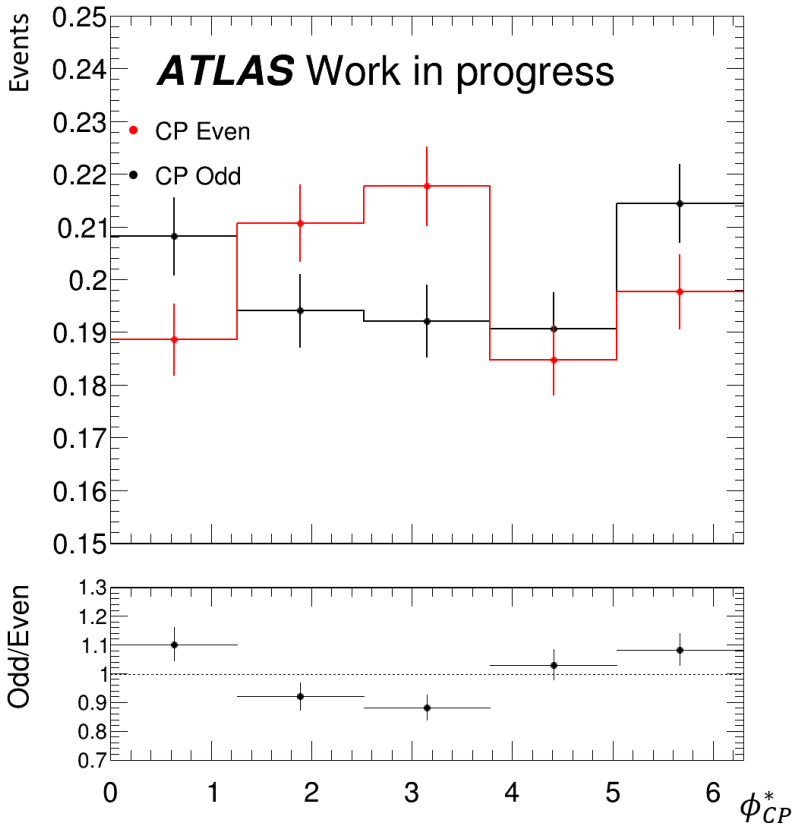


Figure 4.9: Calculated phistar candidate using type-1 method. Two $\pi^{\pm}$ with the same charge are selected to form the fork component.

fork component	y	D_{KL} ($\times$10000)
The $\pi^{\pm}$ pair with higher p_T	$y_{\rho\rho}$	68.28
The $\pi^{\pm}$ pair with invariant mass closer to 775 MeV	$y_{\rho\rho}$	63.03
The $\pi^{\pm}$ pair with lower invariant mass	$y_{\rho\rho}$	51.94
The $\pi^{\pm}$ pair with higher invariant mass	$y_{\rho\rho}$	31.47
The $\pi^{\pm}$ pair with invariant mass farther from	$y_{\rho\rho}$	21.99
The $\pi^{\pm}$ pair with lower p_T	$y_{\rho\rho}$	20.63

Table 4.4: Top 6 tested observables using the type-2 selection strategy, ranking with the KL-divergence between the CP-even and CP-odd distribution of the proposed observable.

observables and ranks them with the KL divergence. The plots of the distributions of these observables are shown in appendix A.

Figure 4.10 shows the distribution of the best performing observable constructed with the type-2 selection strategy. The 3-prong decayed $\tau-$lepton contains three $\pi^{\pm}$s, there are two same charged $\pi^{\pm}$s and one with the opposite charge. From the two same charged $\pi^{\pm}$s, the one with higher p_T is selected to form the fork component

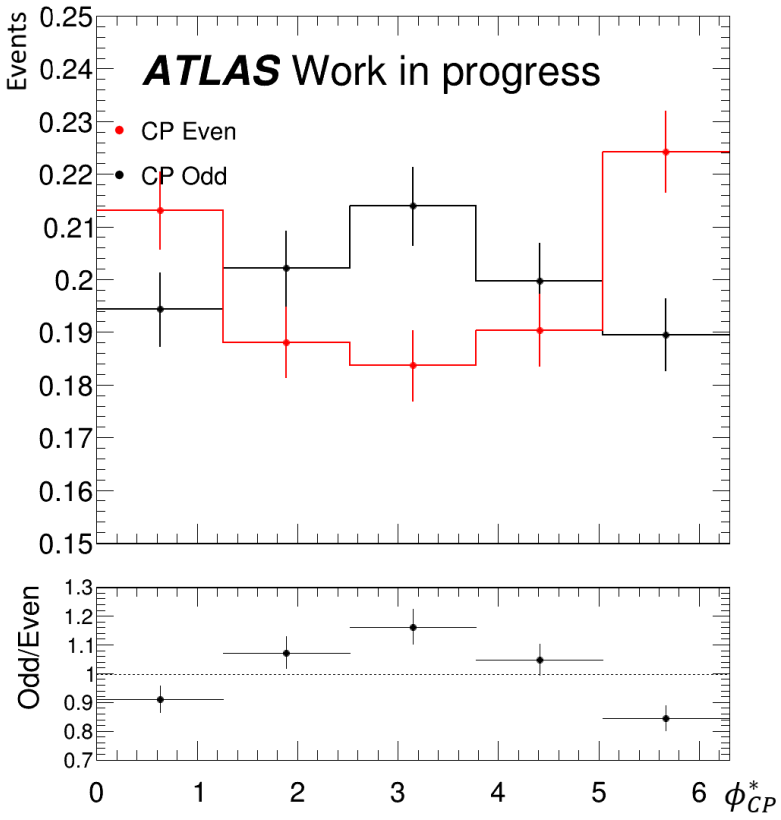


Figure 4.10: Observable constructed with the type-2 strategy. From the two same charged tracks, the one with higher p_T is selected to form the fork component.

with the opposite-charged $\pi^{\pm}$.

Chapter 5

Measurement of CP properties in the $H \to \tau\tau$ coupling

5.1 Experimental data and Monte Carlo simulation

Data

This analysis is performed using ATLAS Run-2 data collected from 2015 to 2018. Table 5.1 lists the integrated luminosity recorded in each year during Run-2.

Monte Carlo simulation

There are four main Higgs production processes that dominate the Higgs boson production at the LHC: gluon-gluon fusion (ggH), vector boson fusion (VBF), associated VH and ttH production processes. The expected cross sections of these Higgs production processes are listed in table 5.2. The Powheg [82] and Pythia [83] programs are used for event generation and modeling of parton showering, respectively. Table 5.3 lists the packages used in the simulation of each production process. The "ME" column gives the generator for calculating the matrix elements. The "PS" column shows the package for generating parton showering [98]. The ggH process is simulated at the next-to-next-to-leading-order (NNLO) accuracy [99, 100], with a normalization to the cross section calculated in the next-to-next-to-next-to-leading-order (N^3LO) QCD and NLO electroweak corrections. The VBF and VH processes are simulated to next-to-leading order (NLO) accuracy, with a normalization to the cross section calculated in the NNLO QCD and NLO electroweak corrections [101, 102, 103, 104, 105, 106, 107, 108, 109, 110]. The ttH process is generated at NLO accuracy.

Year	$\int \mathcal{L}\mathrm{d}t$ (fb^{-1})
2015	3.22
2016	32.99
2017	44.31
2018	58.45

Table 5.1: Integrated luminosity in each year during the ATLAS Run-2 [97].

Higgs production process	Expected cross section(pb)
ggF	$48.5^{+2.2}_{3.3}$
VBF	$3.78^{+0.08}_{-0.08}$
VH	$2.26^{+0.03}_{-0.03}$
$t\bar{t}H$	$0.51^{+0.03}_{-0.05}$

Table 5.2: Expected cross section for the four main Higgs production processes at $m_H = 125.09$ GeV [41].

The "PDF set" column lists the parton distribution function packages used in generating parton showering. The parton distribution function represents the probability density to find a parton carrying a momentum fraction at a certain energy scale. The "PDF4LHC15" [111], "CTEQ6L1" [85] and "NNPDF2.3" [112] are parton distribution function sets calculated in previous researches.

A "tune" is a specific modification for a Monte Carlo generator. The Monte Carlo simulation models depends on several free parameters. The simulation requires optimization so that the result produced matches the ATLAS experimental result [113]. The tune packages are developed for different PDF sets accordingly. The "AZNLO" [114] and "A14" [113] are the tune packages developed for the ATLAS Monte Carlo simulations.

The spin correlation of $\tau-$lepton is simulated with the TauSpinner package in all simulated signal events. TauSpinner is a tool to generate different CP hypotheses from generated Higgs [115, 116, 117]. For each event, TauSpinner generates a weight for each CP hypothesis. By using the truth-level kinematics information, it can add spin-effect. The simulated Monte Carlo $H \to \tau\tau$ with TauSpinner contains the kinematics of the particles, as well as TauSpinner weight for different CP hypotheses. By applying these weights to events while generating a distribution, it will give a reweighted distribution for a certain CP hypothesis. Figure 5.1 and figure 5.2 demonstrate the validation of the TauSpinner package. Three samples are generated with Pythia8 [118], corresponding to the CP-even, CP-odd, and unpolarised samples, respectively. The unpolarized sample does not have spin-effect.

Process	Generator		PDF set		Tune	Order
	ME	PS	ME	PS		
$H \to \tau\tau$						
ggF	Powheg	Pythia8	PDF4LHC15	CTEQ6L1	AZNLO	NNLO
VBF	Powheg	Pythia8	PDF4LHC15	CTEQ6L1	AZNLO	NLO
VH	Powheg	Pythia8	PDF4LHC15	CTEQ6L1	AZNLO	NLO
$t\bar{t}H$	Powheg	Pythia8	NNPDF2.3		A14	NLO
Background						
V + jets	Sherpa 2.2.1		NNPDF30		Sherpa	NNLO
$t\bar{t}$	Powheg	Pythia8	NNPDF2.3		A14	NLO
Single top	Powheg	Pythia8	NNPDF2.3		A14	NLO
Di-Boson	Sherpa 2.2.1		NNPDF30		Sherpa	NNLO

Table 5.3: Overview of the MC generators used for the main signal and background samples. The "ME" column gives the generator to calculate Matirx Elements. The "PS" column shows the package for generating Parton Showering.

It can be reweighted to a CP-state with the corresponding TauSpinner weight. In each of the plots, the blue dots label the samples simulated with the CP-even or CP-odd hypotheses in Pythia8, and the red dots label the samples reweighted from an unpolarized sample with a TauSpinner weight. The distributions of the kinematic observables and the φ^*_{CP} in the reweighted samples show similar behavior with the samples simulated with the same CP state.

Table 5.3 also lists the generators used for the simulation of the background samples. The background with V + jets and di-boson final states are generated with Sherpa 2.2.1 [119]. The background with $t\bar{t}$ and single top quark final states are generated with Powheg and Pythia8.

All simulated samples passed through a full simulation of the ATLAS detector response using the Geant4 package [120, 121]. The effects of multiple interactions in the same and nearby bunch crossings are modeled by overlaying minimum-bias events simulated with Pythia8. Events are weighted to match the distribution of the average number of interactions per bunch crossing [39].

5.2 Event selection

After event reconstruction, the events containing at least one $\tau-$leptons are selected [54]. A set of offline p_{T} thresholds are designed to further select the events. Table 5.4 lists the p_{T} thresholds for different decay modes. The lepton candidates are required to have minimum transverse momenta to ensure trigger operation at the plateau efficiency.

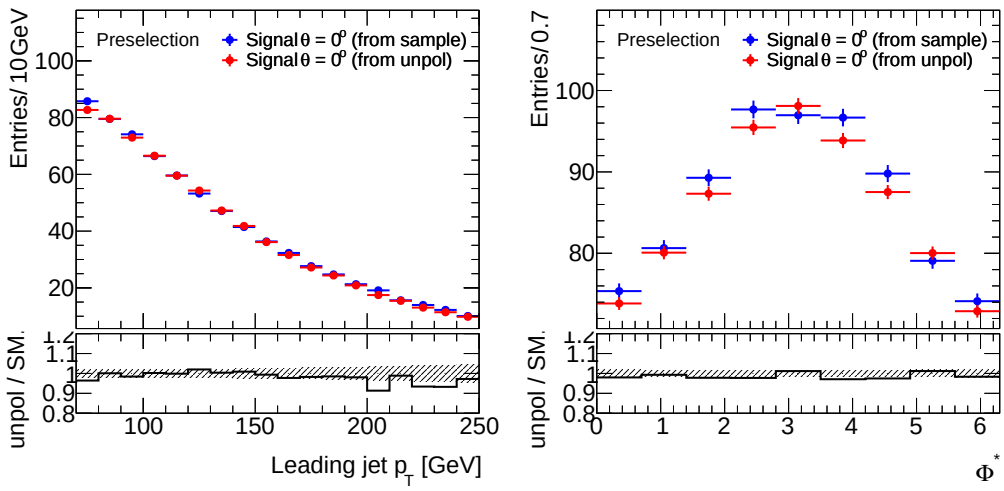


Figure 5.1: Validation of TauSpinner reweighting between SM CP-even and unpolarised samples after TauSpinner reweighting.

Trigger chain	Data period	Trigger p_{T} threshold	Offline p_{T} threshold
Single electron	2015	p_{T} (e) > 24 GeV	p_{T} (e) > 25 GeV
	2016-2018	p_{T} (e) > 26 GeV	p_{T} (e) > 27 GeV
Single muon	2015	p_{T} (μ) > 20 GeV	p_{T} (μ) > 21 GeV
	2016-2018	p_{T} (μ) > 26 GeV	p_{T} (μ) > 27.3 GeV
Ditau	2015-2018	p_{T} (τ_L) > 35 GeV	p_{T} (τ_L) > 40 GeV
		p_{T} (τ_S) > 25 GeV	p_{T} (τ_S) > 30 GeV

Table 5.4: The p_{T} thresholds applied at trigger and offline stage for the selected electrons, muons and $\tau_{\text{had-vis}}$. τ_L is the leading $\tau_{\text{had-vis}}$. τ_S is the sub-leading $\tau_{\text{had-vis}}$.

The two $\tau-$lepton candidates are required to have opposite electric charges in each event. The angular distance between the lepton candidates is required to be $\Delta R_{\tau\tau} < 2.5$, $\Delta\eta_{\tau\tau} < 1.5$ in the $\tau_{lep}\tau_{had}$ channel. For the $\tau_{had}\tau_{had}$ channel, $0.6 < \Delta R_{\tau\tau} < 2.5$ and $\Delta\eta_{\tau\tau} < 1.5$ are required. Each event in the $\tau_{lep}\tau_{had}$ channel requires at least one jets with $p_{\mathrm{T}} > 40$ GeV. The requirement in the $\tau_{had}\tau_{had}$ channel is tightened to contain a jet with $p_{\mathrm{T}} > 40$ GeV, and $|\eta| < 3.2$. The E_T^{miss} is required to be greater than 20 GeV in both decay channels. The $\tau_{lep}\tau_{had}$ channel further requires the transverse mass smaller than 70 GeV, where the transverse mass m_T is defined by $m_T^2 = E^2 - p_z^2$.

The selected events are categorized to the VBF and boosted regions. The VBF region contains events with a sub-leading jet with $p_T > 30$ GeV. The two leading jets are required to have the following kinematics: $m_{jj} > 400$ GeV, $|\Delta\eta_{jj}| > 3.0$; $\eta(j_0) \times \eta(j_1) < 0$. The pseudorapidity η of the two lepton candidates lies between the η of the two leading jets. The Boosted region contains the events with $p_{\mathrm{T}}(\tau\tau) >$

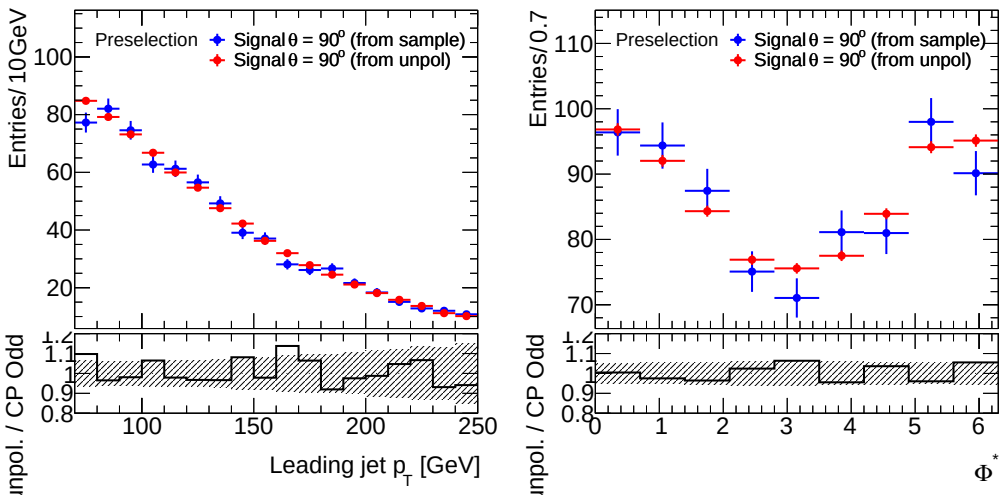


Figure 5.2: Validation of TauSpinner reweighting between SM CP-odd and unpolarised samples after TauSpinner reweighting.

100 GeV and not passing the selection of the VBF region. The VBF region is split into two regions: VBF_1 and VBF_0, based on the output of a BDT-based VBF tagger [122]. The events in the VBF_1 region is more likely to be generated by VBF Higgs boson production. The boost region is also split into Boost_1 and Boost_0 regions. The Boost_1 contains the events with $\Delta R_{\tau\tau} \leq 1.5$ and $p_{\mathrm{T}}(\tau\tau) \geq 140$ GeV.

For each region, the signal region and control region are defined with the invariant mass of the $\tau-$lepton pair, which is estimated with a likelihood-based algorithm called Missing Mass Calculator (MMC) [123]. The invariant mass estimated by the MMC is called $m_{\tau\tau}^{\mathrm{MMC}}$. The signal region is defined with $110 < m_{\tau\tau}^{\mathrm{MMC}} < 150$ GeV. The control region is defined with $60 < m_{\tau\tau}^{\mathrm{MMC}} < 110$ GeV. Table 5.5 lists the selection criteria for the signal regions. The selection criteria are the same for $\tau_{lep}\tau_{had}$ and $\tau_{had}\tau_{had}$ channels.

Within each region, the signal region is divided to high, medium and low regions, depending on the decay channels and selection criteria. The high region contains the events from decay channels: l-1p0n, l-1p1n, 1p0n-1p0n, 1p0n-1p1n and 1p1n-1p1n. In each decay channel, part of the events with higher sensitivity are selected by the selection criteria listed in table 5.6 and table 5.6. The medium region contains the events from l-1pXn, l-3p0n, 1p0n-1pXn, 1p1n-1pXn and 1p1n-3p0n. These decay channels have lower sensitivity than the decay channels in the high region. All the events not passing the selections in the high and medium regions are grouped into the low region.

Signal regions ($110 < m_{\tau\tau}^{\text{MMC}} < 150$ GeV)			
VBF		**Boost**	
Sub-leading jet $p_T > 30$ GeV $m_{jj} > 400$ GeV $\lvert\Delta\eta_{jj}\rvert > 3.0$; $\eta(j_0) \times \eta(j_1) < 0$ Both τ_{had} candidates must lie between the two leading jets in pseudorapidity		Not VBF $p_{\text{T}}(\tau\tau) > 100$ GeV	
VBF 1	**VBF 0**	**Boost 1**	**Boost 0**
BDT(VBF) > 0.0	BDT(VBF) < 0.0	$\Delta R_{\tau\tau} \leq 1.5$ and $p_{\text{T}}(\tau\tau) \geq 140$ GeV	Not Boost 1

$Z \to \tau\tau$ Control regions ($60 < m_{\tau\tau}^{\text{MMC}} < 110$ GeV)			
VBF 1 ZCR	**VBF 0 ZCR**	**Boost 1 ZCR**	**Boost 0 ZCR**
BDT(VBF) > 0.0	BDT(VBF) < 0.0	$\Delta R_{\tau\tau} \leq 1.5$ and $p_{\text{T}}(\tau\tau) \geq 140$ GeV	Not Boost 1

Table 5.5: Summary of selection criteria for the VBF and Boost regions in the analysis.

Channel	Signal region	Decay mode combination	Selection criteria
$\tau_{lep}\tau_{had}$	High	l-1p0n	$d_0^{\text{sig}}(e) > 2.5$ or $d_0^{\text{sig}}(\mu) > 2.0$ $d_0^{\text{sig}}(\tau_{1p0n}) > 1.5$
		l-1p1n	$d_0^{\text{sig}}(e) > 2.5$ or $d_0^{\text{sig}}(\mu) > 2.0$ $\lvert y(\tau_{1p1n})\rvert > 0.1$
	Medium	l-1pXn	$d_0^{\text{sig}}(e) > 2.5$ or $d_0^{\text{sig}}(\mu) > 2.0$ $\lvert y(\tau_{1pXn})\rvert > 0.1$
		l-3p0n	$d_0^{\text{sig}}(e) > 2.5$ or $d_0^{\text{sig}}(\mu) > 2.0$ $\lvert y(\tau_{3p0n})\rvert > 0.6$
	Low	All above	Not passing selection criteria

Table 5.6: Summary of selection criteria for the signal regions in the $\tau_{lep}\tau_{had}$ channel.

Channel	Signal region	Decay mode combination	Selection criteria
$\tau_{had}\tau_{had}$	High	1p0n-1p0n	$d_0^{\text{sig}}(\tau_1) > 1.5$ and $d_0^{\text{sig}}(\tau_2) > 1.5$
		1p0n-1p1n	$d_0^{\text{sig}}(\tau_{1p0n}) > 1.5$ and $\lvert y(\tau_{1p1n})\rvert > 0.1$
		1p1n-1p1n	$\lvert y(\tau_1)y(\tau_2)\rvert > 0.2$
	Medium	1p0n-1pXn	$d_0^{\text{sig}}(\tau_{1p0n}) > 1.5$ and $\lvert y(\tau_{1pXn})\rvert > 0.1$
		1p1n-1pXn	$\lvert y(\tau_{1p1n})y(\tau_{1pXn})\rvert > 0.2$
		1p1n-3p0n	$\lvert y(\tau_{1p1n})\rvert > 0.1$ and $\lvert y(\tau_{3p0n})\rvert > 0.6$
	Low	All above	Not passing selection criteria

Table 5.7: Summary of selection criteria for the signal regions in the $\tau_{had}\tau_{had}$ channel.

5.3 Background estimation

The background processes in the signal region is estimated using Monte Carlo simulation and data-driven techniques. The dominant process in the background is the $Z/\gamma^* \to \tau\tau + jets$. This process is estimated through Monte Carlo simulation with the generators listed in the table 5.3. The control region in each category is used to extract the normalizations of the $Z \to \tau\tau + jets$ background.

The second most significant background is the events with jets misidentified as a $\tau-$lepton, labeled as "fake" in the rest of this thesis. The contribution of the misidentified $\tau-$lepton is estimated with a data-driven method called the fake-factor method. The fake-factor method estimates the contribution of fake events in various regions using data and MC simulated samples. For each signal region, a corresponding anti-τ region is selected. An anti-τ region contains the events passing all the selections in the analysis except for the $\tau-$lepton identification. The distribution of the observables is estimated with the events within this region. The expected yield of the fake events in the signal region is calculated by:

$$N_{fakes}^{SR} = (N_{Data}^{anti-\tau} - N_{MC,\ no\ jet\to\tau}^{anti-\tau}) \times \mathcal{F}, \tag{5.1}$$

where $N_{Data}^{anti-\tau}$ represents the number of events in the anti-τ region of data. The $N_{MC,\ no\ jet\to\tau}^{anti-\tau}$ is the number of non-fake events in the anti-τ region of Monte Carlo simulated samples, normalized to expected yield. The difference between them is multiplied by a fake factor $\mathcal{F}$, which is defined as:

$$\mathcal{F} = R_W F_W + R_{Top} F_{Top} + R_{QCD} F_{QCD}, \tag{5.2}$$

where F_i is calculated separately in the control regions for different sources. The fake events mostly consist of the events with final states of $W+$jets, top quark, and

multiple jets produced by quantum chromodynamics (QCD). The control regions for these background sources are defined by inverting some of the selection criteria so that the region is enriched with the fake events [124, 125]. The fake factor is calculated as:

$$F_i = \frac{N_{Data}^{pass-\tau,CR_i} - N_{MC,\ no\ jet\to\tau}^{pass-\tau,CR_i}}{N_{Data}^{anti-\tau,CR_i} - N_{MC,\ no\ jet\to\tau}^{anti-\tau,CR_i}}, \tag{5.3}$$

where $N_{Data}^{pass-\tau,CR_i} - N_{MC,\ no\ jet\to\tau}^{pass-\tau,CR_i}$ is the difference between the number of events passing the τ identification in the data and MC samples. It is divided by the difference between the number of events failing the $\tau-$lepton identification in the data and MC samples. These calculated fake factors are summed up with a weight R_i, which is the expected fractional contribution of that process.

The fake factor method is based on the following assumptions. The fake factor calculated from different data periods can be combined. The fake factor calculated for different decay channels $\tau_e\tau_{had}$ and $\tau_\mu\tau_{had}$ can be combined. The fake factor calculated in the control region can be used in the signal region.

Besides the $Z \to \tau\tau + jets$ and fake events, there are some processes with small contribution to the background, such as $Z \to ll{+}jets$ and $Z \to \tau\tau{+}jets$, $H \to WW$, and top quark production processes. Background events from these sources are summarised as "others". These are estimated with Monte Carlo simulation and are normalized to their theoretical expectations. All the MC simulated samples are generated with the packages listed in table 5.3.

5.4 Expected event number

Table 5.8 and Table 5.9 shows the Standard Model expected event number in each region of the $\tau_{\text{lep}}\ \tau_{\text{had}}$ and $\tau_{\text{had}}\ \tau_{\text{had}}$ channels in the ATLAS Run-2 data, respectively.

5.5 Systematic uncertainties

Systematic uncertainties generally affect the yields in the signal and control regions as well as the distribution shape of the main fit observable φ^*_{CP}. They can be grouped into three types: the experimental uncertainties, the theoretical uncertainties, and the $\tau-$lepton decay reconstruction uncertainties.

	$H \to \tau\tau$	$Z \to \tau\tau$	Fake	Other Bkgs
VBF 1 high	17.90±2.32	9.61±4.99	2.33±1.02	0.23±0.26
VBF 1 medium	8.13±1.64	6.36±3.63	1.90±0.52	0.06±1.21
VBF 1 low	60.37±6.98	56.68±14.48	30.83±10.82	15.71±7.91
VBF 0 high	11.52±2.04	40.55±9.89	8.53±3.44	5.45±8.90
VBF 0 medium	5.02±0.85	21.28±6.55	4.75±1.15	0.61±1.43
VBF 0 low	38.50±5.72	190.49±43.06	101.27±38.42	61.27±23.13
Boost 1 high	61.30±12.61	269.44±47.49	19.10±3.96	13.78±4.16
Boost 1 medium	25.99±5.17	120.21±19.79	11.00±2.48	5.50±1.63
Boost 1 low	194.74±38.81	1080.25±162.75	288.86±44.61	297.89±27.84
Boost 0 high	50.23±8.24	387.97 ±66.45	79.52 ±11.66	21.93 ±6.92
Boost 0 medium	23.07 ±3.62	196.48 ±36.32	41.73 ±7.28	8.01 ±8.10
Boost 0 low	181.13 ±28.06	1886.67 ±281.04	1063.31 ±127.73	539.24 ±86.25
VBF 1 ZCR	12.50 ±1.66	263.09 ±20.80	22.71 ±5.58	9.10 ±6.85
VBF 0 ZCR	9.47 ±1.39	699.14 ±62.05	74.58 ±17.29	28.18 ±11.62
Boost 1 ZCR	54.09 ±12.22	10844.00 ±582.14	370.50 ±48.92	339.81 ±20.71
Boost 0 ZCR	37.74 ±6.78	7256.64 ±435.14	665.00 ±71.62	217.85 ±37.13

Table 5.8: Expected event number for the signal and contol regions in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel in the ATLAS Run-2 data.

	$H \to \tau\tau$	$Z \to \tau\tau$	Fake	Other Bkgs
VBF 1 high	17.75±2.87	18.05±14.62	8.47±3.15	0.62±1.66
VBF 1 medium	13.70±2.50	12.26±5.03	8.11±2.38	1.86±1.22
VBF 1 low	31.02±4.72	26.90±8.78	19.92±6.55	1.27±1.14
VBF 0 high	12.52±2.54	56.00±19.65	25.35±7.27	6.69±3.25
VBF 0 medium	9.40±1.75	36.20±14.11	20.75±4.90	5.42±2.58
VBF 0 low	21.76±4.25	90.16±22.55	99.66±19.60	15.15±5.83
Boost 1 high	78.97±18.08	426.64±80.49	50.76±12.63	25.50±7.88
Boost 1 medium	58.63±12.46	308.78±59.94	45.87±8.52	20.61±4.10
Boost 1 low	134.43±29.50	764.51±136.35	122.43±26.91	63.86±11.01
Boost 0 high	60.79±12.39	645.85±101.76	258.41±51.44	64.79±16.68
Boost 0 medium	48.49±8.79	494.10±79.63	261.17±41.70	44.96±7.76
Boost 0 low	109.63±20.22	1167.99±191.39	894.93±171.16	174.55±33.03
VBF 1 ZCR	7.07±1.21	188.26±17.31	40.56±9.80	2.70±1.13
VBF 0 ZCR	6.03±1.20	575.28±60.81	140.55±25.16	11.21±3.40
Boost 1 ZCR	44.71±10.62	9854.75±722.34	495.04±106.27	210.02±25.79
Boost 0 ZCR	25.32±5.54	6309.88±496.63	1687.33±326.82	121.54±18.45

Table 5.9: Expected event number for the signal and contol regions in the τ_{had} τ_{had} channel in the ATLAS Run-2 data.

Experimental uncertainties

Experimental uncertainties are estimated following the recommendations from the ATLAS combined performance groups. The uncertainties considered in the analysis are listed below.

- Muon: Energy scale and resolution, reconstruction, identification, isolation, triggers [126].
- Electron: Energy scale and resolution, reconstruction, identification, isolation, triggers, charge identification [127].
- Tau: Energy scale and resolution, reconstruction, identification, triggers, misidentified $\tau-$lepton estimation [128].
- Jet: Energy scale and resolution, flavor, Jet-Vertex-Tagger, b-tagging [129].
- Missing transverse energy: Resolution, energy scale [130].
- Luminosity [131].

Jet uncertainties including jet energy resolution and scale have the largest impact among all the systematic uncertainties. Systematic uncertainties in the jet energy scale for central jets ($|\eta| < 1.2$) vary from 1% for a large range of high-p_T jets $250 < p_T < 2000$ GeV, to 5% at very low p_T (20 GeV) and 3.5% at very high p_T ($p_T >$ 2.5 TeV). The absolute uncertainty on the relative jet energy resolution is found to be 1.5 at 20 GeV decreasing to 0.5 at 300 GeV [132].

The uncertainties from misidentified $\tau-$lepton estimation arise from the statistical uncertainties in the fake factors and their relative weighting. The uncertainties are assigned per $\tau-$lepton pair decay combination in the $\tau_{\text{lep}}\,\tau_{\text{had}}$ and $\tau_{\text{had}}\,\tau_{\text{had}}$ channels.

Theoretical uncertainties

This analysis follows closely the full Run-2 $H \to \tau\tau$ cross-section analysis [97], where most of the $H \to \tau\tau$ and $Z \to \tau\tau$ theoretical uncertainties are inherited from. The theoretical uncertainties are applied to the signal and background processes.

The signal prediction uncertainties mainly come from three main sources: the QCD scale due to missing higher orders calculation. the parton shower and hadronisation model, renormalization and factorization scales [97] and the PDF [111, 85]. These uncertainties are applied to the production cross-sections, which only affect the signal normalization.

For MC background, the uncertainties are considered for renormalization and factorization scales, resummation scale variations, jet-to-parton matching scheme [133], choice of the value of strong coupling constant α_S, and PDF uncertainties [112].

Analysis-specific uncertainties

Besides the experimental uncertainties and the theoretical uncertainties, there are some uncertainties considered specifically in this analysis. These uncertainties arise from the decay mode classification of $\tau-$lepton, the uncertainties affecting the φ^*_{CP} shape, and the fake background estimation.

The uncertainties on the $\tau-$lepton decay mode classification is derived from the decay mode reconstruction efficiency and the event migration between decay modes. The scale factors are derived for a $\tau-$lepton identification working point to correct the MC sample to match with data. In this analysis, the $\tau-$lepton decay mode classification uncertainties are taken from the preliminary measurement hadronic tau substructure in the ATLAS Tau CP working group using the 2016-2017 data[134].

The shape of observable φ^*_{CP} is affected by the uncertainties of the objects such as the charged particles and neutral pions. Most of these uncertainties have already been considered in the experimental uncertainties. Within this analysis, additional uncertainties on the energy scales and angular resolution of the π^0 are measured. The invariant mass of the $\pi^\pm - \pi^0$ system $m(\pi^\pm, \pi^0)$ is found sensitive to π^0 energy scale and angular resolution. The distribution of the observable $m(\pi^\pm, \pi^0)$ is measured in $Z \to \tau\tau$ control regions. For the $\tau_{\mathrm{lep}}\ \tau_{\mathrm{had}}$ decay modes, the $Z \to \tau\tau$ control region consists of $l - 1p1n$ events. For the $\tau_{\mathrm{had}}\ \tau_{\mathrm{had}}$ decay modes, the $Z \to \tau\tau$ control region consists of $1p1n - 1p1n$ events. The uncertainties on the energy scales and angular resolution of the π^0 are determined from the max-likelihood fit in these control regions.

The uncertainty of the TauSpinner reweighting is also assigned. The uncertainties of TauSpinner are calculated for each di-τ decay mode and production mode combination separately. These uncertainties are estimated by comparing the distribution of the truth φ^*_{CP} between the simulated CP-even/odd samples and the templates reweighted by TauSpinner from an unpolarized sample.

5.6 Maximum-likelihood fit

The CP-mixing angle ϕ_τ is measured with the binned maximum likelihood fit. The maximum likelihood fit uses 24 signal regions (SR) and 10 control regions (CR). These signal regions and control regions are listed below.

- Signal regions (12 SRs per $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ channels)
  - VBF 1: High, Medium, Low;
  - VBF 0: High, Medium, Low;
  - Boost 1: High, Medium, Low;
  - Boost 0: High, Medium, Low.
- Z($\to \tau\tau$)+jets control regions (4 CRs per $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ channels)
  - VBF 1;
  - VBF 0;
  - Boost 1;
  - Boost 0.
- ρ-constraint Z($\to \tau\tau$)+jets control regions (1 CR per $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ channels)
  - Inclusive ℓ-1p1n or 1p1n-1p1n.

The Z($\to \tau\tau$)+jets control regions (ZCR) is divided into two part, ZCR 0 and ZCR 1. The ZCR 0 is split to the 8 Z($\to \tau\tau$)+jets control regions for VBF 1, VBF 0, Boost 1, Boost 0 in the $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ channels. Each of the control regions in ZCR 0 contains only one bin. These control regions are used to extract the Z$\to \tau\tau$ normalizations from data.

The ZCR 1 for the $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ channels are called ρ-constrainted Z($\to \tau\tau$)+jets control regions in the analysis. These control regions are used to constrain the π^0 related uncertainties with the shape of ρ invariant mass distributions. In the Standard Model $\tau-$lepton decay, the 1p1n decay products $\pi^{\pm}$ and π^0 are decayed from a ρ particle. The $m(\pi^{\pm}, \pi^0)$ calculated from 4-momentum of $\pi^{\pm}$ and π^0 could constrain the uncertainties related to π^0.

The φ^*_{CP} is an observable within a range between 0 and 2π. In each signal region, the φ^*_{CP} from different decay channels are filled in the same histogram with 9 evenly distributed bins, except for the high and medium regions of VBF 1 with 7 bins and all the low regions with 4 bins. The high and medium regions of VBF 1 have limited sample size. The low regions have low sensitivity to the φ^*_{CP}.

The number of simulated signal and background events are normalized to the therotical predicted values [135]. The 5 normalization factors (NF) are defined to leave the normalization of the $H \to \tau\tau$ signal and Z$\to \tau\tau$ background samples free

floating in the fit. There is one NF for signal: NF H ($\mu_{H\to\tau\tau}$). And there are 4 NF for $Z \to \tau\tau$ background, corresponding to the 4 Z control regions: VBF_1, VBF_0, Boost_1, Boost_0, respectively.

The event number in each bin is expected to follow the Poisson distribution. The likelihood function is defined as a product of conditional probabilities P over binned distributions of the discriminating observables in each event category [136]:

$$\mathcal{L}(x;\phi_\tau,\theta) = \prod_{i}^{N_{bin}} P(n_i|S_i(x;\phi_\tau,\theta) + B_i(\theta)) \prod_{m}^{N_{nuisance}} C_m(\theta). \tag{5.4}$$

In this equation, n_i represents the observed number of events in each bin. $S_i(x;\phi_\tau,\theta)$ and $B_i(\theta)$ are the number of events for the estimated signal and background. Constraints on the nuisance parameters corresponding to the systematic uncertainties are represented by the functions $C_m(\theta)$ [137]. The binned likelihood fit is implemented in the `TRExFitter` package [138], which is a fitting software interfaced with `HistFactory` [139], `RooFit` [140], and `RooStats` [141].

The parameter of interest ϕ_τ is chosen such that the CP-even SM hypothesis is realised for $\phi_\tau = 0$, while the CP-odd hypothesis corresponds to $\phi_\tau = \pm 90$. Other ϕ_τ values correspond to other in-between CP-mixed couplings. A set of nineteen signal templates corresponding to ϕ_τ values from -90 to 90 in intervals of 10 are created by reweighting the simulated unpolarized $H \to \tau\tau$ signal samples with TauSpinner [117], as described in Section 5.1. A simple one-dimensional morphing with linear interpolation [142] on the parameter of interest ϕ_τ is employed in the fit for the CP hypothesis templates. The morphing is a built-in function of `TRExFitter` [138] package that generates templates on the parameter of interest by the linear combinations of given templates.

The normalizations of the CP-even or CP-mixed signal sample and $Z \to \tau\tau$ background sample are introduced as unconstrained nuisance parameters in the likelihood model. Other nuisance parameters enter the fit account for various uncertainties as described in Section 5.5.

Test statistic and quantifying sensitivity

The binned likelihood function $\mathcal{L}(x;\phi_\tau,\theta)$ is a function of the data x, the free-floating mixing angle ϕ_τ, and nuisance parameters θ corresponding to the systematic uncertainties. The likelihood function is evaluated for each ϕ_τ hypothesis using the reweighted signal templates, with the background processes unchanged. The negative log-likelihood (NLL) is defined as the test statistic and an NLL curve can

Fitted parameters	Observed value	Expected value
ϕ_τ	$9 \pm 16^\circ$	$0 \pm 28^\circ$
$\mu_{\tau\tau}$	1.02 ± 0.20	1.00 ± 0.21
$NF_{Z\to\tau\tau}^{VBF1}$	1.04 ± 0.08	1.00 ± 0.08
$NF_{Z\to\tau\tau}^{VBF0}$	0.95 ± 0.07	1.00 ± 0.08
$NF_{Z\to\tau\tau}^{Boost1}$	1.01 ± 0.05	1.00 ± 0.04
$NF_{Z\to\tau\tau}^{Boost0}$	1.02 ± 0.05	1.00 ± 0.05

Table 5.10: Fitted result for the free-floating parameters in the measurement. Observed and expected values are shown for the CP-mixing angle ϕ_τ, the signal strength $\mu_{\tau\tau}$ and the background normalizations $NF_{Z\to\tau\tau}$.

be constructed as a function of ϕ_τ. The test statistic is based on a profile likelihood ratio [143]:

$$q = -2\ln\frac{\mathcal{L}(\vec{\kappa}, \hat{\hat{\vec{\theta}}}(\vec{\kappa}))}{\mathcal{L}(\vec{\kappa}, \hat{\vec{\theta}}(\vec{\kappa}))} = -2\ln\lambda(\vec{\kappa}), \tag{5.5}$$

with the conditional and the unconditional maximum-likelihood estimators in the numerator and the denominator, respectively. The NLL adopted in the analysis is equivalent to $q/2$. The frequentist central confidence interval $[\hat{\phi_\tau} - \sigma_{\hat{\phi_\tau}}, \hat{\phi_\tau} + \sigma_{\hat{\phi_\tau}}]$ can be directly read off the NLL function using Neyman constructions [144]:

$$-\log\mathcal{L}(\hat{\phi_\tau} \pm \mathrm{N}\sigma_{\hat{\phi_\tau}}) = -\log\mathcal{L}_{\max} + \frac{\mathrm{N}^2}{2}, \tag{5.6}$$

where N quantifies the confidence level. N = 1 refers to standard double-sided Gaussian quantiles "1σ". The equation 5.6 means that after constructing the NLL curve by calculating the NLL value for each ϕ_τ hypothesis and a specific dataset $\mathbf{x}$, the 68% central confidence interval (N = 1 refers to standard double-sided gaussian quantiles "1σ") can be determined from the best estimator $\hat{\phi_\tau}$, at which the NLL curve is minimal, by reading off the Δ NLL = NLL - $\mathrm{NLL}_{\min}$ at 0.5. This interval should contain the true value of ϕ_τ in 68% of all cases. The expected central confidence interval can be determined under the assumption that the pseudo-measured dataset contains the pure $\mathcal{CP}$-even SM-like signal and the expected background.

5.7 Results

The post-fit distributions of the observable φ_{CP}^* are presented in this section, where the parameters are set to the value that maximizes the likelihood to observed data. Figure 5.3 and figure 5.4 show the distributions of φ_{CP}^* in the high signal region of

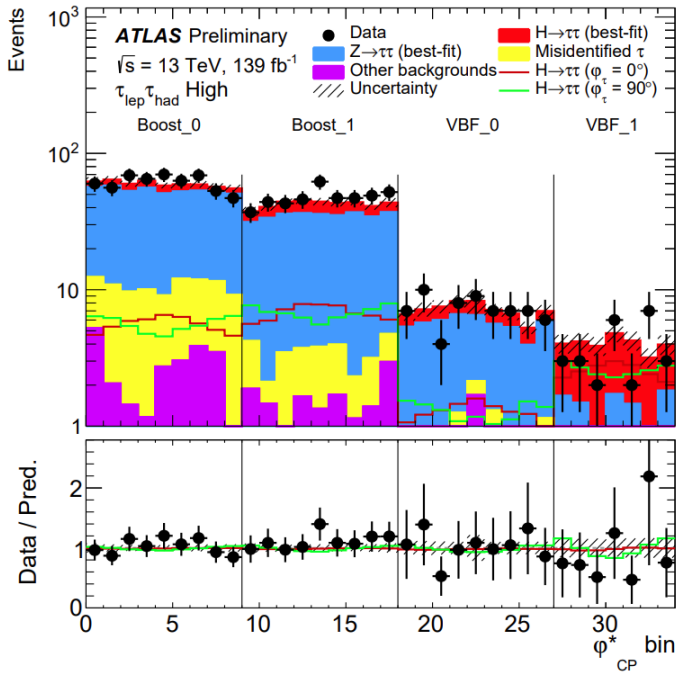


Figure 5.3: Post-fit distributions of φ^*_{CP} in the high signal regions in the $\tau_{\mathrm{lep}}\,\tau_{\mathrm{had}}$ channel.The φ^*_{CP} bins are counted incrementally through all signal regions and cover the range $[0, 360]^\circ$ for each signal region. Other backgrounds include $W+$jets, di-boson, top, $Z \to ll$, and $H \to WW$. The uncertainty band includes all sources of considered uncertainties.

the $\tau_{\mathrm{lep}}\,\tau_{\mathrm{had}}$ and $\tau_{\mathrm{had}}\,\tau_{\mathrm{had}}$ channels respectively. The horizontal axis is bined in observable φ^*_{CP}, and divided into four sections corresponding to the four signal regions VBF_1, VBF_0, Boost_1, and Boost_0. In each of these sections, the bins are evenly arranged to cover the range $[0^\circ, 360^\circ]$. The black dots in the figures represent the observed data. The estimated signal and background events are stacked and colored. The best fit $H \to \tau\tau$ signal is represented by the red color. The dominant background $Z \to \tau\tau$ is represented by the blue color, and the misidentified τ background is represented by the yellow color. The background from $W+$jets, di-boson, top, $Z \to ll$, and $H \to WW$ events contribute much less than the dominant two. These backgrounds are combined and labeled as "Other" and colored with purple. The uncertainty is represented with shaded area. Within the figures, the distributions of φ^*_{CP} in the CP-even and CP-odd states are represented by red and green lines. The post-fit distributions of medium and low signal region in $\tau_{\mathrm{lep}}\,\tau_{\mathrm{had}}$ and $\tau_{\mathrm{had}}\,\tau_{\mathrm{had}}$ channels are shown in figure 5.5, 5.6, 5.7, and 5.8 with the same notations.

Figure 5.9 and figure 5.10 show the post-fit distributions of the $\pi^\pm$ and π^0 total invariant mass in the Z control regions in the $\tau_{\mathrm{lep}}\,\tau_{\mathrm{had}}$ and $\tau_{\mathrm{had}}\,\tau_{\mathrm{had}}$ channels. In the $\tau_{\mathrm{lep}}\,\tau_{\mathrm{had}}$ channel, the l-1p1n events are used. In the $\tau_{\mathrm{had}}\,\tau_{\mathrm{had}}$ channel, the 1p1n-1p1n events are used. In the 1p1n-1p1n channel, there are two choices of ($\pi^\pm$,π^0)

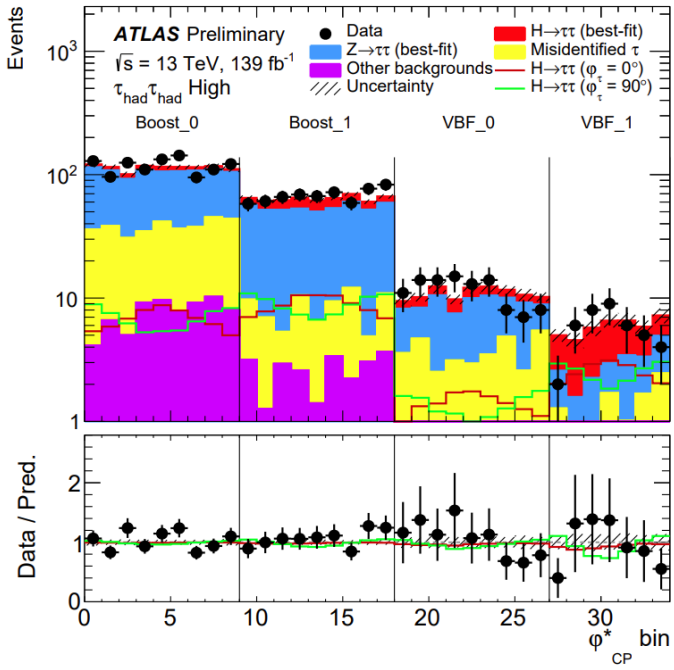


Figure 5.4: Post-fit distributions of φ^*_{CP} in the high signal regions in the τ_{had} τ_{had} channel.The φ^*_{CP} bins are counted incrementally through all signal regions and cover the range $[0, 360]^\circ$ for each signal region. Other backgrounds include W + jets, di-boson, top, $Z \to ll$, and $H \to WW$. The uncertainty band includes all sources of considered uncertainties.

pair. The ($\pi^{\pm}$,π^0) pair from $\tau-$lepton with larger transverse momentum is selected to calculate $m(\pi^{\pm}, \pi^0)$. Data shows a good agreement with the prediction in the distribution of $\pi^{\pm} - \pi^0$ invariant mass.

Table 5.10 lists the fitted parameters in the real data and the Asimov data. The best-fit value of the CP-mixing angle ϕ_τ is $9^\circ \pm 16^\circ$ in the observation and $0^\circ \pm 28^\circ$ in the expectation at 68% confidence level. The pure CP-odd hypothesis is disfavoured by the observation at 3.4σ, while the expected exclusion limit is 2.1σ, where σ is the standard deviation. A larger value of standard deviation refers to a lower possibility that the observation is caused by statistical fluctuation. Figure 5.11 shows the 1D likelihood scan of ϕ_τ. The difference between the observed and expected sensitivities of the ϕ_τ can be attributed to statistical fluctuation in data. The uncertainties of the ϕ_τ measurement are highly dependent on the size of the modulation amplitude of the φ^*_{CP}. The analysis is sensitive to the shape of φ^*_{CP} distribution. This measurement is compatible with the measurement of the CMS Collaboration on the same mixing-angle parameter, where the observed CP-mixing angle is $-1^\circ \pm 19^\circ$ and $0^\circ \pm 21^\circ$ in the expectation at 68% confidence level [136].

The best fit value of signal strength is 1.02 ± 0.20 in the observation and 1.00 ± 0.21 in the expectation. The normalizations for background in each signal region are

Set of nuisance parameters	Impact on ϕ_τ ($^\circ$)
Jet	4.3
E_T^{miss}	0.4
Electron	0.3
Muon	0.9
τ_{had} reconstruction	1.0
Fake τ	0.6
τ_{had} decay mode reconstruction	0.3
π^0 angular resolution and energy scale	0.2
Trajectories and impact parameter	0.7
Flavour Tagging	0.2
Luminosity	0.1
Theory uncertainty in $H \to \tau\tau$ processes	1.5
Theory uncertainty in $Z \to \tau\tau$ processes	1.1
Simulated background sample statistics	1.4
Signal normalization	1.4
Background normalization	0.6
Total systematic uncertainty	5.2
Data sample statistics	15.6
Total	16.4

Table 5.11: Impact of different sources of uncertainty on the ϕ_τ measurement.

shown in table 5.10. These results are compatible with the SM expectation within the measured uncertainties.

Table 5.11 summarises the impact of the uncertainties. The total uncertainty is dominated by the statistical uncertainties of the data sample. The dominant contributions to the systematic uncertainties are from jets, followed by limited sample size of the background simulations, uncertainties from the free-floating normalization factors and theory uncertainties. Effects from other sources are no larger than 1°.

The expected sensitivities in excluding pure CP-odd hypothesis is 1.1σ in the τ_{lep} τ_{had} channel and 1.7σ in the τ_{had} τ_{had} channel. The high signal regions contribute the most. In the τ_{lep} τ_{had} channel, the high signal region contributes 1.0σ, and the high region in the τ_{had} τ_{had} channel contribute 1.4σ. Other signal regions contribute little.

Figure 5.12 shows the 2D likelihood scan on the CP-mixing angle ϕ_τ and the signal strength $\mu_{\tau\tau}$. The solid line shows the 1σ confidence level region. The dashed line shows the 2σ confidence level region. No strong correlation is observed between ϕ_τ and $\mu_{\tau\tau}$. The SM prediction with $\phi_\tau = 0$ and $\mu_{\tau\tau} = 1$ is compatible with the observation within the 1σ confidence region.

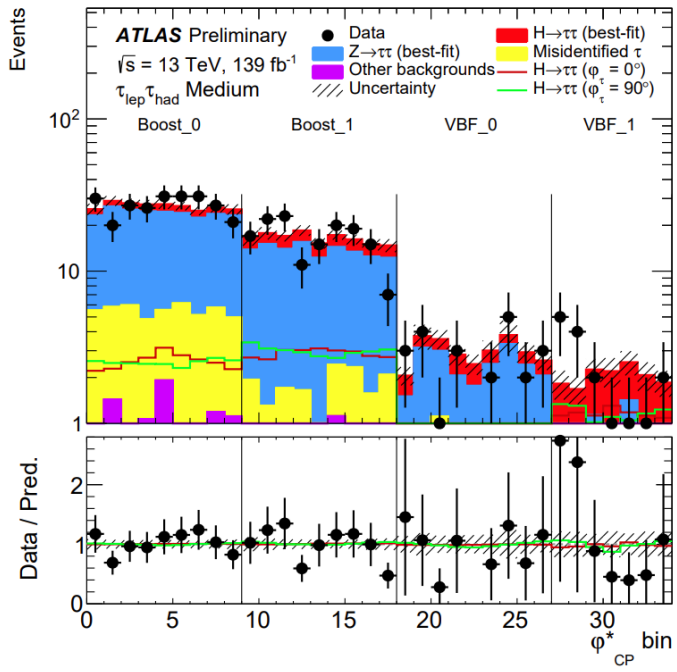


Figure 5.5: Post-fit distributions of φ^*_{CP} in the medium signal regions in the $\tau_{\mathrm{lep}}\,\tau_{\mathrm{had}}$ channel.The φ^*_{CP} bins are counted incrementally through all signal regions and cover the range $[0, 360]^\circ$ for each signal region. Other backgrounds include W + jets, di-boson, top, $Z \to ll$, and $H \to WW$. The uncertainty band includes all sources of considered uncertainties.

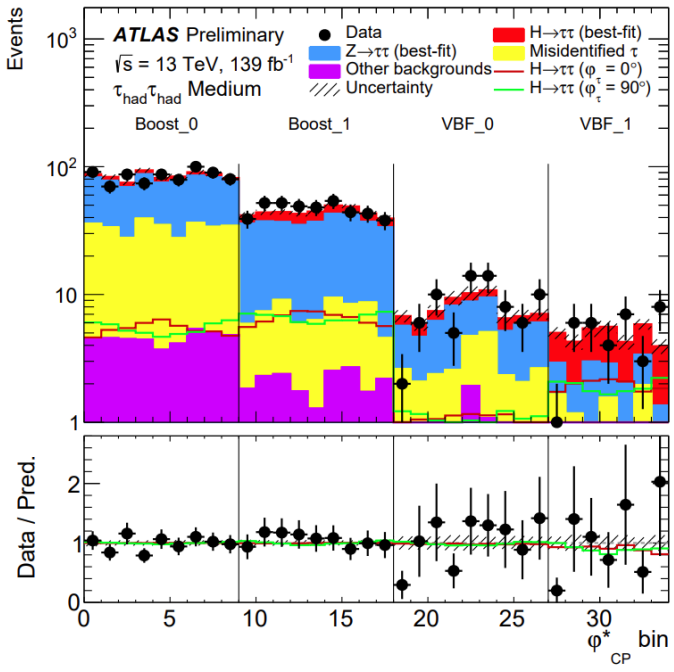


Figure 5.6: Post-fit distributions of φ^*_{CP} in the medium signal regions in the $\tau_{\mathrm{had}}\,\tau_{\mathrm{had}}$ channel.The φ^*_{CP} bins are counted incrementally through all signal regions and cover the range $[0, 360]^\circ$ for each signal region. Other backgrounds include W + jets, di-boson, top, $Z \to ll$, and $H \to WW$. The uncertainty band includes all sources of considered uncertainties.

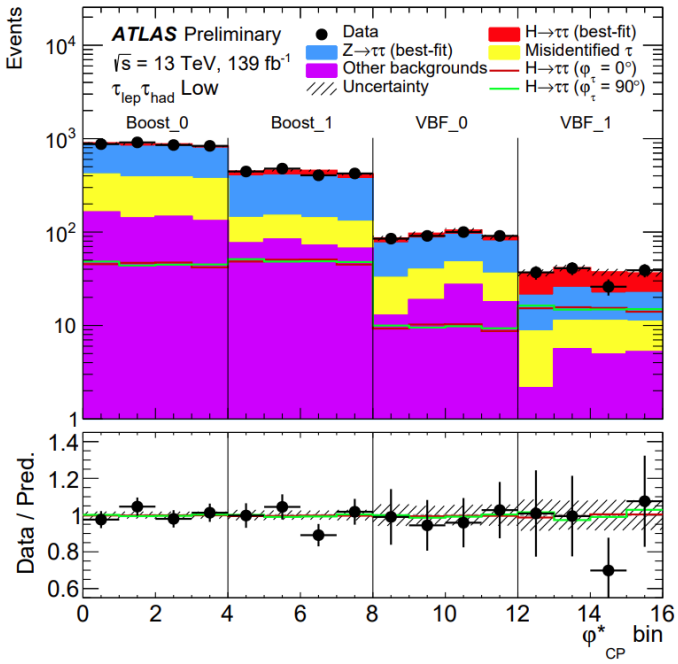


Figure 5.7: Post-fit distributions of φ^*_{CP} in the low signal regions in the $\tau_{\mathrm{lep}}\,\tau_{\mathrm{had}}$ channel. The φ^*_{CP} bins are counted incrementally through all signal regions and cover the range $[0, 360]^\circ$ for each signal region. Other backgrounds include W + jets, di-boson, top, $Z \to ll$, and $H \to WW$. The uncertainty band includes all sources of considered uncertainties.

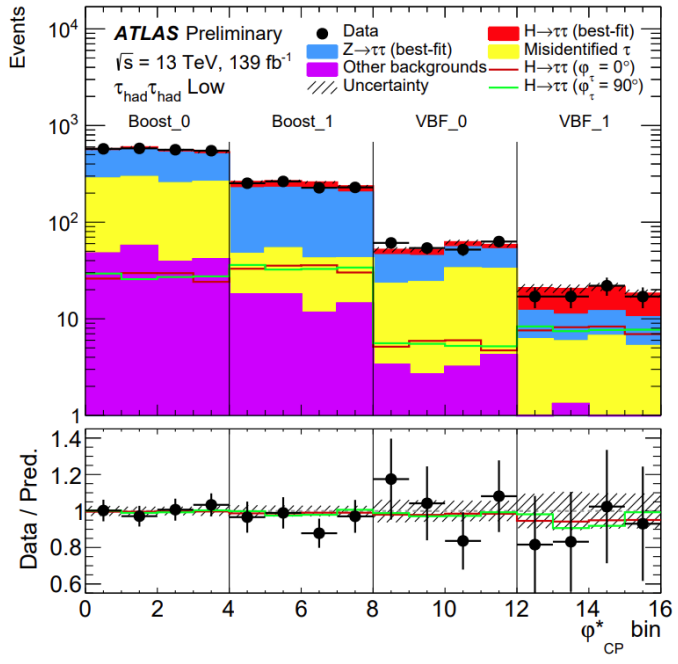


Figure 5.8: Post-fit distributions of φ^*_{CP} in the low signal regions in the $\tau_{\mathrm{had}}\,\tau_{\mathrm{had}}$ channel.The φ^*_{CP} bins are counted incrementally through all signal regions and cover the range $[0, 360]^\circ$ for each signal region. Other backgrounds include W + jets, di-boson, top, $Z \to ll$, and $H \to WW$. The uncertainty band includes all sources of considered uncertainties.

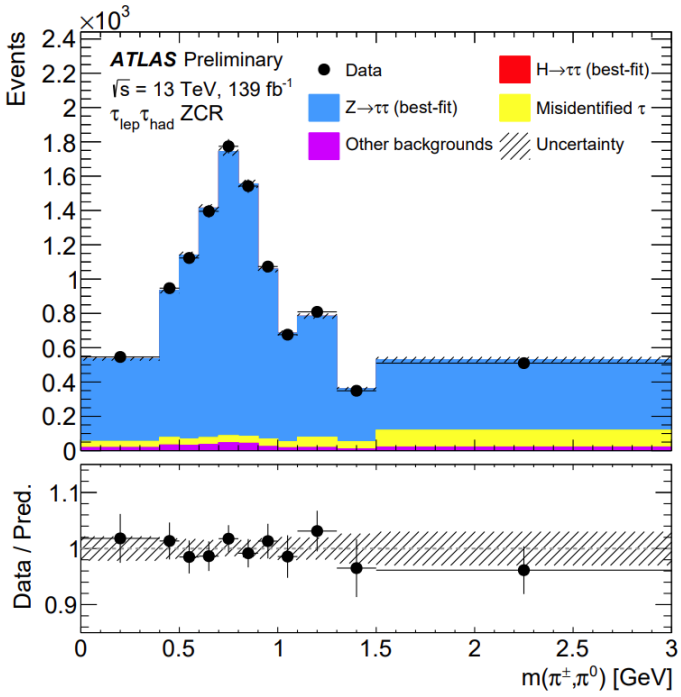


Figure 5.9: Post-fit distributions of $\pi^{\pm} - \pi^{0}$ invariant mass in the Z control regions in the τ_{lep} τ_{had} channel. The l-1p1n events are used in the fitting. Other backgrounds include W + jets, di-boson, top, $Z \to ll$ and $H \to WW$. The uncertainty band includes all sources of considered uncertainties.

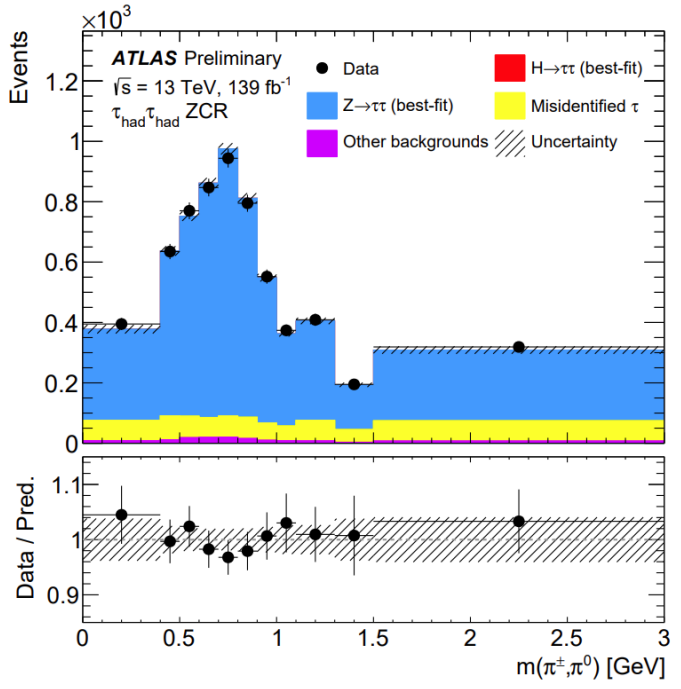


Figure 5.10: Post-fit distributions of $\pi^{\pm} - \pi^{0}$ invariant mass in the Z control regions in the τ_{had} τ_{had} channel.The 1p1n-1p1n events are used in the fitting. Only the τ_{had} with highest p_{T} is selected. Other backgrounds include W + jets, di-boson, top, $Z \to ll$, and $H \to WW$. The uncertainty band includes all sources of considered uncertainties.

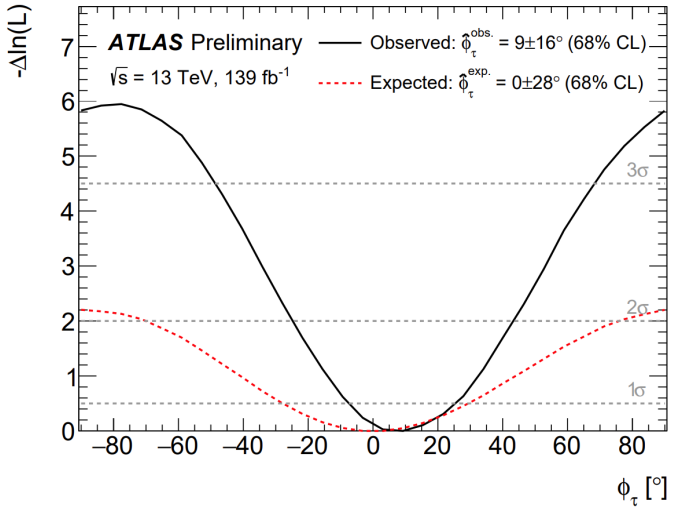


Figure 5.11: 1-D likelihood scan for the CP-mixing angle ϕ_τ. The observed (expected) value of ϕ_τ is $9^\circ \pm 16^\circ$ ($0^\circ \pm 28^\circ$) at 68% confidence level. The value of ϕ_τ is $9^\circ \pm 16^\circ$ at 68% confidence level. The CP-odd hypothesis is rejected at 3.4σ level in the observation and 2.1σ level in the expectation.

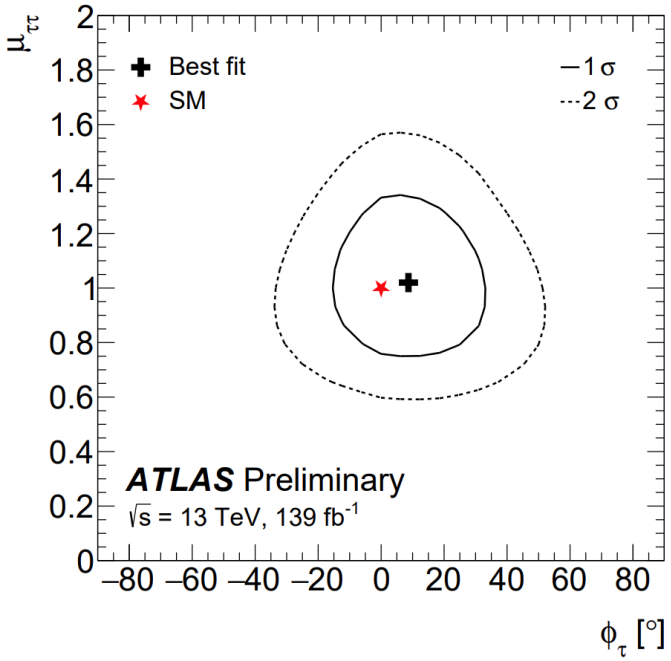


Figure 5.12: The Observed 2D likelihood scan of ϕ_τ and signal strength $\mu_{\tau\tau}$. The solid line shows the 1σ confidence level region. The dashed line shows the 2σ confidence level region.

Chapter 6

Further Studies

The measurement of CP-mixing can be improved in future studies. The LHC Run-3 has started since July-2022, at an unprecedented center of mass energy of 13.6 TeV [145]. The uncertainties of the CP-mixing measurement will be improved with the Run-3 data samples, which are expected to deliver larger integrated luminosity. Some improvements can also be done in the method of the measurement. The analysis introduced in this thesis contains the $\tau_{\text{lep}}\ \tau_{\text{had}}$ and $\tau_{\text{had}}\ \tau_{\text{had}}$ channels, while the $\tau_{\text{lep}}\ \tau_{\text{lep}}$ channels are not being taken into account. The events with 3p1n decay or more than one τ_{had} decay into 3p0n or 1pXn are also not considered in the current analysis. These decay channels can contribute to the measurements even if they showed low sensitivity in the reconstructed Run-2 samples. Besides, some studies on the algorithms are being performed. This chapter introduces one of the ongoing studies on exploring the CP-sensitive 3p0n-1p1n observables with machine learning techniques.

6.1 Investigation of CP-sensitive 3p0n-1p1n observables

The observable φ^*_{CP} in 3p0n-1p1n channel is constructed with the a1-ρ method as described in chapter 4. There exists an arbitrariness in how to determine the intermediate ρ in the τ decay product. In the full analysis, several different observables are tested. The one using the lowest p_T π pair shows the best expected pure CP-odd exclusion limit in an MC simulated data fit, without considering the nuisance parameters. Other observable does not show considerable improvement to the lowest p_T method. The methods of constructing the φ^*_{CP} in the 1p1n-3p0n channel are based on assumptions, which have numerous possibilities. Section 4.3 lists some of these

attempts. Machine learning techniques can help searching for optimal observables with better efficiency.

Several published articles have shown the potential of machine-learning based methods to study the CP-violation in $H \to \tau\tau$ decay, including 3p0n-1p1n channel events [146, 147]. In these articles, neural network models are designed to evaluate the preference of a single $H \to \tau\tau$ decay event to be CP-even and CP-odd. The neural network models take the 4-momentum and the charges of all 4 $\pi^{\pm}$ and 1 π^0 in a 3p0n-1p1n τ_{had} τ_{had} decay as the input, and provide a number between 0 and 1 as the output. The input events are generated with Pythia8 [118] and the spin correlation is simulated by TauSpinner [115, 116, 117]. These MC samples are shared by the particle physics group in the Institute of Physics, Jagellonian University and used in this study. The output of the model determines whether an event has a higher TauSpinner weight in the CP-even or CP-odd state. Their result published in 2017 demonstrated the possibility of predicting the spin correlation by the reconstructable information of visible decay products of $\tau-$lepton [146].

In the ATLAS $H \to \tau\tau$ decay CP analysis, the observable φ^*_{CP} is measured in a binned histogram. The histogram is reweighted to different CP hypotheses. This study designed a neural network to train a CP-sensitive observable for the 3p0n-1p1n decay channel. The neural network model is required to take the reconstructable information of $\pi^{\pm}$ and π^0 as the input and to provide an output variable, defined between 0 and 2π. The model is trained to achieve better separation between the histograms reweighted by the TauSpinner weight of the CP-even and CP-odd states. considering that there is no known model that provides the best separation, the model is trained with reinforcement learning instead of supervised learning.

6.2 Main types of machine learning techniques

Machine learning technique has become widely used in particle physics analysis. There are mainly three types of machine learning techniques: supervised learning, unsupervised learning and reinforcement learning.

Supervised Learning

Supervised Learning is a method used to map input dataset with a set of labels. In the training process, it requires a set of pairs, consisting of an input object and a desired output value. A supervised learning algorithm tends to learn a model to give an output for each input object and make it as close as possible to the desired

output value. Therefore the supervised learning may design the loss function as a function of the desired value and algorithm output value.

Supervised learning is widely used in the binary classification. Both ATLAS and CMS are using supervised learning model in the $H \to \tau\tau$ analysis, for instance, the particle identification [79]. In a binary classification, the output of the model is a number. A working point is selected within the range of model output and the sign of the difference between the output number and the work point determines the classification of the input. The Jet-vertex-tagger described in Chapter 4 is a typical example of the supervised learning based algorithm.

Unsupervised Learning

Unsupervised learning is a type of machine learning algorithm that learns patterns in an unlabeled dataset. Even though the dataset is unlabeled, the unsupervised learning may learn the feature to separate the input events into clusters. Sometimes it is mixed with supervised learning, and called semi-supervised learning. Semi-supervised learning extract feature from dataset that only part of the dataset has desired output value.

Reinforcement Learning

Reinforcement learning is a kind of machine learning algorithm that trains an intelligent agent to take action to some state in an environment. Figure 6.1 shows the working principle of a typical reinforcement learning algorithm. During the training of a reinforcement learning model, the agent detects states of the environment and takes actions based on the model. The actions may affect the environment and change its state. The actions grant rewards to the agent. The actions may come to an end and a final reward is calculated. The model will be updated based on the final reward.

6.3 Searching for 3p0n-1p1n observable with reinforcement learning

Instead of raising and testing methods manually as demonstrated in section 4.3, the reinforcement learning may test the methods automatically with a much higher speed. A neural network ϕ_{CP}^{NN} is designed to replace the $\varphi_{\mathrm{CP}}^{*}$ in 3p0n-1p1n channel. A reinforcement learning model is designed as below. The environment is the training dataset and two histograms of the same size, which are initially empty. The agent

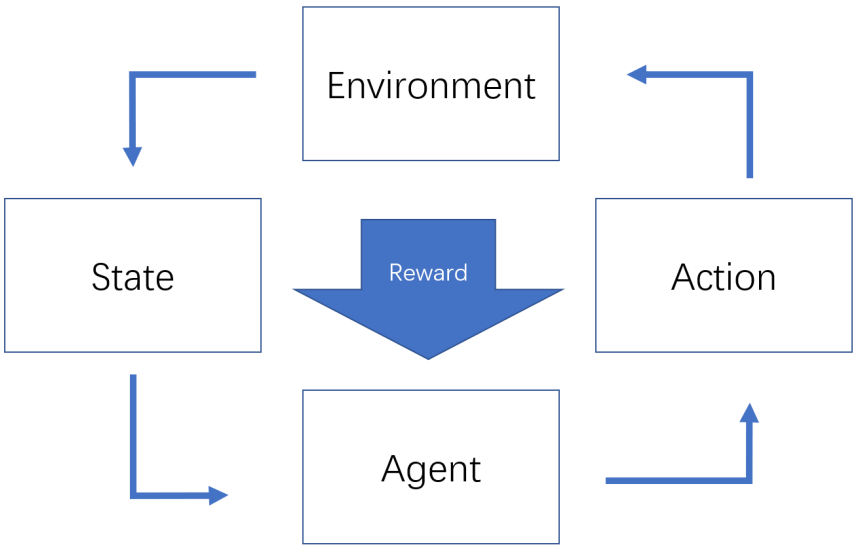


Figure 6.1: The typical structure of a Reinforcement learning algorithm.

receives a state that contain an event from the training dataset and the status of the histograms. The agent provides a number n as the output of the ϕ_{CP}^{NN} model, which determines the nth bin of the two histograms that will be filled. The two histograms are corresponding to CP-even and CP-odd separately. The nth bin of the CP-even histogram is filled with the CP-even TauSpinnner weight, while the nth bin of the CP-odd histogram is filled with the CP-odd TauSpinnner weight. The filling ends when the designed upper limit is reached. The DNLL value is calculated between the CP-odd histogram and the CP-even histogram, as described in chapter 5. The DNLL score is used as the reward to update the ϕ_{CP}^{NN} model.

Monte Carlo sample and input structure

The machine learning model in this study is trained with the MC samples provided by the group that published the articles [146, 147]. These MC samples are generated using Pythia8 and Tauspinner to simulate the spin correlation. Two sets of MC samples are used. One of them only contains $H \to \tau\tau$ events. The other one only contains $Z \to \tau\tau$ events. Each event in either set contains 4-momenta and charges of four $\pi^{\pm}$ and one π^0 in 3p0n-1p1n τ_{had} τ_{had} decay. Besides, each event contains two weight numbers. For $H \to \tau\tau$ events, the two numbers correspond to TauSpinner weight of CP-even and CP-odd state. Since no spin correlation is expected in the simulated $Z \to \tau\tau$ events, both weight numbers are set to 1 in $Z \to \tau\tau$ events.

Layers	Matrix size	bias vector size	dropout rate	activation function
First Layer	88×59	88	No drop out	$tanh(x)$
Second Layer	12×88	12	0.1	$tanh(x)$
Third Layer	2×12	2	No drop out	N/A

Table 6.1: Structure of layers in the neural network model ϕ_{CP}^{NN}.

The input of neural network contains the charge of each pion, its kinetic information, the $\varphi_{\mathrm{CP}}^{*}$ and y calculated with 5 different methods. The kinematic information of each pions are boosted to the ZMF of all visible decay products. Then the direction of the particle is expressed as a 3-dimensional vector with unit length. The kinematic information of a pion is converted to a vector of dimension 5, containing energy, p_T and the 3-dimensional direction vector. The input contains CP-sensitive observables calculated using 5 different methods. One of them is the method used in current analysis, which uses the two $\pi^{\pm}$ with the lowest p_T as neutral component. The other four methods are selected by assuming only opposite charge $\pi^{\pm}$ pair can be selected to be neutral component. There are only two ways to select neutral component from opposite charge $\pi^{\pm}$ pair. Each of the two ways provides two $\varphi_{\mathrm{CP}}^{*}$ calculated by $a_1 - \rho$ method and ρ_{int} method separately. Each of the 5 methods contributes 5 numbers in the input of neural network: $\varphi_{\mathrm{CP}}^{*}$ calculated with corresponding method, y, $2\pi - \varphi_{\mathrm{CP}}^{*}$, $\varphi_{\mathrm{CP}}^{*}$ without the correction of y, and $\varphi_{\mathrm{CP}}^{*}$ with an opposite correction $-y$. The $\varphi_{\mathrm{CP}}^{*}$ observable takes value from 0 to 2π. It is represented as a coordinate on a unit circle to avoid discontinuity. This representation is applied to all $\varphi_{\mathrm{CP}}^{*}$ related input variables. Adding these CP-sensitive observables to the input may help the neural network to learn the progress made by previous studies. Since the currently used $\varphi_{\mathrm{CP}}^{*}$ is one of the input variable, the neural network will be trained to obtain an observable better than the known one.

Design of neural network

A neural network ϕ_{CP}^{NN} is designed to replace the CP sensitive observable $\varphi_{\mathrm{CP}}^{*}$ in 3p0n-1p1n τ_{had} decay channel. It is a fully connected neural network, with three layers. Each layer contains a matrix and a bias vector. All the elements in these matrices and bias vectors are floating. They are updated during the training. Chapter 4 has shown how a neural network works. Each layer receives an input vector. The vector pass through the matrix and the result is summed up with the bias vector. All the elements in the result vectors go through the activation function. The result is then delivered to the next layer, or is delivered as the output if it is the last layer.

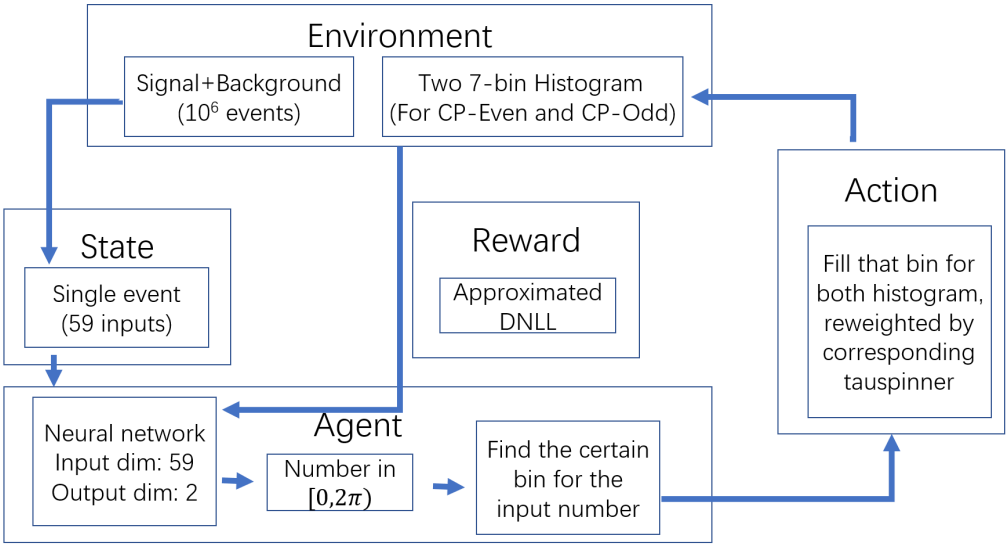


Figure 6.2: Structure of the designed Reinforcement learning algorithm.

Table 6.1 shows the structure of layers with the size of matrix and bias vectors. These layers use $tanh(x)$ as the activation function.

Technique called dropout [148] is applied to the training of ϕ_{CP}^{NN} model. It is a computationally cheap and effective method to reduce overfitting in the training of neural network. The dropout is applied to some layers in a neural network. When a layer is set to a dropout layer with dropout rate r_d, the elements in the input vector of that layer are randomly set to be 0 by a possibility r_d. Those elements not set to 0 are rescaled by $\frac{1}{1-r_d}$. The second layer of neural network ϕ_{CP}^{NN} is a dropout layer with a dropout rate $r_d = 0.1$. The dropout is only applied in the training. No input is muted when the neural network is applied to a test.

The output of the neural network ϕ_{CP}^{NN} is a 2-D vector. The 2-D vector is normalized and mapped to a unit circle, where each point represents an angle within 0 to 2π. The output angle is the observable designed to replace the CP-sensitive observable $\varphi_{\mathrm{CP}}^{*}$. The action of the reinforcement learning is to select a bin index and fill the corresponding bin in two histograms with CP-even and CP-odd TauSpinner weight, where both histograms have 7 evenly divided bins.

The neural network updates the parameters and cleans the histograms after every 3000 input events. The process between each update is called an epoch. The 3000 events contains 1000 $H \to \tau\tau$ events and 2000 $Z \to \tau\tau$ events. The reward is defined as the negative number of the loss function of neural network. It is defined as an approximated log-likelihood of CP-odd histogram to the CP-even histogram.

Loss function

In the training of neural network, the loss function is required to be differentiable. The likelihood function without floating signal strength and nuisance parameters can be expressed as [143]:

$$L = \prod_{j=1}^{N} \frac{(s_j + b_j)^{n_j}}{n_j!} e^{-(s_j+b_j)}, \tag{6.1}$$

where the N is the number of bins of the histograms. Both histograms must have same number of bins. The quantity n_j is the number of events in the j-th bin of the tested histogram. In this study, it is the histogram of CP-odd sample, containing both the $H \to \tau\tau$ signal and the $Z \to \tau\tau$ background. The quantities s_j and b_j are the numbers of signal and background events in the j-th bin of the base histogram, respectively. In this study, the base histogram is the histogram of of CP-even sample, containing both the $H \to \tau\tau$ signal and the $Z \to \tau\tau$ background.

The loss function is defined as the log-likelihood function where the factorial function is replaced by the Stirling's approximation [149]:

$$n! \approx \sqrt{2\pi n}(\frac{n}{e})^n \approx cn^{n+\frac{1}{2}}e^{-n}, \tag{6.2}$$

where $c = \sqrt{2\pi} \approx 2.5066$. The approximated log-likelihood LL_{approx} can be expressed as:

$$LL_{approx} = \sum_m (N_m^{odd} - N_m^{even} + N_m^{odd} ln(N_m^{even}) - ln(c) - (N_m^{odd} + \frac{1}{2}) ln(N_m^{odd})), \tag{6.3}$$

where N_m^{even} and N_m^{odd} are the content in the bin m of the histograms reweighted by CP-even and CP-odd TauSpinner weight. This approximation has high accuracy for a large enough N_m^{odd} and the N_m^{even} is not related to the accuracy. The difference between LL_{approx} and log-likelihood function is:

$$ln(S(n)) - ln(n!) = \frac{1}{12n} + \frac{1}{288n^2} + \cdots, \tag{6.4}$$

where $S(n)$ is the stirling's function. The loss function is the negative number of the approximated delta negative log-likelihood:

$$L_f = -DNLL(odd, even) = LL_{approx}(odd, even) - LL_{approx}(even, even). \tag{6.5}$$

During the training of neural network ϕ_{CP}^{NN}, the contribution to the loss function from bins with entries less than 4 is set to zero. The output of the loss function is called loss.

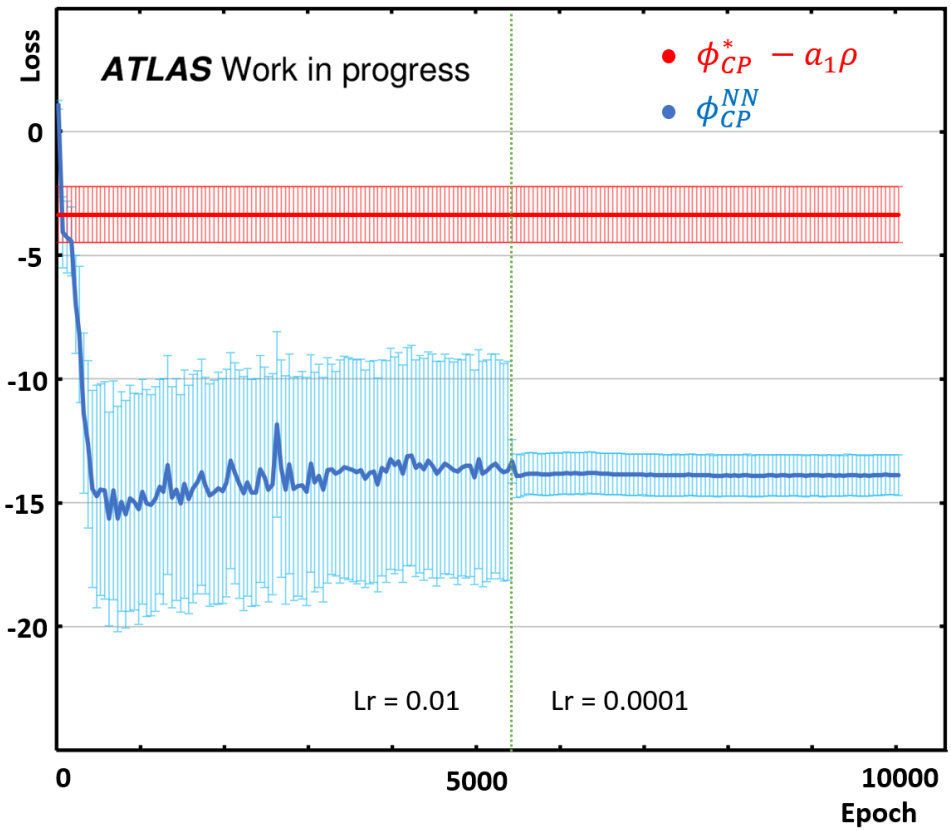


Figure 6.3: The changes of the loss during the training. The red dots represent the loss calculated for φ^*_{CP}. The blue dots represent the loss calculated for the observable trained by ϕ^{NN}_{CP}.

6.4 Results

The model of neural network ϕ^{NN}_{CP} converged after a training with 18000 epochs. The model is trained with samples randomly selected from 10^6 events generated with Pythia8 and TauSpinner. After every 50 epochs, the model is tested with the ATLAS MC sample described in Chapter 5. The test is repeated 800 times with 1000 events randomly selected from $H \to \tau\tau$ samples and 2000 events selected from $Z \to \tau\tau$ samples. Figure 6.3 shows the changes of loss during the whole training. The learning rate was set to 0.01 at the beginning and it was adjusted to 0.0001 after the 5400 epochs. Each pair of points and error bar represents the mean and the standard deviation of the 800 tested results. The red dots in figure 6.3 represent the loss calculated with φ^*_{CP} in the test sample. The blue dots represent the loss of the neural network ϕ^{NN}_{CP} calculated in each epoch. The testing loss of the neural network quickly dropped below the loss for φ^*_{CP} and it became stable when the learning rate is manually set to 0.0001.

Figure 6.4 compares the DNLL value calculated for the observable trained with neural network ϕ^{NN}_{CP} and the observable in different CP hypothesis. The number

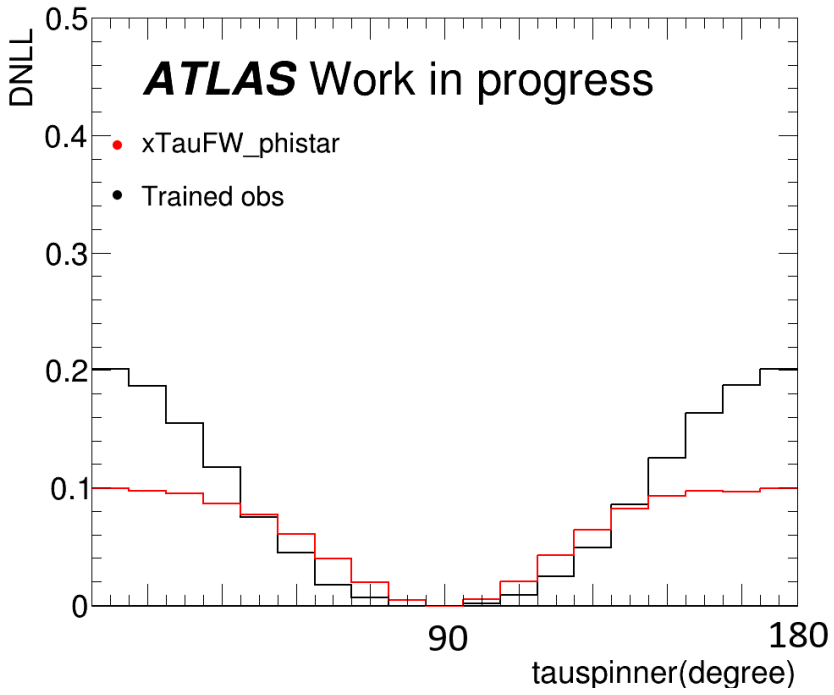


Figure 6.4: Comparison between the DNLL score of trained observable and φ^*_{CP}. The number of $H \to \tau\tau$ events is scaled to the expected yield in Run-2 with a luminosity of 136 fb^{-1}.

of $H \to \tau\tau$ is scaled to the expected yield in Run-2 with a luminosity of 136 fb^{-1}. The background is estimated by $Z \to \tau\tau$ sample with twice the number of $H \to \tau\tau$ events. This figure shows that the model trained with CP-even and CP-odd sample performs well in the hypothesis with ϕ_τ close to $\pm\frac{\pi}{2}$, and performs worse when ϕ_τ close to 0.

Chapter 7

Summary

A measurement of the CP-properties of the interaction between the Higgs boson and $\tau-$leptons has been presented. The measurement was performed using $139 fb^{-1}$ proton–proton collision data collected at a center-of-mass energy of $\sqrt{s} = 13$ TeV. The data was collected by the ATLAS experiment at the LHC in 2015-2018.

This study measures the CP-properties via observables φ^*_{CP} constructed from the $\tau-$lepton decay product. The CP-sensitive observable φ^*_{CP} is defined by IP-method, ρ-method, IP-ρ method or a_1-method, depending on the decay mode of the $\tau-$lepton. The CP-violating interactions are parametrized by the CP-mixing angle ϕ_τ. The distribution of the observable φ^*_{CP} is affected by the CP-mixing angle ϕ_τ. This measurement constrains the ϕ_τ with the maximum likelihood fit to the observable φ^*_{CP}.

The analysis uses the $\tau_{\mathrm{lep}}\,\tau_{\mathrm{had}}$ events from l-1p0n, l-1p1n, l-1pXn and l-3p0n decay modes, and the $\tau_{\mathrm{had}}\,\tau_{\mathrm{had}}$ events from 1p0n-1p0n, 1p0n-1p1n, 1p1n-1p1n, 1p0n-1pXn, 1p1n-1pXn and 1p1n-3p0n decay modes. The events that pass the selection are filled into histograms in the 24 signal region and 10 control regions depending on their kinematic properties and decay mode of $\tau-$lepton pair. The maximum likelihood fit determines the best fit value of ϕ_τ as well as the free floating signal strength and background normalization factors.

Using the Monte Carlo simulated data, the expectation value of the ϕ_τ for the collected dataset is constrained to be $0^\circ \pm 28^\circ$ at 68% confidence level, and $0^{\circ\,+75^\circ}_{\;-70^\circ}$ at 95.5% confidence level. The observed value of ϕ_τ is $9^\circ \pm 16^\circ$ at 68% confidence level, and $9^\circ \pm 34^\circ$ at 95.5% confidence level. The best fit value of the signal strength is 1.02 ± 0.20. The pure CP-odd hypothesis is disfavoured at 3.4 standard deviations. This observation is consistent with the Standard Model expectation.

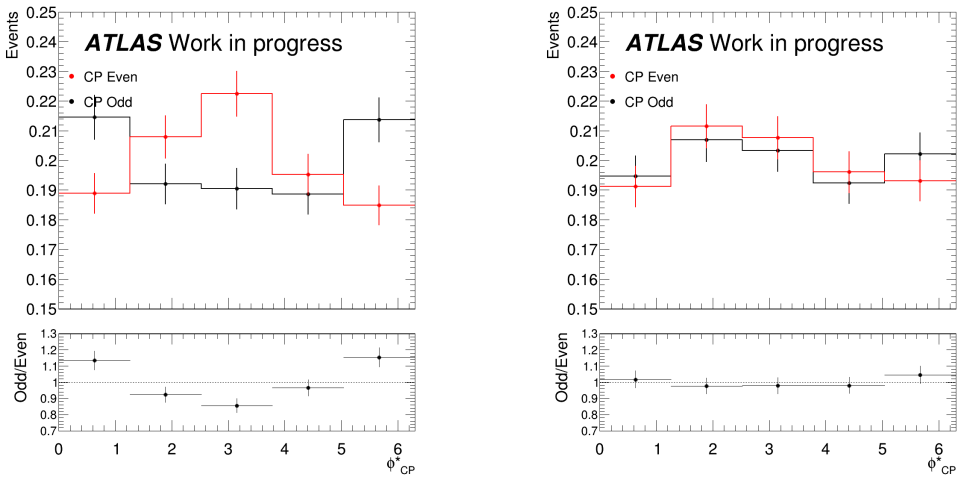
The definition of φ^*_{CP} in the 1p1n-3p0n decay mode is not unique. Several methods of constructing φ^*_{CP} has been presented in this thesis. The performance of these observables is evaluated with Kullback–Leibler divergence. The a_1-method is selected as observable used in this analysis due to its high CP-sensitivity and no requirement on the charge information of $\pi^{\pm}$. A machine learning method is developed in a simplified dataset, which only consists of $H \to \tau\tau$ signal and $Z \to \tau\tau$ background data with twice the event amount of the signal events. This machine learning method performs better in rejecting the pure CP-odd hypothesis compared to the φ^*_{CP} constructed using a_1-method in the simplified dataset. The machine learning method can be further improved with a more complete background estimation.

This thesis also presents an improvement to the pile-up jet tagging algorithm. The Jet-Vertex-Tagger (JVT) algorithm is a commonly used tool in the analysis performed at the ATLAS. It is used to distinguish the hard-scattering jet from the pile-up jets. The study presented in this thesis retrained the JVT with the K-Nearest-Neighbor method using the Monte Carlo simulated sample. The original JVT has the pile-up rejection rate 0.915, 0.956 and 0.964 for the signal efficiency 97% (loose), 92% (medium), 85% (tight), respectively. The retrained JVT improves the pile-up rejection rate to 0.916, 0.961, 0.968 for the loose, medium and tight selections, respectively. A new JVT algorithm using neural network is trained using the same dataset. It improves the pile-up rejection rate to 0.931, 0.966, 0.973 for the loose, medium and tight selections, respectively. This new JVT algorithm has been adopted as the new default jet-vertex-tagging algorithm in the ATLAS Run-3 data analysis.

The precision of the presented analysis is limited by the statistical uncertainty of the data sample. The Run-3 of the Large Hadron Collider has started in Spring 2022. This period is proposed to deliver 280 fb^{-1} total integrated luminosity, which is much larger than the 156 fb^{-1} total integrated luminosity in the Run-2. It will provide more data and help improving the statistical uncertainty. Improvements on the measurement method and reconstruction software may also improve the result.

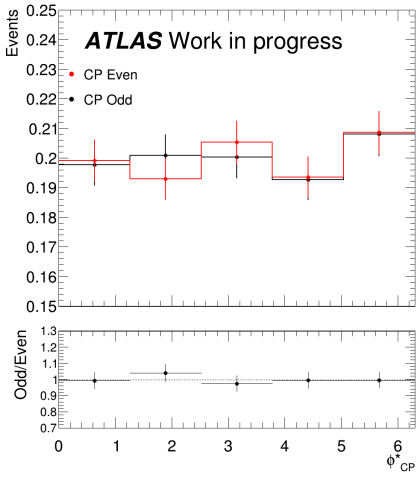
Appendices

A 3-prong observables

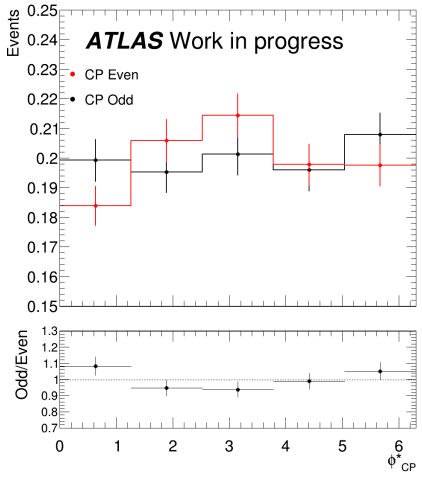


(a)

(b)

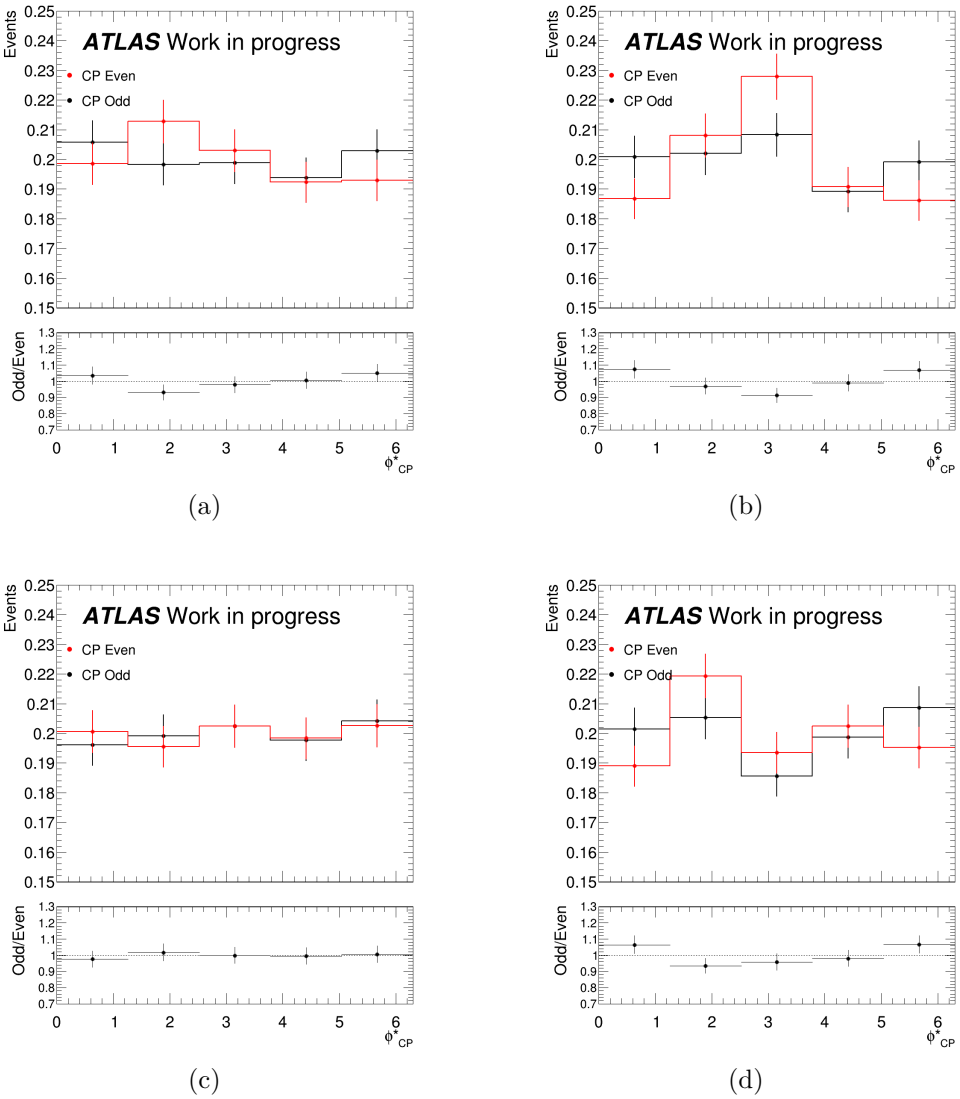


(c)



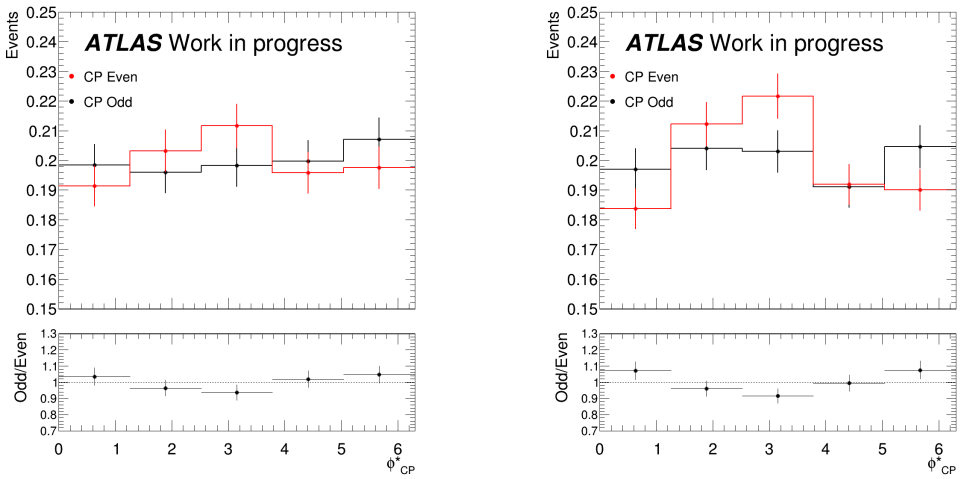
(d)

Figure 1: Type 1 observables in di-τ visible ZMF, the observables are corrected with y_{a_1}. (a) fork component = smallest p_T pair; (b) fork component = largest p_T pair; (c) fork component = pair with closest mass to ρ mass; (d) fork component = opposite charge pair farest to ρ mass.



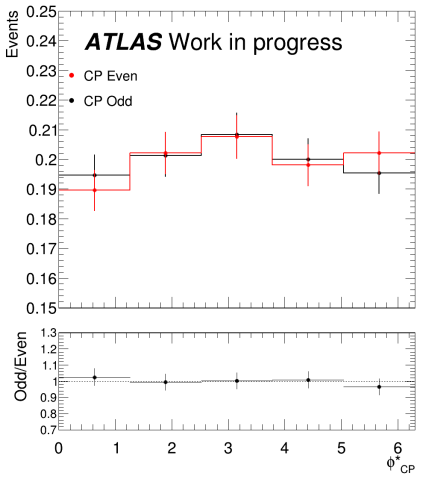
(a) (b) (c) (d)

Figure 2: Type 1 observables in di-τ visible ZMF, the observables are corrected with y_{a_1}. (a) fork component = opposite charge pair with smaller mass; (b) fork component = same charge pair; (c) fork component = opposite charge pair closer to ρ mass; (d) fork component = opposite charge pair farther to ρ mass.

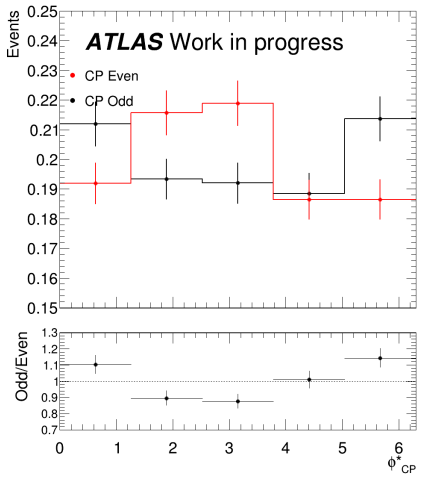


(a)

(b)



(c)



(d)

Figure 3: Type 1 observables in di-τ visible ZMF, the observables are corrected with $y_{\rho-like}$. (a) fork component = smallest p_T pair; (b) fork component = largest p_T pair; (c) fork component = pair with closest mass to ρ mass; (d) fork component = opposite charge pair farest to ρ mass.

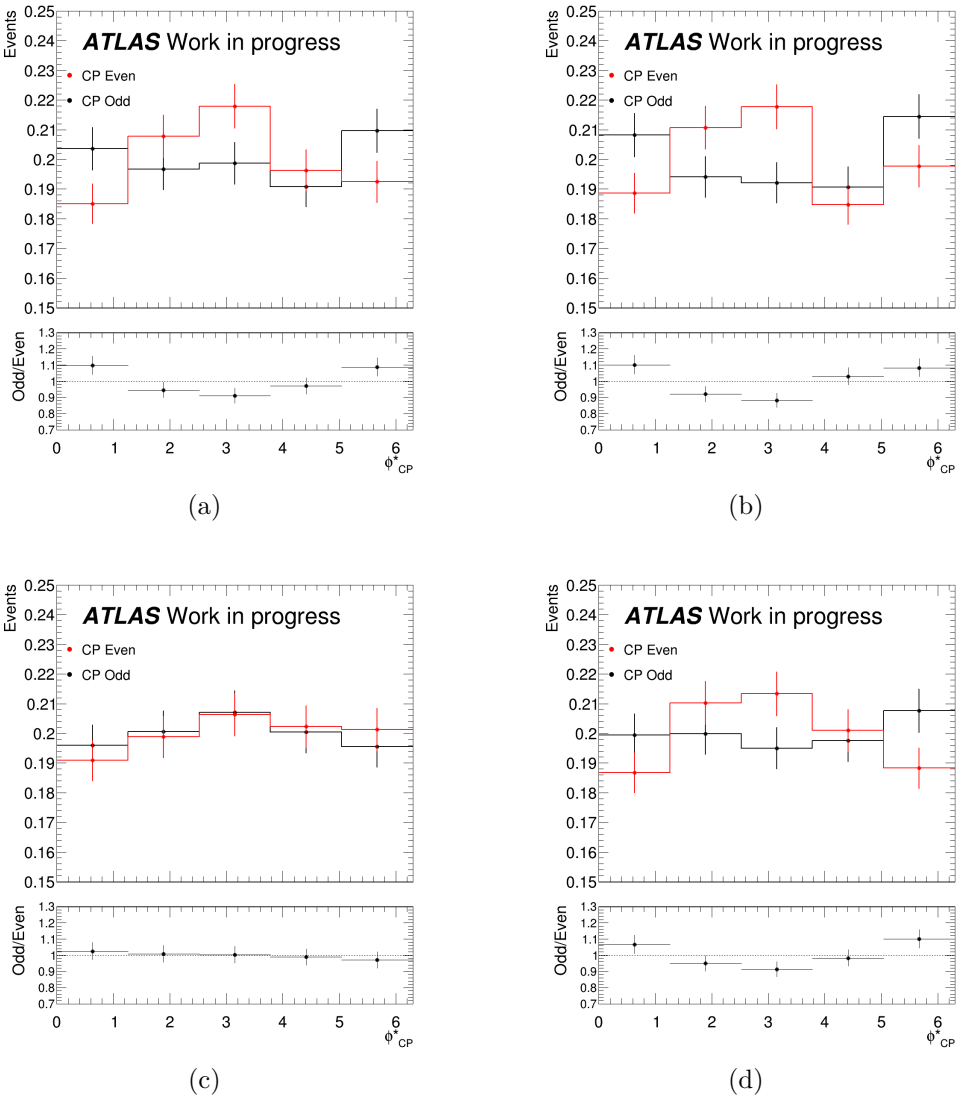
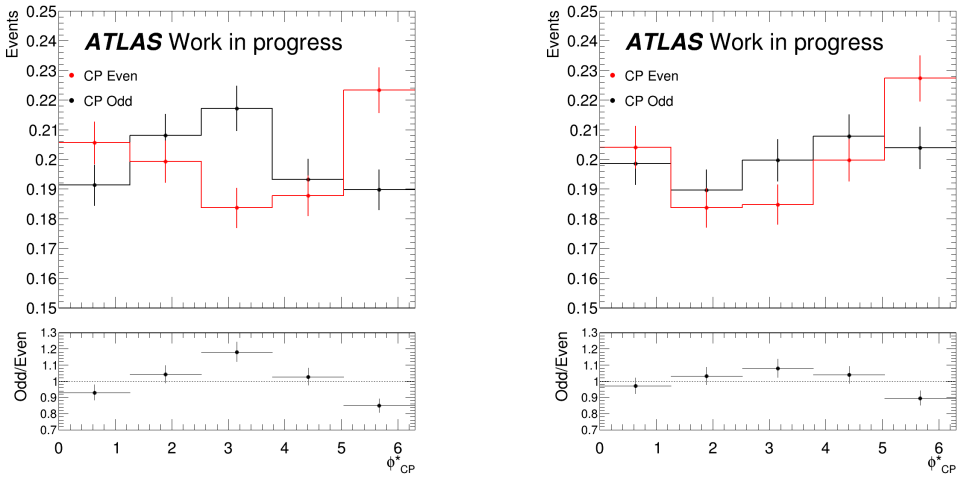
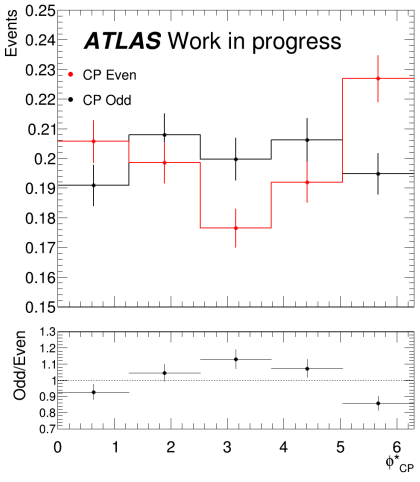


Figure 4: Type 1 observables in di-τ visible ZMF, the observables are corrected with $y_{\rho-like}$. (a) fork component = opposite charge pair with smaller mass; (b) fork component = same charge pair; (c) fork component = opposite charge pair closer to ρ mass; (d) fork component = opposite charge pair farther to ρ mass.



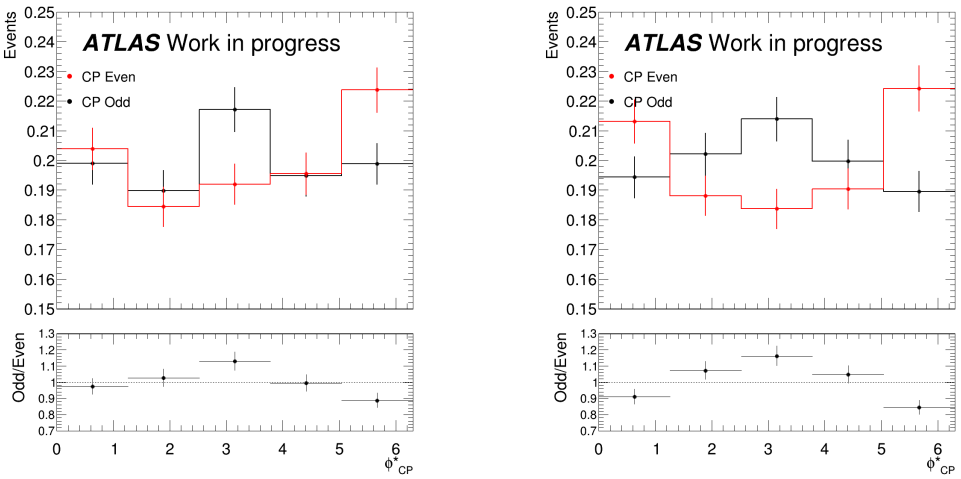
(a)

(b)



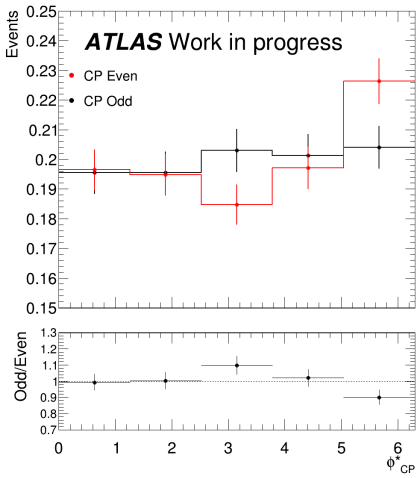
(c)

Figure 5: Type 2 observables in di-τ visible ZMF, the observables are corrected with $y_{\rho\rho}$. (a) fork component = opposite charge pair with mass closer to ρ mass; (b) fork component = opposite charge pair with mass farther to ρ mass; (c) fork component = opposite charge Pair with smaller mass.



(a)

(b)



(c)

Figure 6: Type 2 observables in di-τ visible ZMF, the observables are corrected with $y_{\rho\rho}$. (a) fork component = opposite charge Pair with larger mass; (b) fork component = opposite charge pair with larger p_T; (c) fork component = opposite charge Pair with smaller p_T.

References

[1] **ATLAS** Collaboration, "Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc," *Physics Letters B* **716** (2012) 1–29.

[2] **CMS** Collaboration, "Observation of the higgs boson decay to a pair of τ leptons with the cms detector," *Phys. Lett. B* **779** (2018) 283–316.

[3] **ATLAS** Collaboration, "Cross-section measurements of the higgs boson decaying into a pair of τ-leptons in proton-proton collisions at $\sqrt{s} = 13$ tev with the atlas detector," *Phys. Rev. D* **99** (2019) 072001.

[4] **CMS** Collaboration, "Constraints on the spin-parity and anomalous HVV couplings of the Higgs boson in proton collisions at 7 and 8 TeV," *Phys. Rev. D* (2015) 012004.

[5] C. Zhang and S. Willenbrock, "Effective-Field-Theory Approach to Top-Quark Production and Decay," *Phys. Rev. D* **83** (2011) 034006.

[6] **Planck** Collaboration, "Planck 2015 results. XIII. Cosmological parameters," *Astron. Astrophys.* (2016) A13.

[7] **CDF** Collaboration, "Observation of top quark production in $\bar{p}p$ collisions," *Phys. Rev. Lett.* **74** (1995) 2626–2631.

[8] S. W. Herb *et al.*, "Observation of a Dimuon Resonance at 9.5-GeV in 400-GeV Proton-Nucleus Collisions," *Phys. Rev. Lett.* **39** (1977) 252–255.

[9] P. M. Watkins, "Discovery of the W and Z bosons," *Contemp. Phys.* **27** (1986) 137–163.

[10] S. L. Wu and G. Zobernig, "A Method of Three Jet Analysis in e+ e- Annihilation," *Z. Phys. C* **2** (1979) 107.

[11] J. Ellis, "The Discovery of the Gluon," *Int. J. Mod. Phys. A* **29** (2014) 1430072.

[12] **ATLAS** Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," *Phys. Lett. B* **716** (2012) 1–29.

[13] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons," *Phys. Rev. Lett.* **13** (1964) 508–509.

[14] O. C. Stoica, "The Standard Model Algebra - Leptons, Quarks, and Gauge from the Complex Clifford Algebra $\mathbb{C}\ell_6$," *Adv. Appl. Clifford Algebras* **28** (2018) 52.

[15] M. E. Peskin, "An introduction to quantum field theory," *CRC press* (2018).

[16] S. Ordek, "Investigation of the cp properties of vbf higgs production in hadronic final states of $h \to \tau\tau$ decays with the atlas detector," `CERN-THESIS-2020-277`. (2021).

[17] **Particle Data Group**, "Review of Particle Physics," *PTEP* (2020) 030001.

[18] **CDF** Collaboration, "High-precision measurement of the W boson mass with the CDF II detector," *Science* **376** (2022) 170–176.

[19] S. Schael *et al.*, "Precision electroweak measurements on the Z resonance," *Phys. Rept.* **427** (2006) 257–454.

[20] J. M. Cline, "Baryogenesis," in *Les Houches Summer School - Session 86: Particle Physics and Cosmology: The Fabric of Spacetime*, p. 012005. (2006).

[21] R. H. Cyburt, "Primordial nucleosynthesis for the new cosmology: Determining uncertainties and examining concordance," *Phys. Rev. D* **70** (2004) 023505.

[22] L. Bennett, M. Halpern, and et al., "Searching for Faraday rotation in cosmic microwave background polarization," *The Astrophysical Journal Supplement Series* **148** (2003) 3089–3099.

[23] M.-C. Chen, "TASI 2006 Lectures on Leptogenesis," in *Theoretical Advanced Study Institute in Elementary Particle Physics: Exploring New Frontiers Using Colliders and Neutrinos*. 2007.

[24] M. S. Turner, "Big band baryosynthesis and grand unification," in *Weak Interactions as Probes of Unification*, vol. 72. 1981.

[25] A. D. Sakharov, "Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe," *Pisma Zh. Eksp. Teor. Fiz.* **5** (1967) 392.

[26] A. Riotto and M. Trodden, "Recent progress in baryogenesis," *Ann. Rev. Nucl. Part. Sci.* **49** (1999) 35–75.

[27] T. D. Lee and C. N. Yang, "Question of parity conservation in weak interactions," *Phys. Rev.* **104** (1956) 254–258.

[28] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, "Evidence for the 2π Decay of the K_2^0 Meson," *Phys. Rev. Lett.* **13** (1964) 138–140.

[29] **Belle** Collaboration, "Observation of large CP violation in the neutral B meson system," *Phys. Rev. Lett.* **87** (2001) 091802.

[30] **BaBar** Collaboration, "Observation of CP violation in the B^0 meson system," *Phys. Rev. Lett.* **87** (2001) 091801.

[31] **LHCb** Collaboration, "Evidence for CP violation in time-integrated $D^0 \to h^- h^+$ decay rates," *Phys. Rev. Lett.* **108** (2012) 111602.

[32] M. Kobayashi and T. Maskawa, "CP Violation in the Renormalizable Theory of Weak Interaction," *Prog. Theor. Phys.* **49** (1973) 652–657.

[33] L. Wolfenstein, "Parametrization of the Kobayashi-Maskawa Matrix," *Phys. Rev. Lett.* **51** (1983) 1945–1947.

[34] **CMS** Collaboration, "Study of the Mass and Spin-Parity of the Higgs Boson Candidate Via Its Decays to Z Boson Pairs," *Phys. Rev. Lett.* **110** (2013) 081803.

[35] **ATLAS** Collaboration, "Evidence for the spin-0 nature of the Higgs boson using ATLAS data," *Phys. Lett. B* **726** (2013) 120–144.

[36] M. J. Dolan *et al.*, "Constraining CP-violating Higgs Sectors at the LHC using gluon fusion," *Phys. Rev. D* **90** (2014) 073008.

[37] **ATLAS** Collaboration, "CP Properties of Higgs Boson Interactions with Top Quarks in the $t\bar{t}H$ and tH Processes Using $H \to \gamma\gamma$ with the ATLAS Detector," *Phys. Rev. Lett.* **125** (2020) 061802.

[38] **CMS** Collaboration, "Measurements of tt$H$ Production and the CP Structure of the Yukawa Interaction between the Higgs Boson and Top Quark in the Diphoton Decay Channel," *Phys. Rev. Lett.* **125** (2020) 061801.

[39] **ATLAS** Collaboration, "Luminosity determination for low-pileup datasets at $\sqrt{s} = 5$ and 13 TeV using the ATLAS detector at the LHC," ATLAS-CONF-2020-023. (2020).

[40] "Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc," *Phys. Lett. B* **716** (2012) 30–61.

[41] **LHC Higgs Cross Section Working Group**, "Handbook of LHC Higgs Cross Sections," CERN-2017-002-M. (2016).

[42] **ATLAS** Collaboration, ATLAS, "Observation of $h \to b\bar{b}$ decays and vh production with the atlas detector," *Phys. Lett. B* **786** (2018) 59–86.

[43] **CMS** Collaboration, "Observation of the Higgs boson decay to a pair of τ leptons with the CMS detector," *Phys. Lett. B* **779** (2018) 283–316.

[44] **ATLAS** Collaboration, "A search for the dimuon decay of the standard model higgs boson with the ATLAS detector," *Phys. Lett. B* **812** (2021) 135980.

[45] **CMS** Collaboration, "Measurement of Higgs boson decay to a pair of muons in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$," tech. rep., CERN, (2020).

[46] M. Mangano, "Introduction to the LHC," in *9th Baikal Summer School on Physics of Elementary Particles and Astrophysics*. 2009.

[47] L. Evans and P. Bryant, "LHC Machine," *JINST* **3** (2008) S08001.

[48] M. Lamont *et al.*, "LHC beam and luminosity lifetimes revisited," tech. rep., CERN, Geneva, 2014.

[49] O. S. Bruning *et al.*, "LHC Design Report Vol.1: The LHC Main Ring," CERN-2004-003. (2004).

[50] G. Apollinari *et al.*, "High Luminosity Large Hadron Collider HL-LHC," *CERN Yellow Rep.* CERN-2015-005.1. (2015).

[51] **CMS** Collaboration, "The CMS Experiment at the CERN LHC," *JINST* **3** (2008) S08004.

[52] **LHCb** Collaboration, "The LHCb Detector at the LHC," *JINST* **3** (2008) S08005.

[53] **ALICE** Collaboration, "The ALICE experiment at the CERN LHC," *JINST* **3** (2008) S08002.

[54] **ATLAS** Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider," *JINST* **3** (2008) S08003.

[55] K. Potamianos, "The upgraded Pixel detector and the commissioning of the Inner Detector tracking of the ATLAS experiment for Run-2 at the Large Hadron Collider," *PoS* (2015) 261.

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

[57] **ATLAS** Collaboration, "Alignment of the ATLAS Inner Detector in Run-2," *Eur. Phys. J. C* **80** (2020) 1194.

[58] A. Ahmad *et al.*, "The Silicon microstrip sensors of the ATLAS semiconductor tracker," *Nucl. Instrum. Meth. A* **578** (2007) 98–118.

[59] **ATLAS TRT** Collaboration, "The ATLAS TRT barrel detector," *JINST* **3** (2008) P02014.

[60] M. Gupta, "Calculation of radiation length in materials," tech. rep., CERN. PH-EP-Tech-Note-2010-013 (2010).

[61] **ATLAS** Collaboration, "ATLAS tile calorimeter: Technical design report," `CERN-LHCC-96-42`. (1996).

[62] **ATLAS** Collaboration, "ATLAS muon spectrometer: Technical design report," `CERN-LHCC-97-022`. (1997).

[63] **ATLAS** Collaboration, M. Aaboud *et al.*, "Performance of the ATLAS Trigger System in 2015," *Eur. Phys. J. C* **77** (2017) 317.

[64] **ATLAS** Collaboration, "Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s} = 13$ TeV," *Eur. Phys. J. C* **79** (2019) 639.

[65] **ATLAS** Collaboration, "Concepts, Design and Implementation of the ATLAS New Tracking," `ATL-SOFT-PUB-2007-007`. (2007).

[66] **ATLAS** Collaboration, F. Meloni, "Primary vertex reconstruction with the ATLAS detector," *JINST* **11** (2016) C12060.

[67] **ATLAS** Collaboration, "Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run 2 data collected in 2015 and 2016," *Eur. Phys. J. C* **79** (2019) 205.

[68] **ATLAS** Collaboration, "Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung," `ATLAS-CONF-2012-047`. (2012).

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

[70] **ATLAS** Collaboration, "Atlas muon combined performance with 2017 and 2016 dataset," `ATL-PHYS-PROC-2017-246`. (2017).

[71] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm," *JHEP* **04** (2008) 063—063.

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

[73] L. R. F. Castillo, "The search and discovery of the higgs boson," *Morgan Claypool* (2015) 275–313.

[74] **ATLAS** Collaboration, "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1," *Eur. Phys. J. C* **77** (2017) 490.

[75] ATLAS Collaboration, "Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using *pp* collisions at $\sqrt{s} = 13$ TeV." ATLAS-CONF-2017-029. (2017).

[76] **ATLAS** Collaboration, "Identification of hadronic tau lepton decays using neural networks in the ATLAS experiment," `ATL-PHYS-PUB-2019-033`. (2019).

[77] **ATLAS** Collaboration, "Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using *pp* collisions at $\sqrt{s} = 13\ TeV$," `ATLAS-CONF-2017-029`. (2017).

[78] **ATLAS** Collaboration, "Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV," *Eur. Phys. J. C* **78** (2018) 903.

[79] **ATLAS** Collaboration, "Tagging and suppression of pileup jets," `ATLAS-CONF-2014-018`. (2014).

[80] T. Dorigo *et al.*, "Deep Regression of Muon Energy with a K-Nearest Neighbor Algorithm," `arXiv: 2203.02841`. (2022).

[81] **ATLAS** Collaboration, "Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV." (2015).

[82] T. Melia, P. Nason, R. Rontsch, and G. Zanderighi, "W+W-, WZ and ZZ production in the POWHEG BOX," *JHEP* **11** (2011) 078.

[83] R. Corke and T. Sjostrand, "Multiparton Interactions with an x-dependent Proton Size," *JHEP* **05** (2011) 009.

[84] J. Pumplin *et al.*, "New generation of parton distributions with uncertainties from global QCD analysis," *JHEP* **07** (2002) 012.

[85] D. Stump, J. Huston, J. Pumplin, W.-K. Tung, H. L. Lai, S. Kuhlmann, and J. F. Owens, "Inclusive jet production, parton distributions, and the search for new physics," *JHEP* **10** (2003) 046.

[86] A. Aly *et al.*, "Cost estimation of spatial k-nearest-neighbor operators," *EDBT* (2015).

[87] S.-F. Ge *et al.*, "CP-violating Higgs Di-tau Decays: Baryogenesis and Higgs Factories," *Phys. Rev. D* **103** (2021), no. 9, 095027.

[88] V. D. Barger *et al.*, "New Constraints on the Charged Higgs Sector in Two Higgs Doublet Models," *Phys. Rev. D* **41** (1990) 3421–3441.

[89] S. Berge, W. Bernreuther, and S. Kirchner, "Prospects of constraining the Higgs boson's CP nature in the tau decay channel at the LHC," *Phys. Rev. D* **92** (2015) 096012.

[90] S. Berge, W. Bernreuther, and S. Kirchner, "Determination of the Higgs CP-mixing angle in the tau decay channels," *Nucl. Part. Phys. Proc.* **273-275** (2016) 3164.

[91] S. Berge and W. Bernreuther, "Determining the CP parity of Higgs bosons at the LHC in the tau to 1-prong decay channels," *Phys. Lett. B* **671** (2009) 470–476.

[92] S. Berge, W. Bernreuther, B. Niepelt, and H. Spiesberger, "How to pin down the CP quantum numbers of a Higgs boson in its tau decays at the LHC," *Phys. Rev. D* **84** (2011) 116003.

[93] S. Berge, W. Bernreuther, and S. Kirchner, "Determination of the Higgs CP-mixing angle in the tau decay channels at the LHC including the Drell–Yan background," *Eur. Phys. J. C* **74** (2014) 3164.

[94] S. Berge, W. Bernreuther, and H. Spiesberger, "Higgs CP properties using the τ decay modes at the ILC," *Phys. Lett. B* **727** (2013) 488–495.

[95] R. Józefowicz, E. Richter-Was, and Z. Was, "Potential for optimizing the Higgs boson CP measurement in H $\to \tau\tau$ decays at the LHC including machine learning techniques," *Phys. Rev. D* **94** (2016) 093001.

[96] J. Shlens, "Notes on kullback-leibler divergence and likelihood," `arXiv: 1404.2000`. (2014).

[97] **ATLAS** Collaboration, "Measurement of the Higgs boson coupling properties in the H$\to \tau\tau$ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector," tech. rep. (2020).

[98] **ATLAS** Collaboration, "ATLAS tunes of PYTHIA 6 and Pythia 8 for MC11," `ATL-PHYS-PUB-2011-009`. (2011).

[99] K. Hamilton, P. Nason, E. Re, and G. Zanderighi, "NNLOPS simulation of Higgs boson production," *JHEP* **10** (2013) 222.

[100] K. Hamilton, P. Nason, and G. Zanderighi, "Finite quark-mass effects in the NNLOPS POWHEG+MiNLO Higgs generator," *JHEP* **05** (2015) 140.

[101] M. Ciccolini, A. Denner, and S. Dittmaier, "Strong and electroweak corrections to the production of Higgs + 2jets via weak interactions at the LHC," *Phys. Rev. Lett.* **99** (2007) 161803.

[102] M. Ciccolini, A. Denner, and S. Dittmaier, "Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC," *Phys. Rev. D* **77** (2008) 013002.

[103] P. Bolzoni, F. Maltoni, S.-O. Moch, and M. Zaro, "Higgs production via vector-boson fusion at NNLO in QCD," *Phys. Rev. Lett.* **105** (2010) 011801.

[104] M. L. Ciccolini, S. Dittmaier, and M. Kramer, "Electroweak radiative corrections to associated WH and ZH production at hadron colliders," *Phys. Rev. D* **68** (2003) 073003.

[105] O. Brein, A. Djouadi, and R. Harlander, "NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders," *Phys. Lett. B* **579** (2004) 149–156.

[106] O. Brein, R. Harlander, M. Wiesemann, and T. Zirke, "Top-Quark Mediated Effects in Hadronic Higgs-Strahlung," *Eur. Phys. J. C* **72** (2012) 1868.

[107] L. Altenkamp, S. Dittmaier, R. V. Harlander, H. Rzehak, and T. J. E. Zirke, "Gluon-induced Higgs-strahlung at next-to-leading order QCD," *JHEP* **02** (2013) 078.

[108] A. Denner, S. Dittmaier, S. Kallweit, and A. Mück, "HAWK 2.0: A Monte Carlo program for Higgs production in vector-boson fusion and Higgs strahlung at hadron colliders," *Comput. Phys. Commun.* **195** (2015) 161–171.

[109] O. Brein, R. V. Harlander, and T. J. E. Zirke, "vh@nnlo - Higgs Strahlung at hadron colliders," *Comput. Phys. Commun.* **184** (2013) 998–1003.

[110] R. V. Harlander, A. Kulesza, V. Theeuwes, and T. Zirke, "Soft gluon resummation for gluon-induced Higgs Strahlung," *JHEP* **11** (2014) 082.

[111] J. Butterworth *et al.*, "PDF4LHC recommendations for LHC Run II," *J. Phys. G* **43** (2016) 023001.

[112] **NNPDF** Collaboration, R. D. Ball *et al.*, "Parton distributions for the LHC Run II," *JHEP* **04** (2015) 040.

[113] A. Buckley, "ATLAS Pythia 8 tunes to 7 TeV data," in *6th International Workshop on Multiple Partonic Interactions at the LHC*. `ATL-PHYS-PROC-2014-273`. (2014).

[114] **ATLAS** Collaboration, "Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector," *JHEP* **09** (2014) 145.

[115] Z. Czyczula, T. Przedzinski, and Z. Was, "TauSpinner Program for Studies on Spin Effect in tau Production at the LHC," *Eur. Phys. J. C* **72** (2012) 1988.

[116] A. Kaczmarska *et al.*, "Application of TauSpinner for Studies on τ-Lepton Polarization and Spin Correlations in Z, W and H Decays at the LHC," *Acta Phys. Polon. B* **45** (2014) 1921–1946.

[117] T. Przedzinski, E. Richter-Was, and Z. Was, "TauSpinner: a tool for simulating CP effects in $H \to \tau\tau$ decays at LHC," *Eur. Phys. J. C* **74** (2014) 3177.

[118] Sjöstrand *et al.*, "An Introduction to PYTHIA 8.2," *Comput. Phys. Commun.* **191** (2015) 159–177.

[119] E. Bothmann *et al.*, "Event generation with sherpa 2.2," *SciPost Physics* **7** (2019).

[120] **ATLAS** Collaboration, "The ATLAS Simulation Infrastructure," *Eur. Phys. J. C* **70** (2010) 823–874.

[121] S. Agostinelli, J. Allison, K. Amako, and et al., "Geant4—a simulation toolkit," *Nucl. Instrum. Meth. A* **506** (2003) 250–303.

[122] **ATLAS** Collaboration, "Measurements of Higgs boson production cross-sections in the $H \to \tau^+\tau^-$ decay channel in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector," `CERN-EP-2021-217`. (2022).

[123] A. Elagin, P. Murat, A. Pranko, and A. Safonov, "A New Mass Reconstruction Technique for Resonances Decaying to di-tau," *Nucl. Instrum. Meth. A* **654** (2011) 481–489.

[124] **ATLAS** Collaboration, "Measurements of Higgs boson production cross-sections in the $H \to \tau^+\tau^-$ decay channel in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector," `CERN-EP-2021-217`. (2022).

[125] **ATLAS** Collaboration, M. Aaboud *et al.*, "Cross-section measurements of the Higgs boson decaying into a pair of τ-leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," *Phys. Rev. D* **99** (2019) 072001.

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

[127] **ATLAS** Collaboration, G. Aad *et al.*, "Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data," *JINST* **14** (2019) P12006.

[128] **ATLAS** Collaboration, M. Hübner, "Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at $\sqrt{s} = 13$ TeV," *PoS* **LHCP2018** (2018) 024.

[129] **ATLAS** Collaboration, M. Aaboud *et al.*, "Jet reconstruction and performance using particle flow with the ATLAS Detector," *Eur. Phys. J. C* **77** (2017) 689.

[130] **ATLAS** Collaboration, "$E_{\mathrm{T}}^{\mathrm{miss}}$ performance in the ATLAS detector using 2015-2016 LHC *pp* collisions," `ATLAS-CONF-2018-023`. (2018).

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

[132] **ATLAS** Collaboration, "Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," *Eur. Phys. J. C* **81** (2021) 689.

[133] J. K. Anders *et al.*, "V+Jets theoretical uncertainties estimation via a parameterisation method," tech. rep., CERN, Geneva, 2016.

[134] **ATLAS** Collaboration, "Reconstruction of hadronic decay products of tau leptons with the ATLAS experiment," *Eur. Phys. J. C* **76** (2016) 295.

[135] **ATLAS** Collaboration, "Measurement of the Higgs boson coupling properties in the H$\to \tau\tau$ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector," `ATL-COM-PHYS-2020-721`. (2020).

[136] **CMS** Collaboration, "Analysis of the CP structure of the Yukawa coupling between the Higgs boson and τ leptons in proton-proton collisions at $\sqrt{s} =$ 13 TeV," *JHEP* **06** (2022) 012.

[137] J. Conway, "Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra," `arXiv: 1103.0354`. (2011).

[138] K. Choi *et al.*, "Towards Real-World Applications of ServiceX, an Analysis Data Transformation System," *EPJ Web Conf.* (2021) 02053.

[139] **ROOT** Collaboration, K. Cranmer *et al.*, "HistFactory: A tool for creating statistical models for use with RooFit and RooStats," tech. rep. (2012).

[140] W. Verkerke and D. P. Kirkby, "The RooFit toolkit for data modeling," *eConf* **C0303241** (2003).

[141] **RooStats** Collaboration, "RooStats for Searches," in *PHYSTAT 2011*. CERN, 2011.

[142] A. L. Read, "Linear interpolation of histograms," *Nucl. Instrum. Meth. A* **425** (1999) 357–360.

[143] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics," *Eur. Phys. J. C* **71** (2011) 1554.

[144] J. Neyman, "Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability," *Phil. Trans. Roy. Soc. Lond. A* **236** (1937) 333–380.

[145] S. Fartoukh *et al.*, "LHC Configuration and Operational Scenario for Run 3," tech. rep., CERN, Geneva, Nov, 2021.

[146] E. Barberio *et al.*, "Deep learning approach to the Higgs boson CP measurement in $H \to \tau\tau$ decay and associated systematics," *Phys. Rev. D* **96** (2017) 073002.

[147] K. Lasocha, E. Richter-Was, D. Tracz, Z. Was, and P. Winkowska, "Machine learning classification: Case of Higgs boson CP state in $H \to \tau\tau$ decay at the LHC," *Phys. Rev. D* **100** (2019) 113001.

[148] G. E. Hinton *et al.*, "Improving neural networks by preventing co-adaptation of feature detectors." arXiv: 1307.1347 (2012).

[149] K. Pearson, "Historical note on the origin of the normal curve of errors," *Biometrika* **16** (1924) 402–404.