Effects of tau decay product reconstruction in a Higgs CP analysis with the ATLAS experiment

Michael Hübner

Masterarbeit in Physik
angefertigt im Physikalischen Institut

vorgelegt der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität
Bonn

September 2016

I hereby declare that the work presented here was formulated by myself and that no sources or tools other than those cited were used.

Bonn, ...
Date Signature

1. Gutachter: Prof. Dr. Klaus Desch
2. Gutachter: Prof. Dr. Jochen Dingfelder

Acknowledgements

I would like to thank Prof. Klaus Desch for the opportunity to write my thesis in his group and for his support. I would also like to thank Peter Wagner and Philip Bechtle for their support and all the insightful discussions. For the technical support and the deployment of the xTauFramework I want to thank its authours, especially Christian Grefe who also provided me with the plotting framework used as the basis for my plots. I also want to thank the whole working group for the nice atmosphere. Finally, I want to thank my whole family and Tascha for their encouragment and support over the past years.

Contents

CHAPTER 1

Introduction

In 2012 the experiments ATLAS and CMS at the LHC discovered a new resonance, supposedly the Higgs boson. [1] [2] With a mass of $125\,\mathrm{GeV/c^2}$ this new particle could act as an, up to then, missing puzzle piece of the Standard Model by confirming the Higgs mechanism proposed by Peter Higgs, François Englert and Robert Brout. Without this mechanism all particles in the Standard Model would have to be massless which stands in complete contrast to experimental observations. Peter Higgs and François Englert were awarded with the nobel prize in physics for this mechanism in 2013. [3]

In order to confirm that the new resonance is indeed the Higgs boson further measurements on the quantum numbers of the resonance were performed. So far, all measurements are in agreement with the Standard Model predictions and seem to confirm the SM Higgs boson.

Nevertheless, there is still room for the possibility of a non-SM Higgs. The measurements performed in the coupling to gauge bosons exclude a pure $\mathcal{CP}$-odd Higgs boson at a given confidence level but they don't exclude the possibility of a $\mathcal{CP}$-mixed Higgs. [4] Small admixtures of a $\mathcal{CP}$-odd state to the $\mathcal{CP}$-even state are still possible and would be a direct hint to physics beyond the Standard Model. Some beyond the Standard Model theories predict the $\mathcal{CP}$-mixing to happen in the coupling to fermions. The mixing in this sector is not limited by the performed measurements.

This thesis discusses a method that is capable of measuring $\mathcal{CP}$-admixtures, i.e. it is possible to measure a mixing angle ϕ_τ of the two states. To do so, the angle between the planes spanned by the decay products of the Higgs boson, namely τ leptons, is used. An event selection adapted from the $H \to \tau\tau$ coupling analysis is used and its impact on the measurement is investigated. Furthermore, reconstruction effects in phase space regions defined by the event selection are studied. The reconstruction effects are studied separately for the different parts of the method used to measure the mixing angle ϕ_τ. Following the reconstruction effect studies, an attempt to improve the method's performance at reconstruction level is investigated.

In this context different ditau-mass reconstruction techniques are discussed. The most promising technique, the Missing Mass Calculator, for the purpose of improving the method's performance at reconstruction level is studied in more detail. This study includes a new tune for the MMC which is part of the DiTauMassTools package (tag 00-00-24) and also part of the AnalysisBase release 2.4.14. The main goal of the MMC is to discriminate $Z \to \tau\tau$ against $H \to \tau\tau$. Hence, it is used in different analyses, see for example ref. [5] and ref. [6].

Throughout this thesis the convention is used that sets the speed of light to $c = 1$ and gives masses, momenta and energies in units of GeV. All studies are based on Monte Carlo simulations.

CHAPTER 2

Theory

A short introduction of the Standard Model including its short-comings will be given in the following chapter. Followed by a motivation for the search of $\mathcal{CP}$-violation in the Higgs sector the theoretical background for a measurement of such a $\mathcal{CP}$-violation will be discussed.

2.1 Standard Model

The Standard Model of Particle Physics, in the following called SM, is a gauge theory $SU(3)_C \times SU(2)_L \times U(1)_Y$ [7] that describes all known elementary particles and their interactions, i.e. the electromagnetic, weak and strong force but not gravity. All forces are mediated via so-called gauge bosons, where bosons mean that they have integer spin $S = 0, 1$. Contrary to the bosons, fermions have spin $S = \frac{1}{2}$ and build the basis for all known visible matter. The particle content of the SM is shown in figure 2.1 including some additional information of the particles like their masses. All shown fermions have corresponding anti-particles which have the same mass but opposite electric, weak and colour charge. Figure 2.1 also includes the Higgs boson which is a result of the self-interaction of the Higgs field and will be further described in section 2.2.

The fermions can be grouped together in three generations where the second generation and the third generation are heavier copies of the first generation. Subgroups of the fermions are the leptons and the quarks. All colour-neutral objects, like the leptons, can move freely whereas the colour-charged objects, like the quarks, cannot. This leads to the fact that colour-charged objects always build colour-neutral states. These states are composed of grouped-together quarks and are then called hadrons. Hadrons can be composed of a different number of quarks, like mesons for a bound state of two quarks and baryons for a bound state of three quarks. Section 2.1.2 gives some additional information on this behaviour and its source, the Quantum Chromo Dynamics. Recently, there were hints discovered for possible tetraquark- [9] [10] [11]- and pentaquark-states [12] [13]. In the SM the gauge bosons mediate a force by coupling to fermions. Furthermore, it is also possible for the gauge bosons to couple among themselves in specific combinations.

2.1.1 Quantum Electrodynamics and Electroweak Unification

The electromagnetic force is one of the three fundamental forces included in the SM and acts on all particles that carry an electric charge. It is also the force which is most prominent in everyday-life as it holds together atoms and molecules and describes electromagnetic fields. Its theoretical description

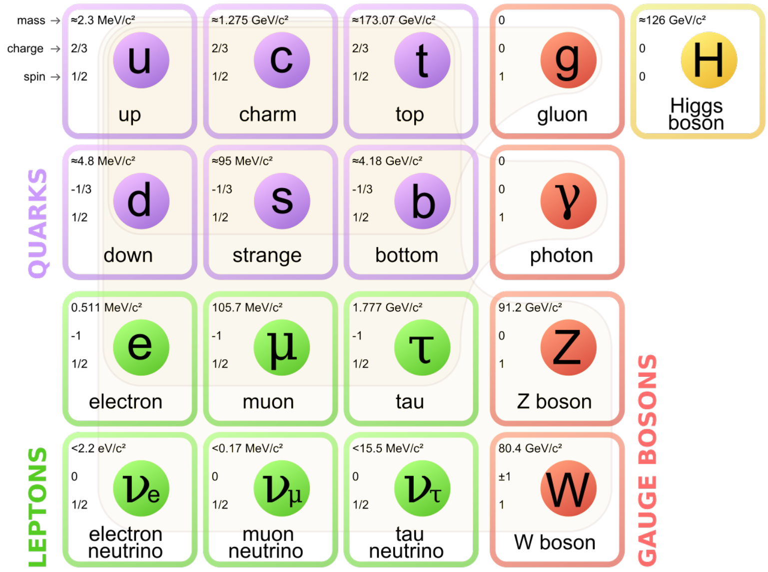


Figure 2.1: Particle content of the Standard Model of Particle Physics. Taken from ref. [8]

is called Quantum Electrodynamics, or short QED. The photon γ is the mediator of this force. It is electrically neutral and massless and does not couple to itself unlike the gluon. All forces couple with a specific strength to particles. In case of the electromagnetic force, this is the coupling constant α. However, it is not a constant like its name suggests but rather a function of the transferred momentum in a process, see fig. 2.2. It can also be expressed in an analytic formula where Q is the momentum transfer and m_e is the mass of the electron: [14]

$$\alpha\left(Q^2\right) = \frac{\alpha\left(\mu^2\right)}{1 - \alpha\left(\mu^2\right)\frac{1}{3\pi}\ln\frac{q^2}{\mu^2}} \tag{2.1}$$

The phenomenological explanation for this running is called ”screening”. Responsible for this running of α is vacuum polarisation, i.e. virtual electron-positron-pairs, which shield the electric charge of a particle. The closer one gets to a particle, the smaller is the effect of these vacuum polarisations and the higher the effective charge gets.

As a consequence of the required local gauge invariance for such theories, in order to keep them renormalisable, a new field has to be introduced to maintain this local invariance under a phase transformation. This new field can be identified as the photon and the corresponding Lagrangian of QED reads as follows: [14]

$$\mathcal{L}_{\mathcal{QED}} = \bar{\psi}\left(i\gamma^\mu\partial_\mu - m_e\right)\psi + e\bar{\psi}\gamma^\mu\psi A_\mu - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.2}$$

If the photon was massive one would also find a mass-term like $\frac{1}{2}m_\gamma^2 A_\mu A^\mu$ in this Lagrangian aside from the kinetic terms and interaction terms.

The weak force has a significantly smaller coupling constant which leads to longer lifetimes of particles that can only decay via a weak interaction compared to particles that can decay via an electromagnetic interaction or strong interaction. In everyday-life the weak force is responsible for the radioactive β-decay. It is the only force that solely couples to left-handed fermions and right-handed antifermions, thus it is maximally $\mathcal{P}$-violating but $\mathcal{CP}$-conserving. Additionally, it is the only force that allows

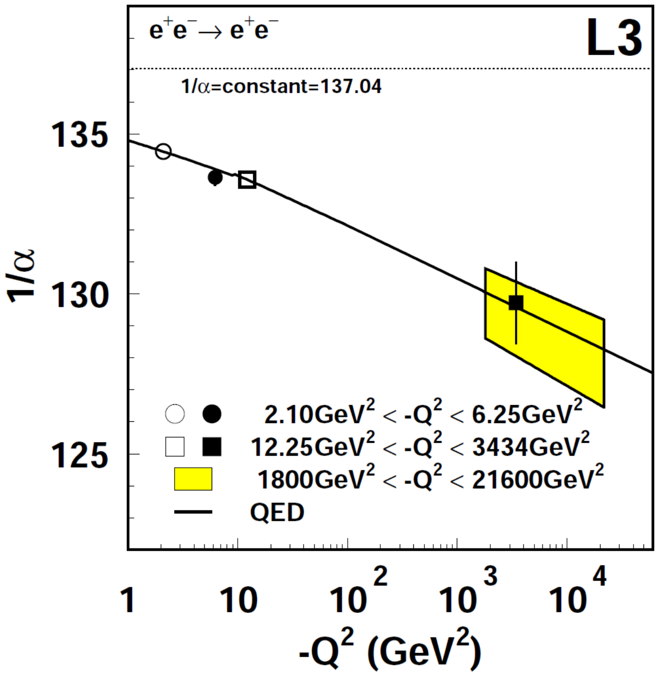


Figure 2.2: Variation of α at different Q^2 scales. Taken from ref. [15]

transitions between fermion-generations. Namely the charged current is the only possibility to change the flavour of fermions, while the neutral current, which is also part of the weak force, cannot change the flavour of fermions. The neutral current, however, is the only option for neutrinos to participate in scattering processes. In the 1960s Glashow, Weinberg and Salam combined the electromagnetic and weak interaction in the Electroweak Unification theory, represented by $SU(2) \times U(1)$. [14] This theory introduced the weak Isospin T, representing the part $SU(2)$, where its third component T_3 is connected to the electric charge Q, and the weak hypercharge Y_W, representing the part $U(1)$, via: [14]

$$Q = T_3 + \frac{Y_W}{2}. \tag{2.3}$$

The weak hypercharge is a conserved quantity in weak interactions, leading to a change of the third component of the weak Isospin. Left-handed particles have $|T_3| = \frac{1}{2}$ and right-handed particles $|T_3| = 0$. The right-handed singlets are not affected by the gauge-transformation of the weak interaction and thus don't couple to the gauge bosons mediating the weak force. [14] Hence, the electroweak symmetry has to be broken if one introduces masses to the particles. This symmetry breaking causes the hypothetical W^0-boson, which would be responsible for interactions where T_3 does not change, and the B-Boson from weak hypercharge to mix and produce thereby the physical Z^0-Boson and the photon. More details are given in ref. [14] which is the source of this subsection.

2.1.2 Quantum Chromo Dynamics

The strong force is mediated by the gluons, which carry colour charge, and is responsible for the existence of, for example, protons and neutrons. Quantum Chromo Dynamics, QCD, is the theoretical framework

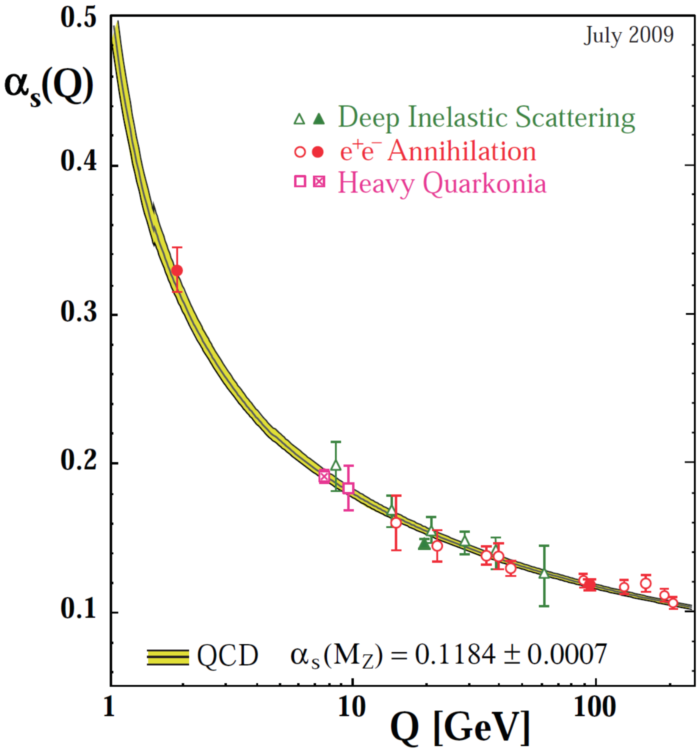


Figure 2.3: Variation of α_S at different Q scales. Taken from ref. [16]

for the strong force. Its coupling constant α_s is also not a constant. Similar to the coupling constant of the electromagnetic force, it varies with the momentum transfer as follows: [14]

$$\alpha_S\left(Q^2\right) = \frac{\alpha_S\left(\mu^2\right)}{1 + B\alpha_S\left(\mu^2\right)\ln\frac{q^2}{\mu^2}} \quad \text{with} \quad B = \frac{11N_C - 2N_f}{12\pi}. \tag{2.4}$$

Here N_C is the number of colours and N_f is the number of quark flavours accessible at the energy Q^2. Unlike α, α_s decreases with increasing momentum transfer leading to the so-called asymptotic freedom of quarks, see figure 2.3. Due to the self-interaction of gluons there are not only fermionic but also bosonic loops which influence the development of α_S at higher energy scales. These bosonic loops add negative contributions which causes α_S to decrease at high Q.

This means that the closer two quarks get, the less they are influenced by each other via the strong force. However, if two quarks move away from each other, the binding force between these two quarks increases strongly, leading to the so-called confinement. If the two quarks in this scenario reach a certain distance there is so much energy stored in the binding that the connection breaks up and a new quark-antiquark pair is produced. In high-energy collider physics this process is called ”hadronisation”. This is also the reason why there are no free quarks but only colour-neutral objects observed in nature. QCD processes can also lead to initial state radiation and final state radiation at the LHC which result in additional jets in an event. The strong force couples to left- and right-handed particles in the same way as opposed to the weak force which only couples to left-handed fermions and right-handed anti-fermions. Generally, there is a $\mathcal{CP}$-phase in the Lagrangian of QCD but this phase is experimentally tested to be consistent with

Decay mode	Branching ratio in %
leptonic	35.24
$e^- \bar{\nu}_e \nu_\tau$	17.83
$\mu^- \bar{\nu}_\mu \nu_\tau$	17.41
hadronic 1-prong	46.68
`1p0n`	10.81
`1p1n`	25.52
`1pXn`	10.35
hadronic 3-prong	13.93
`3p0n`	9.31
`3pXn`	4.62
others	4.15

Table 2.1: Branching ratios of the possible decays of a τ^-. The decays for a τ^+ are the same just with conjugated charge. [4] `XpYn` stands for a decay into `X` $\pi^\pm$ and `Y` π^0 and the corresponding τ neutrino.

zero meaning there is no $\mathcal{CP}$-violation in QCD. For a more complete introduction in QCD please consult ref. [14] which is the source for this subsection.

2.1.3 The τ Lepton

The τ lepton plays a major role in this thesis as the $H \rightarrow \tau\tau$ decay is investigated. It is the heaviest of all leptons and thus short-lived with a lifetime of $\tau_\tau = 290 \times 10^{-15}$ s [4]. Due to this short lifetime it decays inside the ATLAS experiment and it is only possible to detect its decay products. These decay products are the key to the measurement introduced in section 5. Since the mass of the τ is $m_\tau = 1\,776.82$ MeV [4], it can decay into a whole variety of other particles and even into hadrons. The most abundant of these decay modes are listed in table 2.1.

If the τ decays into an electron or muon with the corresponding neutrinos the decay is called a leptonic decay. All other decays are called hadronic decays since there are always quarks involved in the decay products and will be referred to as τ_{had} in the following. The most abundant hadronic decays are the so-called `1p0n` and `1p1n` decays. `1p0n` means that the τ decays into one $\pi^\pm$ and zero π^0 with the corresponding ν_τ whereas `1p1n` means that the τ decays into one $\pi^\pm$ and one π^0 with the corresponding ν_τ via the ρ-resonance. A schematic view of this decay is shown in figure 2.4.

2.2 Higgs Mechanism

The Higgs mechanism, for which Peter Higgs and François Englert were awarded with the nobel prize in physics in 2013 [3], introduces the spontaneous symmetry breaking of the electroweak group to give the particles their masses. A short introduction of the mechanism will be given here, for further details see ref. [17] and ref. [14].

The local gauge invariance in the electroweak unification would be broken if mass-terms for the bosons were directly added to the Lagrangian. Since only theories that are locally gauge invariant are renormalisable, the Higgs mechanism is a very important component of the SM. The local gauge invariance is maintained by adding the mass terms of the bosons through the symmetry breaking via the Higgs mechanism. Fermion mass terms are introduced by an interaction with the Higgs field. In the SM,

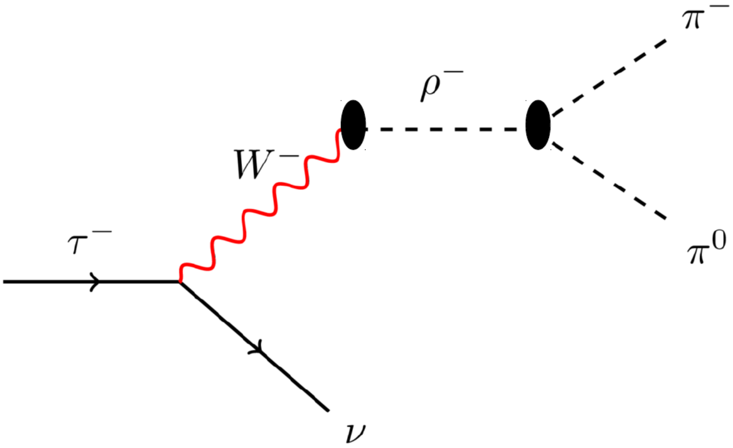


Figure 2.4: Schematic view of the τ decay into one $\pi^\pm$ and one π^0 via the ρ resonance.

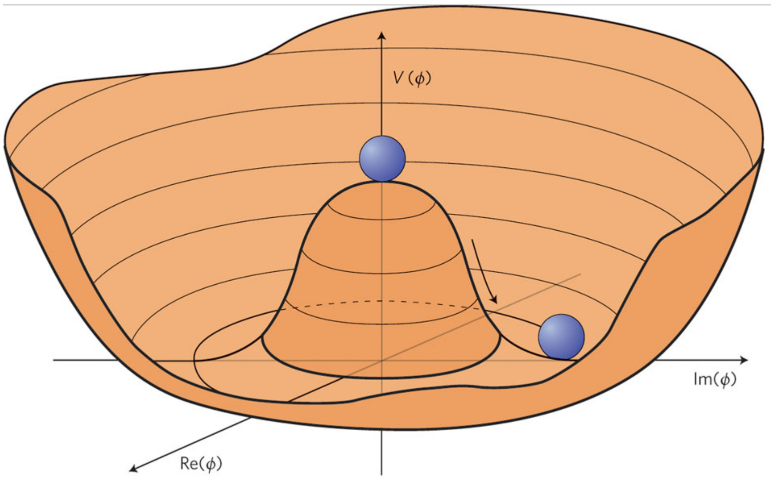


Figure 2.5: Potential of the Higgs field. Taken from ref. [18]

the Higgs field can be written as a doublet of two complex fields: [14]

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{2.5}$$

with the corresponding Lagrangian:

$$\mathcal{L} = (\partial_m \phi)^\dagger (\partial^\mu \phi) - V(\phi) = (\partial_m \phi)^\dagger (\partial^\mu \phi) - \mu^2 \phi^\dagger \phi - \lambda \left(\phi^\dagger \phi\right)^2 \tag{2.6}$$

where the minima of the potential $V(\phi)$ for $\mu^2 < 0$, see figure 2.5, satisfy:

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

$v = \sqrt{-\frac{\mu^2}{\lambda}} = 246\,\text{GeV}$ is the vacuum expectation value. For the photon to remain massless, this has to correspond to a minimum for the neutral field ϕ^0. The spontaneous symmetry breaking is encoded in the existence of a non-zero vacuum expectation value. After this spontaneous symmetry breaking the Higgs fields reads:

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \tag{2.8}$$

This is obtained by expanding the Higgs field about the vacuum expectation value and gauging away the Goldstone-Bosons which are introduced by the spontaneous symmetry breaking. In order to keep the local gauge symmetry in $SU(2)_L \times U(1)_Y$ the derivatives have to be replaced by covariant derivatives which then introduce new fields B^μ and $W^{(1-3),\mu}$. The linear combinations of the fields $W^{(1),\mu}$ and $W^{(2),\mu}$ correspond to the physical $W^\pm$-bosons and acquire mass through the longitudinal polarisations caused by the gauging away of two Goldstone-bosons, meaning the $W^\pm$-bosons absorb their degree of freedom. B^μ and $W^{(3)}$ mix and the resulting fields, A^μ and Z^μ, are the physical photon and Z-boson respectively. The mass of the $W^\pm$-bosons can be written as:

$$m_W = \frac{1}{2} g_W v \tag{2.9}$$

and the mass of the photon and Z-boson reads:

$$m_A = 0 \quad \text{and} \quad m_Z = \frac{1}{2} v \sqrt{g_W^2 + g'^2}. \tag{2.10}$$

The only free parameter introduced by this mechanism, that is not fixed by any measurement, is the mass of the Higgs boson itself. It can be written as:

$$m_H^2 = 2\lambda v^2 \tag{2.11}$$

where λ is undetermined. Since this parameter is directly connected to the mass of the Higgs-boson, a measurement of the latter determines the value for λ. Fermions acquire mass through the interaction with the Higgs field. The coupling to this field, called Yukawa coupling, is then proportional to the fermion's mass. However, the neutrinos remain massless since they don't interact with the Higgs field. The Yukawa coupling can be written as: [14]

$$g_f = \sqrt{2} \left(\frac{m_f}{v} \right). \tag{2.12}$$

The vacuum expectation value v can be determined by measuring the mass of the Higgs boson. Thus, the coupling constant g_f can be fixed by measuring the fermion's mass.

2.3 Higgs Boson

2.3.1 Production

The Higgs boson can be produced via different production channels at the LHC which are depicted in figure 2.6 including the corresponding Feynman diagrams. The production cross sections are the result of different considerations. First of all one has to consider the parton distribution functions (PDFs) for the protons that collide at the LHC. Figure 2.7 shows the PDFs at $Q^2 = 1 \times 10^4\,\mathrm{GeV}^2$. It can be seen that the probability for gluons to carry a specific fraction of the proton's momentum grows very steeply for a decreasing fraction of the proton's momentum. Hence, the most dominant production channel for the Higgs boson at the LHC, the so-called gluon-gluon-fusion or ggF, includes gluon interactions. Secondly, the Higgs boson only couples to particles that are massive resulting in no direct coupling to gluons. Thus,

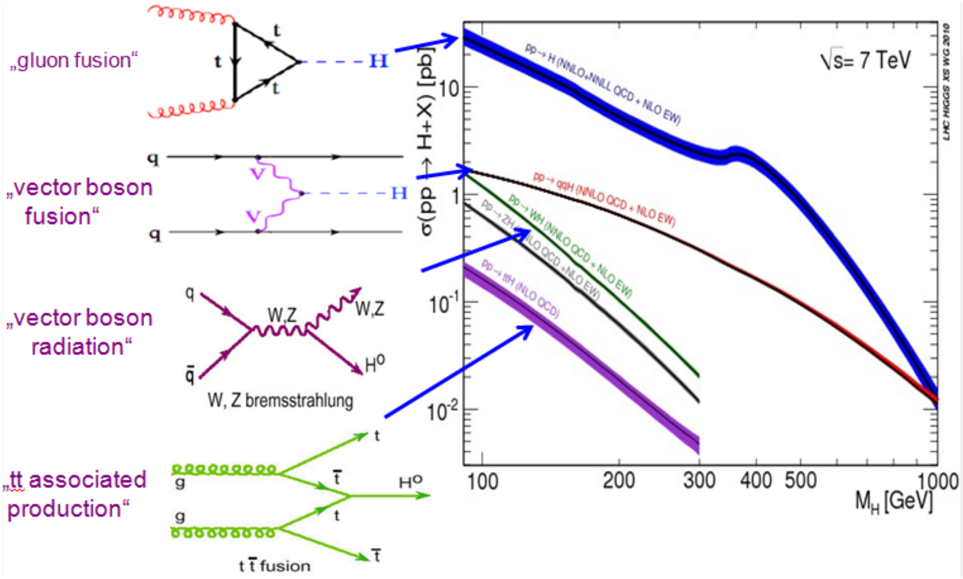


Figure 2.6: Possible production channels of the Higgs boson at the LHC. Taken from ref. [18]

gluon-gluon-fusion can only occur via a loop with massive particles where the highest contribution come from top-quarks as they are the heaviest particles in the SM.

2.3.2 Decay

Since the Higgs boson couples to all particles with mass and the coupling strength is proportional to the particle’s mass, the Higgs boson will most likely decay into the the heaviest possible particle. Figure 2.8 shows the branching ratios of different decay modes as a function of the Higgs mass. For the measured Higgs mass, it is most likely that the Higgs boson will decay into a pair of b-quarks. They are the heaviest particles that can be produced as a real pair in the Higgs decay, i.e. none of the b-quarks is a virtual particle. The decay into two top-quarks is strongly suppressed as the resonance for two real tops is at about three times the mass of the Higgs. Even though, the resonance for two real Z or $W^{\pm}$ bosons is also at a higher mass than the mass of the Higgs boson it is still more likely that the Higgs will decay into one real and one virtual of these bosons than into a top quark pair as their masses are about half as big. This thesis will focus on the $H \to \tau\tau$ decay mode. τ leptons are only the next to heaviest fermions, not considering the top-quark, but have the advantage of a much smaller QCD background compared to the decay into two b-quarks.

2.3.3 Discovery and Properties

The discovery of a new boson at a mass of $m \approx 125\,\text{GeV}$, observed by the CMS collaboration [1] and the ATLAS collaboration [2] in 2012, indicates that the Higgs mechanism is realised in nature. The new boson is compatible with a SM Higgs boson within all its observed properties. The SM Higgs is predicted to be neutral, $C\mathcal{P}$-even, scalar and spin zero stating $J^{PC} = 0^{++}$, see [4] chapter 11. An explanation of what $C\mathcal{P}$-even means can be found in section 2.4. These quantum numbers arise from a variety of measurements performed. Since the charge conjugation C is multiplicative the measurement of the decay channel $H \to \gamma\gamma$, $C_\gamma = -1$, immediately results in $C_H = 1$. It also follows that the new boson has to be neutral. This decay channel also excludes that the new boson has spin $J = 1$ or more generally odd

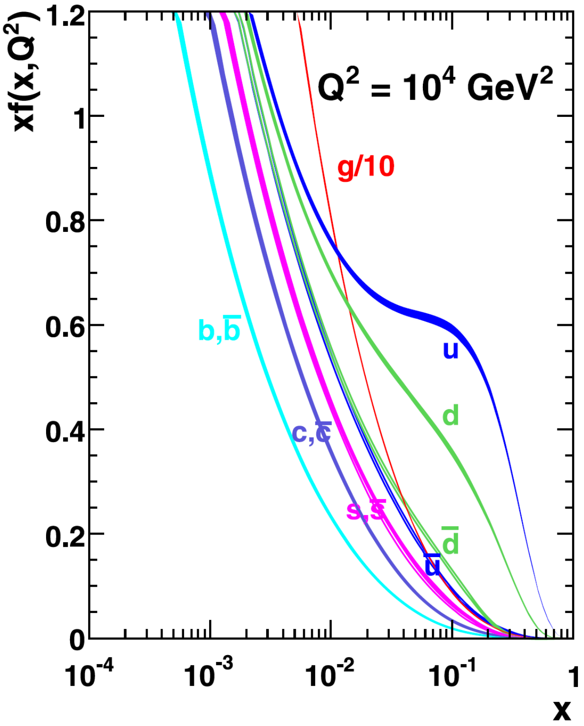


Figure 2.7: Parton distribution function for the proton at $Q^2 = 1 \times 10^4$ GeV^2. Taken from ref. [19]

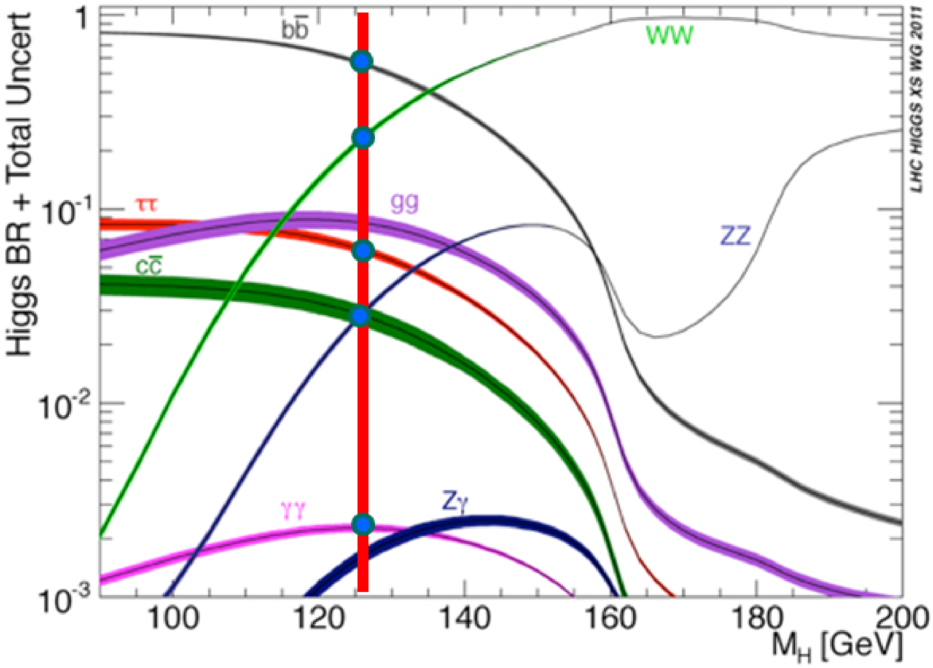


Figure 2.8: Possible decay modes for the Higgs boson. Taken from ref. [18]

spins with $P = -1$ due to the Landau-Yang-Theorem [20]. The spin zero hypothesis seems the most likely since hypotheses like 2^+ et cetera are excluded at high confidence levels via measurements of the production angles and decay angles in $H \to \gamma\gamma$, $H \to WW^* \to l\nu l\nu$ and $H \to ZZ^* \to 4l$, see [4] chapter 11. These measurements also exclude $J^P = 0^-$. Thus, the quantum numbers are compatible with the SM prediction $J^{PC} = 0^{++}$. However, this does not mean that the Higgs boson has to be purely $\mathcal{CP}$-even. A boson with Spin zero and $\mathcal{CP}$-even properties is called a scalar, while it would be called a pseudoscalar if it had $\mathcal{CP}$-odd properties and a vector boson if the spin was equal to one. A general $\mathcal{CP}$-violating scenario is not fully excluded at this point while a pure $\mathcal{CP}$-odd scenario is excluded in the $H \to ZZ^* \to 4l$ channel [21], for the method used see [22]. Additionally, $\mathcal{CP}$-violation in the Higgs coupling to fermions has a different theoretical description than in the coupling to gauge bosons [23] as the two particle families couple in different ways to the Higgs boson. This leaves the possibility for a higher $\mathcal{CP}$-odd contribution in the leptonic decay modes. A possible phenomenological Lagrangian for $\mathcal{CP}$-violation may look like the following: [24]

$$\mathcal{L}_Y = -g_\tau \left(\cos\phi_\tau \bar{\tau}\tau + \sin\phi_\tau \bar{\tau} i\gamma_5 \tau\right) h, \quad \phi_\tau \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right]. \tag{2.13}$$

The different $\mathcal{CP}$-scenarios are encoded in the mixing phase ϕ_τ. The different scenarios would have the following phases:

$$\text{scalar}: \quad \phi_\tau = 0 \tag{2.14}$$

$$\text{pseudoscalar}: \quad \phi_\tau = \tfrac{\pi}{2} \tag{2.15}$$

$$\text{maximal} \quad \mathcal{CP} - \text{violating}: \quad \phi_\tau = \pm\tfrac{\pi}{4}. \tag{2.16}$$

A possible $\mathcal{CP}$-violating scenario is realised in the Minimal Supersymmetric Model (MSSM), for further information on this please have a look at ref. [25].

2.4 Open Questions in the SM and CP-Violation

Although the Standard Model is an impressively successful theory there are some open questions that can not be answered with the SM alone. One of these unanswered questions is what dark matter and dark energy are. As only roughly 5% of the universe is visible and consists of the SM particles there has to be some other source of matter to explain the rotational behaviour of galaxies, called dark matter. Also the universe is expanding at an increasing speed and the energy/force driving this is called dark energy. More on these topics can be found in ref. [26] and ref. [27].

Another question left open in the SM is whether all forces can be combined into one Grand Unified Theory (GUT) and whether it is possible to also include gravity in this theoretical framework. A possible way to unify the three forces of the SM is a supersymmetric model. More on GUTs can be found in ref. [28].

A last example is the asymmetry between matter and antimatter that can be seen in the universe. The conditions that have to be fulfilled for the observed asymmetry are called Sakharov conditions and read: [29]

1. baryon number violation,
2. $C-$ and $\mathcal{CP}$-violation,
3. deviation from thermal equilibrium.

The measurement discussed in this thesis is a $\mathcal{CP}$ measurement. Thus, only the second condition will be further discussed. $\mathcal{CP}$ is the combination of parity $\mathcal{P}$ and charge conjugation $\mathcal{C}$. A process respects parity if it looks equal for a particle and its mirrored self. Similarly, it respects charge conjugation if it looks the same for a particle and its anti-particle. A $\mathcal{CP}$-violating process looks different for a particle and its mirrored antiparticle. If a process violates parity it still can conserve $\mathcal{CP}$.

The SM provides some $\mathcal{CP}$ violation for example in the decay of B mesons. In the SM the $\mathcal{CP}$ violation is proportional to $\sin 2\beta$. In 2001 the BaBar collaboration showed evidence for $\mathcal{CP}$ violation in the B^0 meson system with a non-zero value of $\sin 2\beta = 0.59 \pm 0.14(\text{stat}) \pm 0.05(\text{syst})$. [30] However, the strength of these processes alone is not sufficient to fulfill the Sakharov condition. Additional $\mathcal{CP}$ violation in the Higgs sector might be a step towards the realisation of said condition.

2.5 Theory of the measurement and Observables

The goal of the measurement studied in this thesis is to determine the $\mathcal{CP}$ quantum numbers of the found resonance at the LHC, or more precisely to determine whether the resonance is a pure $\mathcal{CP}$-even boson or an admixture of $\mathcal{CP}$-even and $\mathcal{CP}$-odd states which would be a direct evidence for new physics. How this measurement is motivated and performed will be discussed in the following.

First, one has to choose a decay channel of the Higgs boson to study these quantum numbers. Although, the decay $H \to b\bar{b}$ is the most abundant, see figure 2.8, it will not be used for the measurement due to a high and difficult to suppress background and depolarisation effects in the hadronisation of the quarks. [31] The depolarisation effects are introduced by the fragmentation of the quarks in pseudoscalar B mesons. [32] Instead, the $H \to \tau\tau$ channel will be used as it is a much cleaner channel with a still sufficiently high branching ratio.

The general idea is to define a model-independent observable which could separate between the different $\mathcal{CP}$-scenarios. Because the decay probability of the Higgs boson is sensitive to transverse spin correlations of the decay products: [31]

$$\Gamma\left(H/A^0 \to \tau^+\tau^-\right) \propto 1 - s_{\parallel}^{\tau^+} s_{\parallel}^{\tau^-} \pm s_{\perp}^{\tau^+} s_{\perp}^{\tau^-}, \tag{2.17}$$

it is possible to distinguish between a $\mathcal{CP}$-even state and $\mathcal{CP}$-odd state via angular correlations of the τ decay products. Equation (2.17) assumes a simplified scenario where there is no $\mathcal{CP}$-admixture possible, i.e. the Higgs couples either purely $\mathcal{CP}$-even or $\mathcal{CP}$-odd. The polarisation vectors given in that equation are given in their respective τ frame and denote the spin component parallel/perpendicular to the Higgs momentum in that frame. In the following this simplified scenario will be used to introduce the observable and then to generalise it for a possible $\mathcal{CP}$-mix scenario.

2.5.1 Observable to distinguish $\mathcal{CP}$-even and $\mathcal{CP}$-odd

The information of the decay probability can be used to write down a differential decay probability. For the τ decay studied in this thesis, $H \to \tau^+\tau^- \to \rho^+\bar{\nu}_\tau\rho^-\nu_\tau \to \pi^+\pi^0\bar{\nu}_\tau\pi^-\pi^0\nu_\tau$, the corresponding term reads: [31]

$$\frac{1}{\Gamma}\frac{\mathrm{d}\Gamma\left(H/A^0 \to \rho^+\bar{\nu}_\tau\rho^-\nu_\tau\right)}{\mathrm{d}\cos\theta^+\mathrm{d}\cos\theta^-\mathrm{d}\phi^*} = \frac{1}{8\pi}\left[1 + \frac{\left(m_\tau^2 - 2Q^2\right)^2}{\left(m_\tau^2 + 2Q^2\right)^2}\left[\cos\theta^-\cos\theta^+ \mp \sin\theta^+\sin\theta^-\cos\phi^*\right]\right], \tag{2.18}$$

where ϕ^* is the azimuthal angle between the decay planes and $\theta^\pm$ are the polar angles between the $\rho^\pm$

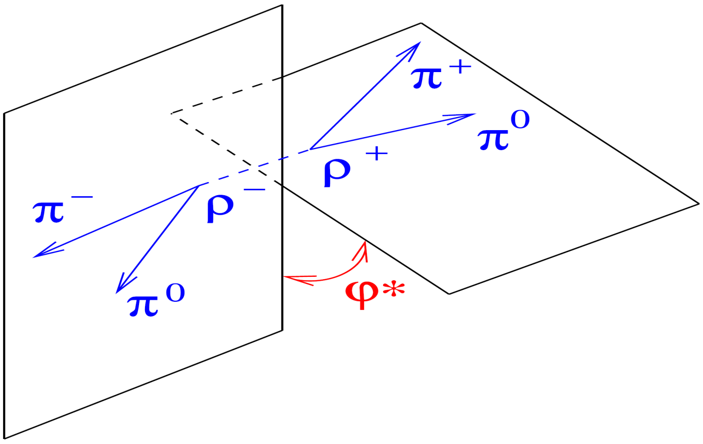


Figure 2.9: Definition of the angle φ^* in the $\rho\rho$ frame. Taken from ref. [31]

and the τ^- in the $\tau^\pm$ rest frames. This can also be done for other decay channels, e.g. $H \rightarrow \tau\tau \rightarrow \pi^+\bar{\nu_\tau}\pi^-\nu_\tau$. In contrast to this decay channel, the channel $H \rightarrow \tau\tau \rightarrow \pi^+\bar{\nu_\tau}\pi^-\nu_\tau$ does not show a term proportional to the hadronic system mass Q^2. This term leads to a degradation in the angular correlation term. The dependence on $\theta^\pm$ can be integrated to one factor. As the ρ mesons themselves decay, additional information can be gained by using the pion momenta in their decays. These additional information are encoded in the so-called polarimetric vector: [31]

$$h^i = \mathcal{N}\left(2\,(q\cdot N)\,q^i - q^2N^i\right), \tag{2.19}$$

where $\mathcal{N}$ is the normalisation, $q = p_{\pi^\pm} - p_{\pi^0}$ and N is the four-momentum of the neutrino and both of them are defined in their respective τ rest frames. If it is assumed that $m_{\pi^\pm} = m_{\pi^0}$ one can exploit the part that is not proportional to the neutrino momentum in the τ rest frame due to the relation: [31]

$$q\cdot N = (E_{\pi^\pm} - E_{\pi^0})\,m_\tau. \tag{2.20}$$

If the angle φ^* is defined as the angle between the decay planes in the $\rho\rho$ rest frame as shown in figure 2.9 and a separation into the two categories [31]

$$1: \quad y_0 y_1 > 0 \tag{2.21}$$

$$2: \quad y_0 y_1 < 0 \tag{2.22}$$

is applied where [31]

$$y_0 = \frac{E_{\pi^+} - E_{\pi^0}}{E_{\pi^+} + E_{\pi^0}} \tag{2.23}$$

$$y_1 = \frac{E_{\pi^-} - E_{\pi^0}}{E_{\pi^-} + E_{\pi^0}} \tag{2.24}$$

the $\mathcal{CP}$-even and $\mathcal{CP}$-odd scenarios can be separated via their φ^* distribution. The resulting distributions can be seen in figure 2.10. The thick lines denote the distribution for a $\mathcal{CP}$-even Higgs and the thin lines denote the distribution for a $\mathcal{CP}$-odd Higgs. Figure 2.10(a) shows the distributions for the first category,

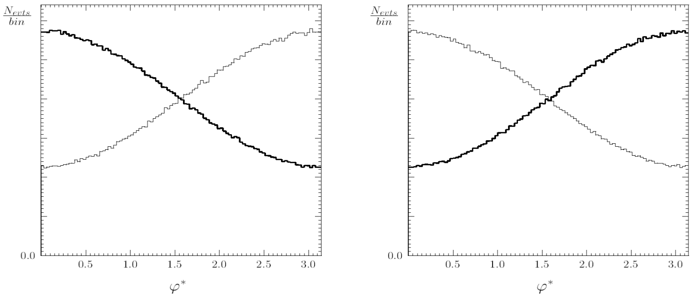


Figure 2.10: Distribution of φ^* for a $\mathcal{CP}$-even Higgs shown with the thick line and $\mathcal{CP}$-odd Higgs shown with the thin line. The left distribution shows the first category, i.e. $y_0 y_1 > 0$, and the right plot the second category, i.e. $y_0 y_1 < 0$. Taken from ref. [31]

i.e. $y_0 y_1 > 0$, and figure 2.10(b) for the second category, i.e. $y_0 y_1 < 0$.

2.5.2 Observable for the generalised possibility of a $\mathcal{CP}$-mixture

In the most general case, which also includes $\mathcal{CP}$-admixtures, the Yukawa interactions of the Higgs with τ's can be written down as: [33]

$$\mathcal{L}_y = -\frac{m_\tau}{v}\kappa_\tau \left(\cos\phi_\tau \bar{\tau}\tau + \sin\phi_\tau \bar{\tau} i\gamma_5 \tau\right) h, \quad \text{with} \quad \kappa_\tau > 0. \tag{2.25}$$

ϕ_τ is the parameter that encodes the $\mathcal{CP}$-mixture and is the one that is measured with the introduced method. The basic ideas behind this method follow the same scheme as in the simplified case in section 2.5.1. In contrast to the differential decay probability in the simplified scenario, it additionally contains information on the $\mathcal{CP}$-admixture in this general discussion: [33]

$$\mathrm{d}\Gamma_{h\to\tau^+\tau^-} \propto 1 - s_z^- s_z^+ + \cos(2\phi_\tau)\left(\mathbf{s}_T^- \cdot \mathbf{s}_T^+\right) + \sin(2\phi_\tau)\left[\left(\mathbf{s}_T^+ \times \mathbf{s}_T^-\right)\cdot \hat{\mathbf{k}}^-\right] \tag{2.26}$$

assuming $\beta_\tau = \sqrt{1 - \frac{4m_\tau^2}{m_h^2}} \approx 1$. $\hat{\mathbf{k}}^-$ is the normalised three-momentum of the τ^- in the Higgs rest frame and $\mathbf{s}$ are the normalised spin vectors defined in their respective τ frames. The transverse component of $\mathbf{s}$ and the z component are defined with respect to the $\hat{\mathbf{k}}^-$. [33] This equation shows that the angle ϕ_τ is encoded in the transverse spin correlations and can be measured via the angle between the planes spanned by the vectors $\hat{\mathbf{k}}^- \leftrightarrow \hat{\mathbf{s}}^-$ and $\hat{\mathbf{k}}^- \leftrightarrow \hat{\mathbf{s}}^+$. The measurement of this angle is performed as follows. First, all four-momenta of all pions, i.e. π^+, π^- and the two π^0s, are boosted in the $\pi^+\pi^-$ rest frame which defines the common axis for both planes. This definition of rest frame is chosen to be consistent with the method for the τ decay to one charged pion and zero neutral pions such that both measurements could be combined in the end. By doing so, the distributions of the final observable for each individual decay channel can be combined into one distribution by just adding them. The angle between the two decay planes can then be measured by: [33]

$$\varphi^* = \arccos\left(\hat{\mathbf{q}}_\perp^{*0+} \cdot \hat{\mathbf{q}}_\perp^{*0-}\right), \quad 0 \leq \varphi^* \leq \pi \tag{2.27}$$

where $\hat{\mathbf{q}}_\perp^{*0\pm}$ are the transverse components of the normalised three-momenta of the neutral pions with respect to their associated charged pions in the $\pi^+\pi^-$ rest frame. Figure 2.11 shows a schematic view of the angle φ^*.

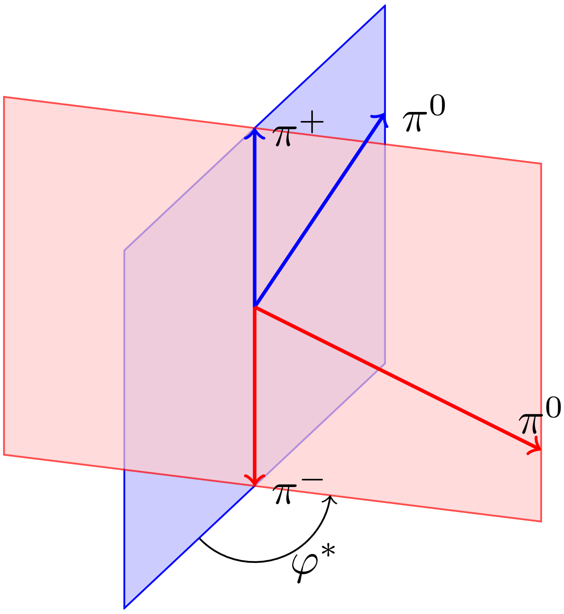


Figure 2.11: Schematic view of the definition of the angle φ^* in the $\pi^+\pi^-$ rest frame.

To be sensitive to $\mathcal{CP}$-odd contributions in the angular distribution, it is necessary to include the $\mathcal{CP}$-odd triple correlation which is defined by: [33]

$$\mathcal{O}^* = \hat{\mathbf{q}}^{*-} \cdot \left(\hat{\mathbf{q}}_\perp^{*0+} \times \hat{\mathbf{q}}_\perp^{*0-}\right), \quad -1 \leq \mathcal{O}^* \leq +1. \tag{2.28}$$

Based on this, the signed angle variable is defined as: [33]

$$\varphi_{CP}^* = \begin{cases} \varphi^*, & \text{if } \mathcal{O}^* \geq 0 \\ 2\pi - \varphi^*, & \text{if } \mathcal{O}^* < 0 \end{cases}. \tag{2.29}$$

As it was shown for the simplified case in section 2.5.1 the polarimetric vector can be used to increase the discrimination of $\mathcal{CP}$-even versus $\mathcal{CP}$-odd. In this general scenario the polarimetric vector has to be used to obtain a non-flat φ_{CP}^*-distribution. Following the notation in ref. [33] the differential decay probability when including possible $\mathcal{CP}$-violation for the τ decays reads: [33]

$$\Gamma_a^{-1} \mathrm{d}\Gamma_a \left(\tau^\mp \left(\hat{\mathbf{s}}^\mp\right) \to a^\mp \left(q^\mp\right) + X\right) = n\left(E_\mp\right)\left[1 \pm b\left(E_\mp\right) \hat{\mathbf{s}}^\mp \cdot \hat{\mathbf{q}}^\mp\right] \mathrm{d}E_\mp \frac{\mathrm{d}\Omega_\mp}{4\pi}, \tag{2.30}$$

where $a^\mp \in e^\mp, \mu^\mp, \pi^\mp, a_1^{L,T,\mp}$. However, this thesis only focusses on the decay chain $H \to \tau^+\tau^- \to \rho^+\bar{\nu}_\tau\rho^-\nu_\tau \to \pi^+\pi^0\bar{\nu}_\tau\pi^-\pi^0\nu_\tau$. The normalised spin vectors are denoted by $\hat{\mathbf{s}}$, the energies by E and the normalised three-momenta by $\hat{\mathbf{q}}$. All of them are defined in their respective τ frame. n and b are spectral functions which depend on the energy of the charged pions. The spectral function or spin analyzer function b carries similar information to the polarimetric vector defined in equation (2.20) in the simplified case. Thus, it will also lead to a categorisation of the angular distribution into two categories. This can also be motivated by the functional dependence of the spin analyzer b on the charged pion energy, see figure 2.12. In particular the sign-flip is the reason why the categorisation is needed.

Just like the polarimetric vector before the spin analyzer is also proportional to $E_{\pi^\pm} - E_{\pi^0}$ [33] which

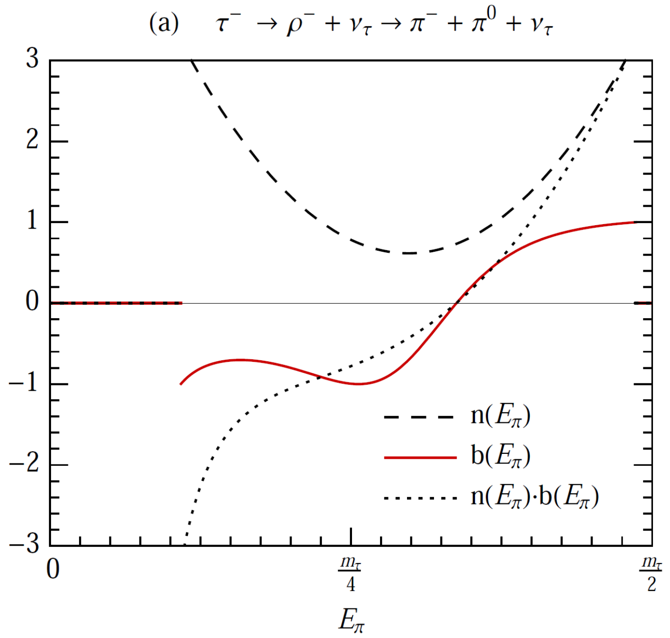


Figure 2.12: Functional dependence of the spin analyzer function b on the charged pion energy in the τ frame. Taken from ref. [34]

is why the same categorisation is used with a slightly different notation: [33]

$$1: \quad y_0 y_1 > 0 \tag{2.31}$$

$$2: \quad y_0 y_1 < 0 \tag{2.32}$$

where [33]

$$y_0 = \frac{E_{\pi^+} - E_{\pi^0}}{E_{\pi^+} + E_{\pi^0}} \tag{2.33}$$

$$y_1 = \frac{E_{\pi^-} - E_{\pi^0}}{E_{\pi^-} + E_{\pi^0}}. \tag{2.34}$$

If $y_0 y_1 < 0$ φ^*_{CP} is replaced by $\varphi^*_{CP} + \pi$ else it stays the same. The final φ^*_{CP} distribution can be seen in figure 2.13 for a $\mathcal{CP}$-even Higgs, a $\mathcal{CP}$-odd Higgs and for the Z boson. As the Z boson is the largest background in a $H \to \tau\tau$ analysis it is fortunate for this measurement that the φ^*_{CP} distribution for the aforementioned boson is flat modulo statistical fluctuations.

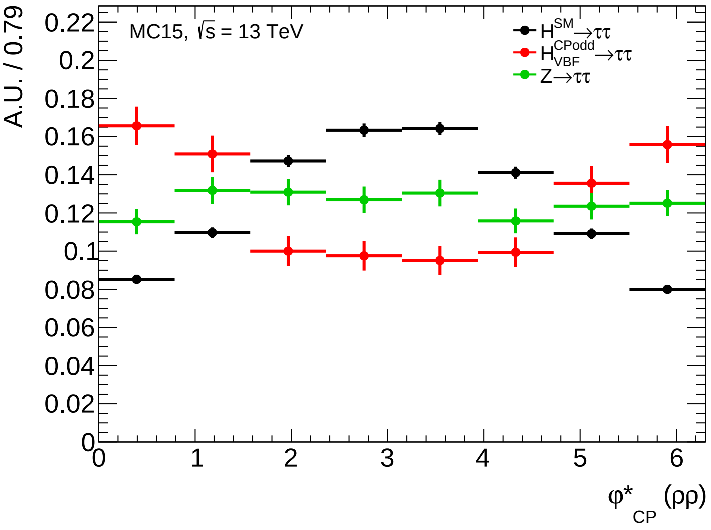


Figure 2.13: φ^{*}_{CP} distribution for a $\mathcal{CP}$-even Higgs boson, a $\mathcal{CP}$-odd Higgs boson and the Z boson.

CHAPTER 3

The Experiment

The studies performed in this thesis are based on simulated data of the ATLAS experiment at the LHC. This means that the simulated data considers characteristics of the LHC beam and of the ATLAS experiment itself. Examples for such characteristics for the LHC beam are pile-up and the centre-of-mass-energy, whereas for the ATLAS detector examples are dead detector regions and detector resolutions. In the following section both machines are shortly introduced and the event reconstruction is explained.

3.1 LHC

The Large Hadron Collider, short LHC, is situated near Geneva at the Swiss-French-Border. It is currently the world's largest circular hadron collider with a circumference of about 27 km and the whole ring is built about 100 m underground. [35] The LHC is operated by the Conseil Européen pour la Recherche Nucléaire, CERN. It is designed to deliver an instantaneous luminosity of about $\mathcal{L} = 1 \times 10^{-34}\,\mathrm{cm}^{-2}\mathrm{s}^{-2}$ and to reach centre-of-mass energies of $\sqrt{s} = 14\,\mathrm{TeV}$ at a bunch crossing rate of 25 ns. [36] The instantaneous luminosity delivered by the LHC can be expressed in terms of machine parameters: [37]

$$\mathcal{L} = \frac{n_b f_r n_1 n_2}{2\pi \Sigma_X \Sigma_Y} \tag{3.1}$$

where n_b is the number of bunches colliding per cycle, f_r the revolution frequency, n_1 and n_2 the number of protons in the bunches of each beam, and Σ_X and Σ_Y are the convolved horizontal and vertical beam widths. The total delivered and recorded data is not given in numbers of events but rather as the integrated luminosity $L = \int \mathcal{L} \mathrm{d}\,t$. Figure 3.1 shows the integrated luminosity delivered by the LHC and the recorded luminosity by ATLAS in 2016.

ATLAS is one of the four big experiments at the LHC and will be further introduced in the next section. The LHC is conceptualised to perform precision measurements of the SM and to search for new physics. Its first goal, the discovery of the Higgs boson, was reached in 2012. As the tasks of this machine are so heterogeneous there are many possible processes that can occur in a collision. The known processes and their cross-sections are shown in figure 3.2 as a function of the centre-of-mass energy $\sqrt{s}$.

As it can be seen the cross-sections for processes that involve the Higgs boson grows more steeply than the cross-section for Z boson production with increasing $\sqrt{s}$. This is of advantage for the method discussed in this thesis as the Z boson is a major background processes.

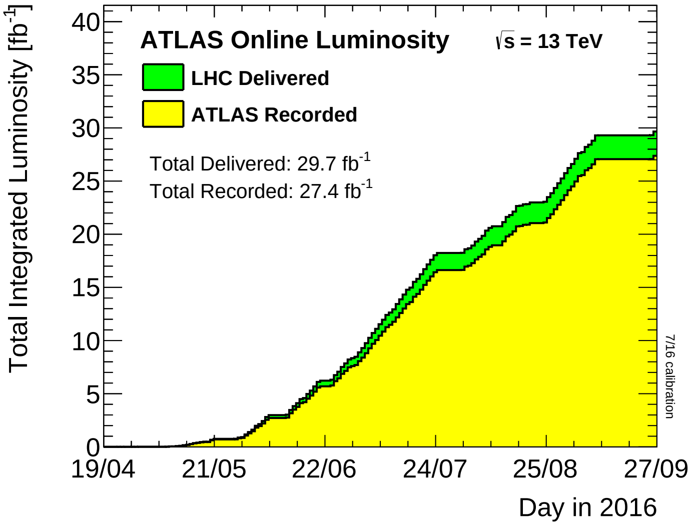


Figure 3.1: Delivered integrated luminosity by the LHC and recorded integrated luminosity by ATLAS in 2016. Taken from ref. [38]

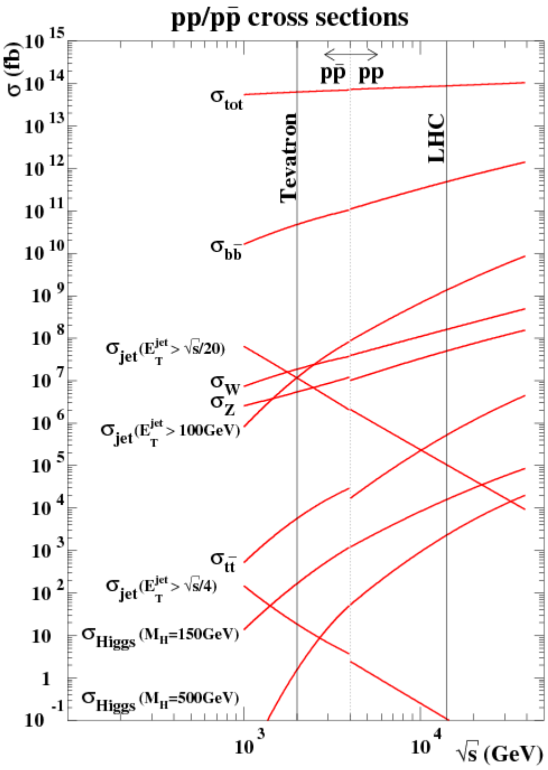


Figure 3.2: Cross-sections for different processes at the LHC as a function of the centre-of-mass energy $\sqrt{s}$. Taken from ref. [39]

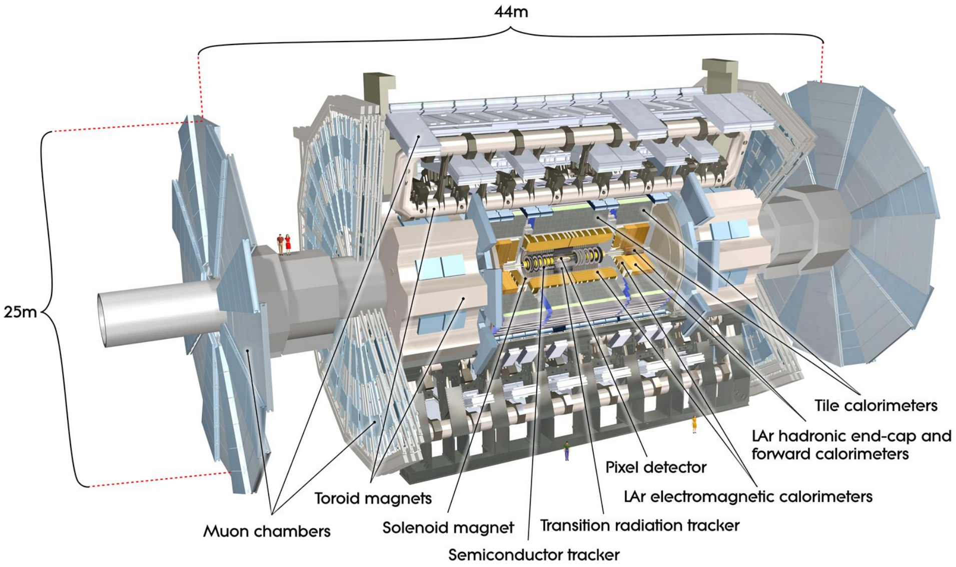


Figure 3.3: View on the ATLAS experiment with all its components labelled. Taken from ref. [40]

3.2 ATLAS

The ATLAS experiment, **A** **T**oroidal **LHC** **A**pparatu**S**, is one of the two multipurpose detectors at the LHC, the other being CMS. The experiment is approximately 25 m high and 44 m long and weighs about 7 000 t. [40] Figure 3.3 shows the detector and its components which will be further described in the following. A full description can be found in ref. [40] and the technical design report [41] on which the subsections on the detector components are based.

The instantaneous luminosity recorded by ATLAS is given by: [37]

$$\mathcal{L} = \frac{\mu_{vis} n_b f_r}{\sigma_{vis}} = \frac{\epsilon \mu n_b f_r}{\epsilon \sigma} \tag{3.2}$$

where ϵ is the efficiency of the ATLAS detector and the event reconstruction algorithm, μ is the average number of interactions per bunch crossing and σ is the total inelastic scattering cross section. Hence, one can determine the efficiency of a detector and algorithm by measuring the visible cross section of inelastic scattering processes.

3.2.1 Tracking System

In the tracking system the trajectories of charged particles are recorded. This part of the detector is inside of a 2 T magnetic field which bends these trajectories and makes it possible to measure the particle momenta via their curvature. The energy resolution of this measurement depends on the momentum itself and is given by: [40]

$$\frac{\sigma_{p_\mathrm{T}}}{p_\mathrm{T}} = 0.05\% p_\mathrm{T} \oplus 1\%. \tag{3.3}$$

In total the tracking system is build out of three components: the first component close to the beam pipe is the silicon pixel detector. This detector is highly granulated to allow for a precise vertex reconstruction. It is composed of three layers which are all segmented in $R - \phi$ and z with a minimum pixel size of

$50 \times 400\,\mu\text{m}^2$. This component and the next surrounding component, the SCT, both cover a range of $\eta < 2.5$. Since the radius of the beam pipe is reduced for Run 2 the pixel detector is upgraded for Run 2 with an insertable barrel layer, short IBL, which serves as an additional pixel detector layer between the old detector component and the new beam pipe. [42] In the end cap regions the same modules are used as in the barrel regions. The difference is that they are aligned as five discs on each side to cover the range $1.7 < \eta < 2.5$.

The semiconductor tracker (SCT) is the next layer and is build as a silicon microstrip tracker. It consists of eight layers with a set of strips in each layer aligned parallel to the beam axis. Each strip measures the $R - \phi$ coordinate with a precision of 17 µm and 580 µm in the z-direction from the position of the layer. The SCT continues in the end-cap regions with strips aligned in circles around the beam pipe.

The third and last layer is the transition radiation tracker, TRT. This detector is a gaseous detector with straw tubes that measure the trajectories of particles with the principle of a drift chamber. Transition radiation of bypassing particles can be used for electron identification. The TRT covers a range of $\eta < 2.0$ and can only provide $R - \phi$ information with a precision of 130 µm. Just like the SCT, the TRT is radially arranged in the end cap region.

3.2.2 Calorimeter System

The calorimeter system covers a range of $\eta < 4.9$ and is composed of different technologies over this range. The electromagnetic calorimeter, EMcal, is divided into a barrel region and the end-cap regions. For both the underlying technique is the same, i.e. both are sampling calorimeters with lead as absorber material and liquid argon as active material. Radially, the EMcal consists of different layers with different granularities. The first layer has the finest granularity in order to separate nearby showers. The calorimeter's total thickness in the barrel region is > 22 radiation lengths X_0 and $> 24X_0$ in the end-cap regions to prevent punch-through in the muon system. The radiation length X_0 is defined as the thickness of a material an electron has to pass to reduce the energy of the electron by the factor $\frac{1}{e}$. [43] In order to be able to reconstruct the complete ϕ-range the calorimeter's active material is shaped like an accordion. In the barrel and end-cap region the EM calorimeter reaches an energy resolution of: [40]

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 0.7\%. \tag{3.4}$$

The hadronic calorimeter, Hcal, consists of three parts. The tile calorimeter is built directly outside the EMcal. It is also a sampling calorimeter with steel as the absorber material and scintillating tiles, which are read out by photomultipliers, as the active material. The liquid argon hadronic end-cap calorimeter lies directly behind the EMcal end-caps and slightly overlaps with the tile calorimeter and the forward calorimeter to avoid a drop in material density in the transition regions from one part of the calorimeter to another part. This sampling calorimeter uses copper as the absorber material and liquid argon as the active material. The forward calorimeter is situated inside the end-caps around the beam pipe. It consists of a first layer where the absorber material is copper which is designed to be sensitive to electromagnetic processes and a second and third layer with tungsten as absorbing material for strong process measurements. The active material again is liquid argon inside a tube. Just like the EM calorimeter's thickness can be described in radiation lengths, the Hcal's thickness can be described in interaction lengths. The interaction length is defined as the mean free path a particle can travel before undergoing an inelastic scattering process. [44] The complete forward calorimeter has material of about 10 interaction lengths. The hadronic calorimeter reaches an energy resolution in the barrel and end-cap region of: [40]

$$\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\% \tag{3.5}$$

and in the forward region of: [40]

$$\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{E}} \oplus 10\%. \tag{3.6}$$

3.2.3 Muon System

The muon system consists of drift tubes and resistive-plate chambers [41]. Muons are the only particles that can reach this part of the detector, except for neutrinos, since the thicknesses of the calorimeters are constructed so that the probability of punch-through pions is low. The muons are bent by a magnetic field caused by toroidal magnets and their tracks are measured. The magnetic field lines are perpendicular to the beam pipe as opposed to the field lines caused by the solenoid magnet system which point along the beam axis. The muon system reaches a total resolution of: [40]

$$\frac{\sigma_{p_\mathrm{T}}}{p_\mathrm{T}} = 10\% \quad \text{at} \quad 1\,\mathrm{TeV}. \tag{3.7}$$

3.2.4 Coordinate System

Hadron colliders show some characteristics which are represented in the variables and the coordinate system. Since the colliding objects, at the LHC these are protons, are not point-like but rather are composite objects and only their partons partake in the actual collision one has to consider their parton distribution functions. Figure 2.7 in section 2.3.1 shows such PDFs for the proton. As it can be seen the actual colliding particles carry statistically distributed fractions of the protons energy which makes it impossible to know the exact initial energy of these colliding particles. Furthermore, in most cases the produced particles are boosted in one direction along the beam axis since it is most likely that the energies of the colliding partons are not balanced.

Since the initial energy along the beam axis is not known it is common to use only the transverse momentum which is well defined as: [14]

$$p_\mathrm{T} = \sqrt{p_X^2 + p_Y^2}. \tag{3.8}$$

Also, the boost of the objects along the beam axis is one reason that a quantity named rapidity is introduced: [14]

$$y = \frac{1}{2}\ln\left(\frac{E + p_Z}{E - p_Z}\right). \tag{3.9}$$

This can further be simplified and directly related to the polar angle if it is assumed that the energy of the produced particle is large compared to its mass. This new quantity is called pseudorapidity: [14]

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right). \tag{3.10}$$

Hence, the four-momenta of the measured particles are typically given as $p = (p_\mathrm{T}, \eta, \phi, M)^\mathrm{T}$. The distance between two particles is given in a measure ΔR that combines the azimuthal angle ϕ and the

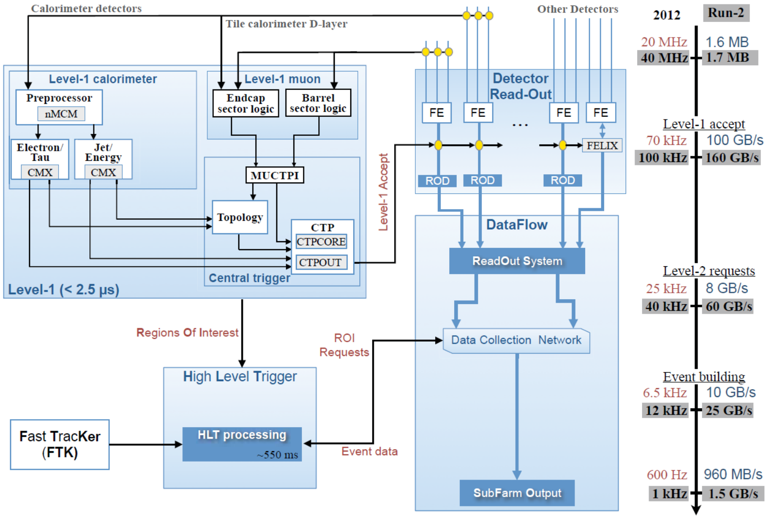


Figure 3.4: Schematic overview of the ATLAS trigger system and data acquisition. Taken from ref. [45]

advantages of the pseudorapidity η:

$$\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}. \tag{3.11}$$

The nominal interaction point has the coordinates $(0, 0, 0)^T$. The x-axis points to the centre of the LHC ring, the y-axis points upwards and the z-axis goes along the beam pipe.

3.2.5 Trigger system

As the bunch spacing in Run 2 is designed to be 25 ns the data taking rate would be 40 MHz if all events were recorded. Most of these events are not of major interest, for example events where only jets are produced in SM processes. Additionally, this amount of data would be impossible to handle and store. Hence, the ATLAS detector uses a trigger system which makes fast decisions whether an event is of interest or not. A more detailed discussion of the upgrades done with respect to Run 1 can be found in ref. [45]. The main challenge was to optimise the trigger system such that it can handle a theoretical increase of roughly a factor of 5 in the trigger rates as this would lead to too much data output. To be able to handle the event rate of 40 MHz a level one, L1, trigger stage is used at hardware level that finds regions of interest and reduces the data rate to about 100 kHz. On the L1 trigger stage a new topological trigger was introduced for Run 2. The new topological trigger is capable of using topological quantities like the angular separation of trigger objects to perform trigger decisions. [46] The old L1 trigger was only able to access p_T or E_T for the trigger decision. A second trigger stage, the high level trigger (HLT), is software based and combines the two HLT stages of Run 1 for Run 2. This trigger stage reduces the data rate to approximately 1 kHz. The events that pass this HLT stage are the events that are saved on tape. There are different triggers used for different physics requirements of different analyses. For example the $H \to \tau\tau$ coupling analysis performed in Run 1 used a number of different triggers. Among others they used a single electron, single muon and di-τ_{had} trigger where τ_{had} are τ leptons that decay hadronically. [5]

CHAPTER 4

Object Reconstruction and Datasets

This section aims to give an overview of the particle reconstruction used in ATLAS. Afterwards, the datasets that are used in this thesis are shortly introduced.

4.1 Event Reconstruction

As different particles leave different signatures in the ATLAS detector it is possible to reconstruct different types of particles from the detector's response, i.e. tracks and calorimeter clusters. Due to a variety of reasons it is a complex field to reconstruct these particles, one of them being that different particles can leave quite similar signatures, for example jets and taus. In the following the basic concepts of the particle reconstruction will be discussed based on ref. [47] if not stated otherwise.

4.1.1 Electrons and Photons

For the electron there are two different algorithms, one that begins with information from the electromagnetic calorimeter, EMcal, and one that begins with track information. The standard algorithm is the one seeded by the EMcal. First, an EMcal cluster with $p_{\mathrm{T}} > 2.5\,\mathrm{GeV}$ [48] in a tower is searched and then matched to tracks which do not belong to photon conversions. The matched track then has to lie in the same direction as the cluster within a window of $\Delta\eta \times \Delta\phi = 0.05 \times 0.1$ [48] on the side where the track gets bent by the magnetic field and $\Delta\eta \times \Delta\phi = 0.05 \times 0.05$ [48] on the other side. If more than one track is matched, tracks with hits in the silicon tracker and with the smallest ΔR are chosen. Electrons in the central tracking region $\eta < 2.5$ are assigned with the energy from the calorimeter and η and ϕ from the best matched track at the vertex. In the forward region, namely $2.5 < \eta < 4.9$, there are no tracking information available which is why the electrons are solely reconstructed from the calorimeter and only if they leave a small amount of energy in the hadronic calorimeter. For the analysers there are five different possible likelihood-based working points `VeryLoose`, `Loose`, `Loose+B-Layer`, `Medium` and `Tight`, each of which has to fulfill tighter criteria subsequently. The working points identify electrons based on different criteria one of them being the ratio of the measured energy in the calorimeter and the momentum measured with the track $\frac{E}{p}$ for the `Tight` working point. In general the variables that are considered for the identification are track quality related, shower shape related or calorimeter related. The reconstruction efficiency is strongly dependent on the η region since the tracking is different in these regions.

As electrons and photons both deposit their energy in the EMcal and it is possible for both to leave tracks in the detector their reconstruction is performed simultaneously. While electrons always leave a track in the detector, photons only do so if they interact with the detector material and create an electron-positron pair, called photon conversion. Just like the electron algorithm the photon algorithm starts with the search for an EMcal cluster with $p_\mathrm{T} > 2.5\,\mathrm{GeV}$ [49] and performs a loose matching of tracks with these clusters in order to be able to reconstruct photon conversions. In this case the tracks have to lie in a window of $\Delta\eta \times \Delta\phi = 0.05 \times 0.05$ in the direction of the bending of the track and $\Delta\eta \times \Delta\phi = 0.05 \times 0.2$ in the opposite direction. Finally, an algorithm decides whether the cluster originates from an electron, an unconverted photon or a converted photon, for more on this please see ref. [49]. The photon identification only has two working points `Loose` and `Tight`, respectively. The main difference between these two is that the `Tight` working point considers information from the EMcal strip layer. Again, the efficiency is dependent on η due to the calorimeter geometry and the resulting different shower shapes.

4.1.2 Muons

For the reconstruction of muons, information of the inner detector, the calorimeter and the muon system are used. First, track candidates of muons are reconstructed independently in the inner detector and different segments, or layers, of the muon system. These tracks are then combined to a muon object for the analyses in different algorithms. A combined muon (CB) is first reconstructed in the muon system and its track is extrapolated to the inner detector to form a combined track within a fitting procedure. A segment-tagged muon (ST) typically has low p_T or is in a region of small acceptance in the muon system. These muons are first reconstructed in the inner detector and then tagged as a muon if a matching segment in the muon system is found. Calorimeter-tagged muons (CT) have the lowest purity as these muons are solely reconstructed from the inner detector and energy deposits in the calorimeters compatible with a minimum-ionizing particle. These muons are mostly reconstructed in the region $|\eta| < 0.1$ since there the most material needed for operation, i.e. dead material, is found. Their reconstruction is optimized for a range of $15\,\mathrm{GeV} < p_\mathrm{T} < 100\,\mathrm{GeV}$. The last type of muons, extrapolated muons (ME), are reconstructed in the region $2.5 < \eta < 2.7$ where there are no information from the inner detector available which is why these muons are reconstructed from the interaction point and the muon system. If there is an overlap between different muon types the following priority order is used:

1. CB muons
2. ST muons
3. CT muons.

If there is an overlap with a ME muon the muon with the larger number of hits and better fit quality is chosen. [50]

4.1.3 Jets

As discussed in section 2.1.2 quarks do not occur freely in nature and thus undergo ”hadronisation” processes at ATLAS. The result of this process is a collimated number of hadronisation products, for example pions. All hadronisation products grouped together are called a jet. Jets at ATLAS are typically reconstructed using the anti-k_t jet clustering algorithm, see ref. [51]. For a more detailed introduction to the clustering algorithm please refer to ref. [52]. The algorithm starts by building three-dimensional topological clusters of energy deposits in the calorimeter from a seed cell. Using a specific distance measure, see ref. [52], the neighbouring cells are added into the cluster. All of the cells clustered in the algorithm have to have higher energies than expected from noise. The cluster can split into subclusters when cells are found that pass the criteria for a seed cell from which the algorithm is initialised again.

Figure 4.1: Idealised showers of a τ and a hadronic jet to emphasize the differences between them. Taken from ref. [54]

Typically, these clusters are formed in a radius $\Delta R < 0.4$ for the anti-k_t algorithm. The jet clustering algorithm has to take account for possible pile-up in the event which would falsely increase the measured jet energy. For more information on pile-up, see section 4.2.2.

4.1.4 Taus

As mentioned in section 2.1.3 the τ lepton has a very short lifetime, thus only its decay products can be seen in the ATLAS detector. As stated in table 2.1 the τ decays in approximately 35% of all events leptonically and in the other 65% hadronically. The reconstruction of leptonic τ decays is complicated as there are two neutrinos in the decay and a lepton, i.e. an electron or muon, which is difficult to distinguish from a direct lepton production. Hence, the reconstruction of these τ's is in fact the reconstruction algorithm of electrons or muons, respectively. Thus, some analyses choose a di-τ channel where one τ decays hadronically and the other leptonically and the latter decay is used as a tag. In the case where both τ's decay leptonically one can choose the decay into one electron and one muon which greatly reduces the background from Drell-Yan-Processes, one example of such an analysis can be found in ref. [53].

The algorithm that reconstructs hadronic τ's starts with all jets reconstructed as described in section 4.1.3 and then performs a test whether the jet is actually a τ or not, based on a Boosted Decision Tree (BDT) and finally calculates its four-momentum. Typically, jets consist of more hadronisation products than there are τ lepton decay products. Furthermore, the decay products of the τ are more collimated than the jet's hadronisation products. Figure 4.1 shows a schematic view of these differences in the shower shape which is one of the inputs for the Boosted Decision Tree (BDT). There are two ways of reconstructing the full τ. On the one hand TauRec which was used in Run 1 and does not provide any information on the individual decay products of the τ and on the other hand the Substructure Reconstruction which can also identify the hadronic decay products of the τ, i.e. also the decay mode, and their four momenta. Both will be shortly introduced in the next subsections.

TauRec

A detailed description of TauRec can be found in ref. [55]. The algorithm starts by selecting all jets reconstructed with the anti-k_t algorithm with a distance parameter of $R = 0.4$ as possible τ_{had} candidates. These seed jets have to fulfill the criteria $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ and $\eta < 2.5$. The latter criterion ensures that the τ_{had} lies in the tracking region since only 1- or 3-prong τ_{had}'s are considered, i.e. τ_{had}'s with one or three tracks. The four momentum is calculated from η and ϕ of the topological clusters of the seed jets where each cluster is assumed to be massless. Similarly, the transverse momentum is calculated from the seed jets within a cone of $\Delta R < 0.2$ around the initial direction to reduce the effect of pileup in the calculation. As the τ_{had} candidates are assigned no mass the transverse momentum and the transverse energy $E_{\mathrm{T}} = E\cos\theta$ are identical. In the final energy scale calibration the difference in the composition

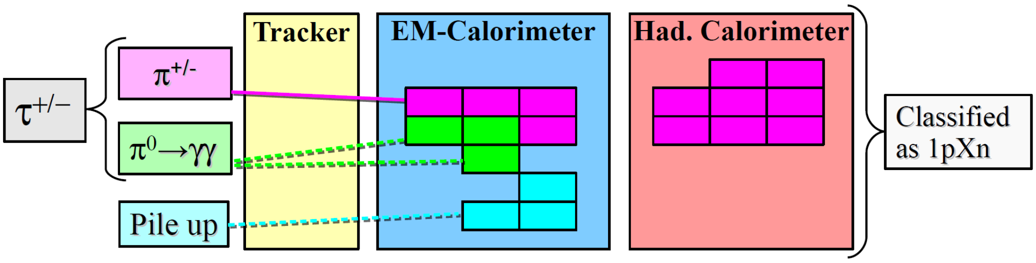


Figure 4.2: Schematic 1p1n decay of a τ identified as a 1pXn decay. Kinematic variables are used to distinguish them. Taken from ref. [57]

of the τ_{had} jets compared to QCD jets is taken into account by an additional correction term of the energy scale for τ_{had} leptons. The correction term is derived from simulated data. Tracks within $\Delta R < 0.2$ are associated to the τ_{had} candidate and must also pass selection criteria that ensure good track quality. Tracks within $0.2 < \Delta R < 0.4$ are kept for discrimination against jets and electrons. This discrimination is done by a BDT which takes for example the shower shape and the number of tracks into account.

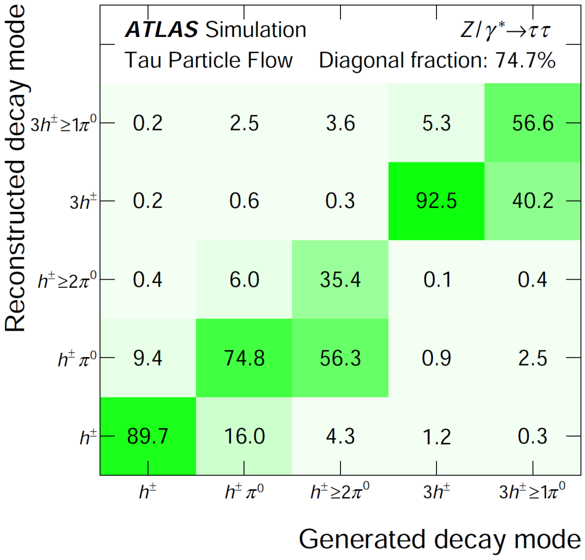
Substructure Reconstruction

The idea of the Substructure Reconstruction is to use the components of the detector which give the best resolution for each decay product of the τ_{had}'s, i.e. the tracking system for the charged pions and the EM calorimeter for the neutral pions. It is aimed to improve the resolution for τ_{had}'s in the a range of $15\,\text{GeV} < p_{\text{T}} < 100\,\text{GeV}$ which is the relevant range for electroweak processes and studies with the SM Higgs boson. More details on the Substructure Reconstruction can be found in ref. [56] and ref. [57].

Just like TauRec, this algorithm takes the jets defined in section 4.1.4 plus the additional requirement of $1 \leq N_{trk} \leq 5$ as possible τ_{had} candidates. It is assumed that hadronic τ's decay almost exclusively into charged pions and neutral pions. First, the charged pions are reconstructed from the tracks and their mass is set to $m_{\pi^\pm}$ due to the just mentioned assumption. The next step is to subtract their energies from the EM energy depositions such that in an ideal scenario only the energy of the neutral pions is left: [56]

$$E_{h^\pm}^{EM} = E_{h^\pm}^{trk} - E_{h^\pm}^{HAD}. \tag{4.1}$$

This energy is subtracted from the closest neutral pion candidate within $\Delta R < 0.04$. Noise and pile-up is removed from the π^0 candidates by using minimum requirements on kinematic variables. In order to avoid classifying $\pi^\pm$ remnants from the subtraction as π^0, a BDT is used that exploits for example the different shower shapes of $\pi^\pm$ and π^0. Neutral pions show a more regular shower shape. For the decay mode classification it is important to reconstruct each individual π^0. Neutral pions almost always decay into two photons which are individually identified in the first layer of the EM calorimeter. With increasing p_{T} of the τ_{had} it gets harder to identify each individual π^0 as the decay products are more collimated. After all charged pions and neutral pions have been identified a first estimate of the decay mode and the visible four momentum of the τ can be given. In a last step kinematic variables are used in a separate BDT which always tests two decay mode hypotheses, of the five possible decay modes in the Substructure Reconstruction, against each other. This, then yields the final decay mode classification and the visible four momentum of the τ_{had}. One example for a misclassified decay mode can be seen in figure 4.2 where pile-up is identified as a π^0. The decay mode efficiency matrix for the five way decay



	$h^\pm$	$h^\pm\pi^0$	$h^\pm\geq2\pi^0$	$3h^\pm$	$3h^\pm\geq1\pi^0$
$3h^\pm\geq1\pi^0$	0.2	2.5	3.6	5.3	56.6
$3h^\pm$	0.2	0.6	0.3	92.5	40.2
$h^\pm\geq2\pi^0$	0.4	6.0	35.4	0.1	0.4
$h^\pm\pi^0$	9.4	74.8	56.3	0.9	2.5
$h^\pm$	89.7	16.0	4.3	1.2	0.3

Figure 4.3: Matrix of the decay mode classification efficiency. Taken from ref. [56]

mode classification is shown in figure 4.3.

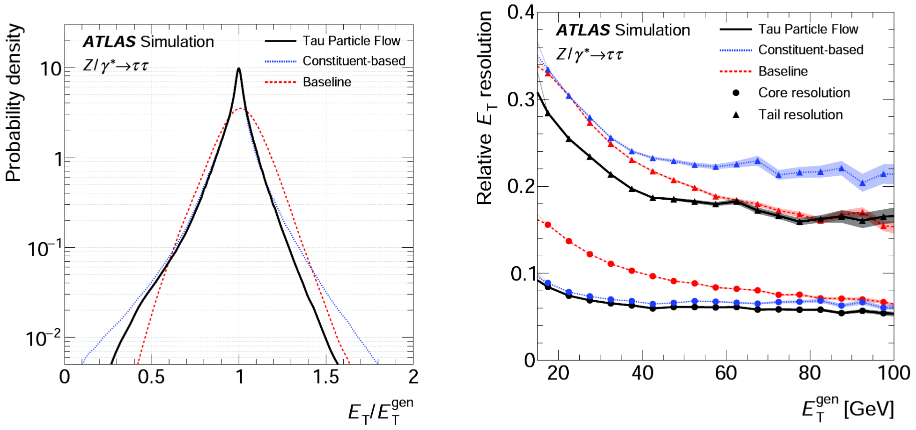
Generally, the more neutral pions there are, the less efficient the τ identification gets. This is due to decay modes with a larger number of neutral pions being broader and looking more like QCD jets. For $p_\mathrm{T} = 20\,\mathrm{GeV}$ the Substructure Reconstruction achieves a core energy resolution of 8% while TauRec has a core energy resolution of 15% as it can be seen in figure 4.4(b). As the Substructure Reconstruction performance is better than the TauRec performance in the core regions of the resolution and TauRec is better in the tails for high p_T, a combination of both is used in this thesis. The idea behind the combination is to use the reconstruction that works better in a specific region. To do so, both reconstructions are weighted with their inverse-square of their p_T-dependent core p_T resolutions. [56] If the two reconstructed transverse momenta differ by more than five times their combined core resolutions the TauRec value is used. [56] In the following this combination will simply be called Substructure Reconstruction. Figure 4.4 shows the advantage of the combination of both reconstructions. The core resolution is greatly improved compared to TauRec while the tail resolutions are also under control.

4.1.5 $E_\mathrm{T}^\mathrm{miss}$

The reconstruction performance of the missing transverse energy, $E_\mathrm{T}^\mathrm{miss}$, depends on all other particle reconstruction performances. Causes for missing transverse momentum are for example neutrinos that leave the detector unseen. Its definition reflects its physical meaning: [58]

$$E_{x(y)}^{\mathrm{miss}} = E_{x(y)}^{\mathrm{miss},e} + E_{x(y)}^{\mathrm{miss},\gamma} + E_{x(y)}^{\mathrm{miss},\tau} + E_{x(y)}^{\mathrm{miss,\ jets}} + E_{x(y)}^{\mathrm{miss},\mu} + E_{x(y)}^{\mathrm{miss,\ soft}}. \tag{4.2}$$

As the initial momentum in a collision perpendicular to the beam is zero, the vectorial sum of all momenta that are detected have to add to zero if there is no undetected particle in the process. First the algorithm needs all calibrated physics objects that occur in equation (4.2). For all terms, except for the soft term, the $E_\mathrm{T}^\mathrm{miss}$ components are calculated as the negative sum of all calibrated objects. The soft term is the sum of all transverse momenta that are not associated with any of the other physics objects



(a) Relative E_T residual distribution.

(b) Half-widths of the 68% and 95% quantiles of the relative E_T residual distribution as a function of E_T.

Figure 4.4: E_T resolution for TauRec (Baseline), Substructure Reconstruction (Constituent-based) and the combination of both (Tau Particle Flow). Taken from ref. [56]

after their reconstruction. It is either possible to reconstruct this soft term from the calorimeter which is called Calorimeter Soft Term (CST) or from the tracks which is called Track Soft Term (TST). This analysis uses the TST terms. It is possible to calculate the transverse direction and the magnitude $E_\mathrm{T}^\mathrm{miss}$ from $E_{x(y)}^\mathrm{miss}$ as follows:

$$E_\mathrm{T}^\mathrm{miss} = \sqrt{\left(E_x^\mathrm{miss}\right)^2 + \left(E_y^\mathrm{miss}\right)^2} \tag{4.3}$$

$$\phi^\mathrm{miss} = \arctan\left(\frac{E_y^\mathrm{miss}}{E_x^\mathrm{miss}}\right). \tag{4.4}$$

The total transverse energy ΣE_T can be used to quantify the activity in an event and to crosscheck the $E_\mathrm{T}^\mathrm{miss}$ modelling with one reference term for each term listed in (4.2):

$$\Sigma E_\mathrm{T} = \Sigma p_\mathrm{T}^e + \Sigma p_\mathrm{T}^\gamma + \Sigma p_\mathrm{T}^\tau + \Sigma p_\mathrm{T}^\mathrm{jets} + \Sigma p_\mathrm{T}^\mu + \Sigma p_\mathrm{T}^\mathrm{soft} \tag{4.5}$$

This crosscheck is used in chapter 7.1. More on $E_\mathrm{T}^\mathrm{miss}$ reconstruction and its performance can be found in ref. [58] and ref. [59].

4.2 Datasets

This thesis is completely based on Monte Carlo events. The used samples will be shortly introduced in this section.

4.2.1 Samples

All Monte Carlo samples used in this thesis are generated with the `POWHEG` event generator interfaced to `PYTHIA8`. [5] More on these Monte Carlo generators can be found in ref. [60], [61], [62] and [63].

Process	Sample name	N_{events}	$\int \mathcal{L} \mathrm{d}t$
$Z \to \tau\tau$	mc15_13TeV.361108.PowhegPythia8EvtGen_ AZNLOCTEQ6L1_Ztautau.merge.DAOD_ HIGG4D3.e3601_s2726_r7326_r6282_p2608	1475615	0.78 fb^{-1}
$H_{gg} \to \tau\tau$	mc15_13TeV.341124.PowhegPythia8EvtGen_ CT10_AZNLOCTEQ6L1_ggH125_tautauhh.merge. DAOD_HIGG4D3.e3935_s2608_s2183_ r6765_r6282_p2608	564913	18.71 fb^{-1}
$H_{VBF} \to \tau\tau$	mc15_13TeV.341157.PowhegPythia8EvtGen_ CT10_AZNLOCTEQ6L1_VBFH125_tautauhh.merge. DAOD_HIGG4D3.e3888_s2608_s2183_ r6765_r6282_p2608	1068230	278.91 fb^{-1}

Table 4.1: Samples used for the MMC studies in section 6 and section 7.

Process	Sample name	N_{events}	$\int \mathcal{L} \mathrm{d}t$
$H_{gg} \to \tau\tau$ $\mathcal{CP}$-even	mc15_13TeV.341124.PowhegPythia8EvtGen_CT10_ AZNLOCTEQ6L1_ggH125_tautauhh.merge.DAOD_ HIGG4D3.e3935_s2608_s2183_r6765_r6282_p2524	568771	18.84 fb^{-1}
$H_{gg} \to \tau\tau$ unpolarised	mc15_13TeV.341907.PowhegPythia8EvtGen_CT10_ AZNLOCTEQ6L1_ggH125_tautauhh_unpol.merge. DAOD_HIGG4D3.e4103_s2608_r6765_r6282_p2524	384235	12.73 fb^{-1}
$H_{VBF} \to \tau\tau$ $\mathcal{CP}$-even	mc15_13TeV.341157.PowhegPythia8EvtGen_CT10_ AZNLOCTEQ6L1_VBFH125_tautauhh.merge.DAOD_ HIGG4D3.e3888_s2608_s2183_r6765_r6282_p2524	1081053	282.26 fb^{-1}
$H_{VBF} \to \tau\tau$ unpolarised	mc15_13TeV.341913.PowhegPythia8EvtGen_CT10_ AZNLOCTEQ6L1_VBFH125_tautauhh_unpol.merge. DAOD_HIGG4D3.e4103_s2608_r6765_r6282_p2524	527564	137.75 fb^{-1}

Table 4.2: Samples used for the $\mathcal{CP}$ and performance studies in section 5 and section 8.

Table 4.1 lists the samples used for the MMC studies performed in section 6 and 7. The samples used for the performance studies and studies on φ^*_{CP}, section 5 and section 8, are listed in table 4.2.

In order to increase the available statistics for the performance studies and $\mathcal{CP}$ studies, the statistically independent Monte Carlo samples of the $\mathcal{CP}$-even $H_{gg} \to \tau\tau$ process, the unpolarised $H_{gg} \to \tau\tau$ process, the $\mathcal{CP}$-even $H_{VBF} \to \tau\tau$ process and the unpolarised $H_{VBF} \to \tau\tau$ process are combined. The unpolarised samples are reweighted with `TauSpinner` to a $\mathcal{CP}$-even distribution if not stated otherwise. `TauSpinner` provides weights which allows to include transverse spin correlations in τ decays in the calculation, for further detail see ref. [64]. If $\mathcal{CP}$-mixed states are studied, the $\mathcal{CP}$-even samples are reweighted to that specific $\mathcal{CP}$-mix by first dividing by the $\mathcal{CP}$-even `TauSpinner` weight and then multiplying with the corresponding `TauSpinner` $\mathcal{CP}$-mix weight. In section 5.1 the relative amplitude and the phase of the φ^*_{CP} distribution are studied for the different Higgs production modes which justifies the aforementioned combination of the samples.

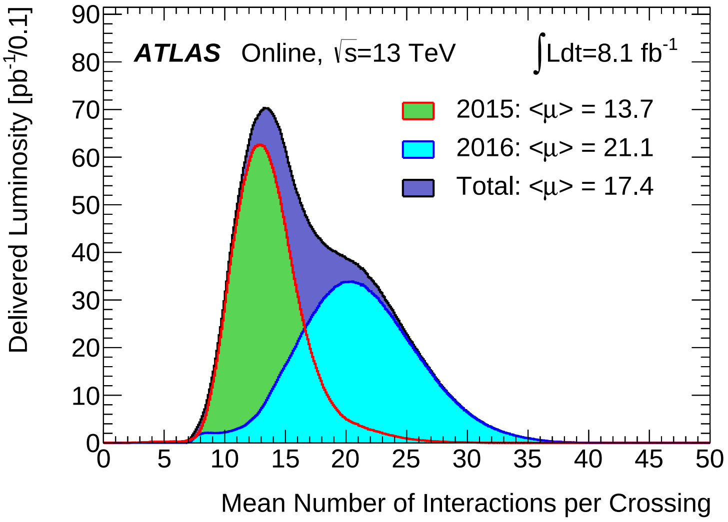


Figure 4.5: μ profile for 2015 and 2016 and the combination. Taken from ref. [38]

4.2.2 Pile-up

Signatures from other processes than the one of interest happening in the same bunch crossing are generally referred to as pile-up. For example if a Higgs boson is produced in a bunch crossing and in the same crossing another process like the hard production of jets takes place but is unrelated to the production of this Higgs. This specific example is called in-time pile-up. Out-of-time pile-up are additional collisions in neighbouring bunch-crossings as it is possible that detector-components take longer than 25 ns to accumulate the signal. In addition different background processes can occur. One of them being neutrons and photons that leave random signatures in the muon system, called cavern backgrounds. Furthermore, beam halo events can occur when the protons from the bunches scrape an up-stream collimator and result in muons approximately parallel to the beam axis. Lastly, beam gas events are events where the protons interact with gas remnants in the beam pipe which happens most of the time not in the center. More on pile-up and its simulation can be found in ref. [65]

The profile of mean interactions per bunch crossing, μ, which is also a measure for the pile-up, is shown in figure 4.5. The difference between 2015 and 2016 can be explained by a higher bunch luminosity as the definition of μ is:

$$\mu = \frac{L_{bunch} \times \sigma_{inel}}{f_r}. \tag{4.6}$$

To get agreement between Monte Carlo and recorded data a pile-up-reweighting is applied to the Monte Carlo events. This reweighting ensures that the pile-up-profile is identical for both such that the simulation can be compared to actual data in a measurement. The samples listed in table 4.1 are pile-up reweighted for the studies in section 5 and section 8 while the samples in table 4.2 are not reweighted for the studies in section 5 and section 8.

CHAPTER 5

Introduction of the $\mathcal{CP}$ measurement

In the following the event selection for a $H \rightarrow \tau\tau$ analysis is introduced. Also, a short discussion of the major backgrounds and their estimation is given.Afterwards, the performance of the Substructure Reconstruction is investigated. A special focus lies on the performance as a function of the observables introduced in section 2.5.2 and their resolutions. Additionally, the performance of the method introduced in section 2.5 is studied as a function of different τ-ID requirements. For studies that investigate the impact of reconstruction effects on φ^*_{CP} a fit to the distribution with the function $f(x) = u \cdot \cos(x + v) + w$ is performed and the relative amplitude is defined as rel. amp. = $\frac{u}{w}$ and the phase shift in units of π is v.

5.1 Event Selection and Backgrounds in a $H \rightarrow \tau\tau$ analysis at the LHC

The event selection used in this thesis is inspired by the one used in the analysis that provided evidence for the Yukawa coupling of the Higgs boson to fermions, see ref. [5]. As this thesis concentrates on the $H \rightarrow \tau\tau$ decay where both τ's decay hadronically the di-hadronic channel is relevant here. Backgrounds are not investigated in the scope of this thesis. However, to see the effects of a full analysis selection the background suppression cuts are applied. The full event selection and the different categories together with the fraction of events passing the cuts are summarized in table 5.1. This event selection follows the cut-based analysis performed in the aforementioned $H \rightarrow \tau\tau$ coupling analysis. Even though, the coupling analysis also performs a multivariate analysis by using a BDT, this thesis concentrates on the cut-based selection. In the cut-based approach it is easier to understand how the cuts act on the phase space defined by the signal regions whereas it is not as easy to understand how the input variables in the BDT are weighted and how they affect the signal regions. However, some input variables of the BDT are identical to the variables used in the cut-based analysis of this thesis. These are m_{MMC}, ΔR, $\Delta\eta_{jj}$ and $E_{\mathrm{T}}^{\mathrm{miss}}$-$\phi$-centrality.

The triggers used are the single-τ and di-τ triggers. p_{T} cuts are applied on both τ's to ensure that all selected τ's are in the trigger plateau. In an actual analysis the identification criteria are applied to suppress the background from jets that are falsely classified as τ's. The requirement on the number of tracks further reduces the jets background while allowing the Substructure Reconstruction to reconstruct all identifiable decay modes. Thus, there is no bias in the decay mode classification. As the investigated τ_{had}'s originate from a Higgs decay which is a neutral particle both τ_{had}'s are required to be oppositely charged with an absolute charge of one respectively. The last cut for the Baseline selection is a cut on the decay mode which is required to be `1p1n` as the method introduced in section 2.5 relies on the

Category	Cut	$\frac{N_{pass}}{N_{all}}$
Baseline	ditau_tau0_HLT_tau35_medium1_tracktwo_tau25_medium1 _tracktwo_L1TAU20IM_2TAU12IM!=0	
	ditau_tau1_HLT_tau35_medium1_tracktwo_tau25_medium1 _tracktwo_L1TAU20IM_2TAU12IM!=0	100%
	$p_{\mathrm{T}}^{\tau_0} > 40\,\mathrm{GeV}$ and $p_{\mathrm{T}}^{\tau_1} > 30\,\mathrm{GeV}$	82.16%
	one `tight` and one `medium` identified τ	40.65%
	$N_{trk} == 1 \parallel N_{trk} == 3$	38.07%
	$q_{\tau_{0,1}} == 1$ && $q_{\tau_0} \times q_{\tau_1} < 0$	37.54%
	`1p1n` decay mode (reconstructed and generated when studying resolutions)	5.18%
	truth matched τ's	5.14%
Preselection	$\Delta\eta < 1.5$	4.86%
	$0.8 < \Delta R < 2.4$	2.42%
	$E_{\mathrm{T}}^{\mathrm{miss}} > 20\,\mathrm{GeV}$	2.02%
	$E_{\mathrm{T}}^{\mathrm{miss}}$-$\phi$-centrality: $E_{\mathrm{T}}^{\mathrm{miss}}$ points in between the two τ_{had}'s or $\min\left(\Delta\phi\left(\tau_{\mathrm{had}}, E_{\mathrm{T}}^{\mathrm{miss}}\right)\right) < \frac{\pi}{4}$	1.86%
VBF Signal Region	$N_{jets} > 1$	1.59%
	$p_{\mathrm{T}}^{jet_0} > 50\,\mathrm{GeV}$	1.48%
	$p_{\mathrm{T}}^{jet_1} > 30\,\mathrm{GeV}$	1.17%
	$\Delta\eta_{jj} > 2.$	0.88%
	$100\,\mathrm{GeV} < m_{MMC} < 150\,\mathrm{GeV}$	0.77%
Boosted Signal Region	fails VBF Signal Region cuts	1.09%
	$p_{\mathrm{T}}^{H} > 100\,\mathrm{GeV}$	0.34%
	$100\,\mathrm{GeV} < m_{MMC} < 150\,\mathrm{GeV}$	0.27%

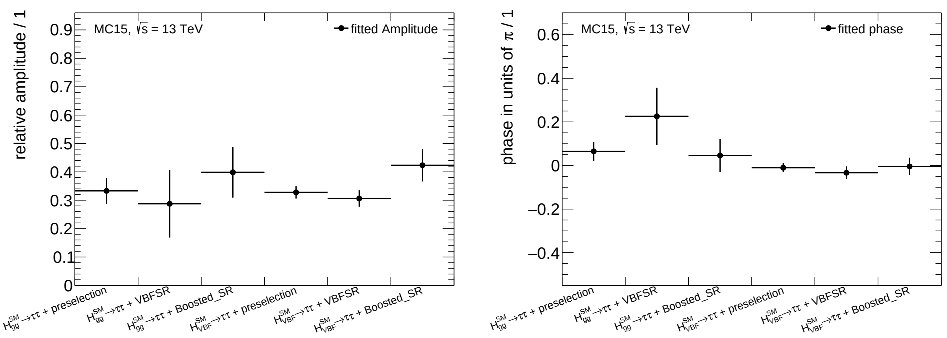
Table 5.1: Event selection and the different selection categories used in this thesis. The fraction of events passing the cuts is also given.

τ_{had}'s decaying into one charged pion and one neutral pion. Truth matching is required because of the resolution studies that are performed in this thesis.

The cuts applied for the Preselection category further reduce the backgrounds, specifically multijet backgrounds.

For the analysis both signal regions of the aforementioned coupling analysis are investigated. A Boosted Signal Region and a VBF Signal Region. The corresponding Feynman diagrams of the production processes are shown in figure 2.6 in section 2.3.1. Gluon-gluon-fusion and vector boson fusion are the two dominant production processes of the Higgs boson at the LHC. The two signal regions aim to exploit characteristic event topologies of each process, for example two additional highly separated and high p_{T} jets in the VBF signal region. Both signal regions include a cut on the ditau-mass reconstructed with the Missing Mass Calculator, specifically the MLM output, which will be further introduced in section 7. This mass window is used to discriminate against background from the $Z \to \tau\tau$ process. Figure 6.7 in section 6.4 shows the mass distributions for $H \to \tau\tau$ and $Z \to \tau\tau$ which motivate the cut on the mass window.

The VBF Signal Region targets events where the Higgs boson is produced via vector boson fusion. As it can be seen in figure 2.6 in section 2.3.1 there are two additional quarks in this process which



(a) Relative amplitude at generator level.

(b) Phase shift at generator level.

Figure 5.1: Relative amplitude and phase shift for the $H_{gg} \to \tau\tau$ sample and $H_{VBF} \to \tau\tau$ sample in the different event selection regions at generator level.

hadronise in the detector and produce two additional high-p_{T} jets which are typically separated by a large $\Delta\eta$. In this selection it is assumed that the two highest-p_{T} jets originate from this process while it allows additional jets from pile-up processes. Although the VBF Signal Region is defined for Higgs bosons produced via vector boson fusion, also Higgs bosons produced via gluon-gluon-fusion possibly pass this selection. In order to avoid overlap between the two signal regions the events passing the Boosted Signal Region are required to fail the VBF Signal Region cuts.

As the Boosted Signal Region is designed for Higgs boson events produced via gluon-gluon-fusion it exploits that the Higgs in this process is typically heavily boosted. Thus, the vectorial sum of $E_{\mathrm{T}}^{\mathrm{miss}}$ and the p_{T} of the two τ_{had}'s is called p_{T}^{H} and is required to be larger than 100 GeV.

As described in section 4.2.1 different samples, including different production modes of the Higgs boson, are combined in order to increase the available statistics. Figure 5.1 shows that the relative amplitudes and the phase shifts are compatible within the statistical uncertainties for the $H_{gg} \to \tau\tau$ sample and $H_{VBF} \to \tau\tau$ sample at generator level. Thus, they are combined. However, this combination makes it impossible to give a total luminosity for the combined sample.

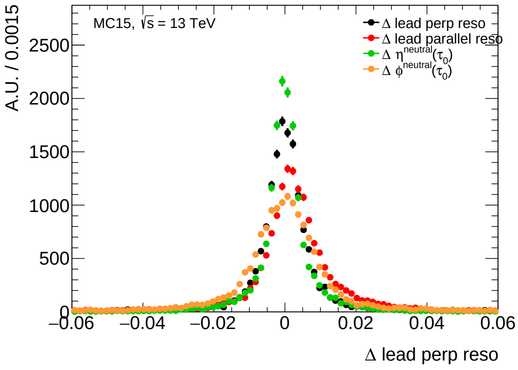
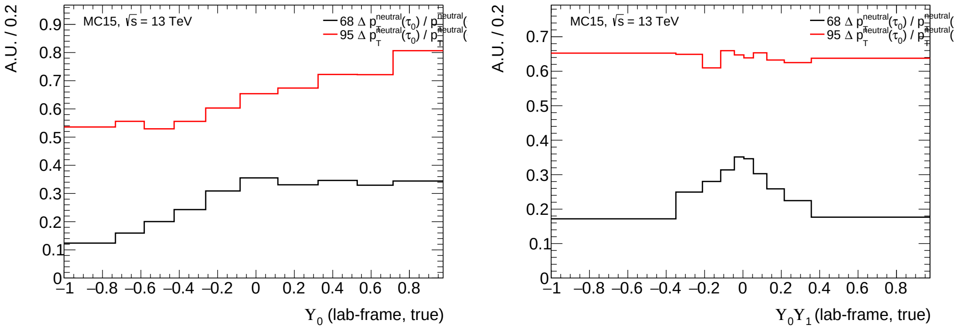


Figure 5.2: Angular resolution of neutral pions of leading τ_{had} in `1p1n` decays along the axis (Δleadparallelreso) which connects the charged pion and the neutral pion in the η-ϕ-plane and perpendicular to this axis (Δleadperpreso). For completeness, the η and ϕ resolution of the neutral pions are shown.

5.2 Substructure Performance at 13 TeV

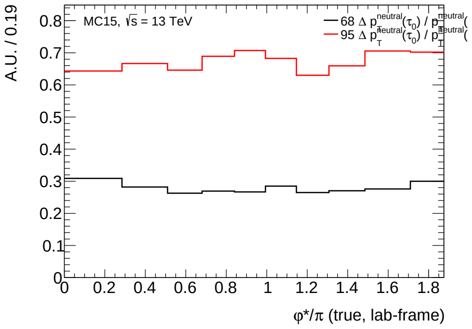
All observables of the $\mathcal{CP}$ analysis rely on the reconstruction of the individual decay products of the τ_{had}. Thus, the performance of the neutral pion reconstruction is studied in this section as they are harder to reconstruct than the charged pions. For the reconstruction of the charged pions, information from the tracking system is used which gives a smaller four-momentum resolution than the calorimeter information needed for the neutral pion reconstruction. These resolution studies are performed with the Baseline selection from table 5.1 applied and with the combined Higgs dataset as described in section 4.2.1. Considering the $\mathcal{CP}$ measurement introduced in section 2.5 the angular resolution of the neutral pions is of peculiar interest. The most relevant component of the neutral pion angular resolution is the one perpendicular to the axis that connects the neutral pion and the charged pion in the η-ϕ-plane. Although, the component of the resolution along this axis can in principle have an impact on the observables, it presumably has a small influence compared to the component perpendicular to this axis. Figure 5.2 shows this resolution in addition with the η and ϕ resolutions of the neutral pion of the leading τ_{had}. The angular resolution is defined as the difference between the reconstructed angle and the generated angle. As it can be seen, the resolution along the previously defined axis is broader than the resolution perpendicular to this axis and shifted by about 0.0027 with respect to zero. This is presumably caused by charged pion remnants from the subtraction in the neutral pion reconstruction. These remnants can introduce a bias of the neutral pion resolution in the direction towards the charged pion. This effect does not occur perpendicular to this direction. In order to have a measure for the resolution that is mostly independent of non-gaussian effects the 68% and 95% quantiles of the resolution curves are used. These measures where used in the paper describing the new τ reconstruction. [56] The 68% (95%) quantile of the resolution along the aforementioned axis is 0.008 (0.029) and for the resolution perpendicular to this axis 0.006 (0.025). This means that the resolution in the core region is approximately 33% better for the perpendicular component while it is approximately 16% better in the tail region. For comparison, the corresponding quantiles for the η and ϕ resolutions are 0.005 (0.020) and 0.009 (0.030) respectively.

It is also important to know how the reconstruction performs as a function of the observables defined in section 2.5 to make sure possible performance degradations are understood. Figure 5.3 shows the relative p_{T} resolution of the neutral pions as a function of these observables. Again, the 68% and the 95% quantiles are used to describe the core and tail resolutions. The corresponding angular resolutions for the perpendicular component with respect to the previously defined axis is shown in figure 5.4. All



(a) p_T resolution as a function of y_0. A similar behavior is seen for y_1.

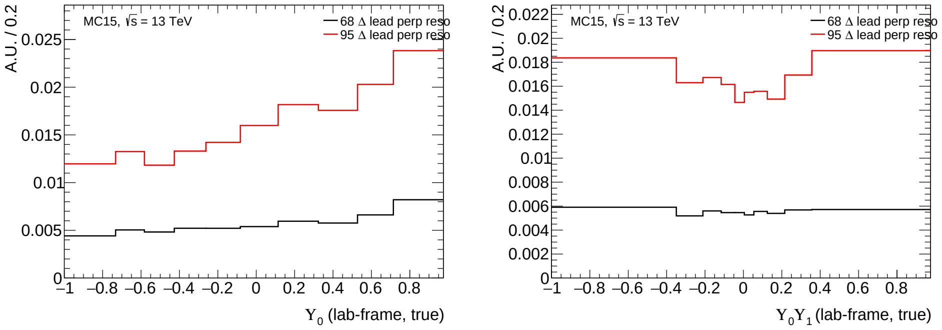
(b) p_T resolution as a function of $y_0 y_1$.



(c) p_T resolution as a function of φ^*_{CP}.

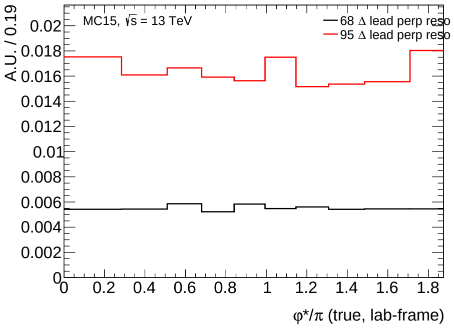
Figure 5.3: 68% and 95% quantiles of the p_T resolution of the neutral pion as a function of the observables.

resolutions are independent of φ^*_{CP}. When looking at y_0 though the resolutions are not independent anymore but get broader with increasing values of y_0. As higher values of y_0 correspond to higher energetic charged pions and less energetic neutral pions, see equation (2.24), this dependence is expected due to a less precise neutral pion reconstruction in this scenario. The same can be observed for increasing values of y_1 which can be found in figure A.1 in appendix A.1. For $y_0 y_1 \approx 0$ the p_T resolution in the 68% quantile is broader since the resolution flattens out for $y_{0,1} > 0$, see figure 5.3(a), and negative values of $y_{0,1} \approx -1$ play a less significant role in this region.



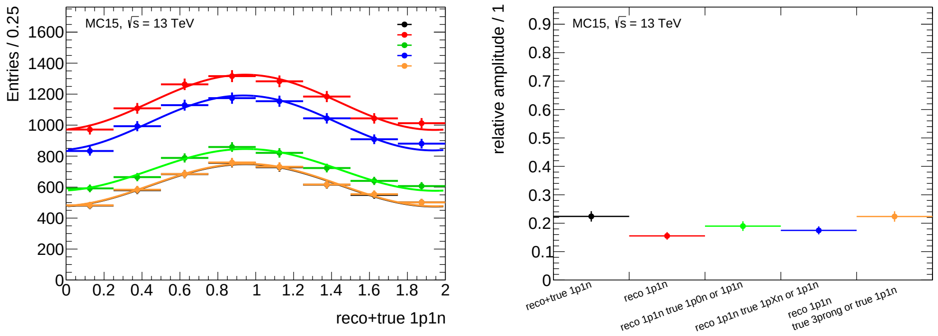
(a) Angular resolution as a function of y_0. A similar behavior is seen for y_1.

(b) Angular resolution as a function of y_0y_1.



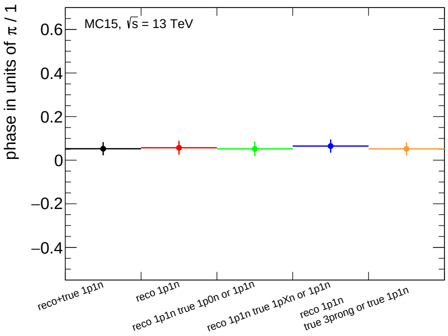
(c) Angular resolution as a function of φ^*_{CP}.

Figure 5.4: 68% and 95% quantiles of the angular resolution of the neutral pion perpendicular to the axis that connects the charged pion and neutral pion in the η-ϕ-plane as a function of the observables.



(a) φ^*_{CP} distribution.

(b) Relative amplitude extracted from the φ^*_{CP} distribution.



(c) Phase shift extracted from the φ^*_{CP} distribution.

Figure 5.5: Effect of falsely classified decay modes on the relative amplitude of the φ^*_{CP} distribution, studied by adding only one possible misidentified decay mode to the generated `1p1n` to check its influence on the `1p1n`-only φ^*_{CP} distribution. The Preselection is applied.

5.3 Effects of decay mode classification on the φ^*_{CP} performance

One major difference between TauRec and the Substructure Reconstruction is the possibility to classify the τ_{had}'s decay mode. Although, the classification efficiency is very high it is not 100%. Thus, the effect of falsely classified decay modes on φ^*_{CP} and the decay mode classification efficiency and purity are studied in this section. Figure 5.5 shows the relative amplitude and the phase shift of φ^*_{CP} for different combinations of generated decay modes. The Preselection is applied and the decay mode requirements are as follows. It is always required that the reconstructed decay mode is `1p1n` and the generated decay mode is always `1p1n` in addition with another generated decay mode to see what impact the misclassification of the additional decay mode has. Between the truth matched `1p1n` requirement and only the reconstructed requirement the amplitude is degraded from 22% to 15.5%. The biggest contribution here comes from misreconstructed `1pXn` decays which is also the biggest contribution to the `1p1n` decay mode purity. The phase shift is constant as a function of the misreconstructed decay modes. The constant shift is caused by other reconstruction effects, see section 8.1.4.

To study whether the contribution of the misclassified decay modes is indeed flat, it can be required that one of the two τ_{had}'s or both τ_{had}'s are misclassified. Figure 5.6 shows this after the Preselection. Only for `1pXn` decays that are reconstructed as `1p1n` the contribution is not completely flat. This is presumably caused by decays where one of the τ_{had}'s is a correctly reconstructed `1p1n` decay. Additionally, in `1pXn`

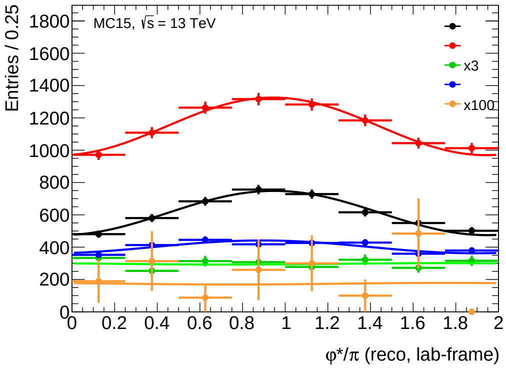


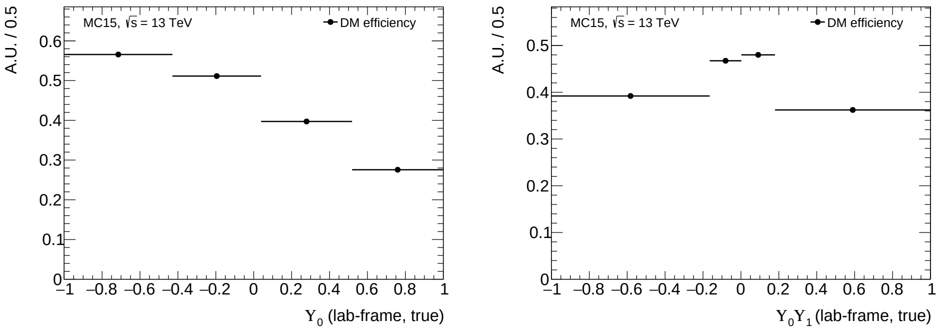
Figure 5.6: φ^*_{CP} distributions for different decay mode requirements. It is required that either one of the two τ_{had}'s or both τ_{had}'s are misclassified. The Preselection is applied and the same colour-coding as in figure 5.5 is used. The misreconstructed `1p0n` and 3-prong decays are scaled by 3 and 100 respectively.

decays the charged pions also carry spin analyser information, see ref. [34], which then can lead to a small φ^*_{CP} amplitude for these decays.

For the systematic uncertainties it would be ideal if the decay mode classification efficiency and purity would be independent of φ^*_{CP} and $y_0 y_1$. The definition for the efficiency and purity are as follows:

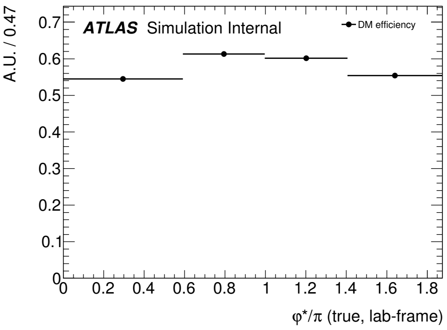
$$\text{Eff.} = \frac{N_{matched}}{N_{generated}}, \qquad \text{Pur.} = \frac{N_{matched}}{N_{reconstructed}}. \tag{5.1}$$

Figure 5.7 shows the dependence of the efficiency on the observables at generator level and figure 5.8 shows the corresponding dependence of the purity on the observables at reconstruction level. For both studies the Baseline selection is used and both τ_{had}'s have to satisfy the corresponding decay mode requirement. y_0 and y_1 are defined for `1p1n` decays. When these are calculated for misreconstructed decays, namely for reconstructed `1p1n` decays, they are set to −2.0 if there is no neutral pion in the generated decay mode or the neutral pion with the highest p_{T} is used for decay modes with more than one. The decay mode purity shows no significant dependence on either of the observables. For the decay mode efficiency the same is observed as a function of φ^*_{CP}. However, the decay mode efficiency shows a significant drop for increasing y_0. Figures 5.7 (e-f) show the fraction of true `1p1n` decays that are reconstructed as `1p0n` decays or as `1pXn` decays as a function of y_0. For higher generated values of y_0 the neutral pion energy decreases and thus more neutral pions are not reconstructed. This can happen for example when they are so low energetic that they cannot pass the minimum p_{T} cut in the reconstruction. Hence, these events are reconstructed as `1p0n` decays and the decay mode efficiency decreases. The fraction of true `1p1n` decays that are reconstructed as `1pXn` is constant.

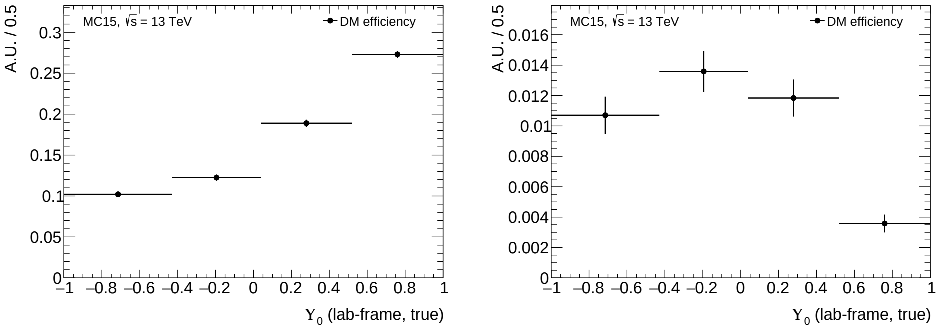


(a) Decay mode efficiency as a function of y_0. A similar behavior is seen as a function of y_1.

(b) Decay mode efficiency as a function of $y_0 y_1$.



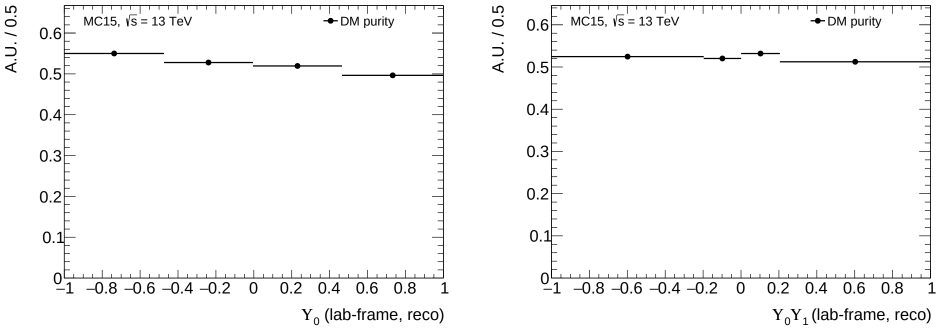
(c) Decay mode efficiency as a function of φ^*_{CP}.



(d) Fraction of true `1p1n` decays that are reconstructed as `1p0n` decays as a function of y_0. A similar behavior is seen as a function of y_1.

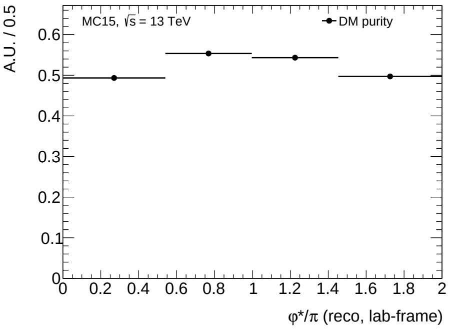
(e) Fraction of true `1p1n` decays that are reconstructed as `1pXn` decays as a function of y_0. A similar behavior is seen as a function of y_1.

Figure 5.7: Decay mode efficiency as a function of the observables.



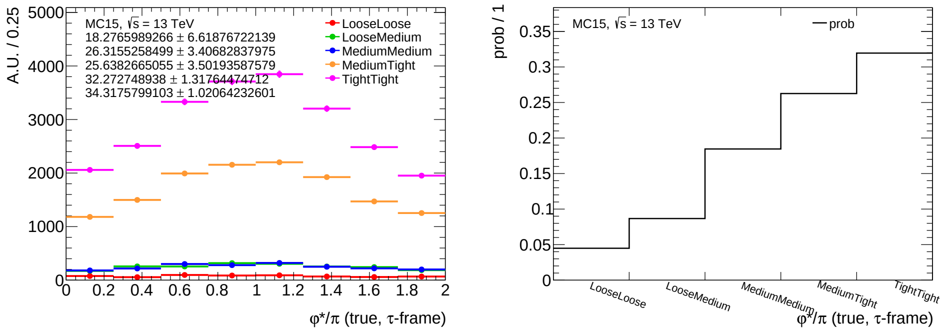
(a) Decay mode purity as a function of y_0. A similar behavior is seen as a function of y_1.

(b) Decay mode purity as a function of $y_0 y_1$.



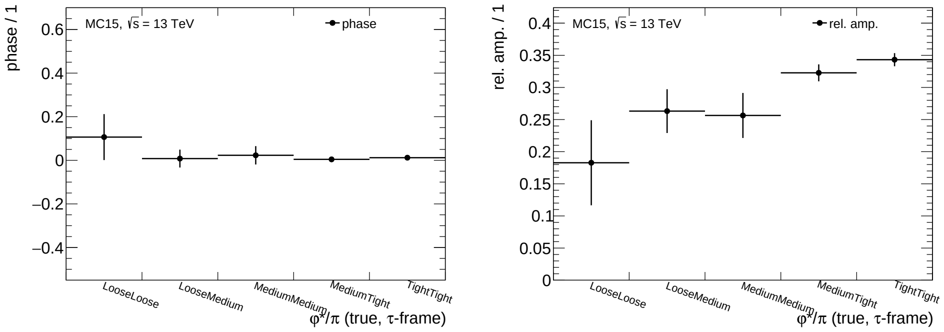
(c) Decay mode purity as a function of φ^*_{CP}.

Figure 5.8: Decay mode purity as a function of the observables.



(a) φ^*_{CP} distribution.

(b) Fit probabilities.



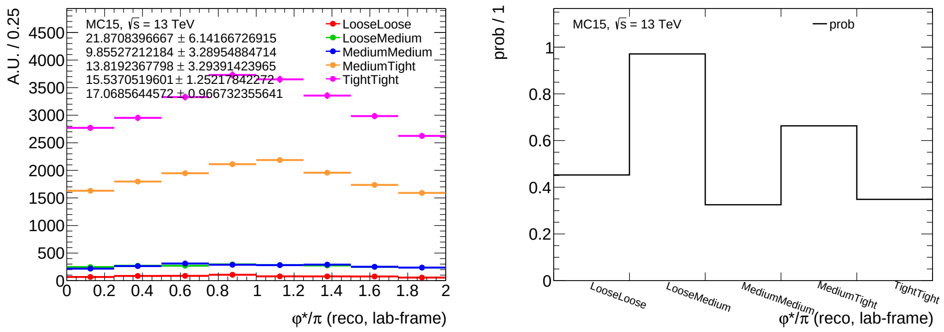
(c) Phase shift.

(d) Relative amplitude.

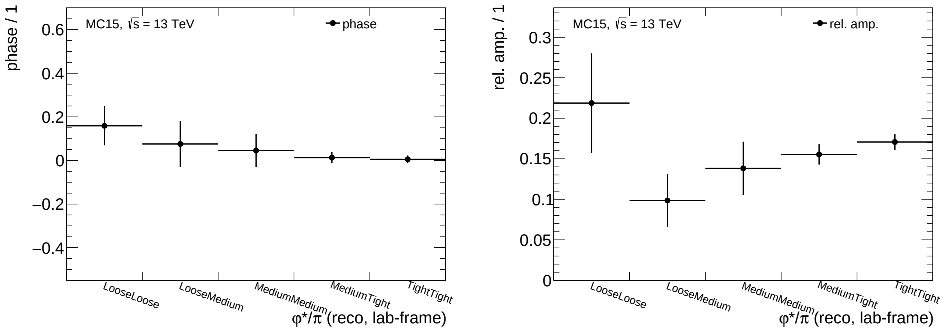
Figure 5.9: Behaviour of the generated φ^*_{CP} distribution for different ID requirement combinations on the τ_{had}'s.

5.4 Effects of τ-ID on the φ^*_{CP} performance

In this section the φ^*_{CP} performance as a function of the object identification is studied in order to find and understand possible performance degradations in the analysis. The effect of the different ID requirements on the generated φ^*_{CP} distribution is studied in figure 5.9 and in figure 5.10 for the reconstructed distribution. Only a very basic selection is applied, i.e. opposite signs of the τ_{had}'s, a trigger match for both and the reconstructed decay mode is required to be `1p1n`. As a τ_{had} that would be identified with the `tight` criteria would also be identified as `medium` and `loose`, τ_{had}'s that are identified as for example `loose` are also required to not pass the `medium` and `tight` criteria. These requirements are supposed to ensure that the different bins are as independent of each other as possible. Figure 5.9 shows that there is no systematic bias in the phase introduced except for the loosest combination and that the relative amplitude increases with tighter requirements. This increase is diminished by reconstruction effects in figure 5.10 such that the relative amplitude is almost stable as a function of the ID requirements and the only bias in the phase is seen for the combination `LooseLoose`.



(a) φ^*_{CP} distribution. (b) Fit probabilities.



(c) Phase shift. (d) Relative amplitude.

Figure 5.10: Behaviour of the reconstructed φ^*_{CP} distribution for different ID requirement combinations on the τ_{had}'s.

CHAPTER 6

Ditau-mass reconstruction techniques

Since $Z \to \tau\tau$ and $H \to \tau\tau$ have similar event topologies the reconstructed invariant ditau-mass of the event is an important discriminant between these two processes. This section introduces different techniques to reconstruct this invariant mass including the technique used in this thesis, the Missing Mass Calculator. Figure 2.4 in section 2.1.3 show the schematic view of a typical τ decay. The neutrino in this decay is invisible for the detector. Thus, the major goal of the following techniques is to approximate the neutrino four-momenta. If it was possible to measure the neutrinos the invariant mass of for example the process $H \to \tau\tau$ could be measured by:

$$m_H = \sqrt{(p_{\mathrm{vis},\tau_0} + p_{\nu_0} + p_{\mathrm{vis},\tau_1} + p_{\nu_1})^2} \tag{6.1}$$

where p_{vis,τ_i} denotes the complete four-momentum of all visible decay products.

6.1 Visible mass

Using the visible mass m_{vis} of the ditau-system is the simplest approach. This approach simply neglects the momenta carried away by the neutrinos which leads to the definition:

$$m_{\mathrm{vis}} = \sqrt{(p_{\mathrm{vis},\tau_0} + p_{\mathrm{vis},\tau_1})^2} \tag{6.2}$$

As the magnitude of the neutrino momenta are statistically distributed, this neglection leads to a smeared out mass distribution which in turn leads to a degradation in the separation power of $H \to \tau\tau$ and $Z \to \tau\tau$. Figure 6.1 shows the m_{vis} distributions for $H \to \tau\tau$ and $Z \to \tau\tau$ respectively and also the overlap of both distributions with the Baseline selection applied. The overlap is used to quantify the separation of the distributions and is defined as the ratio of the bin by bin sum of the bins with the lower entries divided by the sum of the bins with higher entries:

$$\text{overlap} = \frac{\Sigma bin_{\mathrm{lower}}}{\Sigma bin_{\mathrm{higher}}}. \tag{6.3}$$

This definition of the overlap will be used throughout this thesis as a measure for such histograms to quantify how good the separation between two distributions is.

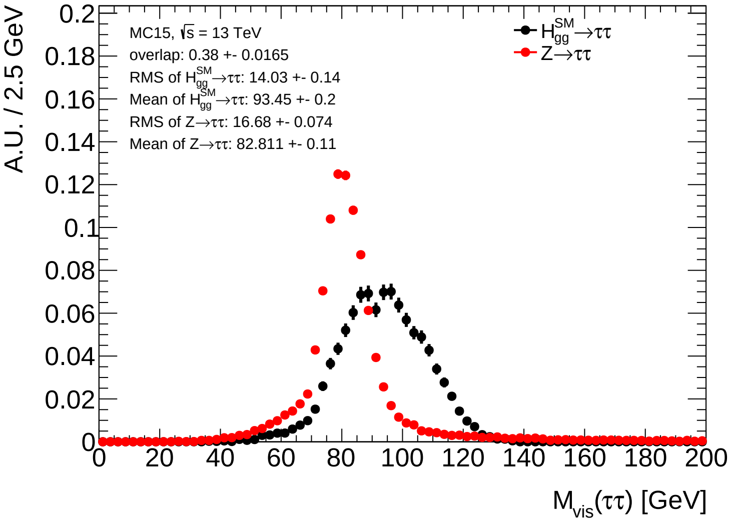


Figure 6.1: Distributions of the visible mass for $H \rightarrow \tau\tau$ and $Z \rightarrow \tau\tau$ with the Baseline selection applied.

6.2 Transverse Mass

A first approach to include the neutrinos in the invariant mass calculation is to consider the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector measured in an event. This approach is called the Transverse Mass and does not make any other assumptions on the event kinematics but that $E_{\mathrm{T}}^{\mathrm{miss}}$ is solely due to the neutrinos and perfectly measured. The Transverse Mass is then defined as follows: [66]

$$\begin{aligned} M_T^2\left(\tau_{\mathrm{vis},0}, \tau_{\mathrm{vis},1}, E_{\mathrm{T}}^{\mathrm{miss}}\right) =& m_{\mathrm{vis},0}^2 + m_{\mathrm{vis},1}^2 \\ &+ 2\cdot\left(\sqrt{m_{\mathrm{vis},0}^2 + p_{\mathrm{vis},0}^2}\sqrt{m_{\mathrm{vis},1}^2 + p_{\mathrm{vis},1}^2} + E_{\mathrm{T}}^{\mathrm{miss}}\sqrt{m_{\mathrm{vis},0}^2 + p_{\mathrm{vis},0}^2} + E_{\mathrm{T}}^{\mathrm{miss}}\sqrt{m_{\mathrm{vis},1}^2 + p_{\mathrm{vis},1}^2}\right) \\ &- 2\cdot\left(\vec{p}_{\mathrm{vis},0}\cdot\vec{p}_{\mathrm{vis},1} + \vec{p}_{\mathrm{vis},0}\cdot\overrightarrow{E_{\mathrm{T}}^{\mathrm{miss}}} + \vec{p}_{\mathrm{vis},1}\cdot\overrightarrow{E_{\mathrm{T}}^{\mathrm{miss}}}\right). \end{aligned} \tag{6.4}$$

$p_{\mathrm{vis},i}^2$ is the squared magnitude of the momentum of the visible decay products of the τ, $\overrightarrow{E_{\mathrm{T}}^{\mathrm{miss}}} = \left(E_{\mathrm{T,x}}^{\mathrm{miss}}, E_{\mathrm{T,y}}^{\mathrm{miss}}, 0\right)^T$ and $E_{\mathrm{T}}^{\mathrm{miss}}$ is the magnitude of this vector. The resulting mass distributions can be seen in figure 6.2. The Baseline selection is applied. It shows that the overlap of the distributions for $H \rightarrow \tau\tau$ and $Z \rightarrow \tau\tau$ is higher than for the visible mass as the width is broader for both. The mean of the distributions has not changed with respect to the visible mass. These observations are the motivation for another approach of including the missing transverse energies caused by the neutrinos.

6.3 Collinear Approximation

The Collinear Approximation is another basic approach of including the neutrinos' four-momenta in the invariant mass calculation. As it is not possible to measure the neutrinos directly some assumptions are applied in this approximation. Figure 6.3 illustrates the assumed direction of the neutrinos in the τ decays. The method is designed for τ decays where the τ's originate from $Z \rightarrow \tau\tau$ or $H \rightarrow \tau\tau$ processes such that it can be assumed that the mass of its mother particle is much higher than the τ mass itself, i.e. $m_{H/Z} >> m_\tau$. This results in heavily boosted τ's and the approximation that the τ's and their decay

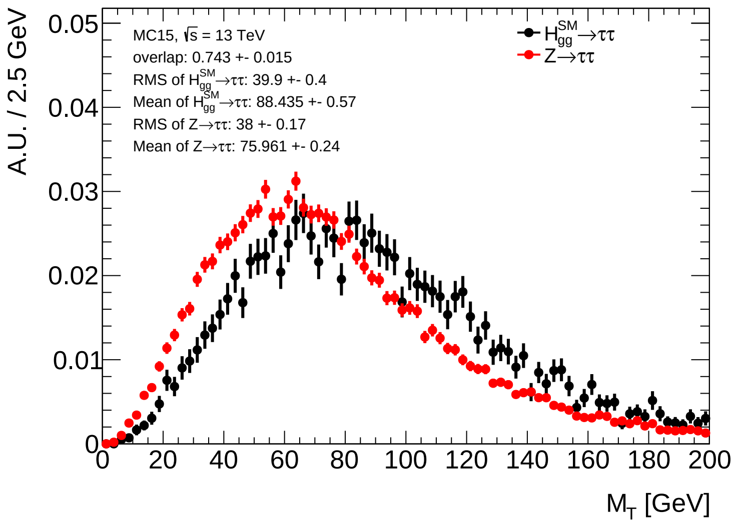
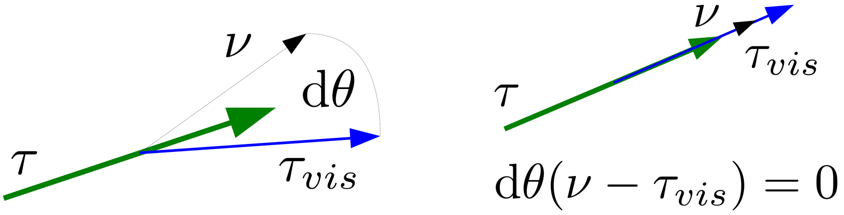


Figure 6.2: Distributions of the transverse mass, M_T, for $H \to \tau\tau$ and $Z \to \tau\tau$ with the Baseline selection applied.



(a) Angular distance $\mathrm{d}\theta$ between the visible τ and the neutrino after the boost.

(b) Angular distance $\mathrm{d}\theta$ between the visible τ and the neutrino after the boost.

Figure 6.3: Illustration of the assumed direction of the neutrinos in the Collinear Approximation.

products go in the same direction, i.e. the angle between the visible τ and the neutrino in each decay is zero $\mathrm{d}\theta = 0$.

Furthermore, it is assumed that $E_{\mathrm{T}}^{\mathrm{miss}}$ is only due to the neutrinos in the τ decays and is perfectly measured. These assumptions lead to two equations with two unknowns which can be solved analytically: [66]

$$E_{\mathrm{T,x}}^{\mathrm{miss}} = p_{\nu_0} \cdot \sin(\theta_0) \cdot \cos(\phi_0) + p_{\nu_1} \cdot \sin(\theta_1) \cdot \cos(\phi_1) \tag{6.5}$$

$$E_{\mathrm{T,y}}^{\mathrm{miss}} = p_{\nu_0} \cdot \sin(\theta_0) \cdot \sin(\phi_0) + p_{\nu_1} \cdot \sin(\theta_1) \cdot \sin(\phi_1)\,. \tag{6.6}$$

p_{ν_i} represent the magnitude of the neutrino momenta and the angles are the measured angles for the visible decay products of the τ's. The resulting ditau-mass distribution with the Baseline selection applied is shown in figure 6.4. It can be seen that the peaks of both distributions are closer to the nominal values of a Z and H boson respectively. Although the width is still very broad and the overlap is bigger than for the visible mass, it is smaller than for M_T. More details on the Collinear Approximation can be found in ref. [66].

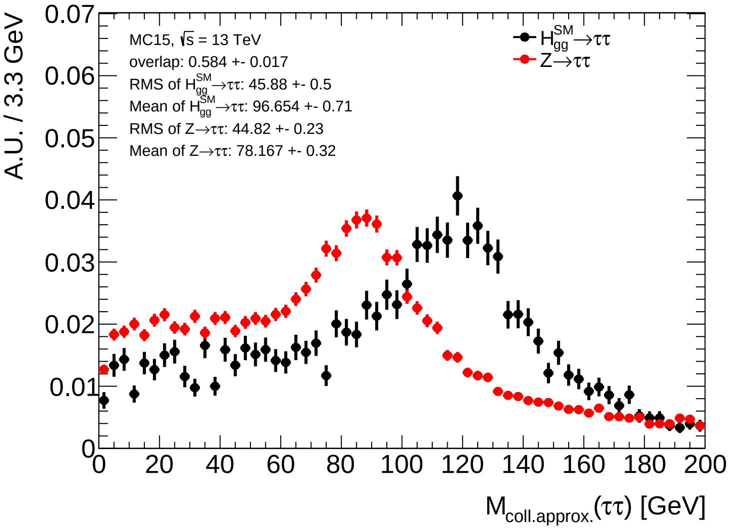


Figure 6.4: Distributions of the collinear mass for $H \rightarrow \tau\tau$ and $Z \rightarrow \tau\tau$ with the Baseline selection applied.

6.4 Missing Mass Calculator

The M(issing) M(ass) C(alculator) is a more sophisticated way of estimating the ditau-mass. Similar to the Collinear Approximation it assumes that the mass of the particle that the τ's originate from is heavy compared to the τ mass, i.e. $m_{H/Z} >> m_\tau$. This again results in a boost of the τ's which means that the τ itself and its visible and invisible decay products are collimated, see figure 6.3. Unlike the Collinear Approximation, the MMC takes into account that in general the τ and its decay products do not go in the exact same direction, i.e. $\mathrm{d}\theta \neq 0$. For the further procedure it will be assumed that the distribution of this angular distance $\mathrm{d}\theta$ is only dependent on the type of τ decay and its four-momentum. The different τ types in this context are either that both τ's decay leptonically, one leptonically and one hadronically decaying τ or that both τ's decay hadronically. Additionally, it is differenced between 1-prong and 3-prong τ's for hadronic decays. Furthermore, it is still assumed that $E_{\mathrm{T}}^{\mathrm{miss}}$ is only caused by the neutrinos in the τ decay. However, the resolution of $E_{\mathrm{T}}^{\mathrm{miss}}$ is taken into account in the calculation. This leads to a parametrisation of $E_{\mathrm{T}}^{\mathrm{miss}}$. MMC is not only able to give an estimate of the ditau-mass but also an estimation of the four-momenta of the neutrinos or the neutrino-system for each τ respectively if the τ decays leptonically.

As $\mathrm{d}\theta$ can in principle differ from zero the number of unknown parameters in the equations used to calculate the neutrino momenta increases. Following the notation in ref. [66] these equations read:

$$E_{\mathrm{T,x}}^{\mathrm{miss}} = p_{\nu_0} \cdot \sin(\theta_{\nu_0}) \cdot \cos(\phi_{\nu_0}) + p_{\nu_1} \cdot \sin(\theta_{\nu_1}) \cdot \cos(\phi_{\nu_1}) \tag{6.7}$$

$$E_{\mathrm{T,y}}^{\mathrm{miss}} = p_{\nu_0} \cdot \sin(\theta_{\nu_0}) \cdot \sin(\phi_{\nu_0}) + p_{\nu_1} \cdot \sin(\theta_{\nu_1}) \cdot \sin(\phi_{\nu_1}) \tag{6.8}$$

$$M_{\tau_0}^2 = m_{\nu_0}^2 + m_{\mathrm{vis}_0}^2 + 2\sqrt{p_{\mathrm{vis}_0}^2 + m_{\mathrm{vis}_0}^2}\sqrt{p_{\nu_0}^2 + m_{\nu_0}^2} - 2p_{\mathrm{vis}_0}p_{\nu_0}\cos(d\theta) \tag{6.9}$$

$$M_{\tau_1}^2 = m_{\nu_1}^2 + m_{\mathrm{vis}_1}^2 + 2\sqrt{p_{\mathrm{vis}_1}^2 + m_{\mathrm{vis}_1}^2}\sqrt{p_{\nu_1}^2 + m_{\nu_1}^2} - 2p_{\mathrm{vis}_1}p_{\nu_1}\cos(d\theta). \tag{6.10}$$

For each leptonically decaying τ the invariant mass of the neutrino system $m_{\nu_i}^2$ is an additional unknown. This set of equations can no longer be solved analytically leading to the necessity of a numerical solution.

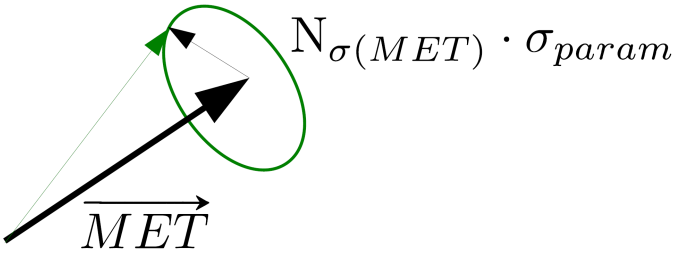


Figure 6.5: Schematic view of the constrain on the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector.

A Markov chain is used to scan the parameter space and a likelihood approach is used to find the best solution.

The parameter scanned by the Markov chain is primarily the distance in the azimuthal angle between the neutrino and visible decay products for each τ. The polar angle itself is not scanned because $E_{\mathrm{T}}^{\mathrm{miss}}$ cannot provide any information on the polar angle of the neutrinos. As $E_{\mathrm{T}}^{\mathrm{miss}}$ has a non-negligible resolution, the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector is scanned, i.e. it is varied in the x- and y-direction. Thus, the polar angle between the neutrino and the visible τ decay product is inherently varied. The code also features a τ energy scan. However, as previous studies showed no improvement using the τ energies, this scan is deactivated by default. For each randomly chosen point in this parameter space the above equations are solved and a weight is applied to that solution for the calculation of the likelihood. As the Markov chain would in principle also scan points that are not a physical solution the parameter space is constrained in order to save computing time. For one event the computing time is approximately ≤ 0.1 s [67]. Since only roughly 1% of all possible points have a physical solution [67] constraining the parameter space has a big impact on the overall computing time.

To constrain the parameter space $\Delta\phi\left(\tau_{\mathrm{vis},i}, \nu_i\right)$, $\mathrm{d}\theta$ and $\overrightarrow{E_{\mathrm{T}}^{\mathrm{miss}}}$ are only allowed to vary between specific values. $\Delta\phi\left(\tau_{\mathrm{vis},i}, \nu_i\right)$ can either be constrained to be ≤ 0.4 or by a function that is dependent on the τ decay type and the visible four-momentum. The default is $\Delta\phi\left(\tau_{\mathrm{vis},i}, \nu_i\right) \leq 0.4$.

The three dimensional angular distance $\mathrm{d}\theta$ can either be constrained to be between 0 and 10, which is the default, or again by function that depends on the type of the τ decay and the total τ four-momentum.

As $E_{\mathrm{T}}^{\mathrm{miss}}$ is not perfectly measured but rather has a non-negligible resolution, its resolution is parametrised, see 7. This parametrisation is used to define the boundaries in which $E_{\mathrm{T}}^{\mathrm{miss}}$ is scanned in its x- and y-component. The number of σ, where σ is defined by the parametrisation, can be freely chosen. This is illustrated in figure 6.5. The variation is allowed in a circular cone about the input vector with no preferred direction. If the τ energy scan is included the energy is allowed to vary upwards and downwards the same number of σ used for the $E_{\mathrm{T}}^{\mathrm{miss}}$ variation. The σ used for the τ energy variation is a τ energy resolution, TER, that is parametrised inside the MMC code.

Whenever one point in parameter space is sampled that fulfills these constraints, its solution is filled into a histogram with a specific weight that accounts for the probability of that point being the correct solution. This weight is composed of different sub-weights, i.e.: [67]

$$\text{weight} = P\left(\Delta\theta_0\right) \cdot P\left(\Delta\theta_1\right) \cdot P\left(\Delta E_{\mathrm{T,x}}^{\mathrm{miss}}\right) \cdot P\left(\Delta E_{\mathrm{T,y}}^{\mathrm{miss}}\right) \cdot P\left(R_0\right) \cdot P\left(R_1\right). \tag{6.11}$$

The underlying distributions used to parametrise this weight calculation can be seen in figure 6.6 for illustration purposes. Each distribution needs to be parametrised separately for the different τ decay types. The distribution in the middle shows the three dimensional angular distance between the visible τ and the

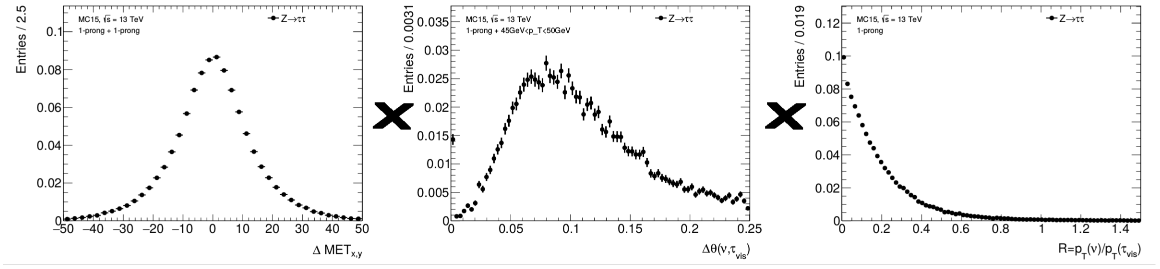


Figure 6.6: Distributions that are being parametrised for the weight calculation in the MMC for hadronically decaying τ's. Only for illustrating purposes.

neutrino or neutrino-system for leptonic decays, respectively. This also has to be separated into different ranges of τ p_{T} since the two particles are more collimated the higher the p_{T} of the τ's. For the angular distance a combination of a gaussian distribution and a Landau distribution is assumed and the fit of this distribution is the function that calculates the sub-weight. [68] The input values for the function are the τ decay type, the angle between the visible τ and the neutrino and total τ four-momentum.

The left distribution shows the parametrisation of the difference between the generator level $E_{\mathrm{T}}^{\mathrm{miss}}$ and the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ meaning it depends on the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution. This distribution is only parametrised as a function of the τ decay type and not as a function of the τ p_{T}. It is not expected to depend on the τ p_{T} specifically but rather on the overall activity in the event, see chapter 7.2. However, the τ decay type does have an impact on the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution as the type determines the number of neutrinos in the τ decay. The input for the weight calculation is the $E_{\mathrm{T}}^{\mathrm{miss}}$ that was measured and the new $E_{\mathrm{T}}^{\mathrm{miss}}$ obtained by the parameter scan. It is calculated for $E_{\mathrm{T,x}}^{\mathrm{miss}}$ and $E_{\mathrm{T,y}}^{\mathrm{miss}}$ separately. The resolution parametrisation is used for constraining the parameter space and for the weight calculation.

The right distribution shows the ratio between the neutrino(-system) energy and the visible τ energy. This parametrisation again is only separated into the different τ decay types.

For leptonic τ decays there is another parametrisation and sub-weight that is not shown here. This sub-weight parametrises the invariant mass of the neutrino-system. It has to be kept in mind that all parametrisations for the weight calculation are strongly dependent on the event selection.

When all found solutions are filled into histograms with their corresponding weight, there are three different possible outputs for the MMC:

1. MAXW: maximum weight
2. MLM: most likely mass
3. MLNU3P: most likely neutrino three-momentum.

The first output, MAXW, simply returns the solution with the highest weight which is equivalent to a maximised likelihood. It also provides an estimate of the three-momentum of the neutrinos.

The second output, MLM, returns the most likely mass by choosing the solution with the highest weight and scaling the obtained mass to the most likely mass. Due to this scaling of the mass it is no longer possible to give an estimate of the neutrinos' three-momenta as the solution of the equations have changed.

The third output, MLNU3P, fits the histograms filled with the solutions and searches for the maximum of these histograms. This maximum is returned which corresponds to the most likely three-momentum of the neutrinos. The resulting ditau-mass distribution for the MLM output and the Baseline selection can be seen in figure 6.7. This output is shown here and used throughout the thesis to discriminate $H \rightarrow \tau\tau$

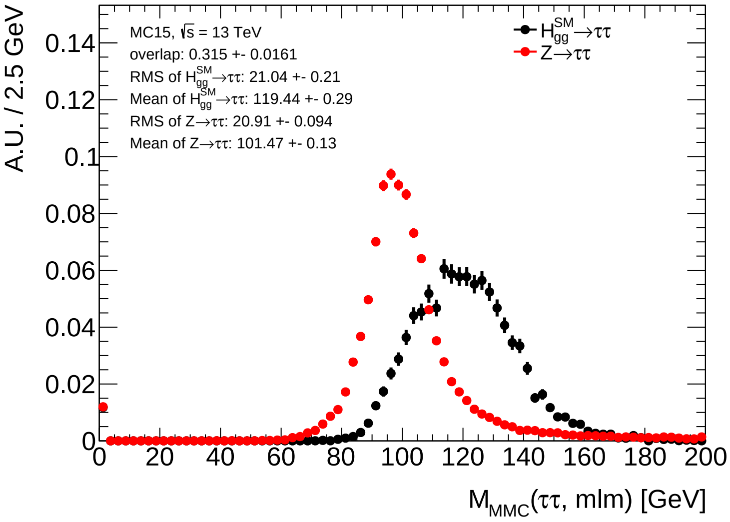


Figure 6.7: Distributions of the MMC mass for $H \to \tau\tau$ and $Z \to \tau\tau$. The Baseline selection is applied.

against $Z \to \tau\tau$ because it yields mass distributions that are the closest to the nominal mass peaks of all possible outputs. Furthermore, the overlap does not differ between the different outputs. The overlap using MMC is the smallest for all reconstruction techniques. More details on the MMC can be found in ref. [66] and ref. [67].

6.5 Comparison of the different techniques

A direct comparison of all mass reconstruction techniques for $H \to \tau\tau$ can be seen in figure 6.8. The Baseline selection has been applied. It shows that the ditau mass resolution is comparable when using the visible mass and the MMC. The mass resolution for the collinear approximation and M_T is broader in comparison. Considering the resolutions and the overlap for the visible mass (38%) and the MMC (31.5%), it becomes clear why m_{MMC} is used as the discriminating variable for $H \to \tau\tau$ against $Z \to \tau\tau$ in this thesis.

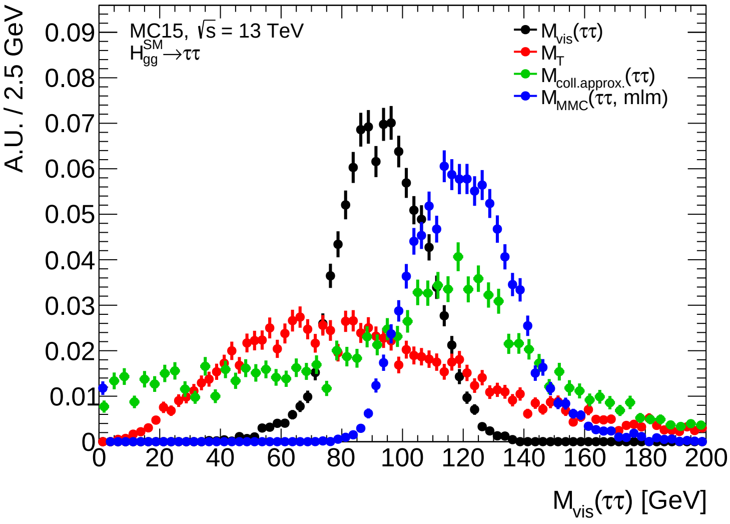


Figure 6.8: Mass distributions for the different reconstruction techniques for $H \to \tau\tau$. The Baseline selection is applied.

CHAPTER 7

MMC Tuning

As mentioned in chapter 6.4 the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution plays a major role in the performance of the MMC. Its parametrisation determines by how much $E_{\mathrm{T}}^{\mathrm{miss}}$ can vary during the scan of the parameter space and also influences the weight calculation for the $E_{\mathrm{T}}^{\mathrm{miss}}$ sub-weight. The Missing Mass Calculator was used for the first time in 2011, see ref. [6], and its $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution has been reparametrised since. Details on the reparametrisation can be found in ref. [69]. The parametrisation is also referred to as tune. Since this parametrisation is dependent on the object selection and event selection, a new parametrisation for the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution considering the object/event selection used in this thesis is derived in this chapter. Furthermore, the τ reconstruction has changed since Run 1 which possibly also affects the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution parametrisation. Only the $\tau_{\mathrm{had}}\,\tau_{\mathrm{had}}$ channel is considered in this chapter and for $E_{\mathrm{T}}^{\mathrm{miss}}$ the TST term is used. Two categories are used for the reparametrisation, one with jets and one without jets. This is done since the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution is expected to be different in the presence of jets compared to the absence of jets. For all following studies a $Z \to \tau\tau$ sample, a $\mathcal{CP}$-even $H_{gg} \to \tau\tau$ sample and a $\mathcal{CP}$-even $H_{VBF} \to \tau\tau$ sample are used separately, see section 4.2.1.

7.1 $E_{\mathrm{T}}^{\mathrm{miss}}$ modelling

First the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction used in the code is validated against externally provided distributions. In order to do so the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ terms are compared with reference terms which are provided by the Jet/EtMis-Group of the ATLAS collaboration for a specific object selection. The object selection close to the one of the Jet/EtMis-Group can be found in table 7.1 together with the object selection used in this thesis.

The resulting ΣE_{T} distributions for the different terms playing a role in $E_{\mathrm{T}}^{\mathrm{miss}}$ are shown in figure 7.1 for the process $H_{gg} \to \tau_{\mathrm{had}}\tau_{\mathrm{had}}$. These distributions are shown for the TauRec reconstruction algorithm. Figure A.2 in appendix A.2 shows the same distributions for the Substructure Reconstruction and are similar to the TauRec distributions.

As it can be seen the distributions for the jet terms and the τ terms show quite significant deviations between the reconstructed distribution and the reference distribution. This can be explained by the object selection applied. A local test with the object selection of the Jet/EtMis-Group and TauRec results in the distributions shown in figure 7.2. The only deviation is seen for the jet term because of an inherent cut on the jet p_{T} of $\geq 20\,\mathrm{GeV}$ in the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction software. The differences seen before are presumably caused by the applied p_{T} cuts on the objects. It is very likely that the τ selection plays an important role in this scenario since the jets used in the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction are all jets that are not classified as τ's in

Cut	Jet/EtMis-Group selection	thesis selection
photon selection	$p_T^\gamma > 10\,\text{GeV}$ `tight` ID / /	$p_T^\gamma > 10\,\text{GeV}$ `tight` ID allowed authors 1, 3 $\eta < 2.47$
electron selection	$p_T^e > 10\,\text{GeV}$ `medium` electron likelihood good object quality not in η crack of calo no muon overlap allowed authors 1, 3	$p_T^e > 15\,\text{GeV}$ `loose` electron likelihood good object quality not in η crack of calo no muon overlap allowed authors 1, 3
muon selection	$p_T^\mu > 2.5\,\text{GeV}$ `loose` quality accepted by MuonSelectionTool passed ID cuts	$p_T^\mu > 20\,\text{GeV}$ / accepted by MuonSelectionTool passed ID cuts
τ selection	$p_T^\tau > 20\,\text{GeV}$ accepted by TauSelectionTool trigger matched with trigger of table 5.1 `medium` τ-ID not passing electron `medium` BDT muon veto	$p_T^\tau > 18\,\text{GeV}$ accepted by TauSelectionTool trigger matched with trigger of table 5.1 `medium` τ-ID / /

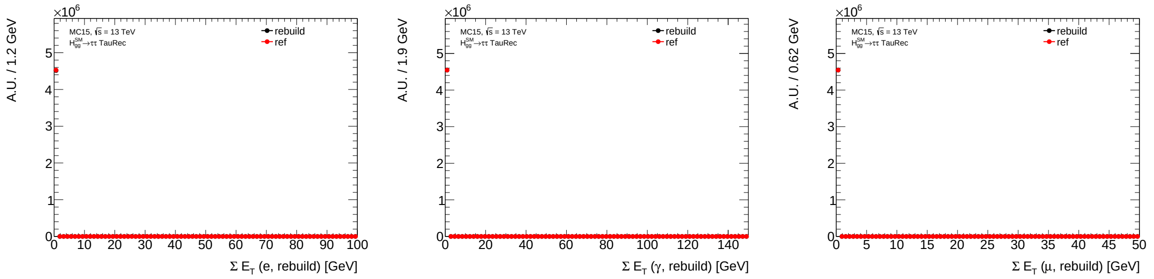
Table 7.1: Object selections of the JetEtMis-Group and the one used in this thesis for the E_T^{miss} reconstruction.

the object selection described in table 7.1.

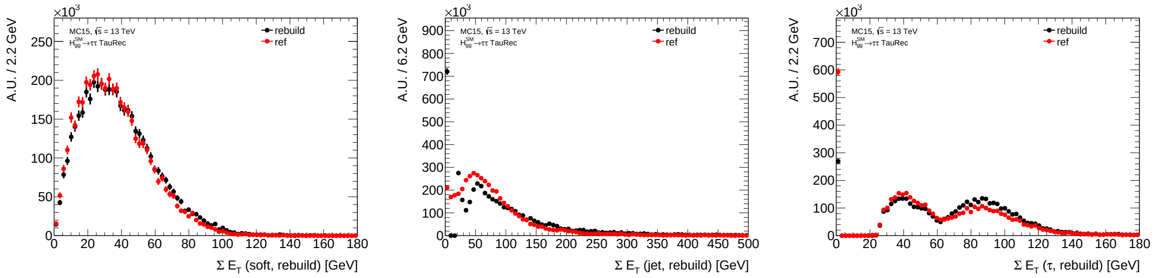
7.2 Parametrisation of the E_T^{miss} resolution as a function of $\sqrt{\Sigma E_T}$

The parametrisation of E_T^{miss} is done for two different categories, one category with jets and one category without jets. The event selection for both categories is listed in table 7.2. Only a very basic selection for the τ_{had}'s is applied and a cut that separates the two categories. The applied selection is inspired by table 5.1 and left at a very basic level in order to allow other analyses, like the $H \to \tau\tau$ coupling analysis, to use the derived tune. As E_T^{miss} is dependent on ΣE_T it makes sense to parametrise the E_T^{miss} resolution as a function of the latter. To do so, $\Delta E_T^{\text{miss}} = E_{\text{T,x(y),reco}}^{\text{miss}} - E_{\text{T,x(y),true}}^{\text{miss}}$ is plotted in 15 bins of $\sqrt{\Sigma E_T}$ with equal statistics in each bin for different samples, i.e. for $H_{gg} \to \tau_{\text{had}}\tau_{\text{had}}$, $H_{VBF} \to \tau_{\text{had}}\tau_{\text{had}}$ and $Z \to \tau_{\text{had}}\tau_{\text{had}}$. Since resolutions have a gaussian core, a Gaussian is fitted to their distributions in the range $[\text{Mean} - 1.5 \cdot \text{RMS}, \text{Mean} + 1.5 \cdot \text{RMS}]$ for each bin. An example of this fit can be seen in figure 7.3 for the last bin of $\sqrt{\Sigma E_T}$ in the $Z \to \tau_{\text{had}}\tau_{\text{had}}$ sample.

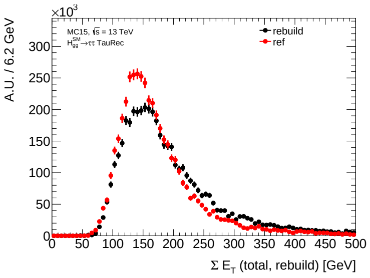
The width of these gaussian fits is then plotted against $\sqrt{\Sigma E_T}$ as shown in figure 7.4 for all three samples and the Substructure Reconstruction. Figure A.3 in appendix A.3 shows the parametrisations for the TauRec algorithm. The distributions are fitted with the function $f\left(\sqrt{\Sigma E_T}\right) = a + b * \log\left(\sqrt{\Sigma E_T} - c\right)$ [69] except for the no jets category of the $H_{VBF} \to \tau_{\text{had}}\tau_{\text{had}}$ sample. In this category and for this sample the statistics is too low for 15 bins. Which is expected, as there should only be events with jets in the $H_{VBF} \to \tau_{\text{had}}\tau_{\text{had}}$ sample per definition. The resulting distribution is fitted with a straight line even though the distribution is not meaningful. Although, a fit to this distribution is performed, the result is



(a) Electron term. (b) Photon term. (c) Muon term.



(d) Soft term. (e) Jet term. (f) Tau term.

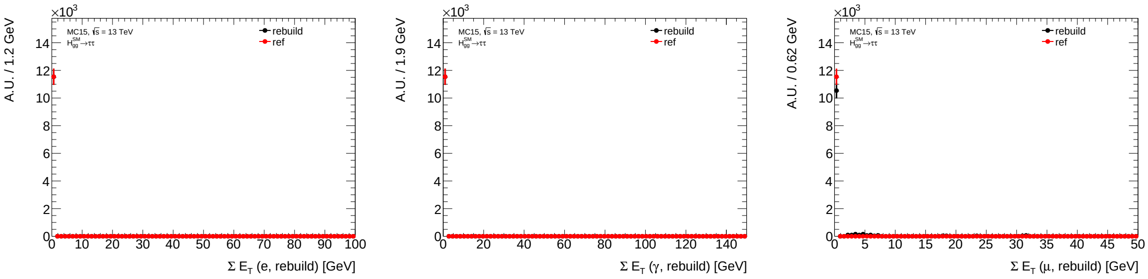


(g) Total term.

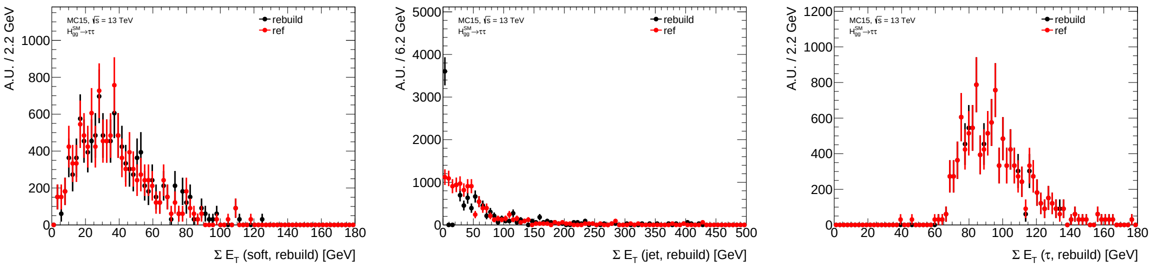
Figure 7.1: Comparison of the reconstructed terms used in the E_T^{miss} rebuilding and the reference terms provided by the Jet/EtMis-Group. The TauRec algorithm is used for the τ reconstruction.

Category	Cut
Both categories	ditau_tau0_HLT_tau35_medium1_tracktwo_tau25_medium1_tracktwo_L1TAU20IM_2TAU12IM!=0 ditau_tau1_HLT_tau35_medium1_tracktwo_tau25_medium1_tracktwo_L1TAU20IM_2TAU12IM!=0 truth matched τ's two `medium` identified and hadronic τ $q_{\tau_{0,1}} == 1 \&\& q_{\tau_0} \times q_{\tau_1} < 0$ $p_T^{\tau_0} > 40\,\text{GeV}$ and $p_T^{\tau_1} > 30\,\text{GeV}$
$2\tau_{\text{had}}$+no jets	$N_{jets} = 0$
$2\tau_{\text{had}}$+no jets	$N_{jets} > 0$

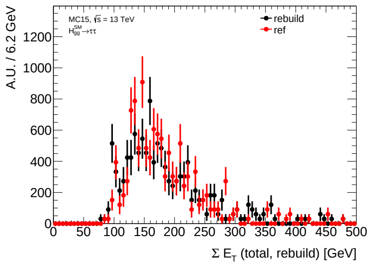
Table 7.2: Event selection for the two categories for the tuning of the MMC.



(a) Electron term.
(b) Photon term.
(c) Muon term.



(d) Soft term.
(e) Jet term.
(f) Tau term.



(g) Total term.

Figure 7.2: Comparison of the reconstructed terms used in the $E_{\mathrm{T}}^{\mathrm{miss}}$ rebuilding and the reference terms provided by the Jet/EtMis-Group with the Jet/EtMis-Group object selection. The TauRec algorithm is used for the τ reconstruction.

not used in the final parametrisation.

As the MMC is tuned to the Z mass peak by convention the parametrisation obtained from the $Z \to \tau\tau$ sample is used and reads:

$$\text{no jets:} \quad f\left(\sqrt{\Sigma E_{\mathrm{T}}}\right) = \begin{cases} -8.51 + 8.54 \cdot \log\left(\sqrt{\Sigma E_{\mathrm{T}}} - 3.97\right), & \text{if } \Sigma E_{\mathrm{T}} > 80\,\mathrm{GeV} \\ 5, & \text{if } \Sigma E_{\mathrm{T}} \leq 80\,\mathrm{GeV} \end{cases} \tag{7.1}$$

$$\text{with jets:} \quad f\left(\sqrt{\Sigma E_{\mathrm{T}}}\right) = \begin{cases} 1.2 + 5.62 \cdot \log\left(\sqrt{\Sigma E_{\mathrm{T}}} - 4.21\right), & \text{if } \Sigma E_{\mathrm{T}} > 115\,\mathrm{GeV} \\ 12.1, & \text{if } \Sigma E_{\mathrm{T}} \leq 115\,\mathrm{GeV} \end{cases} \tag{7.2}$$

When implementing the tune the functional dependence is only used down to a lower boundary. Values of ΣE_{T} below this boundary of 80 GeV in the no jets category and below 115 GeV in the jets category are assigned a constant $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution. The constant terms are obtained from figures 7.4(c) and 7.4(f) respectively. This prevents the resolution to become unrealistically small.

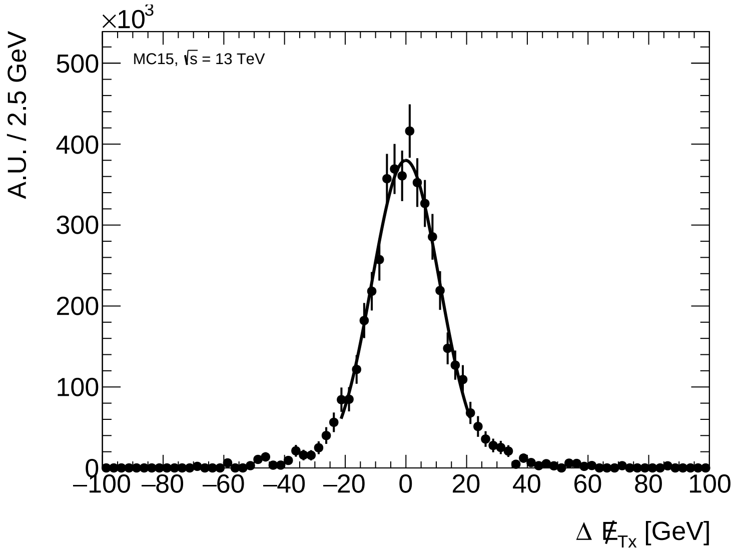
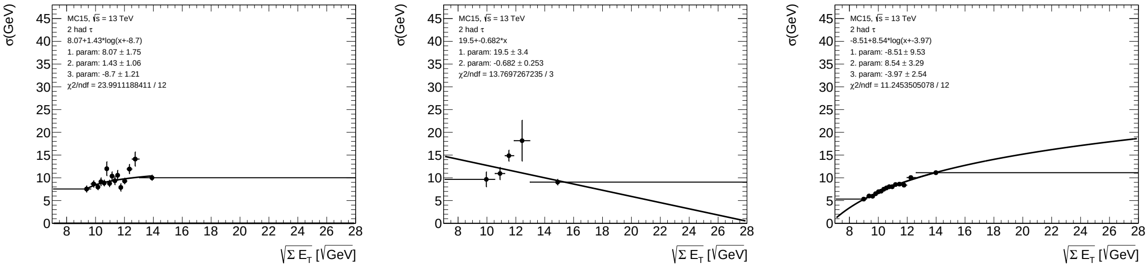


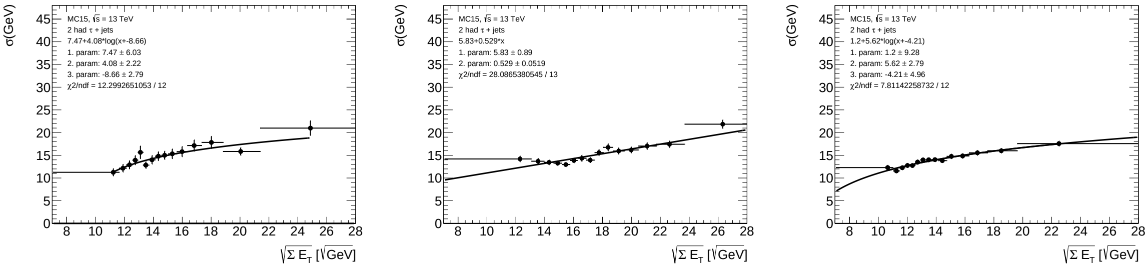
Figure 7.3: Gaussian fit of the $E_{\mathrm{T}}^{\text{miss}}$ resolution for the last $\sqrt{\Sigma E_{\mathrm{T}}}$ bin in the no jets category in the $Z \to \tau_{\text{had}}\tau_{\text{had}}$ sample. The Substructure Reconstruction is used for the τ reconstruction.



(a) $H_{gg} \to \tau_{\text{had}}\tau_{\text{had}}$ sample in the no jets category.

(b) $H_{VBF} \to \tau_{\text{had}}\tau_{\text{had}}$ sample in the no jets category.

(c) $Z \to \tau_{\text{had}}\tau_{\text{had}}$ sample in the no jets category.



(d) $H_{gg} \to \tau_{\text{had}}\tau_{\text{had}}$ sample in the jets category.

(e) $H_{VBF} \to \tau_{\text{had}}\tau_{\text{had}}$ sample in the jets category.

(f) $Z \to \tau_{\text{had}}\tau_{\text{had}}$ sample in the jets category.

Figure 7.4: $E_{\mathrm{T}}^{\text{miss}}$ resolution as a function of $\sqrt{\Sigma E_{\mathrm{T}}}$ for the different samples and categories. The Substructure Reconstruction is used for the τ reconstruction.

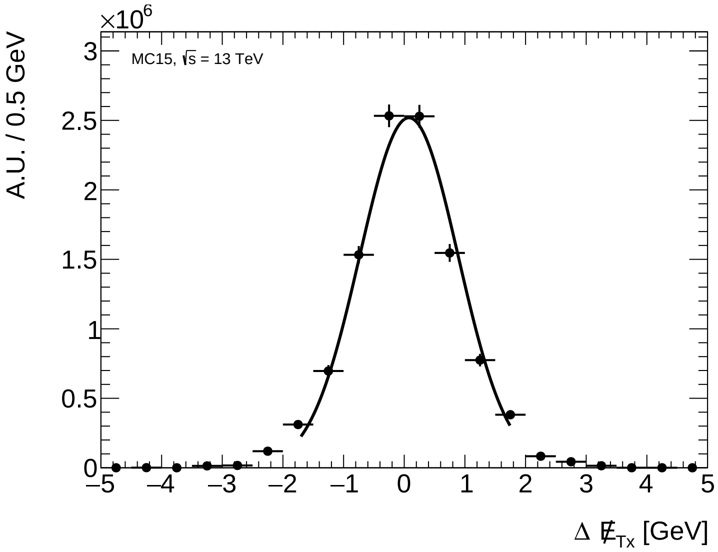


Figure 7.5: Gaussian fit of $\Delta E_{\mathrm{T}}^{\mathrm{miss}}$ for the last $\Delta\phi\left(\tau_{\mathrm{had}}, \tau_{\mathrm{had}}\right)$ bin in the no jets category for the $Z \rightarrow \tau_{\mathrm{had}}\tau_{\mathrm{had}}$ sample. The Substructure Reconstruction is used for the τ reconstruction.

7.3 Correction term for the parametrisation

In contrast to the previous implementation of the MMC, this tune includes a correction of the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution as a function of the angular distance between the two τ's, i.e. of $\Delta\phi\left(\tau_{\mathrm{had}}, \tau_{\mathrm{had}}\right)$. This correction term is added since the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution varies as a function of $\Delta\phi\left(\tau_{\mathrm{had}}, \tau_{\mathrm{had}}\right)$. Previously, this was studied by Antonio de Maria from the university of Göttingen. [69] The idea is to calculate the deviation of the expected resolution event by event:

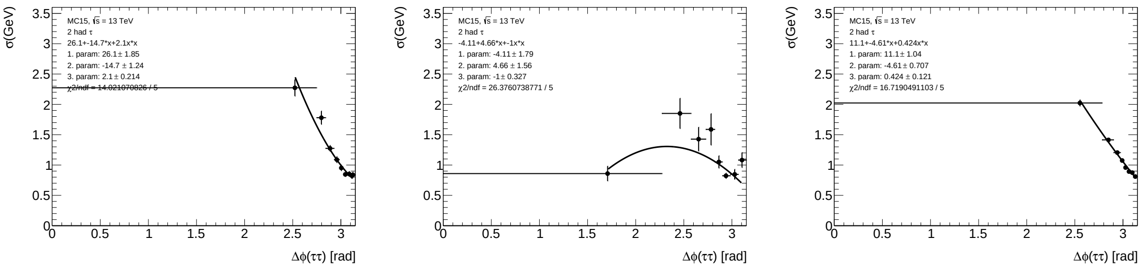
$$\Delta E_{\mathrm{T}}^{\mathrm{miss}} = \frac{E_{\mathrm{T,x(y),reco}}^{\mathrm{miss}} - E_{\mathrm{T,x(y),true}}^{\mathrm{miss}}}{f\left(\sqrt{\Sigma E_{\mathrm{T}}}\right)}. \tag{7.3}$$

With this new definition of $\Delta E_{\mathrm{T}}^{\mathrm{miss}}$ the same procedure as in section 7.2 is repeated. Here, $\Delta\phi\left(\tau_{\mathrm{had}}, \tau_{\mathrm{had}}\right)$ is divided into 8 bins of equal statistics for each sample and each bin is fitted with a Gaussian, see figure 7.5. The fit results for each bin are then plotted against $\Delta\phi\left(\tau_{\mathrm{had}}, \tau_{\mathrm{had}}\right)$ and a polynomial function of second order [69] is fitted as shown in figure 7.6 for all three samples and the Substructure Reconstruction.

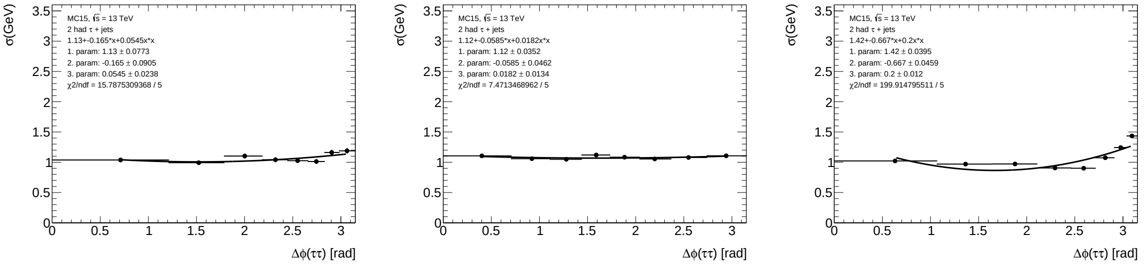
In order to avoid undefined behaviour the fitted function is only used down to a lower boundary of $\Delta\phi\left(\tau_{\mathrm{had}}, \tau_{\mathrm{had}}\right) = 2.5$ in the no jets category and $\Delta\phi\left(\tau_{\mathrm{had}}, \tau_{\mathrm{had}}\right) = 0.6$ in the jets category. The parametrisation is again derived from the $Z \rightarrow \tau\tau$ sample as the MMC is tuned to the Z mass peak the corresponding distributions are shown in figure 7.6(c) and figure 7.6(f). The correction term of the tune in the two categories read:

$$\text{no jets:} \quad f\left(\Delta\phi\left(\tau_{\mathrm{had}}, \tau_{\mathrm{had}}\right)\right) = \begin{cases} 11.1 - 4.61 \cdot x + 0.424 \cdot x^2, & \text{if } \Delta\phi\left(\tau_{\mathrm{had}}\tau_{\mathrm{had}}\right) > 2.5 \\ 2, & \text{if } \Delta\phi\left(\tau_{\mathrm{had}}\tau_{\mathrm{had}}\right) \leq 2.5 \end{cases} \tag{7.4}$$

$$\text{with jets:} \quad f\left(\Delta\phi\left(\tau_{\mathrm{had}}, \tau_{\mathrm{had}}\right)\right) = \begin{cases} 1.42 - 0.667 \cdot x + 0.2 \cdot x^2, & \text{if } \Delta\phi\left(\tau_{\mathrm{had}}\tau_{\mathrm{had}}\right) > 0.6 \\ 1.02, & \text{if } \Delta\phi\left(\tau_{\mathrm{had}}\tau_{\mathrm{had}}\right) \leq 0.6 \end{cases} \tag{7.5}$$



(a) $H_{gg} \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ sample in the no jets category. (b) $H_{VBF} \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ sample in the no jets category. (c) $Z \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ sample in the no jets category.



(d) $H_{gg} \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ sample in the jets category. (e) $H_{VBF} \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ sample in the jets category. (f) $Z \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ sample in the jets category.

Figure 7.6: $\Delta E_{\text{T}}^{\text{miss}}$ as a function of $\Delta\phi(\tau_{\text{had}}, \tau_{\text{had}})$ for the different samples and categories. The Substructure Reconstruction is used for the τ reconstruction.

The complete parametrisation for the τ_{had} τ_{had} channel was included in the DiTauMassTools package in the tag 00-00-24 which is also part of the AnalysisBase release 2.4.14.

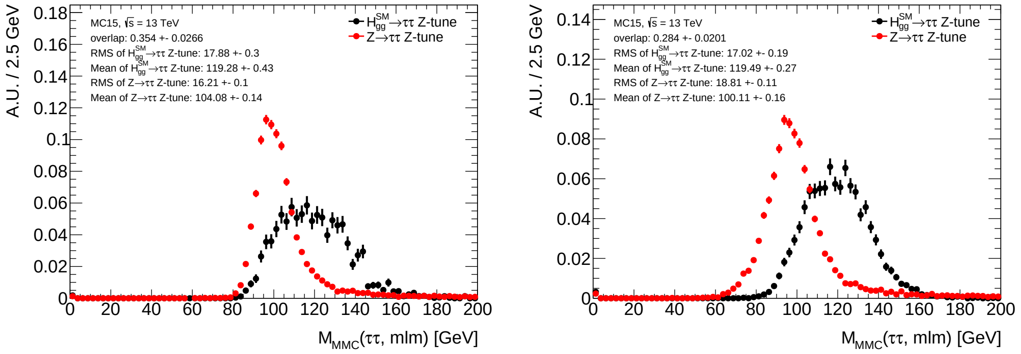
7.4 Results using the new Substructure Reconstruction

The resulting mass distributions (MLM output) for $H \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ and $Z \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ are shown in figure 7.7 for the Substructure Reconstruction. The corresponding distributions for TauRec are shown in figure A.5 in appendix A.4. In order to compare these results with the old tune the overlap of $Z \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ and $H \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ for the old tune is shown in figure 7.8. All results are shown in the no jets and jets categories defined for the derivation of the tune. Within the statistical uncertainties both tunes perform comparable. Even though, the τ resolution improved with the new reconstruction, the MMC performance did not. Nevertheless, a new tune is needed since it is not obvious how the new τ reconstruction affects the Missing Mass Calculator.

7.5 Possibilities for further improvements of the MMC

The performed new parametrisation of the $E_{\text{T}}^{\text{miss}}$ resolution influences by how much $E_{\text{T}}^{\text{miss}}$ can vary when scanning the parameter space and additionally the weight calculation for the $E_{\text{T}}^{\text{miss}}$ sub-weight. Since this new parametrisation alone did not result in an improved performance of the MMC it is interesting what else can be done to possibly improve the result.

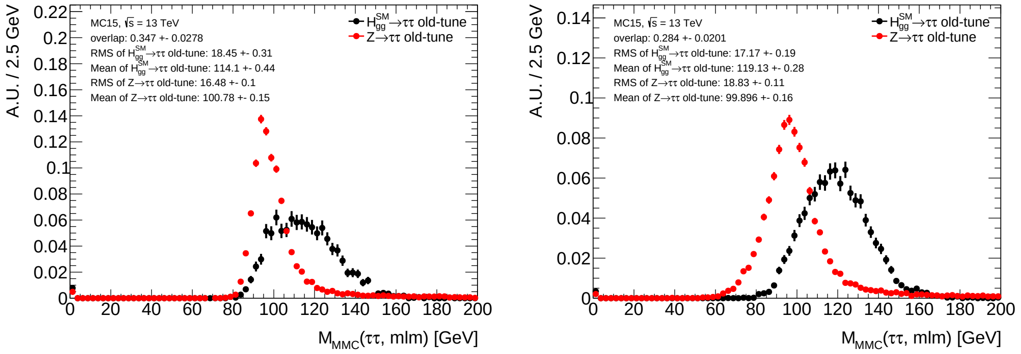
Directly related to the performed parametrisation is the possibility of parametrising the $E_{\text{T}}^{\text{miss}}$ resolution separately along the axis defined by the highest p_{T} jet and perpendicular to this axis. Due to the



(a) No jets category.

(b) Jets category.

Figure 7.7: Mass distributions for $H \to \tau_{\text{had}}\tau_{\text{had}}$ and $Z \to \tau_{\text{had}}\tau_{\text{had}}$ with the new tune for the τ_{had} τ_{had} channel. The Substructure Reconstruction is used for the τ_{had} reconstruction and the MLM output is shown.



(a) No jets category.

(b) Jets category.

Figure 7.8: Mass distributions for $H \to \tau_{\text{had}}\tau_{\text{had}}$ and $Z \to \tau_{\text{had}}\tau_{\text{had}}$ with the old tune for the τ_{had} τ_{had} channel. The Substructure Reconstruction is used for the τ_{had} reconstruction and the MLM output is shown.

reconstruction algorithm of $E_{\mathrm{T}}^{\text{miss}}$ it is highly sensitive to mismeasurements of the jet energy and hence the resolution of $E_{\mathrm{T}}^{\text{miss}}$ is different for the aforementioned axes.

The new Substructure Reconstruction still has more potential that can be used than what has been done up to now as it provides a five-way decay mode classification. Thus, it might be useful to make new parametrisations for all distributions relevant for the weight-calculation and maybe even divide these into the different decay modes. Lukas Weise, a former master's student from Göttingen [70], already studied a new parametrisation of the angular distance between the τ_{had}'s and their neutrinos with the Substructure Reconstruction. This parametrisation resulted in a shift of the mass distributions closer to the nominal mass peaks but did not improve the width of the distributions.

The separation between $Z \to \tau\tau$ and $H \to \tau\tau$ might increase, and maybe even the mass resolution, if all sub-weights of the MMC were reparametrised using the same τ reconstruction. Also, it is possible that the τ energy scan helps with the new improved τ resolution.

Another possibility might be to include the spin of the H and Z boson in the MMC calculation. This could be used for example for two different neutrino distribution parametrisations in the weight calculation which are then used for two separate runs of the MMC and the result with the highest weight could be used as the result. Thus, the separation between $Z \to \tau\tau$ and $H \to \tau\tau$ might improve. [71]

CHAPTER **8**

Impact of selection cuts on φ^*_{CP}

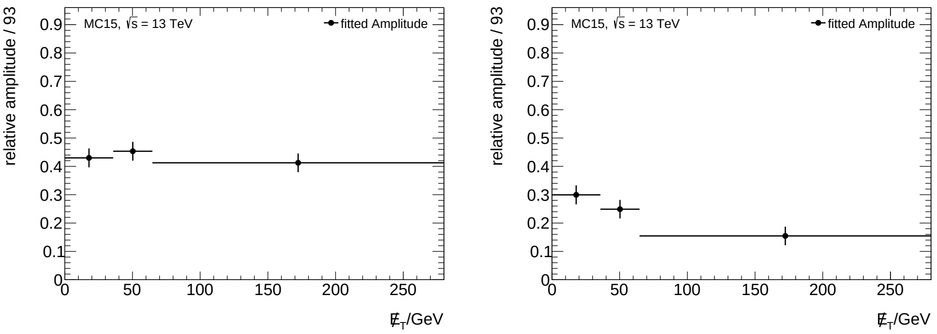
In the following the effects of object and event selection requirements, i.e. of different phase space regions, on the φ^*_{CP} distribution are studied. For all studies in this section the reconstructed decay mode and the generated decay mode are required to be `1p1n`. To study the effects, all cuts of the event selection region, e.g. the Preselection, that the variable under investigation is part of are applied. Except for the cut on that variable being studied. In the following this will be referred to as $N-1$ principle. The distribution of that variable is split into three regions of equal statistics. In each region the relative amplitude and phase shift of φ^*_{CP} are studied by fitting the φ^*_{CP} distribution with a cosine function and extracting the relevant information from the fit, see section 5. Additionally, the difference between generator level and reconstruction level is investigated, i.e. the reconstruction effects responsible for e.g. a drop in the amplitude are searched for. These kinds of studies can unveil sources of systematic biases or uncertainties on the measurement in the different phase space regions defined by the event selection. Thus, the results can be used to determine whether it is useful to change a cut even if it decreases the efficiency of the background suppression. Furthermore, these studies give insight into the sources of possible systematic biases and hence can be used to find possibilities for further improvements of the measurement either in the theory of the observables itself or specific components of the event reconstruction. In order to increase the significance of the results different samples are combined as described in section 4.2.1.

8.1 Effects of event selection variables

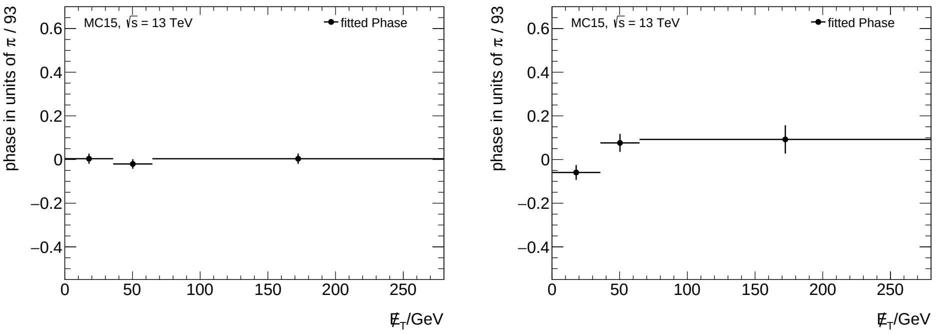
All variables that are listed in table 5.1 have been investigated with the procedure described above. Due to a lack of statistics for this study in the signal regions, the signal region variables have been studied after requiring the Preselection cuts not following the $N-1$ principle. In the following the variables that showed the strongest dependence on the phase space or the most significant reconstruction effects will be discussed. The other variables did not show any significant impact on φ^*_{CP}. For completeness the φ^*_{CP} distributions, fit probabilites, relative amplitudes and phase shifts are shown for the studied variables in appendix A.5. Additionally, the information of the variables discussed in this section that are not shown, like the fit probabilities, can be found in appendix A.5.

8.1.1 $E_{\mathrm{T}}^{\mathrm{miss}}$

The most significant degradation of the relative amplitude in φ^*_{CP} at reconstruction level with respect to generator level is observed with increasing values of $E_{\mathrm{T}}^{\mathrm{miss}}$. Figure 8.1 shows the relative amplitude



(a) Relative amplitude at generator level. (b) Relative amplitude at reconstruction level.



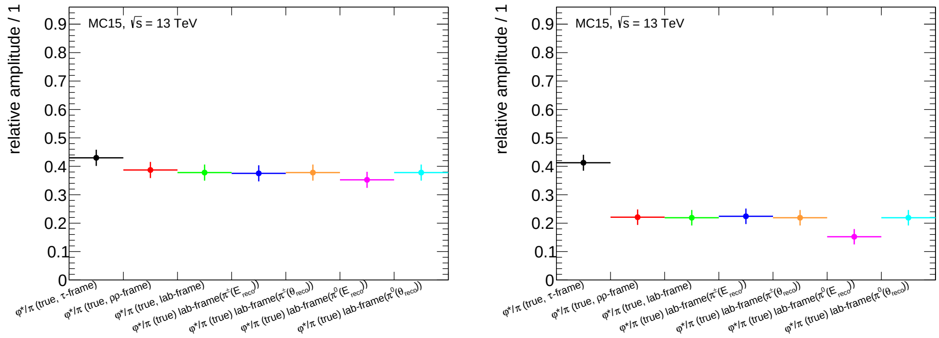
(c) Phase shift at generator level. (d) Phase shift at reconstruction level.

Figure 8.1: Comparison of the dependence of φ^*_{CP} at generator level and reconstruction level as a function of E_T^{miss}. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.

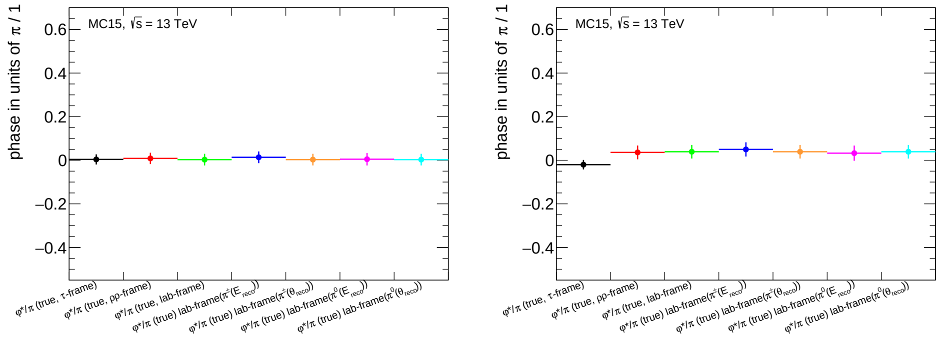
and the phase shift as a function of E_T^{miss}. To spot the part of the method that is influenced the most by reconstruction effects for high E_T^{miss} the calculation introduced in chapter 2.5 is split into separate parts. One part only considers the reconstruction effects for the $y_0 y_1$ calculation and leaves the decay plane reconstruction and calculation of the triple correlation at generator level while the other part does the opposite. As discussed previously it is not possible to reconstruct the neutrinos in the τ_{had} decays. Thus, it is not possible to use the respective τ_{had} frame for the $y_0 y_1$ calculation at reconstruction level and the laboratory frame is used instead.

Figure 8.2 shows the reason for the drop in the relative amplitude. For increasing E_T^{miss} the difference between the total τ_{had} four-momentum and the visible τ_{had} four-momentum increases. Hence, the approximations by other frames that don't include the neutrino four-momentum get worse and the relative amplitude decreases as a consequence. This explains why the relative amplitude degrades at reconstruction level but not at generator level. The phase shift for high E_T^{miss} cannot be explained by the separate reconstruction effects in the separate parts of the φ^*_{CP} calculation. Figure 8.3 shows the phase shift for the reconstruction effects in the decay plane reconstruction. Considering the phase shifts shown in figure 8.2(d-f) for the $y_0 y_1$ calculation, these effects are not able to explain the phase shift seen in figure 8.1(d). The combination of the decay plane reconstruction and the $y_0 y_1$ calculation shifting the φ^*_{CP} distribution is presumably responsible for the observed effect.

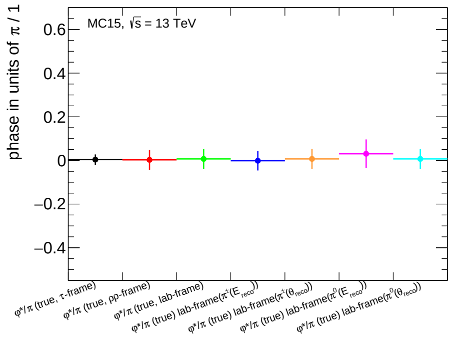
As the frame which is used for the $y_0 y_1$ calculation has the strongest impact on φ^*_{CP} compared to other reconstruction effects, the effects of using other frames than the laboratory frame have been studied. Figure 8.4 shows the relative amplitude for different reconstructed frames for the $y_0 y_1$ calculation while leaving everything else of the φ^*_{CP} calculation at generator level and the generated respective τ_{had}



(a) Relative amplitude for $E_{\mathrm{T}}^{\mathrm{miss}} < 35.8\,\mathrm{GeV}$. (b) Relative amplitude for $E_{\mathrm{T}}^{\mathrm{miss}} \geq 64.7\,\mathrm{GeV}$.

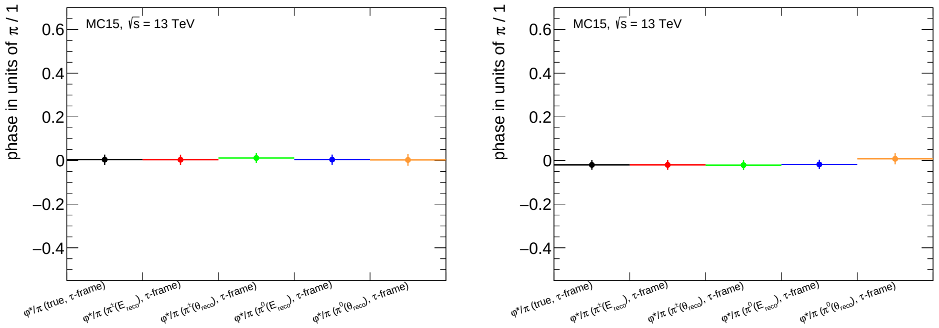


(c) Phase shift for $E_{\mathrm{T}}^{\mathrm{miss}} < 35.8\,\mathrm{GeV}$. (d) Phase shift for $35.8\,\mathrm{GeV} \leq E_{\mathrm{T}}^{\mathrm{miss}} < 64.7\,\mathrm{GeV}$.



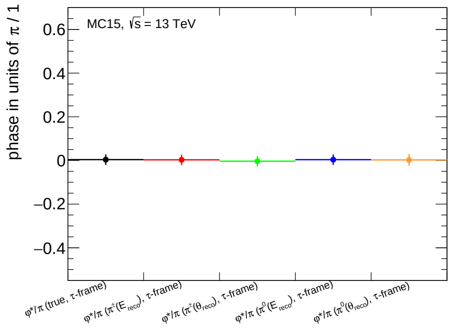
(e) Phase shift for $E_{\mathrm{T}}^{\mathrm{miss}} \geq 64.7\,\mathrm{GeV}$.

Figure 8.2: Effects of different frames for the $y_0 y_1$ calculation and reconstruction effects in the same on the relative amplitude and the phase shift in the different $E_{\mathrm{T}}^{\mathrm{miss}}$ regions.



(a) Phase shift for $E_{\mathrm{T}}^{\mathrm{miss}} < 35.8\,\mathrm{GeV}$.

(b) Phase shift for $35.8\,\mathrm{GeV} \leq E_{\mathrm{T}}^{\mathrm{miss}} < 64.7\,\mathrm{GeV}$.

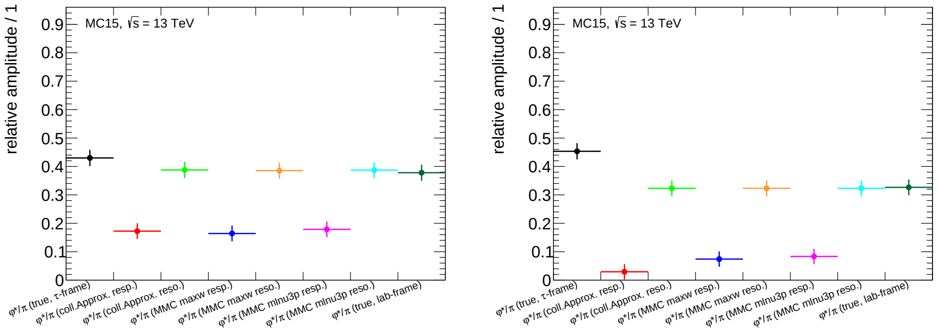


(c) Phase shift for $E_{\mathrm{T}}^{\mathrm{miss}} \geq 64.7\,\mathrm{GeV}$.

Figure 8.3: Effects of reconstruction effects in the decay plane reconstruction and triple correlation calculation on the relative amplitude and the phase shift in the different $E_{\mathrm{T}}^{\mathrm{miss}}$ regions.

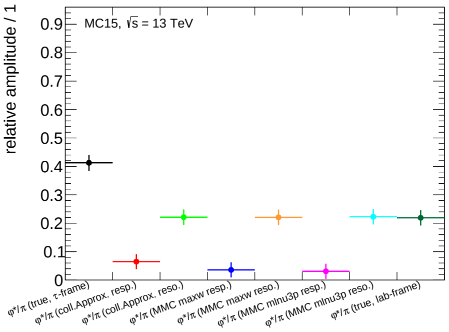
frame as a reference. The collinear approximation and the different outputs of the MMC, that provide neutrino information, have been used to approximate the respective τ_{had} frames (resp.) and the resonance frame (reso.). As it can be seen, none of the investigated frames show significant improvement over the laboratory frame when using generator level quantities for all other parts of the φ^*_{CP} calculation. Figure 8.5 shows the relative amplitudes and phase shifts for the different reconstructed frames with all parts of the φ^*_{CP} calculation at reconstruction level. Although the relative amplitudes and phase shifts for the resonance frames are compatible to the laboratory frame within 1σ, a systematic trend for the resonance frames obtained from the MMC and the Collinear Approximation towards higher relative amplitudes and smaller phase shifts can be observed. Hence, the resonance frame obtained from the Collinear Approximation might be the frame to chose compared to the laboratory frame and more sophisticated methods like the MMC.

The frame reconstruction is the main cause of a degradation of the amplitude for most variables. Thus, the approximation of the τ_{had} frames with the Collinear Approximation and the MMC has been studied for all variables. The results are similar to the results shown in figure 8.4 when the other parts of the φ^*_{CP} calculation are calculated at generator level. All in all, it has to be studied whether the cut on $E_{\mathrm{T}}^{\mathrm{miss}}$ can be loosened in a way that the background suppression is still efficient enough to not introduce another source of degradation in the φ^*_{CP} performance.



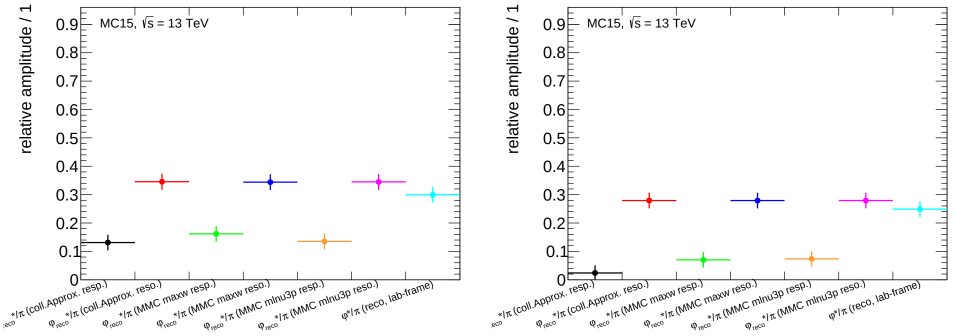
(a) Different reconstructed frames for $E_{\mathrm{T}}^{\mathrm{miss}} <$ 35.8 GeV.

(b) Different reconstructed frames for 35.8 GeV $\leq E_{\mathrm{T}}^{\mathrm{miss}} <$ 64.7 GeV.



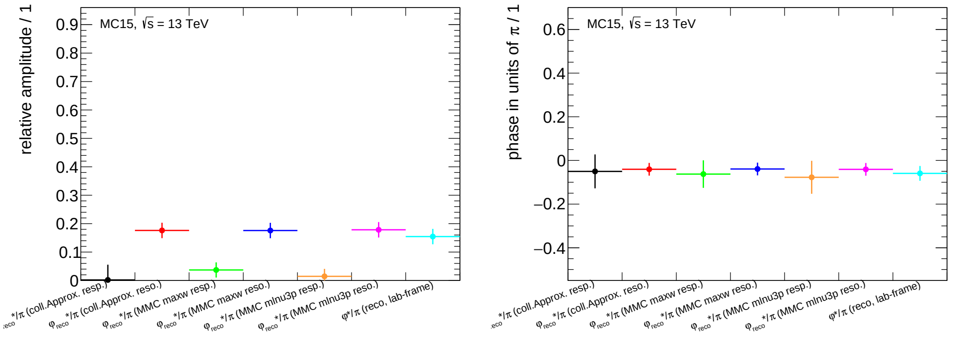
(c) Different reconstructed frames for $E_{\mathrm{T}}^{\mathrm{miss}} \geq$ 64.7 GeV.

Figure 8.4: Relative amplitudes using different reconstructed frames for the $y_0 y_1$ calculation while leaving everything else at generator level. The first bin shows the relative amplitude when using the respective τ_{had} frame as reference.



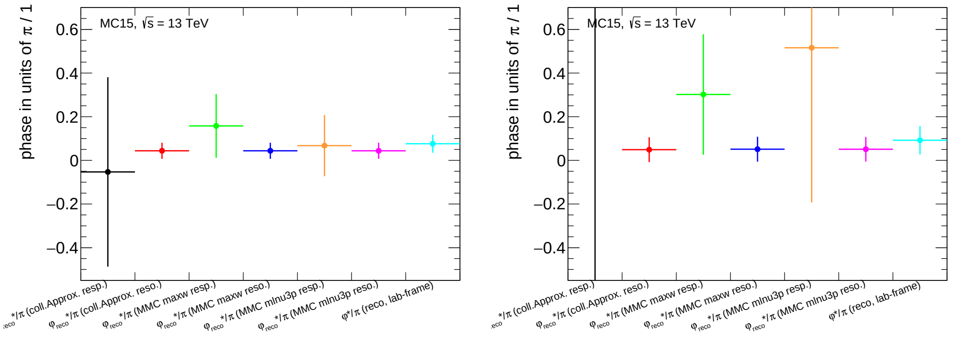
(a) Relative amplitudes for different reconstructed frames for $E_{\mathrm{T}}^{\mathrm{miss}} < 35.8\,\mathrm{GeV}$.

(b) Relative amplitudes for different reconstructed frames for $35.8\,\mathrm{GeV} \leq E_{\mathrm{T}}^{\mathrm{miss}} < 64.7\,\mathrm{GeV}$.



(c) Relative amplitudes for different reconstructed frames for $E_{\mathrm{T}}^{\mathrm{miss}} \geq 64.7\,\mathrm{GeV}$.

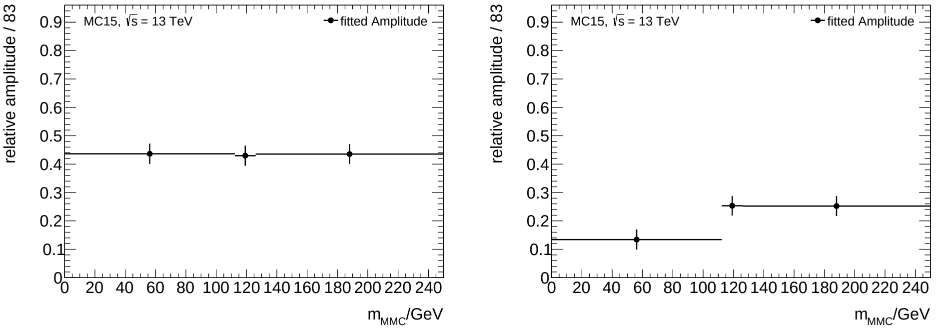
(d) Phase shifts for different reconstructed frames for $E_{\mathrm{T}}^{\mathrm{miss}} < 35.8\,\mathrm{GeV}$.



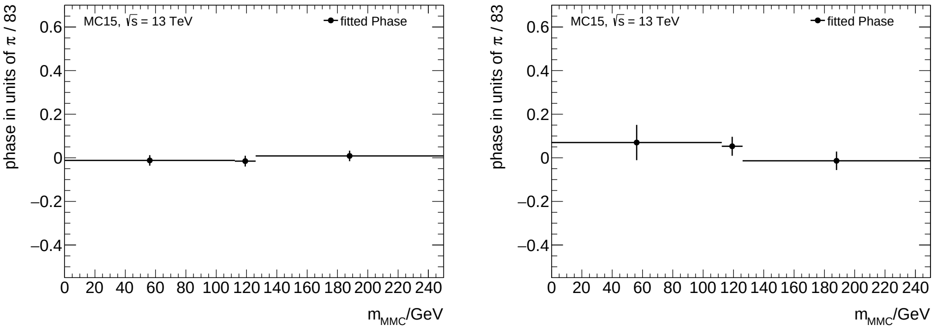
(e) Phase shifts for different reconstructed frames for $35.8\,\mathrm{GeV} \leq E_{\mathrm{T}}^{\mathrm{miss}} < 64.7\,\mathrm{GeV}$.

(f) Phase shifts for different reconstructed frames for $E_{\mathrm{T}}^{\mathrm{miss}} \geq 64.7\,\mathrm{GeV}$.

Figure 8.5: Using different reconstructed frames for the $y_0 y_1$ calculation with reconstructed information for all parts of the φ^*_{CP} calculation.



(a) Relative amplitude at generator level. (b) Relative amplitude at reconstruction level.

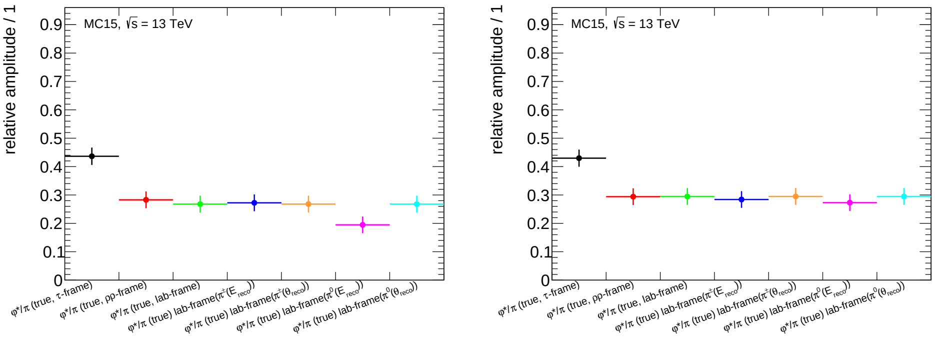


(c) Phase shift at generator level. (d) Phase shift at reconstruction level.

Figure 8.6: Comparison of the dependence of φ^*_{CP} at generator level and reconstruction level as a function of m_{MMC}. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.

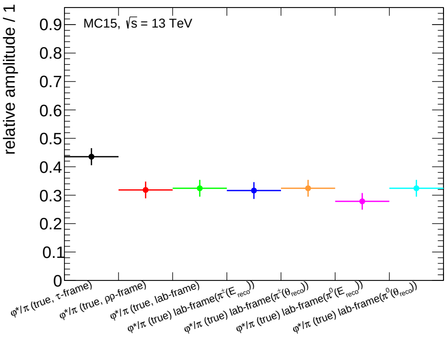
8.1.2 m_{MMC}

The MLM mass output of the MMC is used to discriminate between $Z \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ and $H \rightarrow \tau_{\text{had}}\tau_{\text{had}}$. As it is part of both signal regions, the Preselection is applied but no cuts from either signal region. m_{MMC} is always calculated from the reconstructed quantities. Figure 8.6 shows the opposite trend as $E_{\text{T}}^{\text{miss}}$ in the relative amplitude which is again caused by a worse frame approximation for the $y_0 y_1$ calculation and the reconstruction of the neutral pion energy, see figure 8.7. For low values of m_{MMC} this worse frame reconstruction is presumably caused by a more significant role of the neutrinos in the τ_{had} decays. The phase shift as a function of m_{MMC} is consistent with zero at reconstruction level and generator level as shown in figure 8.6. To exploit this dependence of the relative amplitude of φ^*_{CP} the required mass window in the event selection already seems reasonable.



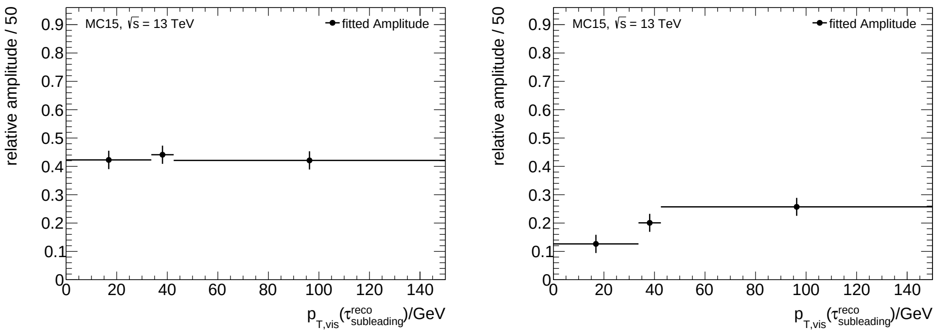
(a) Phase shift for $m_{MMC} < 112.3\,\text{GeV}$.

(b) Phase shift for $112.3\,\text{GeV} \leq m_{MMC} < 126\,\text{GeV}$.

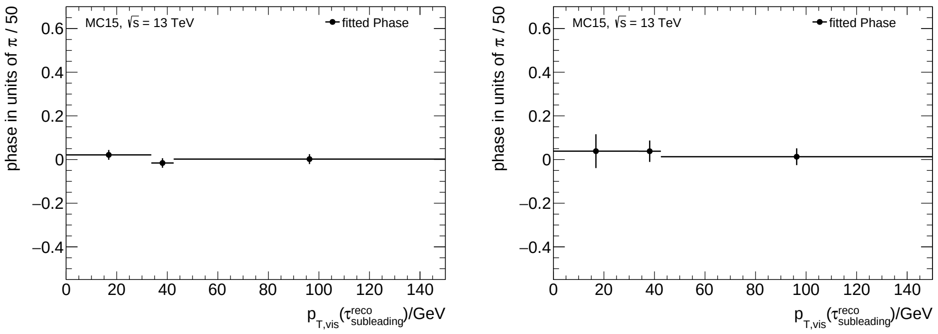


(c) Phase shift for $m_{MMC} \geq 126\,\text{GeV}$.

Figure 8.7: Effects of different frames for the $y_0 y_1$ calculation and reconstruction effects in the same on the relative amplitude and the phase shift in the different m_{MMC} regions.



(a) Relative amplitude at generator level. (b) Relative amplitude at reconstruction level.

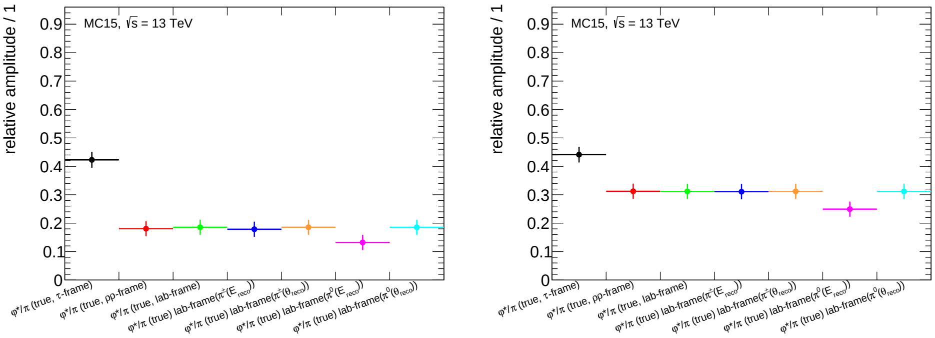


(c) Phase shift at generator level. (d) Phase shift at reconstruction level.

Figure 8.8: Comparison of the dependence of φ^*_{CP} at generator level and reconstruction level as a function of p_{T} of the subleading τ_{had}. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.

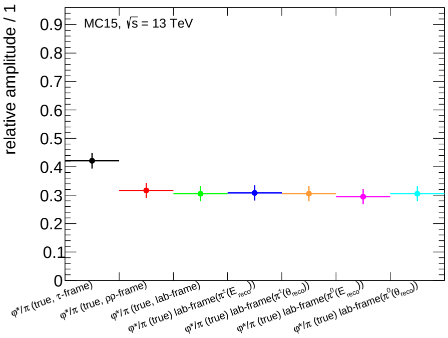
8.1.3 p_{T} of the subleading τ_{had}

Figure 8.8 shows that a similar trend to the one seen for m_{MMC} can be observed for the p_{T} of the subleading τ_{had}. This behaviour is also caused by the frame reconstruction, see figure 8.9. Increasing p_{T} of the subleading τ_{had} leaves less energy for the neutrinos. Hence, the frame approximation with the laboratory frame gets closer to the respective τ_{had} frame. The phase shift is again consistent with zero at reconstruction level and generator level. As it can be seen in table 5.1, a cut on the p_{T} of the subleading τ_{had} is already part of the event selection. To further exploit the studied dependence of φ^*_{CP}, a reweighting of the events could be considered.



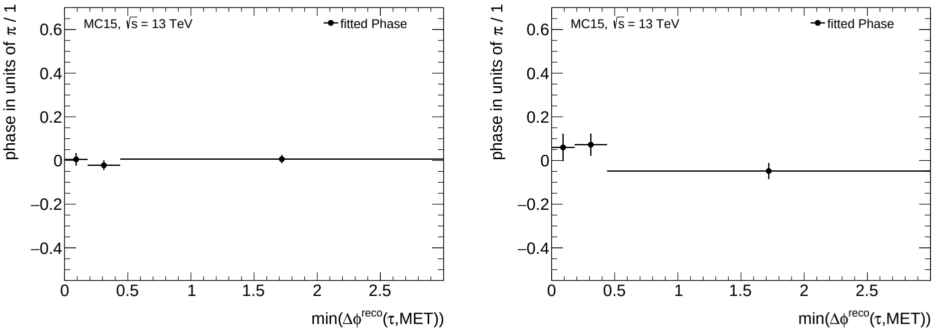
(a) Phase shift for $p_{\mathrm{T}}^{\tau_{sub}} < 33.7\,\mathrm{GeV}$.

(b) Phase shift for $33.7\,\mathrm{GeV} \leq p_{\mathrm{T}}^{\tau_{sub}} < 42.5\,\mathrm{GeV}$.



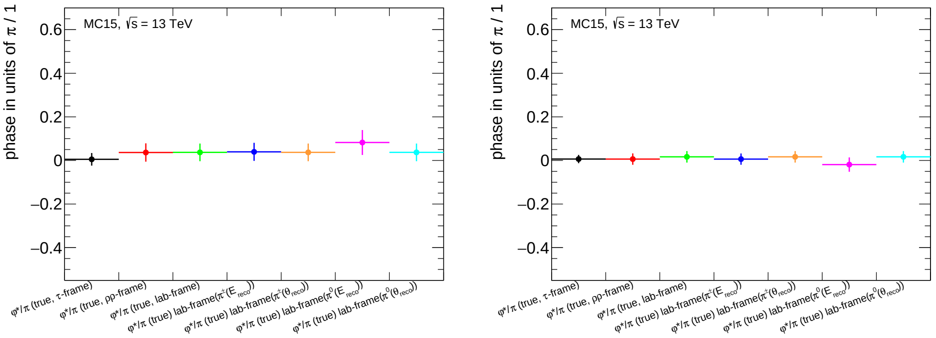
(c) Phase shift for $p_{\mathrm{T}}^{\tau_{sub}} \geq 42.5\,\mathrm{GeV}$.

Figure 8.9: Effects of different frames for the $y_0 y_1$ calculation and reconstruction effects in the same on the relative amplitude and the phase shift in the different regions of the subleading τ_{had} p_{T}.



(a) Phase shift at generator level. (b) Phase shift at reconstruction level.

Figure 8.10: Comparison of the dependence of φ^*_{CP} at generator level and reconstruction level as a function of $\min\left(\Delta\phi\left(\tau_{\mathrm{had}}, E_{\mathrm{T}}^{\mathrm{miss}}\right)\right)$. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.



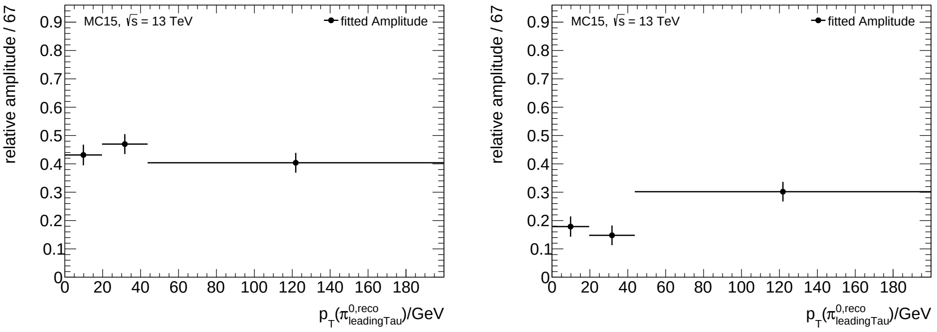
(a) Phase shift for $\min\left(\Delta\phi\left(\tau_{\mathrm{had}}, E_{\mathrm{T}}^{\mathrm{miss}}\right)\right)<0.18$. (b) Phase shift for $\min\left(\Delta\phi\left(\tau_{\mathrm{had}}, E_{\mathrm{T}}^{\mathrm{miss}}\right)\right)>0.44$.

Figure 8.11: Effects of different frame for the $y_0 y_1$ calculation and reconstruction effects in the same on the relative amplitude and the phase shift in the different $\min\left(\Delta\phi\left(\tau_{\mathrm{had}}, E_{\mathrm{T}}^{\mathrm{miss}}\right)\right)$ regions.

8.1.4 $\min\left(\Delta\phi\left(\tau_{\mathrm{had}}, E_{\mathrm{T}}^{\mathrm{miss}}\right)\right)$

As listed in table 5.1 the $E_{\mathrm{T}}^{\mathrm{miss}}$-$\phi$-centrality cut is composed of two cuts of which at least one has to be fulfilled. Figure 8.10 shows the phase shift as a function of $\min\left(\Delta\phi\left(\tau_{\mathrm{had}}, E_{\mathrm{T}}^{\mathrm{miss}}\right)\right)$.

The cause of this phase shift is the frame reconstruction and the reconstructed neutral pion energy in the $y_0 y_1$ calculation, see figure 8.11.



(a) Relative amplitude at generator level. (b) Relative amplitude at reconstruction level.

Figure 8.12: Comparison of the dependence of φ^*_{CP} at generator level and reconstruction level as a function of the neutral pion p_T of the leading τ_{had}. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.

8.2 Effects of other variables

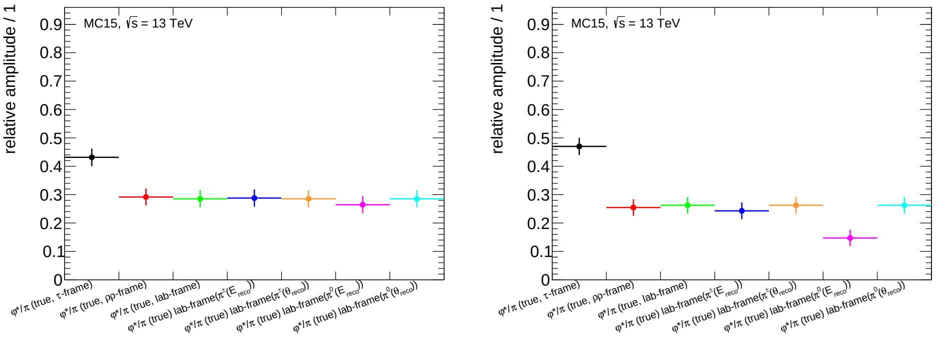
The variables discussed in this section are not part of the event selection. To see how they would behave if they were part of the event selection, the Preselection is applied in the following studies.

8.2.1 Neutral pion p_T of the leading τ_{had}

The neutral pion p_T is not part of the event selection inspired by the Higgs coupling analysis, partly because it was not available with TauRec. Figure 8.12 shows that the relative amplitude increases with increasing values of the neutral pion p_T of the leading τ_{had}, motivating a possible new cut to enhance the performance of the measurement.

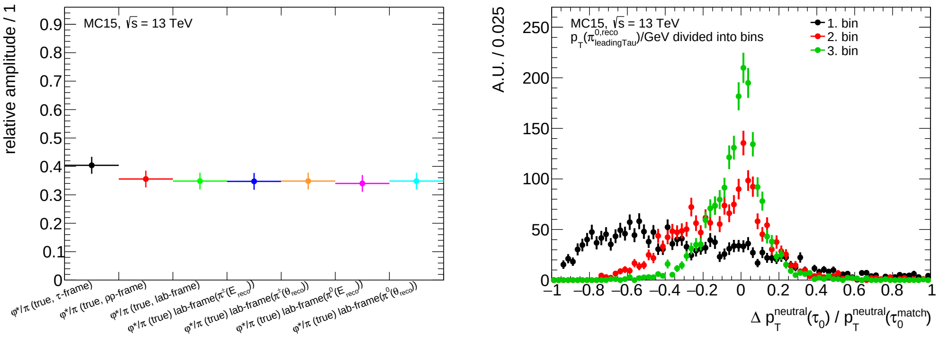
An explanation for this increase in the relative amplitude can be seen in figure 8.13. Due to a selection bias the p_T resolution seen in 8.13(d) is broader and shifted with respect to zero in the region $p_T < 19.6\,\mathrm{GeV}$. Since the Preselection requires a minimum p_T for the τ_{had}'s, events in this region often contain charged pions with high p_T compared to the neutral pion p_T. For these kinds of events, high energetic charged pions and low energetic neutral pions, the error in the subtraction needed for the reconstruction of the neutral pion four-momentum increases. The resolution in this p_T region directly translates to a broadened $y_0 y_1$ distribution in this region, compare figure 8.13(e-f). For the middle region $19.6\,\mathrm{GeV} \leq p_T < 43.7\,\mathrm{GeV}$ the $y_0 y_1$ distribution at generator level and reconstruction level is in comparison sharply peaked around zero. Thus, mismeasurements in the energy measurement of the pions result more often in a sign-flip of $y_0 y_1$ at reconstruction level for the middle region than for the low p_T region. As seen in figure 8.13(a-c) these more abundant sign-flips are the reason for the most significant drop of amplitude due to the frame reconstruction in the middle region.

Since the overall relative amplitude in the low p_T region and the middle region are comparable within statistical uncertainties other reconstruction than the aforementioned ones have to cause the low amplitude in the middle region. Figure 8.14 shows the reconstruction effects not related to the $y_0 y_1$ calculation. A stronger degradation in the relative amplitude for the low p_T region than for the middle region can be observed which fully explains the behaviour seen in figure 8.12. The stronger degradation in the low p_T region for these reconstruction effects is directly related to the previously discussed p_T resolution of the neutral pion. As the angular resolutions of the neutral pion of the leading τ_{had}, i.e. η and ϕ resolutions, do not show a different behaviour in the different p_T bins, see figure A.7 in appendix A.5.2,



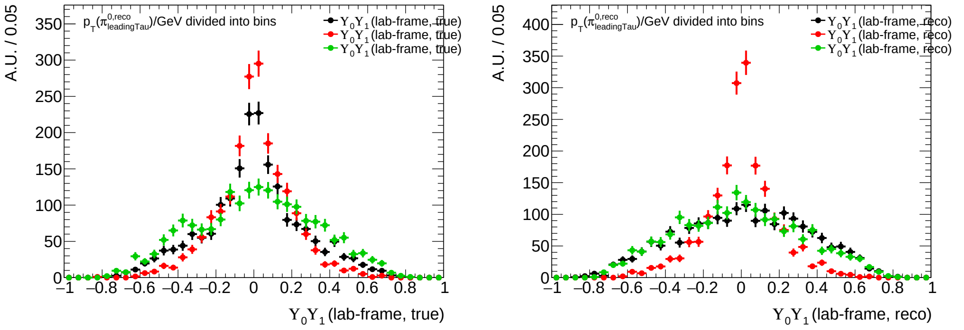
(a) Relative amplitude for $p_{\mathrm{T}} < 19.6\,\mathrm{GeV}$.

(b) Relative amplitude for $19.6\,\mathrm{GeV} \leq p_{\mathrm{T}} < 43.7\,\mathrm{GeV}$.



(c) Relative amplitude for $p_{\mathrm{T}} \geq 43.7\,\mathrm{GeV}$.

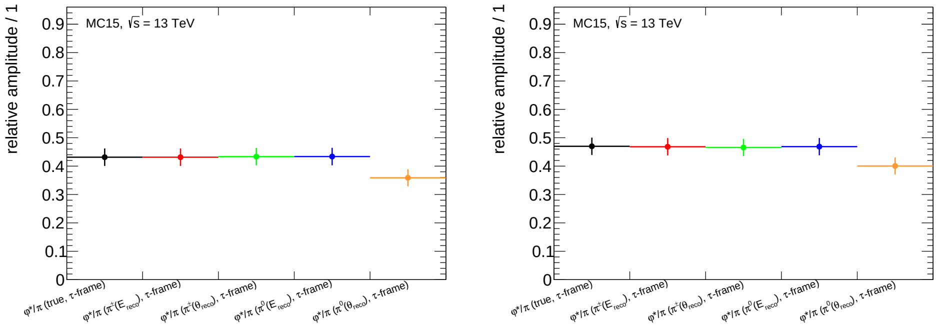
(d) Neutral pion p_T resolution in the different regions.



(e) $y_0 y_1$ distribution at generator level, using the laboratory frame, for the different regions.

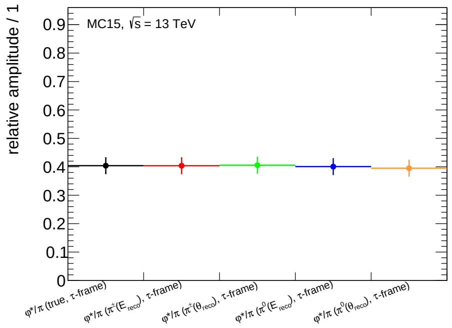
(f) $y_0 y_1$ distribution at reconstruction level for the different regions.

Figure 8.13: Effects of different frame for the $y_0 y_1$ calculation and reconstruction effects in the same on the relative amplitude and the phase shift in the different the neutral pion p_{T} of the leading τ_{had} regions (a-c). Resolution of the neutral pion p_{T} (d) and the $y_0 y_1$ distributions at generator level and reconstruction level in the same regions (e-f).



(a) Relative amplitude for $p_{\mathrm{T}} < 19.6\,\mathrm{GeV}$.

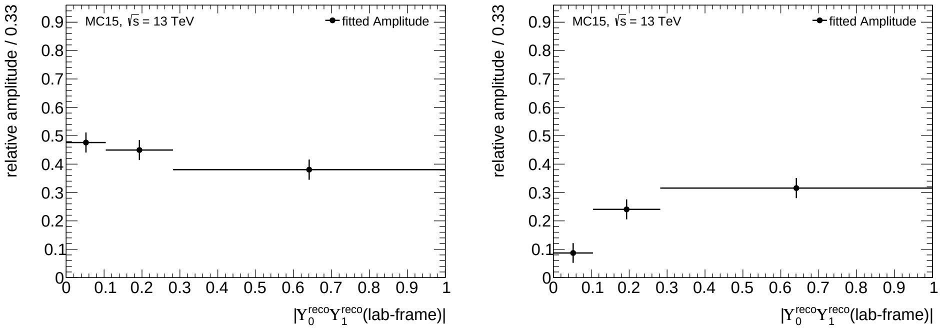
(b) Relative amplitude for $19.6\,\mathrm{GeV} \leq p_{\mathrm{T}} < 43.7\,\mathrm{GeV}$.



(c) Relative amplitude for $p_{\mathrm{T}} \geq 43.7\,\mathrm{GeV}$.

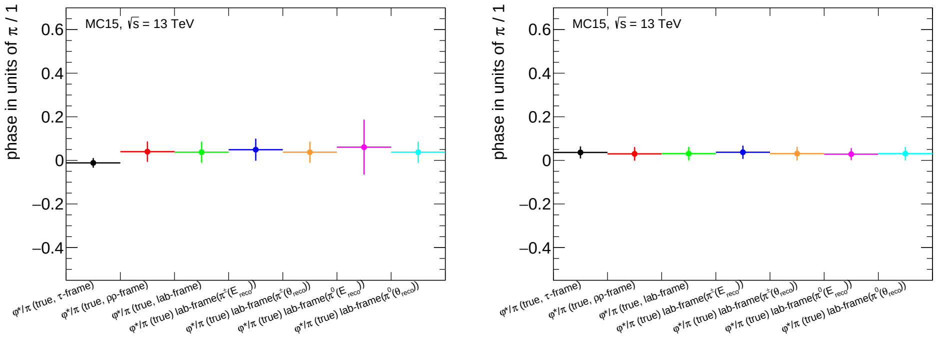
Figure 8.14: Effects of reconstruction effects in the decay plane reconstruction and triple correlation calculation on the relative amplitude and the phase shift in the different the neutral pion p_{T} of the leading τ_{had} regions.

the overall dependence of φ^*_{CP} on the neutral pion p_{T} could be exploited. This could be done for example by weighting the events according to their neutral pion p_{T} of the leading τ_{had} in order to not loose events.



(a) Relative amplitude at generator level. (b) Relative amplitude at reconstruction level.

Figure 8.15: Comparison of the dependence of φ^*_{CP} at generator level and reconstruction level as a function $|y_0y_1|$. At generator level y_0y_1 is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.



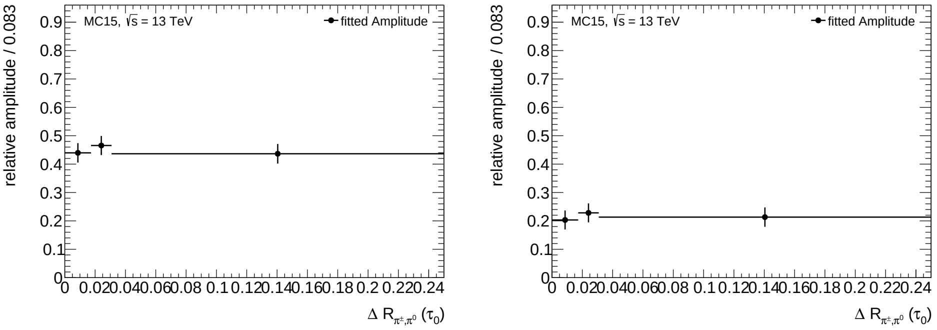
(a) Relative amplitude for $|y_0y_1| < 0.10$. (b) Relative amplitude for $|y_0y_1| > 0.28$.

Figure 8.16: Effects of different frame for the y_0y_1 calculation and reconstruction effects in the same on the relative amplitude and the phase shift in the different $|y_0y_1|$ regions.

8.2.2 y_0y_1

As seen before the sign-flip when reconstructing y_0y_1 has the biggest impact on φ^*_{CP} out of all reconstruction effects. Supposedly, a cut on this variable might increase the performance of the measurement. Figure 8.15 shows that with increasing absolute values the relative amplitude indeed does increase. Thus, it might be worth it to take a cut on $|y_0y_1|$ or a reweighting according to $|y_0y_1|$ into consideration.

The increase of the amplitude observed here is rather an effect of the y_0y_1 calculation itself than of frame reconstruction effects. Events with values of y_0y_1 close to zero are more sensitive to mismeasurements of the pion energies which lead to a sign-flip of y_0y_1. This sign-flip leads to a shift of φ^*_{CP} by π when it shouldn't be shifted and thus decreases the relative amplitude.

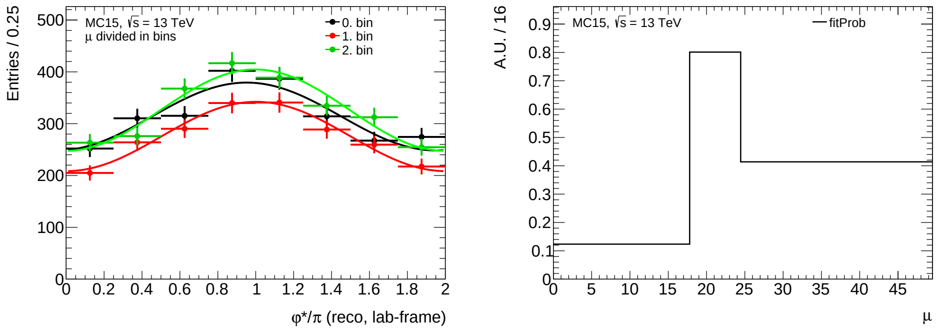


(a) Relative amplitude at generator level. (b) Relative amplitude at reconstruction level.

Figure 8.17: Comparison of the dependence of φ^*_{CP} at generator level and reconstruction level as a function of $\Delta R\left(\pi^\pm, \pi^0\right)$. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.

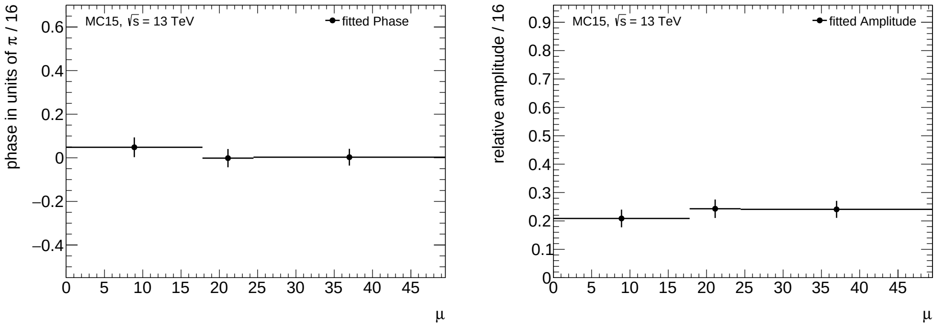
8.2.3 $\Delta R\left(\pi^\pm, \pi^0\right)$

As seen in figure 5.2 the resolution of the neutral pions is better perpendicular to the axis that connects the charged pion and neutral pion in the η-ϕ-plane. Furthermore, the performance of the measurement increases for better neutral pion resolutions as seen in section 8.2.1. Thus, it has been studied whether a cut on the angular distance between the charged pion and the neutral pion, ΔR, could increase the relative amplitude of φ^*_{CP} as the reconstruction of the neutral pions presumably performs better for a higher separation of both pion flavours. Figure 8.17 shows that the relative amplitude shows no statistically significant dependence on $\Delta R\left(\pi^\pm, \pi^0\right)$. Thus, it is of no further interest for an event selection.



(a) φ^*_{CP} distribution.

(b) Fit probabilities.



(c) Phase shift.

(d) Relative amplitude.

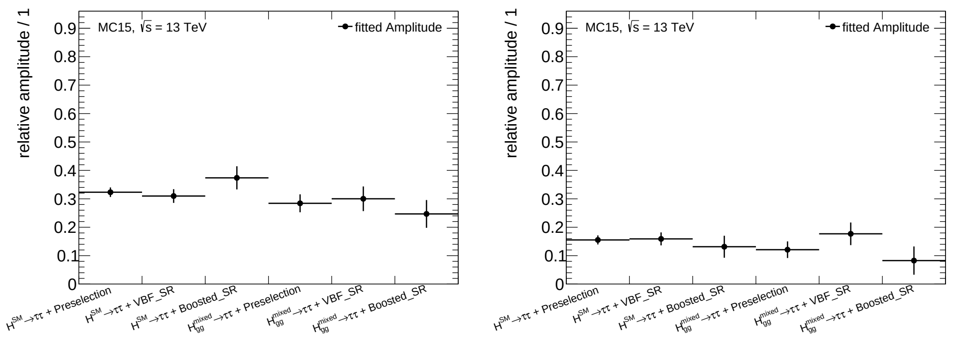
Figure 8.18: Behaviour of the reconstructed φ^*_{CP} as a function of the average number of interactions per bunch crossing, μ.

8.3 Effects of pile up

Section 4.2.2 showed that pile-up has to be considered when performing measurements at ATLAS. Although the average number of interactions per bunch crossing, which is a measure for the pile up, μ, is not part of the event selection, it is important that the method used to measure the $\mathcal{CP}$-state of the Higgs boson is stable as a function of the same. Figure 8.18 shows that the method performs similar in all regions of μ. The Preselection is applied and the reconstructed decay mode and generated decay mode have to be `1p1n`.

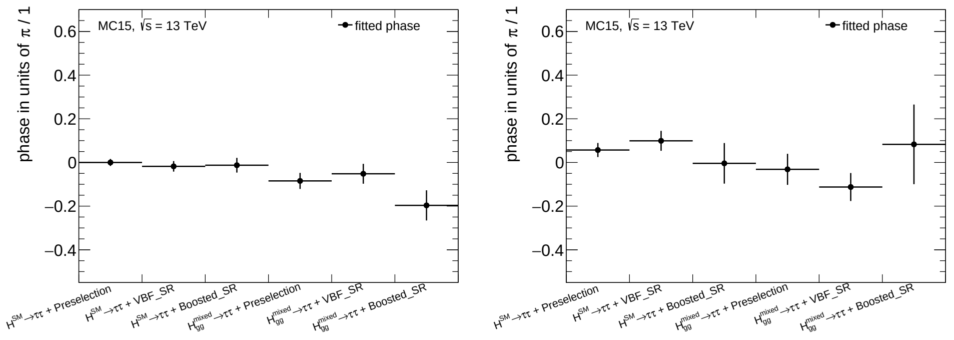
8.4 Performance comparison for $\mathcal{CP}$-even and $\mathcal{CP}$-mixed scenarios

As the method is designed to measure possible $\mathcal{CP}$-admixtures it is important that it performs comparable for all $\mathcal{CP}$-states in the different signal regions. Figure 8.19 shows the relative amplitude and the phase shift of the φ^*_{CP} distribution for the two extreme scenarios of a pure $\mathcal{CP}$-even Higgs and a maximum admixture in all selection categories. Within the statistical uncertainties the relative amplitude is similar for both $\mathcal{CP}$ scenarios in all categories, at generator level and reconstruction level respectively. For the $\mathcal{CP}$-mixed case a bias in the phase at generator level can be observed in the Boosted Signal Region which is not seen at reconstruction level, supposedly because it gets washed out by reconstruction effects. All in all, the method performs similar for both scenarios in all selection categories.



(a) Relative amplitude of the generated φ^*_{CP} distribution.

(b) Relative amplitude of the reconstructed φ^*_{CP} distribution.



(c) Phase shift of the generated φ^*_{CP} distribution.

(d) Phase shift of the reconstructed φ^*_{CP} distribution.

Figure 8.19: Comparison of the relative amplitude and phase shift for a $\mathcal{CP}$-even Higgs and maximal $\mathcal{CP}$-violating Higgs scenario in the different selection categories.

CHAPTER 9

Conclusions

A measurement to determine possible $\mathcal{CP}$-admixtures in the resonance observed at ATLAS and CMS in 2012 was introduced. This measurement uses the angle between the planes spanned by the decay products of each τ_{had} which is then called φ^*_{CP}. The event selection for such a measurement, which is adapted from the $H \to \tau\tau$ coupling analysis, was discussed and used as the reference to investigate its impact on the same.

To do so, all cuts from the specific event selection region were applied but not the cut on the variable under investigation. The phase space of the aforementioned variable was split into three regions of equal statistics and the φ^*_{CP} distribution was fitted with a cosine function:

$$f(x) = u \cdot \cos(x + v) + w \tag{9.1}$$

from which the relative amplitude, rel. amp. $= \frac{u}{w}$, and the phase shift, v, was extracted. For each variable listed in table 5.1 the relative amplitude and the phase shift was compared between the φ^*_{CP} distribution at generator level and reconstruction level for each of the three phase space regions. Regions that showed strong impact on either the relative amplitude or the phase shift were studied more closely by splitting the calculation of the decay angle φ^*_{CP} into separate parts. One part that considers reconstruction effects in the $y_0 y_1$ calculation and one part that considers reconstruction effects in the decay plane reconstruction.

This study showed that the choice of the frame used for the $y_0 y_1$ calculation overall has the most significant impact on the performance of the measurement. Ideally, the respective τ_{had} frame would be used but this frame cannot be accessed since the neutrinos are not measurable by the ATLAS detector. Instead the laboratory frame is used. The second significant reconstruction effect that caused for example a degradation in the relative amplitude was the reconstruction of the neutral pion direction in the decay plane reconstruction.

Furthermore, general reconstruction effects in the measurement have been studied, like the impact of the decay mode classification and the choice of the τ_{had}-ID. It was shown that `1pXn` decays that are classified as `1p1n` decays provide the strongest degradation in the relative amplitude which was expected since this misclassification also provides the biggest contamination. For tighter requirements on the τ_{had}-ID the relative amplitude improved at generator level but it has to be considered that in the actual analysis τ_{had}'s that are identified as `medium` might also be identified as `tight`, meaning that the bins are not as independent in the actual analysis from each other as they are in this study.

Since the choice of the frame in which $y_0 y_1$ is calculated in has the strongest impact on φ^*_{CP}, it has also been studied if the neutrinos in the τ_{had} decays can be approximated by ditau-mass reconstruction techniques. In this context different techniques have been discussed and their performance in reconstructing the ditau-mass have been investigated. These techniques play a major role in the discrimination of $Z \to \tau\tau$ against $H \to \tau\tau$. However, the MMC and the Collinear Approximation also give estimates of the neutrino momenta which can be used for a frame approximation for the φ^*_{CP} calculation. As the Missing Mass Calculator seemed the most promising approach, a new tune for the MMC has been derived and can be used for future analyses like the $H \to \tau\tau$ coupling analysis in Run 2. The tune contains a parametrization of the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution as a function of $\sqrt{\Sigma E_{\mathrm{T}}}$ and a correction term as a function of $\Delta\phi\,(\tau_{\mathrm{had}}, \tau_{\mathrm{had}})$.

The Collinear Approximation and the MMC with the new tune were used to approximate the respective τ_{had} frames and the resonance frame for the $y_0 y_1$ calculation. The results did not improve significantly compared to the usage of the laboratory frame when the rest of the φ^*_{CP} calculation is performed at generator level. However, if everything is calculated at reconstruction level a systematic shift towards better relative amplitudes and smaller phase shifts can be observed for the resonance frame reconstructed with the MMC and the Collinear Approximation. Even though, the laboratory frame and the resonance frames are compatible within 1σ the resonance frame obtained from the Collinear Approximation might be the frame to choose over the laboratory frame or a more sophisticated technique like the MMC.

For possible improvements it could be investigated whether the decay mode classification of the τ_{had}'s or the angular resolutions of the neutral pions can be further improved.

To summarize the frame reconstruction for the $y_0 y_1$ calculation has been identified as the most crucial aspect in the φ^*_{CP} performance and a re-tuning has been performed for the Missing Mass Calculator which is used to discriminate $Z \to \tau\tau$ and $H \to \tau\tau$ in different analyses.

Bibliography

[1] S. Chatrchyan et al.,
Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC,
Phys. Lett. **B716** (2012) 30, arXiv: `1207.7235 [hep-ex]`.

[2] G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1, arXiv: `1207.7214 [hep-ex]`.

[3] Nobelprize.org, *The Nobel Prize in Physics 2013*, 2016,
URL: `http://www.nobelprize.org/nobel_prizes/physics/laureates/2013/` (visited on 14/07/2016).

[4] K. A. Olive et al., *Review of Particle Physics*, Chin. Phys. **C38** (2014) 090001.

[5] G. Aad et al.,
Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector,
JHEP **04** (2015) 117, arXiv: `1501.04943 [hep-ex]`.

[6] *Search for neutral MSSM Higgs bosons decaying to tau+tau- pairs in proton-proton collisions at sqrt(s) = 7 TeV with the ATLAS detector*, tech. rep. ATLAS-CONF-2011-132, CERN, 2011,
URL: `https://cds.cern.ch/record/1383835`.

[7] A. Pich, "The Standard model of electroweak interactions", *2004 European School of High-Energy Physics, Sant Feliu de Guixols, Spain, 30 May - 12 June 2004*, 2005 1,
arXiv: `hep-ph/0502010 [hep-ph]`,
URL: `http://doc.cern.ch/yellowrep/2006/2006-003/p1.pdf`.

[8] Wikimedia, *Standard Model*, 2016, URL: `https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg` (visited on 08/07/2016).

[9] CERN, *LHCb unveils new particles*, 2016,
URL: `https://home.cern/about/updates/2016/07/lhcb-unveils-new-particles`
(visited on 09/07/2016).

[10] R. Aaij et al., *Observation of $J/\psi\phi$ structures consistent with exotic states from amplitude analysis of $B^+ \to J/\psi\phi K^+$ decays*, (2016), arXiv: `1606.07895 [hep-ex]`.

[11] R. Aaij et al., *Amplitude analysis of $B^+ \to J/\psi\phi K^+$ decays*, (2016),
arXiv: `1606.07898 [hep-ex]`.

[12] CERN, *CERN's LHCb experiment reports observation of exotic pentaquark particles*, 2016,
URL: `https://press.cern/press-releases/2015/07/cerns-lhcb-experiment-%20reports-observation-exotic-pentaquark-particles` (visited on 09/07/2016).

[13] R. Aaij et al.,
Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays,
Phys. Rev. Lett. **115** (2015) 072001, arXiv: `1507.03414 [hep-ex]`.

[14] M. Thomson, *Modern particle physics*, New York: Cambridge University Press, 2013, ISBN: 9781107034266, URL: `http://www-spires.fnal.gov/spires/find/books/www?cl=QC793.2.T46::2013`.

[15] P. Achard et al., *Measurement of the running of the electromagnetic coupling at large momentum-transfer at LEP*, Phys. Lett. **B623** (2005) 26, arXiv: `hep-ex/0507078 [hep-ex]`.

[16] S. Bethke, *The 2009 World Average of alpha(s)*, Eur. Phys. J. **C64** (2009) 689, arXiv: `0908.1135 [hep-ph]`.

[17] A. Djouadi, *The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model*, Phys. Rept. **457** (2008) 1, arXiv: `hep-ph/0503172 [hep-ph]`.

[18] W. de Boer, "The Discovery of the Higgs Boson with the CMS Detector and its Implications for Supersymmetry and Cosmology", *Time and Matter 2013 (TAM2013) Venice, Italy*, 2013, arXiv: `1309.0721 [hep-ph]`, URL: `https://inspirehep.net/record/1252561/files/arXiv:1309.0721.pdf`.

[19] K. Jakobs, "Physics at the LHC – From Standard Model measurements to Searches for New Physics", *Proceedings, 2011 European School of High-Energy Physics (ESHEP 2011)*, 2014 309, arXiv: `1206.7024 [hep-ex]`, URL: `https://inspirehep.net/record/1120513/files/arXiv:1206.7024.pdf`.

[20] K. Jakobs and C. Seez, *The Higgs Boson discovery*, Scholarpedia **10**.9 (2015) 32413.

[21] S. Chatrchyan 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**.8 (2013) 081803, arXiv: `1212.6639 [hep-ex]`.

[22] R. M. Godbole, D. J. Miller and M. M. Muhlleitner, *Aspects of CP violation in the H ZZ coupling at the LHC*, JHEP **12** (2007) 031, arXiv: `0708.0458 [hep-ph]`.

[23] R. Harnik et al., *Measuring CP violation in $h \to \tau^+\tau^-$ at colliders*, Phys. Rev. **D88**.7 (2013) 076009, arXiv: `1308.1094 [hep-ph]`.

[24] S. Berge, W. Bernreuther and S. Kirchner, *Determination of the Higgs CP-mixing angle in the tau decay channels at the LHC including the Drell–Yan background*, Eur. Phys. J. **C74**.11 (2014) 3164, arXiv: `1408.0798 [hep-ph]`.

[25] A. Arbey et al., *Exploring CP Violation in the MSSM*, Eur. Phys. J. **C75**.2 (2015) 85, arXiv: `1410.4824 [hep-ph]`.

[26] M. Kamionkowski, "Dark Matter and Dark Energy", *Amazing Light: Visions for Discovery: An International Symposium in Honor of the 90th Birthday Years of Charles H. Townes Berkeley, California, October 6-8, 2005*, 2007, arXiv: `0706.2986 [astro-ph]`, URL: `https://inspirehep.net/record/753633/files/arXiv:0706.2986.pdf`.

[27] B. Novosyadlyj et al., *Dark Energy: Observational Evidence and Theoretical Models*, Kyiv: Academperiodyka, 2013, ISBN: 978-966-360-239-4, 978-966-360-240-0, arXiv: `1502.04177 [astro-ph.CO]`, URL: `https://inspirehep.net/record/1344979/files/arXiv:1502.04177.pdf`.

[28] W. de Boer, *Grand unified theories and supersymmetry in particle physics and cosmology*, Prog. Part. Nucl. Phys. **33** (1994) 201, arXiv: `hep-ph/9402266 [hep-ph]`.

[29] W. Buchmuller, "Baryogenesis: 40 Years Later", *Proceedings on 13th International Symposium on Particles, strings, and cosmology (PASCOS 2007): London, UK, July 2-7, 2007*, 2007, arXiv: `0710.5857 [hep-ph]`, URL: `https://inspirehep.net/record/766316/files/arXiv:0710.5857.pdf`.

[30] B. Aubert et al., *Observation of CP violation in the B^0 meson system*, Phys. Rev. Lett. **87** (2001) 091801, arXiv: `hep-ex/0107013 [hep-ex]`.

[31] M. Worek, *Higgs CP from H/A0 —> tau tau decay*, Acta Phys. Polon. **B34** (2003) 4549, arXiv: `hep-ph/0305082 [hep-ph]`.

[32] V. D. Barger et al., *Higgs bosons: Intermediate mass range at e+ e- colliders*, Phys. Rev. **D49** (1994) 79, arXiv: `hep-ph/9306270 [hep-ph]`.

[33] S. Berge, W. Bernreuther and S. Kirchner, *Prospects of constraining the Higgs boson's CP nature in the tau decay channel at the LHC*, Phys. Rev. **D92** (2015) 096012, arXiv: `1510.03850 [hep-ph]`.

[34] S. Berge et al., *How to pin down the CP quantum numbers of a Higgs boson in its tau decays at the LHC*, Phys. Rev. **D84** (2011) 116003, arXiv: `1108.0670 [hep-ph]`.

[35] C. Press, *facts-figures*, 2016, URL: `http://press.cern/backgrounders/facts-figures` (visited on 18/08/2016).

[36] O. S. Brüning et al., *LHC Design Report*, Geneva: CERN, 2004, URL: `https://cds.cern.ch/record/782076`.

[37] G. Aad et al., *Improved luminosity determination in pp collisions at sqrt(s) = 7 TeV using the ATLAS detector at the LHC*, Eur. Phys. J. **C73**.8 (2013) 2518, arXiv: `1302.4393 [hep-ex]`.

[38] E. Torrence, *LuminosityPublicResultsRun2*, 2016, URL: `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2` (visited on 28/09/2016).

[39] G. Weiglein et al., *Physics interplay of the LHC and the ILC*, Phys. Rept. **426** (2006) 47, arXiv: `hep-ph/0410364 [hep-ph]`.

[40] T. A. Collaboration et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3**.08 (2008) S08003, URL: `http://stacks.iop.org/1748-0221/3/i=08/a=S08003`.

[41] A. Airapetian et al., *ATLAS: Detector and physics performance technical design report. Volume 1*, (1999).

[42] A. La Rosa, *ATLAS IBL Pixel Upgrade*, Nucl. Phys. Proc. Suppl. **215** (2011) 147, arXiv: `1104.1980 [physics.ins-det]`.

[43] M. Gupta, *Calculation of radiation length in materials*, (2010), URL: `https://cds.cern.ch/record/1279627/files/PH-EP-Tech-Note-2010-013.pdf` (visited on 16/07/2016).

[44] S. Tavernier, *Experimental Techniques in Nuclear and Particle Physics*, Berlin, Germany: Springer, 2010, ISBN: 9783642008283, 3642008283.

[45] Y. Nakahama, *The ATLAS Trigger System: Ready for Run-2*, Journal of Physics: Conference Series **664**.8 (2015) 082037, URL: `http://stacks.iop.org/1742-6596/664/i=8/a=082037`.

[46] E. Simioni, *The Topological Processor for the future ATLAS Level-1 Trigger: from design to commissioning*, (2014), arXiv: `1406.4316 [physics.ins-det]`.

[47] G. Aad et al., *Expected performance of the ATLAS experiment: detector, trigger and physics*, Geneva: CERN, 2009, URL: `https://cds.cern.ch/record/1125884`.

[48] G. Aad et al., *Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data*, Eur. Phys. J. **C72** (2012) 1909, arXiv: `1110.3174 [hep-ex]`.

[49] M. Aaboud et al., *Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run-1 data*, (2016), arXiv: `1606.01813 [hep-ex]`.

[50] G. Aad et al., *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at* $\sqrt{s}$ *=13 TeV*, Eur. Phys. J. **C76**.5 (2016) 292, arXiv: `1603.05598 [hep-ex]`.

[51] A. Collaboration, *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at* $\sqrt{s} = 13$ *TeV*, tech. rep. ATL-PHYS-PUB-2015-015, CERN, 2015, URL: `http://cds.cern.ch/record/2037613`.

[52] M. Cacciari, G. P. Salam and G. Soyez, *The Anti-k(t) jet clustering algorithm*, JHEP **04** (2008) 063, arXiv: `0802.1189 [hep-ph]`.

[53] *First Observation of the process* $Z \rightarrow \tau\tau \rightarrow e\mu + 4\nu$ *with the ATLAS Detector*, tech. rep. ATLAS-CONF-2011-045, CERN, 2011, URL: `http://cds.cern.ch/record/1338574`.

[54] Y. Sakurai, "The ATLAS Tau Trigger Performance during LHC Run 1 and Prospects for Run 2", *Proceedings, 2nd Conference on Large Hadron Collider Physics Conference (LHCP 2014)*, 2014, arXiv: `1409.2699 [hep-ex]`, URL: `https://inspirehep.net/record/1315688/files/arXiv:1409.2699.pdf`.

[55] *Performance of the Reconstruction and Identification of Hadronic Tau Decays with ATLAS*, tech. rep. ATLAS-CONF-2011-152, CERN, 2011, URL: `https://cds.cern.ch/record/1398195`.

[56] G. Aad et al., *Reconstruction of hadronic decay products of tau leptons with the ATLAS experiment*, Eur. Phys. J. **C76**.5 (2016) 295, arXiv: `1512.05955 [hep-ex]`.

[57] C. Limbach, *Development of a new reconstruction and classification method for Tau leptons and its application in the ATLAS detector at the LHC*, BONN-IR-2015-03, PhD Thesis: University of Bonn, 2015, URL: `http://hss.ulb.uni-bonn.de/2015/4007/4007.htm`.

[58] B. H. Brunt et al., *Expected performance of missing transverse momentum reconstruction for the ATLAS detector at* $\sqrt{s} = 13$ *TeV*, tech. rep. ATL-COM-PHYS-2015-347, CERN, 2015, URL: `https://cds.cern.ch/record/2013489`.

[59] A. Collaboration, *Performance of missing transverse momentum reconstruction for the ATLAS detector in the first proton-proton collisions at at* $\sqrt{s}$*= 13 TeV*, tech. rep. ATL-PHYS-PUB-2015-027, CERN, 2015, URL: `https://cds.cern.ch/record/2037904`.

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

[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 [hep-ph]`.

[62] S. Alioli et al., *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043, arXiv: `1002.2581 [hep-ph]`.

[63] T. Sjostrand, S. Mrenna and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv: `0710.3820 [hep-ph]`.

[64] T. Przedzinski, E. Richter-Was and Z. Was, *TauSpinner: a tool for simulating CP effects in $H \to \tau\tau$ decays at LHC*, Eur. Phys. J. **C74**.11 (2014) 3177, arXiv: `1406.1647 [hep-ph]`.

[65] Z. Marshall, *Simulation of Pile-up in the ATLAS Experiment*, J. Phys. Conf. Ser. **513** (2014) 022024.

[66] A. Elagin et al., *A New Mass Reconstruction Technique for Resonances Decaying to di-tau*, Nucl. Instrum. Meth. **A654** (2011) 481, arXiv: `1012.4686 [hep-ex]`.

[67] S. Pranko, “Mass Reconstruction with MMC”, ATLAS/CMS Workshop, Bonn, 2012, URL: `https://indico.desy.de/getFile.py/access?contribId=8&resId=0&materialId=slides&confId=7444`.

[68] M. Pitt, *Search for the Standard Model Higgs boson in the $H \to \tau\tau$ channel with the ATLAS detector*, tech. rep. ATL-PHYS-PROC-2013-141, CERN, 2013, URL: `http://cds.cern.ch/record/1554090`.

[69] A. De Maria, *Study of the H boson in the H->tautau decay in Run2 in the ATLAS experiment, prelim. title*, PhD Thesis: University of Göttingen, to-be-published.

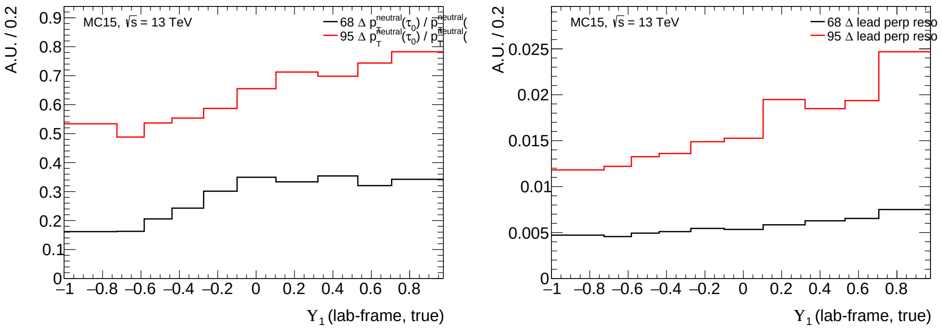
[70] L. Weise, *Optimisation of the reconstruction of the $H \to \tau\tau$ process in the ATLAS experiment*, II.Physik-UniGö-MSc-2015/04, Master Thesis: Geog-August-Universität Göttingen, 2015.

[71] P. Wagner, personal communication, 2016.

APPENDIX A

Appendix

A.1 Substructure Performance at 13 TeV

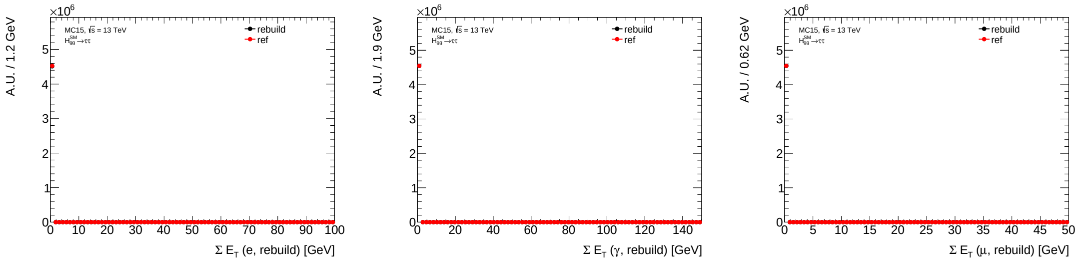


(a) p_{T} resolution as a function of y_1. A similar behavior is seen for y_1.

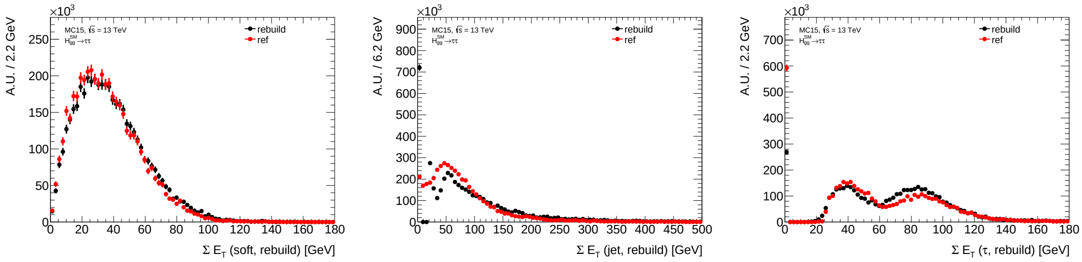
(b) Angular resolution as a function of y_1. A similar behavior is seen for y_1.

Figure A.1: 68% and 95% quantiles of the p_{T} resolution of the neutral pion as a function of y_1. Additionally, the quantiles of the angular resolution of the neutral pion perpendicular to the axis that connects the charged pion and neutral pion in the η-ϕ-plane is shown as a function of y_1.

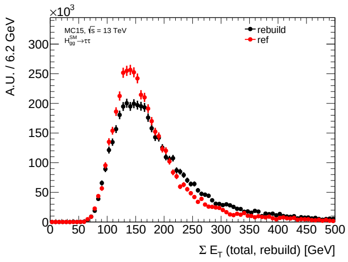
A.2 E_T^{miss} rebuilding



(a) Electron term.
(b) Photon term.
(c) Muon term.



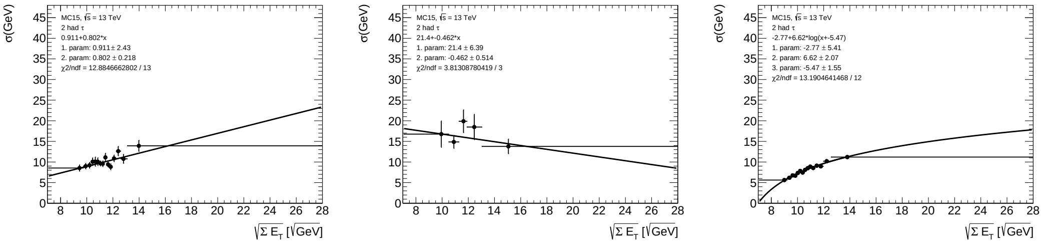
(d) Soft term.
(e) Jet term.
(f) Tau term.



(g) Total term.

Figure A.2: Comparison of the reconstructed terms used in the E_T^{miss} rebuilding and the reference terms provided by the Jet/EtMis-Group. The Substructure Algorithm is used for the τ reconstruction.

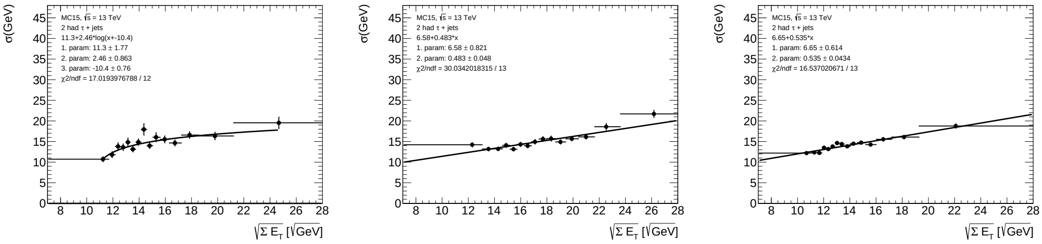
A.3 Parametrization as a function of $\sqrt{\Sigma E_{\mathrm{T}}}$



(a) $H_{gg} \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ sample in the no jets category.

(b) $H_{VBF} \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ sample in the no jets category.

(c) $Z \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ sample in the no jets category.

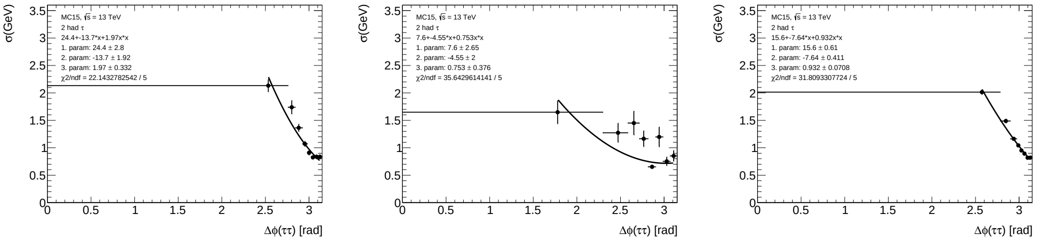


(d) $H_{gg} \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ sample in the jets category.

(e) $H_{VBF} \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ sample in the jets category.

(f) $Z \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ sample in the jets category.

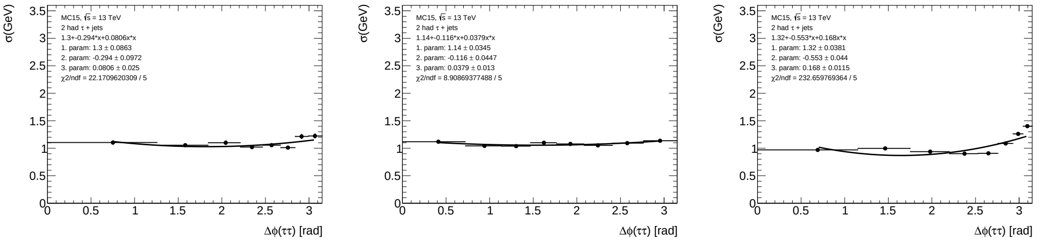
Figure A.3: $E_{\mathrm{T}}^{\text{miss}}$ resolution as a function of $\sqrt{\Sigma E_{\mathrm{T}}}$ for the different samples and categories. The TauRec algorithm is used for the τ reconstruction.



(a) $H_{gg} \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ sample in the no jets category.

(b) $H_{VBF} \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ sample in the no jets category.

(c) $Z \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ sample in the no jets category.



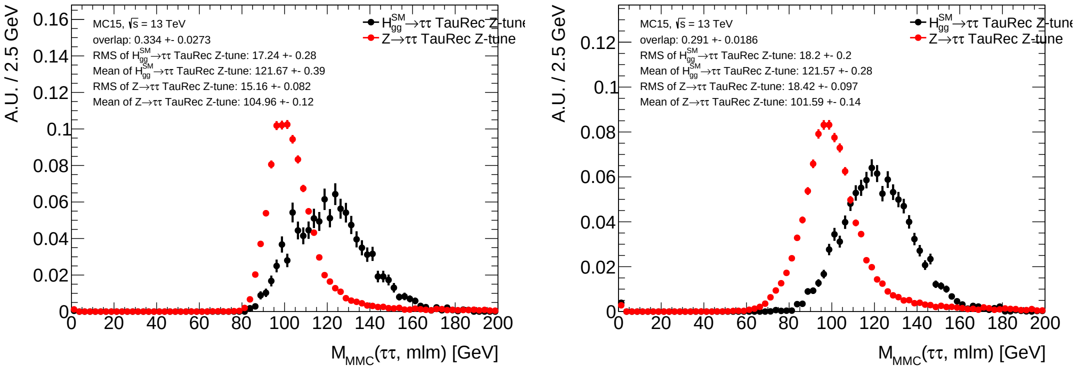
(d) $H_{gg} \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ sample in the jets category.

(e) $H_{VBF} \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ sample in the jets category.

(f) $Z \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ sample in the jets category.

Figure A.4: $\Delta E_{\mathrm{T}}^{\text{miss}}$ as a function of $\Delta\phi\,(\tau_{\text{had}}, \tau_{\text{had}})$ for the different samples and categories. The TauRec algorithm is used for the τ reconstruction.

A.4 Results for TauRec



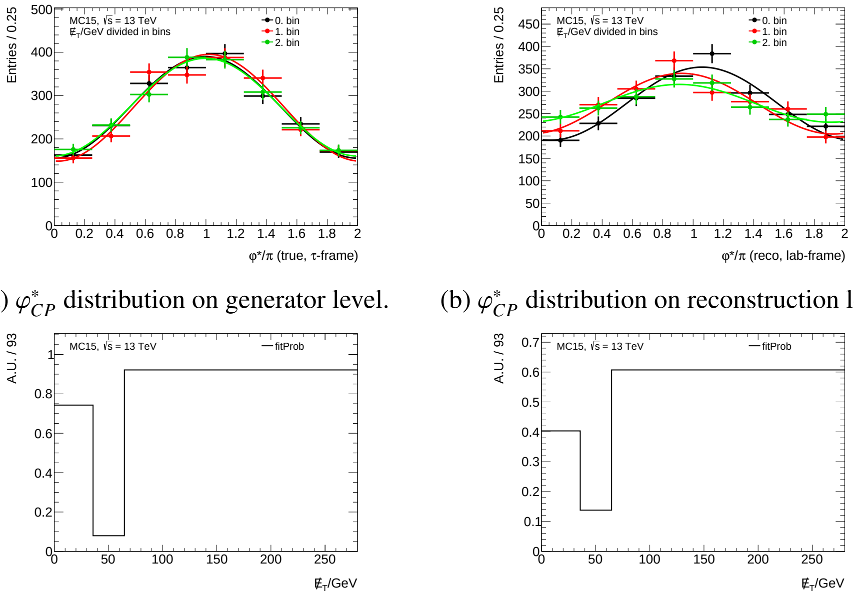
(a) No jets category.

(b) Jets category.

Figure A.5: Mass distributions for $H \to \tau_{\text{had}}\tau_{\text{had}}$ and $Z \to \tau_{\text{had}}\tau_{\text{had}}$ with the new tune for the τ_{had} τ_{had} channel. The TauRec algorithm is used for the τ_{had} reconstruction and the MLM output is shown.

A.5 Impact of selection cuts on φ^*_{CP}

A.5.1 $E_{\mathrm{T}}^{\mathrm{miss}}$

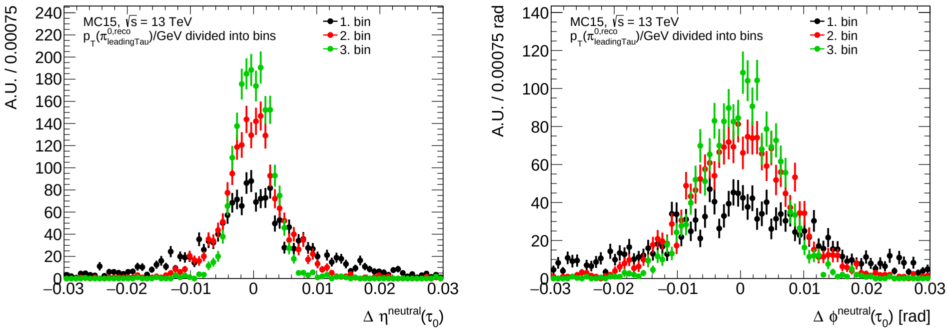


(a) φ^*_{CP} distribution on generator level. (b) φ^*_{CP} distribution on reconstruction level.

(c) Fit probability on generator level. (d) Fit probability on reconstruction level.

Figure A.6: φ^*_{CP} distribution and fit probability at generator level and at reconstruction level as a function of $E_{\mathrm{T}}^{\mathrm{miss}}$. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.

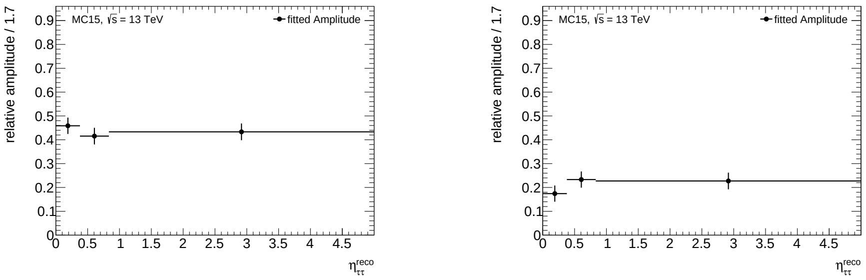
A.5.2 Neutral pion p_{T} of the leading τ_{had}



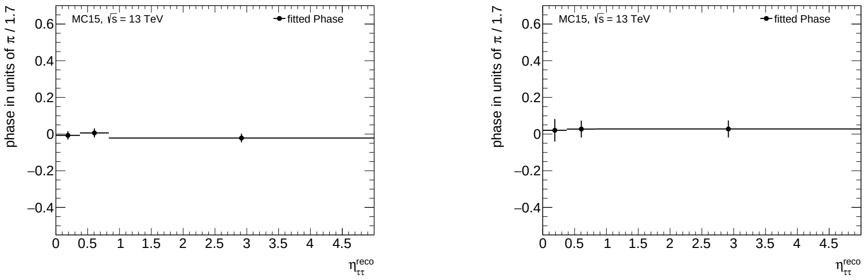
(a) Neutral pion η resolution in the different regions. (b) Neutral pion ϕ resolution in the different regions.

Figure A.7: η and ϕ resolution of the neutral pion in the different p_{T} regions defined in section 8.2.1.

A.5.3 $\Delta\eta$

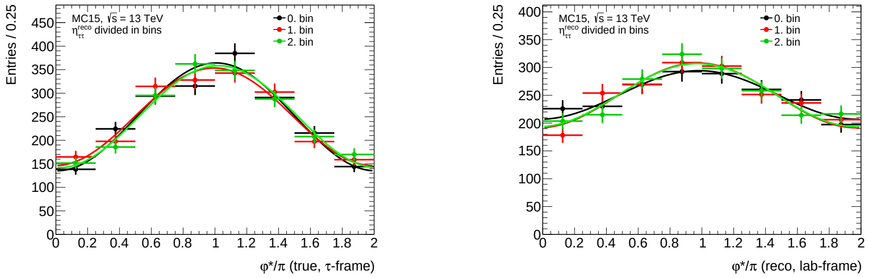


(a) Relative amplitude at generator level. (b) Relative amplitude at reconstruction level.

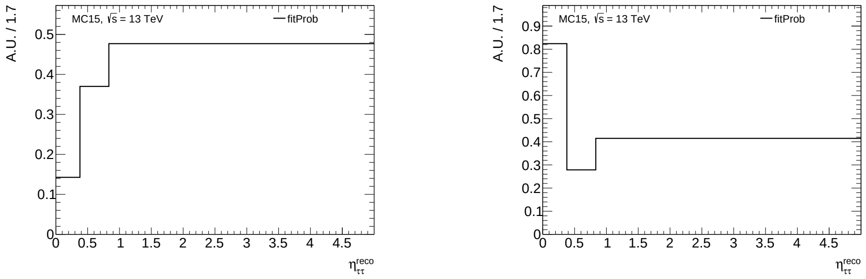


(c) Phase shift at generator level. (d) Phase shift at reconstruction level.

Figure A.8: Comparison of the dependence of φ^*_{CP} at generator level and reconstruction level as a function of $\Delta\eta$. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.



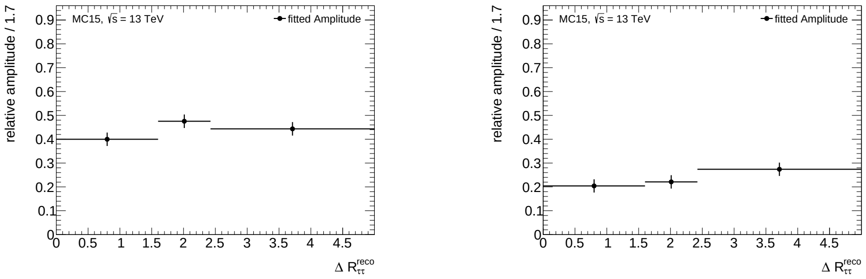
(a) φ^*_{CP} distribution at generator level. (b) φ^*_{CP} distribution at reconstruction level.



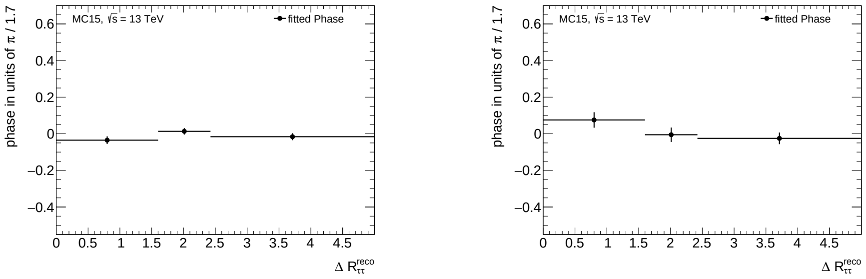
(c) Fit probability at generator level. (d) Fit probability at reconstruction level.

Figure A.9: φ^*_{CP} distribution and fit probability at generator level and at reconstruction level as a function of $\Delta\eta$. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.

A.5.4 ΔR

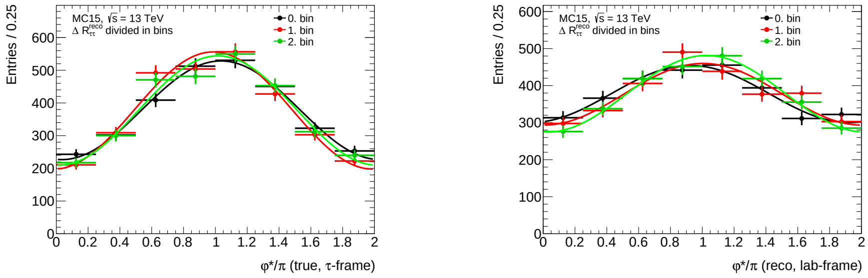


(a) Relative amplitude at generator level. (b) Relative amplitude at reconstruction level.

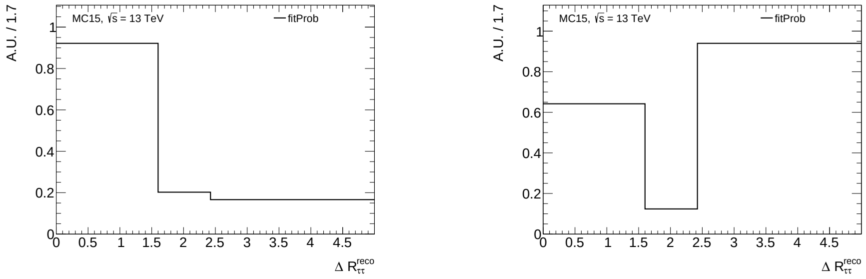


(c) Phase shift at generator level. (d) Phase shift at reconstruction level.

Figure A.10: Comparison of the dependence of φ^*_{CP} at generator level and reconstruction level as a function of ΔR. At generator level y_0y_1 is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.



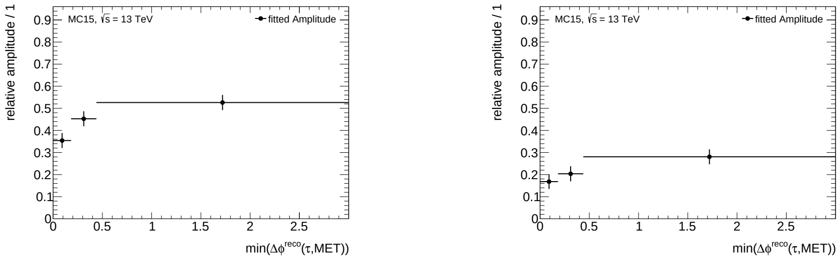
(a) φ^*_{CP} distribution at generator level. (b) φ^*_{CP} distribution at reconstruction level.



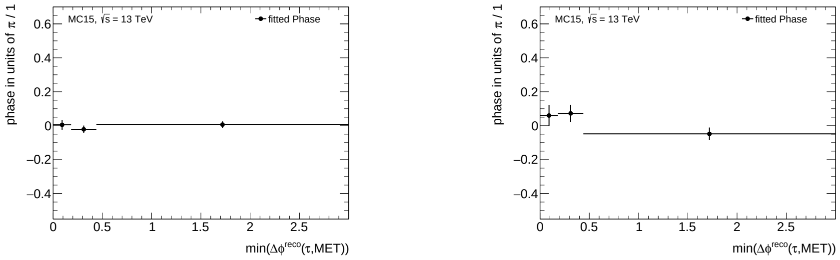
(c) Fit probability at generator level. (d) Fit probability at reconstruction level.

Figure A.11: φ^*_{CP} distribution and fit probability at generator level and at reconstruction level as a function of ΔR. At generator level y_0y_1 is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.

A.5.5 $\min\left(\Delta\phi\left(\tau_{\text{had}}, E_{\text{T}}^{\text{miss}}\right)\right)$

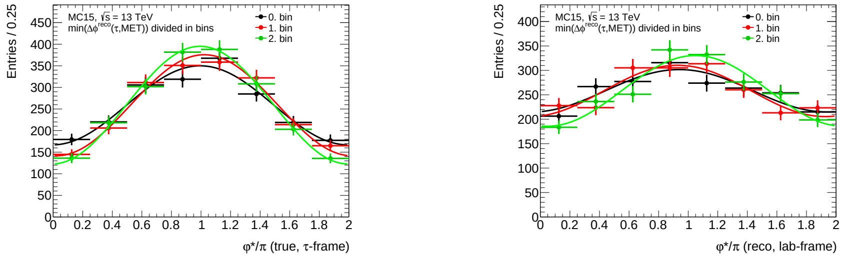


(a) Relative amplitude at generator level. (b) Relative amplitude at reconstruction level.

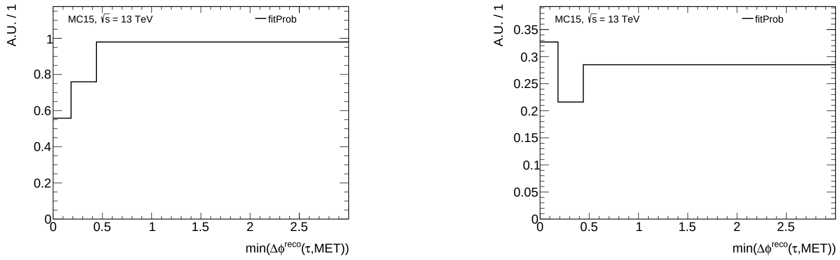


(c) Phase shift at generator level. (d) Phase shift at reconstruction level.

Figure A.12: Comparison of the dependence of φ^*_{CP} at generator level and reconstruction level as a function of $\min\left(\Delta\phi\left(\tau_{\text{had}}, E_{\text{T}}^{\text{miss}}\right)\right)$. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.



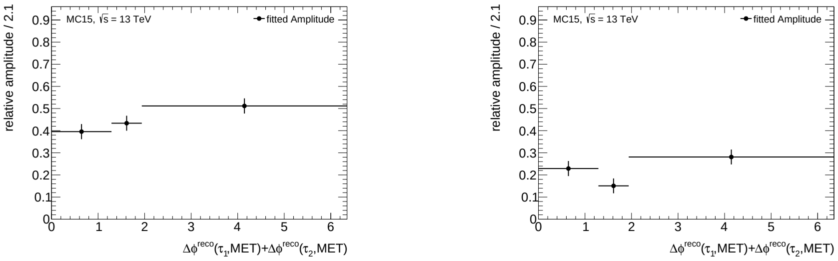
(a) φ^*_{CP} distribution at generator level. (b) φ^*_{CP} distribution at reconstruction level.



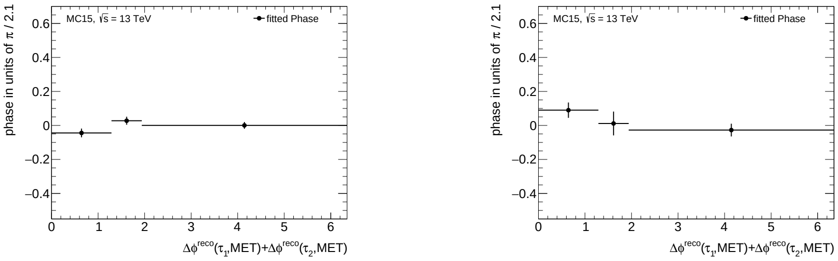
(c) Fit probability at generator level. (d) Fit probability at reconstruction level.

Figure A.13: φ^*_{CP} distribution and fit probability at generator level and at reconstruction level as a function of $\min\left(\Delta\phi\left(\tau_{\text{had}}, E_{\text{T}}^{\text{miss}}\right)\right)$. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.

A.5.6 $\Sigma\left(\Delta\phi\left(\tau_{\mathrm{had}}, E_{\mathrm{T}}^{\mathrm{miss}}\right)\right)$

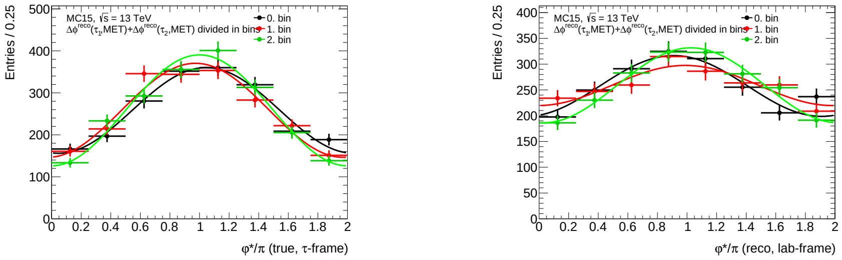


(a) Relative amplitude at generator level. (b) Relative amplitude at reconstruction level.

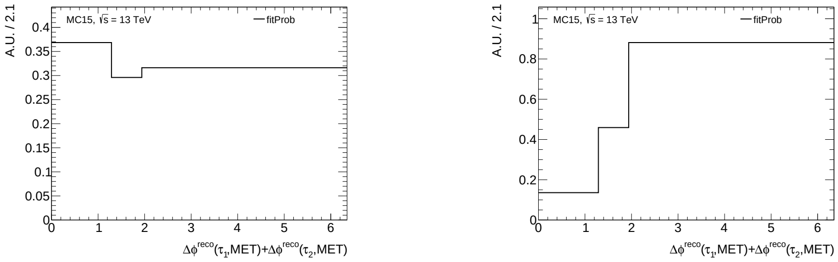


(c) Phase shift at generator level. (d) Phase shift at reconstruction level.

Figure A.14: Comparison of the dependence of φ^*_{CP} at generator level and reconstruction level as a function of $\Sigma\left(\Delta\phi\left(\tau_{\mathrm{had}}, E_{\mathrm{T}}^{\mathrm{miss}}\right)\right)$. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.



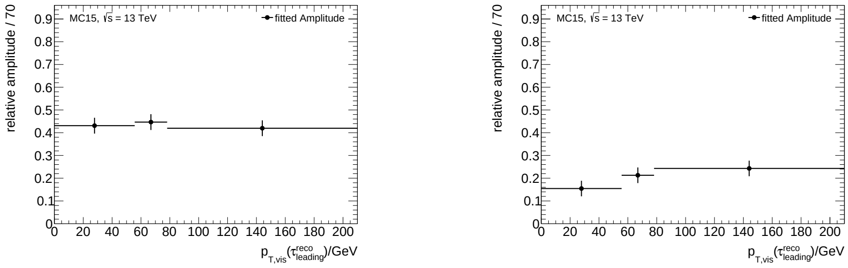
(a) φ^*_{CP} distribution at generator level. (b) φ^*_{CP} distribution at reconstruction level.



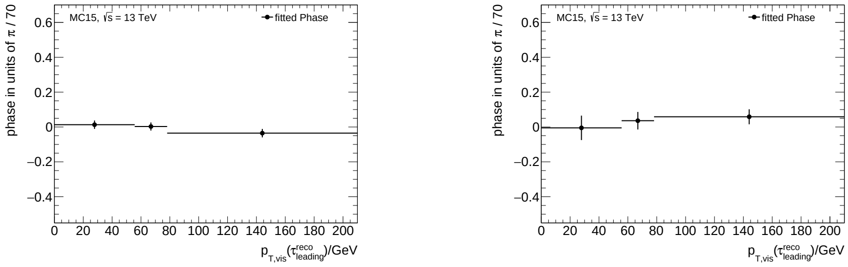
(c) Fit probability at generator level. (d) Fit probability at reconstruction level.

Figure A.15: φ^*_{CP} distribution and fit probability at generator level and at reconstruction level as a function of $\Sigma\left(\Delta\phi\left(\tau_{\mathrm{had}}, E_{\mathrm{T}}^{\mathrm{miss}}\right)\right)$. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.

A.5.7 $p_{\mathrm{T}}^{\tau_0}$

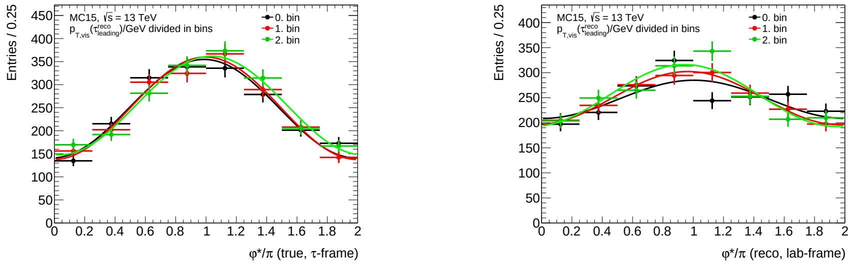


(a) Relative amplitude at generator level.
(b) Relative amplitude at reconstruction level.

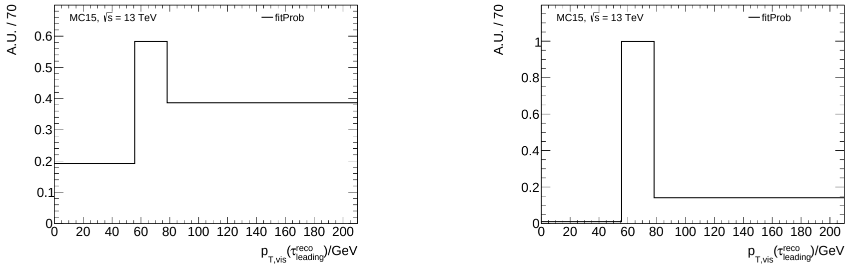


(c) Phase shift at generator level.
(d) Phase shift at reconstruction level.

Figure A.16: Comparison of the dependence of φ^*_{CP} at generator level and reconstruction level as a function of $p_{\mathrm{T}}^{\tau_0}$. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.



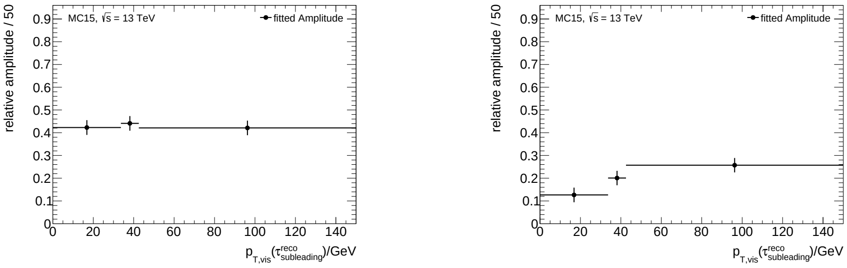
(a) φ^*_{CP} distribution at generator level.
(b) φ^*_{CP} distribution at reconstruction level.



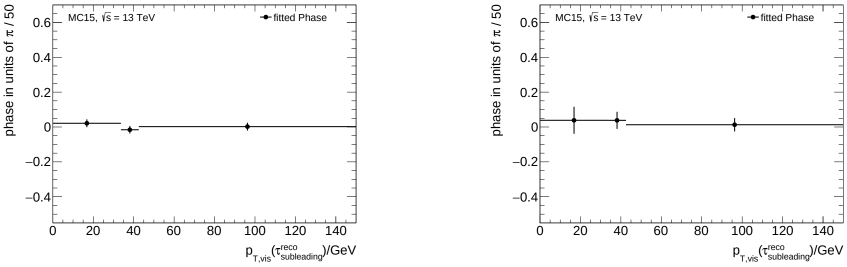
(c) Fit probability at generator level.
(d) Fit probability at reconstruction level.

Figure A.17: φ^*_{CP} distribution and fit probability at generator level and at reconstruction level as a function of $p_{\mathrm{T}}^{\tau_0}$. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.

A.5.8 $p_{\mathrm{T}}^{\tau_1}$

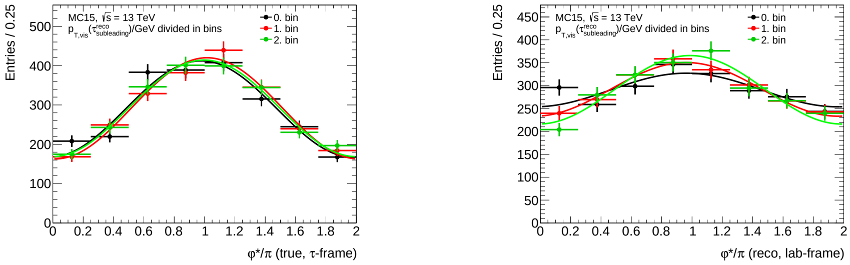


(a) Relative amplitude at generator level. (b) Relative amplitude at reconstruction level.

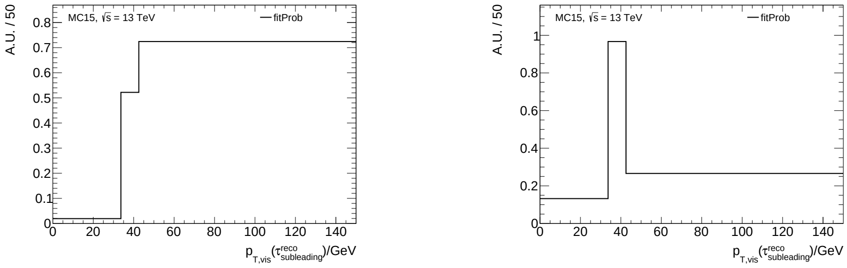


(c) Phase shift at generator level. (d) Phase shift at reconstruction level.

Figure A.18: Comparison of the dependence of φ^*_{CP} at generator level and reconstruction level as a function of $p_{\mathrm{T}}^{\tau_1}$. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.



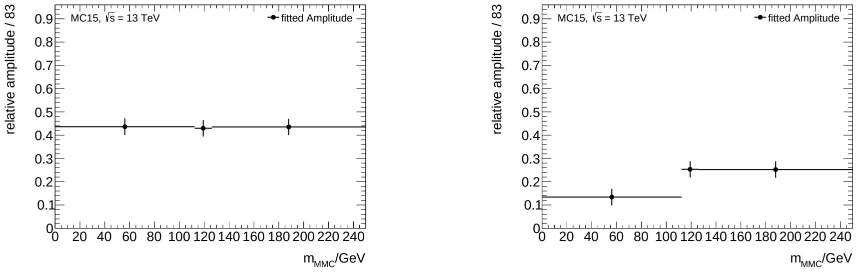
(a) φ^*_{CP} distribution at generator level. (b) φ^*_{CP} distribution at reconstruction level.



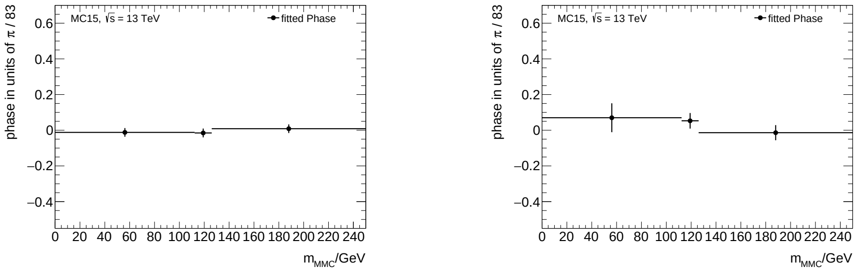
(c) Fit probability at generator level. (d) Fit probability at reconstruction level.

Figure A.19: φ^*_{CP} distribution and fit probability at generator level and at reconstruction level as a function of $p_{\mathrm{T}}^{\tau_1}$. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.

A.5.9 m_{MMC}

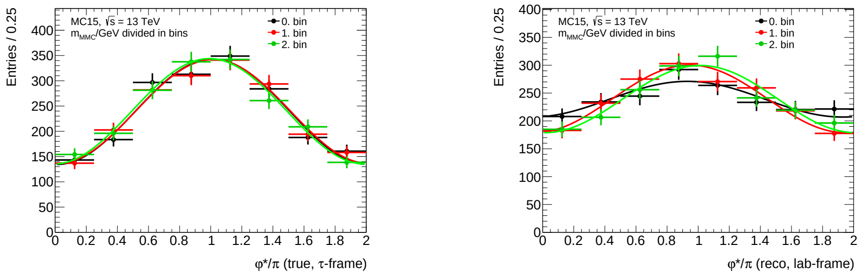


(a) Relative amplitude at generator level. (b) Relative amplitude at reconstruction level.

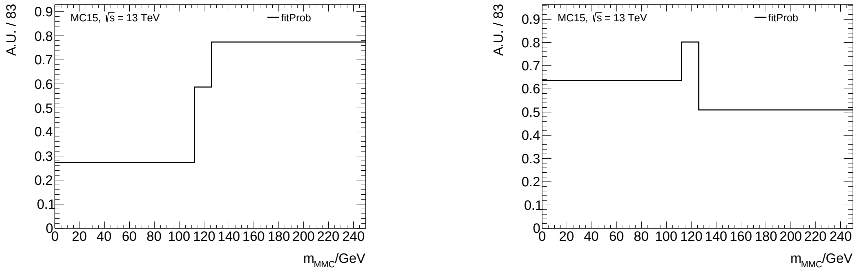


(c) Phase shift at generator level. (d) Phase shift at reconstruction level.

Figure A.20: Comparison of the dependence of φ^*_{CP} at generator level and reconstruction level as a function of m_{MMC}. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.



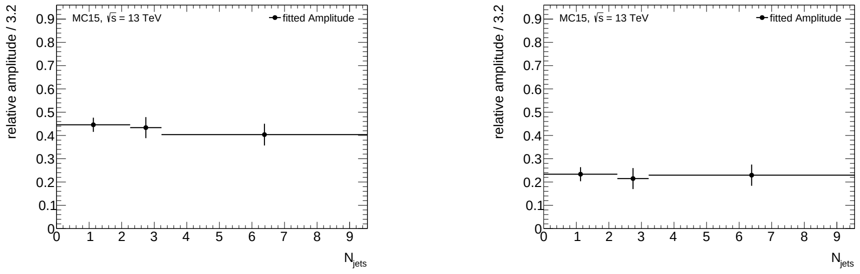
(a) φ^*_{CP} distribution at generator level. (b) φ^*_{CP} distribution at reconstruction level.



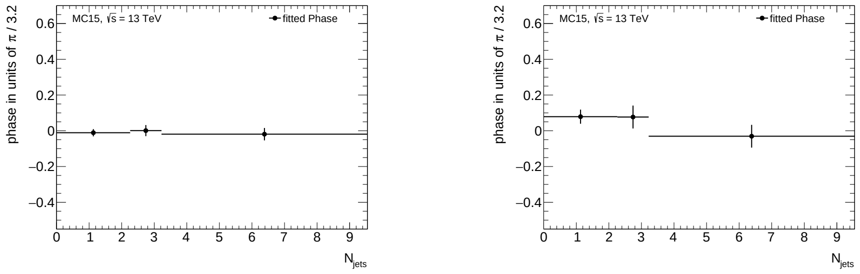
(c) Fit probability at generator level. (d) Fit probability at reconstruction level.

Figure A.21: φ^*_{CP} distribution and fit probability at generator level and at reconstruction level as a function of m_{MMC}. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.

A.5.10 N_{jets}

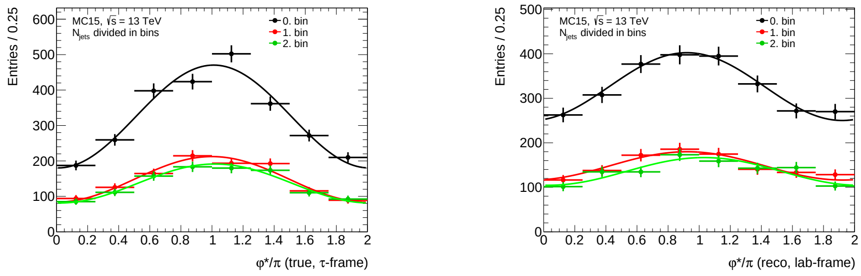


(a) Relative amplitude at generator level. (b) Relative amplitude at reconstruction level.

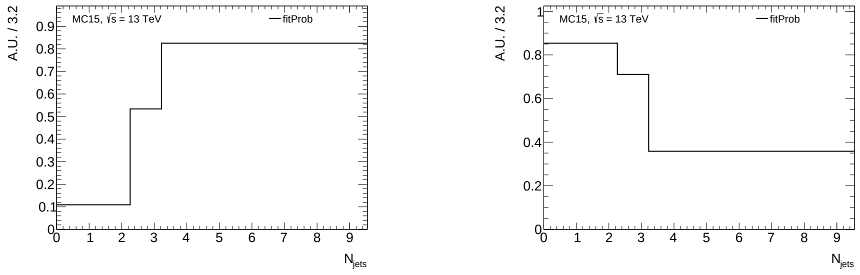


(c) Phase shift at generator level. (d) Phase shift at reconstruction level.

Figure A.22: Comparison of the dependence of φ^*_{CP} at generator level and reconstruction level as a function of N_{jets}. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.



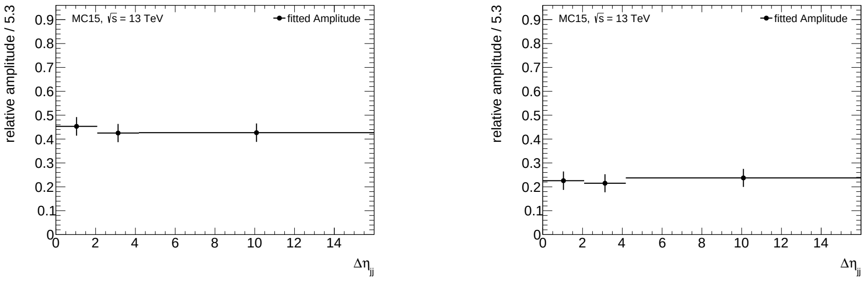
(a) φ^*_{CP} distribution at generator level. (b) φ^*_{CP} distribution at reconstruction level.



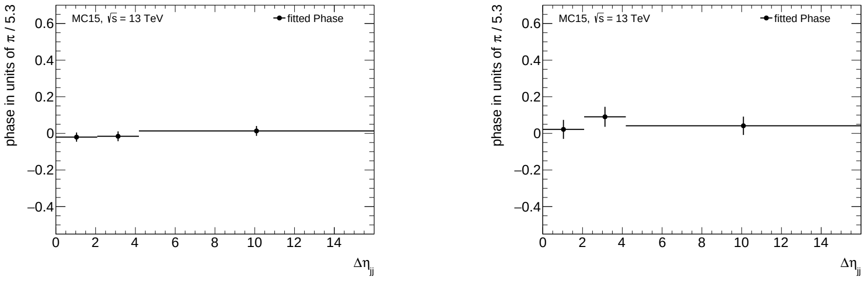
(c) Fit probability at generator level. (d) Fit probability at reconstruction level.

Figure A.23: φ^*_{CP} distribution and fit probability at generator level and at reconstruction level as a function of N_{jets}. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.

A.5.11 $\Delta\eta_{jj}$

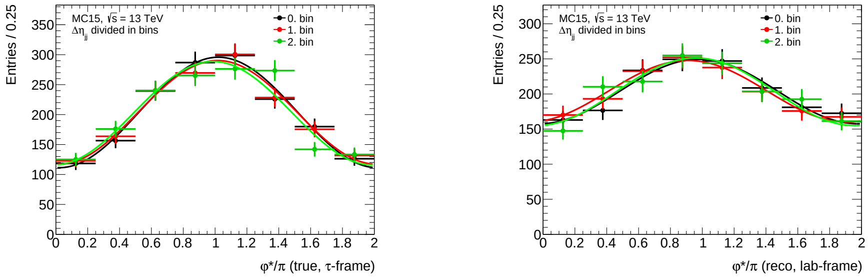


(a) Relative amplitude at generator level. (b) Relative amplitude at reconstruction level.

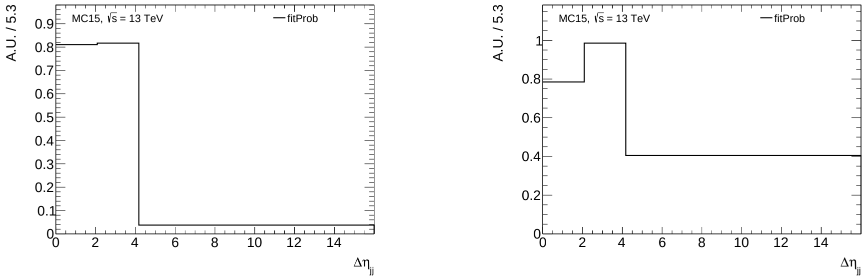


(c) Phase shift at generator level. (d) Phase shift at reconstruction level.

Figure A.24: Comparison of the dependence of φ^*_{CP} at generator level and reconstruction level as a function of $\Delta\eta_{jj}$. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.



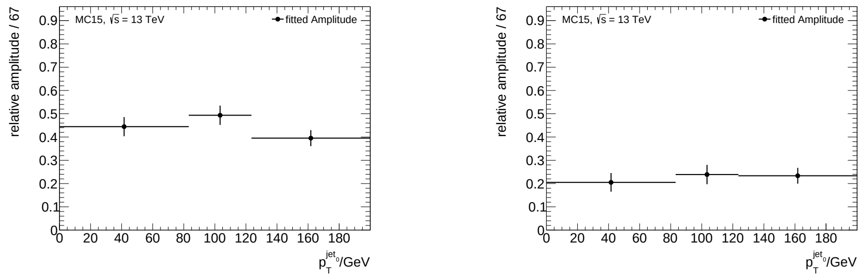
(a) φ^*_{CP} distribution at generator level. (b) φ^*_{CP} distribution at reconstruction level.



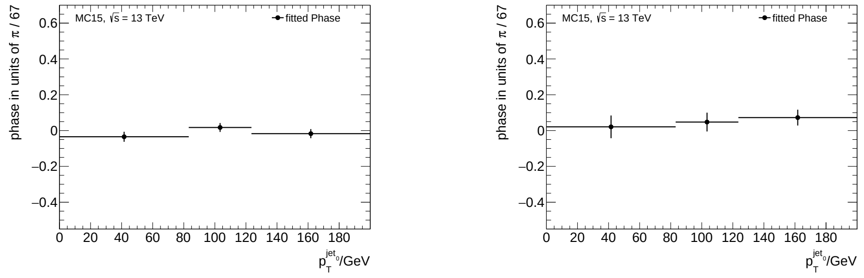
(c) Fit probability at generator level. (d) Fit probability at reconstruction level.

Figure A.25: φ^*_{CP} distribution and fit probability at generator level and at reconstruction level as a function of $\Delta\eta_{jj}$. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.

A.5.12 $p_{\mathrm{T}}^{jet_0}$

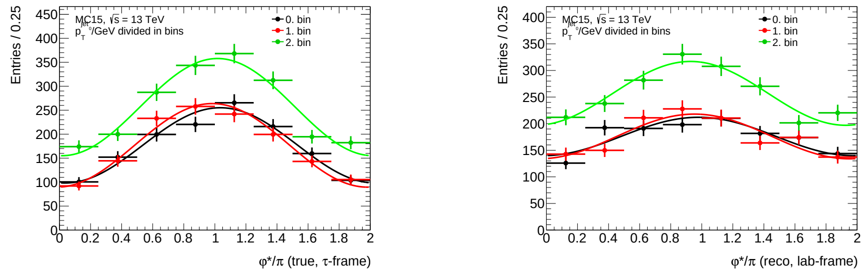


(a) Relative amplitude at generator level. (b) Relative amplitude at reconstruction level.

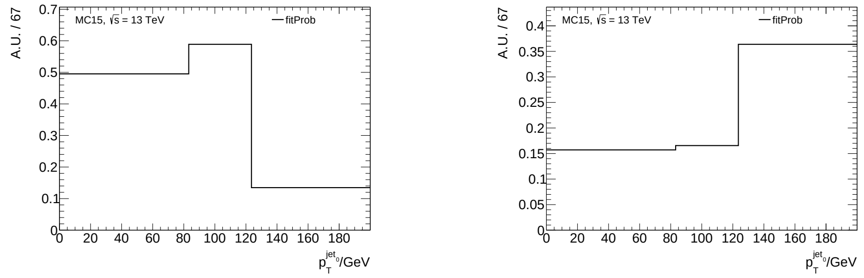


(c) Phase shift at generator level. (d) Phase shift at reconstruction level.

Figure A.26: Comparison of the dependence of φ^*_{CP} at generator level and reconstruction level as a function of $p_{\mathrm{T}}^{jet_0}$. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.



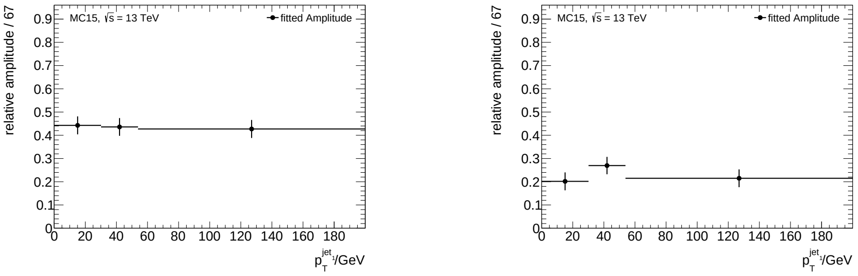
(a) φ^*_{CP} distribution at generator level. (b) φ^*_{CP} distribution at reconstruction level.



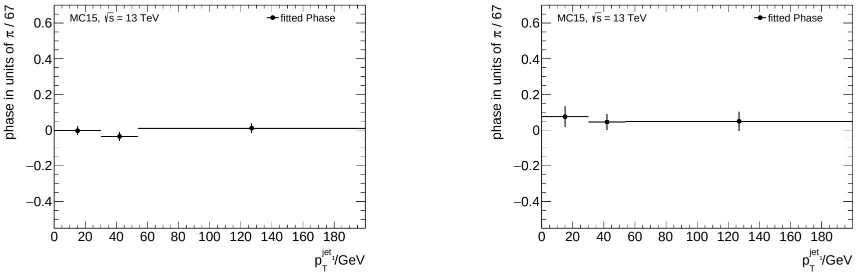
(c) Fit probability at generator level. (d) Fit probability at reconstruction level.

Figure A.27: φ^*_{CP} distribution and fit probability at generator level and at reconstruction level as a function of $p_{\mathrm{T}}^{jet_0}$. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.

A.5.13 $p_{\mathrm{T}}^{jet_1}$

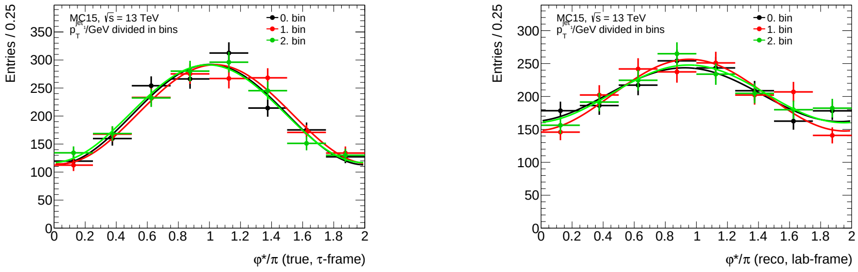


(a) Relative amplitude at generator level. (b) Relative amplitude at reconstruction level.

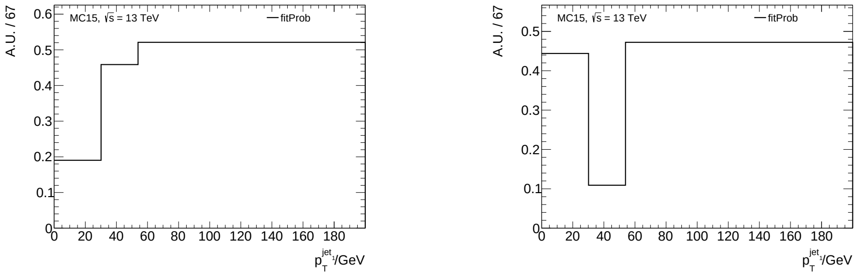


(c) Phase shift at generator level. (d) Phase shift at reconstruction level.

Figure A.28: Comparison of the dependence of φ^*_{CP} at generator level and reconstruction level as a function of $p_{\mathrm{T}}^{jet_1}$. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.



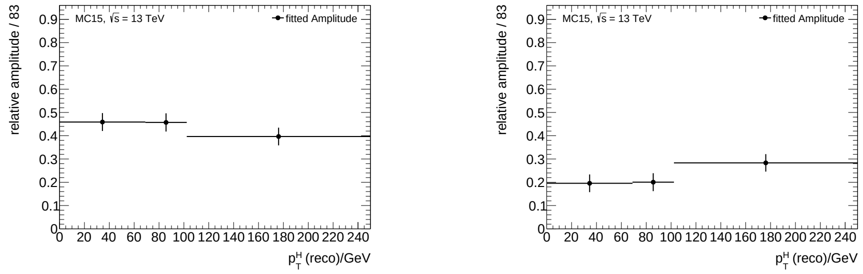
(a) φ^*_{CP} distribution at generator level. (b) φ^*_{CP} distribution at reconstruction level.



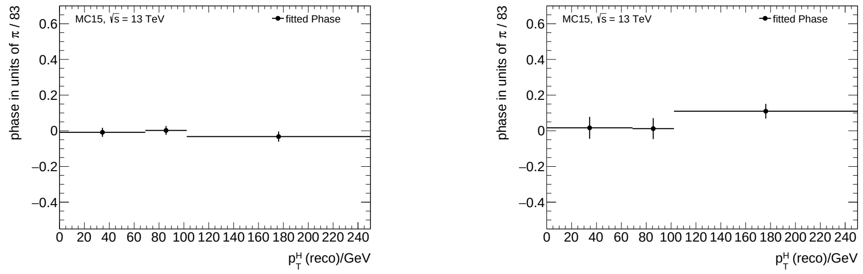
(c) Fit probability at generator level. (d) Fit probability at reconstruction level.

Figure A.29: φ^*_{CP} distribution and fit probability at generator level and at reconstruction level as a function of $p_{\mathrm{T}}^{jet_1}$. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.

A.5.14 p_{T}^{H}

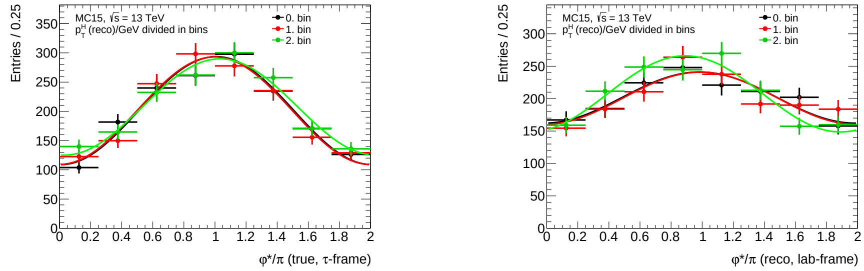


(a) Relative amplitude at generator level. (b) Relative amplitude at reconstruction level.

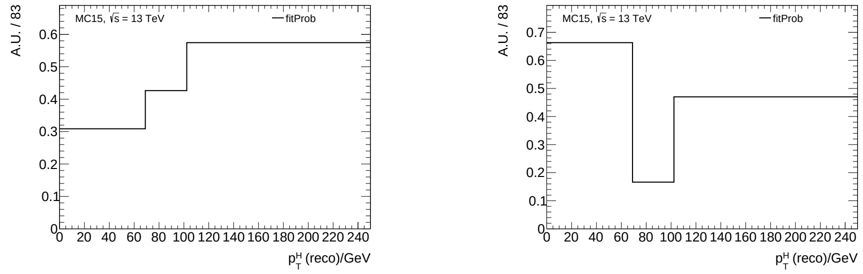


(c) Phase shift at generator level. (d) Phase shift at reconstruction level.

Figure A.30: Comparison of the dependence of φ^*_{CP} at generator level and reconstruction level as a function of p_{T}^{H}. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.



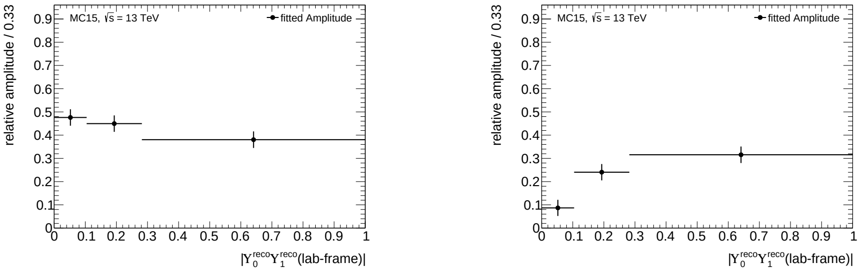
(a) φ^*_{CP} distribution at generator level. (b) φ^*_{CP} distribution at reconstruction level.



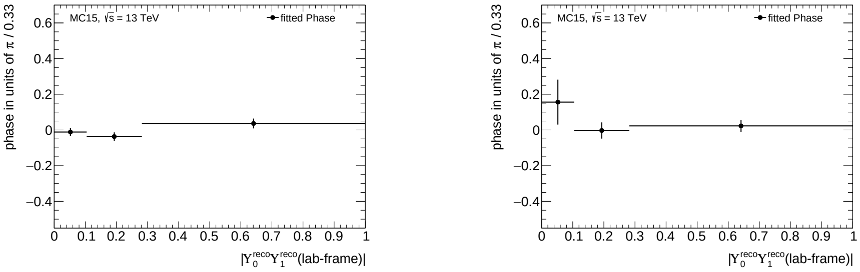
(c) Fit probability at generator level. (d) Fit probability at reconstruction level.

Figure A.31: φ^*_{CP} distribution and fit probability at generator level and at reconstruction level as a function of p_{T}^{H}. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.

A.5.15 $|y_0 y_1|$

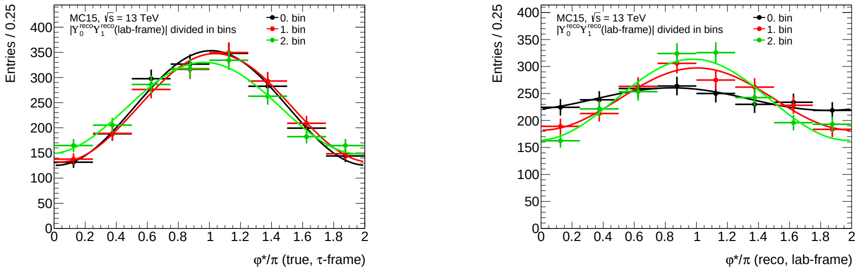


(a) Relative amplitude at generator level. (b) Relative amplitude at reconstruction level.

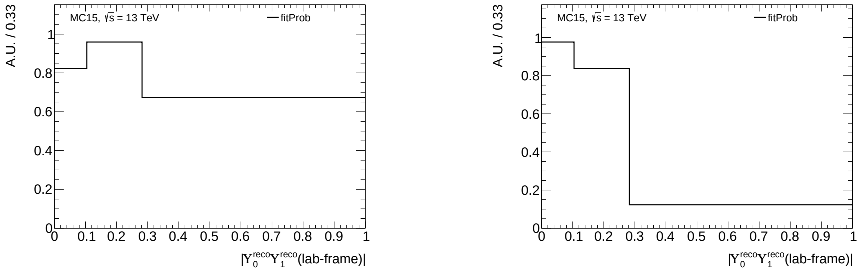


(c) Phase shift at generator level. (d) Phase shift at reconstruction level.

Figure A.32: Comparison of the dependence of φ^*_{CP} at generator level and reconstruction level as a function of $|y_0 y_1|$. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.



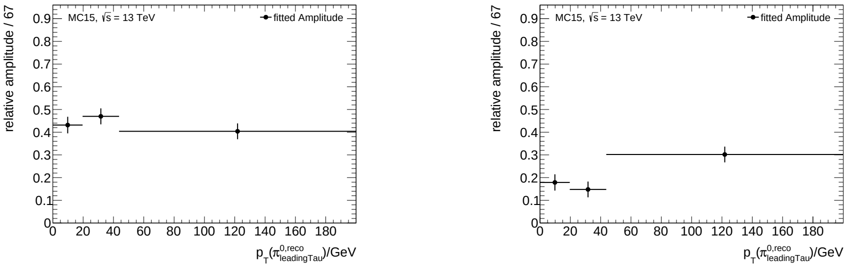
(a) φ^*_{CP} distribution at generator level. (b) φ^*_{CP} distribution at reconstruction level.



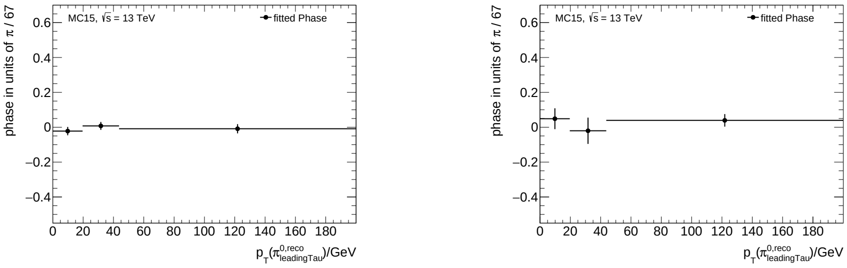
(c) Fit probability at generator level. (d) Fit probability at reconstruction level.

Figure A.33: φ^*_{CP} distribution and fit probability at generator level and at reconstruction level as a function of $|y_0 y_1|$. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.

A.5.16 $p_{\mathrm{T}}^{\pi^0_0}$

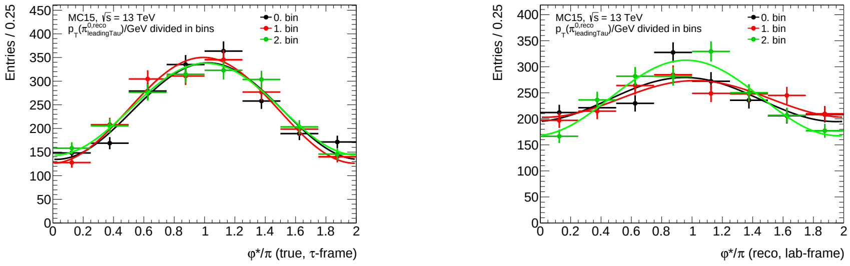


(a) Relative amplitude at generator level. (b) Relative amplitude at reconstruction level.

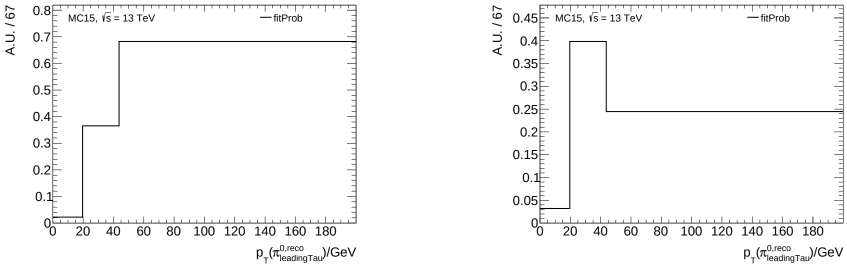


(c) Phase shift at generator level. (d) Phase shift at reconstruction level.

Figure A.34: Comparison of the dependence of φ^*_{CP} at generator level and reconstruction level as a function of $p_{\mathrm{T}}^{\pi^0_0}$. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.



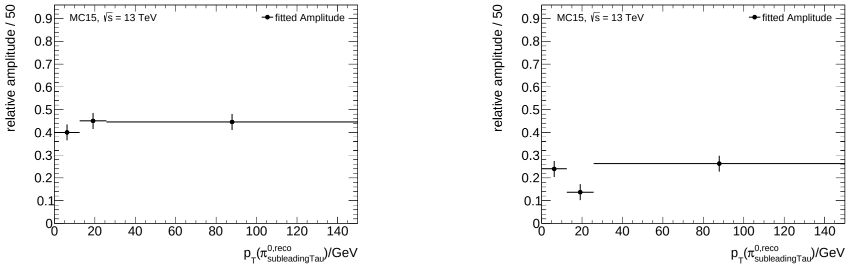
(a) φ^*_{CP} distribution at generator level. (b) φ^*_{CP} distribution at reconstruction level.



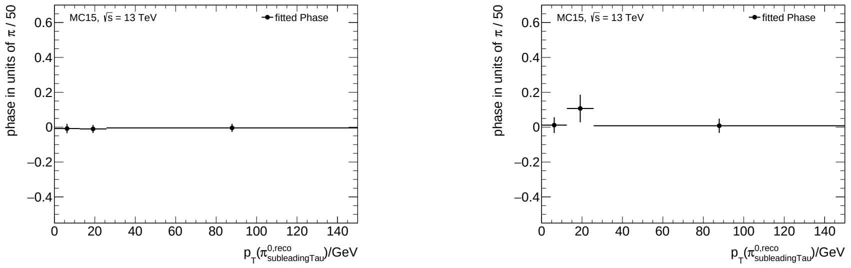
(c) Fit probability at generator level. (d) Fit probability at reconstruction level.

Figure A.35: φ^*_{CP} distribution and fit probability at generator level and at reconstruction level as a function of $p_{\mathrm{T}}^{\pi^0_0}$. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.

A.5.17 $p_{\mathrm{T}}^{\pi_1^0}$

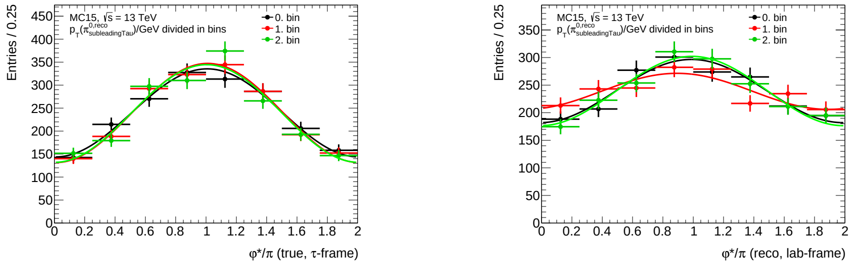


(a) Relative amplitude at generator level. (b) Relative amplitude at reconstruction level.

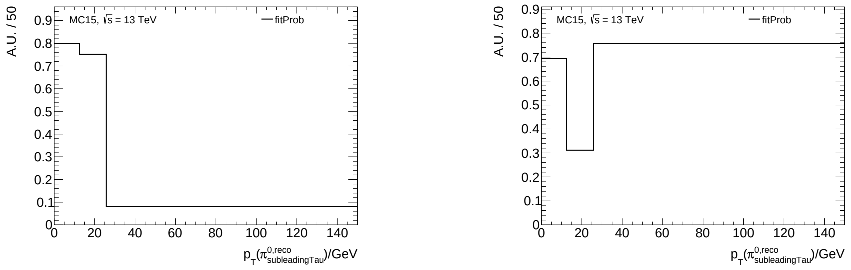


(c) Phase shift at generator level. (d) Phase shift at reconstruction level.

Figure A.36: Comparison of the dependence of φ^*_{CP} at generator level and reconstruction level as a function of $p_{\mathrm{T}}^{\pi_1^0}$. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.



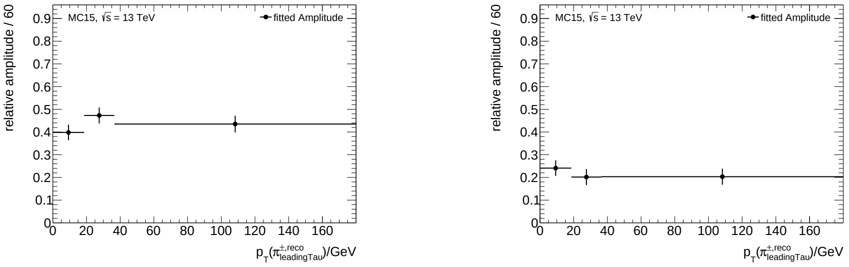
(a) φ^*_{CP} distribution at generator level. (b) φ^*_{CP} distribution at reconstruction level.



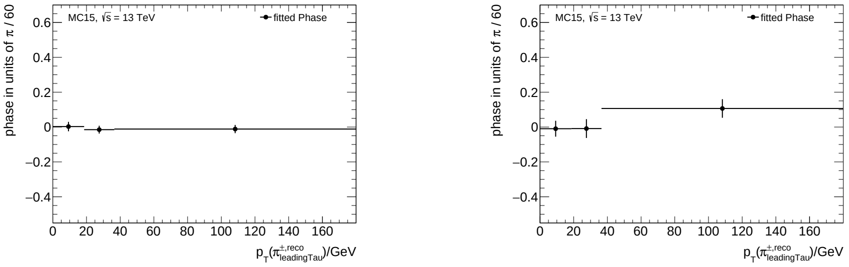
(c) Fit probability at generator level. (d) Fit probability at reconstruction level.

Figure A.37: φ^*_{CP} distribution and fit probability at generator level and at reconstruction level as a function of $p_{\mathrm{T}}^{\pi_1^0}$. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.

A.5.18 $p_{\mathrm{T}}^{\pi^{\pm}_{0}}$

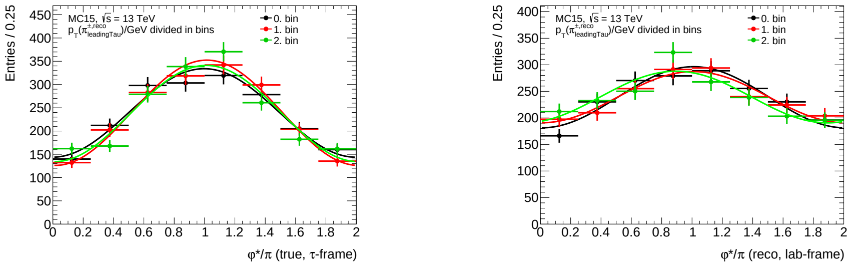


(a) Relative amplitude at generator level. (b) Relative amplitude at reconstruction level.

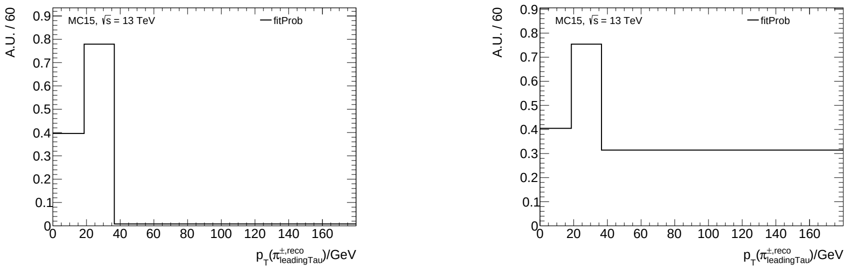


(c) Phase shift at generator level. (d) Phase shift at reconstruction level.

Figure A.38: Comparison of the dependence of φ^*_{CP} at generator level and reconstruction level as a function of $p_{\mathrm{T}}^{\pi^{\pm}_{0}}$. At generator level y_0y_1 is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.



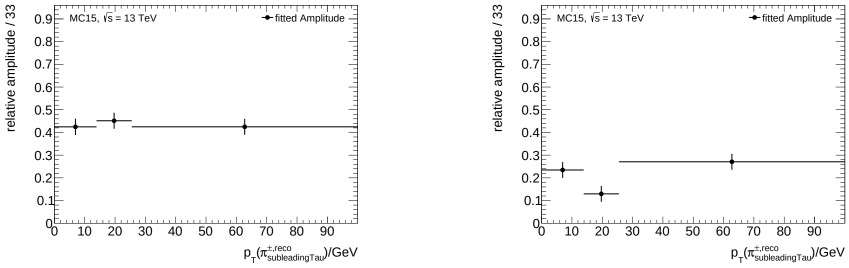
(a) φ^*_{CP} distribution at generator level. (b) φ^*_{CP} distribution at reconstruction level.



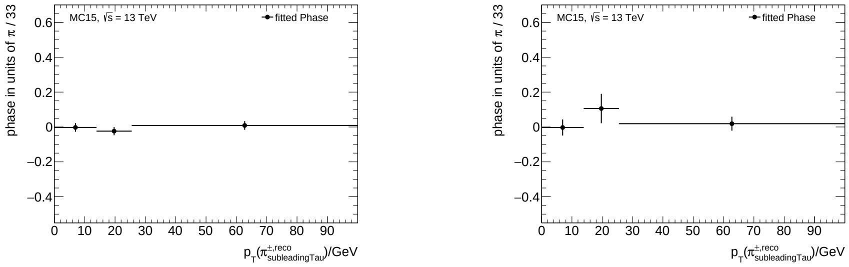
(c) Fit probability at generator level. (d) Fit probability at reconstruction level.

Figure A.39: φ^*_{CP} distribution and fit probability at generator level and at reconstruction level as a function of $p_{\mathrm{T}}^{\pi^{\pm}_{0}}$. At generator level y_0y_1 is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.

A.5.19 $p_{\mathrm{T}}^{\pi_1^\pm}$

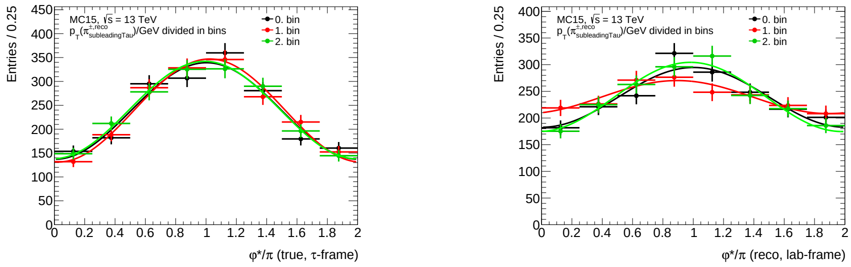


(a) Relative amplitude at generator level. (b) Relative amplitude at reconstruction level.

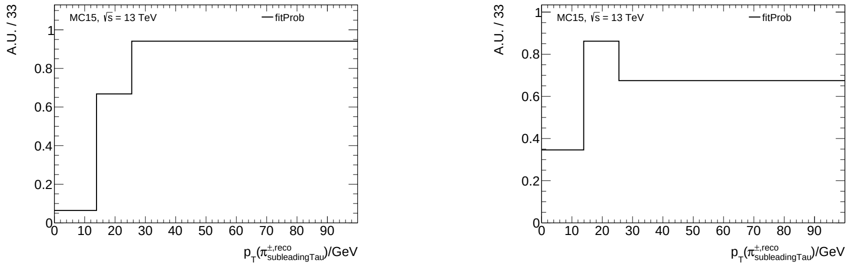


(c) Phase shift at generator level. (d) Phase shift at reconstruction level.

Figure A.40: Comparison of the dependence of φ^*_{CP} at generator level and reconstruction level as a function of $p_{\mathrm{T}}^{\pi_1^\pm}$. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.



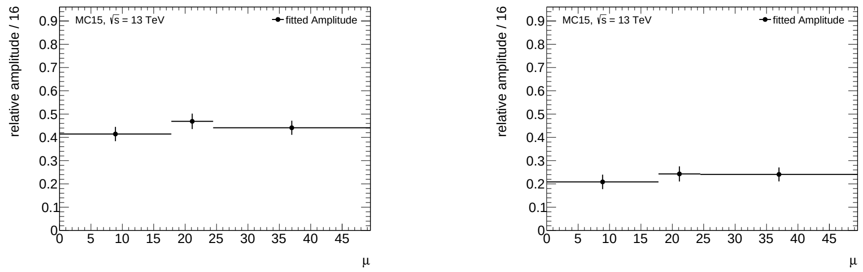
(a) φ^*_{CP} distribution at generator level. (b) φ^*_{CP} distribution at reconstruction level.



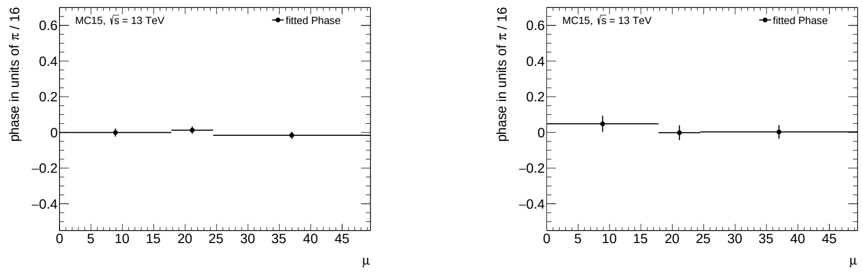
(c) Fit probability at generator level. (d) Fit probability at reconstruction level.

Figure A.41: φ^*_{CP} distribution and fit probability at generator level and at reconstruction level as a function of $p_{\mathrm{T}}^{\pi_1^\pm}$. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.

A.5.20 μ

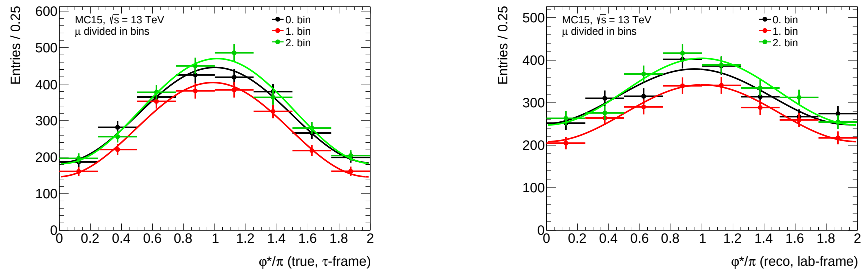


(a) Relative amplitude at generator level. (b) Relative amplitude at reconstruction level.

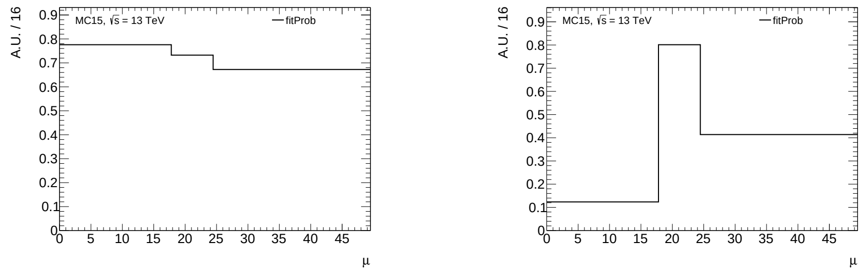


(c) Phase shift at generator level. (d) Phase shift at reconstruction level.

Figure A.42: Comparison of the dependence of φ^*_{CP} at generator level and reconstruction level as a function of μ. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.



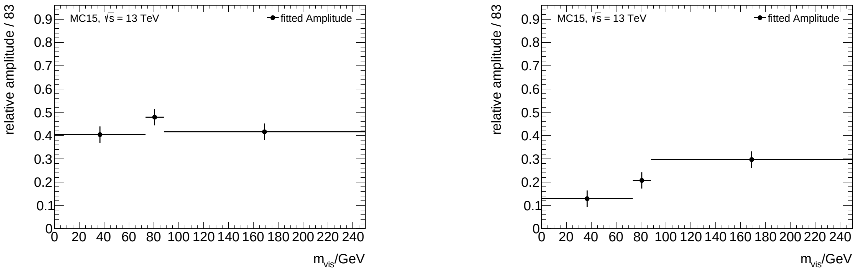
(a) φ^*_{CP} distribution at generator level. (b) φ^*_{CP} distribution at reconstruction level.



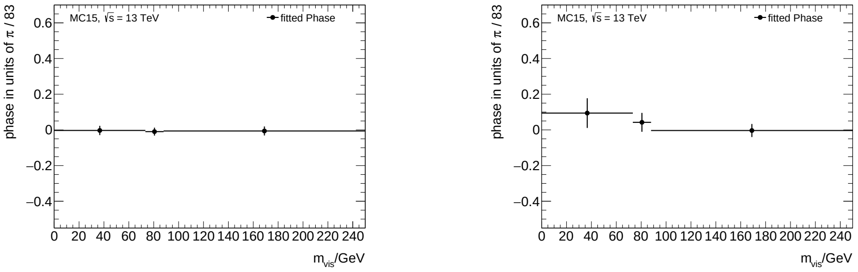
(c) Fit probability at generator level. (d) Fit probability at reconstruction level.

Figure A.43: φ^*_{CP} distribution and fit probability at generator level and at reconstruction level as a function of μ. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.

A.5.21 m_{vis}

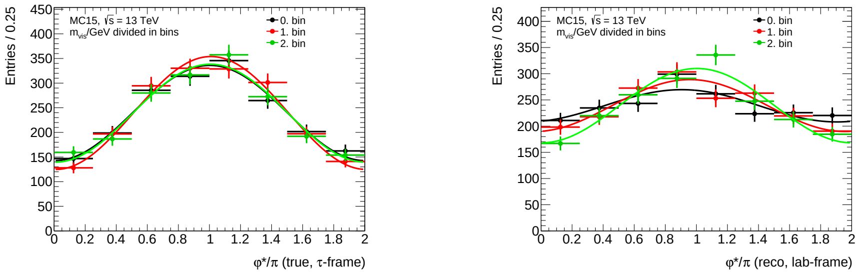


(a) Relative amplitude at generator level. (b) Relative amplitude at reconstruction level.

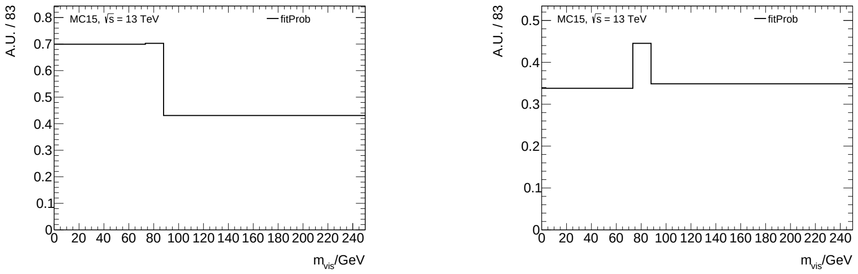


(c) Phase shift at generator level. (d) Phase shift at reconstruction level.

Figure A.44: Comparison of the dependence of φ^*_{CP} at generator level and reconstruction level as a function of m_{vis}. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.



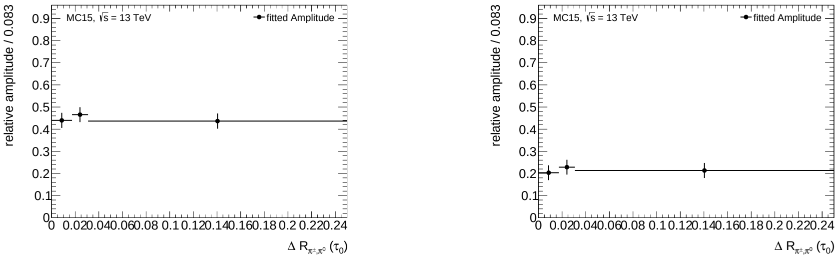
(a) φ^*_{CP} distribution at generator level. (b) φ^*_{CP} distribution at reconstruction level.



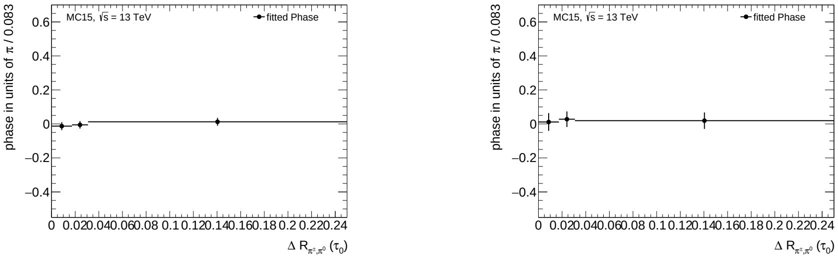
(c) Fit probability at generator level. (d) Fit probability at reconstruction level.

Figure A.45: φ^*_{CP} distribution and fit probability at generator level and at reconstruction level as a function of m_{vis}. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.

A.5.22 $\Delta R\left(\pi^{\pm},\pi^{0}\right)$

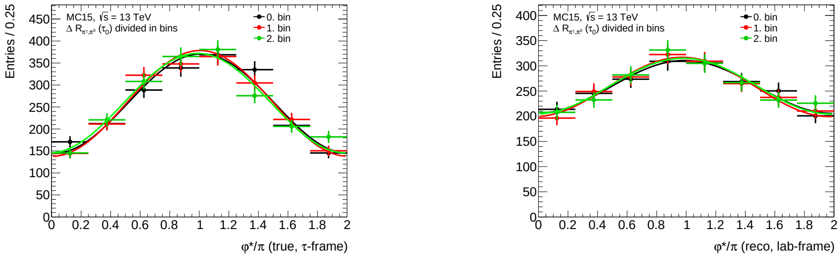


(a) Relative amplitude at generator level. (b) Relative amplitude at reconstruction level.

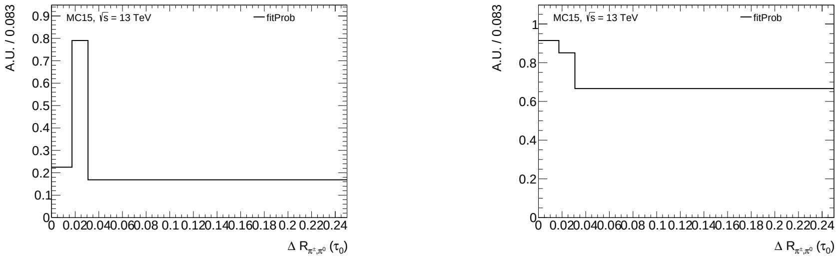


(c) Phase shift at generator level. (d) Phase shift at reconstruction level.

Figure A.46: Comparison of the dependence of φ^*_{CP} at generator level and reconstruction level as a function of $\Delta R\left(\pi^{\pm},\pi^{0}\right)$. At generator level y_0y_1 is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.



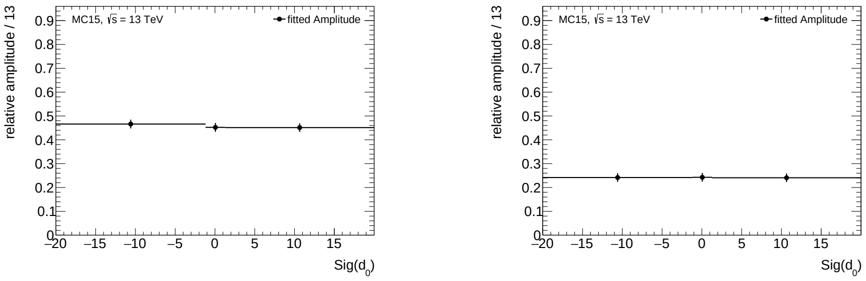
(a) φ^*_{CP} distribution at generator level. (b) φ^*_{CP} distribution at reconstruction level.



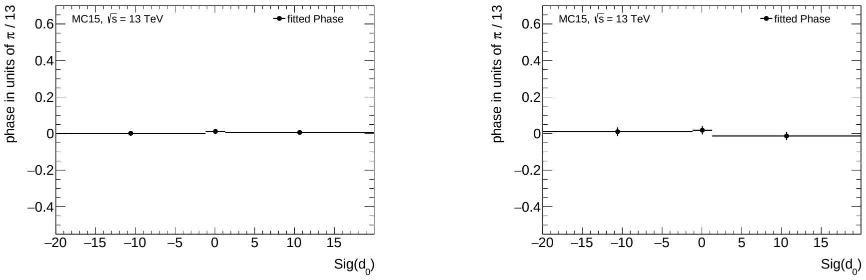
(c) Fit probability at generator level. (d) Fit probability at reconstruction level.

Figure A.47: φ^*_{CP} distribution and fit probability at generator level and at reconstruction level as a function of $\Delta R\left(\pi^{\pm},\pi^{0}\right)$. At generator level y_0y_1 is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.

A.5.23 d_0

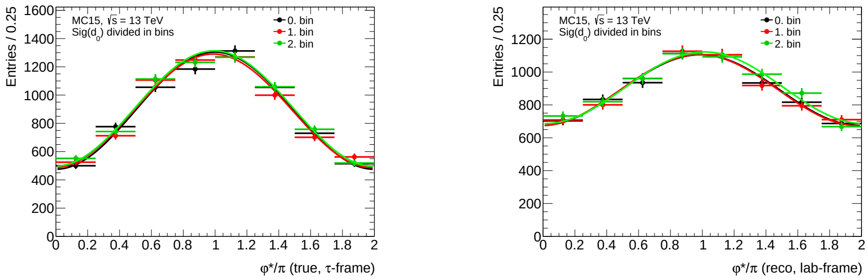


(a) Relative amplitude at generator level. (b) Relative amplitude at reconstruction level.

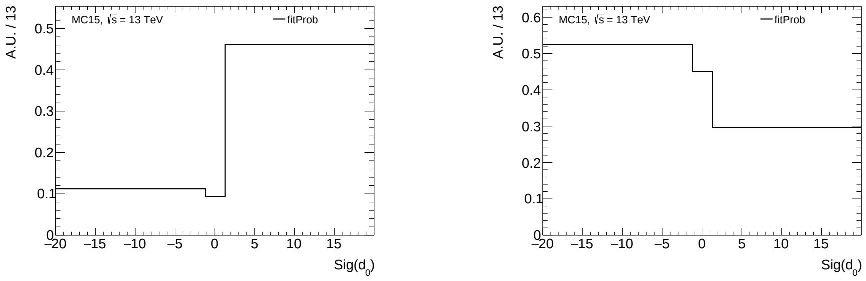


(c) Phase shift at generator level. (d) Phase shift at reconstruction level.

Figure A.48: Comparison of the dependence of φ^{*}_{CP} at generator level and reconstruction level as a function of d_0. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.



(a) φ^{*}_{CP} distribution at generator level. (b) φ^{*}_{CP} distribution at reconstruction level.



(c) Fit probability at generator level. (d) Fit probability at reconstruction level.

Figure A.49: φ^{*}_{CP} distribution and fit probability at generator level and at reconstruction level as a function of d_0. At generator level $y_0 y_1$ is calculated in the respective τ_{had} frame and at reconstruction level in the laboratory frame.

List of Figures

List of Tables
