Search for CP Violation in the Tau Yukawa Coupling with the CMS Run 3 Data

Master Thesis

angefertigt am Deutschen Elektronen-Synchrotron und am
Institut pluridisciplinaire Hubert Curien,
vorgelegt der Fakultät für Mathematik, Informatik und
Naturwissenschaften, Fachbereich Physik, der Universität Hamburg

vorgelegt von

Mathilde Lodovica Witt

aus

Potsdam, Deutschland

Hamburg, 2025

Mathilde Lodovica Witt

Search for CP Violation in the Tau Yukawa Coupling with the CMS Run 3 Data

Masterarbeit der Physik

Universität Hamburg
Institut für Experimentalphysik

Erstgutachterin und -betreuerin: Prof. Dr. Elisabetta Gallo
Zweitgutachterin und -betreuerin: Dr. Anne-Catherine Le Bihan

Abstract

This thesis presents a precision measurement of the CP properties of the Higgs tau Yukawa coupling, focusing on the $H \to \tau\tau \to \mu a_1^{3pr} + 3\nu$ decay channel with ggF and VBF production. The study uses proton–proton collisions data recorded by CMS in 2022 and 2023 at $\sqrt{s} = 13.6\,\text{TeV}$, corresponding to an integrated luminosity of $62.4\,\text{fb}^{-1}$. The CP-mixing angle $\alpha^{H\tau\tau}$ is determined via the coplanarity angle ϕ_{CP}. In the μa_1^{3pr} channel, IP and $\vec{h}_{PV}$ methods are combined for the first time, potentially allowing a twofold improvement in sensitivity compared to the DP method used for hadronic tau decays in Run 2. The tau momenta are reconstructed stepwise, first correcting the amplitudes with the fastMTT di-tau mass reconstruction algorithm and subsequently the angles with a Global Event Fit (GEF). The expected sensitivity to $\alpha^{H\tau\tau}$ is improved within the μa_1^{3pr} channel from $0.41\,\sigma$ to $0.49\,\sigma$, corresponding to a 20% enhancement. The result from this channel will contribute to a combined analysis across all di-tau channels.

Zusammenfassung

Diese Arbeit präsentiert eine Präzisionsmessung der CP-Eigenschaft der Higgs-Tau-Yukawa-Kopplung, spezifisch im $H \to \tau\tau \to \mu a_1^{3pr} + 3\nu$ Zerfall mit den Produktionsmodi ggF und VBF. Es werden Proton-Proton-Kollisionsdaten des CMS-Detektors verwendet, die 2022 und 2023 bei $\sqrt{s} = 13{,}6\,\text{TeV}$ aufgenommen wurden und einer integrierten Luminosität von $62{,}4\,\text{fb}^{-1}$ entsprechen. Der CP-Mischungswinkel $\alpha^{H\tau\tau}$ wird über den koplanaren Winkel ϕ_{CP} gemessen. Für den μa_1^{3pr}-Kanal werden erstmals die IP- und $\vec{h}_{PV}$-Methoden kombiniert, wodurch potenziell eine Verdopplung der Sensitivität gegenüber der in Run 2 eingesetzten DP-Methode für hadronische Tau-Zerfälle erreicht werden kann. Die Tau-Impulse werden schrittweise rekonstruiert, indem zuerst die Amplituden mit dem fastMTT-Algorithmus zur di-Tau-Massenrekonstruktion und anschließend die Winkel mit dem Global Event Fit (GEF) korrigiert werden. Die erwartete Sensitivität auf $\alpha^{H\tau\tau}$ im μa_1^{3pr}-Kanal konnte von $0{,}41\,\sigma$ auf $0{,}49\,\sigma$ gesteigert werden, was einer Verbesserung um 20% entspricht. Das Ergebnis dieses Kanals wird zu einer kombinierten Analyse über alle di-Tau-Zerfallskanäle beitragen.

Résumé

Cette thèse présente une mesure de précision des propriétés CP de l'interaction Yukawa Higgs–tau, centrée sur le canal de désintégration $H \to \tau\tau \to \mu a_1^{3pr} + 3\nu$ dans les modes de production ggF et VBF. L'analyse repose sur les données de collisions proton–proton enregistrées par le détecteur CMS en 2022 et 2023 à $\sqrt{s} = 13{,}6\,\text{TeV}$, correspondant à une luminosité intégrée de $62{,}4\,\text{fb}^{-1}$. L'angle de mélange CP $\alpha^{H\tau\tau}$ est déterminé à partir de l'angle de coplanarité ϕ_{CP}. Pour le canal μa_1^{3pr}, les méthodes IP et $\vec{h}_{PV}$ sont combinées pour la première fois, permettant potentiellement un doublement de la sensibilité par rapport à la méthode DP utilisée pour les désintégrations hadroniques du tau lors du Run 2. La reconstruction des impulsions du tau est effectuée progressivement, en corrigeant d'abord les amplitudes à l'aide de l'algorithme de reconstruction de masse di-tau fastMTT, puis les angles avec le Global Event Fit (GEF). La sensibilité attendue sur $\alpha^{H\tau\tau}$ dans le canal μa_1^{3pr} est ainsi portée de $0{,}41\,\sigma$ à $0{,}49\,\sigma$, soit une amélioration de 20%. Le résultat de ce canal contribuera à une analyse combinée de l'ensemble des canaux di-tau.

Acknowledgements · Danksagung · Remerciements

I would like to express my sincere gratitude to my supervisors, Elisabetta and Anne-Catherine, for their guidance, trust, and constant support throughout this work. I am deeply thankful to Elisabetta for her remarkable availability and dedication. Despite an extremely full schedule, she always found time to discuss my work, even while I was abroad. I always had the feeling that she knew exactly where I stood and what I was working on, which gave me both confidence and direction. I am equally grateful to Anne-Catherine for welcoming me into her group for six months. Thanks to her, I had a smooth and motivating introduction to the subject, and I found my place in Strasbourg, which I now truly consider home. I am very thankful to have been part of their groups. À Strasbourg, je voudrais remercier chaleureusement mes collègues de labo, et plus largement tous ceux qui ont rendu le quotidien si agréable. En particulier, merci à Raph pour m'avoir tout de suite intégrée au groupe et pour m'avoir appris le français avec beaucoup d'humour. Merci à Océane pour avoir toujours été là et pour tous les coups de main. Merci à Paul, mon frère, master de ping-pong, pour les innombrables parties partagées. In Hamburg, I owe special thanks to Stepan, Jacopo, and Aliya. Stepan, thank you for bringing so many skills and projects into the office, everything always seemed possible with you. Jacopo, thank you for your humour and the lightness you brought to our days. You three not only helped me solve the problems I encountered throughout the project, but also made a point of explaining where they came from and how to approach them properly. Together we turned the office into a café, sharing ideas, languages, and many memorable moments. I am grateful to Aliya for always making me feel welcome, and I very much look forward to working together in the future. Thank you to Haoyu, the voice of reason in the office. I also thank Alexei, Caroline, Daina, Di, Gourab, Lakshmi, Oza, Sven, Suman, Thomas, and everyone else in both offices, with you, every day was lively, joyful, and a pleasure. Enfin, mes remerciements les plus sincères vont bien sûr à mes proches, à Strasbourg comme à Hambourg, chez qui j'ai toujours trouvé bien plus qu'un toit. Merci à Victor, Amaury, Mariana, Jeanne, Dominik, Helena, Julia et Roberta pour votre présence et votre amitié. Danke, an meine Freunde aus dem Wohnheim und meine zweite Familie in Oldenburg und Göttingen. Meiner ganzen Familie möchte ich danken, die mir stets Zuversicht entgegenbringt, dass ich meinen Weg in der Wissenschaft gehen werde und mir ihre Freude, Motivation und Disziplin vorgelebt hat. Danke.

Contents

1 Introduction **1**

2 The $H \to \tau\tau$ Process in the Standard Model of Particle Physics **3**
2.1 The Standard Model of Particle Physics . . . 3
2.1.1 Quarks and Mesons . . . 7
2.1.2 Lepton Universality . . . 9
2.1.3 The Tau Lepton . . . 10
2.1.4 Spin, Polarisation, and its Propagation . . . 12
2.1.5 The Higgs Boson at the LHC . . . 15
2.2 Motivation for this CP Violation Analysis . . . 19
2.2.1 Meson Sector . . . 20
2.2.2 Summary of the Chapter and Notation . . . 20

3 Experimental Setup **23**
3.1 The Large Hadron Collider . . . 24
3.2 The CMS Experiment . . . 26
3.2.1 Coordinate System . . . 27
3.2.2 Silicon Tracker . . . 28
3.2.3 Calorimetry . . . 29
3.2.4 Solenoid . . . 30
3.2.5 Muon Chambers . . . 30
3.2.6 Trigger System . . . 31
3.3 Particle Reconstruction . . . 31
3.3.1 Track Reconstruction . . . 32
3.3.2 Vertex Reconstruction . . . 33
3.3.3 Muon Reconstruction . . . 34
3.3.4 Jet Reconstruction . . . 34
3.3.5 Missing Transverse Momentum Reconstruction . . . 35
3.3.6 Tau Lepton Reconstruction . . . 35

4 Strategy of the CP Measurement in $H \to \tau\tau \to \mu a_1^{3pr}$ Decays **39**
4.1 Data and Simulation Samples . . . 40
4.1.1 Recorded Data Samples . . . 40
4.1.2 Generated Monte Carlo Samples . . . 40
4.2 Event Selection . . . 43
4.2.1 Corrections . . . 43

4.2.2 Selection of the $\mu\tau_{had}$ and μa_1^{3pr} Channel . . . 45
4.3 Event Categorisation . . . 46
4.4 Measuring the ϕ_{CP} Angle . . . 47
4.4.1 Impact Parameter Method . . . 49
4.4.2 Decay Plane Method . . . 50
4.4.3 Polarimetric Vector Method . . . 51
4.5 Kinematic Reconstruction . . . 51
4.5.1 fastMTT Method . . . 52
4.5.2 Global Event Fit . . . 55
4.6 Estimation of the Sensitivity to the CP observable . . . 60
4.6.1 Generator-level Observables . . . 60
4.6.2 Model for the ϕ_{CP} Distributions . . . 61
4.6.3 Spin Analysing Power . . . 62
4.7 The ColumnFlow Analysis Tool . . . 62

5 CP Measurement in $H \to \tau\tau \to \mu a_1^{3pr}$ Decays 65
5.1 Distribution of Observables . . . 65
5.2 Event Selection . . . 68
5.3 Signal and Background Estimation BDT . . . 69
5.4 Kinematic Reconstruction . . . 70
5.4.1 fastMTT Method . . . 70
5.4.2 Global Event Fit . . . 73
5.4.3 Secondary Vertex Corrections . . . 73
5.5 Measurement of the ϕ_{CP} angle . . . 74
5.5.1 Reconstruction Stages . . . 74
5.5.2 Statistical Interpretation . . . 77

6 Conclusion 79

Appendix List of Tables 81

Appendix List of Figures 82

Appendix A List of Abbreviations 85

Appendix B Columnflow Taskgraph 89

Appendix C Event Yields of the DesyTau framework 91

Appendix D Further Distributions for the CP Measurement 93
D.1 Distributions of Observables . . . 93
D.2 ϕ_{CP} Distributions for VBF Production . . . 93

Bibliography 93

Chapter 1

Introduction

At the time of the Big Bang, matter and antimatter are generally assumed to have been produced in equal amounts. The overwhelming dominance of matter observed in today's Universe therefore points to a violation of Charge Parity (CP) symmetry. Although the Standard Model (SM) of particle physics does incorporate CP-violating effects, their magnitude is not sufficient to account for the observed asymmetry, indicating the presence of additional sources of CP violation Beyond the Standard Model (BSM).

Following the discovery of the Higgs boson in 2012, by the ATLAS and CMS collaborations, a precise investigation of its properties has become a central component of the physics programme, conducted with the Large Hadron Collider (LHC). While the interactions of the Higgs boson with vector bosons have been shown to be compatible with a purely CP-even structure, its Yukawa couplings to fermions still allow for a CP-mixed component. In this context, a precise determination of the CP nature of the Higgs tau Yukawa coupling probes the SM, either reinforcing its description of this sector or revealing indications of new phenomena.

A comprehensive study of the CP properties of the Higgs boson in di-tau final states was performed for the Run 2 phase of the LHC operation. In that analysis, the CP-mixing angle $\alpha^{H\tau\tau} \in [-90, 90]^\circ$, for which the SM predicts a value of zero, was constrained to

$$\alpha^{H\tau\tau} = -1 \pm 19(\text{stat.}) \pm 1(\text{syst.}) \pm 2(\text{bin-by-bin.}) \pm 19(\text{theo.})^\circ \,@\, 68.3\,\%\,\text{CL},$$

and the purely CP-odd hypothesis was excluded at the level of 3σ for the tau lepton, top quark, and vector boson channels. This work is carried over to Run 3 in a collaboration between the groups of the Deutsches Elektronen-Synchrotron (DESY), Institut pluridisciplinaire Hubert Curien (IPHC), and Imperial College London (IC). The CP structure of the tau Yukawa coupling manifests itself in characteristic angular distributions of the decay products, induced by transverse spin correlations of the tau leptons. Consequently, the angle between the decay planes of the two tau leptons provides a sensitive observable for this measurement, and finally access to the CP-mixing angle $\alpha^{H\tau\tau}$.

This thesis focuses on the $H \to \tau\tau \to \mu a_1^{3pr} + 3\nu$ final state. The channel is particularly well suited for the measurement, since muons can be reconstructed with high precision within the Compact Muon Solenoid (CMS) detector, while the hadronic decay of the other tau lepton retains sensitivity to the spin correlations. From Run 2 to Run 3, the reconstruction method of the hadronic decay plane was changed from the Decay Plane (DP) to

the Polarimetric Vector ($\vec{h}_{PV}$) approach. While the DP approach already exploited most of the available experimental information, the $\vec{h}_{PV}$ method can, in principle, nearly double the achievable sensitivity. In contrast, it is strongly dependent on the accuracy of the tau momentum reconstruction, which requires improvement. For this reason, this thesis is dedicated to the improvement of the tau momentum reconstruction. By applying kinematic and angular correction techniques based on fastMTT and the GEF, the sensitivity achieved in Run 2 is first recovered and subsequently surpassed, enabling an improved determination of the CP-mixing angle $\alpha^{H\tau\tau}$.

This thesis is structured as follows. Chapter 2 outlines the SM and presents the theoretical concepts relevant for the interpretation of this thesis. The experimental scene is set in chapter 3, introducing the LHC accelerator, the CMS detector, and the reconstruction of physics objects from the recorded data that form the basis for the analysis. The analysis strategy is presented in chapter 4. It introduces the datasets and simulated samples used, describes the event selection of final-state candidates, and the categorisation of events into signal and background. It then discusses in detail the reconstruction of the decay planes on the leptonic and hadronic sides, as well as the tau momentum correction methods. Furthermore, the construction and evaluation of the ϕ_{CP} observable at generator and multiple reconstruction levels are explained, followed by the introduction of the analysis framework employed in this work. The results of the studies are presented and discussed in chapter 5, leading to the final conclusions and an evaluation of the improvements achieved, as summarised in chapter 6.

Chapter 2

The $H \to \tau\tau$ Process in the Standard Model of Particle Physics

2.1 The Standard Model of Particle Physics 3
2.1.1 Quarks and Mesons 7
2.1.2 Lepton Universality 9
2.1.3 The Tau Lepton 10
2.1.4 Spin, Polarisation, and its Propagation 12
2.1.5 The Higgs Boson at the LHC 15
2.2 Motivation for this CP Violation Analysis 19
2.2.1 Meson Sector 20
2.2.2 Summary of the Chapter and Notation 20

Particle physics seeks to characterise the fundamental constituents of matter, the forces governing them, and their interactions through observation, precise measurement, and theory. The Standard Model (SM), introduced at the beginning of this chapter, remains the most accurate theory to date, providing a unified description of physics at the smallest scales. Seventeen particles and their antiparticles have been confirmed within the SM. They are classified into fermions, namely six quarks and six leptons, which constitute matter, and bosons, which mediate three of the four fundamental interactions. The most recently discovered boson, the Higgs boson, postulated in 1964 and observed in 2012, arises from the electroweak symmetry breaking, also known as the mass-generating mechanism. As Higgs boson measurements have now entered the precision era, this thesis focuses on studying its Charge Parity (CP) properties in the tau Yukawa coupling, whose theoretical foundation is presented in the second part of this chapter.

2.1 The Standard Model of Particle Physics

The Standard Model (SM) is a finite theory founded on quantum mechanics and special relativity. So far, no deviations from its predictions have been observed [1]. While it accounts for the described phenomena with high precision, several fundamental questions

remain unanswered. Among the most pressing are: How can gravity be incorporated? What is the nature of dark matter, observed only indirectly? Neutrinos are massless in the SM, so what mechanism is responsible for their experimentally observed non-zero masses? And, given the assumed balance between matter and antimatter at the Big Bang, why did it evolve into the asymmetry observed today?

To understand the scope and limitations of the SM, it is first introduced as the theoretical basis from which the Brout–Englert–Higgs (BEH) mechanism and the concept of CP symmetry are developed. This overview sets out the principles and symmetries that underpin particle interactions and thereby establishes the basis for the subsequent discussion.

The SM is constructed on the principle of gauge symmetries: $U(1)_Y \times SU(2)_L$ for the electroweak sector and $SU(3)_C$ for the strong interaction. Its Lagrangian formalism, inherited from Quantum Electrodynamics (QED), describes currents and charges and is postulated to be invariant under local gauge transformations. Hereafter, the term Lagrangian will be used to mean Lagrangian density, as is common in the field of particle physics. According to Noether's theorem, each such symmetry directly implies a conserved current associated with the corresponding charge, as shown in Table 2.1 [2]. Each fundamental gauge interaction is therefore characterised by a symmetry group, a charge q, a coupling constant g, a set of gauge fields V_μ^a, and the associated generators T_a. These features are summarised in Table 2.1, which unifies the parameters of the three gauge sectors into a single, coherent framework.

Table 2.1: Summary of the three gauge groups $U(1)_Y$, $SU(2)_L$, and $SU(3)_C$: symmetries, rotations, charges, couplings, gauge fields, components, and generators. The charges are hypercharge Y, weak isospin T_3, and colour charge C. For the field and generator notations, μ denotes the spacetime indices in 4-vector notation, a indexes the components of the gauge field, and σ_a and λ_a represent the Pauli and Gell-Mann matrices, respectively, while Y is the hypercharge operator.

Symmetry	**Rotation**	**Charge**	**Coupling**	**Gauge field**	$\boldsymbol{V_\mu^a}$	$\boldsymbol{a}$	**Gen.** $\boldsymbol{T_a}$
$U(1)_Y$	phase	Y	g'	B	B_μ	1	$\frac{1}{2}Y$
$SU(2)_L$	weak isospin	T_3	g	W_1, W_2, W_3	W_μ^a	1,2,3	$\frac{1}{2}\sigma_a$
$SU(3)_C$	colour	C	g_s	8 gluons	G_μ^a	1,...,8	$\frac{1}{2}\lambda_a$

Gauge invariance is implemented through the introduction of couplings g and the corresponding gauge bosons V_μ^a, when the partial derivative $\partial\mu$ is replaced by the covariant one $D_\mu = \partial_\mu + igV_\mu^a T_a$. Hence, the fermion–gauge field interaction reads

$$\mathcal{L}_f = i\bar{\psi}_f\gamma^\mu\partial_\mu\psi_f \xrightarrow{\partial\mu\to D_\mu} i\bar{\psi}_f\gamma^\mu D_\mu\psi_f = \mathcal{L}_{fermion}. \tag{2.1}$$

However, left handed (LH) and right handed (RH) fermions couple differently to the weak force (see subsection 2.1.3 paragraph chirality). Therefore, distinct gauge representations are assigned to them. LH states form $SU(2)_L$ doublets, while RH states are singlets under $SU(2)_L$. Historically, additional particles were postulated to complete the multiplets, as done for the top quark [3]. In the lepton sector, the left-handed doublets and right-handed

singlets can be represented as:

$$\begin{pmatrix}\nu_e\\ e\end{pmatrix}_L,\ \begin{pmatrix}\nu_\mu\\ \mu\end{pmatrix}_L,\ \begin{pmatrix}\nu_\tau\\ \tau\end{pmatrix}_L,\ e_R,\ \mu_R,\ \tau_R$$

By adding the gauge field kinetic term $\mathcal{L}_V$ (Equation 2.2), one obtains the complete Lagrangian for a massless theory:

$$\mathcal{L}_V = -\frac{1}{4}V^a_{\mu\nu}V_a^{\mu\nu} \tag{2.2}$$

The field strength tensor $V^a_{\mu\nu}$ is defined as $V^a_{\mu\nu} = \partial_\mu V^a_\nu - \partial_\nu V^a_\mu + g f_{abc} V^b_\mu V^c_\nu$. Here, μ and ν denote spacetime indices. The first two terms represent the kinetic part, describing the spacetime variation of the gauge field. The last term, governed by the structure constants f_{abc}, arises from gauge invariance, where the non-abelian field strength tensors of $SU(2)$ and $SU(3)$ contain additional self-interaction terms of the gauge bosons, in contrast to the abelian $U(1)$ case, where such contributions are absent.
Note that the symmetry requirement forbids both gauge bosons and fermions in chiral theories from acquiring a mass, contrary to measurements [1]. Thus, an electroweak symmetry breaking $\mathcal{L}_\phi$ is introduced, giving rise to the Higgs field ϕ. This mechanism will be discussed in detail below, before combining the resulting Lagrangians of the different theories into the complete SM Lagrangian. To conclude the discussion of gauge theory, the interactions would be formulated in terms of the mass eigenstates of the gauge bosons, obtained after electroweak symmetry breaking, leading to the physical photon γ, Z, and $W^\pm$ bosons. Finally, fermion masses arise from Yukawa couplings, as introduced below.

Spontaneous Symmetry Breaking by the Brout-Englert-Higgs Mechanism

The Brout–Englert–Higgs (BEH) mechanism breaks the electroweak $SU(2)_L \times U(1)_Y$ symmetry of the SM down to the $U(1)_{EM}$. As a result, the W and Z bosons acquire mass, while photons remain massless. The fermion masses arise from Yukawa interactions described earlier in paragraph 2.1.
First, consider a complex scalar doublet $\phi = \begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} = \begin{pmatrix}\phi_1 + i\phi_2\\ \phi_3 + i\phi_4\end{pmatrix}$ which transforms as an $SU(2)_L$ field. It contains one charged component and one neutral component, allowing the neutral field to acquire a vacuum expectation value (VEV) ν without breaking the electromagnetic $U(1)_{EM}$ symmetry. Physically, ϕ represents a spinless scalar, therefore its dynamics are described by a Klein–Gordon Lagrangian $\mathcal{L}_\phi = |D^\mu\phi|^2 - V(\phi)$. The first summand, where D^μ denotes the covariant derivative of $SU(2)_L \times U(1)_Y$, encodes both the kinetic energy of the doublet and its interactions with gauge bosons. The potential is chosen to be the most general renormalisable form:

$$V(|\phi|) = \mu^2|\phi|^2 + \lambda|\phi|^4. \tag{2.3}$$

Here, μ^2 has dimensions of GeV, while λ is dimensionless. The parameter μ^2 is the first quantity of the SM that carries a mass scale. All particle masses depend on this constant. $|\phi|^3$ terms are removed by the nomenclature of ϕ, where terms proportional to $|\phi|^1$ are absent as they would be unphysical. Higher-order terms $|\phi|^n$ with $n \geq 4$ are excluded to preserve renormalisability.
The proposed potential is depicted in Equation 2.3 and Figure 2.1. For the potential to admit a non-zero VEV, μ^2 must be negative. A positive μ^2 would yield a single minimum at $|\phi| = 0$, which is physically unsuitable. For small field values, the quadratic

term dominates the potential, while at large values, the quartic term controls its behaviour. Furthermore, the parameter $\lambda \geq 0$ ensures the potential is bounded from below, allowing for stable ground states. Since ground states exist, the potential $V(|\phi|)$ must be even. The potential in Equation 2.3 has four base axis, one for every dimension: $\mathrm{Im}(\phi^+), \mathrm{Im}(\phi^0), \mathrm{Re}(\phi^+), \mathrm{Re}(\phi^0)$. For simplicity of illustration, only 2 are drawn. The postulated ground state is a circle in the complex plane with a real and an imaginary axis. Since the field has a $SU(2)$ symmetry, it does not matter which minimum is taken.

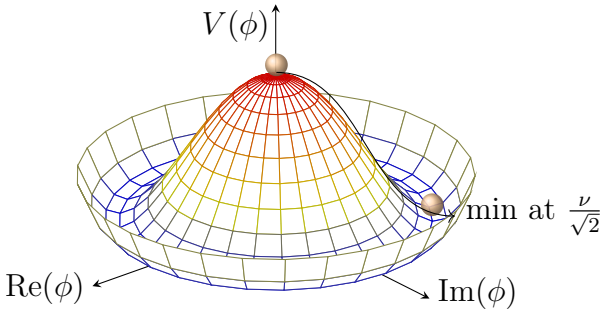


Figure 2.1: Tree-dimensional Higgs potential $V(\phi)$ showing the ground state ring at the vacuum expectation value $\frac{\nu}{\sqrt{2}}$. Modelled after [4].

The symmetry is spontaneously broken by selecting a specific vacuum state $\phi_{vac} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu \end{pmatrix}$. ϕ_1, ϕ_2, and ϕ_4 are set to zero to ensure the electrical neutrality of the vacuum. The complex doublet can then be rewritten as $\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu + H(x) \end{pmatrix}$, introducing a single real scalar field $H(x)$. It represents excitations of the potential relative to the minimum, which itself is the physical Higgs particle. Substituting this expression into the potential yields

$$\begin{aligned} V(\phi) &= \frac{1}{2}\mu^2(\nu + H)^2 + \frac{1}{4}\lambda(\nu + H)^4 \\ &= (\mu^2 + \lambda\nu^2)\nu H + \lambda\nu^2 H^2 + \lambda\nu H^3 + \frac{\lambda}{4}H^4 \\ &= -\frac{\lambda}{4}\nu^4 + \lambda\nu^2 H^2 + \lambda\nu H^3 + \frac{\lambda}{4}H^4. \end{aligned} \tag{2.4}$$

The constant H^0 term corresponds to the vacuum energy, which is an offset in the entire universe[1], and therefore does not contribute to the equations of motion. The quadratic term represents the Higgs mass term, analogous to the mass term in the Klein–Gordon equation, giving

$$m_H^2 = 2\lambda\nu^2 = -2\mu^2. \tag{2.5}$$

Consequently, the VEV, measured to be approximately 246.22 GeV [5], can be expressed in terms of the fundamental parameters μ and λ as

$$\nu = \sqrt{-\frac{\mu^2}{\lambda}}. \tag{2.6}$$

The cubic and quartic H-terms describe the Higgs self-interactions, corresponding respectively to three-point and four-point vertices. Substituting λ from Equation 2.5 into the

[1] Its physical interpretation remains unresolved, as its magnitude of order 10^{-61} vastly exceeds any observed cosmological energy scale.

potential (Equation 2.4) gives its form in dependence of the Higgs mass as

$$V(H(x), m_H) = -\frac{1}{8}\nu^2 m_H^2 + \frac{1}{2}m_H^2 H^2 + \frac{1}{2}\frac{m_H^2}{\nu}H^3 + \frac{1}{8}\frac{m_H^2}{\nu^2}H^4. \tag{2.7}$$

The Tau Yukawa Coupling

The SM formulated using only $\mathcal{L}_{\text{fermion}}$ and $\mathcal{L}_{\text{gauge}}$ cannot generate massive particles, which are, however, observed experimentally. To account for particle masses, the BEH mechanism introduced above postulates a Higgs doublet ϕ, which contains the physical Higgs field H in the form $\phi = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \nu + H \end{pmatrix}$. The coupling between fermions and the Higgs boson, known as the Yukawa coupling after Hideki Yukawa, can be expressed as follows, as illustrated by the example of a down type quark d below. There, ψ_L denotes the left-handed doublet, while d_R refers to the right-handed singlet. The Yukawa constants c_f are new constants of nature, one for each fermion f, and determine the strength of the Yukawa interaction. In contrast to the gauge couplings, the Yukawa couplings change the chirality of fermions:

$$\begin{aligned}
\mathcal{L}_{\text{Yuk}} &= -c_d(\bar{\psi}_L \phi d_R + \bar{d}_R \phi^\dagger \psi_L) = -c_d \bar{\psi}_L \phi d_R + \text{h.c.} \\
&\overset{\text{BEH}}{=} -c_d(\bar{\psi}_L(\nu + H)d_R + \bar{d}_R(\nu + H)\psi_L) \\
&= -\frac{c_d \nu}{\sqrt{2}}\Big(\bar{d}_L d_R + \bar{d}_R d_L\Big) - \frac{c_d H}{\sqrt{2}}\underbrace{\Big(\bar{d}_L d_R + \bar{d}_R d_L\Big)}_{\text{Higgs int.: } H d\bar{d}} \\
&\overset{m_d = c_d \frac{\nu}{\sqrt{2}}}{=} -m_d \bar{d}d - \frac{m_d}{\nu} H \bar{d}d
\end{aligned} \tag{2.8}$$

As shown, fermion masses $m_f = c_f \frac{\nu}{\sqrt{2}}$ arise naturally alongside the Yukawa couplings. For up-type quarks, the charge-conjugated Higgs doublet $\phi_c = -i\sigma_2 \phi^* = \frac{1}{\sqrt{2}}\begin{pmatrix} \nu + H \\ 0 \end{pmatrix}$ is used instead.

Combining the describing effects, the complete SM Lagrangian takes the form of:

$$\mathcal{L}_{SM} = \mathcal{L}_{fermion} + \mathcal{L}_{gauge} + \mathcal{L}_{Yukawa} + \mathcal{L}_{Higgs} \tag{2.9}$$

Here, $\mathcal{L}_{\text{gauge}}$ collects all $\mathcal{L}_V$ terms associated with the three gauge fields B, W, and G, as summarised in Table 2.1. The Higgs contribution is written as $\mathcal{L}_{\text{Higgs}} = \mathcal{L}_\phi$, where the notation has been adjusted for improved readability of Equation 2.9. Finally, $\mathcal{L}_{\text{fermion}}$ and $\mathcal{L}_{\text{Yukawa}}$ denote the corresponding terms introduced above.

In summary, the gauge groups and their associated charges, along with the resulting particle content introduced above, as well as the postulated Higgs boson, are brought together in Figure 2.2, which offers a concise overview of the SM.

2.1.1 Quarks and Mesons

Mesons, composed of quark-antiquark pairs, constitute one of the primary observable signatures in the Higgs boson decay process studied in this thesis. Therefore, the following discussion first focuses on quarks, before considering the role of mesons specific to this analysis.

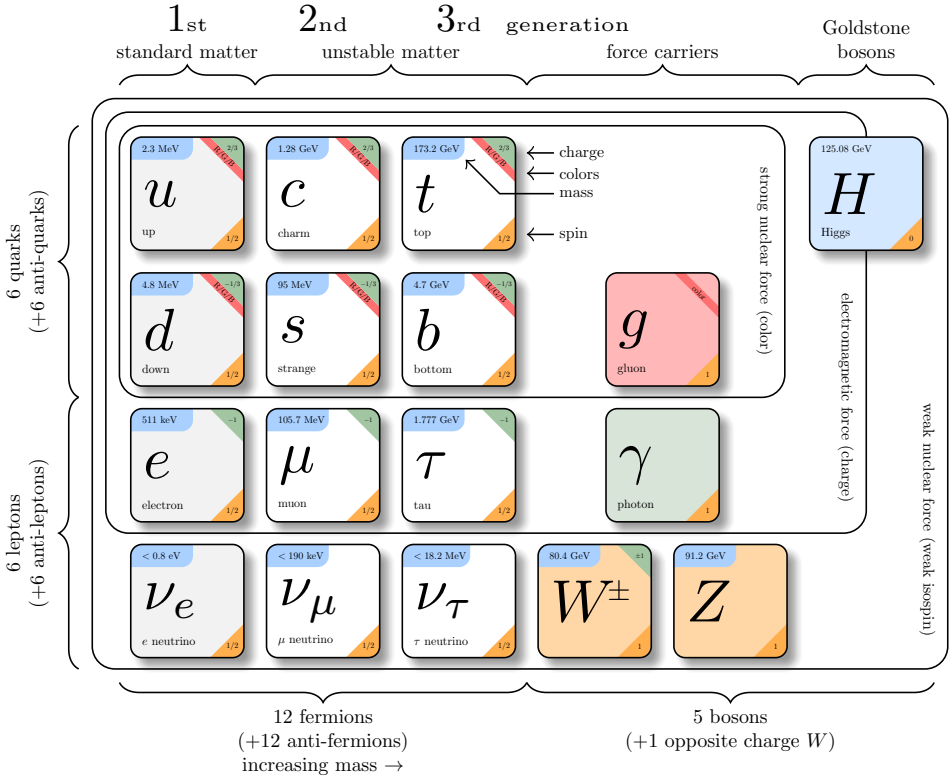


Figure 2.2: Standard Model of fundamental particles. Fermions are separated into quarks (upper left) and leptons (lower left), while bosons are divided into gauge and scalar bosons (right) [1]. Modelled after [6].

Quarks

The six quarks are fundamental constituents of matter in the SM, as illustrated in Figure 2.2. They are classified according to their electric charge. Up-type quarks: up (u), charm (c), and top (t), carry a charge of $+\frac{2}{3}e$, whereas down-type quarks: down (d), strange (s), and bottom/ beauty (b), have a charge of $-\frac{1}{3}e$. e denotes the elementary charge of the electron, 0.303 in natural units or approximately 1.603×10^{-19} C in SI units. All quarks are fermions with a half-integer spin. They interact via all fundamental gauge forces: the strong force, mediated by gluons, for which they carry a colour charge (red, blue, or green); the electromagnetic force, due to their electric charge; and the weak force, responsible for flavour-changing processes. Each quark also has mass, with the heaviest masses arising from the strongest coupling to the Higgs boson (see subsection 2.1.5). Quarks of the first generation (u, d) form the stable matter that makes up the visible universe. In contrast, heavier quarks are unstable and decay into lighter ones. Every quark has a corresponding antiquark, which carries opposite electric and colour charges.

Mesons

Mesons are colour-neutral quark-antiquark pairs ($q\bar{q}'$), bound by the strong force, whose spin determines their classification. The combination of two spin-$\frac{1}{2}$ quarks results in a total spin of 0 or 1. Pseudo-scalar mesons, such as the pion (π), have spin 0, whereas

vector mesons, like the ρ resonance, and axial-vector mesons, such as the a_1 resonance, have spin 1 and differ in parity (see section 2.2). Depending on their quark content, mesons can participate in all three gauge interactions.
Pions are the lightest mesons. While the charged states predominantly decay into muons and their corresponding neutrinos, the neutral pion mainly decays into a photon pair. The ρ and a_1 resonances are heavier excited states and thus unstable, decaying predominantly into two and three pions, respectively. All three mesons, $m \in \{\pi, \rho, a_1\}$, are composed from up- and down quarks, such that $m^+ = u\bar{d}$, $m^- = \bar{u}d$, and $m^0 = \frac{1}{\sqrt{2}}(u\bar{u} - d\bar{d})$. The differences in mass between them arise from their spin-parity configurations and the dynamics of the strong interaction. Pions, as pseudo-scalar mesons, are the lightest due to their approximate nature as pseudo-Goldstone bosons of chiral symmetry breaking. In contrast, vector and axial-vector mesons are heavier, reflecting the additional binding energy associated with spin alignment and excitation states in the quark-antiquark system. Their precise masses, composition, J^{PC} quantum numbers, dominant decay channels, and branching ratios (BR) are summarised in Table 2.2.

Table 2.2: Main properties as mass and quark content of pions (π), ρ and a_1 resonances. Spin J (setting $L = 0$), parity P, and charge-conjugation for neutral states C quantum numbers are given. Dominant decay modes, and branching ratios from the hadronically decaying tau lepton (τ^-) decay (BR) [1] are included.

Meson	**Mass [MeV]**	**Quark content**	J	P	C	**Dominant decays**	**BR [%]**
π^+	139.570 39(18)	$u\bar{d}$	0	+		$\pi^+ \to \mu^+\nu_\mu$	
π^-	139.570 39(18)	$\bar{u}d$	0	−		$\pi^- \to \mu^-\bar{\nu}_\mu$	10.82(5)
π^0	134.9768(5)	$\frac{1}{\sqrt{2}}(u\bar{u} - d\bar{d})$	0	−	+	$\pi^0 \to \gamma\gamma/\gamma e^+e^-$	
ρ^+	775.11(34)	$u\bar{d}$	1	−		$\rho^+ \to \pi^+\pi^0$	
ρ^-	775.11(34)	$\bar{u}d$	1	−		$\rho^- \to \pi^-\pi^0$	25.49(9)
ρ^0	775.26(23)	$\frac{1}{\sqrt{2}}(u\bar{u} - d\bar{d})$	1	−	−	$\rho^0 \to \pi^\pm\pi^\mp$	
a_1^+	1230(40)	$u\bar{d}$	1	+		$a_1^+ \to \rho^{+,0}\pi^{0,+} \to (3\pi)^+$	
a_1^-	1230(40)	$\bar{u}d$	1	+		$a_1^- \to \rho^{-,0}\pi^{0,-} \to (3\pi)^-$	9.31(5)
a_1^0	1230(40)	$\frac{1}{\sqrt{2}}(u\bar{u} - d\bar{d})$	1	+	+	$a_1^0 \to (3\pi)^0$	

In collider analyses, the number of *prongs* refers to the number of charged decay products reconstructed in the tracking detector. Consequently, the $\tau^- \to \pi^-\nu_\tau$ and $\tau^- \to \pi^-\pi^0\nu_\tau$ modes are both classified as 1-prong decays. Although the latter involves two mesons, only the charged pion (π^-) leaves a track, while the neutral pion (π^0) decays into photons detected in the calorimeter. Analogously, the decay $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ ($a_1^{3pr}(1260)$) involves three charged tracks and is therefore referred to as a 3-prong decay, indicated by the superscript ^{3pr}.

2.1.2 Lepton Universality

The SM contains six leptons: the tau lepton, muon, and electron, where the latter two are the lighter counterparts of the first, and their respective nearly massless neutrinos.[2] Each lepton is associated with a corresponding neutrino. No difference in the lepton flavours

[2]The non-neutrino lepton masses are $m_e = 0.510\,998\,950\,00(15)\,\text{MeV}$, $m_\mu = 105.658\,375\,5(23)\,\text{MeV}$, and $m_\tau = 1776.93(9)\,\text{MeV}$ [1].

(e, μ, τ) in weak interactions is postulated by the SM. In other words, the coupling of any lepton to a W or Z boson is identical, which can be expressed mathematically as:

$$\frac{g_W^e}{g_W^\mu} = \frac{g_W^\mu}{g_W^\tau} = 1 \tag{2.10}$$

Precision measurements are ongoing [7]. In accordance with the current results, Lepton Universality (LU) is assumed throughout this thesis. In the context of the $H \to \tau_\ell \tau_{had}$ decay, LU implies that, aside from the effects of the different lepton masses on the decay kinematics and phase-space, the analysis results are expected to remain unchanged when the lepton flavour is varied.

2.1.3 The Tau Lepton

The tau lepton, the heaviest among the leptons, is central to this thesis. With a mass of 1776.93(9) MeV, it is the only lepton capable of both hadronic and leptonic decays. Therefore, the third-generation lepton has an extremely short lifetime τ_τ of $290.0(5) \times 10^{-15}$ s [1], resulting in an almost immediate decay. According to the TauPOG[3] recommendation, the minimum momentum a tau lepton can have is 20 GeV to maintain a good reconstruction efficiency of its decay products. Thus, its minimal decay length λ is approximately

$$\lambda = c\tau_\tau \ \beta\gamma = c\tau_\tau \ \frac{p}{m} \approx (3 \times 10^8 \,\mathrm{m/s} \cdot 3 \times 10^{-13} \,\mathrm{s}) \frac{20 \,\mathrm{GeV}}{1.8 \,\mathrm{GeV}} \approx 1 \,\mathrm{mm}.$$

The tau lepton exhibits a rich decay structure: Leptonic modes, $\tau \to e\nu_e\nu_\tau$ and $\tau \to \mu\nu_\mu\nu_\tau$, account for about 17.8 % and 17.4 % of all decays, respectively. The remaining 64.8 % proceed through hadronic channels, in which the tau lepton decays into one or more mesons and a neutrino. The dominant hadronic modes include $\tau \to \pi\nu_\tau$ (10.8 %), $\tau \to \pi\pi^0\nu_\tau$ and higher resonances (46.9 %), as well as multi-pion final states such as $\pi\pi^+\pi^-\nu_\tau$ (9.3 %) [1]. These modes are illustrated in Figure 2.3.
For the following discussion of the Yukawa coupling, it is useful to begin by looking at the concepts of *helicity* λ and *chirality* $P_{L/R}$.

Helicity λ

Helicity λ denotes the spin component parallel to the momentum of the particle under study

$$\lambda = \frac{\vec{S}\vec{p}}{|\vec{p}|} = \frac{1}{2|\vec{p}|} \begin{pmatrix} \vec{\sigma}\vec{p} & 0 \\ 0 & \vec{\sigma}\vec{p} \end{pmatrix}, \tag{2.11}$$

where $\vec{S}$ is the spin operator and $\vec{\sigma}$ denotes the vector of Pauli matrices.
For spin-$\frac{1}{2}$ fermions, helicity takes the eigenvalues $\pm\frac{1}{2}$, corresponding to right- (RH) and left-handed (LH) states. Dirac spinors $u(p)$ and $v(p)$, describing particle and antiparticle solutions of the Dirac equation, each contain two independent spin components, spin-up and spin-down, which become helicity eigenstates when the spin is aligned or anti-aligned with the momentum. Choosing the momentum along the z-axis and working in the Dirac representation reduces the helicity operator to $\lambda_z = \frac{1}{2}\mathrm{diag}(1, -1, -1, 1)$, acting on the

[3] In CMS, a Physics Object Group (POG) is a group responsible for the reconstruction algorithms, calibration, and performance validation of a given physics object, e.g., electrons, muons, jets, tau leptons, etc., ensuring quality and consistency across analyses.

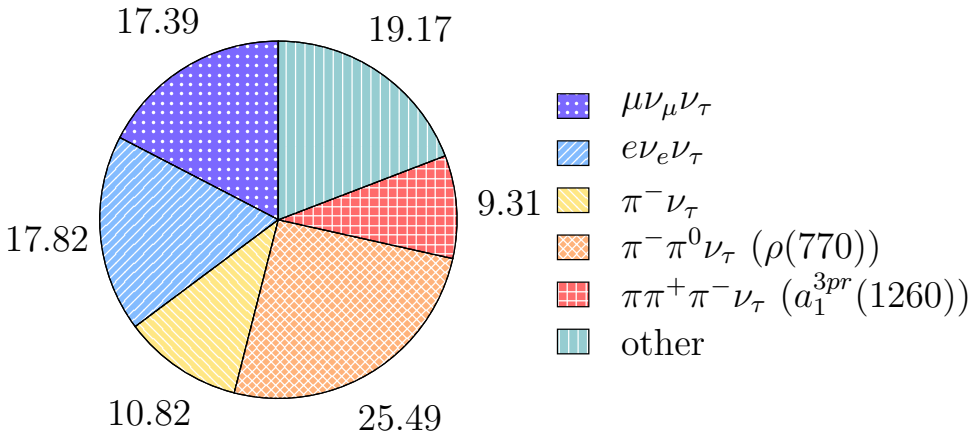


Figure 2.3: Branching fractions of the dominant τ^- decay modes. About 35% of tau leptons decay leptonically, while the hadronic modes ($\approx 65\%$) proceed mainly via the $\rho(770)$ and $a_1(1260)$ resonances, leading to $\pi^-\pi^0\nu_\tau$ and $\pi^-\pi^+\pi^-\nu_\tau$ final states. Subdominant channels are grouped as *other* [1].

four components of the Dirac spinors. The corresponding eigenvalue equations reflect the opposite helicity assignment of particle and antiparticle states:

$$\lambda_z u^{(1)}(p_z) = +\tfrac{1}{2}u^{(1)}(p_z), \quad \lambda_z u^{(2)}(p_z) = -\tfrac{1}{2}u^{(2)}(p_z)$$
$$\lambda_z v^{(1)}(p_z) = -\tfrac{1}{2}v^{(1)}(p_z), \quad \lambda_z v^{(2)}(p_z) = +\tfrac{1}{2}v^{(2)}(p_z)$$

Chirality $P_{L/R}$

The chirality operator is defined as

$$P_{L/R} = \frac{1}{2}(1 \mp \gamma^5), \quad \text{where } \gamma^5 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{2.12}$$

Here, γ^5 is the product of the Dirac matrices one to four. Applying $P_{L/R}$ to a spinor projects it onto its left- or right-handed component. Chirality is conserved in weak decay processes within the SM, so the weak charged current couples exclusively to LH particles and RH antiparticles. Because neutrinos have a near-zero mass, helicity and chirality are approximately equal. So far, only LH neutrinos and RH antineutrinos have been found to exist [8, 9].

Polarisation in the Tau Lepton Decay

The helicity and chirality of particles determine the kinematics of a hadronically decaying tau lepton, which is called a hadronic tau. While the same considerations apply to leptonic decays, the experimental challenges involved are generally less demanding. In contrast, the three-prong hadronic decay considered here is considerably more complex owing to its multi-body and hadronic nature. Consequently, the decay exhibits a more intricate geometry within the detector and the presence of an additional neutrino further complicates its reconstruction.

In the rest frame of the negatively charged tau lepton, the particle itself has no net momentum but retains its intrinsic spin. In consequence of momentum conservation, the decaying neutrino and a_1 meson decay back-to-back. For the sake of argument, all tau leptons in Figure 2.4b are depicted with an illustrative momentum direction from left to right.

Since neutrinos always possess negative helicity, their spin is anti-aligned with their momentum, pointing towards the tau lepton. Depending on the orientation of the tau spin relative to the neutrino spin, the total spin projection of the $\tau^- + \nu_\tau$ system can add up to either 0 or 1. Consequently, the polarisation of the spin-1 a_1 meson aligns such that the total angular momentum is conserved. If the tau spin is approximately aligned with the neutrino spin (Figure 2.4a, Figure 2.4d), the meson is transversely polarised. If the spins point in opposite directions (Figure 2.4b, Figure 2.4c), the meson finds itself longitudinally polarised.
For the charge-conjugate case of a τ^+ lepton decay, the spins, helicities, and chiralities are reversed, whereas the momenta remain unchanged. Consequently, the spin and polarisation configurations are mirrored with respect to those of the τ^-.
These configurations are illustrated in the schematic drawing 2.4. From left to right, the tau spin flips while its momentum direction remains fixed. From the top to the bottom line, the angular distribution of the decay products is mirrored, which in turn influences the meson's polarisation.

2.1.4 Spin, Polarisation, and its Propagation

The CP nature of the Higgs boson, subject of this study, can be probed through the spin correlations of its decay products. These correlations influence the angular distributions of the final-state particles, which can be measured experimentally. To interpret these distributions, it is essential to understand how the spin information of the initial boson is transferred to its decay products, a process characterised through their polarisation.

Spin is an intrinsic quantum property of particles, independent of their spatial motion. It is described through spin states, which are eigenstates of the spin projection operator S_z. For a spin-$\frac{1}{2}$ particle, the spin operator is $\vec{S} = \frac{1}{2} diag(\vec{\sigma}, \vec{\sigma})$, where $\vec{\sigma}$ denotes the Pauli matrices. The eigenstates of S_z satisfy $S_z u^{(1/2)}(p_z) = \pm\frac{1}{2} u^{(1/2)}(p_z)$, and correspond to the conventional spin-up and spin-down configurations. Helicity, by contrast, depends on the particle's kinematics. For massless or ultrarelativistic particles, it becomes a convenient label for describing spin in interactions, as it remains invariant under boosts along the momentum direction. In this limit, spin-up and spin-down states coincide with positive and negative helicity states, respectively. The total angular momentum $\vec{J} = \vec{L} + \vec{S}$ combines the orbital and intrinsic contributions and is conserved in all interactions by construction. Polarisation then provides a practical means of characterising how the spin degrees of freedom are distributed in production and decay processes.

The *polarisation* of a particle quantifies the orientation of its spin relative to a chosen axis, commonly, the particle's momentum direction is chosen. It can be decomposed into a longitudinal and a transverse component,

$$P_L^\tau = \langle 2s_z \rangle = \frac{N_{\mathrm{RH}} - N_{\mathrm{LH}}}{N_{\mathrm{RH}} + N_{\mathrm{LH}}}, \tag{2.13}$$

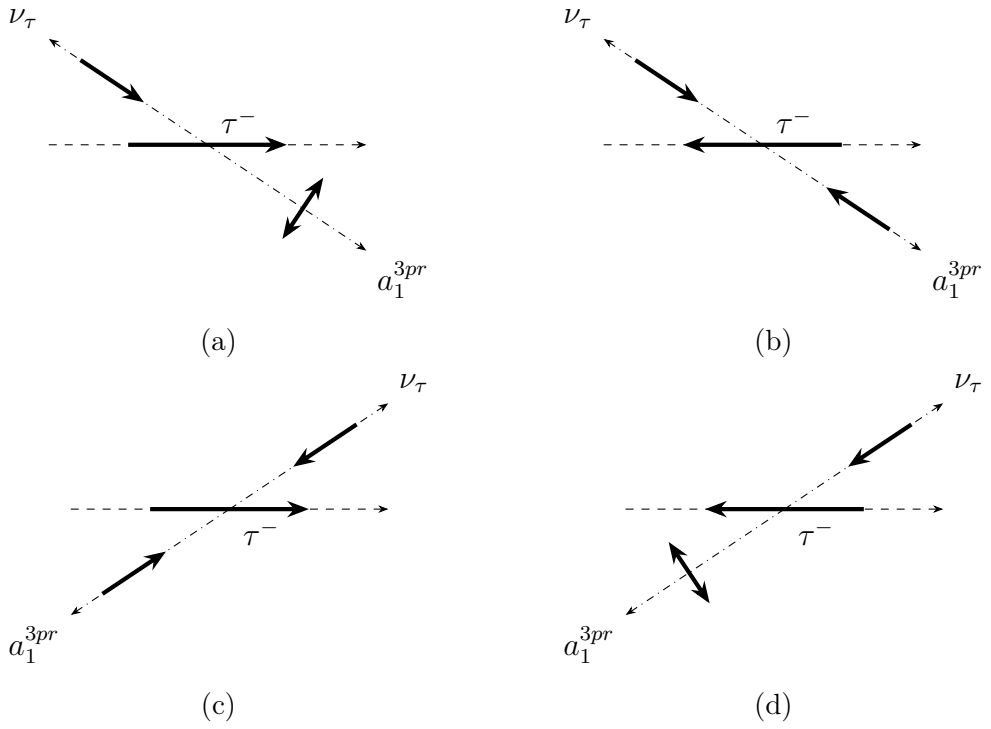


Figure 2.4: Schematic representation of the spin and momentum configurations in the $\tau^- \to a_1^{-,3pr}\nu_\tau$ decay. Momenta are depicted as thin arrows, and spin and polarisation as thick ones. In all four cases, the tau lepton momentum points in the same direction. From left to right, the tau spin flips. From top to bottom, the kinematic distribution of the decay products is mirrored. The neutrino ν_τ always carries negative helicity (spin opposite to its momentum). When the tau spin is aligned with the neutrino spin, the total spin of the $\tau\nu_\tau$ system adds up to 1, resulting in a transversely polarised (spin 0) a_1^- meson, which conserves the spin 0 nature of the overall $\tau\tau$, respectively H, system. Likewise, when the two spins are anti-aligned, the total spin is 0, leading to a longitudinally polarised (spin 1) a_1^- meson. Modelled after [10].

$$P_T^\tau = \sqrt{\langle s_x \rangle^2 + \langle s_y \rangle^2}, \tag{2.14}$$

where N_{RH} and N_{LH} denote the numbers of RH and LH tau leptons, respectively. Fully RH (LH) tau states correspond to a scalar (pseudoscalar) Higgs boson, and thus to a CP-even (CP-odd) Yukawa interaction, because the scalar coupling $\tau\bar{\tau}$ preserves parity, whereas the pseudoscalar coupling $\tau i\gamma^5\bar{\tau}$ flips it. Further details can be found in section 2.2 and [11–13].

Since the Higgs boson has a spin of zero, the di-tau system emerging from its decay must conserve this property. Consequently, spin correlations arise between the two tau leptons. Each tau spin can be decomposed into a longitudinal component, $\vec{s}_L$, along its momentum, and a transverse one, $\vec{s}_T$, perpendicular to it. Because the Higgs boson is isotropic, correlations involving the longitudinal spin components average out over the full phase space. The remaining transverse components, however, define decay planes for each tau lepton, whose relative azimuthal orientation carries information about the CP structure of the interaction. The resulting angular dependence appears in the squared matrix element, which is proportional to the differential decay rate $d\Gamma|_{H\to\tau\tau}$, experimentally

accessible, and named ϕ_{CP}.

$$d\Gamma|_{H\to\tau\tau} \propto 1 - \underbrace{(\vec{s}_T^{\,\tau^+} \cdot \vec{s}_T^{\,\tau^-})}_{\sim\cos(\phi_{CP})} \cos\big(2\alpha^{H\tau\tau}\big) - \tau \underbrace{(\vec{s}_T^{\,\tau^+} \times \vec{s}_T^{\,\tau^-})}_{\sim\sin(\phi_{CP})} \sin\big(2\alpha^{H\tau\tau}\big), \tag{2.15}$$

here $\vec{s}_T^{\,\tau^\pm}$ denote the transverse tau spin vectors in their respective rest frames. The first cosine term represents the parity-even (CP-even) correlation as it is invariant under a parity transformation and therefore also under a CP transformation. The sine term describes the parity-odd (CP-odd) correlation, which geometrically corresponding to the orientation and rotation of the two transverse spin vectors in their decay planes. The parameter $\alpha^{H\tau\tau}$ encodes the CP structure of the Tau Yukawa interaction, as discussed in the following. Experimentally, the differential decay rate is studied as a function of the acoplanarity angle ϕ_{CP}:

$$\frac{d\Gamma}{d\phi_{CP}} \sim 1 - \frac{\pi^2}{16} b(\tau^+) b(\tau^-) \cos\big(\phi_{CP} - 2\alpha^{H\tau\tau}\big), \tag{2.16}$$

where the prefactor includes the spectral functions $b(\tau^\pm)$ of both tau leptons. They quantify how strongly the momentum distribution of the decay products reflects the parent–tau spin orientation, for instance, $b = 1$ for purely hadronic $\tau^\pm \to \pi^\pm \nu_\tau$ decays and smaller values for multi-body or leptonic modes such as $\tau^\pm \to a_1^\pm \nu_\tau \to 3\pi^\pm \nu_\tau$ due to the internal correlations among the three pions.
Because the CP structure of the interaction determines the relative strength of scalar and pseudoscalar contributions to the tau Yukawa coupling, the decay rate can be parametrised via

$$\kappa_\tau = \cos\big(\alpha^{H\tau\tau}\big) \qquad \text{and} \qquad \tilde{\kappa}_\tau = \sin\big(\alpha^{H\tau\tau}\big)$$

assuming $\gamma_\tau \gg 1$ and $\beta_\tau \approx 1$. Thereby, the tau Yukawa Lagrangian from Equation 2.8 can be rewritten as

$$\mathcal{L}_{Yuk}^{H\tau\tau} = -\frac{m_\tau}{v} \bar{\tau}\big(\kappa_\tau + i\gamma^5 \tilde{\kappa}_\tau\big)\tau H, \tag{2.17}$$

where the ratio of the pseudoscalar to scalar coupling constants defines the *CP-mixing angle*

$$\alpha^{H\tau\tau} = \arctan\left(\frac{\tilde{\kappa}_\tau}{\kappa_\tau}\right) = \begin{cases} 0 & : \text{scalar (CP-even)} \\ \frac{\pi}{2} & : \text{pseudoscalar (CP-odd)} \\ \frac{\pi}{4} & : \text{mixed state (CP-maximal mixing)} \\ \text{other} & : \text{mixed state} \end{cases}. \tag{2.18}$$

The angle therefore characterises the CP nature of the Higgs-tau interaction such that $\alpha^{H\tau\tau} = 0$ corresponds to a purely scalar coupling, $\alpha^{H\tau\tau} = \pi/2$ to a purely pseudoscalar one, and any intermediate value signifies a CP-mixed state. The corresponding theoretical predictions for the CP-sensitive observable ϕ_{CP} are shown in Figure 2.5 [14]. In all following plots, the SM CP-even distribution is depicted in red, the CP-maximally mixed state in green, and the BSM CP-odd state in blue. As $\alpha^{H\tau\tau}$ is set to zero for the SM hypothesis. Therefore, it can be seen as a shift, from this null-hypothesis so other states. For comparison, the distribution for Z boson decays into the same final state is also shown. As expected, it is flat, reflecting the absence of any CP-sensitive spin correlations in this process and thus serving as a useful control channel.

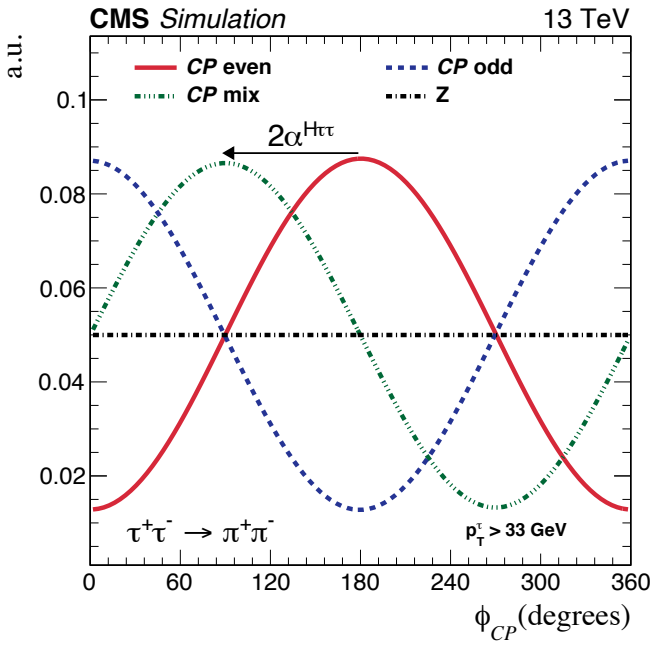


Figure 2.5: Theoretical ϕ_{CP} distributions for a CP-even (red), CP-odd (blue), and CP-maximally mixed (green) Higgs boson coupling to tau leptons. A non-zero CP-mixing angle $\alpha^{H\tau\tau}$ manifests itself as a phase shift of the distribution with respect to the SM hypothesis. The distribution for $Z \to \tau\tau$ decays (black) is flat, reflecting the absence of CP-sensitive spin correlations in this process [14].

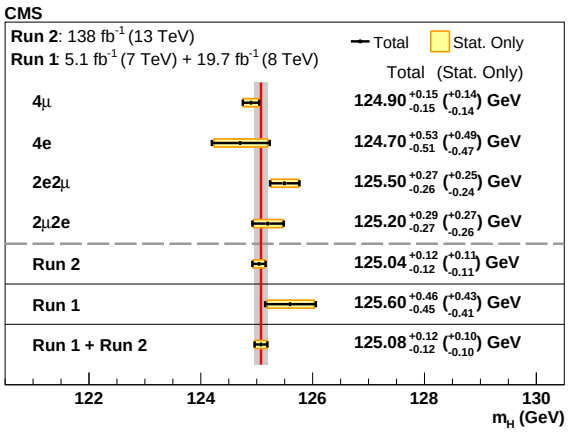
In summary, measuring the coplanarity angle ϕ_{CP} provides a direct experimental handle on the CP structure of the tau Yukawa coupling through the respective spin polarisations. The longitudinal polarisation component constrains the overall scalar or pseudoscalar nature of the interaction, whereas the transverse component reveals the CP-violating phase, together offering a comprehensive picture of the Higgs boson's spin and CP properties. In decays of the Higgs boson to tau lepton pairs, the relative orientation of the tau spins reflects the Higgs boson's underlying CP nature. This correlation determines the spin–momentum structure of the subsequent $\tau \to a_1^{3pr}\nu$ decays and how the spin information is transferred to observable quantities. As a result, the angular correlations among the visible decay products provide direct sensitivity to the CP-mixing angle $\alpha^{H\tau\tau}$.

2.1.5 The Higgs Boson at the LHC

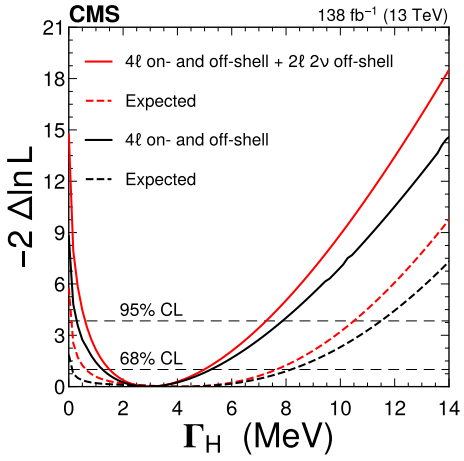
The Higgs boson, the most recent addition to the SM, inaugurates the class of scalar bosons and thereby introduces a new type of interaction. Today, it provides a long-awaited gateway to the discovery of new physics, either through precision measurements, as in the present work, or via potential extensions of the SM.

Before 1964, when it was postulated independently by Peter W. Higgs, Robert Brout, and François Englert [15–17], as well as by Thomas W. B. Kibble, Carl R. Hagen, and Gerald Guralnik [18], the SM gauge theory forbade the W and Z bosons from acquiring mass, while fermions also remained massless through the $SU(2)_L$ interaction. This clearly disagreed with the experimentally observed non-zero masses. Two solutions were proposed:

the first and later excluded [15], was a composite theory, postulating that these particles are not fundamental. The second relied on spontaneous symmetry breaking (SSB), introducing a new scalar field. This Goldstone boson was eventually observed 48 years later, in 2012, by the ATLAS and CMS collaborations [19, 20]. As shown in Figures 2.6a and 2.6b, its mass m_H is measured to be 125.08(12) GeV and its width to be $\Gamma_H = 3.0^{+2.0}_{-1.5}$MeV [21].



(a) Combination of multiple Higgs mass measurements, showing statistical (yellow bands) and total (black bars) uncertainties, with the red line and grey band indicating the Run 1 + Run 2 combined result and its total uncertainty [21].



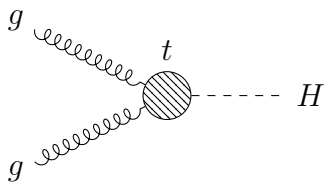
(b) Observed (solid) and expected (dashed) likelihood scans over the Higgs width Γ_H combining on-shell 4ℓ data with off-shell 4ℓ (black), and both added with $2\ell 2\nu$ (red), showing 68 % and 95 % CL regions and excluding the no off-shell hypothesis at $3.6\,\sigma$ [21].

Figure 2.6: Combined results of the Higgs mass (a) and width (b) measurement with the CMS Experiment.

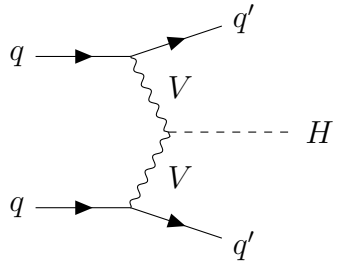
The most prominent Higgs production channels at the 13.6 TeV Large Hadron Collider (LHC) are ordered by their Higgs yield: gluon-gluon fusion (ggF), vector boson fusion (VBF), Higgs boson bremsstrahlung (VH), and top quark associated Higgs boson production (ttH) [22], as portrayed in Figure 2.7. From a phenomenological perspective, the key point is their contribution to the total Higgs yield, which determines how many Higgs bosons are available for study and the resulting statistical sensitivity in the μa_1^{3pr} decay channel.

In 13.6 TeV LHC collisions, ggF (see Figure 2.7a) represents the dominant Higgs production mechanism. This is primarily because gluons carry most of the proton momentum [23], making their interactions frequent. Although gluons are massless and do not couple directly to the Higgs boson, the interaction occurs mainly via a virtual top quark loop in which the top quark exists as an off-shell intermediate state. Because the Higgs coupling scales with mass, consequently the large top quark mass enhances the effective ggF coupling, resulting in the highest Higgs production cross section. VBF constitutes the next most significant production mode (see Figure 2.7b). Here, quarks from the colliding protons emit vector bosons, which subsequently fuse to produce a Higgs boson. The weak coupling of quarks to vector bosons is weaker than the strong vertices in ggF, and the overall production rate is further reduced due to the smaller Higgs–vector boson coupling

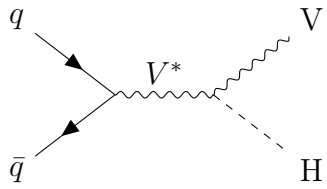
relative to the top Yukawa coupling. To provide a complete overview, VH and ttH are also included in Figures 2.7c and 2.7d. Although not studied further in this thesis, VH is phenomenologically similar to VBF, differing primarily by one fewer quark–vector-boson vertex, and therefore has a comparable but slightly reduced production cross section. The ttH process is suppressed at the LHC relative to ggF and VBF due to the high mass threshold required to produce top quarks.



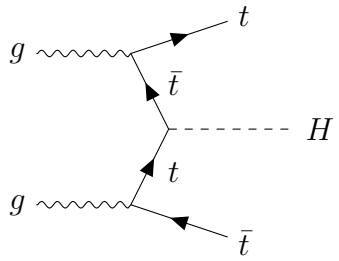
(a) Gluon-gluon fusion (ggF) via a virtual top quark loop.



(b) Vector boson fusion (VBF).



(c) Higgs boson bremsstrahlung (VH)/ radiation of W or Z bosons.



(d) Top quark associated Higgs boson production (ttH).

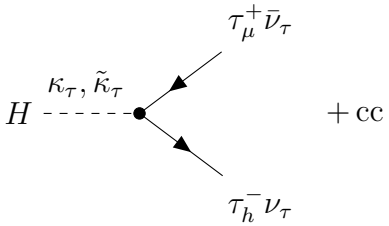
Figure 2.7: Representative Feynman diagrams for the main Higgs boson production mechanisms at the LHC.

To study the Higgs boson CP structure, the thesis focuses on the

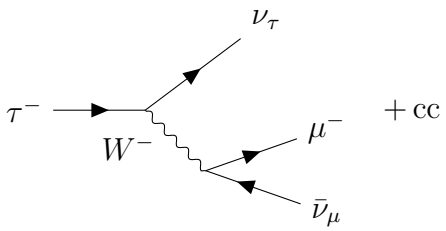
$$H \to \tau^+\tau^- \to (\mu^+\nu_\mu\bar{\nu}_\tau)(\pi^+\pi^-\pi^-\nu_\tau) + cc \equiv \mu a_1^{3pr} \tag{2.19}$$

decay chain. The Higgs bosons considered in this analysis are produced via ggF and VBF as described above.

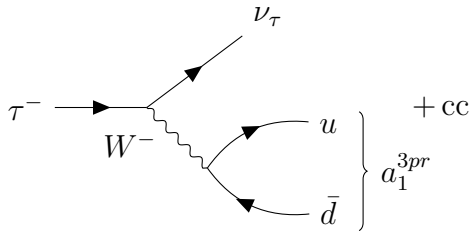
Higgs bosons most frequently decay into b quark pairs (see Figure 2.9), with a branching ratio (BR) of 58.2 %, where the spin information is largely lost during hadronisation. Decays into vector bosons yield the second largest BR. Although CP properties have been studied in Run 1 [24–26], these measurements probe mainly the bosonic couplings and do not directly constrain the Yukawa fermion sector. The latter is studied separately due to the different mass-generation mechanism for fermions and bosons. The first experimentally accessible channel following the BR is the Higgs boson decaying into a tau lepton pair, with a branching ratio of 6.27 %. This channel enables the measurement of the tau Yukawa coupling and, consequently, the CP structure of the Higgs boson, as illustrated in Figure 2.8a. From this point onwards, the focus of this thesis centres on the semi-leptonic decay mode of the di-tau system. Muons (see Figure 2.8b) are highly resolved in the muon chambers, giving the Compact Muon Solenoid (CMS) detector its name, and allows for precise momentum reconstruction. Meanwhile, the tau lepton (see Figure 2.8c) retains



(a) Higgs to di-tau lepton coupling with the reduced Yukawa coupling constants κ_τ (CP even, scalar Higgs boson) and $\tilde{\kappa}_\tau$ (CP odd, pseudoscalar Higgs boson) with two neutrinos in the final state.



(b) Leptonic tau lepton decay to a light lepton and its corresponding neutrino via a short lived W boson.



(c) Hadronic tau lepton decay via a W boson to a quark-anti quark pair, forming an a_1^{3pr} resonance.

Figure 2.8: Representative Feynman diagrams for the Higgs boson decay and the subsequent tau lepton decays of the decay chain under study. Shown are single charge configurations, for which their charge-conjugate processes (cc) are implied.

information about the Higgs spin and parity, which propagate to its decay products as described in subsection 2.1.4. Hence, the decay chain from Equation 2.19 provides a direct means of probing the Higgs CP properties.

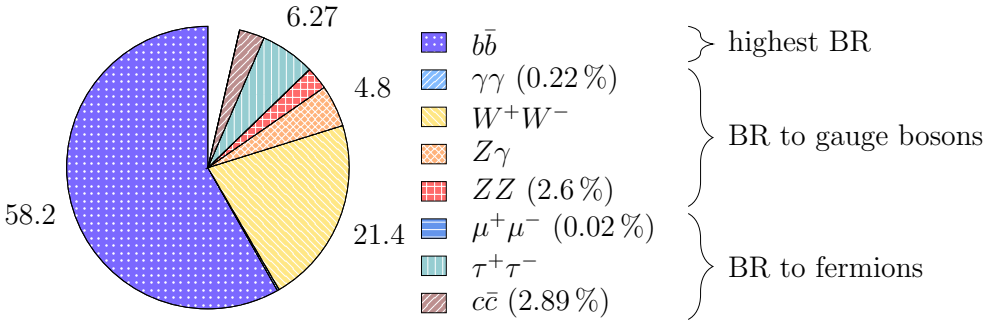


Figure 2.9: Branching fractions of the main Higgs boson decay modes. The dominant decay proceeds to $b\bar{b}$, while bosonic modes include W^+W^-, ZZ, $Z\gamma$ and $\gamma\gamma$. Among the fermionic channels, $H \to \tau^+\tau^-$ is experimentally relevant, followed by $H \to \mu^+\mu^-$ and $H \to c\bar{c}$, which are strongly suppressed due to the small Yukawa coupling [1].

2.2 Motivation for this CP Violation Analysis

At the time of the Big Bang, the universe was extremely hot and dense, containing equal amounts of matter and antimatter. Today, however, the visible universe predominantly consists of matter, an imbalance known as *baryon asymmetry*, which lacks a complete theoretical explanation. Nevertheless, it strongly suggests the existence of CP-violating mechanisms. In the SM, a small CP violation arises from the complex phase of the Cabibbo–Kobayashi–Maskawa (CKM) matrix, but this is insufficient to explain the observed asymmetry [27]. Therefore, studying CP violation in particle decays, such as those of the Higgs boson, thus not only probes new sources, but also tests the SM.

To establish the context for studying CP violation in Higgs decays, the theoretical concept of Charge-Parity-Time (CPT) symmetry, its components, and possible violations are discussed first. Then, it is useful to review its manifestations in mesonic systems. In subsection 2.2.1, the earliest experimental observation of CP symmetry breaking in neutral kaon decays is discussed. Subsection 2.2.2 then presents the formalism of CP violation in Yukawa couplings, which form the basis for studying Higgs decays into tau lepton pairs, both in the $\tau \to a_1^{3pr}\nu_\tau$ and $\tau \to \mu\nu_\mu\nu_\tau$ decay modes.

Although parity transformation (P) and time reversal (T) symmetry have already been formulated in classical mechanics, electrodynamics, and even classical gravity, nature itself is not invariant under these transformations. Macroscopic examples such as molecular chirality or the asymmetric placement of human organs demonstrate that parity is not a manifest symmetry of the world [28]. However, fundamentally, the microscopic origin of this breaking lies in the chiral structure of the weak interaction, which cannot be captured within a purely classical framework. Therefore, a consistent description necessarily requires the tools of relativistic quantum field theory.

The CPT theorem states that every quantum field theory that is local, Lorentz-invariant, and, formulated in a Hermitian Hamiltonian framework, must remain invariant under the combined application charge conjugation (C), P, and T.
A C transformation converts a particle into its corresponding antiparticle by reversing the sign of all its charges, whereas a P operation inverts spatial coordinates. In terms of Lorentz structure, this distinction becomes transparent. Vectors change sign under parity. Axial-/pseudovector currents remain parity-even. Geometrically, parity can be visualised as a spatial mirror reflection followed by a rotation of π. For spinors, these three internal symmetries act as

$$Cu^{(1/2)} = u_C^{(1/2)} = v^{(1/2)} \tag{2.20}$$

$$Pu^{(1/2)} = \gamma^0 u^{(1/2)}(E, \vec{p}) = +u^{(1/2)}(E, -\vec{p}) \tag{2.21}$$

$$Pv^{(1/2)} = \gamma^0 v^{(1/2)}(E, \vec{p}) = -v^{(1/2)}(E, -\vec{p}). \tag{2.22}$$

It follows that spin-$\frac{1}{2}$ fermions possess positive parity and spin-$\frac{1}{2}$ antifermions have negative parity.
Since CPT invariance is assumed to hold in all processes within the SM, the T transformation effectively compensates for the combined action of CP. The consequences of full CPT symmetry for experiments are that particles and antiparticles have equal masses, lifetimes and exactly symmetric charges. Tight upper limits have been set on any deviation from these predictions [1, 28].

2.2.1 Meson Sector

Each discrete symmetry appearing in the CPT theorem can, in principle, be violated separately. The first observation of such a violation occurred in 1957, when Wu and her team demonstrated parity (P) violation in the beta decay of polarised ^{60}Co nuclei. The measured angular asymmetry of the emitted electrons revealed maximal parity violation, since only left-handed neutrinos participate in weak interactions [28–30]. C violation was observed in weak decays of pions to muons, as the experiment [31] showed that replacing all particles by their corresponding antiparticles changes the decay probabilities. This reflects the chiral nature of the weak interaction, which couples only to LH fermions and RH antifermions. In 1964, studies of the neutral Kaon system provided evidence for combined CP violation. The flavour eigenstates K^0 and $\bar{K}^0$ mix to form two mass eigenstates, a long-lived and a short-lived one, denoted K_L^0 and K_S^0, which predominantly decay into three and two pions, respectively. Cronin and Fitch observed decays of the long-lived state into two pions, thereby demonstrating CP violation.
So far, no experimental evidence for CPT violation has been found, thereby extremely stringent limits have been established [1, 28], protecting the CPT theorem. Because of the experimental difficulty of reversing time evolution to test T symmetry directly, and since CPT invariance is generally assumed to hold, measurements of CP violation provide equivalent information on possible T-symmetry breaking.

2.2.2 Summary of the Chapter and Notation

The Yukawa interaction governs Higgs–fermion couplings and fermion mass generation. It also provides a framework to study potential CP-violating effects in the Higgs sector.

Inserting the Higgs doublet into the Yukawa Lagrangian (see Equation 2.8) yields the interaction between the Higgs field and fermions. Its strength is characterised by the Yukawa coupling constant c_f, leading to

$$\begin{aligned}\mathcal{L}_{Yuk} &= -\frac{c_f}{\sqrt{2}}\bar{\psi}_L(\nu + H)\psi_R - \frac{c_f}{\sqrt{2}}\bar{\psi}_R(\nu + H)\psi_L \\ &= -m_f(\bar{\psi}_L\psi_R - \bar{\psi}_R\psi_L) - \frac{m_f}{\nu}(\bar{\psi}_L H\psi_R + \bar{\psi}_R H\psi_L),\end{aligned}$$

where the fermion masses arise naturally as $m_f = \frac{c_f\nu}{\sqrt{2}}$. The linear dependence on H shows that the Higgs–fermion coupling strength scales with the fermion mass.

Next, the fermion fields are projected onto their chiral CP components, introducing the even and odd Yukawa coupling constants y_f and $\tilde{y}_f$, with f denoting any fermion:

$$\mathcal{L}_{Yuk} = -\bar{\psi}_f(y_f + i\gamma^5\tilde{y}_f)H\psi_f$$

Defining the reduced, dimensionless Yukawa couplings as

$$\kappa_f = \frac{\nu}{m_f}y_f \qquad \text{and} \qquad \tilde{\kappa}_f = \frac{\nu}{m_f}\tilde{y}_f$$

the Yukawa Lagrangian can be expressed in the following form

$$\mathcal{L}_{Yuk} = -\frac{m_f}{\nu}\,\bar{\psi}_f(\kappa_f + i\gamma^5\tilde{\kappa}_f)H\psi_f.$$

From the reduced Yukawa coupling, the CP coupling behaviour can be examined via the CP-mixing angle

$$\alpha^{H\tau\tau} = \arctan\left(\frac{\tilde{\kappa}_\tau}{\kappa_\tau}\right) = \begin{cases} 0 & : \text{scalar (CP-even)} \\ \frac{\pi}{2} & : \text{pseudoscalar (CP-odd)} \\ \frac{\pi}{4} & : \text{mixed state (CP-maximal mixing)} \\ \text{other} & : \text{mixed state} \end{cases},$$

which in turn depends on the decay rate per angle between the decay planes, denoted by ϕ_{CP}

$$\frac{d\Gamma}{d\phi_{CP}} \sim 1 - \frac{\pi^2}{16} b(\tau^+) b(\tau^-) \cos\big(\phi_{CP} - 2\alpha^{H\tau\tau}\big).$$

Detailed explanations of the methods used to ascertain these can be found in section 4.4.

Chapter 3

Experimental Setup

3.1 The Large Hadron Collider 24
3.2 The CMS Experiment 26
3.2.1 Coordinate System 27
3.2.2 Silicon Tracker 28
3.2.3 Calorimetry 29
3.2.4 Solenoid 30
3.2.5 Muon Chambers 30
3.2.6 Trigger System 31
3.3 Particle Reconstruction 31
3.3.1 Track Reconstruction 32
3.3.2 Vertex Reconstruction 33
3.3.3 Muon Reconstruction 34
3.3.4 Jet Reconstruction 34
3.3.5 Missing Transverse Momentum Reconstruction 35
3.3.6 Tau Lepton Reconstruction 35

In the late 1940s and early 1950s, the European continent sought to unify its scientific communities and pool national resources in order to secure a leading position in research in particle physics, especially after many scientists had fled during the Second World War. As a result, the European Organisation for Nuclear Research (CERN) complex was proposed and eventually constructed. Today, it embodies freedom, cooperation, and scientific diplomacy among its 25 Member States and a global collaboration of over 600 institutes and universities. Incidentally, many technical achievements are being made that benefit the economy and deliver numerous advances in areas such as medical technology.

This chapter sets out the experimental environment in which the analysis is conducted. The LHC provides high-energy proton–proton (pp) collisions, which are recorded at four main interaction points by the ATLAS, LHCb, ALICE, and CMS experiments. This thesis is based on data collected by the CMS detector in 2022 and 2023, its description therefore follows in this chapter. On a smaller scale, the chapter covers the relevant methods

for reconstructing particles. These interpret detector signals, forming an essential bridge between raw experimental data and the refined observables used in the final analysis.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the most powerful synchrotron ever constructed to study the fundamental laws governing the universe. Situated on the Swiss–French border and operated by CERN, it is housed in a 27 km-long circular tunnel approximately 100 m underground. It was designed to collide composite particles, hadrons, at velocities close to the speed of light at a centre-of-mass energy of $\sqrt{s} = 14\,\text{TeV}$ and an instant luminosity L of the order of $10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$. During the main operation periods, pp collisions dominate. Elsewise, heavy-ion collisions, most often using lead nuclei but also including gold, oxygen, or neon, are performed in dedicated runs to study the conditions of the early universe.

For operation, protons are extracted from a hydrogen source and initially accelerated to 90 keV. They enter the LINAC 2, where they are further accelerated to 50 MeV. Subsequently, their energy is increased step by step to 1.4 GeV in the Proton Synchrotron Booster (PSB), and to 25 GeV in the Proton Synchrotron (PS), where they are grouped into bunches containing about 1.2×10^{11} protons each. These bunches are then accelerated to 450 GeV in the Super Proton Synchrotron (SPS) before being injected into the two opposite LHC beam pipes. Inside the LHC ring, they are further accelerated to the desired collision energy and finally brought into collision at one of the four main interaction points of the LHC shown in Figure 3.1. After the collision moment, the bunches continue to circulate until dumped.

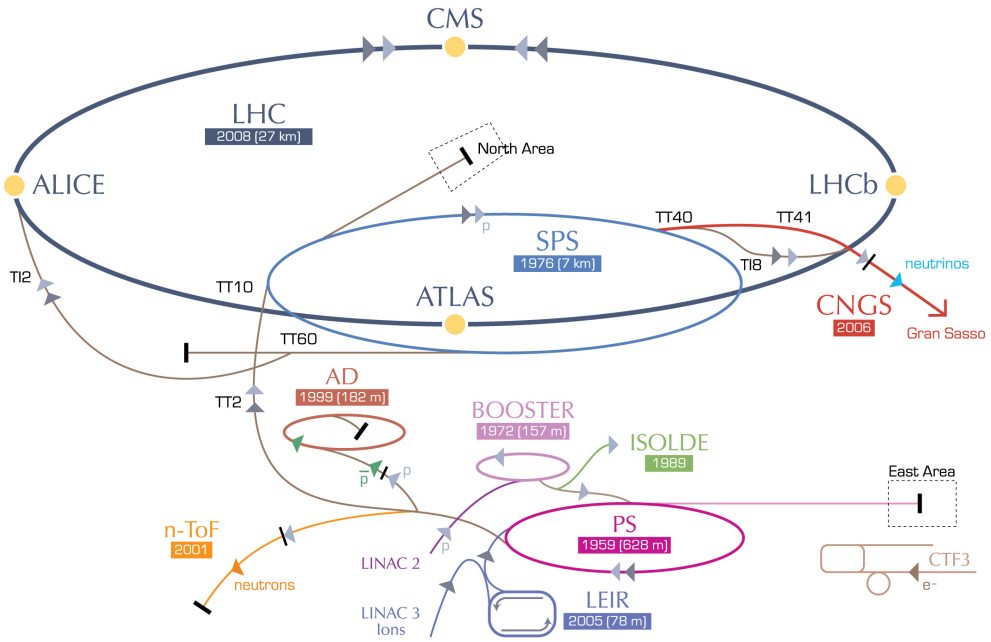


Figure 3.1: The CERN accelerator complex with its four largest experiments ATLAS, LHCb, ALICE, and CMS at the LHC. The diagram illustrates the sequential acceleration stages from LINAC to the LHC ring [32].

Runs at the LHC are characterised by their centre-of-mass energy, luminosity, and total

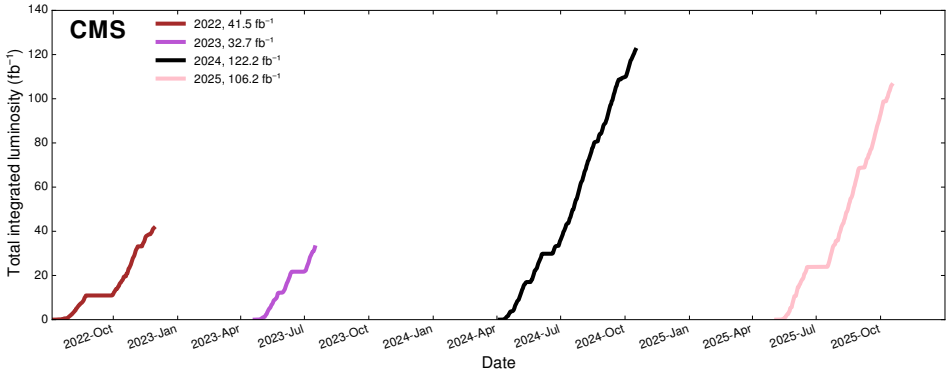


Figure 3.2: Yearly integrated luminosities collected during Run 3. The years are colour-coded with 2022 in brick red, 2023 in purple, 2024 in black, and 2025 in rose, recorded by the CMS Collaboration [33].

cross section σ. In the context of a dual synchrotron accelerating two particles to the same energy, the first is given by the sum of the beam energies of each counter-rotating beam $\sqrt{s} = 2E_{beam}$. The instantaneous luminosity L quantifies the number of particle interactions per unit time and area when two bunches collide. For circular colliders, it takes the form of

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

in which N_b is the number of bunches per beam, which at Run 3 peak operation is 2808. f denotes the circulation frequency, $N_{1/2}$ the number of protons per bunch in each beam, and σ_i the effective propagation of the bunch in the direction of i, as shown in subsection 3.2.1. Integrating the luminosity over time yields the integrated luminosity $L_{int} = \int L(t)dt$. To set an experimental scale, an integrated luminosity of $1\,\text{fb}^{-1}$ equals 10^{14} pp collisions. The cross section σ is the probability for a specific process to occur in a collision and is determined through precision measurements, which can be found, for instance, in [34]. The bunch crossing rate at peak performance is 40 MHz, corresponding to one crossing every 25 ns. The interaction rate is given by:

$$\frac{dN}{dt} = \sigma L \;\leftrightarrow\; N = \sigma L_{int}. \tag{3.2}$$

The LHC operates in distinct running phases, separated by long shutdown periods for technical upgrades and maintenance. This study focuses on the ongoing third run, which began in 2022. The collider currently operates at a centre-of-mass energy of $\sqrt{s} = 13.6\,\text{TeV}$. To date, an integrated luminosity of approximately $515\,\text{fb}^{-1}$[1] has been recorded, of which $302\,\text{fb}^{-1}$ are from the ongoing Run 3 operation. Not all the recorded luminosity is used. For analysis, the certified total integrated luminosities for 2022 and 2023 are $41.5\,\text{fb}^{-1}$ and $32.7\,\text{fb}^{-1}$ respectively. The yearly integrated luminosities collected so far during Run 3 are shown in Figure 3.2. For Run 3 conditions ($\sigma_{inel} \approx 80\,\text{mb}$, $L \approx 20\,\text{pb}^{-1}\,\text{s}^{-1}$), roughly $16 \cdot 10^8$ collisions appear per second. Dividing by the bunch-crossing rate yields an average of about 40 to 80 simultaneous interactions per crossing, referred to as the pileup. The yearly pileup for Run 3 is shown in Figure 3.3.

[1]The SI unit of instantaneous luminosity is $\text{cm}^{-2}\,\text{s}^{-1}$. For integrated luminosity, the inverse barn is commonly used, where $1\,\text{b} = 10^{-28}\,\text{m}^2$.

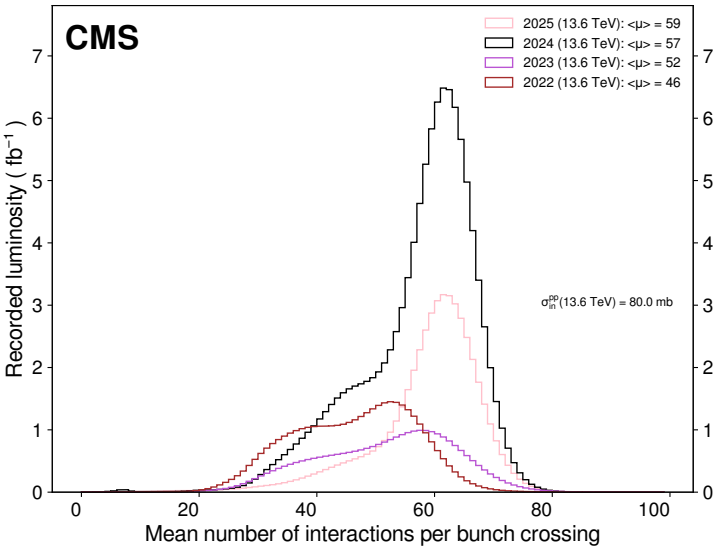


Figure 3.3: Yearly pileup counts collected during Run 3. The years are colour-coded with 2022 in brick red, 2023 in purple, 2024 in black, and 2025 in rose, recorded by the CMS Collaboration [33].

3.2 The CMS Experiment

The Compact Muon Solenoid (CMS) collaboration, alongside the ATLAS experiment, constitutes one of the largest experimental groups worldwide. Both experiments employ general-purpose detectors, designed to explore a broad spectrum of physics phenomena at the TeV scale independently from one another. The nearly hermetic CMS detector [35] is situated near Cessy, France, at Point 5 of the LHC ring. The cylindrical apparatus, illustrated in Figure 3.4, has a diameter of 15 m. This is remarkably small given the level of magnetic field and instrumentation it hosts, thereby justifying the *compact* in its name. The detector consists of concentric subdetectors layered symmetrically around the interaction point in the beam pipe, each designed to measure distinct particle properties with high precision, enabling particle identification and kinematic reconstruction.

The primary objective is to detect stable particles directly through the trigger system and to infer intermediate or unstable particles via precise measurements of momentum, energy, and spatial coordinates of their decay products. The detector identifies electrons, muons, photons, as well as charged and neutral hadrons, while neutrinos, though stable, escape detection due to their minimal interaction cross sections. Subsequently, more complex reconstructed objects and event properties, such as tau leptons, jets, and Missing Transverse Energy (MET) are reconstructed through the Particle Flow Algorithm (PF) (see section 3.3), which combines information from the all-silicon inner tracker, the crystal Electromagnetic Calorimeter (ECAL) to the brass–scintillator Hadronic Calorimeter (HCAL). These subdetectors are enclosed within a 3.8 T superconducting solenoid, causing charged particles to follow curved trajectories, where the outermost detector layers consist of the gas-ionisation muon chambers embedded in the iron return yoke, which closes the magnetic circuit and apparatus setup. Together, the exceptionally strong solenoid and the dedicated Muon System form the defining architectural features from which the detector derives its name.

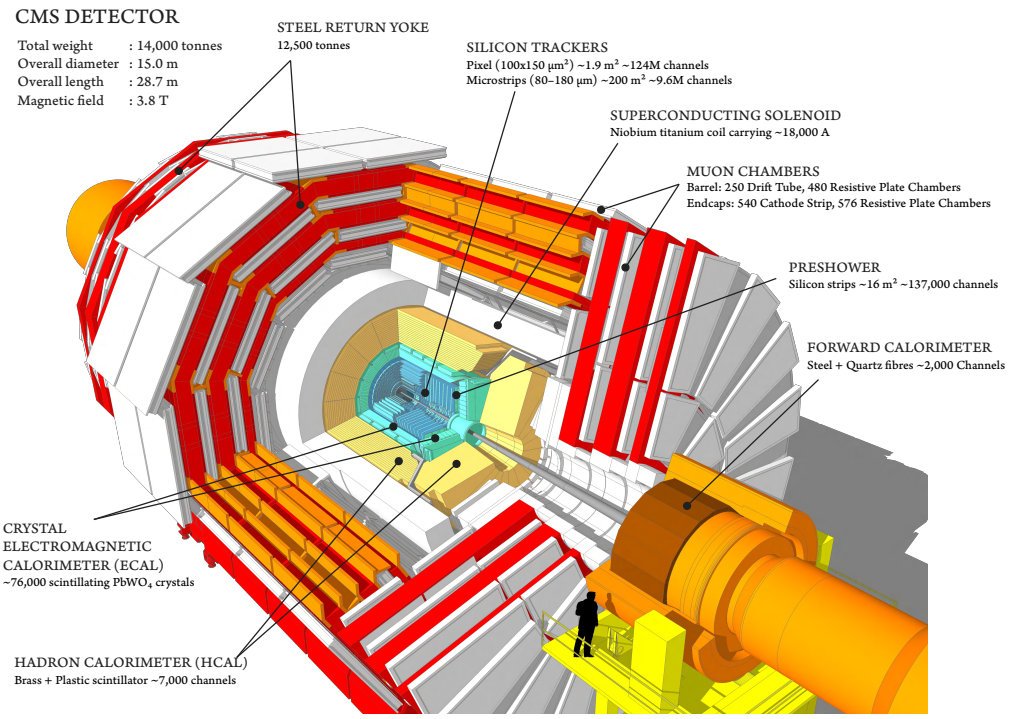


Figure 3.4: Overview of the CMS detector in its Phase-1 configuration. The same detector layout was used during both Run 2 and Run 3, following refurbishment and maintenance during the second Long Shutdown. From the innermost to the outermost layers, the figure shows the Silicon Tracker (blue), the Electromagnetic Calorimeter (ECAL, turquoise), the Hadronic Calorimeter (HCAL, yellow), the Superconducting Solenoid Magnet (white), and the Muon Chambers (grey) embedded in the Return Iron Yoke (red). The detector is nearly hermetically sealed by the Barrel and the two Endcaps [36].

3.2.1 Coordinate System

The CMS collaboration uses a right-handed coordinate system, as illustrated in Figures 3.5 and 3.6. The origin $(0, 0, 0)$ is set to the collision point. The z-axis follows the beam pipe in a positive, counterclockwise direction, the x-axis points radially inward towards the centre of the LHC ring, and the y-axis points vertically upwards, completing the orthogonal basis.

Given the detector geometry, a cylindrical coordinate representation is particularly convenient and sketched in Figure 3.6. Since the radial distance of each hit on a detector component from the interaction point is known, the full three-dimensional track description can be effectively reduced to two dimensions: the azimuthal angle $\varphi \in [-\pi, \pi]$, defined to be zero on the x-axis, $\varphi(\hat{e}_x) = 0$, and the polar angle θ. The latter is replaced by the pseudorapidity η, as defined in Equation 3.3.

Differences in η and φ are invariant under Lorentz boosts along the beam axis and therefore remain unaffected by longitudinal boosts, a pivotal consideration in hadron collisions, where parton momenta vary from event to event.

$$\eta = -\ln \tan \frac{\theta}{2} = -\frac{1}{2} \ln \frac{|\vec{p}| + p_z}{|\vec{p}| - p_z} = \operatorname{arctanh} \frac{p_z}{|\vec{p}|}. \tag{3.3}$$

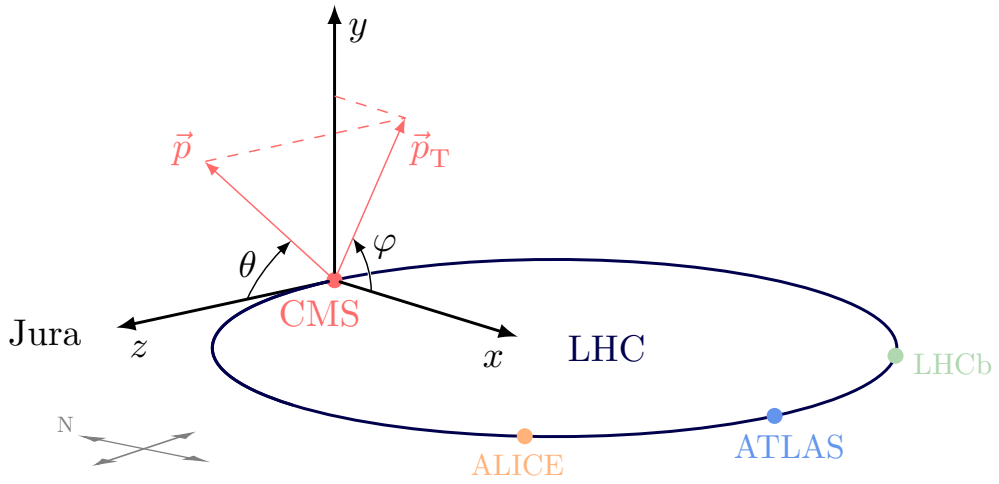


Figure 3.5: Schematic drawing of the Cartesian coordinate system used at CMS, shown in scale with respect to the LHC accelerator. Taken from [37].

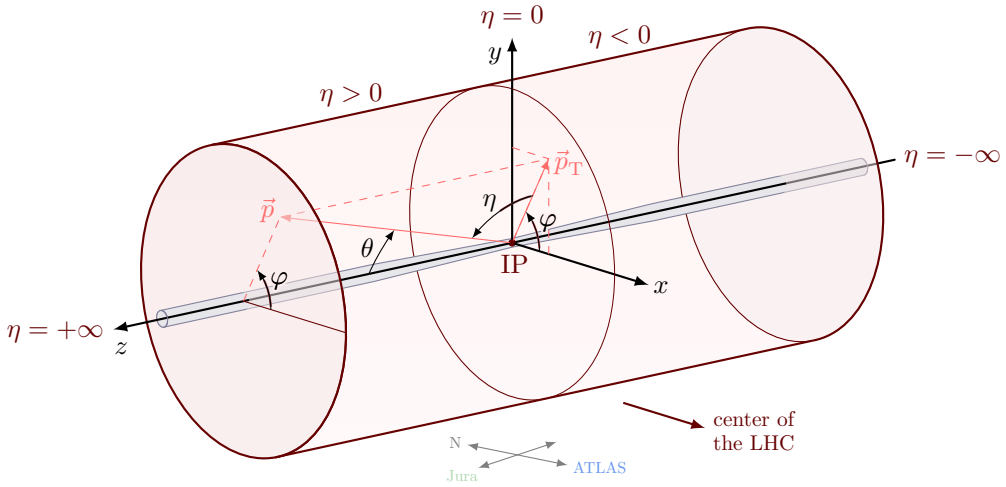


Figure 3.6: Schematic drawing of the CMS cylindrical coordinate system, with reference to the Cartesian system and the surrounding physical environment. Taken from [37].

The angular separation between two objects then is defined in the (η, φ) plane as

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\varphi)^2}. \tag{3.4}$$

This quantity stays invariant under Lorentz boosts along the beam axis.

3.2.2 Silicon Tracker

The CMS tracking system [38, 39] consists of the Silicon Tracker, whose setup is shown in Figure 3.7, and the superconducting solenoid. The system's detection principle is based on silicon sensors that operate as p–n junctions.
The tracker is divided into two complementary subsystems, the Pixel Detector close to the beam pipe and the Strip Detector. Both are described by their cylindrical Barrel section around the beam pipe and two Endcaps, realised as disks orthogonal to the beam axis. Pixel modules provide two-dimensional spatial measurements per hit, whereas strip

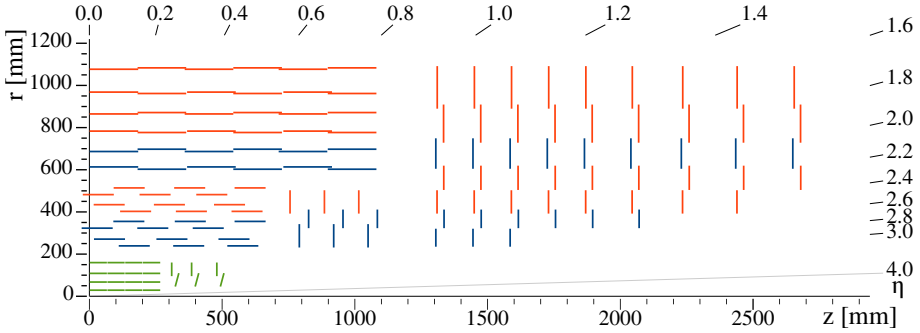


Figure 3.7: Sketch of one quarter of the Phase-1 CMS tracking system in the r-z plane, indicating the pseudorapidity coverage. The Pixel Detector (green) consists of four Barrel layers and three disks in each Endcap, while the strip tracker is realised using single-sided (red) and double-sided (blue) modules. Taken from [40].

modules are elongated and provide one-dimensional information. However, through a sophisticated arrangement, a second coordinate can be inferred.
Following the Phase-1 upgrade for Run 2, the Pixel Detector consists of four layers in the Barrel and three disks in each Endcap, enabling highly precise determination of the primary interaction vertex and achieving a spatial resolution of approximately 10 µm. The first Pixel layer, located just 29 mm from the beam pipe, allows the reconstruction of decay vertices from short-lived particles. At larger radii, extending to roughly 1 m, the Strip tracker provides a long lever arm, which is essential for precise measurement of the transverse momentum p_T of charged tracks, achieving a resolution better than 2 % for particles with $p_T < 100\,\text{GeV}$ and $\eta < 1.0$ with a spatial resolution of 100 µm to 500 µm. The magnet is further described in subsection 3.2.4.
By extrapolating all selected and reconstructed tracks, primary and secondary vertices (PV, SV) are reconstructed to a precision of 25 µm. Moreover, SVs are crucial for b-tagging and the precise reconstruction of decays involving long-lived particles.

3.2.3 Calorimetry

The calorimetric system of the CMS detector surrounds the tracking system and is spatially limited by the solenoid [35]. Its primary purpose is to measure particle energies with high precision by fully absorbing the incoming shower, except in the case of muons and neutrinos, which traverse the detector almost unaffected due to their low bremsstrahlung probability or weak interaction cross section, respectively. The system is divided into two components: the electromagnetic calorimeter (ECAL) and the hadronic calorimeter (HCAL). The first measures the energy of electrons and photons, while the latter detects hadrons as jets and contributes critically to the reconstruction of MET, by being nearly hermetic. Electromagnetic showers are generally shorter than hadronic ones. Consequently, the ECAL is placed closest to the interaction point, followed by the HCAL, which is designed to contain the longer hadronic cascades. Both calorimeters require a high spatial granularity to ensure reliable particle identification throughout the detector, as discussed in section 3.3.

The homogeneous ECAL [35, 41] is composed of dense, transparent $PbWO_4$ crystals. Electrons and photons traversing the crystals initiate electromagnetic showers through successive bremsstrahlung and pair-production processes, producing cascades of secondary

particles. Scintillation light produced during these interactions is collected by photodetectors, providing precise energy measurements and fine spatial resolution for particle identification. The Barrel region covers a pseudorapidity range of $|\eta| < 1.48$, extended by the Endcaps to $|\eta| < 3.0$, providing efficient reconstruction across a wide energy spectrum.

The HCAL is crucial for the measurement of high-energy jets and MET, and provides input for the Level-1 trigger system. Hadronic showers develop when incoming hadrons undergo inelastic collisions with nuclei in the absorber, producing secondary hadrons, nuclear fragments, and electromagnetic sub-showers from neutral pion decays. Therefore, the calorimeter's sampling structure is composed of alternating layers of non-magnetic brass as absorber and plastic scintillators as the active medium, allowing a fraction of the deposited energy to be detected via scintillation light. This sampled energy is proportional to the full hadronic energy, which can be reconstructed through calibration. The HCAL provides nearly hermetic coverage: the Barrel region covers $|\eta| < 1.3$, the Endcaps extend it to $|\eta| < 3.0$, and the forward region reaches $|\eta| < 5.2$, with the latter constructed from radiation-hard steel and quartz fibres to withstand the high radiation near the beam pipe. To contain high-energy hadronic showers that would otherwise escape the inner detector, an additional outer HCAL subdetector is installed in the Barrel ($|\eta| < 1.26$), matching the segmentation of the Muon System.

3.2.4 Solenoid

The cylindrical magnet coil produces a magnetic field of 3.8 T flowing along the z-axis. Owing to its large diameter of 7 m, there is very little material in front of the calorimeters, which improves their performance, but their spatial depth is limited.
Outside the superconducting NbTi solenoid, the magnetic field reverses direction and decreases to 2 T, returning through the iron yoke that is incorporated with the Muon Chambers.
Charged particles are forced onto curved trajectories by the homogeneous magnetic field due to the Lorentz force. From the bending radius R inside the tracker, the transverse momentum p_T is determined by equating the centripetal force of a particle on a curved track with the corresponding magnetic force:

$$p_T = eBR = 0.3 \cdot z \cdot \frac{R}{[\mathrm{m}]} \cdot \frac{B}{[\mathrm{T}]}\,[\mathrm{GeV}] \qquad \text{, with} \qquad \frac{\sigma(|\vec{p}|)}{|\vec{p}|} = \frac{1}{2}|\vec{p}|[\mathrm{GeV}]. \tag{3.5}$$

The relative uncertainty of momentum measurements increases with particle momentum, as higher-momentum tracks bend less in the magnetic field.

3.2.5 Muon Chambers

Muons are minimally ionising particles that traverse the tracker and calorimeters with little energy loss. Energy loss via Bremsstrahlung is strongly suppressed due to their large mass, as $dE/dx|_{\mathrm{brem}} \propto E/m^2$, while ionisation energy loss follows the Bethe-Bloch scaling, roughly $dE/dx|_{\mathrm{ion}} \sim 1/\beta^2$ at sub-relativistic velocities. Very low-momentum charged particles may curl up in the magnetic field and form loopers in the tracker, whereas muons with typical high momenta traverse both the tracker and the calorimeters. For this reason, the Muon Chambers are placed outside the magnet coil and integrated into the iron return yoke, completing the CMS detector. The thick and dense iron layers

ensure that muons traverse sufficient material to produce detectable signals in multiple chamber layers, allowing a precise measurement of their momentum. The muon system is therefore designed to detect muons, not to fully absorb them. It also provides excellent timing resolution.

The Muon Chambers are gaseous detectors, in which passing charged particles ionise the gas. The released charge is then amplified into an electron avalanche and collected by the readout. To match the magnetic field configuration and detector geometry, the system is composed of Drift Tubes (DT) in the Barrel region ($|\eta| < 1.2$), Resistive Plate Chambers (RPC) up to $|\eta| < 1.6$, and Cathode Strip Chambers (CSC) in the Endcaps ($0.9 < \eta < 2.4$). The RPCs and CSCs have a fast signal response, which makes them well suited for precise timing and thus triggering. The full muon track is reconstructed combining the muon chamber and the tracker information, achieving a transverse momentum resolution of better than 6 %.

3.2.6 Trigger System

The trigger system [42] forms the first stage of preselection and is essential to reduce the enormous data amount to one that can realistically be stored. At a bunch-crossing rate of 40 MHz, about 10^9 pp collisions occur per second at Point 5. Since a single event corresponds to roughly 1 MB of data, this would result in close to 950 TB of data per second. However, the data acquisition system can store only a few kHz of events. The raw data stream must therefore be reduced before it is written to disk. Instead of selecting events at random, CMS employs a dedicated trigger system which preferentially keeps those collisions most relevant for its physics goals. CMS uses a two-level trigger structure, composed of the fast hardware-based Level-1 Trigger (L1T) and the more defined High-Level Trigger (HLT).

The L1T is implemented in custom electronics and performs a first, coarse reduction of the event rate to 100 kHz. Its decisions are based on low-level information from the Calorimeters and the Muon System, using reduced spatial granularity. Within a latency of about 1 µs, it reconstructs high-p_T muons, electrons and photons, as well as global variables such as transverse energy sums and missing transverse momentum.
Events accepted by the L1T are passed to the HLT, which performs a software-based refinement and further reduces the recording rate to a few kHz. Since the HLT has access to information from all CMS subdetectors, in particular the tracker, it can apply tighter reconstruction criteria and select specific final states with much greater precision. The final reconstruction accuracy is then reached offline during the analysis stage. The normal data stream corresponds to about 1.5 kHz of events, while additional *parked* data, such as B-physics, VBF, or low-p_T electron/muon events, contributes a few kHz and is reconstructed later when CPU resources are available.

3.3 Particle Reconstruction

After their production in a pp collision, all long-lived particles propagate through the detector before eventually interacting with its various subdetectors. Those that produce visible and characteristic signatures can be reconstructed with high fidelity. These particles include muons μ, electrons e, hadrons, either charged $\pi^{\pm}, K^{\pm}$, or neutral π^0, and

photons γ. Early reconstruction strategies relied on the global energy flow. This concept has since evolved into the sophisticated Particle Flow Algorithm (PF) [43], which combines information from all subdetectors in an iterative procedure. By linking complementary measurements across detector layers, PF achieves a significantly improved energy and momentum resolution and produces so-called *PF candidates*, which form the basis for higher-level objects such as jets, hadronically decaying tau leptons, as well as the reconstruction of the primary and secondary interaction vertices.

The reconstruction proceeds stepwise. First, the low-level information from each subdetector is preprocessed: tracker hits form charged-particle tracks, the calorimeters provide energy clusters, and the muon system yields track signatures. In a second step, these detector-level objects are spatially matched and iteratively linked to form *PF candidates*, which are then interpreted according to their particle identity:

- Electron: A track in the tracker matched to an ECAL energy deposit.
- Photons: ECAL energy deposits without a matched track.
- Charged hadrons: A tracker track associated with clusters in the ECAL and HCAL.
- Neutral hadrons: Energy deposits in both calorimeters without a corresponding track.
- Muons: Tracks reconstructed consistently in both the tracker and the Muon System, with the momentum determined from the full trajectory's curvature in the magnetic field across the sub-detectors.

The particles of interest for this thesis are muons, charged hadrons, and neutrinos, the latter reconstructed in the form of MET. From these visible decay products, the parent tau leptons are inferred. Figure 3.8 illustrates the tau decay modes relevant for this analysis and shows how their visible daughter particles are reconstructed from *PF candidates* by linking detector-level signatures across subdetectors. In the one-prong hadronic decay (DM0), a single charged pion leaves one track in the silicon tracker, accompanied by a hadronic calorimeter deposit. The ρ decay mode (DM1) additionally produces an electromagnetic shower from the neutral pion in the ECAL. Three-prong decays (DM10) yield three charged tracks originating from a common, possibly displaced secondary vertex, while the muonic decay is identified as a minimally ionising track extending into the muon chambers.
This motivates the next reconstruction steps, namely the determination of charged-particle trajectories and their common production vertices, followed by the reconstruction of physics objects from detector information as *PF candidates* such as muons and jets, while neutrinos are inferred from the missing transverse momentum.

3.3.1 Track Reconstruction

Within PF reconstruction, charged-particle trajectories are obtained from successive hits in the Silicon Tracker, forming track candidates that are subsequently associated with calorimeter information to aid particle identification [43, 45].

Tracker hits are first combined into short initial segments, or seeds, which are iteratively

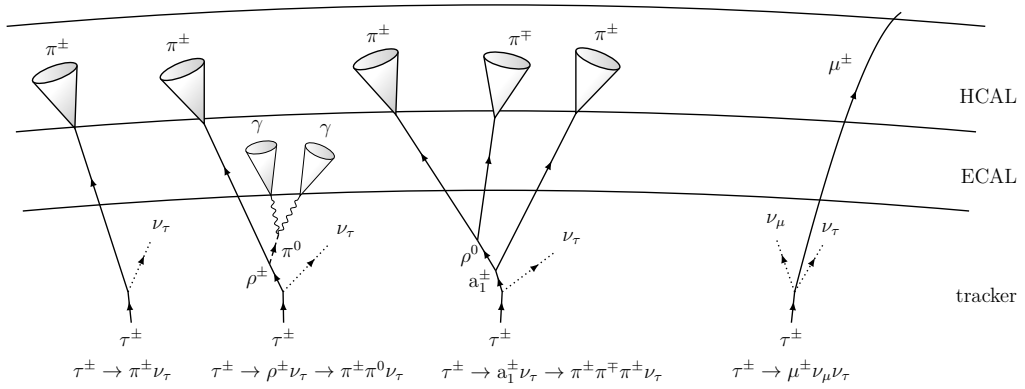


Figure 3.8: Illustration of the chosen tau lepton decay modes in the CMS detector. The tracks and showers represent the visible decay products in the tracker, ECAL, and HCAL subdetectors. Shown are the single-prong hadronic decay $\tau^\pm \to \pi^\pm\nu_\tau$ (DM0), the ρ mode $\tau^\pm \to \rho^\pm\nu_\tau \to \pi^\pm\pi^0\nu_\tau$ (DM1), the three-prong decay $\tau^\pm \to a_1^\pm\nu_\tau \to \pi^\pm\pi^\mp\pi^\pm\nu_\tau$ (DM10), and the leptonic muon decay $\tau^\pm \to \mu^\pm\nu_\mu\nu_\tau$. The curvature of the charged particle tracks corresponds to the bending in the CMS magnetic field. Modelled after [44].

extended to form full trajectories. Tracks from higher-momentum particles are typically reconstructed first, followed by lower-energy or displaced tracks, because the straighter trajectories of high-momentum particles produce less ambiguous hit patterns, simplifying the reconstruction and reducing conflicts when assigning hits to more curved or displaced tracks. Calorimeter clusters are then built from energy deposits by identifying local maxima and adding neighbouring cells with significant energy. Overlapping clusters share energy according to their proximity to each maximum, and nearby contributions from secondary particles, such as Bremsstrahlung photons, are separated when necessary. Finally, tracks and clusters are linked layer by layer, producing *PF candidates*, which are then interpreted as electrons, muons, photons, or hadrons.

3.3.2 Vertex Reconstruction

Precise vertex reconstruction [45] is essential to separate the hard-scatter interaction from pileup and to resolve displaced decay topologies. While most particles produced in proton–proton collisions decay promptly at the interaction point, some particles, such as hadrons containing b-quarks or tau leptons, have a finite lifetime and travel a measurable distance before decaying. Their decay products therefore originate from a point displaced from the primary interaction vertex, forming so-called secondary vertices (SV). Vertices are reconstructed within the PF algorithm. PVs are required for general event reconstruction and interpretation, while SVs play a crucial role in b-tagging, tau lepton reconstruction, and in the identification of other signatures from displaced decays of long-lived particles.

A PV is reconstructed by selecting tracks compatible with the interaction region, clustering them, and fitting a common origin. The vertex with the largest sum of track transverse momenta is identified as the PV, while the remaining ones are classified as *pileup vertices*. SVs are reconstructed by seeding from tracks with a large impact parameter significance. The impact parameter is defined as the minimum distance of approach between the extrapolated trajectory of a charged particle and the primary vertex. A non-zero impact parameter typically indicates that the particle was produced in a displaced

decay. Tracks with a large impact parameter significance are therefore strong indicators of particles originating from decays of long-lived particles, such as b hadrons or tau leptons. Geometrically compatible neighbouring tracks are then clustered, and a robust, outlier-resistant fit is performed to determine the position of the SV.

3.3.3 Muon Reconstruction

Muons are the easiest particles to identify, as they are the only ones that traverse the entire detector up to the Muon Chambers and leave a visible signature. For this reason, PF attempts to reconstruct them first.

Three types of muons are defined [46]. *Tracker muons* are built from tracks in the Tracker System that are geometrically matched to at least one Muon Chamber segment, provided the momentum thresholds of 2.5 GeV total and 0.5 GeV transverse momentum are met. They benefit from high reconstruction efficiency, particularly at low p_T, since low-momentum muons leave sufficient hits in the tracker to form well-measured tracks, even though they may leave only a limited number of hits in the Muon System. However, they are more susceptible to punch-through hadrons, i.e. charged hadrons that traverse the calorimeters without being fully absorbed and leave signals in the Muon System. *Stand-alone muons*, in contrast, are reconstructed solely from muon chamber information. This suppresses fakes originating from the interaction point but admits cosmic-ray muons, i.e. high-energy muons originating from cosmic rays that enter the detector from outside rather than from the collision point. Due to the absence of tracker information, their momentum resolution is comparatively poorer.
The strengths of both approaches are combined in *global muons*, where a stand-alone muon track is matched to a tracker track and combined into a single, optimised trajectory. If several candidates are available, the one with the smallest fit uncertainty is chosen. This yields the best momentum resolution, especially above 200 GeV, while significantly reducing the rate of misidentified muons. The resulting excellent p_T resolution enables the reconstruction of narrow invariant-mass peaks from well-known resonances such as quarkonium states or the Z boson.

3.3.4 Jet Reconstruction

Although jets are not a signal object in this $H \to \tau^+\tau^- \to \mu a_1^{3pr}$ analysis, they must be reconstructed to obtain an accurate MET estimate, event categorisation, and fake rates.

Jets originate from quarks and gluons, which cannot exist as free particles due to colour confinement in QCD[2]. They hadronise into collimated sprays of hadrons, which appear as clustered energy deposits in the calorimeters. Within PF, these deposits are combined into jet candidates that are not already associated with isolated neutral particles, and are subsequently clustered using, most commonly, the anti-k_T algorithm, which is a recombination algorithm that produces conical jets of different radii [47]. The standard choice in

[2] *Quantum Chromodynamics* (QCD) is the Non-Abelian gauge theory of the strong interaction, formulated on the $SU(3)_C$ symmetry group, as introduced for the general case in section 2.1. It describes the dynamics of quarks and gluons, which carry the conserved quantum number known as *colour charge*. Due to colour confinement, these partons cannot exist in isolation and therefore hadronise into colour-neutral bound states.

this analysis are AK4 jets (radius parameter $\Delta R = 0.4$), while larger AK8 jets ($\Delta R = 0.8$) are mainly used in boosted topologies where several partons merge into a single fat jet.

3.3.5 Missing Transverse Momentum Reconstruction

Missing Transverse Energy (MET) quantifies the total momentum carried by particles that escape detection and is a key observable in many analyses.
Genuine MET arises predominantly from neutrinos, whose extremely small interaction cross section prevents them from depositing energy in the detector, but it may also signal invisible particles predicted in theories beyond the SM. Apparent MET, by contrast, can be introduced by reconstruction inefficiencies, inactive detector regions, or mismeasured jets. Experimentally, accurate MET reconstruction relies on hermetic calorimeter coverage, excellent muon reconstruction, and the overall performance of the PF algorithm. In PF, the best resolution from the tracker and calorimeter is used for charged hadrons, improving MET reconstruction.

Since the longitudinal momentum of the colliding partons is unknown, only the transverse plane is considered. The missing transverse momentum vector $\not{E}_T$ is defined as

$$\not{E}_T = \left|\vec{E}_T^{miss}\right| \qquad \text{and} \qquad \vec{E}_T^{miss} = -\sum_{i}^{particles} \vec{p}_{T,i} \ , \tag{3.6}$$

ensuring transverse momentum conservation in the event.
The datasets this thesis considers are taken under high pileup conditions, as discussed in section 3.1. Each additional pileup collision produces charged and neutral particles. Charged ones, such as charged pions, leave tracks in the tracker and can therefore be assigned to their true non-PV vertex and thus removed. Neutral particles, like photons and neutral pions, leave only calorimeter deposits without a track to identify their origin. They therefore appear as if originating from the main hard interaction and hence add spurious transverse energy.
To mitigate this, the PileUp Per Particle Identification (PUPPI) [48] weighting is applied to the PF algorithm. Each *PF candidate* is assigned a weight according to its likelihood of originating from the primary vertex, effectively suppressing pileup contributions and substantially improving the MET resolution.

3.3.6 Tau Lepton Reconstruction

The precise reconstruction of tau leptons and the correct identification of their Decay Mode (DM) are essential for this analysis, as they provide access to the properties of the Higgs boson. Three complementary algorithms are applied to achieve this objective [49]. Initially, the Hadron-plus-Strip Algorithm (HPS) algorithm builds tau candidates from charged and neutral PF objects within reconstructed PF jets and identifies hadronic tau decays. These candidates are then classified and refined by the DeepTau Convolutional Neural Network (DeepTau), which reduces misidentification from jets, electrons, and muons. Finally, the ParticleNet Algorithm (PNet) further improves the assignment of tau decay modes, achieving higher purity than HPS alone.

Hadron-plus-Strip Algorithm

The Hadron-plus-Strip Algorithm (HPS) [50] reconstructs hadronic tau leptons from PF candidates. In addition to identifying charged hadrons, it reconstructs neutral pions, which immediately decay into photon pairs that appear as strips in the ECAL, and also collects electrons from photon conversions, hence the algorithm's name. While the HPS can, in principle, identify tau decay modes, studies of Run 3 data [49] show that the PNet achieves higher purity and efficiency, and has therefore largely superseded it.

Tau reconstruction proceeds in three steps. First, charged and neutral PF candidates are clustered into tau candidates, which are then evaluated under all allowed DM hypotheses. In addition, the tau four-momentum is reconstructed for each hypothesis to determine the most likely assignment. A discriminator is subsequently applied to classify each candidate as a genuine tau or background. Genuine tau jets are typically isolated, have lower particle multiplicity, and narrower energy deposits than jets from quarks or gluons, or electrons and muons. The rate of misidentification can be reduced by applying different Working Points (WPs)[3] defined by the Tau POG.

DeepTau Algorithm

The multiclass deep convolutional neural network DeepTau [51, 52] aims to reduce the misidentification of hadronic tau candidates from jets, muons, and electrons. It relies on tau candidates reconstructed by the HPS, using their constituents, kinematic properties, reconstructed charge, and probability of originating from the primary vertex (PV) as input features. These inputs are arranged in a grid of cells in the (η, φ) plane around the leading tau decay product. Finer inner cells cover the signal cone and coarser outer cells cover the isolation cone. High-level tau and event properties are also used to improve classification. The network processes these features according to their distances from the tau jet axis or the objects/events they represent. For each tau candidate, the network outputs four scores y_{obj}, corresponding to the probability of the candidate being a genuine tau lepton, a jet, an electron, or a muon. For event categorisation, the candidate is associated with the object hypothesis with the highest y_{obj}, which ensures a selection enriched in genuine tau leptons. To study fake rates and allow flexible balancing between efficiency and background rejection, a discriminator is defined as:

$$p_{\tau-vs-obj} = \frac{y_\tau}{y_\tau + y_{obj}} \tag{3.7}$$

which provides the misidentification rate as a function of the efficiency for genuine tau selection.
The performance of the DeepTau discriminator [53] is shown in Figure 3.9, which shows the misidentification probability of jets, electrons, and muons as a function of the genuine hadronic tau identification efficiency for low- and high-p_T candidates. For typical working points, corresponding to genuine hadronic tau identification efficiencies of about 70–80%, DeepTau achieves misidentification probabilities of the order of 10^{-3} to 10^{-4} for jets, and even lower values for electrons and muons, depending on the p_T region and working point choice. Compared to its predecessor, the new DeepTau version 2.5 demonstrates improved

[3] The respective POGs define several WPss, which balance signal efficiency against background rejection. Each physics object has its own identification criteria. For tau leptons, available working points include Vvloose, Vloose, Medium, Tight, VTight, and VVThight.

performance across all commonly used working points, yielding a cleaner selected sample with fewer misidentified background objects and thus improving the overall selection purity in physics analyses.

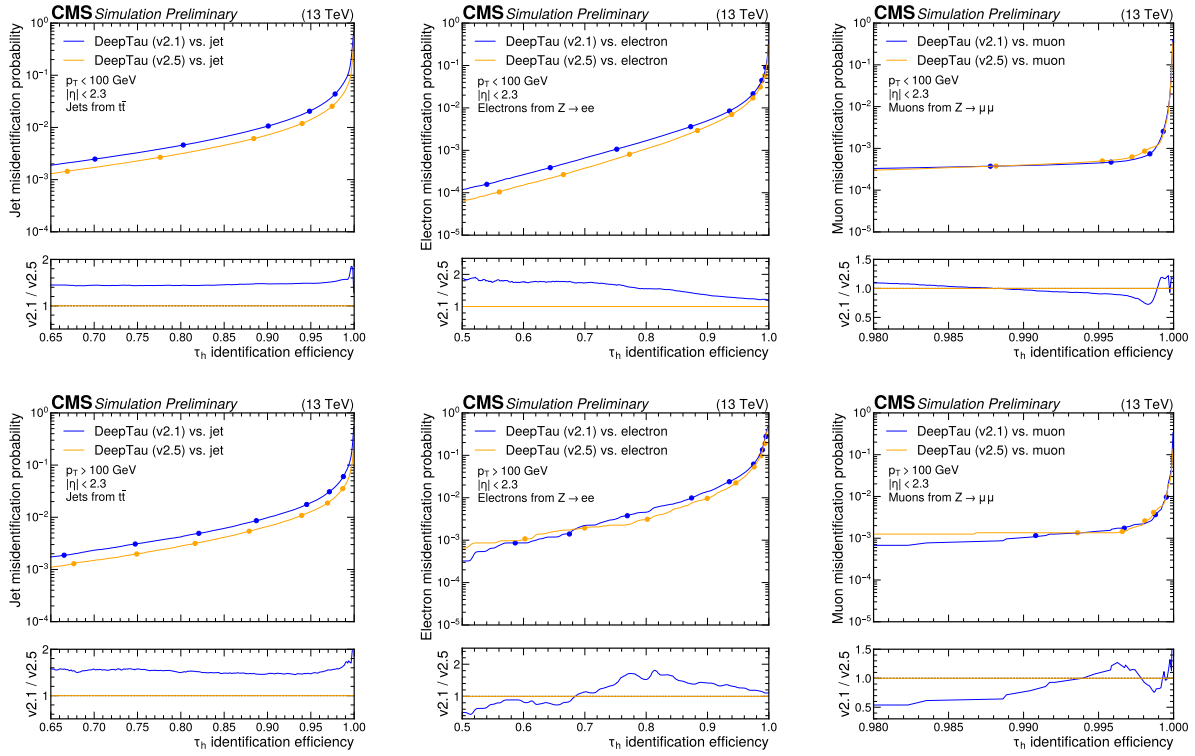


Figure 3.9: Misidentification probability as a function of the genuine hadronic tau identification efficiency for the DeepTau discriminator, evaluated on simulated 2018 datasets. From left to right, the rows show the performance against jets, electrons, and muons, respectively, for low-p_T (top) and high-p_T (bottom) τ_h candidates. The defined working points are indicated as filled circles [53].

ParticleNet Algorithm

In the global Run 3 analysis, the ongoing follow-up of [14], multiple DMs with varying sensitivities are considered, whereas this thesis focuses specifically on the μa_1^{3pr} subset. Accurate DM identification is therefore crucial. Hadronically decaying tau leptons typically appear as narrow, isolated jets in the detector, and their correct classification is essential for distinguishing genuine tau leptons from quark or gluon jets, as well as from electrons or muons. In Run 2, tau identification relied on dedicated Boosted Decision Trees (BDTs) trained on HPS tau candidates to separate genuine taus from background objects.

The ParticleNet Algorithm (PNet) [54] algorithm provides an alternative approach by treating each tau candidate as a *particle cloud* composed of its constituent PF objects. Jets originating from different particles exhibit distinct characteristics, for example, gluon jets typically have broader energy distributions than quark jets, while jets from heavy particles often contain multiple prongs, which the network exploits. Generally, jets can be

represented as sequences, trees, graphs, or unordered sets of particles, motivating the use of a dynamic graph convolutional neural network.
The network captures both local and global jet structures: shallow layers encode local neighbourhood information of particles within the jet, while deeper layers reconstruct global correlations and interconnections. Non-additive quantities such as particle type and kinematic features are preserved throughout the network. In addition, high-level properties of the tau candidate and the event are included, allowing the network to combine fine-grained particle-level topology with global event information. By learning both local and global structures, PNet improves the identification of genuine tau candidates compared to the Run 2 BDT approach.

The PNet performance, evaluated in Run 3 [55], reaches a reconstruction efficiency for genuine hadronic tau leptons of about 90% when using CHS jets. Charged Hadron Subtraction (CHS) removes charged hadrons that are not linked to the PV, giving the network a cleaner set of particles to work with. In comparison, DeepTau using HPS candidates achieves around 70% efficiency. The misidentification of quark and gluon jets is slightly higher at very tight working points, but PNet can recover approximately 25 to 30% of genuine tau leptons missed by HPS. Misidentification rates for electrons and muons remain below 1%, maintaining a good balance between efficiency and purity. As a result, PNet proves especially useful in studies that need to maximize the selection of genuine hadronic tau candidates while controlling backgrounds.

Chapter 4

Strategy of the CP Measurement in $H \to \tau\tau \to \mu a_1^{3pr}$ Decays

4.1 Data and Simulation Samples . . . 40
4.1.1 Recorded Data Samples . . . 40
4.1.2 Generated Monte Carlo Samples . . . 40
4.2 Event Selection . . . 43
4.2.1 Corrections . . . 43
4.2.2 Selection of the $\mu\tau_{had}$ and μa_1^{3pr} Channel . . . 45
4.3 Event Categorisation . . . 46
4.4 Measuring the ϕ_{CP} Angle . . . 47
4.4.1 Impact Parameter Method . . . 49
4.4.2 Decay Plane Method . . . 50
4.4.3 Polarimetric Vector Method . . . 51
4.5 Kinematic Reconstruction . . . 51
4.5.1 fastMTT Method . . . 52
4.5.2 Global Event Fit . . . 55
4.6 Estimation of the Sensitivity to the CP observable . . . 60
4.6.1 Generator-level Observables . . . 60
4.6.2 Model for the ϕ_{CP} Distributions . . . 61
4.6.3 Spin Analysing Power . . . 62
4.7 The ColumnFlow Analysis Tool . . . 62

This chapter presents the analysis setup and the principles underlying the measurement. It begins with a description of the recorded data and simulated MC samples used for both signal and background processes. This is followed by the event selection, including corrections and trigger requirements, leading to the definition of the μa_1^{3pr} candidate sample. To discriminate between signal and background events in data, a Boosted Decision Tree (BDT) classifier is introduced. The measurement of the CP properties of the Higgs

boson is then presented, where the key observable $\alpha^{H\tau\tau}$ is extracted via the measurable coplanarity angle ϕ_{CP} between the decay planes of the tau leptons. Consequently, three different plane reconstruction methods are discussed in detail. To enhance the sensitivity of the spin analysis, the tau momenta are successively corrected, first using a modular correction with fastMTT [56–58], and then with an angular correction based on the Global Event Fit (GEF) [56, 58, 59] approach. Subsequently, the fit of the resulting distributions and the calculation of the corresponding asymmetry variable, used to quantify the analysis performance, are explained. Finally, an overview of the columnar framework tool COLUMNFLOW is provided.

4.1 Data and Simulation Samples

The measurements are based on pp collision data recorded by the CMS detector at a centre-of-mass energy of 13.6 TeV during Run 3 of the LHC operation. In addition to collision data, Monte Carlo (MC) simulated samples are employed to model the signal and background processes, to optimise the event selection, and to evaluate the performance and systematic uncertainties of the analysis. First, the recorded data samples and their corresponding event yields are summarised. This is followed by a description of the generation of signal and background MC samples, including a detailed listing of all considered background processes and their classification into irreducible and reducible contributions.

4.1.1 Recorded Data Samples

All data used in this analysis were recorded during Run 3 of the LHC operation with the CMS detector. The dataset consists of pp collisions at a centre-of-mass energy of 13.6 TeV. They amount to an integrated luminosity of $41.5\,\text{fb}^{-1}$ in 2022 and $32.7\,\text{fb}^{-1}$ in 2023 delivered by the LHC, of which $38.5\,\text{fb}^{-1}$ and $29\,\text{fb}^{-1}$, respectively, were recorded by the CMS experiment. An overview of the data samples is given in Table 4.1. All samples are skimmed with the HLepRare package `NanoProd (v14)` and belong to the `nano_tau_skim_2025_v1` campaign.

4.1.2 Generated Monte Carlo Samples

To understand the backgrounds of the signal processes studied in data, it is crucial to simulate these physics processes as accurately as possible. Monte Carlo (MC) simulated samples also contain the truth-level information, which the analysis aims to extract from the data. The MC generation procedure for both background and signal samples is outlined below. An overview of the simulated event samples used in this analysis can be found in Table 4.2.

Simulating HEP collisions

The simulation of high-energy pp collisions is performed in several consecutive steps, each describing a different physical aspect of the interaction.

First, the hard scattering process is computed at the parton-level using a fixed-order matrix-element calculation. The incoming protons are described in terms of their constituent partons, whose longitudinal momentum distributions are given by parton distri-

Table 4.1: Data samples used in the analysis with their corresponding data-taking eras and number of events. All samples are skimmed with the HLepRare package `NanoProd (v14)` and belong to the `nano_tau_skim_2025_v1` campaign.

Run era	Dataset	Number of events
2022preEE	`/SingleMuon/Run2022C-22Sep2023-v1/MINIAOD`	20 162 441
	`/Muon/Run2022C-22Sep2023-v1/MINIAOD`	138 329 693
	`/Muon/Run2022D-22Sep2023-v1/MINIAOD`	75 440 027
2022postEE	`/Muon/Run2022E-22Sep2023-v1/MINIAOD`	141 480 973
	`/Muon/Run2022F-22Sep2023-v2/MINIAOD`	449 185 088
	`/Muon/Run2022G-22Sep2023-v1/MINIAOD`	76 689 396
2023preBPix	`/Muon0/Run2023C-22Sep2023 v1-v1/MINIAOD`	54 715 896
	`/Muon0/Run2023C-22Sep2023 v2-v1/MINIAOD`	17 063 451
	`/Muon0/Run2023C-22Sep2023 v3-v1/MINIAOD`	20 015 377
	`/Muon0/Run2023C-22Sep2023 v4-v1/MINIAOD`	138 943 783
	`/Muon1/Run2023C-22Sep2023 v1-v1/MINIAOD`	54 621 922
	`/Muon1/Run2023C-22Sep2023 v2-v1/MINIAOD`	17 059 895
	`/Muon1/Run2023C-22Sep2023 v3-v1/MINIAOD`	20 010 429
	`/Muon1/Run2023C-22Sep2023 v4-v2/MINIAOD`	138 834 244
2022postBPix	`/Muon0/Run2023D-22Sep2023 v1-v1/MINIAOD`	100 211 533
	`/Muon0/Run2023D-22Sep2023 v2-v1/MINIAOD`	21 462 916
	`/Muon1/Run2023D-22Sep2023 v1-v1/MINIAOD`	100 281 976
	`/Muon1/Run2023D-22Sep2023 v2-v1/MINIAOD`	21 463 645
$\sum$		1 605 972 685

bution functions (PDFs). The cross section is obtained by convoluting the partonic matrix element with the PDFs. The dependence on the renormalisation and factorisation scales is used to estimate the impact of missing higher-order corrections. This is generated commonly by one of the generators: Powheg [60–63], aMCatNLO, or MadGraph [64, 65]. The outgoing partons produced in the hard interaction are then evolved through a parton shower simulation, which approximates higher-order QCD radiation. Since coloured partons cannot exist as free particles due to colour confinement, the parton shower is followed by a hadronisation model, which converts the partons into colour-neutral hadrons. Additional soft interactions between the proton remnants, commonly referred to as the underlying event, are simulated alongside the hard process. The common tool for this is Pythia [66]. In order to avoid double counting of radiation processes between the fixed-order matrix-element calculation and the parton shower, matching procedures as FxFx [67] or MLM [68] are applied, depending on the generator configuration. Furthermore, for specific processes, spin correlations in particle decays can be introduced via the TauSpinner or MadSpin tools [69–71]. Finally, the generated particles are passed through a detailed simulation of the CMS detector based on Geant4 [72, 73], which models their interactions with the detector material and magnetic field. From this stage onwards, the same reconstruction and analysis chain is applied to simulated events and collision data.

The simulation strategy described above is implemented as follows. The signal corresponds to Higgs boson production with subsequent decay into a pair of tau leptons. The dominant production modes considered in this analysis are ggF and VBF. The hard scat-

tering process is generated using POWHEG, and the Higgs boson decays into di-tau pairs without spin correlations. The effects of the tau lepton spin correlations are subsequently introduced using the TAUSPINNER package. Thereby, three signal hypotheses are constructed by event-by-event reweighting of the events to CP-mixing angles $\alpha^{H\tau\tau} = 0°$, 45°, and 90°, corresponding to the CP-even, CP-maximally mixed, and CP-odd scenarios, respectively. The parton showering, hadronisation, and underlying event are simulated using PYTHIA, which is used consistently for all simulated processes in this analysis.

All Drell-Yan events, generated inclusively or with up to two jets, are simulated using the MADGRAPH5 generator. To match events where final-state radiation occurs and is modelled twice, FXFX matching is performed. The W+jets background is generated, inclusively or with up to four outgoing partons, at the matrix-element-level using MADGRAPH, and the matching to the parton shower is performed using the MLM scheme. Top-quark pair production and single-top processes are generated using POWHEG. For single-top production, spin correlations in the top-quark decay are modelled using MADSPIN. Finally, multi-boson processes are simulated using PYTHIA.

Background Processes

The identification of tau leptons is of central importance in increasing the signal sensitivity, as the data samples contain contributions from many physics processes with similar experimental signatures, primarily due to the presence of numerous background processes that can imitate the signal. Backgrounds are commonly classified into two categories: irreducible and reducible, depending on whether the final state contains genuine tau leptons or objects misidentified as taus.

Irreducible backgrounds, in this analysis, originate from electroweak processes that produce genuine tau lepton pairs and therefore closely resemble the signal in both kinematics and topology, making them indistinguishable from the signal based on the visible decay products alone. The most important processes in the $H \to \tau\tau$ analysis are:

- **Drell-Yan ($Z \to \tau\tau$):** This decay constitutes the dominant irreducible background. Its final state is almost identical to the signal, with the invariant mass distribution of the di-tau system being the primary discriminating observable. This distribution peaks at the Z boson mass, rather than at the Higgs boson mass.

- **Di- and tri-boson production (VV, VVV, ($V = W, Z$)):** These processes can produce tau leptons when at least one gauge boson decays leptonically ($W \to \tau\nu$, $Z \to \tau\tau$) to a tau lepton. Their production cross sections are smaller than that of inclusive Z production. Due to their very small cross sections, the overall contribution from VVV processes remains limited.

Reducible backgrounds, in contrast, originate from processes such as QCD multijet or light-lepton production, in which jets or leptons are misidentified as tau leptons and thus fake the signal signature:

- **W+jets:** Events in which a W boson decays into a tau lepton are accompanied by QCD jets misidentified as a hadronic tau.

- **Single-top production (tW):** Single-top processes produce a W boson that can decay into a tau lepton, with additional jets in the final state that can be misidentified as hadronic taus.

- **$t\bar{t}$ production:** Top quark pair production leads to final states containing two W bosons, which can decay into tau leptons. In addition, the presence of associated b-jets further complicates the reconstruction if they are misidentified as tau leptons.

- **QCD multijet:** These events are predominantly composed of gluon or light-quark jets, making this background particularly difficult to model and typically requiring data-driven estimation techniques. However, high-energy jets can mimic the signature of hadronic tau decays.

4.2 Event Selection

Events are selected to isolate a clean and well-defined sample of $H \to \tau\tau \to \mu\tau_{had} + 3\nu$ events, with particular emphasis on the $\tau_{had} \to a_1^{3pr}\nu$ decay mode. The event and pair selection applied in the $\mu\tau_{had}$ channel is summarised in Table 4.4. It is designed to suppress the dominant reducible backgrounds, in particular QCD multijet production, W+jets, and Drell-Yan processes, while maintaining a high signal efficiency. In addition, a consistent treatment of data and simulated events is ensured through the application of the appropriate corrections.

4.2.1 Corrections

Corrections are necessary because neither data nor MC simulations provide an accurate description of the physics processes or the detector effects. They can be categorised as object-level (e.g. leptons, taus, jets), event-level (e.g. pileup, MET), and process-level (e.g. top p_T) corrections.

Discrepancies in data arise from imperfect detector response, finite detector resolution, inefficiencies in reconstruction and identification, and the gradual ageing of detector components. For MC samples, some global event properties, such as MET, pileup, or the number of reconstructed jets, cannot be modelled perfectly due to limitations in simulating physics processes and detector effects. To account for these differences, scale factors are applied to correct object identification efficiencies, energy scales, etc..
Therefore, the following set of calibrations is applied within the analysis pipeline:

- **Tau identification efficiency and energy scale corrections:** To correct for differences in tau reconstruction and identification between data and simulation, and to ensure an accurate modelling of both signal and background yields.

- **Tau trigger efficiency corrections:** To correct for residual mismodelling of the tau trigger response in simulation.

- **Impact Parameter (IP) vector and significance corrections for leptons:** To improve the modelling of decay planes and to enhance the separation between prompt and non-prompt leptons.

Table 4.2: Overview of the simulated signal and background samples used in this analysis. Datasets are listed with their corresponding physics processes, including phase-space restrictions where applicable. Signal samples are divided by production mode and by CP hypothesis: even (SM), odd (PS), and maximally mixed (MM). The generator and parton-shower setups are given, with spin correlations in brackets and multi-jet matching schemes where applicable. Labels and colours are defined consistently for all plots throughout the thesis. Signal samples include ggFs and VBFs Higgs production with different CP configurations, while background samples cover Drell–Yan, W+jets, single-top, $t\bar{t}$, and Di-Boson processes. The Generators are Powheg, aMCatNLO, and MadGraph, with Pythia 8 for parton showering. The Spin-correlation tools are MadSpin and TauSpinner. The matching tools are FxFx and MLM.

Dataset / Process / Phase space		**Generator**	**Label**	**Colour**
$ggF \to H \to \tau\tau$	CP even CP odd CP maxmix	Powheg_MiNLO (+ TauSpinner) + Pythia 8	ggH SM ggH PS ggH MM	
$VBF \to H \to \tau\tau$	CP even CP odd CP maxmix	Powheg (+ TauSpinner) + Pythia 8	qqH SM qqH PS qqH MM	
$Z/\gamma^* \to \tau\tau$ $m_{\tau\tau} > 50\,\text{GeV}$	inclusive 0 Jets 1 Jet 2 Jets	aMCatNLO + FxFx matching + Pythia 8	$Z \to \tau\tau$ + jet fakes	
$Z/\gamma^* \to \ell\ell$ $m_{\ell\ell} > 50\,\text{GeV}$	inclusive 0 Jets 1 Jet 2 Jets	aMCatNLO + FxFx matching + Pythia 8	$Z \to \ell\ell$	
$Z/\gamma^* \to \ell^+\ell^-$ $m_{\ell\ell} \in (10, 50)\,\text{GeV}$	inclusive			
W+jets $W \to l\nu$	inclusive 1 Jet 2 Jets 3 Jets 4 Jets	MadGraph + MLM matching + Pythia 8	W + jets	
Single $t/\bar{t}$	$t \xrightarrow{W^+} \ell^+\nu\, b$ $\bar{t} \xrightarrow{W^-} \ell^-\bar{\nu}\,\bar{b}$	Powheg (+ MadSpin) + Pythia 8	Single $t/\bar{t}$	
$t(\xrightarrow{W} b)W$	$(\ell^+\nu\, b)(\ell^-\bar{\nu})$ $(\ell^-\bar{\nu}\,\bar{b})(\ell^+\nu)$ $(\ell^+\nu\, b)(q\bar{q}')$ $(\ell^-\bar{\nu}\,\bar{b})(q\bar{q}')$	Powheg + Pythia 8		
$t\bar{t}$	WW leptonic WW semi-leptonic WW hadronic	Powheg + Pythia 8	$t\bar{t}$	
WW WZ ZZ	inclusive inclusive inclusive	Pythia 8	Di-Boson	

- **MET recoil corrections:** To improve the modelling of MET-dependent observables, recoil corrections derived from $Z \to \mu\mu$ events in data, are applied to the simulated samples. This ensures a more accurate description of the hadronic recoil and, consequently, of the MET distribution.

- **Pileup reweighting:** To match the luminosity profile observed in data and to correct for mismodelling of the event multiplicity in simulation.

- **Jet energy scale and resolution corrections:** To reproduce the measured jet energy response and resolution in data, ensuring that jet kinematics and related observables are accurately modelled.

- **Lepton (muon) identification, isolation, and trigger scale factors:** To account for differences in muon selection and trigger efficiencies between data and simulation, ensuring a consistent description of muon observables.

4.2.2 Selection of the $\mu\tau_{had}$ and μa_1^{3pr} Channel

This analysis targets the $H \to \tau\tau \to \mu\tau_{had} + 3\nu$ final state, where one tau lepton decays leptonically into a muon and the other decays hadronically into three charged pions via the intermediate a_1 resonance, both accompanied by one and two neutrinos, respectively. This channel provides a favourable compromise between a clean experimental signature, ensured by the presence of a high-p_T muon, and a good sensitivity to the tau decay kinematics through the hadronic three-prong system. In particular, the presence of three charged tracks allows the reconstruction of the SV of the hadronic tau decay, providing additional constraints on the tau decay geometry and kinematics.

The data used in this analysis are collected with a single-muon trigger requiring an isolated muon with $p_T > 24\,\text{GeV}$ and $\eta < |2.4|$. Offline, the reconstructed muon in matched to the triggered one. Its transverse momentum is required to exceed 26 GeV to operate in the plateau region of the trigger efficiency. The object selection is performed according to the recommendations of the *TauPog* for the muons [74] and tau leptons [75]. Tables 4.3 and 4.4 summarise the selection cuts applied to the reconstructed muons, hadronic tau leptons, and the pairs they form, from which the $\mu\tau_{had}$ candidate pairs are built. It is performed on the following: The relative muon isolation, `PF_RelIso_0.4_all`, which is defined as the sum of the transverse momenta of charged hadrons, neutral hadrons, and photons within a cone of $\Delta R = 0.4$, corrected for pileup, divided by the muon p_T. The IP significance is the ratio of the transverse IP to its uncertainty. The DeepTau discriminators against jets, electrons, and muons (`VSjet`, `VSe`, `VSmu`) are used to reduce background from misidentified objects, with working points chosen according to the *TauPog*. The visible mass is the invariant mass of the reconstructed di-tau system on detector-level, and the transverse mass m_T is calculated from the muon transverse momentum and MET.

If such candidates are identified as $DM10$ by the PNet algorithm and carry the `has_refit_sv` flag, they are classified as μa_1^{3pr} pairs, for which a reconstructed SV of the hadronic tau decay is available. The resulting event yield per channel and data-taking era is shown in Table C.1.

Table 4.3: Selection criteria for muon and hadronic tau candidates in the $\mu\tau_{had}$ channel.

Requirement	**Muon**	**Hadronic τ**
Transverse momentum	$p_T > 26\,\text{GeV}$	$p_T > 20\,\text{GeV}$
Pseudorapidity	$\lvert\eta\rvert < 2.4$	$\lvert\eta\rvert < 2.3$
Identification	Medium ID	ParticleNet
Allowed decay modes	–	10 (0, 1, 2, 11)
Relative isolation	$PF_RelIso_0.4_all < 0.15$	–
DeepTau VSjet	–	VTight
DeepTau VSmu	–	Tight
DeepTau VSe	–	VVLoose
Transverse impact parameter	$\lvert d_{xy}\rvert < 0.045\,\text{cm}$	–
Longitudinal impact parameter	$\lvert d_z\rvert < 0.2\,\text{cm}$	–
IP significance	$IPsig > 1.25$	$IPsig_{\pi^0} > 1.0$

Table 4.4: Selection criteria applied at the pair and event level in the $\mu\tau_{had}$ channel.

Requirement	**Value**
Lepton pair	One $\mu\tau_{had}$ pair
Charge	Opposite sign
Visible mass	$m_{vis} > 40\,\text{GeV}$
Transverse mass	$m_T < 65\,\text{GeV}$
Angular separation	$\Delta R(\mu, \tau) > 0.5$
Tau GenPartFlavor	1 to 5
Extra lepton veto	No additional leptons in the event
Di-lepton veto	No di-lepton pairs of the same light-lepton flavour

4.3 Event Categorisation

After calibration and selection of μa_1^{3pr} candidate pairs, they are categorised according to the physical process from which they originate. This is necessary because the selection only applies cuts on the measured final-state particles and event variables, which can be similar for different physics processes. Consequently, the selected sample contains not only the Higgs signal but also background events.

To distinguish these processes, the analysis by the Institut pluridisciplinaire Hubert Curien (IPHC), Imperial College London (IC), and Deutsches Elektronen-Synchrotron (DESY) employs a BDT classifier, trained by IC, to classify events into three categories. Each event is assigned a score between 0 and 1, reflecting its probability of belonging to the respective category:

- **Signal:** Events classified as signal, originating from Higgs boson production via ggF or VBF. VH and ttH are also classified as signal, but are not considered in this thesis.

- **Genuine tau leptons:** Events where both selected candidates originate from true tau leptons, such as in $Z \to \tau\tau$, $t\bar{t}$, or di-boson processes.

- **Jets faking taus:** Events where one or both selected candidates are misidentified jets or leptons, corresponding to backgrounds such as $Z \to \ell\ell$, $Z \to$ jets, $t\bar{t}$, $W+$jets, di-boson, and QCD processes.

4.4 Measuring the ϕ_{CP} Angle

As described in subsection 2.1.4 and summarised in subsection 2.2.2, the CP nature of the tau Yukawa coupling, and thus of the Higgs boson, can be investigated through the acoplanarity angle ϕ_{CP}, defined as the angle between the decay planes of the Higgs boson's decay products. This angle is experimentally accessible because the differential decay width, and therefore the production cross section, exhibits a characteristic dependence on ϕ_{CP}, which originates from the spin correlations of the decay products. Although the decay $\tau \to a_1^{3pr}\nu \to 3\pi\nu$ corresponds to a four-body final state, in practice, it is dominated by the sequential two-body chain $\tau \to a_1^{3pr}\nu$, $a_1^{3pr} \to \rho\nu$, and $\rho \to 2\pi$. In the a_1 rest frame, this effectively constrains the three pions to a common decay plane.

A plane can be uniquely defined by two non-collinear vectors if a common point of origin is specified. In the present analysis, this reference point is taken to be the decay vertex of the Higgs boson decay. These two vectors used to characterise the decay kinematics depend on the decay channel considered and the analysis strategy followed. In all cases, one of these vectors is given by the momentum of the respective visible decay product, while the second vector is constructed from the IP, the Decay Plane (DP), or the Polarimetric Vector ($\vec{h}_{PV}$) method. The construction of these vectors is described below. First, however, the general formalism [76] used to extract the angle ϕ_{CP} between two decay planes is introduced, as it is independent of the specific choice of vector pairs and decay channels.

The decay planes of the two tau leptons are each defined by a pair of vectors:

$$\vec{\mathcal{P}}^{\pm} : \text{momentum of the charged decay product,}$$
$$\vec{\mathcal{R}}^{\pm} : \text{reference vector, with } \vec{\mathcal{R}}^{\pm} \neq \vec{\mathcal{P}}^{\pm}.$$

In order to maximise the sensitivity to CP-violating effects, all vectors are boosted into the centre-of-mass frame (CMF), in which the sum of the particle momenta vanishes. This yields

$$\vec{\mathcal{P}}^{\pm} \xrightarrow{\text{boost}} \vec{P}^{\pm}, \qquad \vec{\mathcal{R}}^{\pm} \xrightarrow{\text{boost}} \vec{R}^{\pm}.$$

In this frame, only the transverse components of the reference vectors carry information on the orientation of the decay planes. Therefore, the vectors are projected onto the plane perpendicular to $\vec{P}^{\pm}$:

$$\vec{R}^{\pm}_{\perp} = \vec{R}^{\pm} - (\vec{R}^{\pm} \cdot \vec{P}^{\pm}) \frac{\vec{P}^{\pm}}{|\vec{P}^{\pm}|^2}.$$

The unsigned acoplanarity angle φ^*, defined as the angle between the two decay planes, is then given by

$$\varphi^* = \arccos\left(\frac{\vec{R}^{-}_{\perp}}{|\vec{R}^{-}_{\perp}|} \cdot \frac{\vec{R}^{+}_{\perp}}{|\vec{R}^{+}_{\perp}|} \right). \tag{4.1}$$

Since this definition alone only yields an angle in the range $[0, \pi]$, the orientation of the two planes is recovered by introducing the triple product $\mathcal{O} = (\vec{R}_\perp^+ \times \vec{R}_\perp^-) \cdot \vec{P}^-$. The acoplanarity angle is then defined as

$$\phi_{CP} = \begin{cases} \varphi^* & \text{if } \mathcal{O} \geq 0, \\ 2\pi - \varphi^* & \text{if } \mathcal{O} < 0. \end{cases} \tag{4.2}$$

In multi-body decays, such as those considered in the $\vec{h}_{PV}$ method, the orientation of the $\vec{h}_{PV}$ can change depending on the energy distribution among the decay products. To account for this, the variable y_i is introduced:

$$y_i = \frac{E_{\text{charged}} - E_{\text{neutral system}}}{E_{\text{charged}} + E_{\text{neutral system}}}, \quad i = +, -,$$

where E_{charged} is the energy of the charged decay product and $E_{\text{neutral system}}$ is the total energy of the associated neutral particles.
For three-prong decays, such as tau leptons decaying via an a_1 resonance into three pions and a neutrino, y_i is defined by the two oppositely charged pions whose invariant mass is closest to the ρ resonance. In this case, y_i reduces to

$$y_{a_1{}^{3pr}} = \frac{E_{\pi^\pm} - E_{\pi^\mp}}{E_{\pi^\pm} + E_{\pi^\mp}}, \tag{4.3}$$

and a phase shift of π is applied to ϕ_{CP} whenever $y_{a_1{}^{3pr}} < 0$, ensuring a consistent orientation of the $\vec{h}_{PV}$.

Table 4.5 summarises the vector pairs, $\vec{P}^\pm$ and $\vec{R}^\pm$, used to define the decay planes of the two tau leptons for each of the three methods considered in this analysis: IP, DP, and $\vec{h}_{PV}$. These methods are described in detail in the subsequent subsections.

Table 4.5: Definition of the vector pairs used to construct the decay planes of the two tau leptons for the three methods Impact Parameter (IP), Decay Plane (DP), and Polarimetric Vector ($\vec{h}_{PV}$), $\vec{h}$, considered in this analysis. $\vec{p}_p$ represents the momentum of the visible decay product.

Method	$\vec{P}^+$	$\vec{R}^+$	$\vec{P}^-$	$\vec{R}^-$	**Phase Shift**
	τ^+		τ^-		
Impact Parameter (IP)	$\vec{p}_p$	$\overrightarrow{IP}$	$\vec{p}_p$	$\overrightarrow{IP}$	-
Decay Plane (DP)	$\vec{p}_\pi$	$\vec{p}_{\pi^0}$	$\vec{p}_\pi$	$\vec{p}_{\pi^0}$	$y_+ \cdot y_- < 0$
Polarimetric Vector ($\vec{h}_{PV}$)	$\vec{p}_\tau$	$\vec{h}$	$\vec{p}_\tau$	$\vec{h}$	$y_{a_1{}^{3pr}} < 0$

As this thesis focuses on the semi-leptonic final state (Equation 2.19) with $\tau_\ell \to \mu\nu_\mu\nu_\tau$ and $\tau_{had} \to a_1^{3pr}\nu_\tau$, the decay plane of the leptonic tau is reconstructed using the IP method. In contrast to the Run 2 analysis, the decay plane of the hadronic tau is reconstructed using the $\vec{h}_{PV}$ approach. The combined reconstruction of both decay planes and the resulting CP-sensitive acoplanarity angle ϕ_{CP} are illustrated in Figure 4.1.

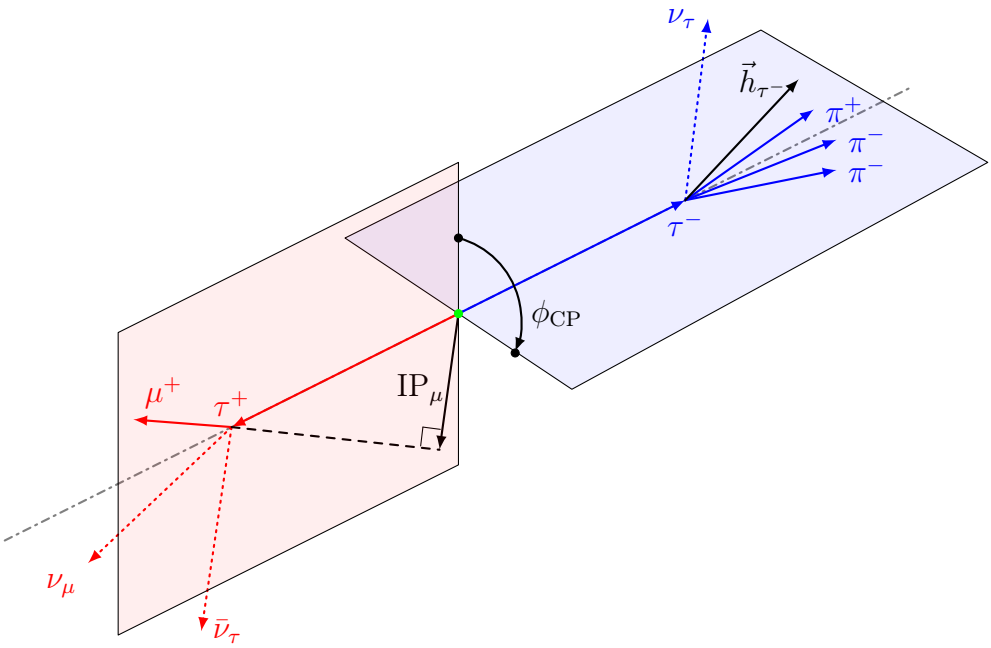


Figure 4.1: Schematic illustration of the decay plane reconstruction in the $\tau^+ \to \mu^+\nu_\mu\bar{\nu}_\tau$ and $\tau^- \to a_1^{3pr}\nu_\tau$ channel. The decay plane of the leptonic tau is reconstructed using the IP method (red), while the decay plane of the hadronic tau is reconstructed using the $\vec{h}_{PV}$ method (blue). The polarimetric vector $\vec{h}_{\tau^-}$ corresponds to the most probable tau spin direction. The acoplanarity angle ϕ_{CP} is defined as the angle between the two reconstructed decay planes in the CMF of the visible tau decay products.

4.4.1 Impact Parameter Method

The Impact Parameter (IP) method is in principle applicable to all tau decays, but it is typically used for one-prong decays such as $\tau \to \mu, \pi, a_1^{1pr} + \nu$. Previous Run 2 studies [14], have shown that for the a_1^{1pr} decays, the reconstruction is largely insensitive to CP effects due to contributions with different a_1 polarisations. Consequently, the DP or $\vec{h}_{PV}$ methods are preferred in this case.

The method is based on the relatively long tau lifetime, which allows the tau to travel a measurable distance from the primary vertex (PV) before decaying. Conceptually, the IP is defined as the vector connecting the tau production vertex PV to the tau decay vertex SV, perpendicular to the momentum of the charged decay product. However, for one-prong decays, the SV cannot be reconstructed experimentally. Instead, the point of closest approach of the extrapolated charged track to the PV is used to define the IP vector. To describe the geometry, the PV, corresponding to the production vertex of the Higgs boson, the estimated SV, and the momentum of the visible decay product $\vec{p}$ are introduced. The direction of the charged decay product is given by the unit vector $\hat{d}$. The vector connecting the tau production vertex to the decay vertex is

$$\vec{L} = \overrightarrow{SV} - \overrightarrow{PV} \tag{4.4}$$

The component of $\vec{L}$ parallel to the particle trajectory is obtained by projecting $\vec{L}$ onto $\hat{d}$,

$$\vec{L}_{\parallel} = (\vec{L} \cdot \hat{d})\,\hat{d}.$$

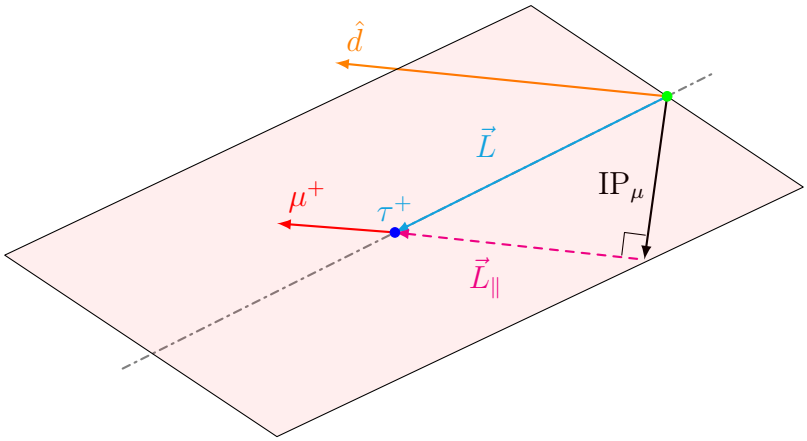


Figure 4.2: Geometrical construction of the impact parameter vector for a one-prong tau decay into a muon ($\tau^+ \to \mu^+\nu_\mu\bar{\nu_\tau}$). The tau lepton is produced at the primary vertex (green point) and decays at the secondary vertex (blue point) into a muon. The vector $\vec{L}$ connects the primary and secondary vertices. The muon momentum (red arrow) defines the unit direction vector $\hat{d}$ (orange arrow). The component of $\vec{L}$ parallel to the track direction (pink arrow), $\vec{L}_\parallel$, is obtained by projecting $\vec{L}$ onto $\hat{d}$. The impact parameter vector $\overrightarrow{IP}_\mu$ (black arrow) is then defined as the orthogonal component of $\vec{L}$ and corresponds to the shortest vector connecting the primary vertex to the extrapolated muon trajectory. The shaded plane indicates the decay plane of the leptonic system, while the neutrinos are omitted for clarity.

The IP vector is then defined as the component of $\vec{L}$ perpendicular to the track direction:

$$\overrightarrow{IP} = \vec{L} - \vec{L}_\parallel \tag{4.5}$$

By construction, the impact parameter vector $\overrightarrow{IP}$ is orthogonal to the momentum direction of the charged decay product and represents the shortest vector connecting the tau production vertex to the extrapolated particle trajectory. The decay plane is then reconstructed following the procedure outlined at the start of this section.
Although the IP method is straightforward in concept, its performance depends on the accuracy of the PV reconstruction. The nominal PV is reconstructed from the event tracks only. This analysis uses the beam spot-corrected PV, which incorporates the known beam geometry and has been shown to improve the reconstruction accuracy [77].

4.4.2 Decay Plane Method

The Decay Plane (DP) method provides a supplementary approach to the reconstruction of decay planes, particularly suited for hadronic tau decays with multiple final-state particles, such as $\tau \to \rho, a_1^{1pr} + \nu$ decays. It can also be applied to three-prong decays, following the Run 2 procedure [14], in which two pions are first combined to form a ρ^0 meson, which is then treated as a single charged prong, as described below.

The DP method is a simplification of the $\vec{h}_{PV}$ method, based on several assumptions. First, the mass difference between the charged and neutral pions is neglected. Under these assumptions, the polarimetric vector $\vec{h}_{PV}$ can be expressed in the simplified form $\vec{h} \propto 2m_\tau(E_{\pi^\pm} - E_{\pi^0})\vec{q} - q^2\hat{n}$, so that both the magnitude and, in particular, the sign of $\vec{h}$

depend solely on the energy difference of the charged and neutral pions. This measurement allows the events to be classified into $y^+ \cdot y^- > 0$ and $y^+ \cdot y^- < 0$. Second, in the tau rest frame, the reduced $\vec{h}_{PV}$ lies in the $\pi^\pm\pi^0\nu$ plane. Third, in the laboratory frame, the neutrino momentum is neglected, i.e. $\vec{p}_\nu \approx 0$, which allows the tau flight direction to be approximated as lying within the $\pi^\pm$–π^0 plane. Consequently,

$$\vec{h} \perp (\vec{p}_{\pi^\pm} \times \vec{p}_{\pi^0}) \perp \vec{p}_\tau.$$

Furthermore, the $\vec{h}_{PV}$ can then be effectively replaced by the neutral pion momentum, under the assumption that

$$\vec{p}_{\pi^\pm} \times \vec{p}_{\pi^0} \parallel \vec{h} \times \vec{p}_\tau.$$

This allows the tau direction to be replaced by $\vec{p}_{\pi^\pm}$ and the $\vec{h}_{PV}$ by $\vec{p}_{\pi^0}$, so that the acoplanarity angle can be calculated in the CMF of the two charged pions, using $\pi^\pm$ as $\vec{P}^\pm$ and π^0 in place of $\vec{h}$ as $\vec{R}^\pm$ (see Table 4.5).

In practice, the decay plane is defined using the neutral pion momentum, or the sum of momenta for multiple neutral pions. To avoid negative interference from differently polarised states, the phase-shift variable is introduced before applying the method described above. While the DP method provides a plane correlated with the true tau decay plane, it should be noted that it does not exactly reconstruct the real decay plane.

4.4.3 Polarimetric Vector Method

The Polarimetric Vector ($\vec{h}_{PV}$) method has not been used in the corresponding Run 2 analysis for the μa_1^{3pr} channel. It exploits the direct correlation between CP-sensitive observables and the tau spin, as expressed by the differential decay rate

$$\mathrm{d}\Gamma \propto (1 + \vec{h} \cdot \vec{s}).$$

Here, $\vec{s}$ denotes the tau spin and $\vec{h}$ the $\vec{h}_{PV}$. Therefore, $\mathrm{d}\Gamma$ is maximal when $\vec{h}$ is aligned with $\vec{s}$, meaning that $\vec{h}$ most likely points in the direction of the tau spin. In practice, the $\vec{h}_{PV}$ can be estimated from the momenta of the hadronic decay products and the decay model of the mediating resonance, as implemented in TAUOLA [78] and calibrated with measurements from the *CLEO* collaboration [79]. It is illustrated in Figure 4.3 for the μa_1^{3pr} case.

The $\vec{h}$ is defined according to the resonance model of the a_1. Due to information loss from undetected neutrinos, the tau momenta reconstruction is improved using a combination of the *fastMTT* method for the modular and the *GEF* algorithm for the angular information, as described in the following subsection 4.5.1 and subsection 4.5.2. Once the $\vec{h}_{PV}$ is computed from the hadronic decay products, the method described at the beginning of this section is applied to extract the CP-sensitive observables.

4.5 Kinematic Reconstruction

As a consequence of neutrinos in the studied $\tau_\ell\tau_{had}$ final state, the tau lepton momenta cannot be fully reconstructed from the visible decay products, as part of the kinematic information is carried away by undetected neutrinos. This loss of information is mitigated

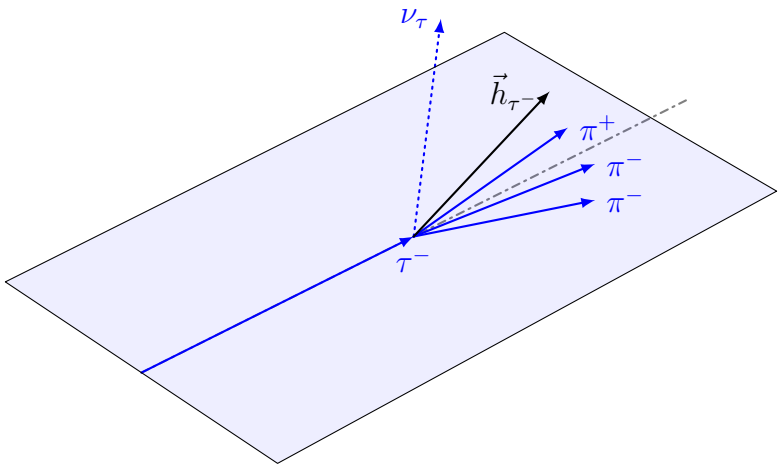


Figure 4.3: Illustration of the polarimetric vector method for a $\tau^- \to (a_1^{3pr})^- \nu_\tau$ decay. The tau lepton decays into three charged pions and a neutrino. The polarimetric vector $\vec{h}_{\tau^-}$ (black arrow) is reconstructed from the momenta of the hadronic decay products using the a_1 resonance model and defines the most probable direction of the tau spin in the tau rest frame. The shaded plane indicates the decay plane of the hadronic system.

using a two-step reconstruction procedure. First, the magnitudes of the tau leptons momenta are estimated using a dynamic likelihood-based scan constrained to the Higgs boson mass. In the second step, the angular components are refined while accounting for kinematic limits being surpassed in some events and maintaining them artificially. The first step is performed using the fastMTT algorithm, while the second step uses the GEF method.

4.5.1 fastMTT Method

Di-tau final states are accompanied by at least two neutrinos. Since neutrinos escape detection due to their extremely small interaction cross section, the complete event kinematics cannot be reconstructed from the visible decay products alone. In recorded collision events, all information about the invisible part of the final state is therefore encoded in the MET observable (see subsection 3.3.5).

The fastMTT algorithm exploits this information by using the reconstructed MET together with the four-momenta of the visible tau decay products to test hypotheses for the invariant mass of the di-tau system, and thereby to estimate the mass of the SM Higgs boson by varying the kinematic configuration of the invisible neutrino system. Historically, this problem was addressed in CMS using a likelihood approach based on the full matrix element of the process: $pp \to H \to \tau\tau$. In order to reduce the prohibitive computational cost of this method, the classic Secondary Vertex Fit (cSVfit) algorithm [58] was introduced. Instead of evaluating the full matrix element, it integrates a likelihood function over all unknown parameters of the neutrino system. This approach achieves an excellent mass resolution, but remains slow for large-scale applications. To further reduce the computational cost while retaining a comparable performance, the fastMTT algorithm [56, 57] was developed as a fast approximation to cSVfit.

Principle of the fastMTT Algorithm

A tau decay can be decomposed into a visible system, with four-momentum $p_{vis} = (E_{vis}, \vec{p}_{vis})$, and an invisible system consisting of neutrinos, described by the measured missing transverse momentum. Its kinematics can be parametrised in terms of the energy fraction carried by the visible system, $x = \frac{E_{vis}}{E_\tau}$, the Gottfried Jackson angle θ_{GJ} (GJ angle), which describes the opening angle between the tau lepton momentum and the direction of the visible decay products (see Figure 4.5), the azimuthal angle φ, and the invariant mass of the neutrino system $m_{\nu\nu}$. For a di-tau system, this results in six free parameters, three for each tau decay.
To motivate the fastMTT approach, the cSVfit method is briefly outlined in the following. It is based on the evaluation of a multidimensional likelihood integral (see Equation 4.6), which computes the likelihood for a given mass hypothesis, such as the SM Higgs boson mass, to be realised given the measured event observables by integrating over all unknown parameters of the neutrino system

$$\mathcal{L}(m_{\tau\tau}^{test} \mid \text{event}) = \int \delta\big(m_{\tau\tau}(\vec{a}) - m_{\tau\tau}^{test}\big)\; p(\vec{y} \mid \vec{a})\, \mathrm{d}\vec{a}, \tag{4.6}$$

where $\vec{a}$ denotes the set of unknown parameters describing the neutrino system and $\vec{y}$ the set of measured event observables. The probability density $p(\vec{y} \mid \vec{a})$ is factorised as

$$p(\vec{y} \mid \vec{a}) = ME(\vec{a}) \cdot TF(\vec{y} \mid \vec{a}) \cdot REG(m_{\tau\tau}), \tag{4.7}$$

where ME is the matrix-element term describing the tau decay kinematics and TF denotes the transfer-function term, which models the detector response and, in practice, is dominated by the MET transfer function. Since MET is the only experimental observable that directly constrains the neutrino system, the MET transfer function describes the probability that a true missing transverse momentum $\vec{E}_T^{\text{miss,hypo}}$ is reconstructed in the detector as $\vec{E}_T^{\text{miss,reco}}$. It is parametrised as a two-dimensional Gaussian using the event-by-event MET covariance matrix. Furthermore, REG is a regularisation term, parametrised as a power law in $m_{\tau\tau}$, which stabilises the likelihood and suppresses unphysical regions of phase space.

The simplification of this procedure relies on the following assumptions and observations:

1. Since the hadronic decay $\tau \to hadrons + \nu_\tau$ involves exactly one neutrino, the invariant mass of the neutrino system vanishes, $m_{\nu\nu} = 0$.

2. The decay is assumed to be collinear, such that $\theta_{GJ} \approx 0$, as the true angles are typically very small. While this approximation significantly reduces the computational cost, it also has a strong impact on the CP sensitivity and is therefore discussed separately in subsection 4.5.2. As a consequence, the azimuthal angle φ is no longer defined and is thus removed from the parameter set.

3. The method assumes isotropic decays and a two-body or three-body phase-space model for hadronic and leptonic tau decays, respectively, with the hadronic system treated as a single particle. The matrix-element term is set to unity, $ME = 1$, as its impact on the likelihood is dominated by the MET transfer function. An empirical regularisation is introduced to compensate for the removal of the matrix-element term, which rescales the mass hypothesis $m_{\tau\tau} \to \alpha\; m_{\tau\tau}$ and applied a power-law

weighting of the reconstructed mass $m_{\tau\tau}^3 \to m_{\tau\tau}^\beta$ with parameters α and β chosen to reproduce the mass scale and resolution of the original cSVfit algorithm:

$$\alpha = \frac{1}{1.15} \qquad \beta = 6 \tag{4.8}$$

4. The computational cost is further reduced by replacing the likelihood integration by a maximisation on a discrete 100×100 grid in the (x_1, x_2) plane.

The remaining degrees of freedom are the energy fractions of the two tau leptons, x_i, and their corresponding four-momenta $p_{\tau,i}$, given by

$$x_i = \frac{E_{\text{vis},i}}{E_{\tau,i}} \qquad \text{and} \qquad p_{\tau,i} = \frac{p_{\text{vis},i}}{x_i}, \qquad i = 1, 2. \tag{4.9}$$

Finally, after removing the matrix-element term, the angular parameters, and the integration, the neutrino momenta, thus the hypothetical missing transverse momentum, are fully determined as a function of (x_1, x_2), and the likelihood takes the simplified form

$$\mathcal{L}(m_{\tau\tau}^{\text{test}} \mid \text{event}) = \max_{x_1,x_2} MET_TF\left(\vec{E}_T^{\text{miss,reco}} \mid \vec{E}_T^{\text{miss,hypo}}(x_1, x_2)\right). \tag{4.10}$$

The fastMTT algorithm is able to achieve a computational speed-up of about two orders of magnitude with respect to cSVfit, while retaining a very similar mass resolution. This makes it particularly well-suited for systematic uncertainty scans, multiple per-event evaluations, and large-scale production workflows.

Implementation of the fastMTT Algorithm

In the fastMTT framework, modifiers x_i are determined for each tau lepton decay, labelled $i = 0, 1$. Using these modifiers, the corrected tau momenta, energies, and masses, accounting for the unobserved neutrinos, are defined as

$$\vec{p}^{\,ij} = \frac{\vec{p}_{vis}^{\,ij}}{x_i}, \qquad E_i = \sqrt{|\vec{p}^{\,i}|^2 + m_\tau^2}, \qquad \text{and} \qquad m_i = \frac{\sqrt{E_i^2 - |\vec{p}^{\,i}|^2}}{x_i}, \tag{4.11}$$

where $j \in \{x, y, z\}$ denotes the spatial components and m_τ is the SM tau mass. The invariant di-tau mass $m_{\tau\tau}$, which for signal events is expected to peak at the Higgs boson mass, is computed as

$$m_{\tau\tau} = \frac{m_{\text{vis}}}{\sqrt{x_0\, x_1}}.$$

The effect of missing neutrinos on the di-tau mass reconstruction, as well as the improvement achieved by fastMTT, is illustrated in Figure 4.4. The visible invariant mass distribution peaks far below the Higgs boson mass, while the fastMTT reconstruction shifts the distribution towards the correct mass scale. The generator-level distribution represents the ideal case and indicates the best achievable resolution.

Mass Constraint in the fastMTT Algorithm

In the fastMTT implementation used in this analysis, the simplified likelihood (Equation 4.10) is supplemented by an empirical regularisation, parametrised by α and β (Equation 4.8). Depending on the analysis purpose, different treatments of the reconstructed

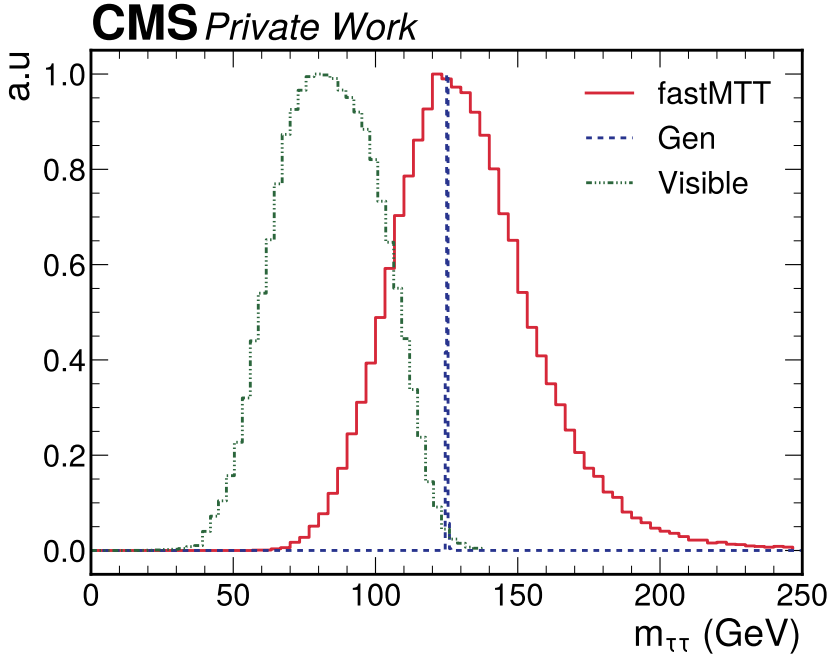


Figure 4.4: Comparison of the expected $m_{\tau\tau}$ distributions at different reconstruction levels. The visible mass (green) peaks far below the Higgs boson mass. The fastMTT-reconstructed mass (red) shifts the distribution towards the Higgs boson mass. The generator-level mass (blue) represents the best-case scenario, with a peak with the generated Higgs width around the Higgs boson mass.

di-tau mass can be applied. When used indirectly, e.g., to optimise the tau momenta for computing ϕ_{CP}, a mass constraint can improve the reconstruction by further constraining the fastMTT-reconstructed mass towards the ideal generator-level value. However, it is not applied when $m_{\tau\tau}$ enters directly, e.g., as an input feature for BDT training, since this would be unphysical. Three common approaches are considered:

- An unconstrained reconstruction,
- A mass-window requirement,

$$m_{\tau\tau} \in [m_{\text{min}}, m_{\text{max}}], \tag{4.12}$$

- Or a Breit–Wigner constraint centred on the Higgs boson mass,

$$BW(m_{\tau\tau}) = \frac{1}{(m_{\tau\tau}^2 - m_H^2)^2 + m_H^2 \Gamma_H^2} \tag{4.13}$$

  with $m_H = 125.08\,\text{GeV}$, and the Higgs boson width Γ_H set either to $4\,\text{MeV}$, as in the generated MC sample, or to $\Gamma_H = 0.01\,m_H$, corresponding to the empirically optimal value determined in cSVfit studies.

4.5.2 Global Event Fit

The Global Event Fit (GEF) method [56, 58, 59] addresses the correction of the angular component of the tau momenta. For many analyses, it is sufficient to assume collinearity between the resonance and the tau lepton, however, in the ultra-relativistic approximation

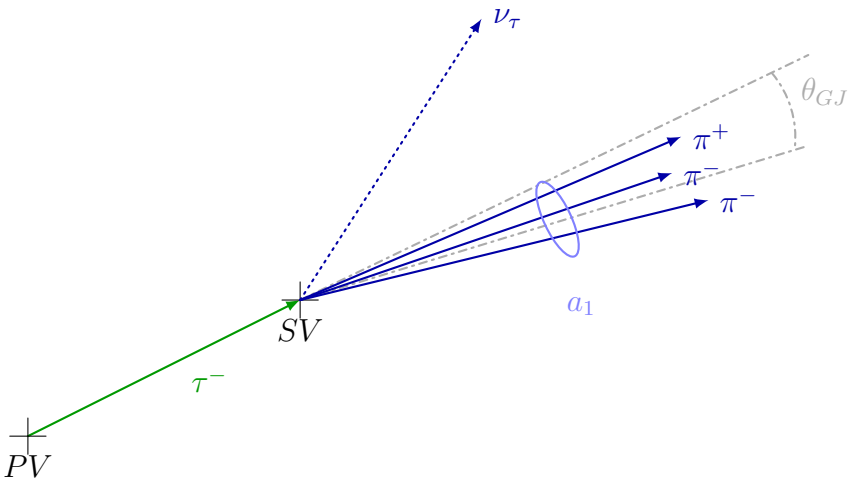


Figure 4.5: Schematic illustration of the Gottfried Jackson angle θ_{GJ} (GJ angle) in a single tau lepton decay to an a_1^{3pr} and a neutrino. The angle GJ angle is defined between the tau momentum direction (green arrow) and the momentum of the visible decay products (blue arrows). The dotted arrow represents the undetected neutrino, and the intermediate a_1 resonance is indicated by the ellipse.

and in decays involving a single neutrino, this assumption is no longer accurate. Events of this type, such as $\tau \to a_1^{3pr}\nu_\tau$ decays, can be illustrated as in Figure 4.5. To implement this correction, the GJ angle is computed from the reconstructed tau and visible decay momenta, with a kinematic upper limit imposed to ensure physically allowed configurations. Events exceeding this limit are rotated to the maximal allowed angle.
The following section describes the construction of the tau and visible momenta from the decay products, the subsequent computation of θ_{GJ}, and the rotation if necessary.

Reconstruction of the Visible and Tau Momenta

To begin with, the three-momenta $\vec{p}_{vis}$ and $\vec{p}_\tau$ are defined by converting the corresponding four-vectors p_{vis} and p_τ. In the considered three-prong tau decay, the visible momentum is the sum of its constituent pions, as illustrated in Figure 4.6:

$$p_{vis} = p_{\pi_{os}} + p_{\pi_{ss1}} + p_{\pi_{ss2}}, \tag{4.14}$$

here, π_{os} denotes the pion with a charge opposite to that of the tau lepton (and thus of the a_1 meson), while π_{ss1} and π_{ss2} are the same-sign pions. The first of these, π_{ss1}, together with π_{os} forms an intermediate ρ^0 resonance.
The tau four-vector is assigned the momentum magnitude corrected by the fastMTT algorithm. Its mass is set to the SM value, and its direction is approximated by the vector connecting the PV to the SV. To achieve the most accurate reconstruction, the beam spot-corrected PV is taken. In p_T, η, φ, m coordinates, the tau four-vector is thus given by

$$p_\tau = (p_T^{fastMTT},\ (SV - PV)_\eta,\ (SV - PV)_\varphi,\ 1.777\,\text{GeV})^T. \tag{4.15}$$

Kinematic Expression of the Gottfried Jackson Angle

As the GJ angle describes the opening between $\vec{p}_{vis}$ and $\vec{p}_\tau$, it is calculated from the cosine

$$\theta_{GJ} = \arccos \hat{n}_{vis} \cdot \hat{n}_\tau, \tag{4.16}$$

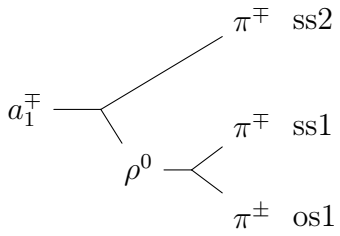


Figure 4.6: Decay chain of the three-prong a_1^{3pr} resonance, illustrating the naming convention of the pions. The decay products consist of one opposite-sign pion (π_{os}), carrying the charge opposite to the tau lepton, and two same-sign pions (π_{ss1}, π_{ss2}). The pions π_{os} and π_{ss1} form the intermediate ρ^0 resonance.

where $\hat{n}_{vis}$ and $\hat{n}_\tau$ are the unit vectors along the directions of the corresponding momenta. The construction in the decay plane is illustrated schematically in Figure 4.7:

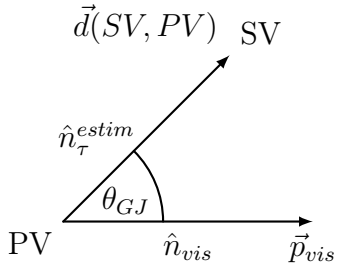


Figure 4.7: Geometrical construction of the Gottfried Jackson angle θ_{GJ}. The angle is defined between the direction of the visible momentum $\vec{p}_{vis}$ and the estimated tau flight direction obtained from the vector connecting the primary and secondary vertices.

To derive its kinematic form, the four-momenta of the tau lepton, the visible system, and the neutrino can be written as $p_i = (E_i, \vec{p}_i)$, $i = \tau, \text{vis}, \nu$, with masses assigned as follows: m_τ as the SM mass, m_{vis} from the reconstructed visible system, and m_ν is approximated as zero. By momentum conservation, $p_\tau = p_{vis} + p_\nu \Rightarrow p_\nu = p_\tau - p_{vis}$, the invariant masses satisfy $p_\tau^2 = m_\tau^2$, $p_{\text{vis}}^2 = m_{\text{vis}}^2$, and that for the neutrino gives:

$$p_\nu^2 = m_\nu^2 = (p_\tau - p_{vis})^2 = p_\tau^2 + p_{vis}^2 - 2p_\tau p_{vis} = m_\tau^2 + m_{vis}^2 - 2p_\tau p_{vis}$$

Solving the latter for the scalar product of the four-momenta yields

$$2p_\tau p_{vis} = m_\tau^2 + m_{vis}^2 - m_\nu^2. \tag{4.17}$$

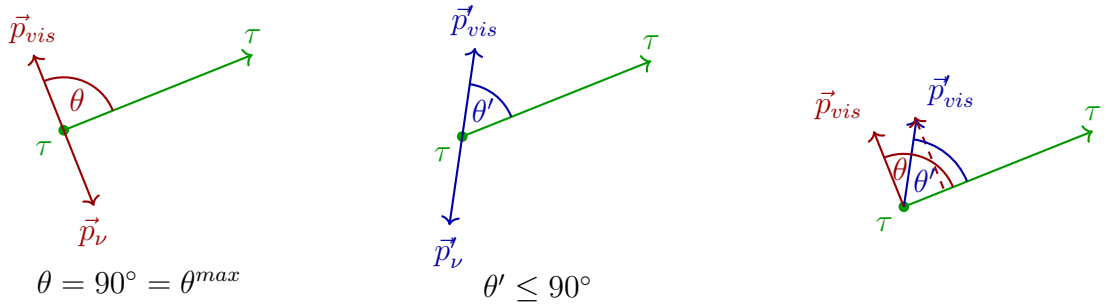
Thus, while expressing the scalar product in terms of energies and three-momenta, $p_\tau p_{vis} = E_\tau E_{vis} - \vec{p}_\tau \vec{p}_{vis}$, the GJ angle angle can be written in its kinematic form as

$$\theta_{GJ} = \arccos \frac{p_\tau p_{vis}}{|p_\tau||p_{vis}|} = \arccos - \frac{m_\tau^2 + m_{vis}^2 - m_\nu^2 \pm 2E_\tau E_{vis}}{2|p_\tau||p_{vis}|}. \tag{4.18}$$

Due to detector effects, measurement uncertainties, and other inaccuracies, it can happen that the kinematic limit is exceeded in some events, resulting in a GJ angle larger than physically allowed.

Derivation of the Kinematic Upper Limit

The kinematic limiting case of the GJ angle can be derived from a geometric argument. Consider the decay in the laboratory (green) and rest frame of the tau lepton (red and blue) as shown in Figure 4.8.



(a) Configuration in the tau RF, where the visible hadronic system and the neutrino are emitted back-to-back. The opening angle between the tau direction and the visible momentum reaches its maximal value of $\theta^{\text{max}} = 90°$.

(b) The same decay after a Lorentz boost to a different RF, which modifies only the angular configuration while leaving the momentum magnitudes unchanged.

(c) Superposition of the visible momentum vectors in different frames, illustrating that the maximal allowed opening angle in the laboratory frame is entirely fixed by kinematics.

Figure 4.8: Kinematic origin of the Gottfried Jackson angle limit (red θ) in the $\tau \to hadrons + \nu_\tau$ decay for different Lorentz boosted systems.

The angle θ is enclosed by the visible momentum and the tau direction, in the tau reference frame (RF) and in the laboratory RF, respectively. The maximal transverse momentum of the visible system with respect to the boost is reached at an angle configuration of 90° as shown in Figure 4.8a. When a different RF of the tau lepton is chosen (blue, Figure 4.8b), only the angular configuration changes, while the magnitudes of the particle momenta remain unchanged.
It follows: $|\vec{p}_{vis}| = const$, $|\vec{p}_\tau| = const$, and $E_{vis} = const$ (Equation 4.19) and furthermore due to the ultra-relativistic case:

$$\vec{p}_{vis} = -\vec{p}_\tau \;\Rightarrow\; |\vec{p}_{vis}| = |\vec{p}_\tau|, \tag{4.19}$$

$$E_{vis} + E_\nu = m_\tau. \tag{4.20}$$

Accounting for the on-shell-mass relations

$$E_{vis} = \sqrt{p_{vis}^2 + m_{vis}^2} \qquad \text{and} \qquad E_\nu = \sqrt{p_\nu^2 + m_\nu^2} \overset{m_\nu=0}{=} |\vec{p}_\nu|,$$

Equation 4.20 can be written as

$$\sqrt{p_{vis}^2 + m_{vis}^2} + |\vec{p}_\nu| = m_\tau. \tag{4.21}$$

Then, substituting Equation 4.19 into Equation 4.21 yields the expression

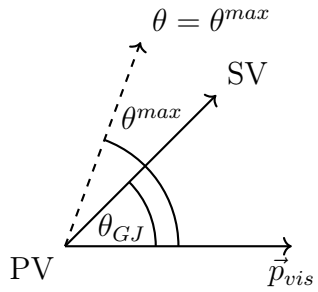
$$|\vec{p}_{vis}| = \frac{m_\tau^2 - m_{vis}^2}{2m_\tau} \tag{4.22}$$

It becomes evident that the vector of visible momentum in the laboratory frame (green) stays unchanged throughout differently boosted regimes (Figure 4.8a and Figure 4.8b), as it is expressed as masses. The vectors can be overlapped as shown in Figure 4.8c and thus the enclosed angle $\theta = \theta^{max}$ can be expressed as

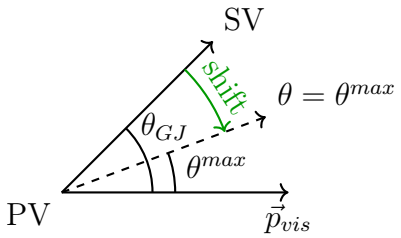
$$\theta^{max} = \arcsin \frac{|\vec{p}_{vis}^{\,\tau\,RF}|}{|\vec{p}_{vis}^{\,lab}|} = \arcsin \frac{m_\tau^2 - m_{vis}^2}{2m_\tau |\vec{p}_{vis}^{\,lab}|}. \tag{4.23}$$

Rotation Procedure for Unphysical Configurations

In the analysis, the observed and maximal angles will be calculated following Equation 4.18 and Equation 4.23. If the angle exceeds the limiting value, the tau momentum is minimally rotated inside the decay plane back to the physical phase-space boundary, while preserving its magnitude and energy as shown below and in Figure 4.9b.



(a) Event configuration in which the opening angle θ_{GJ} between the visible momentum and the reconstructed tau direction is smaller than the maximally allowed value $\theta^{\max}$. No correction to the event is made.



(b) The opening angle θ_{GJ} exceeds the kinematic maximum and therefore, the tau momentum is minimally rotated within the decay plane until the configuration reaches the physical boundary $\theta_{GJ} = \theta^{\max}$, while preserving both its magnitude and energy.

Figure 4.9: Illustration of the correction procedure applied when the reconstructed Gottfried Jackson angle exceeds its kinematic upper bound.

The correction is performed entirely in three-dimensional space using the spatial components of the visible and fastMTT-corrected system momenta $\vec{p}_{vis}$ and $\vec{p}_\tau^{\,fastMTT}$.

To construct the rotation, an orthonormal basis is defined in the decay plane. The first axis is chosen along the visible momentum direction, while the third axis is perpendicular to the plane spanned by the first and $\vec{p}_{vis}$, and the second axis completes the right-handed system:

$$\hat{n}_1 = \frac{\vec{p}_{vis}}{|\vec{p}_{vis}|}, \qquad \hat{n}_3 = \frac{\hat{n}_1 \times \vec{p}_{vis}}{|\hat{n}_1 \times \vec{p}_{vis}|}, \qquad \hat{n}_2 = \hat{n}_3 \times \hat{n}_1. \tag{4.24}$$

The $\vec{p}_\tau^{\,fastMTT}$ is decomposed in this basis where P_1 and P_2 form the in-plane component and $P_\perp$ is the out-of-plane component:

$$\vec{p}_\tau^{\,fastMTT} = \underbrace{P_1\,\hat{n}_1 + P_2\,\hat{n}_2}_{\vec{p}_\tau^{\,in-plane}} + \underbrace{P_\perp\,\hat{n}_3}_{\vec{p}_\tau^{\,out-of-plane}} \tag{4.25}$$

If a rotation is required (Figure 4.9b), the out-of-plane component is left unchanged, while the in-plane component is rotated such that its opening angle with respect to $\hat{n}_1$ becomes exactly the maximal allowed angle, $\theta_{\max}$. This rotation yields two possible in-plane directions:

$$\hat{n}_{\parallel}^{\pm} = \hat{n}_1 \cos\theta_{GJ}^{max} \mp \hat{n}_2 \sin\theta_{GJ}^{max}. \tag{4.26}$$

The two corresponding rotated tau momenta candidates are then reconstructed by combining the unchanged out-of-plane component with each rotated in-plane component:

$$\vec{p}_{\tau}^{\,candidate,(\pm)} = \vec{p}_{\tau}^{\,out-of-plane} + |\vec{p}_{\tau}^{\,in-plane}|\,\hat{n}_{\parallel}^{\pm}. \tag{4.27}$$

Among the two candidate solutions, the one closest to the original tau flight direction is selected, ensuring that the applied correction is minimal.

Finally, the spatial momentum magnitude is kept fixed, and the energy is recomputed from the on-shell condition

$$E_{\tau} = \sqrt{|\vec{p}_{\tau}|^2 + m_{\tau}^2}, \tag{4.28}$$

such that only the direction of the tau momentum is modified, while its mass and energy remain unchanged.

4.6 Estimation of the Sensitivity to the CP observable

The sensitivity of the analysis to the CP nature of the tau Yukawa coupling is assessed by studying the distributions of the CP-sensitive observable ϕ_{CP} at different stages of the reconstruction chain: at generator-level, at visible-level, after the fastMTT correction of the visible-level, and after the angular correction applied to the fastMTT-level.
For the muonically decaying tau leptons, the IP method is used to define the decay plane. For the hadronically decaying tau leptons, the $\vec{h}_{PV}$ method is used by default in combination with the stepwise corrections of the visible-level provided by fastMTT and GEF. To allow a direct comparison with the DP method used in the previous run, the hadronic tau decay is also reconstructed using the DP method.
The overall strategy consists of three main steps. First, generator-level distributions are constructed in order to establish an ideal reference corresponding to perfect measurements without detector or reconstruction effects. Second, the reconstructed ϕ_{CP} distributions are modelled using an analytic parametrisation. Finally, an asymmetry observable is defined to quantify the separation power between different ϕ_{CP} hypotheses and thereby estimate the sensitivity of the reconstruction.

4.6.1 Generator-level Observables

To establish a reference corresponding to ideal detector performance, generator-level ϕ_{CP} distributions are constructed. These distributions represent how the observable would be measured in the absence of detector effects, finite resolution, or reconstruction inefficiencies. Firstly, they define the intrinsic, detector-independent shape of the ϕ_{CP} distribution and thus the maximal CP achievable in an ideal scenario. Secondly, they provide a reference against which the impact of different reconstruction steps and detector effects can be quantified by direct comparison.

For each reconstructed visible particle in the $\tau_\ell\tau_{had}$ final state, the corresponding generator-level object is identified using Monte Carlo truth information. The leptonic and hadronic decay legs are treated separately.
For the muonic decay leg, the reconstructed muon is matched to its generator-level counterpart using the generator particle index `genPartIdx`. For signal events, the muon originates from a tau lepton produced in a Higgs boson decay, $H \to \tau_\ell\tau_{had},\ \tau_\ell \to \mu + \nu$. Due to possible photon or gluon radiation in the decay chain, the muon's direct mother particle is not necessarily the Higgs boson itself, it may be separated from it by intermediate radiative tau decays. Therefore, starting from this generator-level muon, the particle ancestry is followed recursively until either a Higgs boson is identified or, for background events, none is found in the ancestry. If a Higgs boson is found, the corresponding tau lepton originating from the Higgs decay at the generator-level is identified and used as the reference particle for the production kinematics.
This procedure ensures consistent separation of the tau lepton at the production-level and following radiative effects, while retaining full access to the decay kinematics.

For the hadronic decay leg, the reconstructed τ_{had} candidate is first matched to the corresponding generator-level visible tau decay product, `GenVisTau`. As this object only contains the visible decay products, the associated generator-level tau lepton after final-state radiation is identified using the generator particle index. Starting from this post-radiation tau lepton, the full decay tree is reconstructed by accessing its daughter particles `GenTauProds` via `genPartDaughter`. Additionally, the ancestry of this tau lepton after radiation is traced recursively until the Higgs boson is found, thereby identifying the tau lepton produced in the Higgs decay at the generator-level.

The generator-level IP vector is computed for the IP method and IP calibration with Dell-Yan sets.
The direction of the tau lepton is determined by the displacement between the generator-level PV and the tau decay vertex SV,

$$\vec{p}_\tau = \overrightarrow{SV} - \overrightarrow{PV} \qquad \text{, with its unit vector} \qquad \hat{n}_\tau = \frac{\vec{p}_\tau}{|\vec{p}_\tau|}. \tag{4.29}$$

The unit vector along the direction of the visible decay product, constructed from the three-momentum of the visible tau decay product $\vec{p}_{vis}$ and is given by $\hat{n}_{vis} = \frac{\vec{p}_{vis}}{|\vec{p}_{vis}|}$.
The generator-level IP vector is then obtained by projecting the tau momentum onto the plane perpendicular to the visible decay direction

$$\overrightarrow{IP}_{\text{gen}} = \vec{p}_\tau - (\vec{p}_\tau \cdot \hat{n}_{\text{vis}})\ \hat{n}_\tau. \tag{4.30}$$

At the generator-level, particle masses are set to zero in the input samples. This leads to incorrect Lorentz transformations when boosting into the relevant RFs. To obtain physically meaningful boosts for the $\vec{h}_{PV}$ calculation and the following reconstruction of the coplanarity angle ϕ_{CP}, the masses of the τ, μ, and π are manually set to their nominal values.

4.6.2 Model for the ϕ_{CP} Distributions

After event selection and generator-level matching, the ϕ_{CP} distributions are constructed for each event category and signal production mode. For each CP hypothesis (CP-even,

CP-odd, and maximally mixed (CP-maxmix)), the distribution is modelled using a cosine function of the form

$$f(x) = a\cos(x + c) + b\,, \tag{4.31}$$

where a represents the modulation amplitude, b an overall offset, and c a phase shift to match a sinusoidal distribution.
The fit is performed using a least-squares minimisation implemented with the IMINUIT package [80]. The statistical uncertainties of the histogram bins are propagated to the fit and originate from the histogramming step of the COLUMNFLOW-based analysis pipeline, as described in section 4.7 and shown in Appendix B. These uncertainties are used as input to the least-squares minimisation for the fit. IMINUIT also provides a robust and efficient numerical minimisation of the χ^2 function, from which the goodness of fit is quantified using the χ^2/ndof value for each hypothesis.

4.6.3 Spin Analysing Power

In order to quantify how well two CP hypotheses can be distinguished from one another, an asymmetry observable A is defined based on their fitted ϕ_{CP} distributions. For two fitted curves represented by arrays f_1 and f_2, the asymmetry is computed as

$$A = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{f_1(i) - f_2(i)}{f_1(i) + f_2(i)}\right|\,, \tag{4.32}$$

where N is the number of bins.

The uncertainty on the asymmetry is obtained by standard Gaussian error propagation from the uncertainties of the fitted curves. For a single bin i, the asymmetry contribution is denoted by a_i. Assuming the uncertainties on $f_1(i)$ and $f_2(i)$ to be uncorrelated, the corresponding uncertainty σ_{a_i} is given by

$$\sigma_{a_i} = \sqrt{\left(\frac{2f_2(i)}{(f_1(i) + f_2(i))^2}\,\sigma_{f_1(i)}\right)^2 + \left(\frac{2f_1(i)}{(f_1(i) + f_2(i))^2}\,\sigma_{f_2(i)}\right)^2}\,. \tag{4.33}$$

Since the total asymmetry is defined as the normalised sum of a_i over all bins, the corresponding uncertainty is

$$\sigma_A = \frac{1}{N}\sqrt{\sum_{i=1}^{N}\sigma_{a_i}^2}\,. \tag{4.34}$$

A large value of the asymmetry indicates a strong separation power between the two hypotheses under consideration, most commonly between the CP-even and CP-odd cases. This manifests itself in a good spatial separation of their curves in the ϕ_{CP} plots shown in section 5.5. Consequently, this observable provides a direct and quantitative measure of the sensitivity of the analysis to the CP structure of the Higgs to tau coupling.

4.7 The ColumnFlow Analysis Tool

A key feature of this analysis is its implementation within the columnar framework COLUMNFLOW [81]. Unlike traditional event-wise looping analyses, columnar approaches

operate on observables as arrays, or columns, allowing simultaneous processing of all events. This design provides significant computational efficiency and simplifies handling of large datasets, which is crucial for high-statistics Higgs studies.

Data in COLUMNFLOW are organised per-variable rather than per-event. Operations are vectorised over these columns, often using NUMPY and AWKWARD ARRAYS, enabling GPU acceleration and greatly reducing processing time. Columnar storage also allows selective reading of only the required variables (e.g., from `Parquet` files), optimising memory usage. Once the columnar structure is in place, physics logic can be implemented straightforwardly, by referencing object variables directly, e.g.

```
tau.pt = np.sqrt(tau.px**2 + tau.py**2)
```

This results in code that is both clean and highly readable, while keeping the focus on implementing physics logic. In addition, column-wise operations are naturally batchable, which simplifies parallelisation on multi-core, distributed, or GPU systems. Furthermore, given the rapid development of Machine Learning (ML) algorithms, which operate natively on arrays, columnar frameworks are particularly well suited for modern analysis workflows. However, it should be noted that adopting a columnar framework often requires a period of adjustment, especially when transitioning from an event-wise looping approach, as it relies on a different setup basis, using array broadcasting and nested data structures. The nested organisation of arrays can reduce flexibility in certain complex operations. Finally, for very large datasets, careful memory management is essential, since loading long arrays into memory may exceed the available RAM if appropriate chunking strategies are not employed.

The COLUMNFLOW tool (see Appendix B) organises analysis tasks through a structured task graph, providing a clear workflow from raw data to final histograms. The main steps include:

1. Loading datasets: Reading input files.
2. Calibration of events: Applying object- and event-level corrections.
3. Selections: Defining masks to filter events based on physics criteria.
4. Reductions: Merging columns as needed to simplify subsequent operations.
5. Histogram creation: Producing histograms per dataset and era.
6. Merging histograms: Combining histograms across multiple variables per dataset and data-taking era.
7. Plotting: Generating final plots for physics interpretation.

Chapter 5

CP Measurement in $H \to \tau\tau \to \mu a_1^{3pr}$ Decays

5.1 Distribution of Observables 65
5.2 Event Selection 68
5.3 Signal and Background Estimation BDT 69
5.4 Kinematic Reconstruction 70
5.4.1 fastMTT Method 70
5.4.2 Global Event Fit 73
5.4.3 Secondary Vertex Corrections 73
5.5 Measurement of the ϕ_{CP} angle 74
5.5.1 Reconstruction Stages 74
5.5.2 Statistical Interpretation 77

This chapter presents a study of the CP measurement described in chapter 4, using the μa_1^{3pr} channel and the 2022 and 2023 data. It begins with a general overview of the event and objects' kinematics in both simulation and data. The performance of the BDT used to classify data events into signal and background categories is then discussed, as this classification forms the basis for the extraction of the CP observables. Before presenting the final results, the performance of different tau lepton reconstruction methods is evaluated for fastMTT, GEF, and the estimated tau lepton flight direction. Finally, the CP measurement is performed and the results are evaluated within a statistical interpretation framework.

5.1 Distribution of Observables

Following the event selection described in section 4.2, the distributions of the main kinematic and SV-related observables are presented for the μa_1^{3pr} final state. The analysed data correspond to an integrated luminosity of $62.4\,\text{fb}^{-1}$, with $Z \to \tau\tau$ events dominating the background (■). The six signal samples, corresponding to $\{\text{ggF}, \text{VBF}\} \times \{\text{CP even}, \text{CP odd}, \text{CP maxmix}\}$, are shown as unstacked curves, and the data as black

points. Colours follow the convention in Table 4.2. Overall, the agreement between data and MC simulation is good, especially given the specificity of the μa_1^{3pr} channel and the resulting limited statistics. Event yields are reported in Table C.1.

Figure 5.1 and Figure 5.2 show, for the MET-corrected $(\tau_\ell \to)\mu$ and $(\tau_h \to)a_1^{3pr}$ resonance, respectively, the four-vector components p_T, η, ϕ, and m in the first row, and the IP significance and spatial coordinates in the second row. These distributions are obtained using the best-performing reconstruction models, namely fastMTT with a mass-window constraint and GEF, as discussed in section 5.4. They serve as a validation of both the event reconstruction and the modelling of the MC used in the subsequent analysis. For comparison, distributions obtained with fastMTT without a mass constraint, where the di-tau invariant mass is included in the BDT training, are provided in Figure D.1.

The p_T distributions have a minimum p_T due to trigger thresholds, a peak reflecting the typical momentum of the mother particle distributed to both tau leptons, and a high-p_T tail caused by off-shell or highly energetic production. For the fastMTT-reconstructed distributions, the muon and tau p_T peak around 60 GeV. For background without fastMTT (see Figure D.1), muon and tau p_T peak at 50 GeV, whereas signal tau peaks around 55 GeV, consistent with the harder Higgs kinematics.

The η distributions peak around zero, corresponding to the centre of the detector. The ϕ distributions are approximately uniform, showcasing the cylindrical symmetry of the detector. The mass distributions show the reconstructed lepton masses without applying fastMTT, allowing them to reflect the actual reconstructed kinematics rather than being fixed to the SM tau mass as in fastMTT. In particular visible for the hadronic tau mass distribution, it peaks below the nominal SM value, as expected due to the neutrinos escaping detection, whereas the fastMTT-corrected masses would instead exhibit clear peaks around the SM tau mass.

The IP significance distribution exhibits a two-peak structure, also observed in other groups' analyses, where the precise origin is unknown. The IP_{sig} cut at 1.25 for muons improves the sample purity by removing poorly reconstructed background while retaining most signal events. The spatial components $IP_{x,y,z}$ peak symmetrically around small offsets, reflecting the displacement of the SV in cm from the detector centre. Even after applying the IP corrections, residual mismodelling of approximately 10% remains in all spatial components, while IP_{sig} additionally shows a shape discrepancy. For hadronic tau leptons, IP_{sig} is shifted to higher values, as it is followed by a multi-prong decays producing a more displaced SV in the reconstruction. The corresponding cut at 1 for $\tau \to \pi^0$ decays suppresses misidentified background, where neutral pions decay promptly into photons, contributing no charged tracks to the vertex. Events passing the cut are more likely to correspond to genuinely displaced tau decays with well-measured IP vectors. A precise IP reconstruction is crucial for an accurately reconstructed orientation of the tau decay plane, which is essential for employing the IP method.

Figure 5.3 presents additional event-level observables, including the Decay Mode (DM) as reconstructed by the previously used HPS and the currently employed PNet algorithms, the PUPPI MET p_T and ϕ, the angular separation ΔR between the two leptons, and the number of jets within an acceptance of $p_T > 20$ GeV and $|\eta| < 4.7$.

The DM distributions highlight differences between the previous HPS and current PNet algorithms. Events classified as a_1^{3pr} by PNet are sometimes assigned to $\tau \to 4\pi$ by HPS,

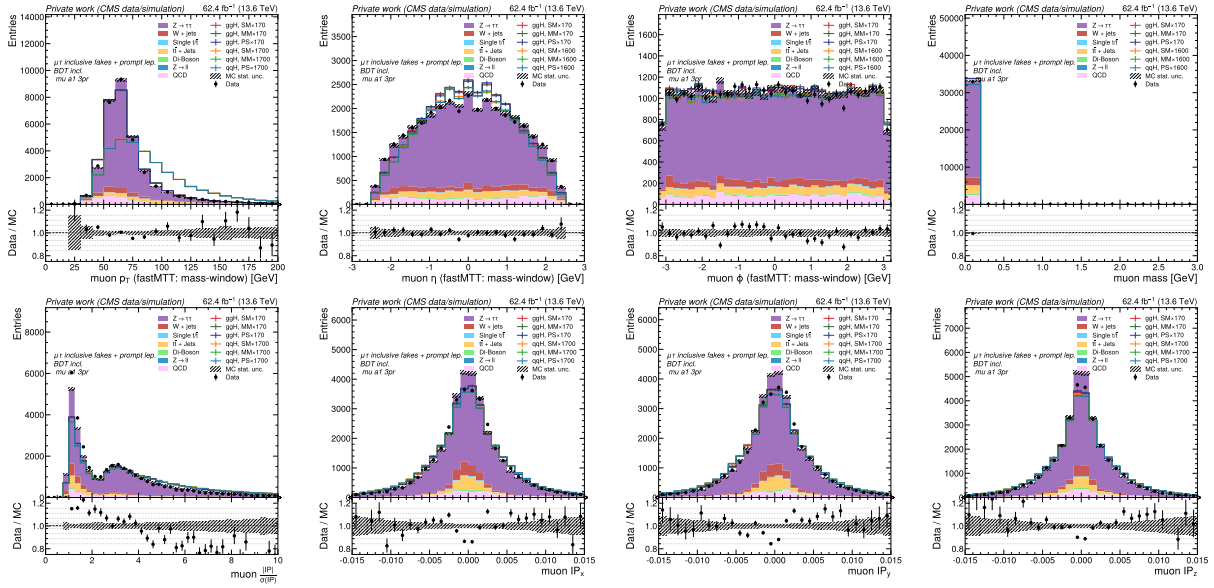


Figure 5.1: Control distributions of the MET corrected $(\tau \rightarrow)\mu$ observables p_T, η, ϕ, and m (top row), as well as the impact parameter significance and spatial components IP_x, IP_y, and IP_z (bottom row), shown for the full 2022 and 2023 datasets. The distributions are obtained within the DesyTau framework using the fastMTT reconstruction with a mass-window constraint and GEF.

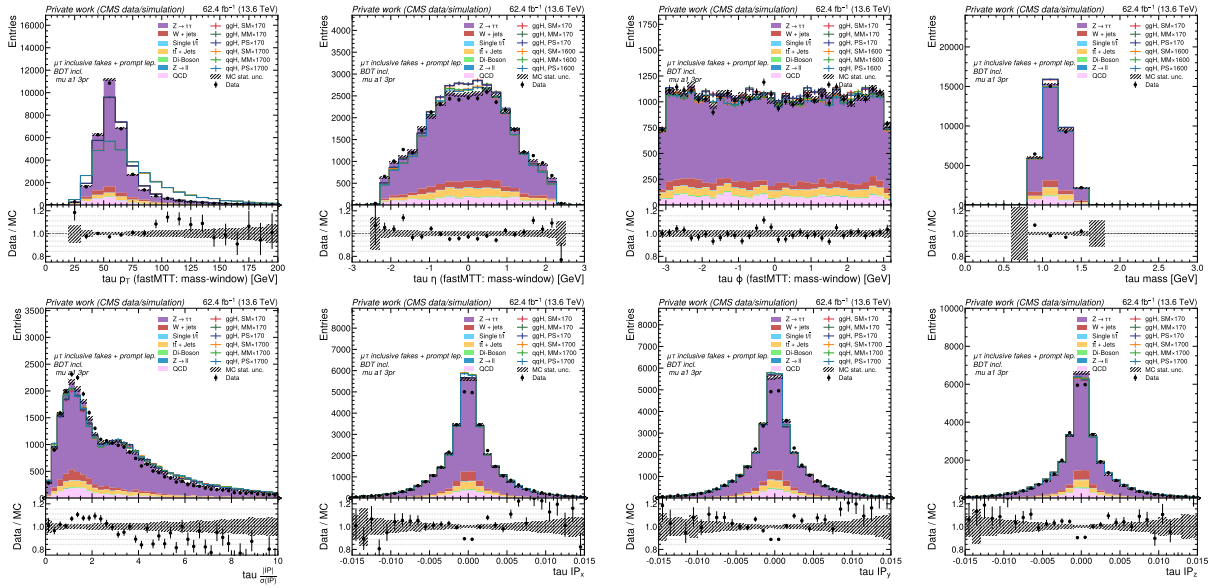


Figure 5.2: Control distributions of the MET corrected hadronic tau a_1^{3pr} observables p_T, η, ϕ, and m (top row), as well as the impact parameter significance and spatial components IP_x, IP_y, and IP_z (bottom row), shown for the full 2022 and 2023 datasets. The distributions are obtained within the DesyTau framework using the fastMTT reconstruction with a mass-window constraint and GEF.

reflecting the improved resolution of multi-prong decays by PNet. The PUPPI MET p_T peaks around 20 GeV, with a flat ϕ distribution, consistent with detector geometry and event topology. Jet multiplicities reveal that most ggF and background events contain zero jets, whereas VBF events show higher multiplicities due to the presence of the two forward

quark jets from the hard scattering. This behaviour correlates with the ΔR distributions, which peak at larger values for background and ggF, while VBF signals peak at low ΔR. The smaller ΔR in VBF arises because the two forward jets carry away part of the momentum from the hard scattering, boosting the Higgs boson in the opposite direction. As a result, the tau leptons stemming from the Higgs decay, are more collimated, leading to smaller angular separation between them.

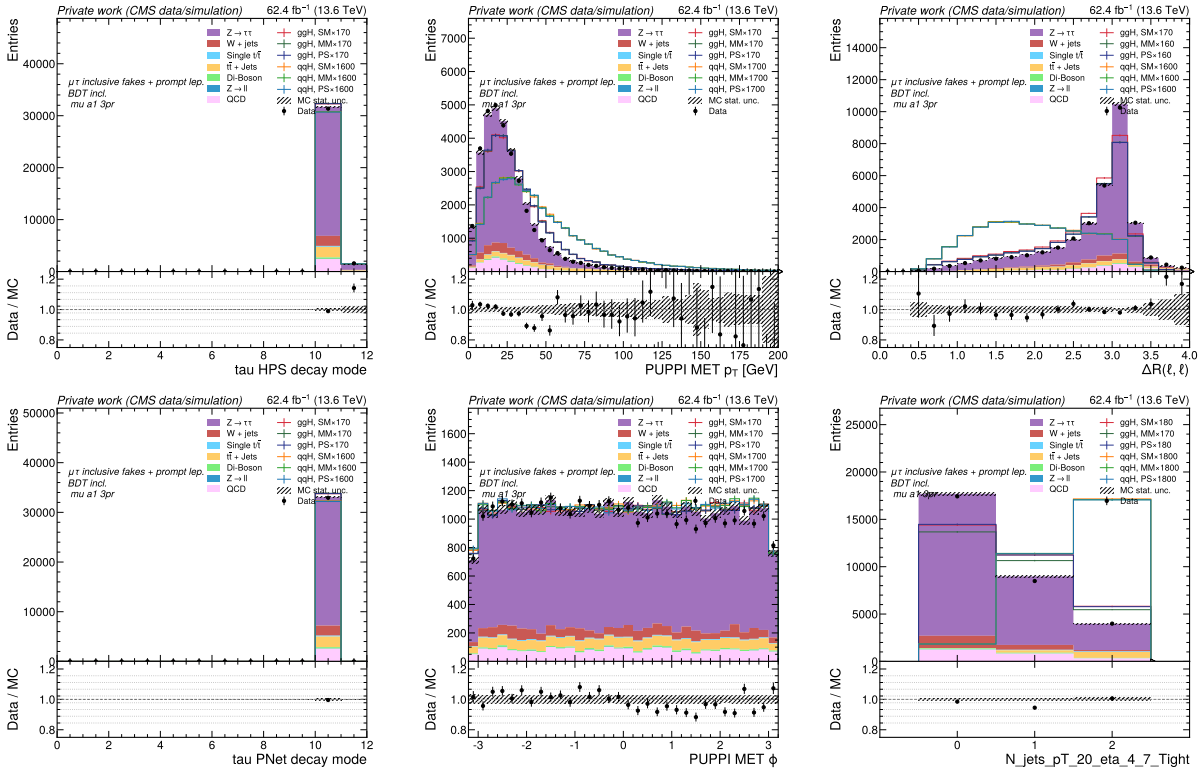


Figure 5.3: Control distributions of the reconstructed hadronic tau DM, comparing the previously used HPS and the currently employed PNet algorithms (left); the PUPPI MET p_T and ϕ (middle); as well as the angular separation ΔR between the two leptons and the number of reconstructed jets with $p_T > 20\,\text{GeV}$ and $|\eta| < 4.7$ (right). The distributions are obtained within the DesyTau framework using the fastMTT reconstruction with a mass-window constraint and GEF, and are shown for the full 2022 and 2023 datasets.

5.2 Event Selection

Figure 5.4 summarises the effect of the event selection on a representative batch of 100 000 events from the ggF SM sample. Each selection retains a fraction of events according to its criteria, gradually refining the dataset.
Unlike traditional loop-based analyses, the steps in the columnar framework are applied simultaneously. Consequently, the relative fractions reflect the individual impact of each cut rather than a cumulative effect. This representation emphasises the selections with the largest influence on the event yield, providing a concise overview of the dataset's refinement. Similar patterns are observed for all signal samples.

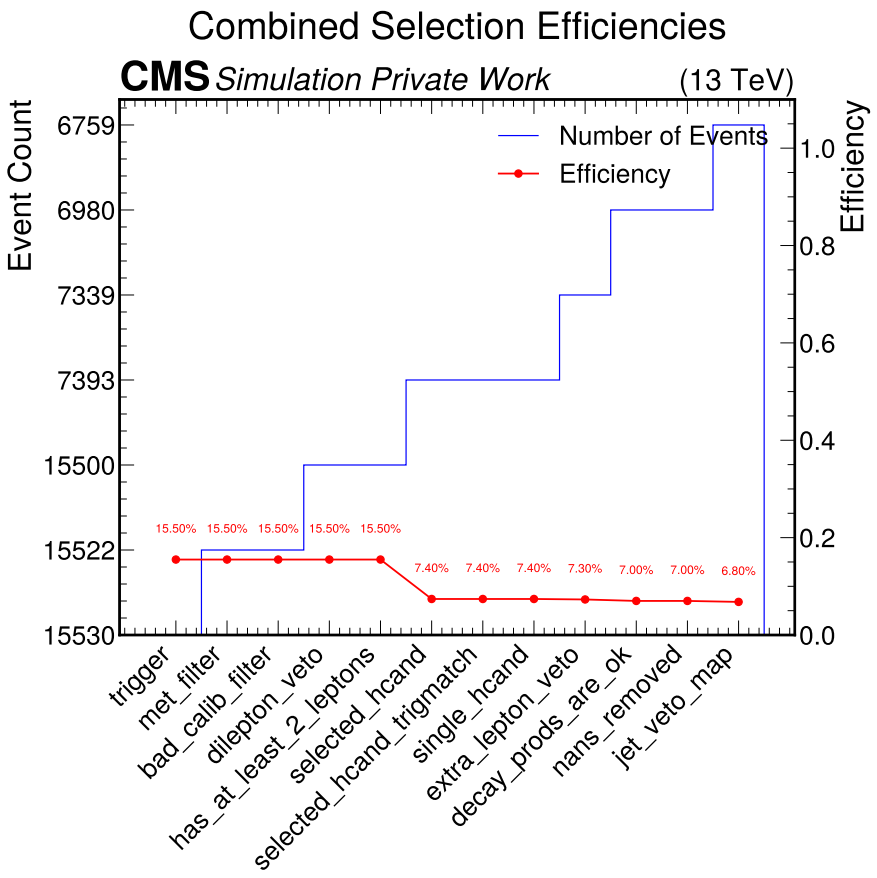


Figure 5.4: Combined selection efficiencies for the SM $ggF \to H \to \tau\tau$ signal set and a chunk size of 100 000.

5.3 Signal and Background Estimation BDT

The trained BDT classifies events into three signal categories (green, orange, and violet in ascending order) and two background categories, corresponding to genuine μa_1^{3pr} final states from tau lepton decays and to signatures arising from misidentified jets. The classifier output, see Figure 5.5, is a continuous variable between 0 and 1, with values closer to 1 indicating a higher probability for an event to belong to the respective category.

The BDT output effectively separates signal-like and background-like events. In the signal categories, background events peak towards 0, while the six signal hypotheses, {ggF, VBF} × {CP-even, CP-odd, CP-maxmix}, cluster near 1. This separation is more pronounced for VBF, owing to the additional quark jets that improve event classification. For the background category with genuine tau leptons, $Z \to \tau\tau$ events peak to the right, as expected, while Higgs signals shift to the left. Misidentified-jet backgrounds peak towards 0 in both signal and background categories, with the effect stronger in VBF due to the presence of forward jets.
Overall, the distributions confirm that the classifier reliably distinguishes signal from background, with the peak positions and shapes reflecting the underlying event topology and production mechanism. This clear separation directly feeds into the statistical model, enhancing the sensitivity to the CP-mixing angle $\alpha^{H\tau\tau}$, as shown in the corresponding

scans in subsection 5.5.2.

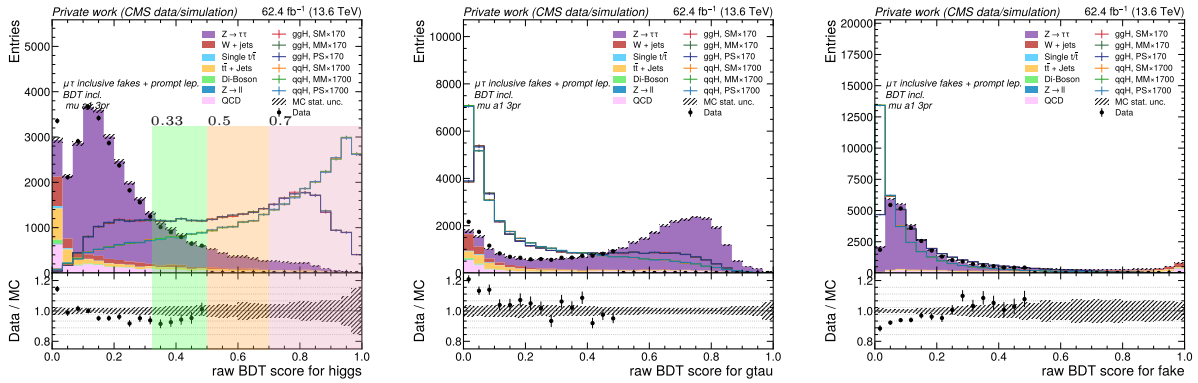


Figure 5.5: Distributions of the BDT output for the three signal categories (green, orange, and violet) and the two background categories, shown for the full 2022 and 2023 datasets. The signal categories correspond to genuine $(H \to)\mu a_1^{3pr}$ events, while the background categories either originate from genuine tau leptons not coming from a Higgs boson or from misidentified jets. The classifier output is a continuous variable between 0 and 1, where values closer to 1 indicate a higher probability for the event to belong to the given category.

5.4 Kinematic Reconstruction

The tau lepton momenta are reconstructed and progressively corrected to account for the missing MET arising from undetected neutrinos. First, the momentum magnitudes are corrected using fastMTT, followed by the angular components via GEF. Finally, the reconstructed tau lepton directions are presented, which were observed to be misaligned in previous Run 2 studies.

5.4.1 fastMTT Method

The fastMTT method (subsection 4.5.1) was tested without a constraint, using a mass-window constraint of $[122.6, 127.6]\,\text{GeV}$, as well as a Breit–Wigner approach with two different Higgs widths Γ_H, namely $0.01\,m_H$, following the best results from the cSVfit reconstruction group, and 4 MeV, corresponding to the simulated Higgs width in the MC samples. The resulting invariant mass distributions, compared to that of the visible system, are shown in Figure 5.6.

As expected, the visible invariant mass m_{vis} distribution peaks below the Higgs mass, compared to the corresponding generator-level distribution (Figure 4.4), due to the presence of undetected neutrinos in the tau decays. With the unconstrained fastMTT reconstruction, both signal and background distributions shift to higher masses and become more symmetric. The signal moves closer to the Higgs mass because the algorithm can find neutrino momenta consistent with the observed MET, effectively reconstructing the full di-tau system. For background events, which do not originate from a Higgs decay, the MET may include contributions from mismeasurements or other neutrinos, and the visible objects do not satisfy the Higgs-like kinematics. As a result, the reconstructed

mass shifts less. The distribution becomes more symmetric because the MET partially pulls low-mass events upward, without fully reproducing a Higgs-like configuration. For the fastMTT reconstruction with a Breit-Wigner width of 1.25 MeV, the distributions are strongly pulled towards the Higgs mass of 125 GeV, producing a narrow peak, as expected from the applied constraint. The mass-window approach also shifts the distribution towards the Higgs mass, but it becomes unphysical, as it peaks at the boundaries.

Nevertheless, as discussed below, the latter configuration provides the best results for certain applications. When the invariant mass is used directly in further analyses, as for the BDT training, only the unconstrained observable ensures physically meaningful distributions. However, if the invariant mass is employed to optimise other quantities, such as the tau momenta, the best-performing reconstruction method, namely the mass-window approach, is chosen.

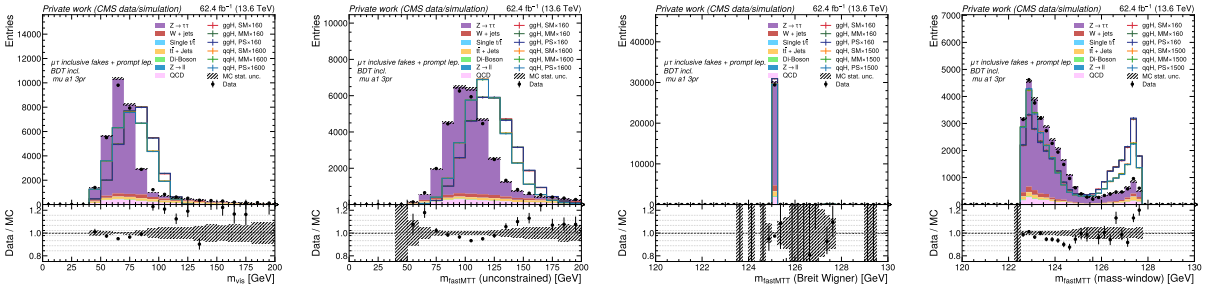


Figure 5.6: Di-tau invariant mass distributions for the full 2022 and 2023 dataset: the visible mass (first column), unconstrained fastMTT (second column), the Breit–Wigner reconstruction with a width of $0.01\,m_H$ (third column), and the mass-window $[122.6, 127.6]$ GeV approach (fourth column).

The resolution plots are shown in Figure 5.7 for both muonic and hadronic tau candidates, comparing the three fastMTT constraint configurations: Breit–Wigner with $\Gamma = 4\,\text{MeV}$, as simulated in the samples, Breit–Wigner with $\Gamma = 0.01\,m_H$, which corresponds to the best-performing cSVfit method, and the mass-window approach. By construction, a perfect reconstruction relative to the generator-level information would yield a resolution distribution sharply peaked at zero.

At the visible-level, all reconstructed objects (blue) exhibit broader distributions, shifted to the right. This indicates a systematic underestimation of the true momentum, with the effect being more pronounced for the muon due to the presence of two undetected neutrinos. Applying the fastMTT reconstruction (red curves) shifts the resolution distributions closer to zero and reduces their widths, demonstrating that the momentum corrections effectively recover the true kinematics. The mean and width values, indicated in the statistical boxes, quantitatively confirm this improvement. In particular, the mass-window approach yields the narrowest width and a mean closest to zero, indicating that this configuration provides the best performance for the fastMTT reconstruction. Consequently, the mass-window configuration is adopted as the optimal fastMTT setup for further studies of ϕ_{CP}.

An additional optimisation study was performed, adjusting the parameters of the likeli-

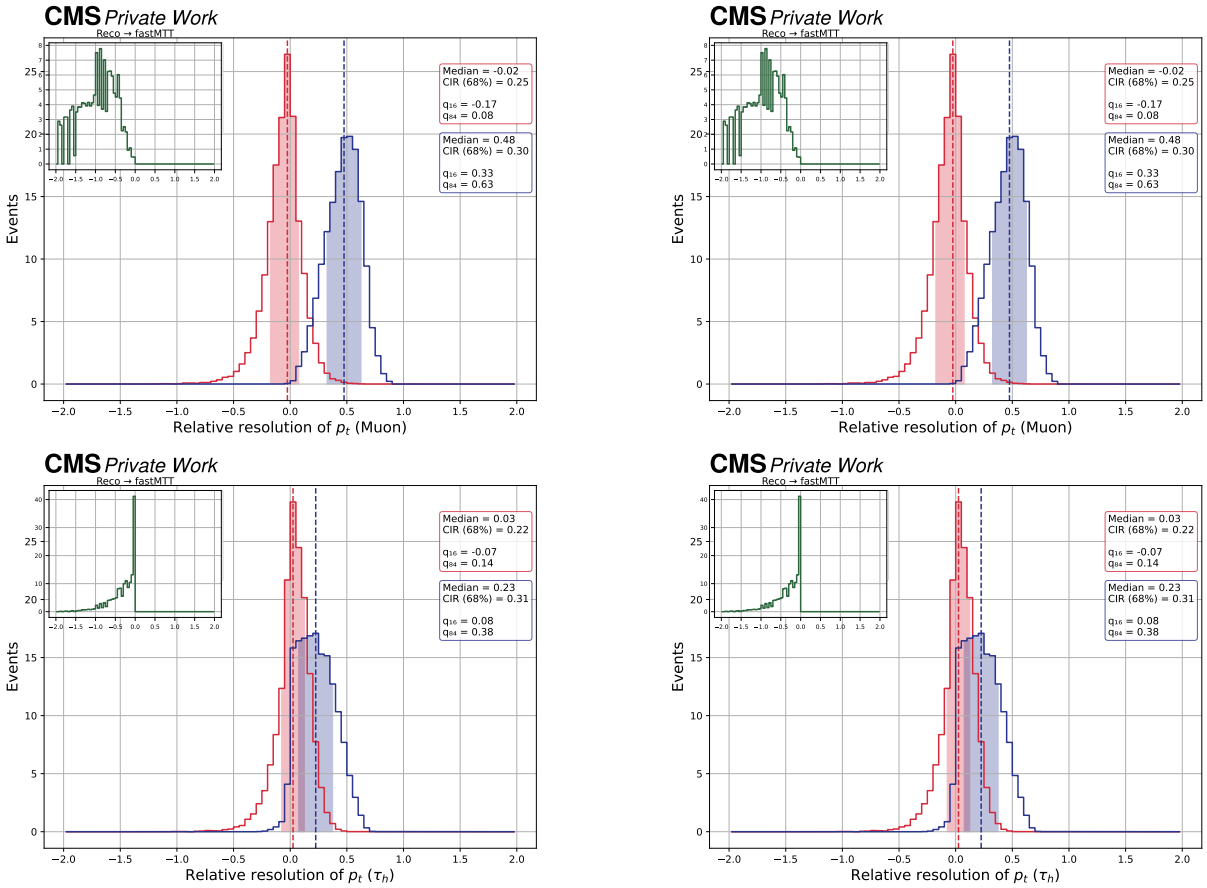


Figure 5.7: Resolution distributions for the different fastMTT constraints configurations: Breit–Wigner with 4 MeV width (first column) and Breit–Wigner with $0.01\,m_H$ width (second column), for the $(\tau \to)\mu$ (top) and $(\tau \to)a_1^{3pr}$ (bottom) planes. The resolution $\frac{p_T^{comp} - p_T^{obj}}{p_T^{comp}}$, is blue for $comp = gen$ and $obj = visible$ information and red when the obj changes to the different fastMTT models. The comparison between $comp = visible$ and $obj = fastMTT$ is shown in green.

hood scan (see subsection 4.5.1) as detailed in Table 5.1. No improvement over the results obtained using the initial mass hypothesis and regularisation parameters was found.

Table 5.1: Parameter variations studied in the fastMTT likelihood function. The empirical regularisation parameters α and β correspond to the rescaling and power-law weighting of $m_{\tau\tau}$ as described in subsection 4.5.1.

α	β	Window width [GeV]	$m_{\tau\tau}^{\text{test}}$ [GeV]
1/1.15	6.0	2.5	125.10
1/1.2	5.5	2.0	125.08
1/1.2	6.5	2.0	125.08
1/1.3	7.0	2.0	125.08
1/1.15	6.0	6.0	125.10

5.4.2 Global Event Fit

The GEF correction is applied following the fastMTT correction for μa_1^{3pr} pair candidates, with the decay planes reconstructed using the IP and $\vec{h}_{PV}$ methods, respectively. Although the rotation is necessary for only a small subset of events, it has a noticeable impact on the CP sensitivity, as discussed in section 5.5.

Figure 5.8 shows the measured GJ angle, the corresponding maximal kinematically allowed angle, and the resulting distribution after the rotation is applied wherever needed. The measured θ_{GJ} distribution is broader than the θ_{GJ}^{max} distribution, indicating that, for some events, the reconstructed angle exceeds the kinematically allowed maximum. The rotation procedure corrects these events, producing the final distribution shown in the third plot, in which all angles lie within the physical range. This demonstrates that the GEF rotation effectively enforces kinematic constraints while preserving the overall shape of the angular distribution.

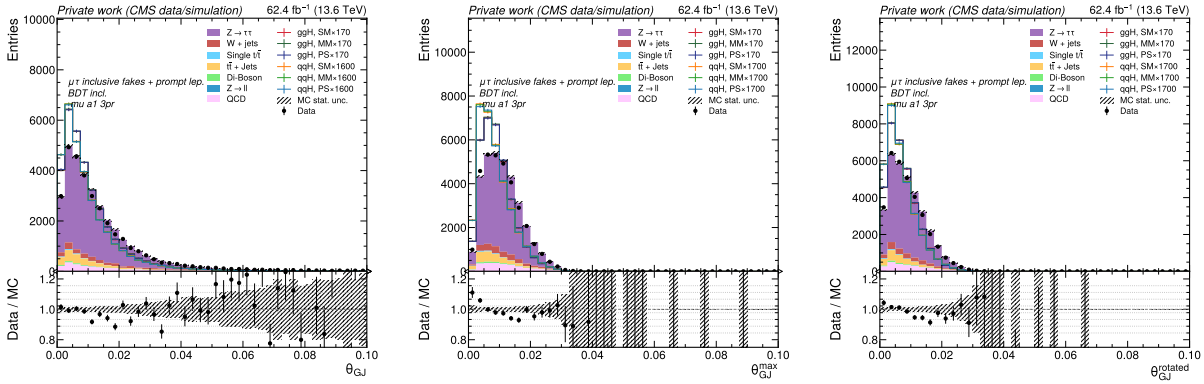


Figure 5.8: Measured and maximal GJ angle distributions for μa_1^{3pr} pair candidates, where the decay planes are reconstructed using the IP and $\vec{h}_{PV}$ methods, respectively. The third plot shows the resulting distribution of angles after applying the GEF rotation where necessary.

5.4.3 Secondary Vertex Corrections

In the corresponding Run 2 analysis, a miscalibration was observed in the estimated tau direction, defined as the vector from the PV to the SV. The distributions of the magnitude and the three spatial components of this vector are shown in Figure 5.9 for 2022 and 2023 data.

The distance between the SV and the PV peaks around zero for all spatial coordinates and falls off symmetrically on both sides, indicating that the estimated tau direction is unbiased. The widths are similar for all three coordinates. A mismodelling of about 10% in both normalisation and shape is observed between data and simulation. Compared to the Run 2 analysis, this discrepancy has been reduced by roughly a factor of two, but it is still expected to have a significant impact on the resulting spin analysing power.

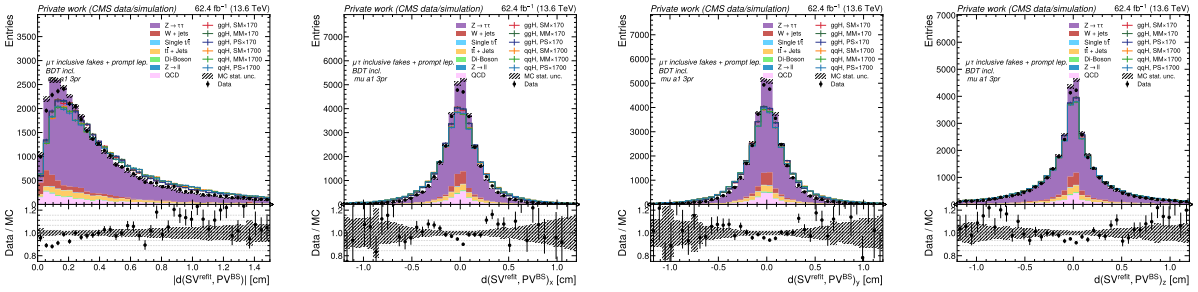


Figure 5.9: Estimated tau direction as the difference between the SV and PV positions, $d(SV, PV) = \overrightarrow{SV} - \overrightarrow{PV}$, shown for the magnitude (1st) and the x (2nd), y (3rd), and z (4th) components.

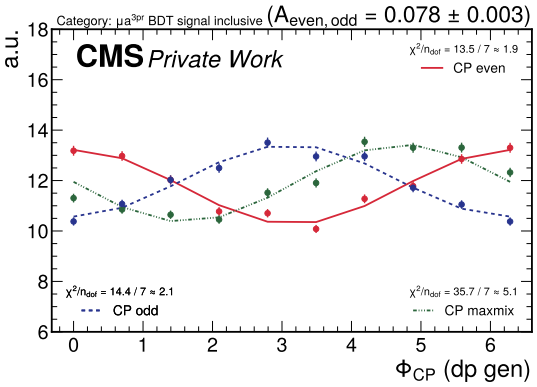
5.5 Measurement of the ϕ_{CP} angle

The coplanarity angle ϕ_{CP} is measured for μa_1^{3pr} pair candidates classified as signal-like by the BDT. In order to avoid unblinding the data, the full signal region (green + orange + violet regions in Figure 5.5) is considered. The angular distributions are computed such that the $(\tau \to)\mu + 2\nu$ decay plane is always reconstructed using the IP method, while the $(\tau \to)a_1^{3pr} + \nu$ decay plane is reconstructed either with the DP method, in order to compare the Run 2 approach within the Run 3 framework and datasets, or with the $\vec{h}_{PV}$ method, which was introduced for this channel in Run 3 and subsequently improved step by step, as discussed in the previous sections. Finally, the sensitivity to the CP-mixing angle $\alpha^{H\tau\tau}$ (see Equation 2.18) is evaluated using a statistical model.

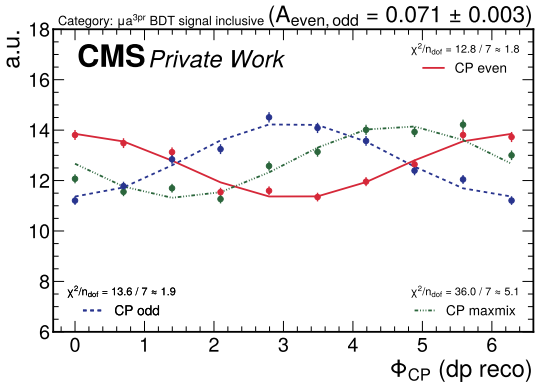
5.5.1 Reconstruction Stages

In this section, all results are shown for the ggF production mode. The corresponding distributions for VBF are provided in section D.2. The overall behaviour and conclusions are found to be very similar for both production modes.

Figure 5.10 shows the coplanarity angle distributions obtained with the DP method employed on the hadronic tau side. The left plot uses generator-level information and therefore represents the maximally achievable sensitivity for this method in this channel. The right plot shows the same distribution reconstructed at the visible-level. The obtained asymmetries of $A = 0.078(3)$ (generator-level) and $A = 0.071(3)$ (visible-level) already indicate that, for this method, the visible decay products retain most of the maximal CP sensitivity. In other words, only a relatively small fraction of sensitivity is lost when going from generator to visible-level.
However, when comparing the generator-level distributions of the DP method to those obtained with the $\vec{h}_{PV}$ method (see Figure 5.11), one observes that the intrinsic sensitivity can be increased by approximately a factor of two, reaching $A = 0.136(3)$. Since this method relies strongly on the reconstruction of the tau momenta, the CP sensitivity is, however, almost entirely lost at the visible-level, where an asymmetry of $A = 0.004(3)$ is obtained. After applying the modular correction with fastMTT, the CP sensitivity of the Run 2 approach can be recovered, yielding $A = 0.068(3)$. A further improvement of about 30% is achieved by additionally applying the GEF angular correction, resulting in an asymmetry of $A = 0.090(3)$.

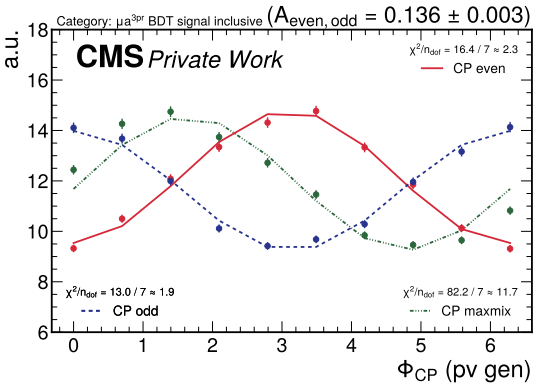


(a) Generator-level.

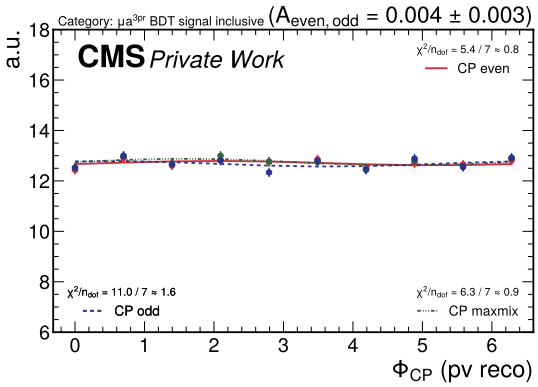


(b) Visible-level.

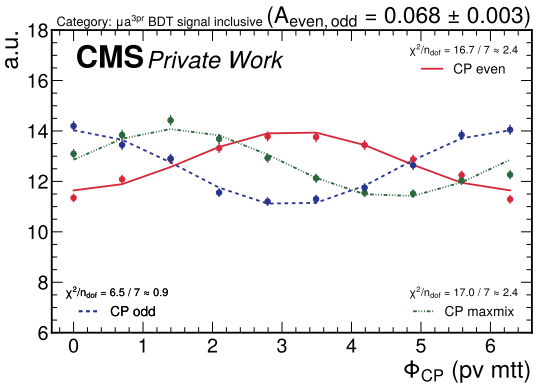
Figure 5.10: Distributions of the coplanarity angle ϕ_{CP} for ggF production and the DP method on the hadronic tau side. (a) shows the distribution at generator-level, corresponding to the maximal achievable sensitivity in this channel, while (b) shows the same observable reconstructed at visible-level.



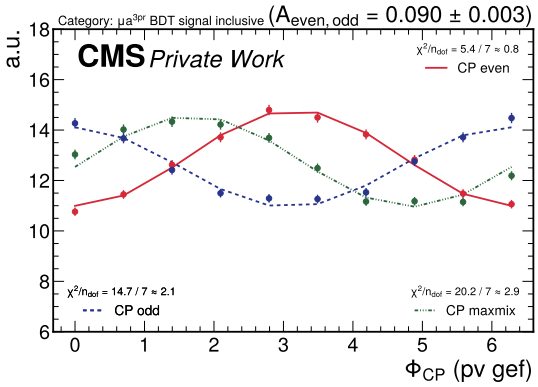
(a) Generator-level.



(b) Visible-level.



(c) fastMTT-level.



(d) GEF-level.

Figure 5.11: Distributions of the coplanarity angle ϕ_{CP} for ggF production and the $\vec{h}_{PV}$ method on the hadronic tau side, shown at generator-level (a), visible-level (b), after the fastMTT modular correction (c), and after the additional GEF angular correction (d).

For all ϕ_{CP} distributions, the fitted functions describe the shapes reasonably well. To facilitate a direct and quantitative comparison of all reconstruction stages and methods, a summary plot, Figure 5.12, is presented, showing all extracted asymmetries for DP

and $\vec{h}_{PV}$ in both production modes. The DP method is shown at generator and visible-level, while the $\vec{h}_{PV}$ method is shown additionally at fastMTT and GEF-levels. It is clear that the $\vec{h}_{PV}$ method systematically improves the spin-analysing power relative to the DP method, and that the corrections, fastMTT and GEF, progressively restore and enhance the sensitivity at visible-level. The trends are consistent for both ggF and VBF production modes, demonstrating that the improvements achieved by the $\vec{h}_{PV}$ reconstruction are robust across different Higgs production mechanisms.

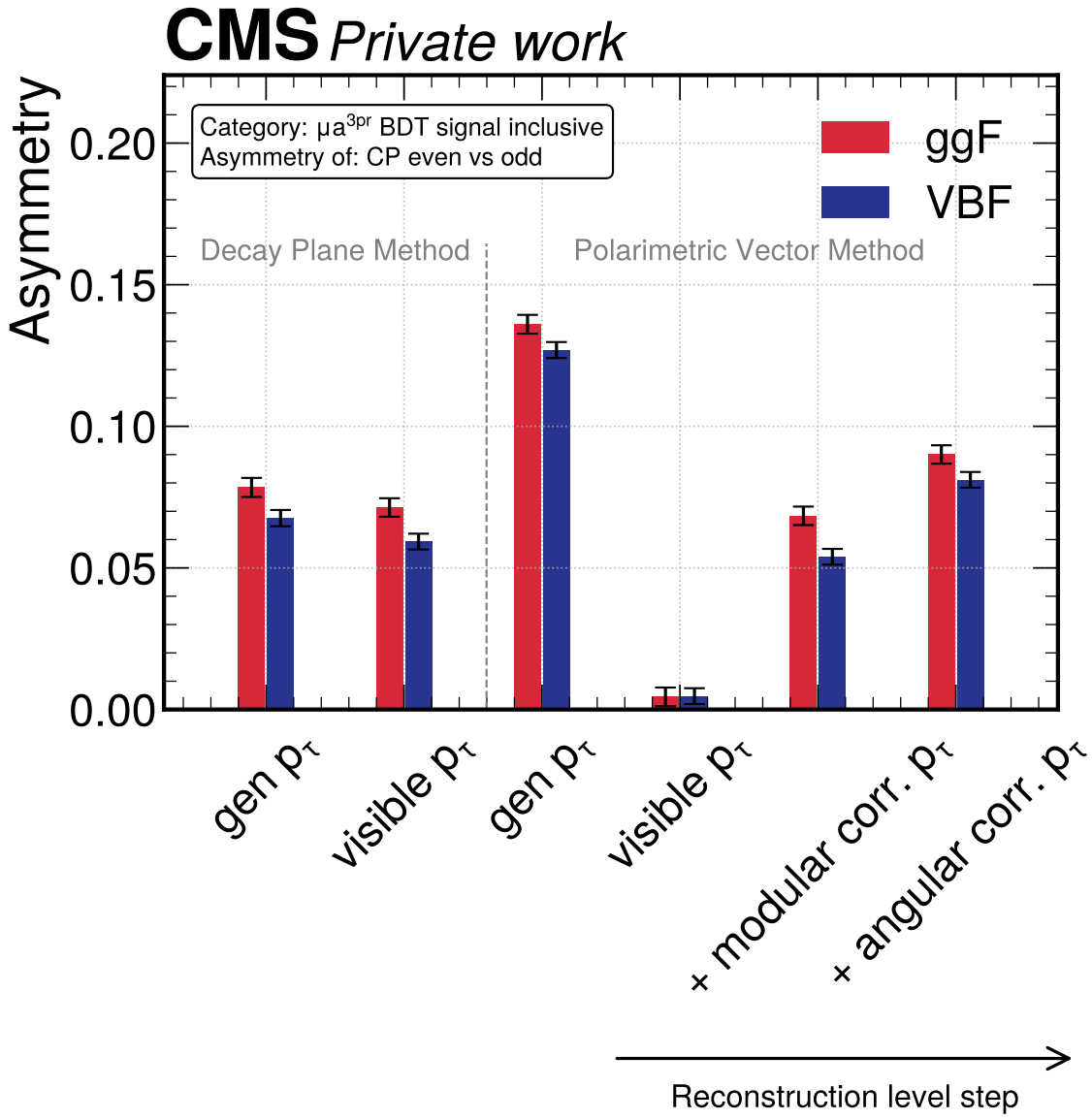


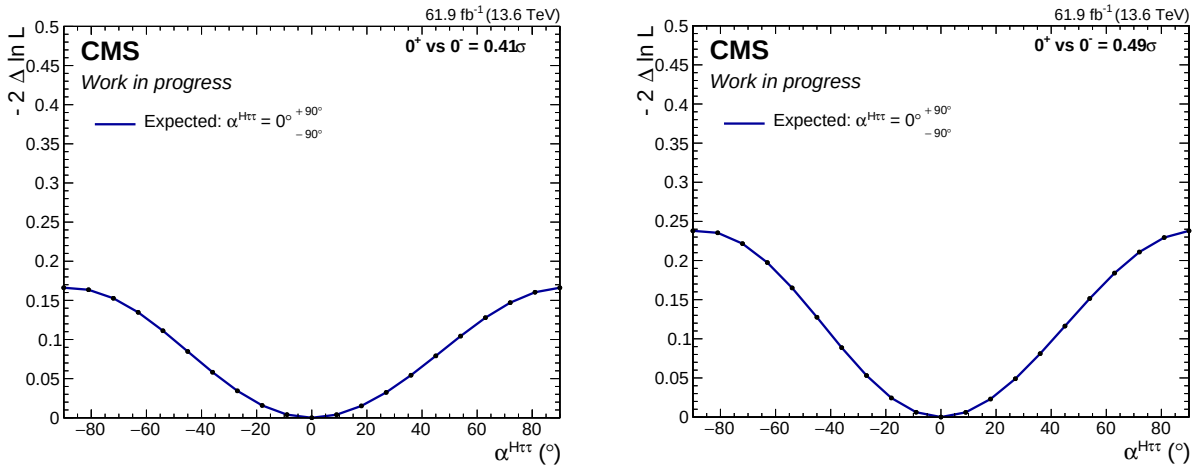
Figure 5.12: Comparison of the measured asymmetries A for the ggF (red) and VBF (blue) production modes, shown for the DP and $\vec{h}_{PV}$ methods at different reconstruction stages. The DP method is calculated at generator and visible-level only. The $\vec{h}_{PV}$ method is shown at generator and visible-level, at the fastMTT-level (visible-level + fastMTT correction), and finally at the GEF-level (visible-level + fastMTT correction + GEF correction), all are measured in the inclusive BDT category.

5.5.2 Statistical Interpretation

Finally, a statistical model is employed to perform sensitivity scans for the CP-mixing angle using the measured coplanarity angle ϕ_{CP}. The model includes statistical and luminosity uncertainties. The signal processes considered are ggF and VBF. The background contributions taken into account are $Z \to \ell\ell$, $Z \to \tau\tau$, di-, tri-boson production, and top-quark processes, QCD multijet events, and W+jets. The statistical inference is performed in the categories defined by the BDT classifier described in Figure 5.5, which is used to enhance the signal-to-background separation.

The expected sensitivity scans for the CP-mixing angle $\alpha^{H\tau\tau}$ are shown in Figure 5.13, obtained from the coplanarity angle ϕ_{CP}. The vertical axis corresponds to $-2\Delta \ln \mathcal{L}$, where $\mathcal{L}$ is the likelihood constructed from both the signal and background processes, while the horizontal axis represents the CP-mixing angle in degrees, quantifying the deviation of the Higgs boson CP structure from the SM CP-even hypothesis. The minimum of each curve indicates the best-fit value of $\alpha^{H\tau\tau}$, and the curvature around the minimum reflects the expected statistical uncertainty.

Panel (a) shows the result obtained with the Run 2 (DP) reconstruction strategy, within the Run 3 framework and with the Run 3 data, for which the mixing angle is estimated to be $\alpha^{H\tau\tau} = 0 \pm 90°$ with a significance of $0.41\,\sigma$. Panel (b) shows the result obtained with the improved $\vec{h}_{PV}$ method, which reaches a significance of $0.49\,\sigma$ in the μa_1^{3pr} channel alone, corresponding to an increase of about 20% in sensitivity. The thresholds for evidence or discovery (3σ or 5σ) correspond to values of 1 and 4 on the chosen scale $-2\Delta \ln \mathcal{L}$. Although neither method reaches these thresholds, the $\vec{h}_{PV}$ curve is deeper and slightly narrower, illustrating the improved sensitivity achieved by the reconstruction strategy developed in this thesis.



(a) Run2 (DP) method.

(b) $\vec{h}_{PV}$ method; as improved by this thesis.

Figure 5.13: Expected sensitivity scans for the CP-mixing angle $\alpha^{H\tau\tau}$ obtained from the coplanarity angle ϕ_{CP}. The left panel (a) shows the result obtained with the Run 2 (DP) reconstruction strategy, while the right panel (b) shows the result obtained with the $\vec{h}_{PV}$ method as improved by this thesis.

Chapter 6

Conclusion

The aim of this thesis was to improve the sensitivity to the CP-mixing angle $\alpha^{H\tau\tau}$ in Higgs boson decays to tau leptons by enhancing the reconstruction of the tau lepton momenta and thereby providing access to the CP-sensitive coplanarity angle ϕ_{CP}. In particular, the spin-analysing power of the transverse spin-correlation observables was studied in the $H \to \tau\tau \to \mu a_1^{3pr} + 3\nu$ decay channel. The analysis used the full Run 3 CMS dataset collected in 2022 and 2023, corresponding to an integrated luminosity of $62.4\,\mathrm{fb}^{-1}$. The momentum reconstruction was performed stepwise, starting with a kinematic correction using fastMTT, followed by an angular correction based on the GEF.

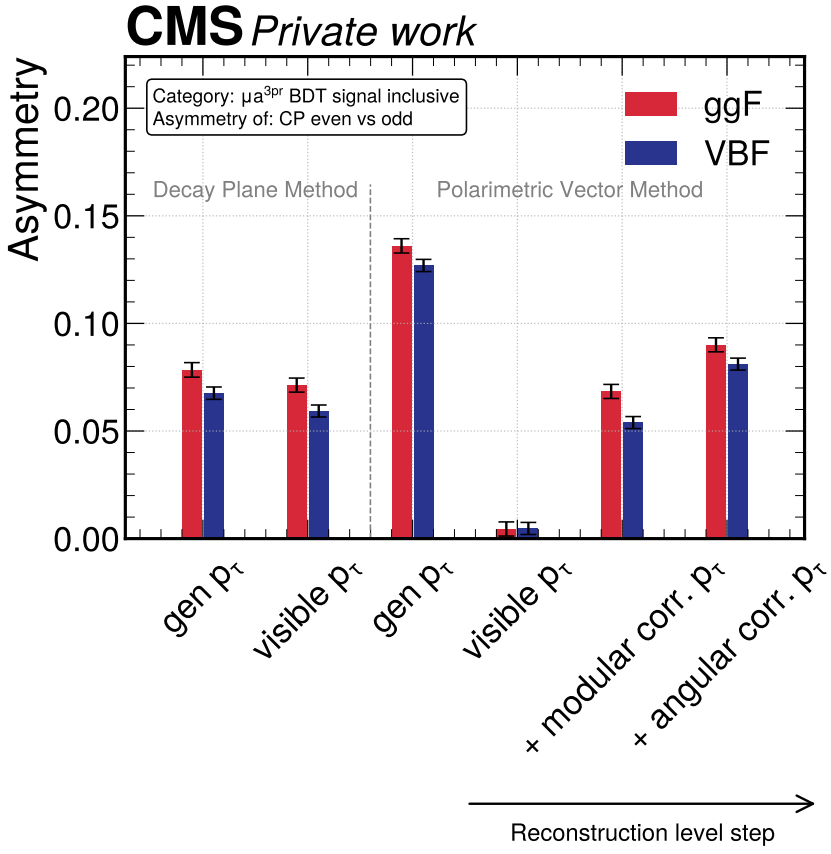


Figure 6.1: The final plot 5.12 is repeated here. Comparison of the measured asymmetries at different reconstruction stages $\{\mathrm{ggF}, \mathrm{VBF}\} \times \{\mathrm{DP}, \mathrm{PV}\} \times \{\mathrm{gen}, \mathrm{reco}, \mathrm{mtt}, \mathrm{gef}\}$.

At generator-level, it was found that the $\vec{h}_{PV}$ method achieves approximately twice the sensitivity of the DP method used in Run 2, as seen in Figure 6.1. However, missing neutrino information and detector effects were found to strongly degrade this sensitivity almost completely at detector-level. Correcting the tau momentum magnitude restored the sensitivity to a level comparable to that obtained in Run 2, while an additional angular correction yielded a further improvement of about 30% with respect to the visible DP method. Overall, the spin-analysing power of the $\vec{h}_{PV}$ method increased by factors of 11 (14) for the ggF (VBF) production mode when going from visible-level to the final GEF-corrected reconstruction.

A BDT was employed to classify signal and background events. Applying this classifier to simulated events resulted in an expected sensitivity to the CP-mixing angle $\alpha^{H\tau\tau}$ of $0.49\,\sigma$, an improvement of about 20% over the previous value of $0.41\,\sigma$. The measured sensitivity scans are consistent with a CP-even Higgs boson, in agreement with the Run 2 results, which excluded a purely CP-odd Higgs boson hypothesis at approximately 3σ. However, the statistical uncertainty in the μa_1^{3pr} channel alone remains large, and the present results alone do not yet allow a stringent constraint on the CP-mixing angle.

These results demonstrate that the $H \to \tau\tau \to \mu a_1^{3pr} + 3\nu$ channel can provide complementary sensitivity to the CP structure of the Higgs sector, although further improvements in event reconstruction and selection are still required. This study is currently limited to simulated samples, and does not yet include a treatment of shape-related uncertainties. Mismodelling of the IP components at the level of 10%, and a shape discrepancy in IP_{sig}, limit the precision of the IP decay plane reconstruction. In addition, the estimation of the primary and secondary vertices could be further improved and is expected to yield additional gains in spin-analysing power.
In order to achieve a significantly higher sensitivity, the present framework should be extended to combine the inclusive $\mu\tau$, $e\tau$, and $\tau\tau$ final states of the Higgs boson decay, and to exploit the full Run 3 dataset once fully available. This would constitute an important step towards a comprehensive measurement of the CP properties of the Higgs boson.

List of Tables

2.1 Main attributes of the $U(1)_Y$, $SU(2)_L$, and $SU(3)_C$ Gauge groups. 4
2.2 Properties and decay modes of π, ρ, and a_1 resonances in tau lepton decays. 9

4.1 Data samples per era with number of events. 41
4.2 Overview of the simulated signal and background samples used in this analysis. 44
4.3 Selection criteria for muon and hadronic tau candidates in the $\mu\tau_{had}$ channel. 46
4.4 Selection criteria applied at the pair and event level in the $\mu\tau_{had}$ channel. . 46
4.5 Vector pairs defining the decay planes of the two tau leptons. 48

5.1 Parameter variations of the fastMTT likelihood function. 72

C.1 Event yields for the DesyTau framework for 2022 and 2023 datasets. 92

List of Figures

2.1 Drawing of the Higgs potential. 6
2.2 Overview of the Standard Model. 8
2.3 Tau lepton decay modes with branching ratios. 11
2.4 Spin and momentum configurations in $\tau^- \to a_1^{-,3pr}\nu_\tau$ decays. 13
2.5 Theoretical ϕ_{CP} distributions for different CP hypotheses. 15
2.6 CMS Higgs boson mass and width measurements. 16
2.7 Dominant Higgs boson production mechanisms at the LHC. 17
2.8 Decay chain Feynman diagrams. 18
2.9 Higgs boson decay modes with branching ratios. 18

3.1 The CERN accelerator complex, showing its four main experiments and the sequential acceleration. 24
3.2 Yearly integrated luminosities collected during Run 3 at CMS. 25
3.3 Yearly pileup during Run 3. 26
3.4 Overview of the CMS detector in its Phase-1 configuration. 27
3.5 Cartesian coordinate system at CMS in scale to the LHC accelerator. . . . 28
3.6 Cylindrical coordinate system at CMS in reference to the Cartesian system and the physical environment. 28
3.7 Sketch of the Phase-1 CMS tracking system in the r-z plane. 29
3.8 Schematic of the reconstruction of tau lepton decay modes in CMS. 33
3.9 DeepTau performance versus jets, electrons, and muons. 37

4.1 Reconstruction of the decay planes in the μa_1^{3pr} channel. 49
4.2 Geometrical construction of the impact parameter. 50
4.3 Polarimetric vector construction for a_1^{3pr} decays. 52
4.4 Comparison of the expected $m_{\tau\tau}$ distribution at visible, fastMTT-reconstructed, and generator levels. 55
4.5 Gottfried Jackson angle illustration for a $\tau^- \to a_1^{3pr}\nu_\tau$ decay. 56
4.6 Naming convention of the decay chain of the a_1^{3pr} resonance. 57
4.7 Construction of the Gottfried Jackson angle in a $\tau \to had + \nu$ decay. . . . 57
4.8 Kinematic origin of the Gottfried Jackson angle limit. 58
4.9 Rotation procedure for different Gottfried Jackson angles. 59

5.1 Control plots for muon observables within the DesyTau framework. 67
5.2 Control plots for hadronic tau observables within the DesyTau framework. 67
5.3 Control plots of event-level observables within the DesyTau framework. . . 68
5.4 Combined selection efficiencies. 69

5.5 BDT output distributions for signal and background categories. 70
5.6 Di-tau invariant mass distributions. 71
5.7 Resolution distributions for different fastMTT constraints. 72
5.8 Measured, maximal, and rotated GJ angle distributions. 73
5.9 Estimated tau direction. 74
5.10 ϕ_{CP} with the DP method at generator and visible-level for ggF. 75
5.11 ϕ_{CP} with the $\vec{h}_{PV}$ method at different reconstruction stages for ggF. 75
5.12 Comparison of the asymmetries for ggF and VBF at different reconstruction stages $\{\text{ggF}, \text{VBF}\} \times \{\text{DP}, \text{PV}\} \times \{\text{gen}, \text{reco}, \text{mtt}, \text{gef}\}$. 76
5.13 Sensitivity scans for $\alpha^{H\tau\tau}$ using the DP and $\vec{h}_{PV}$ methods. 77

6.1 Summary plot of the asymmetries at different reconstruction stages $\{\text{ggF}, \text{VBF}\} \times \{\text{DP}, \text{PV}\} \times \{\text{gen}, \text{reco}, \text{mtt}, \text{gef}\}$. 79

B.1 Columnflow Taskgraph. 89

D.1 Control Plots for muon observables with an unconstrained fastMTT approach within the DesyTau framework. 93
D.2 ϕ_{CP} with the DP method at generator and visible-level for VBF. 94
D.3 ϕ_{CP} with the $\vec{h}_{PV}$ method at different reconstruction stages for VBF. 94

Appendix A

List of Abbreviations

ATLAS A Toroidal LHC Apparatus 26

BDT Boosted Decision Tree 37

BEH Brout–Englert–Higgs 4

BR branching ratio 17

BSM Beyond the Standard Model 14

C charge conjugation 19

CERN European Organisation for Nuclear Research 23

CHS Charged Hadron Subtraction 38

CKM Cabibbo–Kobayashi–Maskawa 19

CMF centre-of-mass frame 47

CMS Compact Muon Solenoid 17

CP Charge Parity 3

CPT Charge-Parity-Time 19

CSC Cathode Strip Chambers 31

DeepTau DeepTau Convolutional Neural Network 35

DESY Deutsches Elektronen-Synchrotron 46

DM Decay Mode 35

DP Decay Plane 47

DT Drift Tubes 31

ECAL electromagnetic calorimeter 29

GEF Global Event Fit 40

ggF gluon-gluon fusion 16

GJ angle Gottfried Jackson angle θ_{GJ} 53

HCAL hadronic calorimeter 29

HLT High-Level Trigger 31

HPS Hadron-plus-Strip Algorithm 35

IC Imperial College London 46

IP Impact Parameter 43

IPHC Institut pluridisciplinaire Hubert Curien 46

L1T Level-1 Trigger 31

LH left handed 4

LHC Large Hadron Collider 16

LU Lepton Universality 10

MC Monte Carlo 40

MET Missing Transverse Energy 26

ML Machine Learning 63

fastMTT fastMTT 94

P parity transformation 19

PDF parton distribution function 40

PF Particle Flow Algorithm 26

PNet ParticleNet Algorithm 35

POG Physics Object Group 10

pp proton–proton 23

PS Proton Synchrotron 24

PSB Proton Synchrotron Booster 24

PUPPI PileUp Per Particle Identification 35

PV primary vertex 36

$\vec{h}_{PV}$ Polarimetric Vector 47

QED Quantum Electrodynamics 4

RF reference frame 58

RH right handed 4

RPC Resistive Plate Chambers 31

SM Standard Model 3

SSB spontaneous symmetry breaking 16

SPS Super Proton Synchrotron 24

SV secondary vertex 33

cSVfit classic Secondary Vertex Fit 52

T time reversal 19

ttH top quark associated Higgs boson production 16

VBF vector boson fusion 16

VEV vacuum expectation value 5

VH Higgs boson bremsstrahlung 16

WP Working Point 36

Appendix B

Columnflow Taskgraph

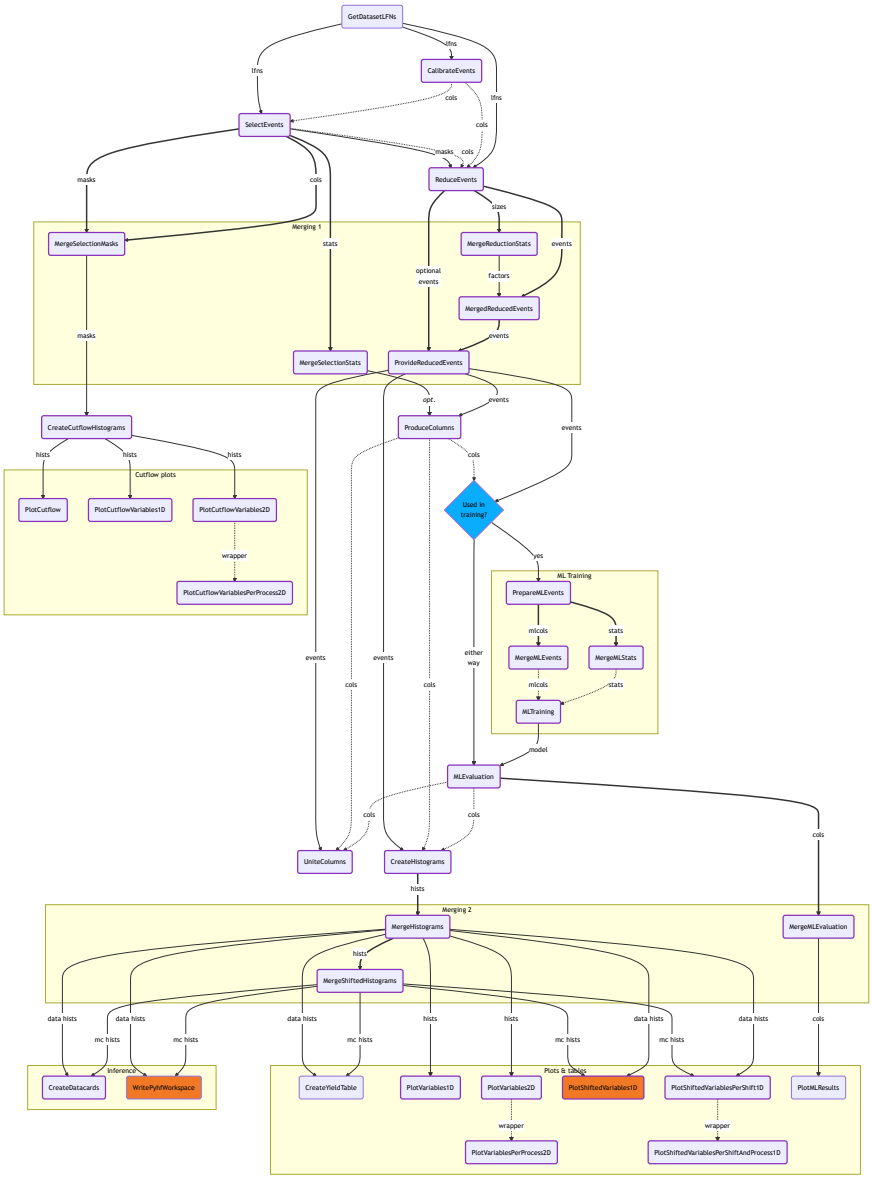


Figure B.1: Columnflow Taskgraph [81].

Appendix C

Event Yields of the DesyTau framework

	ggH, SM	ggH, MM	ggH, PS	qqH, SM	qqH, MM	qqH, PS	$Z \to \tau\tau$ + jet fakes	$Z \to \ell\ell$	Di-Boson	$t\bar{t}$ + Jets	Single $t/\bar{t}$	W + jets	Data
2022preEE													
$\mu\tau$ inclusive	596.8(13)	596.6(13)	595.1(13)	56.71(12)	56.76(12)	56.78(12)	90 780(210)	26 130(130)	9811(27)	87 760(50)	7085(10)	369 600(900)	1 336 500(1200)
$\mu\tau$ inclusive signal region	155.5(7)	155.9(7)	156.2(7)	16.27(6)	16.30(7)	16.32(7)	17 040(90)	212(10)	189(4)	1499(6)	128.2(11)	820(40)	19 150(140)
$\mu\pi$	14.09(22)	14.04(21)	14.17(23)	1.414(19)	1.421(19)	1.416(19)	1317(24)	34(4)	14.5(11)	97.4(14)	8.16(27)	61(11)	1400(40)
$\mu\rho$	51.3(4)	51.7(4)	51.6(4)	5.39(4)	5.42(4)	5.43(4)	5360(50)	39(4)	59.0(22)	477.7(31)	39.8(6)	244(22)	6000(80)
μa_1^{1pr}	9.28(16)	9.27(16)	9.25(16)	0.994(16)	0.978(16)	0.972(16)	977(21)	1.8(16)	13.4(11)	105.5(16)	9.20(28)	52(10)	1162(34)
μa_1^{3pr}	29.34(29)	29.62(29)	29.78(29)	3.000(28)	3.005(28)	3.013(28)	3810(40)	7.2(23)	41.2(18)	304.4(25)	26.6(5)	123(16)	4250(70)
μa_1^{3pr} BDT $D_H \in (0.33, 0.5]$	1.57(7)	1.60(7)	1.58(7)	0.108(5)	0.107(5)	0.109(5)	116(7)	0.0	0.76(23)	8.9(4)	0.74(7)	0.0	148(12)
μa_1^{3pr} BDT $D_H \in (0.5, 0.7]$	6.63(14)	6.73(14)	6.73(14)	0.556(12)	0.554(12)	0.559(12)	239(10)	0.0	3.3(5)	24.7(7)	2.18(13)	0.0	293(17)
μa_1^{3pr} BDT $D_H \in (0.7, 1.0]$	10.41(18)	10.47(18)	10.45(18)	1.690(21)	1.691(21)	1.691(21)	133(7)	0.0	3.0(5)	22.6(7)	2.51(14)	0.0	187(14)
2022postEE													
$\mu\tau$ inclusive signal region	475.7(12)	476.3(12)	476.0(12)	49.76(12)	49.82(12)	49.90(12)	51 930(150)	699(18)	586(7)	4754(9)	407.5(18)	2910(70)	63 660(250)
μa_1^{3pr}	88.7(5)	88.7(5)	88.4(5)	9.07(5)	9.04(5)	9.04(5)	11 560(70)	15(4)	123.1(30)	940(4)	80.8(8)	463(29)	13 100(110)
μa_1^{3pr} BDT $D_H \in (0.7, 1.0]$	33.01(32)	33.02(33)	32.84(32)	5.10(4)	5.08(4)	5.11(4)	391(12)	0.0	8.3(8)	68.9(11)	7.46(23)	0.0	577(24)
2023preBPix													
$\mu\tau$ inclusive signal region	332.7(10)	333.4(10)	333.0(10)	34.16(10)	34.21(10)	34.34(10)	36 130(140)	242(15)	380(6)	3135(13)	269.3(24)	1350(80)	47 410(220)
μa_1^{3pr}	63.3(4)	63.5(4)	63.3(4)	6.38(4)	6.42(4)	6.41(4)	8140(70)	3(5)	86.6(28)	650(6)	55.9(11)	213(27)	10 140(100)
μa_1^{3pr} BDT $D_H \in (0.7, 1.0]$	21.48(25)	21.37(25)	21.29(25)	3.397(30)	3.405(30)	3.401(31)	242(11)	0.0	5.2(7)	48.2(16)	4.84(30)	0.0	392(20)
2023postBPix													
$\mu\tau$ inclusive signal region	177.0(7)	177.4(7)	177.0(7)	18.32(7)	18.25(7)	18.19(7)	19 480(100)	158(11)	200(5)	1687(10)	143.4(18)	750(60)	25 170(160)
μa_1^{3pr}	33.66(32)	33.76(32)	33.64(32)	3.419(30)	3.398(29)	3.363(29)	4450(50)	7.5(29)	47.7(22)	353(4)	29.5(8)	114(21)	5450(70)
μa_1^{3pr} BDT $D_H \in (0.7, 1.0]$	12.24(20)	12.31(20)	12.09(20)	1.861(22)	1.861(22)	1.834(22)	147(8)	0.0	2.8(6)	26.2(12)	3.03(25)	0.0	207(14)
Combined													
$\mu\tau$ inclusive signal region	1140.9(19)	1143.0(19)	1142.2(19)	118.51(16)	118.58(16)	118.75(16)	124 580(250)	1311(25)	1355(11)	11 075(18)	948.4(33)	5830(110)	155 390(360)
μa_1^{3pr}	215.0(8)	215.6(8)	215.1(8)	21.87(6)	21.86(6)	21.83(6)	27 960(110)	32.7(66)	298.6(46)	2247(8)	192.8(15)	913(39)	32 940(170)
μa_1^{3pr} BDT $D_H \in (0.7, 1.0]$	77.14(47)	77.17(48)	76.67(47)	12.05(5)	12.04(5)	12.04(5)	913(17)	0.0	19.3(12)	165.9(23)	17.84(45)	0.0	1363(32)

Table C.1: Event yields for the DesyTau framework for $\mu\tau$ final states in signal and background processes, and data for the 4 configurations 2022preEE, 2022postEE, 2023preBPix, 2023postBPix, and their sum.

Appendix D

Further Distributions for the CP Measurement

D.1 Distributions of Observables

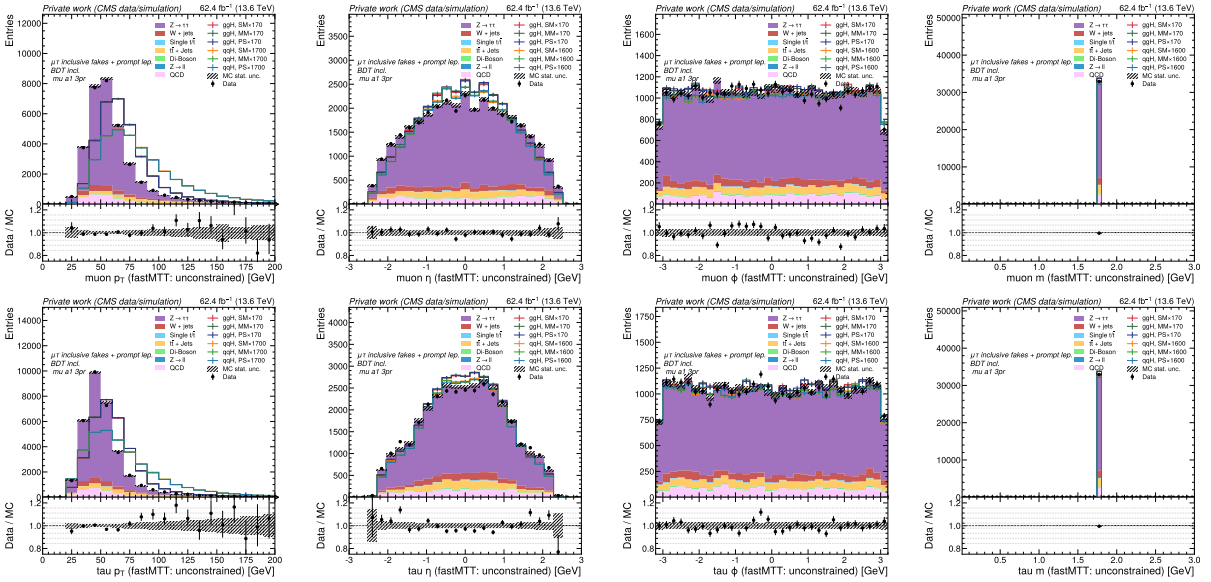
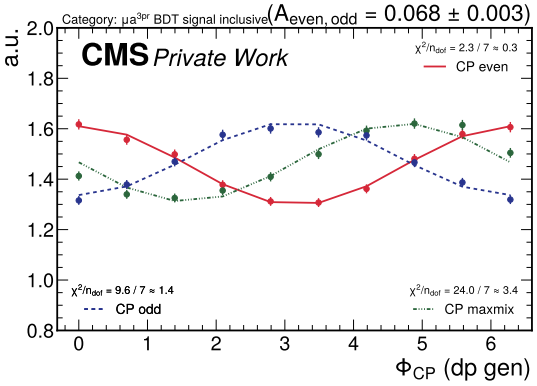
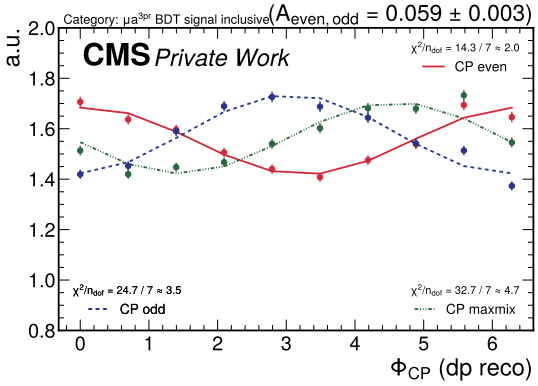


Figure D.1: Control Plots for muon (top) and hadronic tau (bottom) observables p_T, η, ϕ, m, with an unconstrained fastMTT approach, for the full 2022 and 2023 data, in the DesyTau framework.

D.2 ϕ_{CP} Distributions for VBF Production

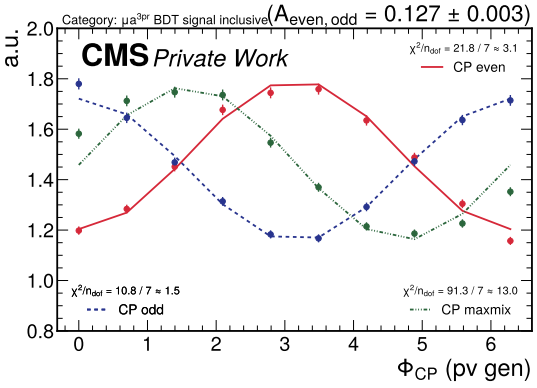


(a) Generator-level.

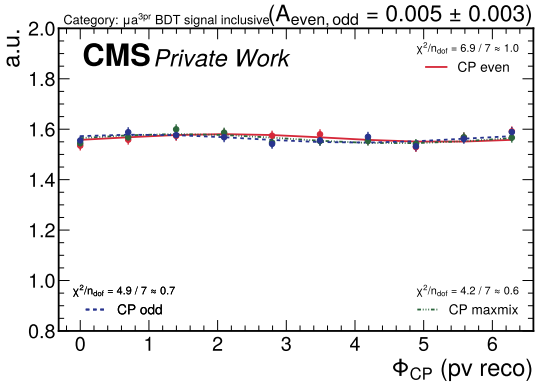


(b) Visible-level.

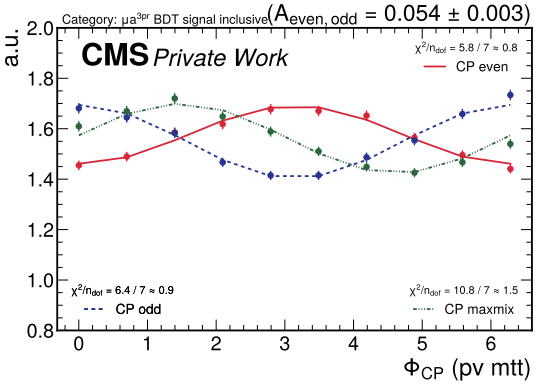
Figure D.2: Distributions of the coplanarity angle ϕ_{CP} in the VBF category for the DP method on the hadronic tau side. (a) shows the distribution at generator-level, corresponding to the maximal achievable sensitivity in this channel, while (b) shows the same observable reconstructed at visible-level.



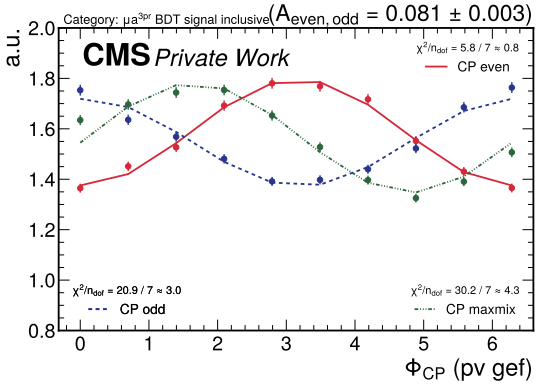
(a) Generator-level.



(b) Visible-level.



(c) fastMTT-level.



(d) GEF-level.

Figure D.3: Distributions of the coplanarity angle ϕ_{CP} in the VBF category obtained with the $\vec{h}_{PV}$ method on the hadronic tau side, shown at generator-level (a), visible-level (b), after the fastMTT modular correction (c), and after the additional GEF angular correction (d).

Bibliography

[1] Particle Data Group Collaboration, S. Navas et al., "Review of particle physics." *Phys. Rev. D* **110** (2024), no. 3 030001.

[2] E. Noether, "Invariante variationsprobleme." *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* **1918** (1918) 235–257.

[3] P. Schleper, "Avanced particle physics." Lecture notes, Universität Hamburg, Germany (2023,2024). PDF verfügbar unter: `https://www.physik.uni-hamburg.de/en/iexp/gruppe-schleper/personen/schleper-peter.html`.

[4] J. Riebesell, "Higgs potential." TikZ.net (2022). Accessed: 20 October 2025.

[5] C. Amsler et al., "Review of particle physics." *Physics Letters B* **667** (2008), no. 1-5 1–6. Bibcode: 2008PhLB..667....1A; Archived 2012-07-12; Accessed 2015-09-04.

[6] C. Burgard and D. Galbraith, "Standard model of physics." TeXample.net (2016). Diagram created at CERN Webfest 2012. Programmed in TikZ by Carsten Burgard. Licensed under CC BY-SA 4.0.

[7] R. Aaij, C. A. Beteta, T. Ackernley, et al., "Test of lepton universality in beauty-quark decays." *Nature Physics* **18** (Mar., 2022) 277–282.

[8] M. Goldhaber, L. Grodzins, and A. W. Sunyar, "Helicity of neutrinos." *Phys. Rev.* **109** (Feb, 1958) 1015–1017.

[9] W. Fetscher, "HELICITY OF THE muon-neutrino IN PI+ DECAY: A COMMENT ON THE MEASUREMENT OF P (MU) XI DELTA / RHO IN MUON DECAY." *Phys. Lett. B* **140** (1984) 117–118.

[10] V. Cherepanov, *Measurement of the polarization of tau-leptons produced in Z decays at CMS and determination of the effective weak mixing angle*, ph.d. thesis, RWTH Aachen University (2016). First measurement of τ polarization in $Z \to \tau\tau$ decays at 8 TeV with CMS.

[11] M. Worek, "Higgs cp from h/a -> tau tau decay." (2003).

[12] CMS Collaboration, A. Hayrapetyan, A. Tumasyan, et al., "Measurement of the τ lepton polarization in z boson decays in proton-proton collisions at $\sqrt{s} = 13$ tev." *J. High Energ. Phys.* **101** (2024) 101.

[13] P. H. Eberhard, B. van Eijk, J. A. Fuster, S. Jadach, A. M. Lutz, E. Richter-Was, P. Rosselet, O. Schneider, and Z. Was, "The tau polarization measurement at LEP.".

[14] A. Tumasyan, W. Adam, J. W. Andrejkovic, T. Bergauer, S. Chatterjee, et al., "Analysis of the cp structure of the yukawa coupling between the higgs boson and leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV." *Journal of High Energy Physics* **2022** (June, 2022).

[15] P. W. Higgs, "Spontaneous symmetry breakdown without massless bosons." *Phys. Rev.* **145** (May, 1966) 1156–1163.

[16] P. W. Higgs, "Broken symmetries, massless particles and gauge fields." *Phys. Lett.* **12** (1964) 132–133.

[17] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons." *Phys. Rev. Lett.* **13** (1964) 321–323.

[18] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global Conservation Laws and Massless Particles." *Phys. Rev. Lett.* **13** (1964) 585–587.

[19] G. Aad, T. Abajyan, B. Abbott, et al., "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** (sep, 2012) 1–29.

[20] S. Chatrchyan, V. Khachatryan, A. Sirunyan, et al., "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC." *Physics Letters B* **716** (sep, 2012) 30–61.

[21] A. Hayrapetyan, A. Tumasyan, W. Adam, et al., "Measurement of the higgs boson mass and width using the four-lepton final state in proton-proton collisions at $\sqrt{s} =$ 13 TeV." *Physical Review D* **111** (May, 2025).

[22] CERN, "Cern yellow reports: Monographs, vol 2 (2017): Handbook of lhc higgs cross sections: 4. deciphering the nature of the higgs sector." (2017).

[23] G. Barr, R. Devenish, R. Walczak, and T. Weidberg, *Particle Physics in the LHC Era*. Oxford Master Series in Particle Physics, Astrophysics, and Cosmology. Oxford University Press, Oxford (2016).

[24] CMS Collaboration, S. Chatrchyan, V. Khachatryan, A. M. Sirunyan, et al., "Study of the mass and spin-parity of the higgs boson candidate via its decays to z boson pairs." *Phys. Rev. Lett.* **110** (Feb, 2013) 081803.

[25] G. Aad, T. Abajyan, B. Abbott, and J. A. others, "Evidence for the spin-0 nature of the higgs boson using atlas data." *Physics Letters B* **726** (2013), no. 1 120–144.

[26] S. Chatrchyan, V. Khachatryan, A. Sirunyan, et al., "Measurement of the properties of a higgs boson in the four-lepton final state." *Physical Review D* **89** (May, 2014).

[27] H. Fritzsch and Z. zhong Xing, "The light quark sector, cp violation, and the unitarity triangle." *Nuclear Physics B* **556** (1999), no. 1 49–75.

[28] G. Branco, L. Lavoura, and J. Silva, *CP violation*. Oxford University Press, United Kingdom (May, 2024). Publisher Copyright: © G. C. Branco, L. Lavoura, and J. P. Silva 1999.

[29] L. B. Okun, *Leptoni e quark*. Editori Riuniti, Rome (1986).

[30] X. Chen and Y. Wu, "Probing the cp-violation effects in the h coupling at the lhc." *Physics Letters B* **790** (Mar., 2019) 332–338.

[31] 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.

[32] C. Lefèvre, "The CERN accelerator complex." 2008.

[33] CMS Collaboration, "Public cms luminosity information." (2024). Accessed: 2025-10-22.

[34] CERN, "Cern yellow reports: Monographs, vol 2 (2017): Handbook of lhc higgs cross sections: 4. deciphering the nature of the higgs sector." (2017).

[35] CMS Collaboration, G. L. Bayatian, S. Chatrchyan, G. Hmayakyan, A. M. Sirunyan, et al., *CMS Physics*. Technical design report. CMS. CERN, Geneva (2006). There is an error on cover due to a technical problem for some items.

[36] CMS Collaboration, T. Sakuma, "Cutaway diagrams of CMS detector.".

[37] I. Neutelings, "Cms coordinate system." (2024). Accessed: 21 October 2025.

[38] CMSTrackerGroupoftheCMS Collaboration, W. CMS Collaboration, Adam, T. Bergauer, et al., "The CMS Phase-1 Pixel Detector Upgrade." *JINST* **16** (2021), no. 02 P02027, [arXiv:2012.14304].

[39] A. Dominguez, D. Abbaneo, K. Arndt, N. Bacchetta, A. Ball, E. Bartz, W. Bertl, G. M. Bilei, G. Bolla, H. W. K. Cheung, M. Chertok, S. Costa, N. Demaria, D. D. Vazquez, K. Ecklund, W. Erdmann, K. Gill, G. Hall, K. Harder, F. Hartmann, R. Horisberger, W. Johns, H. C. Kaestli, K. Klein, D. Kotlinski, S. Kwan, M. Pesaresi, H. Postema, T. Rohe, C. Schäfer, A. Starodumov, S. Streuli, A. Tricomi, P. Tropea, J. Troska, F. Vasey, and W. Zeuner, "CMS Technical Design Report for the Pixel Detector Upgrade." tech. rep., 2012. Additional contacts: Jeffrey Spalding, Fermilab, Jeffrey.Spalding@cern.ch Didier Contardo, Universite Claude Bernard-Lyon I, didier.claude.contardo@cern.ch.

[40] CMS Collaboration, "Cms tracker detector performance results - cms public wiki." Accessed: 22 October 2025.

[41] CMS Collaboration, P. Bloch, R. Brown, P. Lecoq, and H. Rykaczewski, *Changes to CMS ECAL electronics*. Technical design report. CMS. CERN, Geneva (2002).

[42] CMS Collaboration, A. Tapper and D. Acosta, "CMS Technical Design Report for the Level-1 Trigger Upgrade." tech. rep., 2013. Additional contacts: Jeffrey Spalding, Fermilab, Jeffrey.Spalding@cern.ch Didier Contardo, Universite Claude Bernard-Lyon I, didier.claude.contardo@cern.ch.

[43] A. Sirunyan, A. Tumasyan, W. Adam, E. Asilar, et al., "Particle-flow reconstruction and global event description with the cms detector." *Journal of Instrumentation* **12** (Oct., 2017) P10003–P10003.

[44] I. Neutelings, "Hadronic tau decay." (July, 2021). Accessed: 20 October 2025.

[45] T. C. Collaboration, "Description and performance of track and primary-vertex reconstruction with the cms tracker." *Journal of Instrumentation* **9** (Oct., 2014) P10009–P10009.

[46] A. Sirunyan, A. Tumasyan, W. Adam, F. Ambrogi, and orthers, "Performance of the cms muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV." *Journal of Instrumentation* **13** (jun, 2018) P06015.

[47] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-ktjet clustering algorithm." *Journal of High Energy Physics* **2008** (Apr., 2008) 063–063.

[48] A. Sirunyan, A. Tumasyan, W. Adam, F. Ambrogi, et al., "Pileup mitigation at cms in 13 tev data." *Journal of Instrumentation* **15** (sep, 2020) P09018.

[49] CMS Collaboration, "Validation of tau identification algorithms in Run 3 data and simulation.". CMS-DP-2025-074.

[50] CMS Collaboration, "Performance of reconstruction and identification of tau leptons in their decays to hadrons and tau neutrino in LHC Run-2." tech. rep., CERN, Geneva, 2016.

[51] A. Tumasyan, W. Adam, J. Andrejkovic, et al., "Identification of hadronic tau lepton decays using a deep neural network." *Journal of Instrumentation* **17** (jul, 2022) P07023.

[52] CMS Collaboration, "Identification of tau leptons using a convolutional neural network with domain adaptation in the CMS experiment." tech. rep., CERN, Geneva, 2025.

[53] A. Hayrapetyan, V. Makarenko, A. Tumasyan, et al., "Identification of tau leptons using a convolutional neural network with domain adaptation." *Journal of Instrumentation* **20** (Dec., 2025) P12032.

[54] H. Qu and L. Gouskos, "Jet tagging via particle clouds." *Physical Review D* **101** (Mar., 2020).

[55] A. C. for the CMS Collaboration, "Recent advancements in the tau reconstruction and identification techniques in CMS." CMS Conference Report CMS CR-2025/277, TAU2025 Workshop (Nov., 2025). Available at: urlhttps://cds.cern.ch/record/2950104/files/CR2025$_2$77.*pdf*.

[56] A. Kalinowski and W. Matyszkiewicz, "Efficient tau-pair invariant mass reconstruction with simplified matrix element techniques." (2025).

[57] W. Matyszkiewicz and A. Kalinowski, "Tau-pair Invariant Mass Estimation Using Maximum Likelihood Estimation and Collinear Approximation." *Acta Phys. Polon. Supp.* **18** (2025), no. 5 5–A21.

[58] L. Bianchini, J. Conway, E. K. Friis, and C. Veelken, "Reconstruction of the higgs mass in $h \to \tau\tau$ events by dynamical likelihood techniques." *Journal of Physics: Conference Series* **513** (jun, 2014) 022035.

[59] V. Cherepanov and A. Zotz, "Kinematic reconstruction of $z/h \to \tau\tau$ decay in proton-proton collisions." (2018).

[60] P. Nason, "A New method for combining NLO QCD with shower Monte Carlo algorithms." *JHEP* **11** (2004) 040, [hep-ph/0409146].

[61] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method." *JHEP* **11** (2007) 070, [arXiv:0709.2092].

[62] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX." *JHEP* **06** (2010) 043, [arXiv:1002.2581].

[63] T. Ježo and P. Nason, "On the Treatment of Resonances in Next-to-Leading Order Calculations Matched to a Parton Shower." *JHEP* **12** (2015) 065, [arXiv:1509.09071].

[64] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations." *Journal of High Energy Physics* **2014** (July, 2014).

[65] R. Frederix, S. Frixione, V. Hirschi, D. Pagani, H.-S. Shao, and M. Zaro, "The automation of next-to-leading order electroweak calculations." *Journal of High Energy Physics* **2018** (July, 2018).

[66] C. Bierlich, S. Chakraborty, N. Desai, L. Gellersen, I. Helenius, P. Ilten, L. Lönnblad, S. Mrenna, S. Prestel, C. T. Preuss, T. Sjöstrand, P. Skands, M. Utheim, and R. Verheyen, "A comprehensive guide to the physics and usage of pythia 8.3." (2022).

[67] R. Frederix and S. Frixione, "Merging meets matching in mc@nlo." *Journal of High Energy Physics* **2012** (Dec., 2012).

[68] J. Alwall, S. Höche, F. Krauss, N. Lavesson, L. Lönnblad, F. Maltoni, M. Mangano, M. Moretti, C. Papadopoulos, F. Piccinini, S. Schumann, M. Treccani, J. Winter, and M. Worek, "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions." *The European Physical Journal C* **53** (Dec., 2007) 473–500.

[69] P. Artoisenet, R. Frederix, O. Mattelaer, and R. Rietkerk, "Automatic spin-entangled decays of heavy resonances in monte carlo simulations." *Journal of High Energy Physics* **2013** (Mar., 2013).

[70] S. Frixione, E. Laenen, P. Motylinski, and B. R. Webber, "Angular correlations of lepton pairs from vector boson and top quark decays in monte carlo simulations." *Journal of High Energy Physics* **2007** (Apr., 2007) 081–081.

[71] T. Przedziński, E. Richter-Was, and Z. Was, "Tauspinner: a tool for simulating cp effects in $h \to \tau\tau$ decays at lhc." *The European Physical Journal C* **74** (Nov., 2014).

[72] J. Allison, K. Amako, J. Apostolakis, P. Arce, et al., "Recent developments in geant4." *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **835** (2016) 186–225.

[73] S. Agostinelli, J. Allison, K. Amako, J. Apostolakis, et al., "Geant4—a simulation toolkit." *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** (2003), no. 3 250–303.

[74] CMS Muon Physics Object Group, "Muon reconstruction and identification." `https://twiki.cern.ch/twiki/bin/view/CMS/MuonPOG` (2025). CMS internal documentation (legacy TWiki, frozen).

[75] CMS Tau Physics Object Group, "Tau reconstruction and identification." `https://twiki.cern.ch/twiki/bin/viewauth/CMS/Tau` (2025). CMS internal documentation.

[76] A. Cardini, "Methodologies to measure the cp structure of the higgs yukawa coupling to tau leptons." *Universe* **8** (2022), no. 5.

[77] M. Sessini, *Recherche de violation de CP dans le couplage de Yukawa du lepton tau avec la méthode du vecteur polarimétrique dans les données du Run 2 de l'expérience CMS auprès du LHC*, theses, Université de Strasbourg (Dec., 2023).

[78] N. Davidson, G. Nanava, T. Przedziński, E. Richter-Was, and Z. Was, "Universal interface of tauola: Technical and physics documentation." *Computer Physics Communications* **183** (2012), no. 3 821–843.

[79] CLEO Collaboration, D. M. Asner, A. Eppich, J. Gronberg, T. S. Hill, and other, "Hadronic structure in the decay $\tau^- \to \nu_\tau \pi^- \pi^0 \pi^0$ and the sign of the tau neutrino helicity." *Phys. Rev. D* **61** (Dec, 1999) 012002.

[80] H. Dembinski and P. O. et al., "scikit-hep/iminuit.".

[81] columnflow contributors, "Default task graph in the columnflow Wiki." `https://github.com/columnflow/columnflow/wiki/Default-task-graph` (2025). Accessed: 2026-01-10.

12/2024

FB Physik: Eidesstattliche Erklärung für Abschlussarbeiten

Folgende Formulierung inkl. Unterschrift ist in Ihre Abschlussarbeit aufzunehmen.

Eidesstattliche Erklärung

Ich versichere, dass ich die Bachelorarbeit/Masterarbeit selbstständig angefertigt und keine anderen als die angegebenen Hilfsmittel benutzt habe. Alle Stellen, die dem Wortlaut oder dem Sinn nach anderen Werken entnommen sind, habe ich in jedem einzelnen Fall unter genauer Angabe der Quelle deutlich als Entlehnung kenntlich gemacht. Ich versichere, diese Arbeit nicht bereits in gleicher oder ähnlicher Fassung in einem anderen Prüfungsverfahren eingereicht zu haben.

Sofern im Zuge der Erstellung der vorliegenden Bachelorarbeit/Masterarbeit generative Künstliche Intelligenz (gKI) basierte elektronische Hilfsmittel verwendet wurden, versichere ich, dass meine eigene Leistung im Vordergrund stand und dass eine vollständige Dokumentation aller verwendeten Hilfsmittel gemäß der Guten wissenschaftlichen Praxis vorliegt. Ich trage die Verantwortung für eventuell durch die gKI generierte fehlerhafte oder verzerrte Inhalte, fehlerhafte Referenzen, Verstöße gegen das Datenschutz- und Urheberrecht oder Plagiate.

Ich bin damit/~~nicht damit~~* einverstanden, dass die ~~Bachelorarbeit~~/Masterarbeit* veröffentlicht wird.

Hamburg, 28.01.2026 M. Witt

Ort, Datum Unterschrift

**nicht Zutreffendes streichen bzw. weglassen.*

12/2024

Physics Department: Affidavit for final theses

The following formulation including signature must be included in your thesis.

Statutory Declaration

I declare that I have authored the Bachelor's/Master's thesis independently and have not used any aids other than those specified. I have clearly labelled all passages that are taken from other works in terms of wording or meaning as borrowed material, stating the exact source in each case. I confirm that I have not already submitted this thesis in the same or a similar version in another examination procedure.

If electronic aids based on generative artificial intelligence (gAI) were used in the course of the preparation of this Bachelor's/Master's thesis, I confirm that my own work was the main focus and that complete documentation of all aids used is available in accordance with good scientific practice. I am responsible for any incorrect or distorted content, incorrect references, violations of data protection and copyright law or plagiarism generated by the gAI.

I agree~~/disagree~~* that the ~~Bachelor's~~/Master's thesis* may be published.

Hamburg, 28 January 2026 M. Witt

Place, date Signature

**Delete or omit as appropriate.*