Fake Tau Estimation using Monte Carlo Likelihood Techniques in the tHq Lep-Had Channel

Marvin Schmitz

Masterarbeit in Physik
angefertigt im Physikalischen Institut

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

September 2021

I hereby declare that this thesis was formulated by myself and that no sources or tools other than those cited were used.

Bonn, ...
Date Signature

1. Gutachter: Prof. Dr. Ian C. Brock
2. Gutachterin: Priv.-Doz. Dr. Philip Bechtle

Acknowledgements

I would like to thank Prof. Dr. Ian C. Brock for giving me the opportunity to write my thesis as a member of his group and for providing me the circumstances to do my research. I would also like to thank Priv.-Doz. Dr. Philip Bechtle for agreeing to perform the duties of a second referee for my thesis.

I would also like to thank all members of the Brock group, past and present, who supported me along the way. A special thanks is dedicated to Federico Diaz Capriles, Tanja Holm, Christian Kirfel and Dr. Oleh Kivernyk who performed the role as my advisor and developed the research for this thesis with me.

I would also like to thank my friends and family for supporting me over the years and especially Judith Meyer who always supported me when I needed it.

Last but not least, i would like to thank the University of Bonn for providing the conditions and a place for me to learn about physics over the course of my studies.

Contents

CHAPTER 1

Introduction

The investigation of falsely reconstructed tau leptons is an essential part of the background estimation for high energy particle physics analysis. This thesis investigates the abundance of those falsely reconstructed tau leptons. The goal is to determine scale factors that improve the ratio between Monte Carlo simulations and data. This is done by employing binned likelihood techniques. The estimation of tau fakes is a challenging project that requires the investigation and application of a variety of techniques in the hope of finding a satisfying approach to correct the discrepancies between simulation and data. This investigation of methods based on likelihood techniques is thought to be contribution to this effort while at the same time constraining the problems arising from the particular kind of tau fakes encountered in the tHq analysis.

The aforementioned process tHq is our process of interest. Particularly the case in which the Higgs decays into two tau leptons. The channels that are investigated in this process are the two multi lepton channels in which either two hadronic taus and one light lepton are reconstructed or in which two light leptons and one hadronic tau are reconstructed.

The structure of this thesis is as follows. In chapter [2] the theoretical concepts for this thesis are discussed. This includes The Standard Model of particle physics, the tHq process and physics related to tau leptons and tau fakes. In addition to this the nature of QCD jets is investigated and the mathematical machine used to perform the estimation of the fake tau abundances, the likelihood fit, is explained. In the next chapter [3] the Large Hadron Collider and the ATLAS detector are explained. This includes the reconstruction and identification of particles in the ATLAS experiment. After these explanations in chapter [4] the methods used in this thesis are described and the setup is explained. Finally in chapter [5] the results are presented and discussed before concluding the investigations of this thesis. Additional information about the work presented in this thesis can be found in the appendices A, B, C, D, E. A list of figures and tables created and literature usedin this thesis are given as well.

CHAPTER 2

Theoretical Concepts

2.1 The Standard Model of Particle Physics

2.1.1 Introduction

As this thesis investigates the abundance of falsely reconstructed τ leptons in the ATLAS detector, it is prudent to discuss the properties of the underlying theory of particle physics. As such this section describes the Standard Model of particle physics, in a detail that is relevant to this thesis. The motivations and ideas of the Standard Model as a quantum field theory are discussed, before assessing the Standard Model's particle content and the particle's properties and behavior. In the end special attention is given to physics concerning τ leptons and processes related to falsely reconstructed τ's.

The Standard Model of particle physics describes the behavior of nature at its most fundamental level. It predicts the existence of indivisible elementary particles, such as the leptons, quarks and exchange bosons and describes their interactions. The Standard Model is formulated as a quantum field theory that obeys a certain set of gauge symmetries in addition to Lorentz invariance. Gauge symmetries are actions from a group on the set of fields that leave the action of the theory invariant and therefore the equations of motion as well. At the highest level of formulation lies the Standard Model Lagrangian. It contains a mixture of kinematic and interaction terms between the fermionic particle content and the existing gauge bosons. Mass terms arise from the coupling of the fermionic content and the gauge bosons with a spin-0 scalar boson that spontaneously breaks the symmetry due to its potential term. Additionally, there are consistency terms arising from quantum field theory, the gauge properties and the non-abelian nature of the symmetry group. These terms are for example the Faddeev–Popov ghost terms. Io test the properties of the predicted particles their interaction can be studied using scattering experiments. To make predictions about the outcome of these experiments based on the Standard Model a perturbative approach to quantum field theory is used. An intuitive and compact way to describe these interactions in perturbation theory are Feynman diagrams. As major references [5] and [6] have been used.

2.1.2 The Standard Model as a Quantum Field Theory

Why a field theory is necessary?

In non-relativistic quantum mechanics one usually quantizes the dynamics of systems of particles. In quantum field theory one applies the laws of quantum mechanics to a system of fields. When describing processes in high energy particles physics, we need to consider special relativity. The inclusion of special relativity necessitates the usage of a quantum theory of fields and not of particles. To illustrate this point we argue similarly to [5] pp. 13-14. The known equation

$$E = mc^2 \tag{2.1}$$

relates the energy to the mass and the speed of light squared. This allows for the production of particle-antiparticle pairs from the vacuum. Even for energies below the required particle-antiparticle pair production energy, multiparticle states appear. An example for this are intermediate states in perturbation theory. As such a theory of high energy particle physics, has to be a multiparticle theory and therefore a field theory. A more rigorous argument can be made by looking at causality. For example it can be calculated that the probability for finding a freely propagating particle outside its own light-cone is nonzero and, as such violates causality. Quantum field theory solves this by canceling the amplitudes of the particle and antiparticle propagating outside the light-cone since the propagation of a particle across a spacelike interval is indistinguishable from that of an antiparticle propagating in the opposite direction. In this case causality is preserved. As such a high energy particle physics theory, can be motivated as a quantum field theory since it has a natural way of including multiparticle states and allowing transitions between them.

The ideas of Quantum Field Theory

The idea of quantum field theory is to define a Lagrangian density as a function of a set of fields describing the dynamics of the problem. The action can be defined as the integration of that Lagrangian density over an appropiate measure in space-time. The Lagrangian is chosen to fulfill Lorentz invariance and the required symmetries of the theory. By applying the principle of Hamilton [7] pp. $2 - 4$ we can find the equations of motion governing the dynamics of these fields. From these and using canonical commutation relations motivated by classical mechanics similar to [5] p. 20 we can write down solutions for the equation of motions for the quantum fields. Quantum fields can be differentiated by their transformation behavior under Lorentz transformations which can be related to their spin [5]. The Standard Model contains spin-0, spin-1 and spin-1/2 fields. Solutions for the so called Klein-Gordon spin-0 field can be found in [5] pp. 19-21 and for the spin-1/2 Dirac field in [5] pp. 58-62. The solutions for spin-1 gauge boson fields can be found in [5] pp. 294-297.

2.1.3 The Lagrangian of the Standard Model

As mentioned in the introduction, at the highest level of the Standard Model lies its Lagrangian. By constructing it to be invariant under the symmetry group and Lorentz transformations we assure these qualities in the equations of motions and the resulting quantum fields. The symmetry group of the SM is

$$G_{\mathrm{SM}} = U(1)_{\mathrm{Y}} \times SU(2)_{\mathrm{L}} \times SU(3)_{\mathrm{C}} \,. \tag{2.2}$$

In the following we will dissect the symmetry group and describe its parts separately. We will also relate the symmetry group for each part to the experimentally observed physics and phenomenological interpretation before putting it all back together to discuss the SM as a whole.

The $U(1)$ group: Electromagnetic Interaction

The first phenomenological force of nature governing the behavior of matter that we want to discuss is electromagnetism. In the standard model electromagnetism and the weak force are formulated in union and we will take a closer look at this unification later. For now we briefly analyze the important parts of the electromagnetic Lagrangian and its symmetry group. Electric and magnetic forces are known for a relatively long time. Magnetism has been observed in some form since Ancient Greece and electric forces are known for centuries as well. The equations describing the dynamics of electromagnetism have been unified and formulated as a whole by Maxwell in the 19th century they can be found here [8]. The Maxwell Equations can also be derived by assuming a Lagrangian fulfilling Lorentz invariance and invariance under the $U(1)$ gauge transformation [8]. A concise way to formulate the dynamics of the Maxwell equations is by introducing a Lorentz invariant quantity made from the vector potential of the electromagnetic field A_μ [8]. This quantity is called the electromagnetic field tensor

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \ . \tag{2.3}$$

By contracting the field tensor, we have a Lorentz invariant term that will be invariant under a $U(1)$ gauge transformation if the transformation property of A_μ is chosen accordingly. The resulting Lagrangian describing the propagation of the electromagnetic field in a vacuum then follows as

$$\mathcal{L}_{\text{kin, em}} = \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \ . \tag{2.4}$$

The full electromagnetic part of the SM Lagrangian consists of a free Dirac Lagrangian [5] p. 52 and the kinematic field term 2.4. Therefore the quantum electrodynamics part would look like

$$\mathcal{L}_{\text{QED}} = \mathcal{L}_{\text{kin, Dirac}} + \mathcal{L}_{\text{int}} + \mathcal{L}_{\text{kin, em}} = i\overline{\psi}\gamma^\mu D_\mu \psi - e\overline{\psi}\gamma^\mu A_\mu - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \ . \tag{2.5}$$

The covariant derivative D_μ will be motivated later and is given as

$$D_\mu = \partial_\mu - ieA_\mu \ . \tag{2.6}$$

Here ψ is a Dirac spinor as seen for example in [5] pp. 45-47, A_μ is the vector potential of the electromagnetic field and γ^μ are the gamma matrices defined in[5]. The first term is the kinematic part of a free Dirac Lagrangian which leads to the equation of motions [7] by varying ψ

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \overline{\psi} \right)} \right) - \frac{\partial \mathcal{L}}{\partial \overline{\psi}} = i\gamma^\mu \partial_\mu \psi - e\gamma^\mu A_\mu = 0 \ . \tag{2.7}$$

To compare this we can look at the free Dirac equations as seen in [5] which are

$$i\gamma^\mu \partial_\mu \psi - m\psi = 0 \ . \tag{2.8}$$

We can notice two things. First, there is no mass term in eq. (2.7) and second, there is an additional term that describes an interaction between the electromagnetic field and the Dirac spinor. The interaction term in eq. (2.7) follows directly from the interaction term in the Lagrangian in eq. (2.5). Naturally a free Dirac equation is missing interaction terms. The missing mass term in eq. (2.7) though can be attributed to a missing mass term in the Lagrangian. Such a mass term for the Dirac field would look like this

$$\mathcal{L}_\mathrm{m} = m\overline{\psi}\psi = m\left(\overline{\psi}_L\psi_R + \overline{\psi}_R\psi_L\right) \tag{2.9}$$

as can be seen in [5] as well. Here the subscript L and R refer to the projected out left and right handed components of the Dirac field and $\overline{\psi} = \psi^\dagger\gamma_0$. Unfortunately a term like in eq. (2.9) cannot simply be included in the SM Lagrangian. To see this we will take a closer look at the invariance of the Lagrangian under the gauge transformations of the symmetry group eq. (2.2).

The unitary group $U(n)$ is the group of $n \times n$ unitary matrices with matrix multiplication as group operation. In the case of $U(1)$ we have $n = 1$ and the elements in the form of $e^{i\alpha}$. The action of the $U(1)$ gauge group on a set of bispinors $\psi(x)$ is given by

$$U(1):\ \psi(x) \longrightarrow \psi'(x) = e^{i\alpha(x)}\psi(x)\ . \tag{2.10}$$

With $\alpha \in \mathbb{C}$. This motivates the definition of the covariant derivative since the $\mathcal{L}_\text{kin, Dirac}$ term in eq. (2.5) would not be invariant under such a local gauge transformation. To preserve the symmetry we construct the electromagnetic field A_μ such that its transformation under the group cancels the extra term in eq. (2.5). In infinitesimal form the transformation is chosen as

$$A_\mu \longrightarrow A'_\mu = A_\mu - \frac{1}{e}\partial_\mu\alpha(x)\ . \tag{2.11}$$

Including this eq. (2.5) is invariant under a local $U(1)$ transformation. Another more formal way to define the field strength tensor from a quantum field point of view is according to [6] p. 490

$$F_{\mu\nu} = \frac{i}{e}\left[D_\mu, D_\nu\right]\ . \tag{2.12}$$

This allows the covariant derivative to look at the difference of two coordinates which a simple partial derivative cannot sensibly do. As such, A_μ follows from the Lie algebra of the gauge group and is therefore sometimes called a gauge field.

Additionally, we can clearly see that a mass term such as eq. (2.9) does not violate the $U(1)$ symmetry when only regarding electrodynamics. The reason for abandoning the direct approach of including fermion masses for Dirac fields via eq. (2.9) therefore, lie in the extension of the theory to include other interactions.

The $SU(2)$ group: Formulating a Lagrangian Density

Before contemplating the missing mass and the non-abelian parts of the gauge group, we summarize some experimental observations related to the weak interaction. Weak interactions other than strong, electromagnetic and even gravitation in a certain way do not form bound states. They can mainly be observed in particle reactions. One of the most famous examples would be the decay of a neutron into a proton and an electron

$$n \longrightarrow p + e^- + \overline{\nu}_e \ . \tag{2.13}$$

First efforts to formulate a theory for these kinds of reactions have been made by Fermi as seen in [9]. The final SM theoretical formulation combines the electromagnetic and the weak interaction into the electroweak interaction based on the works from Glashow, Salam and Weinberg as seen for example in the works [10], [11] and [12]. They predicted the existence of three massive gauge bosons in addition to one massless gauge boson, the photon. The massive W^+, W^- and Z bosons were first observed in the year 1983 in the UA1 and UA2 experiments at the SuperProton Synchroton at CERN as seen in [13] and [14].

Now to better understand this unification and its effects on the SM, we first take a closer look at $SU(2)$. Analogous to $U(1)$ the group $SU(n)$ consists of those $n \times n$ unitary matrices that have a determinant of one with matrix multiplication as the group operation. The action of $SU(2)$ on a set of spinors $\psi(x)$ is given similar to eq. (2.10) as

$$SU(2): \ \psi_i \longrightarrow \psi_i' = \left(e^{i\alpha(x)^a T_a}\right)_{ij} \psi_j \ . \tag{2.14}$$

With this we can describe the weak interaction. Similar to the $U(1)$ case, it is possible to define a field strength tensor that forms the basis for the Lagrangian density of the weak field

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - ig \left[A^a_\mu, A^a_\nu\right] \ . \tag{2.15}$$

Where A^a_μ is in analogy to the electromagnetic case the vector potential describing the weak field. The index a runs from $a = 1, 2, 3$ and corresponds to the three generators of $SU(3)$. So as we can see, we have three exchange boson fields which are related in some linear combinations to the W^+, W^- and Z bosons. For comparison, in the electromagnetic interaction this index is always one due to $U(1)$ having only one generator which corresponds to one exchange boson, in this case the photon. In the $SU(3)$ case for the strong interaction this group index goes up to eight for eight gluons as exchange particles. We are also reminded again that from the theoretical point of view the vector field A^a_μ is a quantity arising from group theory due to keeping the invariance of the Lagrangian under the group transformations. This is equivalent to the $U(1)$ case. As mentioned above the SM formulation of the weak and electromagnetic interactions are unified in the electroweak interactions. The unifying gauge group would be

$$G_{\text{ew}} = U(1)_{\text{Y}} \times SU(2)_{\text{L}} \ . \tag{2.16}$$

Accordingly the gauge transformation on a set of Dirac fields from the gauge group eq. (2.17) is

$$G_{\text{ew}}: \ \psi_i \longrightarrow \psi_i' = e^{i\beta(x)} \left(e^{i\alpha(x)^a T_a}\right)_{ij} \psi_j \ . \tag{2.17}$$

Where the α from eq. (2.10) has been renamed to β but everything else is analogous to eq. (2.10) and eq. (2.14). It is necessary to talk about the indices Y and L in eq. (2.16) and eq. (2.2). The index Y corresponds to the later fully defined hypercharge in the SM. The group $U(1)_{\text{Y}}$ is isomorphic to $U(1)_{\text{Q}}$ and the structure is the same. The difference exists in the coupling and the way it is chosen. In eq. (2.17) the coupling is given by the hypercharge. The index L refers to the observation that the weak interaction only couples to the left-handed parts of a Dirac field. As such, any right-handedness will be projected out in the Lagrangian formulation. Thus the handedness of a particle and its Dirac field refers to its property under chiral transformation, also called chirality. Chirality is heavily connected to the Lorentz group and the bilinears transformations under it. As argued in [5] pp.43-44 the Lorentz group is reducible and can therefore be decomposed into irreducible representations. Since any Dirac field is in an irreducible representation under the Lorentz group the Dirac bilinear ψ reduces to ψ_L and ψ_R. Those are two irreducible representations that transform each differently under Lorentz transformations. As such, left- and right-handed parts of a particle field can even be interpreted as different particles altogether which further motives the coupling of the weak interaction to only left-handed particles. In terms of group actions chirality is induced via the γ^5 matrices as defined for example in [5] p.50. The projections for left- and right-handed particles are given by

$$\psi_{R/L} = \frac{1 \pm \gamma^5}{2}\psi \ . \tag{2.18}$$

If we write down the electroweak field Lagrangian without Dirac parts with similar notation and motivated by [6] we find

$$\mathcal{L}_{\text{ew}} = -\frac{1}{4}\left(W^a_{\mu\nu}\right)^2 - \frac{1}{4}\left(B_{\mu\nu}\right)^2 + \left(D_\mu H\right)^\dagger \left(D_\mu H\right) + m^2 H^\dagger H - \lambda \left(HH^\dagger\right)^2 \ . \tag{2.19}$$

Where $W^a_{\mu\nu}$ and $B_{\mu\nu}$ are the 4 gauge fields with $a = 1, 2, 3$ being the group index of $SU(3)$. The field H contained in the additional terms is a spin-0 scalar field. It will later be related to the Higgs field. The covariant derivative D_μ will be defined later. This mechanism is necessary to give the gauge bosons and the fermions a mass since a mass term in the form of eq. (2.9) violates the gauge symmetry induced by the transformation eq. (2.17).

The Higgs Mechanism and Spontaneous Symmetry Breaking for Gauge Bosons

The discovery of the Z, W^+ and W^- bosons also came with the experimental observation of them being massive particles. Additionally, a quantum theory of the SM should also include masses for the fermionic particle content in nature. Now, as stated before, the mass term eq. (2.9) is not invariant under the transformation eq. (2.17) as can be quickly seen. Additionally, a mass term for the bosons would also violate the gauge symmetry eq. (2.17). Such a mass term would look like this

$$\mathcal{L}_{\text{m, gauge boson}} = \frac{1}{2}m^2 B_\mu B^\mu \ . \tag{2.20}$$

The mass term for the $SU(3)$ fields would look analogous. We can quickly see after some calculations that such a term is not invariant under the gauge transformation.

So the question remains of how to include gauge boson and also fermion masses into the SM. The idea is to include an additional spin-0 scalar field into the Lagrangian whose potential term is such that this spin-0 scalar field has a non vanishing minimum. Therefore, this scalar field would acquire a non vanishing vacuum expectation value and this minimum state of the potential would spontaneously break the gauge symmetry. This is exactly the case for the last two terms in eq. (2.19). It has been shown that if we have a continous symmetry that breaks the ground state, we have new excitations in the particle spectrum consisting of massles spin-0 scalar bosons for each generator of the gauge group that does not leave the groundstate invariant. Such symmetry breaking is called spontaneous symmetry breaking. The massless particles are called Nambu-Goldstone bosons and the theorem is called Goldstones Theorem [15] and [16]. The Higgs mechanism uses this spontaneous symmetry breaking and the resulting Goldstone bosons to give mass to the gauge bosons.

We will see that in the case of eq. (2.17) the symmetry group decomposes from

$$U(1)_{\mathrm{Y}} \times SU(2)_{\mathrm{L}} \rightarrow U(1)_{\mathrm{Q}} \,. \tag{2.21}$$

So only the generators of $SU(2)$ will not leave the ground state invariant. In order to discuss this mechanism in more detail we take a closer look at the potential term of the scalar field in eq. (2.19). The Higgs Lagrangian is given by

$$\mathcal{L}_{\mathrm{ew}} \supset \mathcal{L}_{\mathrm{Higgs}} = \left(D_\mu H\right)^\dagger \left(D_\mu H\right) - V(H) = \left(D_\mu H\right)^\dagger \left(D_\mu H\right) + m^2 H^\dagger H - \lambda \left(H H^\dagger\right)^2 \,. \tag{2.22}$$

The first term is the kinetic term for the H field. The second term is the potential. The field H is a scalar field under Lorentz transformations that transforms like a spinor under $SU(2)$ to form $SU(2)$ invariant terms in the Lagrangian. This field has a hypercharge of $Y(H) = \frac{1}{2}$. The proper definition of the hypercharge will be given at a later point. We can note that it defines the transformation behavior under the $U(1)_{\mathrm{Y}}$ part of eq. (2.17). As can be seen for example in [5] p.701 this additional group is necessary to include a massless gauge boson into the SM. This boson represents the photon. The field H is given by

$$H = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \,. \tag{2.23}$$

Its transformation under the gauge group in eq. (2.17) will be given as

$$G_{\mathrm{ew}} : \; H \rightarrow e^{i\alpha^a \tau_a} e^{i\beta/2} H \,. \tag{2.24}$$

Here the τ_a are given by the Pauli matrices[1] Since we have the 2-dim. representation of $SU(2)$[2]

$$\tau_a = \sigma_a/2 \,. \tag{2.25}$$

It is now time to take a closer look at the potential of 2.22 and the kinetic Higgs-term in eq. (2.19). For this we define the covariant derivative similar to [6] p. 585 to be

$$D_\mu H = \partial_\mu H - igW_\mu^a \tau_a H - \frac{1}{2} g' B_\mu H \,. \tag{2.26}$$

Here g and g' are two different coupling constants who will later be related to the electric charge. From the potential term $V(H)$ in eq. (2.22) we can see there are an infinite amount of minima for values of $m^2 > 0$ and $\lambda > 0$. All the minima fulfill the relation

$$H^\dagger H = \frac{v^2}{2} \,. \tag{2.27}$$

---

[1] They are also related to angular momentum as the algebra of $SU(2)$ is isomorphic to that of the rotation group $SO(3)$.

[2] The transformation of a Dirac spinor would look slightly different since a Dirac spinor transforms as vector under the Lorentzgroup $SO(1,3)$.

Where v is the vacuum expectation value[3] of the field. We can without loss of generality choose v to be real and to be in the lower component of H. Thus we can expand around the vev

$$H = \exp\left(2i\frac{\pi^a \tau_a}{v}\right)\begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} + \frac{h}{\sqrt{2}} \end{pmatrix} . \tag{2.28}$$

With $v = m/\sqrt{2}$ and $\tau_a = \sigma_a/2$. Analogous to [6] we continue in unitary gauge[4] where π is zero. The field h can be viewed as a real scalar field. After the expansion around the vev the gauge boson masses can be found in the kinetic term. Similar to [6] pp.584-585 we will restrict ourself to the mass terms and leave the terms containing h out. We get

$$\begin{aligned} |D_\mu H|^2 &= \frac{g^2 v^2}{8}\begin{pmatrix} 0 & 1 \end{pmatrix}\begin{pmatrix} \frac{g}{g'}B_\mu + W^3_\mu & W^1_\mu - W^2_\mu \\ W^1_\mu + W^2_\mu & \frac{g}{g'}B_\mu - W^3_\mu \end{pmatrix}\begin{pmatrix} \frac{g}{g'}B_\mu + W^3_\mu & W^1_\mu - W^2_\mu \\ W^1_\mu + W^2_\mu & \frac{g}{g'}B_\mu - W^3_\mu \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= \frac{g^2 v^2}{8}\left[\left(W^1_\mu\right)^2 + \left(W^2_\mu\right)^2 + \left(\frac{g}{g'}B_\mu - W^3_\mu\right)^2\right] . \end{aligned} \tag{2.29}$$

As we can directly see, the fields W^1_μ and W^2_μ are diagonal in the mass eigenstates. The fields B_μ and W^3_μ need to be diagonalized though. After diagonalization we find the fields

$$Z_\mu \equiv \cos\left(\theta_W\right) W^3_\mu - \sin\left(\theta_W\right) B_\mu \quad \text{and} \quad A_\mu \equiv \sin\left(\theta_W\right) W^3_\mu + \cos\left(\theta_W\right) B_\mu . \tag{2.30}$$

Here we defined the Weinberg angle θ_W to be

$$\frac{g'}{g} = \tan\left(\theta_W\right) . \tag{2.31}$$

With this transformation we get a mass term for the Z_μ field but none for the A_μ field. The A_μ field is therefore the photon field and the Z_μ field that of the Z boson. By looking at the interaction of the gauge bosons with each other we can relate the electric charge to the coupling constants g and g'. We find

$$e = g\sin\left(\theta_W\right) = g'\cos\left(\theta_W\right) . \tag{2.32}$$

In a similar manner, we can also relate the electric charge to the gauge generators of $SU(2)$ and $U(1)$. We find accordingly with [5]

$$Q = T^3 + Y . \tag{2.33}$$

To find the proper fields for the W^- and W^+ we need linear combinations of the fields W^1_μ and W^2_μ with a positive and negative charge. We find them to be

$$W^+_\mu = \frac{1}{\sqrt{2}}\left(W^1_\mu - iW^2_\mu\right)^2 \quad \text{and} \quad W^-_\mu = \frac{1}{\sqrt{2}}\left(W^1_\mu + iW^2_\mu\right)^2 . \tag{2.34}$$

---

[3] Shortly called vev.

[4] Other gauges like the 't Hooft-Feynman gauge exist but for simplicity in the results the unitary gauge is chosen. The unitary gauge simplifies tree level calculations but hides scalar degrees of freedom meaning the Goldstone bosons from plain sight.

With this we have written down the gauge bosons of the electroweak interactions.

It may be useful to summarize this result. We find that the electroweak description of the four gauge bosons depend on three distinctive quantities. For once the couplings are described solely by the electron charge e and the new parameter θ_W that is called the Weinberg angle. Furthermore, the gauge boson masses are not independent as seen for example in eq. ([5]) p. 703. In fact, we have

$$m_{\mathrm{Z}} = m_{\mathrm{Z}} \cos\left(\theta_W\right) \ . \tag{2.35}$$

So the parameters that describe the electroweak sector of the SM are e, θ_W and m_{W}. Current measurements of these values can for example be found here [17].

It is prudent to remember that for the gauge boson masses we only looked at the terms without the h field in the expansion in eq. (2.28). Those terms containing this field lead to interaction terms between the gauge boson fields and the Higgs field. The Feynman rules and therefore diagrams originating from these interaction terms will be discussed below.

The Fermion Sector in $U(1)_{\mathrm{y}} \times SU(2)_{\mathrm{L}}$ and the SM Particle Content

We will now take a closer look at the fermion sector in the electroweak interaction. Specifically how fermionic mass terms can be realized.

We start by discussing the particle content and its integration into the gauge theory. In the SM fermionic particle states usually transform in irreducible representations of the gauge group. Experimentally we find two types of fermions. Leptons that interact weakly but not strongly and quarks that form, for example mesons and baryons, and interact weakly and strongly, as well as, electromagnetically. The leptons are separated into three families. Each family consists of a $SU(2)$ doublet for left-handed particles and corresponded singlets for right-handed particles. For example we have a flavor neutrino ν_ℓ and a lepton ℓ, where ℓ can be either a e^-, μ or τ. The neutrinos in terms of the SM only interact weakly and therefore have no electric charge. The quarks are split into three families as well. Similar to before, each family consists of a $SU(2)$ doublet as well and singlets for the right-handed particles under $SU(2)$. Hence, all the left-handed fermions transform in the fundamental representation of $SU(2)$. The gauge bosons W^+, W^- and Z transform in the adjoint representation. The right-handed fermions are all singlets under $SU(2)$. Additionally, in terms of $SU(3)$, as will be discussed further below, only the quarks are charged. There are six quarks which are called up, down, charm, strange, top and bottom [3]. We denote the quarks and leptons[5] the following way

$$L^i = \begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix}, \begin{pmatrix} \nu_{\mu L} \\ \mu_L \end{pmatrix}, \begin{pmatrix} \nu_{\tau L} \\ \tau_L \end{pmatrix}, \qquad Q^i = \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \begin{pmatrix} c_L \\ s_L \end{pmatrix}, \begin{pmatrix} t_L \\ b_L \end{pmatrix} \ . \tag{2.36}$$

Here $i = 1, 2, 3$ runs over the generations. The right-handed fermions we denote as

$$e^i_R = \left\{e_R, \mu_R, \tau_R\right\}, \quad u^i_R = \left\{u_R, c_R, t_R\right\}, \quad d^i_R = \left\{d_R, s_R, b_R\right\} \ . \tag{2.37}$$

These are all singlets under $SU(2)$[6]. Since no right-handed neutrinos have been observed in experiments so far they are excluded in the above list. In order to determine the hypercharge of the

---

[5] Both the quark and lepton left-handed fields transform like Weyl spinors under $SU(2)$ in the $\left(\frac{1}{2}, 0\right)$ representation.

[6] And transform as right-handed Weyl spinors under the Lorentz group.

fermions under the $U(1)_{\mathrm{Y}}$ part of the gauge group we remember that

$$Q = T^3 + Y \; . \tag{2.38}$$

This relates the electric chrge Q with the weak hypercharge Y and the third generator of $SU(2)$. The relation stems from looking at eigenstates of the charge operator when determining the gauge bosons as seen for example in [5] pp. 700-703. Sometimes Y can be found as $Y/2$ depending on the definition of β in eq. (2.24). Using this we can determine the hypercharge knowing the electric charge. For electrons for example we find

$$Y(e^-) = Q(e^-) - T^3(e^-) = -1 - \left(-\frac{1}{2}\right) = -\frac{1}{2} \; . \tag{2.39}$$

For quarks the electric charge is $1/3$ of an electron charge. Denoting the fundamental representation of $SU(2)$ as f and uncharged fields as $-$ we can find the table (2.1).

Field	L	e_R	ν_R	Q	u_R	d_R	H
$U(1)_{\mathrm{Y}}$	$-\frac{1}{2}$	-1	0	$\frac{1}{6}$	$\frac{2}{3}$	$-\frac{1}{3}$	$\frac{1}{2}$
$SU(2)_{\mathrm{L}}$	f	$-$	$-$	f	$-$	$-$	f
$SU(3)_{\mathrm{C}}$	$-$	$-$	$-$	f	f	f	$-$

Table 2.1: Field charges in the SM from [6] p. 593 with the fields defined in eq. (2.36) and eq. (2.37). Here f indicates a transformation in the fundamental representation and $-$ indicates that a field is uncharged

.

The antiparticles transform accordingly and have opposite charges.

We will now take a closer look into the creation of fermion masses via the Higgs field. For this we need to introduce coupling terms between the fermion fields and the Higgs. These terms, which are often called Yukawa coupling terms, need to be constructed in a manner such that they do not violate the gauge symmetry. The first look we take is at the quarks. The Lagrangian can be written as[7]

$$\mathcal{L}_{\text{mass, quarks}} = -Y^d_{ij}\bar{Q}^i H d^j_R - Y^u_{ij}\bar{Q}^i \tilde{H} u^j_R + h.c. \; . \tag{2.40}$$

Where $\bar{H}$ denotes the adjoint and $\tilde{H} \equiv i\sigma_2 H^*$ is a field that transforms in the fundamental representation of $SU(2)$ and has a hypercharge of $-\frac{1}{2}$ to make the term invariant. Each term is invariant under eq. (2.2) even when including $SU(3)$. The matrices Y^d_{ij} and Y^u_{ij} contain the Yukawa couplings. They are generally complex matrices that do not have to be hermitean. As argued in [6] complex masses will not appear since we can always redefine the phase of our fields. After spontaneous symmetry breaking the mass term can be written in matrix form as

$$\mathcal{L}_{\text{mass, quarks}} = -\frac{v}{\sqrt{2}}\left[\bar{u}_L Y_d d_R + \bar{u}_L Y_u u_R\right] + h.c. \; . \tag{2.41}$$

Even though this matrix is in general not hermitean it can be diagonalized as can be seen for example in [6] p. 596. When diagonalizing it we rotate the fields $d_R \rightarrow K_d d_R$, $d_u \rightarrow K_u d_u$ and $u_L \rightarrow U_u u_L$,

---

[7] Where $h.c.$ stands for the hermitean conjugate of the term.

$d_L \rightarrow U_d d_L$. Where $U_{u/d}$ and $K_{u/d}$ are unitary matrices used to diagonalize $Y_{u/d}$. The mass terms appear after diagonalization

$$\mathcal{L}_{\text{mass, quarks in flavor basis}} = -m^d_j \bar{d}^j_L d^j_R - m^u_j \bar{u}^j_L u^j_R + h.c. \, . \tag{2.42}$$

The interactions between the fermions and the gauge bosons are given by the kinetic terms of the fermions like the first term in eq. (2.5), where we only discussed the $U(1)$ group. When discussing the electroweak gauge group and including the covariant derivative eq. (2.26) and the fermion fields eq. (2.36) and eq. (2.37) we can write the kinetic fermion Lagrangian in the flavor basis as

$$\begin{aligned}\mathcal{L}_{\text{flavor basis}} =& L^i \left[i\not\partial + \gamma_\mu \begin{pmatrix} \frac{g'}{6}B_\mu + \frac{g}{2}W^3_\mu & \frac{g}{\sqrt{2}}W^+_\mu \\ \frac{g}{\sqrt{2}}W^-_\mu & \frac{g'}{6}B_\mu - \frac{g}{2}W^3_\mu \end{pmatrix} \right] L_i \\ &+ \bar{u}^i_R \left(i\not\partial + g'\frac{2}{3}\not{B} \right) u^i_R + \bar{d}^i_R \left(i\not\partial - g'\frac{1}{3}\not{B} \right) d^i_R + \mathcal{L}_{\text{mass, quarks}}\end{aligned} \tag{2.43}$$

according to [6]. When rotating to the mass eigenbasis the only terms where the rotation matrices do not drop out are the ones containing W^-_μ and W^+_μ. There we have an additional matrix that mixes the flavor of the interaction called the Cabibbo-Kobayashi-Maskawa[8] matrix. It is given by

$$V \equiv U^\dagger_u U_d \tag{2.44}$$

as can be seen for example in [18]. The degrees of freedom of the CKM matrix and its properties are of vital understanding for the details of the SM as seen in [6]. As mentioned above, the kinetic terms lead to interactions between the fermions and the gauge bosons. Feynman rules and Feynman diagrams will be discussed later.

To the quark masses we can write down mass terms for the leptons as well. Those being

$$\mathcal{L}_{\text{mass, leptons}} = -Y^e_{ij} \bar{L}^i H e^j_R + h.c. \, . \tag{2.45}$$

If we would include a right-handed neutrino field ν^i_R we could form additional Yukawa terms. Those would lead to neutrino masses as well. In the SM neutrino masses are formally excluded even though they have been indirectly experimentally observed through neutrino oscillations [19]. Since we exclude this field, we end up without a mass matrix like the CKM matrix. This pendant would be called the Pontecorvo-Maki-Nakagawa-Sakata[9] matrix. Otherwise after SSB, the mass terms appear similar to the quark case. From the kinetic term of the leptons we can acquire the interaction terms between leptons and gauge bosons as well.

The $SU(3)$ Group and the Strong Interaction

Now that we concluded the discussion of the electroweak sector and the inclusion of gauge boson and fermion masses, we focus on the last remaining part of the gauge group in eq. (2.2) $SU(3)_\text{C}$. The theory of strong interactions in the SM is Quantum Chromodynamics. The first ideas about a short-ranged strong theory have been brought up to explain the stability of the nucleus. Since the

---

[8] In short CKM matrix.

[9] Shortly called PMNS.

protons have been found to be positively charged and some nuclei consist of many protons a force was necessary that binds the nucleons together. This force had to be, at least for the range of the size of the nucleus, stronger than electromagnetism. While doing deep inelastic scattering experiments it was found that protons consist of subsequent composite particles. These particles were described by Feynman as partons [20] and have later been found to be made up mostly of quarks and gluons. The underlying gauge symmetry of QCD was found to be $SU(3)_C$ which is related to there being three color charges. The quarks transform in the fundamental representation under $SU(3)_C$ and behave under the whole gauge group as seen in eq. (2.1). The gluons are the exchange bosons of the strong interaction. Their part in the QCD Lagrangian is similar to that of the other gauge bosons except that they have no mass. They transform in the adjoint representation of $SU(3)_C$. In order to explain how quarks could exist in similar quantum states and form for example a baryon without violating the Pauli exclusion principle, the concept of the color charge was introduced as can be referenced in [3]. Based on this, there exists three colors red, green, and blue and only colorless states can form hadrons. Where red, green and blue together, in analogy to optical color theory, form colorless states. Additionally, the antiquarks carry an anticolor that in combination with a color forms colorless states as well. The property of free states to be colorless is called confinement.

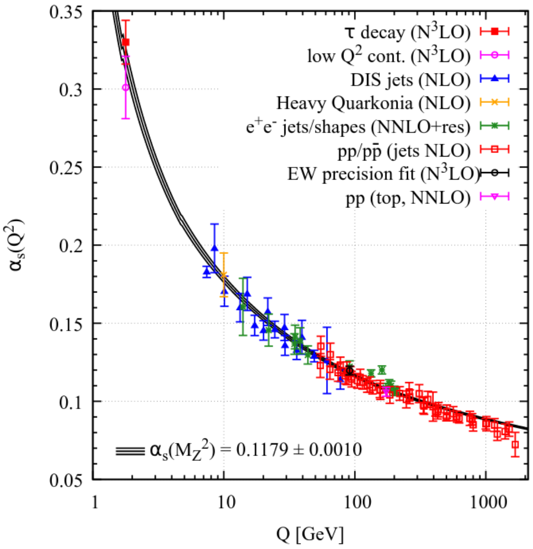


Figure 2.1: Measurements of the strong coupling from [21] where NLO stands fo next-to-leading order and NNLO for next-to-next-to-leading order and so forth.

Another important property of strong interactions is the peculiarity of its coupling compared to electromagnetic interactions. The coupling strength of the strong interactions, as well as, that of the other interactions as well depend on the energy, This is an effect of renormalization and due to the introduction of a running coupling in a renormalized quantum field theory. More on this can be found in [5] part 2. The coupling strength of the strong interaction gets higher for smaller energies as can be seen in figure (2.1). This means that a perturbative treatment of QCD is only reasonable at high energies when the particles are in a state of asymptotic freedom and processes can be calculated using Feynman diagrams. For lower energies other methods like effective field theories and lattice QCD are needed to make theoretical predictions about strong interaction processes. As mentioned above, the running coupling is due to renormalization and higher order processes in perturbation theory. To understand the discrepancy between the running coupling of the electromagnetic and the strong interaction it is necessary to understand that in QCD the gluons themselves are charged with color. This stands in stark contrast to the electromagnetic case where the photon possesses a neutral electric charge. Keeping this in mind we can understand the difference in behavior when imagining the creation of additional particle-antiparticle pairs and an increase in the number of force carriers. In an intuitive image we can say that they create a screen of charges that prevent the flux of the field to propagate undisturbed. Therefore, those particles created in higher orders of perturbation theory can be gluons. Gluons carry a charge which means that for larger distances, meaning lower energy, the effective coupling becomes stronger. This explains the behavior shown in figure (2.1). In the electromagnetic case the opposite happens since virtual positrons are created. They get closer to the electron and screen its charge. Hence, the coupling is lower for longer distances and higher for shorter distances. The behavior of the running coupling is described by the β-function of the renormalization group [5] ch. 12. For QCD it takes the following form

$$\alpha_s(Q^2) = \frac{1}{\beta_0 \ln\left(Q^2/\Lambda\right)} \,. \tag{2.46}$$

Where Λ is the QCD scale at which perturbation theory breaks down.

We now write down the QCD Lagrangian without an explicit mass term. We have

$$\mathcal{L}_{\text{QCD}} = \sum_q \bar{\psi}_{q,a} \left(i\gamma^\mu \partial_\mu \delta_{ab} - g_s \gamma^\mu t^C_{ab} A^C_\mu \right) \psi_{q,b} - \frac{1}{4} G^A_{\mu\nu} G^{\mu\nu A} \,. \tag{2.47}$$

Where this time we denoted all indices that are necessary. We sum over all quarks q and index the colors with a and b. Which run from $a = 1$ to $N_{\text{C}} = 3$. The A^A_μ correspond to the gluon fields and C runs from $C = 1$ to $N_{\text{C}}^2 - 1 = 8$ since the gluons transform in the adjoint representation of $SU(3)$. The matrices $t^C_{ab} = \lambda_{ab}/2$ are given by the eight Gell-Mann matrices that represent the action of $SU(3)$ as seen for example in [22]. The field strength tensor is accordingly given as

$$G^A_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g_s f_{ABC} A^B_\mu A^C_\nu \,. \tag{2.48}$$

Here f_{ABC} are the structure constants of $SU(3)$.

Looking at eq. (2.47) QCD seems to be a relatively simple theory but its discussion in detail becomes quite complex. One reason is the aforementioned behavior of the running coupling and the breakdown of perturbation theory. The properties of strong processes especially those related to jets and those processes faking τ's will be reviewed in section 2.4.

The Complete Lagrangian of the Standard Model

We have the basic components to write down the Lagrangian of the SM as a summary of the observations discussed above. The SM presented in this thesis is not a complete review. Main topics like renormalization have been completely left out. Additionally, due to the behavior of quantized gauge groups additional terms exist in the SM Lagrangian that have not been mentioned. For example, there are the Fadeev-Poppov ghost terms that are necessary for a non-abelian quantum theory. They are related to gauge fixing in the theory. The ghosts themselves are unphysical particles and will not be observed in S-matrix calculations. Including all this is beyond the horizon of this thesis. For further reference one can review [5] and [6].

The complete Lagrangian after SSB can be in short form written down as

$$
\begin{aligned}
\mathcal{L}_{\text{SM}} =& -\frac{1}{2}tr\left(G_{\mu\nu}G^{\mu\nu}\right) - \frac{1}{8}tr\left(W_{\mu\nu}W^{\mu\nu}\right) - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \\
&+ \bar{L}^i\left(i\not{D}_\mu\right)L^i + \bar{e}^i_R\left(i\not{D}_\mu\right)e^i_R + h.c. \\
&- \frac{\sqrt{2}}{v}\left[\bar{L}^i H M^e e^i_R + \bar{e}^i_R \bar{H}\bar{M}^e L^i\right] \\
&+ \bar{Q}^i\left(i\not{D}_\mu\right)Q^i + \bar{u}^i_R\left(i\not{D}_\mu\right)u^i_R + \bar{d}^i_R\left(i\not{D}_\mu\right)d^i_R + h.c. \\
&- \frac{\sqrt{2}}{v}\left[\bar{Q}^i H M^d d^i_R + \bar{Q}^i\tilde{H}M^u u^i_R\right] + h.c \\
&+ \left(D_\mu H\right)^\dagger\left(D_\mu H\right) + m^2\left[H^\dagger H - v^2/2\right]^2/2v^2 \\
&+ \mathcal{L}_{\text{additional terms}}\,.
\end{aligned}
\tag{2.49}
$$

Where we renamed the matrices V^i to M^i to highlight them being mass matrices. Additionally, the covariant derivative changes slightly for the type of fermion. For quarks there is an additional gluon field term and for right handed particles the weak field term drops out. The Higgs field H takes the form of eq. (2.28). The last term includes everything not mentioned in a greater detail as for example the ghost terms and so forth. The other terms have in some form been discussed above. From this form we could read the Feynman Rules for the processes relevant for this thesis.

2.1.4 Feynman Rules and Feynman Diagrams

In this part we want to discuss the role of Feynman diagrams and Feynman Rules as a concept and a way to understand the SM. In physics we often use perturbation theory to find an approximate answer to a given problem. In particle physics this translates to processes whose matrix element can be calculated to different orders of perturbation theory. There are leading order processes and those of higher orders. The matrix elements we are referring to is related to the S-matrix[10]. It is the operator that translates a state from the distant past meaning $t = -\infty$ to the distant future meaning $t = \infty$ and describes the change between these two initial and final states. The S-matrix can be calculated to different orders of perturbation theory where we require higher orders to contribute lesser than lower orders in order for the perturbative expansion to be well defined. This lies to an extent at the heart of the problem of renormalization since this is not always the case. The S-matrix can be related to so called correlation functions as can be seen in [5]. These can in quantum field theory be interpreted as the amplitude for a propagation of a particle from a point x to y in space-time. The theorem that relates the correlation functions to S-matrix elements is called the Lehmann-Symanzik-Zimmermann theorem[11] as can be seen in [23]. Feynman diagrams are a graph theoretical method to calculate correlation functions. Because correlation functions can be related to S-matrix elements, Feynman diagrams therefore also describe physical processes to different orders in perturbation theory. Feynman diagrams can be constructed following a given set of rules and consist of external legs related to in- and outgoing particles, propagators describing the propagation of virtual particles and vertices that

---

[10] S-matrix is short for Scattering matrix

[11] Shortly called LSZ theorem.

are intersection points between those elements. Additionally, it is useful to define the direction in which time flows since different conventions exist. This thesis uses the convention of time flowing from left to right. Lastly, spin conservation is shown through arrows on the lines of a diagram. In the convention used in this thesis those arrows describe the flow of particle nature. Meaning the momentum of particles goes parallel to the arrows while the momentum of antiparticles goes in the opposite direction. Additionally, arrows never only flow in and out of a vertex with time and never collide in it. These conventions are due to spin conservation and would change when applying a 2-component formulation of a given quantum theory where each field is described by a spinor. We will not do this and instead rely on a four component Dirac field. Hence, the above conventions have been described.

The Feynman rules of the SM can be seen in the appendix of [5]. They can be derived from the Lagrangian of the theory. As mentioned above interaction terms play a vital part in this. A Lagrangian in quantum field theory can be viewed as consisting of kinetic and mass terms and additional interaction terms. The kinetic terms relate to the propagation of the particle and defines the propagator. The mass term can be seen as an interaction of the particle with itself and yields the mass. The interaction terms yield the vertices of the Feynman rules. As can be seen in [5] ch. 9 it is possible to derive the Feynman rules through functional derivatives. In short, we can say that when there is an interaction term involving a set of fields, then there is a vertex involving those fields. Furthermore, there are differences between the different kind of fields resulting in different Feynman rules for for Dirac- and Klein Gordon fields. This includes their propagators and also the diagram as a whole.

As an example of a Feynman diagram we can look at figure (2.2). This diagram describes a LO[12] $t\bar{t}$ production process. Two quarks annihilate and produce a virtual gluon which decays into a t and a $\bar{t}$ respectively. The features described above such as vertices, propagators and external legs can be seen exemplified here. Additionally, this diagram as all the following diagrams will be read from left to right.

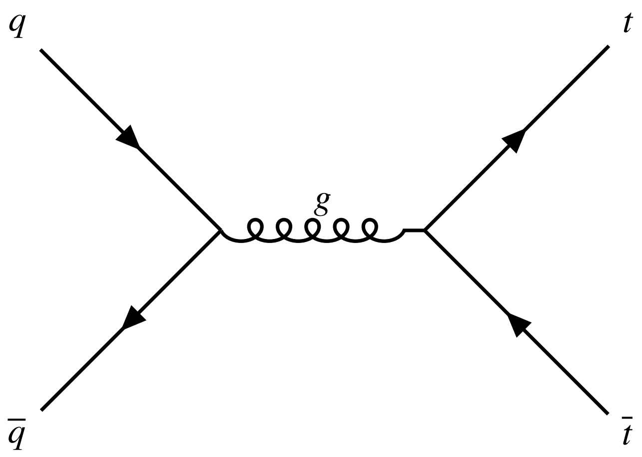


Figure 2.2: Example of a Feynman diagram: $t\bar{t}$ production

---

[12] Meaning leading order.

2.2 Top Quark Physics and the tHq Process

As, in this thesis we are interested in estimating the abundances of falsely reconstructed τ leptons in top quark-related channels, this section will focus on top quark physics. We will discuss the relevant properties of the top quark and then take a closer look at its production and related processes with special consideration given to the tHq production and some important background processes of the tHq analysis.

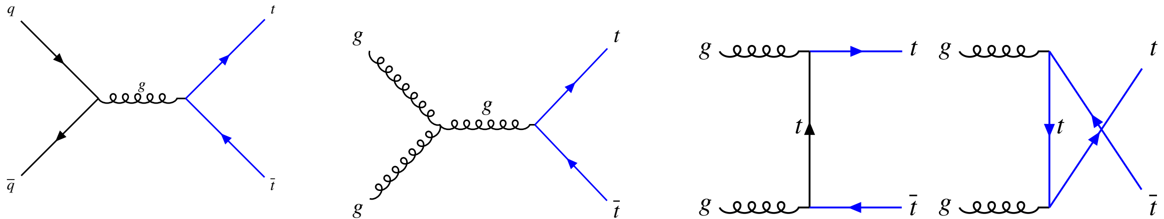
2.2.1 Top Quark Physics

The top quark is the heaviest of the quarks included in the SM. Its behavior under the gauge group transformation can be seen in table (2.1). Together with the b quark, their left-handed part forms a $SU(2)$ doublet. Both have been postulated in the 70's to explain CP violation in Kaon systems. It was first discovered in 1995 at the Tevatron [24]. Its mass and mean lifetime was found to be 173.34 ± 0.76 GeV and 5×10^{-25} s respectively [17]. Due to its large mass being heavier than that of a W boson, it decays semi-weakly into a real W and a b quark. This results in a very short lifetime and prevents the top quark from hadronization. As can be seen in the Yukawa term of the electroweak sector eq. (2.40) the coupling of the quarks to the Higgs is mass dependent. Due to its large mass the only Higgs coupling to a fermion that is of order unity is that of the top quark. Hence the processes related to the top are very sensitive to the physics of the Higgs boson as well.

Top Quark Production Processes

The production of top quarks separates into either the production of a single top quark via the weak interaction or top quark pairs.

The Production of top quark pairs is either through gluon fusion $\left(gg \to t\bar{t}\right)$ or quark antiquark annihilation $\left(q\bar{q} \to t\bar{t}\right)$. At the LHC, gluon fusion contributes to the majority of top quark pair events. The LO diagrams for top pair production can be seen in fig. (2.3).



(a) quark antiquark annihilation in s-channel

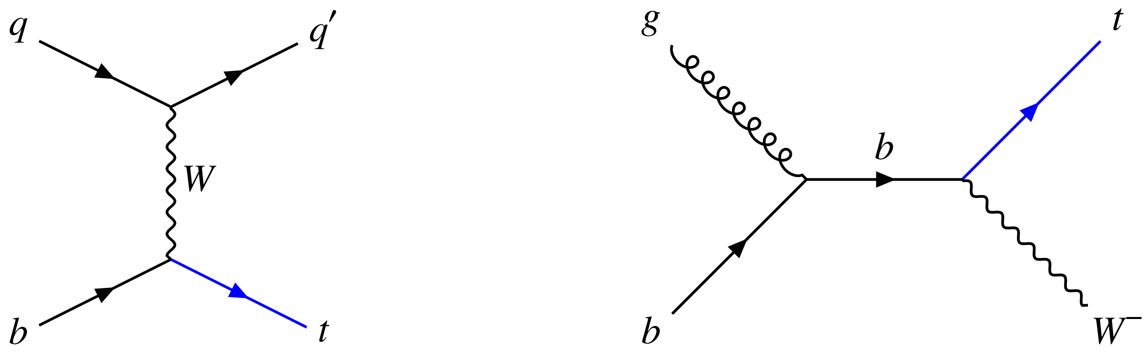
(b) gluon fusion in s-channel

(c) t-channel

(d) u-channel

Figure 2.3: LO top pair production processes.

The production of singletop quarks is mediated partially by the weak interactions. The biggest contributing LO order Feynman diagrams can be seen in fig. (2.4). The t-channel diagram in fig. (2.4) is responsible for about 70% of singletop production processes. Here, a W boson interacts with a b quark to a t quark. The W boson then mediates the change of a spectator q to a different flavored quark q'. The next biggest production process is an interaction between a gluon and a bottom quark. The virtual b quark then decays into a top quark and a W boson. The other relevant LO Feynman diagrams for example for the s-channel have a considerably smaller cross section than those single-top production processes presented in [17].



(a) t-channel

(b) tW-channel

Figure 2.4: Biggest contributing LO singletop production processes.

Because of its large mass, a decay into the W boson is possible. Hence, it dominantly decays into a b quark and a real W boson. The vector boson can then decay leptonically or hadronically. As such, a t quark decay is associated with b quark production which are relatively easily tagged. This simplifies the search for top-quark related processes.

2.2.2 The tHq Process and the $2\ell 1\tau_{\text{had}}$ channel

As this thesis investigates the abundance of τ fakes with a special focus on the $2\ell 1\tau_{\text{had}}$ channel, it is prudent to take a closer look at the relevant processes involved in this channel as well as the signal process whose overall observation is of particular interest. This signal process is the tHq production. As a single-top process it is deeply related to the physics of top quarks and especially the coupling of the Higgs to the electroweak sector. The most important background processes whose τ fakes contribute the most are $t\bar{t}$, Z+jets and to a lesser extent W+jets and Diboson events. As will be covered when discussing the physics of fake τ's.

Of particular interest in Higgs analysis is the coupling Y_{t}^{SM} between the Higgs and the top quark. Different BSM theories predict deviating behavior of this coupling $Y_{\text{t}} = \kappa_{\text{t}} Y_{\text{t}}^{\text{SM}}$ parametrized by κ_{t} from the SM. Current measurements favor the SM prediction [25]. Because the dominant Higgs production processes do not depend on the sign of κ_{t} a flipped sign coupling is possible. A sensitive channel to the sign of this coupling is the tHq channel. There we have destructively interfering singletop quark production diagrams that lead to a small cross section of $\sigma = 18.3$ fb [26]. As can be seen in fig. (2.5). A flip of the sign in the top Higgs coupling could lead to a significant higher cross section. This allows an efficient way to probe for BSM physics using current data.

The tHq process decays into several final state channels related to the Higgs decay mode. In this thesis we are mostly interested in ML[13] final states. these includes specifically the $2\ell 1\tau_{\text{had}}$[14] and $1\ell 2\tau_{\text{had}}$[15]

---

[13] Meaning multi lepton.

[14] Also called lep-had.

[15] Also called had-had.

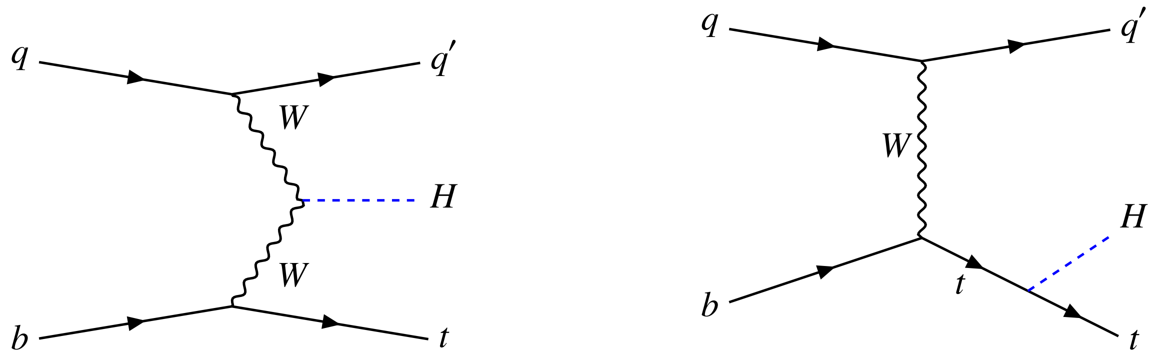


Figure 2.5: Two diagrams representing single top production processes involving a Higgs boson. The destructive interference between those two processes leads to a small tHq cross section in the SM

channels. The Higgs decay modes that contribute to this final state are $H \to \tau\tau$, $H \to ZZ$ and $H \to WW$.

The decay that we are looking for is the $H \to \tau\tau$ decay where the τ leptons either decays hadronically or leptonically. An example for a tHq process in the lephad channel can be seen in 2.6[16]. Here one τ decays hadronically via a decay into a W boson that decays into mesons and the other τ decays leptonically into a W boson that decays into a light lepton and their respective neutrino. The mesons turn out be nearly universally pions in this case. The other light lepton then comes from the decay of the top quark. Additionally, the had-had channel acquires its additional hadronically decaying τ through the $H \to \tau\tau$ decay and the light lepton comes again from the aforementioned top quark.

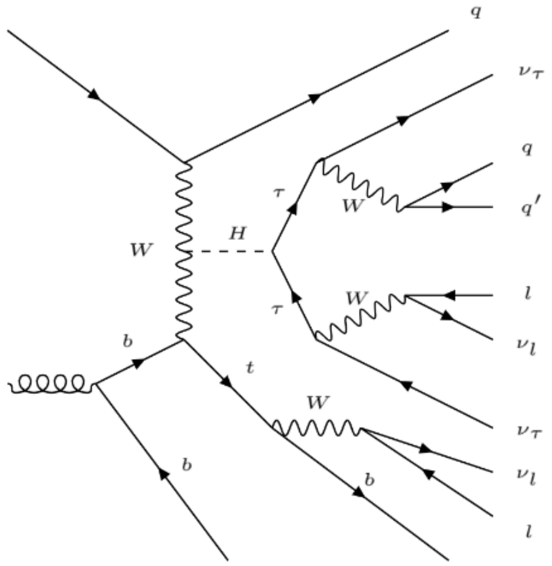


Figure 2.6: An example for a tHq decay into the lephad channel. Taken from [27]

---

[16] There are additional Feynman diagrams for the lephad channel of the tHq process. Especially those processes including Higgs production like the second diagram in fig. (2.5)

2.3 The Physics of τ Leptons and τ Fakes

This section discusses the properties and physics related to τ leptons as well as their associated fakes. The basic properties and importance of τ leptons will be reviewed before talking about the definition of fakes. Additionally, the creation of τ fakes will be discussed in terms of contributing signatures as well as the processes related to τ fakes in the lephad and hadhad channel of the ML final states in the tHq process.

2.3.1 Tau Leptons

The τ lepton is the third lepton of three flavor families of leptons. And its left-handed field together with its neutrino form a doublet under $SU(2)$. The τ lepton is charged under the electromagnetic interaction and is a singlet under $SU(3)$. This means it does not interact via the strong interaction. Its mass and mean lifetime are measured to be $m_\tau = 1776.86 \pm 0.12$ MeV and $t_\tau = (2.903 \pm 0.005)\times 10^{-13}$s [17]. Compared to the electron and muon the tau has a significantly larger mass and shorter mean lifetime. As such, it usually cannot be detected in the detector directly and instead is reconstructed via its decay products. The τ lepton decays either leptonically into light leptons and their respective neutrinos or hadronically via a W boson decay into pairs of quarks. Those quarks later form jets inside the detector. Those jets are distinguished by the number of charged particle tracks that initiate the jet called prongs. As a τ lepton has an electric charge it can only decay into an uneven number of charged mesons. Those mesons are nearly exclusively pions [17]. It is possible though that additional neutral pions are emitted as well. Due to its large mass, compared to the other leptons, the cross section of a Higgs decaying into two taus, is of order unity rather than negligible as in the case of muons and electrons [17]. As such, the τ lepton is an important decay particle for processes sensible to the Higgs sector as discussed in subsection 2.2.2. To illustrate this even more we can look a the branching ratio of the Higgs decay modes in fig. (2.8). In this SM calculation we can see that the $H \to \tau\tau$ has a relatively high branching ratio[17] for an expected Higgs mass of $m_{\mathrm{H}} \approx 125$ GeV. It represents at these masses a branching ratio of about 10%. Processes like $H \to ZZ$ and $H \to WW$ that also can result in the ML final states have alower cross section predicted by the SM.

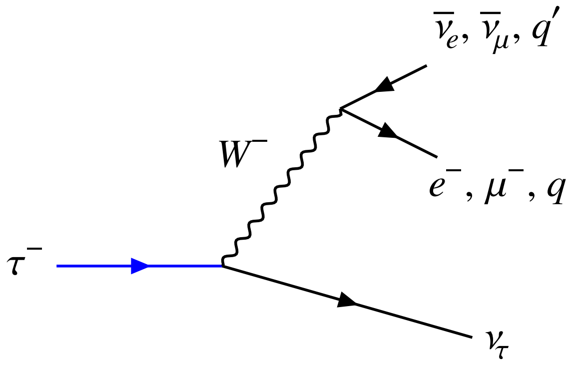


Figure 2.7: The decay of a τ lepton into light leptons or hadronically.

---

[17] At least compared to the processes Higgs to light lepton. This can be seen in fig (2.8). The decay of $H \to bb$ dominates in this Higgs mass region.

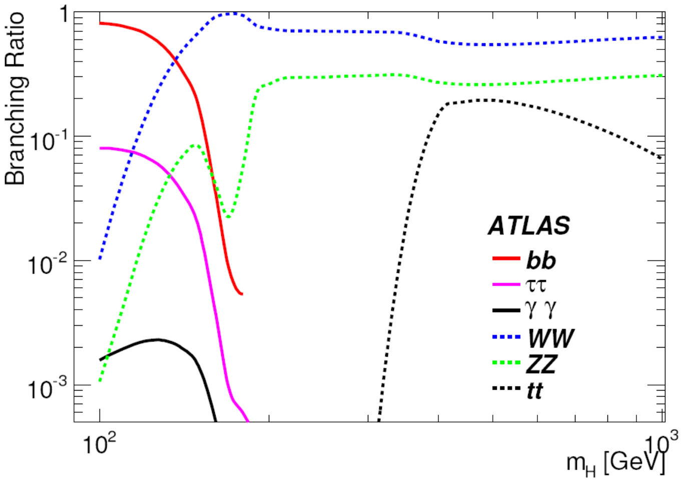


Figure 2.8: Branching ratio of Higgs decay for different Higgs masses m_{H} from [28].

2.3.2 Tau Fakes

In order to deal with tau fakes their proper definition is given in this subsection. Tau fakes are the particles and jets that were wrongly reconstructed in the detector as tau leptons. A more detailed description of τ reconstruction will be given in chapter [3]. The problem at the heart of this thesis can then be formulated as the following.

Assuming there is a signal region[18]. And in this SR there is a region in data that is suspected to be contaminated with tau fakes. The goal is to estimate how many fakes are expected there. One solution is to extract scale factors[19] from so-called control regions (CR) enriched with fakes. The CR has to be topologically similar to our SR where we want to do the physics analysis. This SF then gives us a correction for Monte Carlo simulations in the SR and as such yields a better modeling of fakes. Those fakes come from different sources.

The biggest contributing sources to falsely reconstructed τ leptons[20] are jets, electrons and muons.

From these sources jets are estimated to be the biggest contributing factor. In this thesis we will differentiate between quark initiated and gluon initiated jets. As such, a more detailed description of jet physics and physics concerning quark and gluon jets will be discussed in the next chapter. The reason jets fake τ leptons is because the detcor profile of a jet can be similar to the profile of a hadronically decaying τ lepton.

Light leptons as well are sometimes wrongly reconstructed as τ leptons. This happens mostly in the case of electrons but also muon tau fakes are possible. In this case the light leptons are not properly recognized as light leptons because they fail criteria for reconstruction. In addition to that they are surrounded by jets which in combination with the light leptons create a detector profile that is similar to that of a hadronically decaying taus. This can lead to the misidentification of the light leptons as taus.

---

[18] In short called SR.

[19] Shortly called SF.

[20] Shortly called τ fakes.

Important Background Processes

There are important background processes for the tHq signal whose potential fake τ's contribute the most. We will start by taking a closer look at the $t\bar{t}$ process. Its cross section has been measured to be $\sigma_{t\bar{t}} = 136 \pm 20$ (stat.) $\pm$ 40 (syst) $\pm$ 8 (lumi.) fb [29]. Some LO order $t\bar{t}$ production processes have already been discussed in fig. (2.3). Here one or both top quarks can decay eventually into a τ lepton. Additionally, b quarks are produced as well. Hence, in the $0b$ region we do not expect any or very few contributions from the $t\bar{t}$ process. Since the b tagging efficiency is not perfect $t\bar{t}$ still retains significant importance in the $1b$ region and naturally dominates the $2b$ region. The hadron that would be falsely reconstructed as a τ could have its source from a quark or gluon emissioned jet.

Another important background process that we want to discuss is the $Z+$jets process. This involves a decay of a Z boson and additional jets. The Z boson is fairly well measured as can be seen by its precise mass measurement [17]. Additionally, we are interested in those decays $Z \longrightarrow \ell\ell$ that enter our multi lepton[21] final state channel. Specifically the $1\ell 2\tau_{\text{had}}$ and $2\ell 1\tau_{\text{had}}$ channels. Due to the Z decay we expect big contributions of this process in the $0b$ region and fewer in the regions for higher numbers of b quarks. Additionally, those τ leptons faked in this process are expected to be mostly due to quark initiated jets for kinematic reasons. The behavior of the $W+$jets channel is similar to that of $Z+$jets but differentiates itself in the flavor changing nature of the W boson. This is why we expect a higher importance of $W+$jets events in regions of a larger number of b quarks compared to the $Z+$jets events.

The next important background consists of diboson events. Since those contain W and Z bosons the properties of both influence this process. An example for a simple LO diboson diagram can be seen in fig. (2.9). Here the two bosons can either be a W or a Z boson. To be a relevant background in the ML channel different combinations of decays and possibilities of fakes can be assumed. One example for the $2\ell 1\tau_{\text{had}}$ channel would be a decay of the Z into to light leptons and a decay of the W into a τ and a ν_τ. Again, this τ if decaying hadronically can be faked by jets. For example, by jets of a $Z+$jets decay. There are additional processes whose τ leptons can be faked in the two relevant ML channels $2\ell 1\tau_{\text{had}}$ and $1\ell 2\tau_{\text{had}}$. But their contributions are comparatively small to those discussed above. The principal components of these faking processes are very similar though. Additionally, fake contributions from muons and electrons are also relevant, if less so. Especially those instances of τ fakes from light leptons which relate to the processes $W+$jets and $Z+$jets. Here the process of faking a τ depends on the combinaton of a light lepton being surrounded by jets.

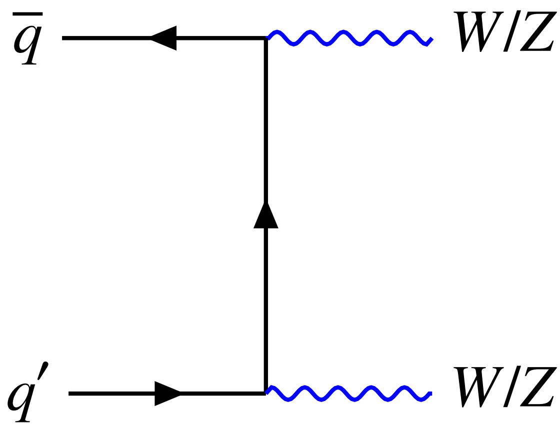


Figure 2.9: An example diboson diagram.

It is to be expected that real world detectors are not perfectly symmetrical in space and in their tracking ability due to small and unavoidable inaccuracies in the hardware. Because of this we need

---

[21] Abbreviated as ML. This a set of decay channels for the tHq process that end in leptons. The channels of interests for this thesis that end in multiple lepton final states are the had-had and lep-had channel.

to consider a dependence of the SF on the region of the detector which we parametrize with $|\eta|$[22]. Additionally, we expect a SF dependence on the transverse momentum p_{T} as well. One reason for this is that the energy deposition and shower shapes change with p_{T}. In fact, for higher transverse momentum it is easier to differentiate between QCD jets and hadronically decaying τ leptons. Because τ leptons decay in uneven numbers of prongs, we also need to take this into account. The reconstruction of three prong τ decays is assumed to be better due to more tracks leading to more information. Additionally, 1 prong decays are more likely to be faked by light leptons than 3 prong decays are. In general the kinematics of 1 prong and 3 prong decays differ. Hence, we expect a dependence of the SF on the number of prongs as well.

When estimating the SF's these dependencies need to be taken into account as well as the difference in physics and kinematics of quark initiated and gluon initiated jets. This will be the focus point of discussion in the next section.

2.4 QCD Jets

As mentioned above, the biggest contributions to the tau fake abundance is due to QCD jets being falsely reconstructed as τ leptons. Therefore, it is prudent to take a closer look at the physics of jets and especially at the difference between quark and gluon initiated jets. As major references [30] and [31] have been used.

2.4.1 Jet Physics

A jet is a narrow cone of hadrons that are produced from quarks and gluons after hadronization [30]. Particles possessing color charge cannot exist in isolation but need to form colorless singlet states.[23] This is called QCD confinement. As such, when as a result of a hard QCD process a quark fragments from a colorful object, the quark creates other colorful objects around him to form a colorless state. These then called colorless states form the jets observed in particles detectors. The states contained in those jets are usually collinear and as such the jets form narrowed tracks in the detector. The production of jets goes through a series of steps that we will discuss in more detail in the following parts of this section. A jet production process often goes from high energy QCD that often is in the asymptotically free perturbative domain of the theory to bound states whose energy and formation cannot be explained properly via perturbation theory due to the nature of the QCD coupling fig. (2.1). Instead, one relies on experimentally determined parton distribution functions[24] and evolution equations to describe this process. Recent developements in computational physics like effective field theories in lattice QCD predict those to some accuracy [30] even though a reliance on experimental measurement is still necessary [32].

Jet Production and Deep Inelastic Scattering

As already mentioned, the coupling of the strong interaction decreases for higher transferred momentum and goes to infinity at the QCD scale as seen in fig. (2.1). This creates a domain in which perturbative

---

[22] Here η is the pseudorapidity and will be defined later. It describes the angle of a particle relative to the beam axis.

[23] As mentioned in 2.1 when discussing QCD.

[24] Which we shortly abbreviate with PDF. Which is not be mistaken for probability density functions as they are also often abbreviated with PDF.

formulation of hard QCD scattering processes using Feynman diagrams is possible at high energies. In particle colliders, especially at the LHC jet production starts off in this domain. After the fragmentation of the color object into different color states like for example quarks, showering occurs and the energy at some point shifts to the domain where the strong coupling diverges and perturbation theory breaks down. At these scales, hadronization, the process to form colorless singlet states consisting of hadrons, occurs [31]. This situation creates a challenge to calculate QCD cross sections for such reactions. Fortunately, the factorization theorem of QCD [33] allows one to deal with this situation. Employing this we can separate a cross section calculation into a short distance perturbative QCD part and a non-perturbative part parametrized through universal process independent functions. These functions, called parton distribution functions, can be measured experimentally. They quantify the probability to find a parton with a longitudinal momentum fraction within a hadron. The functions are usually characterized by the scale of factorization [5]. As such, we can write the cross section of a high transverse momentum pp process similar to [30]

$$\mathrm{d}\sigma_{pp\to X} = \sum_{i,j} \int \mathrm{dxdx}' f_{i/p}(x', \mu_f) \cdot f_{j/p}(x, \mu_f) \times \mathrm{d}\hat{\sigma}_{ij\to X}(x, x', \mu_f, \mu_r, \alpha_s(\mu_r)) \,. \tag{2.50}$$

Here the $f_{i/p}$, $f_{j/p}$ are the proton PDF's as functions of the fractional longitudinal momenta and the factorization scale. The sum goes over all the the initial partons so $i, j \in \{q, \overline{q}, g\}$. The quantity $\mathrm{d}\hat{\sigma}_{ij\to X}$ describes the parton level cross section and depends on the fractions of momenta. The factorization scale defines the resolution below which the information of the process is absorbed into the PDF's that describe the non-perturbative physics [30]. This factorization scale is usually introduced in the $\overline{MS}$ factorization scheme and defined through dimensional regularization [5]. It is therefore related to infrared divergences similar to how the renormalization scale μ_r is related to ultraviolet divergences [5]. This collinear factorization ansatz is at the heart of making quantitative QCD predictions. It partly relies on information extraction from data to determine the PDF's. The evolution of those PDF's for different factorization scales is given by QCD [5] by the Dokshitzer–Gribov–Lipatov–Altarelli–Paris equations[25] [34].

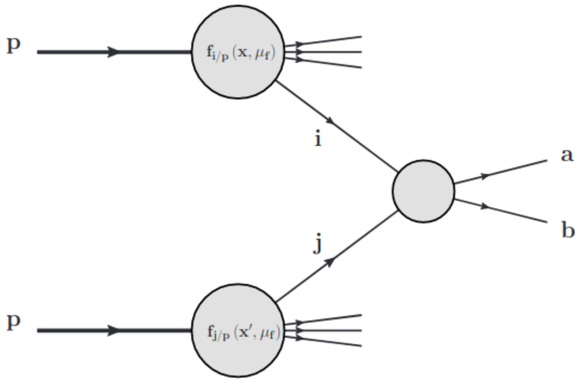


Figure 2.10: Schematic depiction of a high p_T scattering process of two partons from two protons. Taken from [30].

If in a reaction like 2.10 a parton with color charge has been produced its evolution into a color neutral state needs to be described as well. A possible way to describe this evolution into a hadron state, as given in [30], is to define so called fragmentation functions[26] which describe the probability of finding a hadron within a fragmentation product of a given parton. They are therefore similar to the PDF's with the difference being that they

---

[25] Which can be abbreviated as DGLAP equations.

[26] Shortly called FF.

contain information about the final states. As seen in [30] including the FF's into eq. (2.50) we can write

$$\mathrm{d}\sigma_{pp\to X} = \sum_{i,j,k} \int \mathrm{d}x\mathrm{d}x'\mathrm{d}z f_{i/p}(x', \mu_f) \cdot f_{j/p}(x, \mu_f) \times \mathrm{d}\hat{\sigma}_{ij\to k}(x, x', \mu_f, \mu_r, z, \alpha_s(\mu_r)) \times D_{k\to X}(z, \mu_f) \,. \tag{2.51}$$

Here $D_{k\to X}(z, \mu_f)$ are the FF's for finding a hadron X in a parton fragment k. Similar to the PDF's they can only be determined experimentally.

Sometimes it is sufficient to know the energy flow and shapes of an event and then this can be compared to the observations inside the detector. If a detailed understanding of the full hadronic content that makes up the final jet is required then MC event generators are usually employed.

MC Event Generators

This part focuses on the demands that the physical evolution requires from simulations describing the formation of jets and QCD scattering events in high energy physics.

Based on physically motivated probability distributions MC event generators simulate the reaction between two beam particles from their parton level scattering to the final state. This final state consists of specific particle types including their respective four momenta and masses [30]. The simulation beyond this final state, for example the propagation through the detector material, is not part of the event generator simulation and instead is generated by the detector simulation. Taking a simple Drell-Yan process $qq \to \ell\ell$ as an example. A schematic explanation of such a process being generated by MC event generators can be seen in fig. (2.11). Here we have two protons interacting via a parton level process in fig. (2.10) at high p_T. In this case in the central red blob a pair of leptons is created. The leptons do not have a color charge and interact no further. The initial partons though radiate gluons which evolve into a cascades of partons. This process of radiation is described perturbatively by parton showers [35]. These parton showers work as a leading-logarithmic approximation[27] correction to the LO order parton level cross section result [30]. The partons inside the shower need to eventually form colorless states. In fig. (2.11) these are displayed as colorless ellipses which form hadrons. The hadrons are the yellow points in fig. (2.11). This process of fragmentation, which is sometimes called hadronization, has to be treated non-perturbatively. For this models, like the Lund string fragmentation model [36], are used in event generators. Depending on the specific

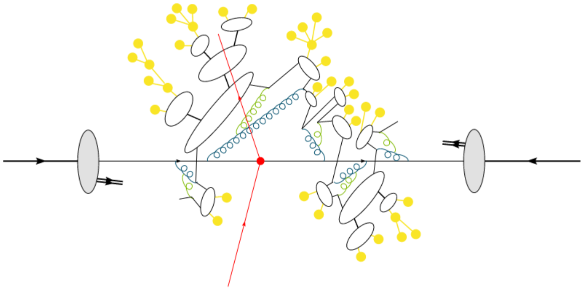


Figure 2.11: Schematic process of a MC event generator simulation of a Drell-Yan process. Where the red center denotes the lepton pair production. The partons radiate soft gluons which produce parton showers. These then form colorless states represented in this picture as empty ellipses. Those states then fragment to the yellow hadron circles in a non-perturbative process. Taken from [30].

---

[27] In short LLA.

kind of particle the hadron itself may decay in an additional step as well.

Jet Algorithms

Building on this understanding of pQCD[28] scattering, the non-perturbative aspects of hadronization and fragmentation and its realization in MC event generators, we will now look more closely at the heart of jet algorithms themselves.

The definition of a jet has been given in the introduction of this section. From an experimental particle physics point of view a jet consists of measured tracks and energy deposits in a particle detector [30]. This is illustrated in fig. (2.12). Here you can see a proton collision and the evolution of the partons into hadrons contrasted against the length scales up until which hadronizations occurs. Further away those collinear colorless state objects then form measured tracks inside the detector and deposit their energy along them. Out of these detector signals particle jets are reconstructed. In order to be able to conclude from a set of detector observables to particle jets jet algorithms are employed. These algorithms group energy deposits together depending on an angular cone centered around the axis of a specific direction. Jet algorithms can be sorted into two classes:

- cone algorithms that assign energy deposits inside a cone based on geometrical criteria to objects that form a jet;
- sequential-recombination algorithms that combine those objects closest together.

These algorithms need to fulfill a certain set of requirements [31]. One of the most important ones is collinear and infrared safety. In order to employ pQCD the outcome of a jet algorithm must neither depend on the splitting or merging of collinear parton four vectors nor on the addition of arbitrarily soft partons [30] through radiation. This can be seen in fig. (2.13) where on the left side in fig. (2.13(a)) the splitting of a jet leads to a transverse momentum that is under the threshold and therefore the jet is not recognized. This can occur especially in cone algorithms where a seed for a jet has to pass a certain p_t threshold. On the right side in fig. (2.13(b)) the radiation of an additional soft gluon merges two jets into one which prevents the cancellation of divergences in

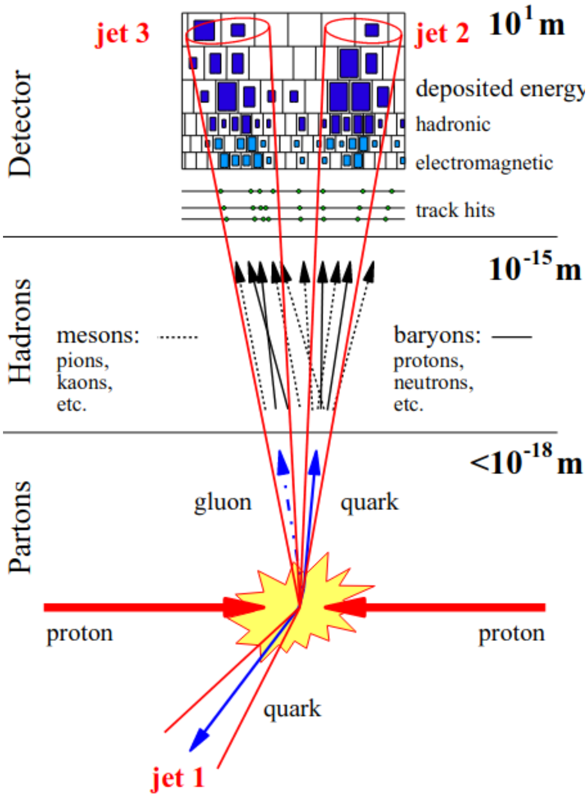
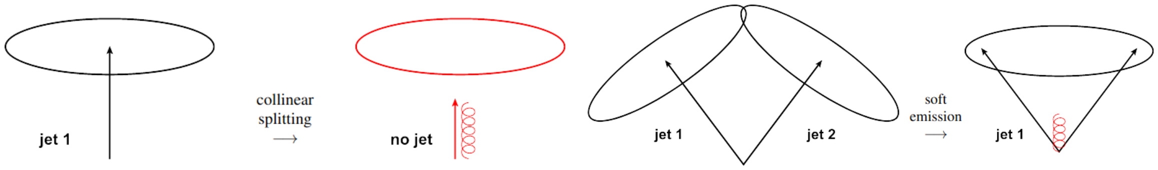


Figure 2.12: Schematic expression of a proton proton collision and the formation of jets and their observable tracks inside the detector from which the jet can be reconstructed. Taken from [30].

---

[28] Here the p stands for perturbative.

virtual corrections. There are additional conditions a well defined jet algorithm needs to fulfill like longitudinal boost invariance and boundary stability [30].



(a) A collinear splitting stops the new seeds from passing the p_t threshold such that the jet disappears.

(b) Soft radiation merges two jets into one which prevents the cancellation of divergences in virtual corrections.

Figure 2.13: Schematic illustration of infrared and collinear safety requirements for jet algorithms. Taken from [30].

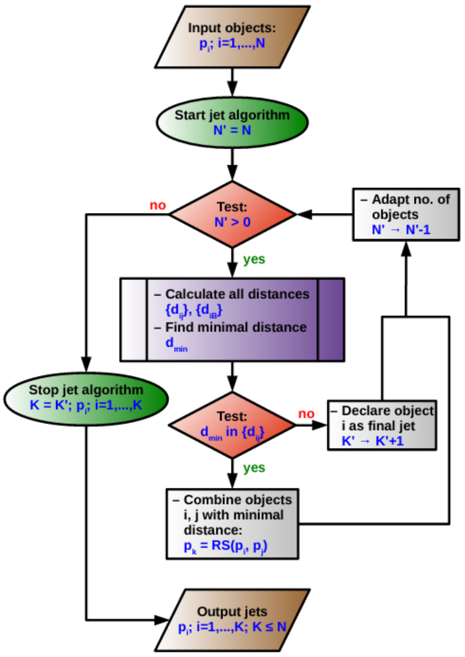


Figure 2.14: Flowchart describing a jet algorithm for sequential clustering of N objects. The output is a cluster of K jets. Taken from [30].

Another requirement for jet algorithms stems from an experimental point of view. In search for very rare processes, like for example tHq or other Higgs related processes, high luminosities or transverse momenta are required. In order achieve those, multiple proton-proton collisions per bunch crossing are piled up. Due to the finite integration time of detector components, energy from bunches get deposited into the events of other bunches. This is called pile-up. The energy from those pile-up collisions needs to be substracted. Jet algorithms are sensitive to pile-up.

The jet algorithm used at the LHC is the cone-like sequential-recombination algorithm anti-k_t. It defines two distances. The relative distance d_{ij} between each pair of input objects and the distance between each input object and the beam direction d_{iB}. They are given as

$$\begin{aligned} d_{ij} &= \min\left(p_{\mathrm{T},i}^{2p}, p_{\mathrm{T},j}^{2p}\right)\frac{\Delta R_{ij}^2}{R^2} \\ d_{iB} &= p_{\mathrm{T},i}^{2p} \\ \Delta R_{ij}^2 &= (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2 \,. \end{aligned} \tag{2.52}$$

Where ΔR_{ij} is the angular distance between objects i and j given by the azimuthal angle ϕ and the rapidity η[29]. The parameter R denotes the cone radius and p decides whether low- or high-p_T objects are clustered first. This is a deciding characteristic between the k_t for $p = 1$, Cambridge/Aachen for $p = 0$ and the anti-k_t algorithm for $p = -1$. As seen in the flow chart fig. (2.14) the algorithm works in the following way. From all distances d_{ij} and beam distances d_{Bi} the minimal distance is found. If $d_{\min} \in \{d_{iB}\}$ the input object is declared a jet [30], it is removed

---

[29] Both are defined properly in 3.1.2.

from the list of objects and added to the list of jets. If $d_{\min} \in \{d_{ij}\}$ the objects are merged and added as one to the object pool and the parent objects i and j are removed from the pool of clustering objects. This is repeated until only objects in the list of jets remain. The usual sizes R for the jets range from $0.4 - 0.6$ in the ATLAS experiment. Therefore, this is sometimes included in the naming scheme.

2.4.2 Quark and Gluon Jets

An important distinction for this thesis are the physical properties and the phenomology of quark initiated and gluon initiated jets. As such, this subsection investigates their properties more deeply and defines an important discriminant used in the likelihood fit to estimate the τ fakes originating from these different jets. This subsection mainly relies on [37], [38] and [39].

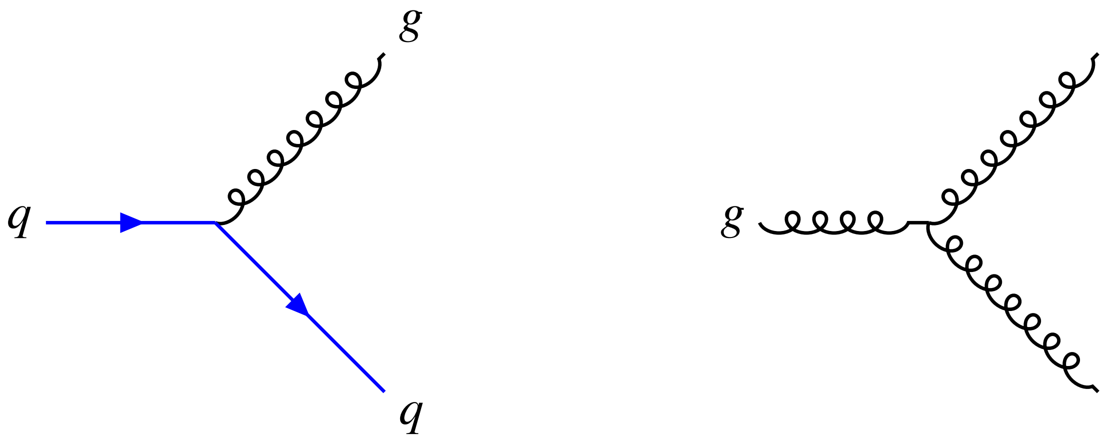
The Definitions of Quark and Gluon Jets

A proper definition of quark and gluon jets is not straightforward. We will discuss definitions of quark jets. Definitions of gluon jets follow accordingly. A naive ansatz would be to refer to a jet initiated by a quark as a quark jet. This assumes though that there exists a one-to-one map between a jet and its initiating parton. This is not correct since additional gluon radiations form an important role in manifesting the jets properties. Therefore such a definition can be regarded as too naive. A series of definitions is discussed in [37]. A common issue with most definitions of quark jets and by extension also gluon jets is that the quark is defined as an object unto itself [37]. But the properties of quarks in these kind of reactions are process dependent. Therefore, such definitions that treat quarks as process independent objects are not precise. It is more sensible to treat quarks, rather than as objects, as attributes. With this mindset [37] proposes to define quark jets as a phase space region that yields an enriched number of quarks. Here the phase space region is defined via cross section measurements into hadronic states and the criteria for quark definitions need be sufficiently suitable. These criteria themselves are ambiguous. To illustrate this further let us assume that α is a quark/gluon discriminant. This discriminant can be measured for any given jet. Let us further assume that small values of α denote gluon jets and high values of α indicate quark jets. As such a quark classification criterion is given. The usual procedure is to first measure the discriminants of interest and then depending on the focus of the analysis does one develope a classification scheme based on the discriminants.

The Properties of Quark and Gluon Jets

Having discussed a proper definition of quark and gluon jets we now focus on the properties of these and the observed phenomenological differences between their jets. Additional, we discuss an important discriminant used for the template likelihood fit.

The physics of quark and gluon jets is defined by their respective couplings to soft gluon radiation processes. This coupling relies on the color factor. Those radiation processes are the dominant contributors to parton showers. Since quarks and gluons both have a color charge in QCD there exists color conservation in vertices. This is similar to spin conservation in Feynman diagrams and often called color flow. Quarks transform inside the fundamental and gluons inside the adjoint representation of $SU(3)$. When calculating QCD S-matrix elements a summation over all group indices and therefore all color indices is needed. This yields a combinatoric factor based on the vertices in the diagram. This factor is called the color factor. When looking at soft gluon radiations from gluons and quarks they are calculated to be $C_A = 3$ for gluons and $C_F = 4/3$ for quarks. The relevant vertices can be seen in fig. (2.15(b)) for C_A and fig. (2.15(a)) for C_F. Because of this we expect an asymptotically 9/4 higher multiplicity in jets from a gluon source compared to jets from a quark source. Most observables in quark gluon tagging efforts are functions of the quotient C_A/C_F [37]. In the following we will discuss a list of different properties between quark and gluon jets.



(a) Radiation of a gluon from a quark. This yields a color factor of $C_F = 4/3$.

(b) Radiation of gluons from a gluon. This yields a color factor of $C_A = 3$.

Figure 2.15: Color factors for both gluon and quark radiations of a gluon.

Due to greater radiation of soft gluons in a gluon jet, compared to a quark jet the width of gluon jets is expected to be broader. Therefore, the energy closer to the jet axis is larger for quark jets than for gluon jets as seen in [39]. An experimental observation of this expectation can be seen in fig. (2.16(a)). Here you can see the differential energy profile of gluon and quark jets and two MC predictions. Those predictions agree well with data and the energy profile for the quark jet is closer to the jet axis.

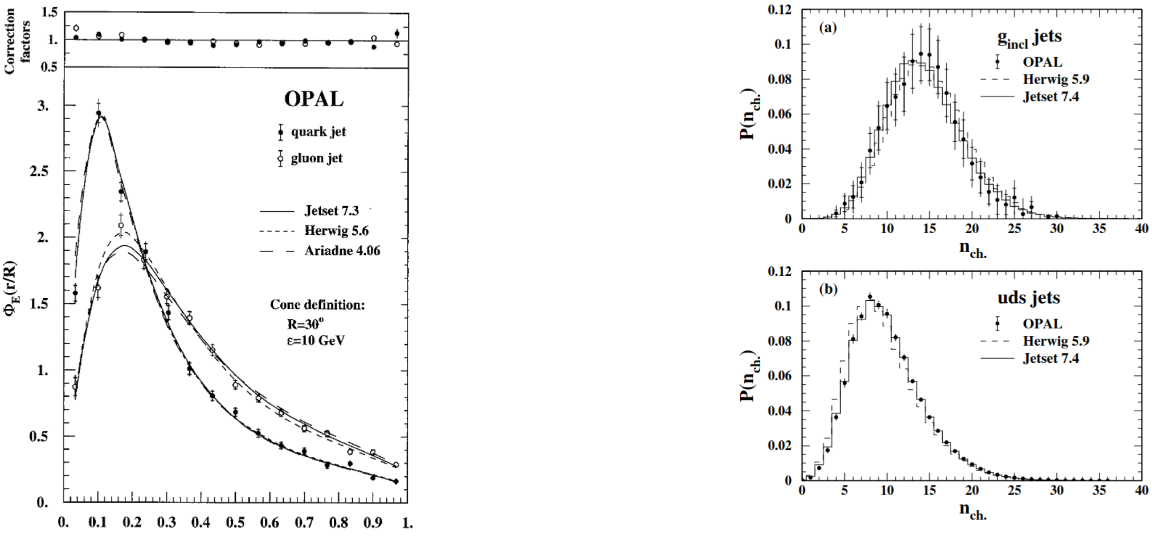
Another direct consequence of the larger soft gluon emission from gluons compared to quarks is the expected hadron multiplicity inside the jet. This is expected to be higher for gluon jets. As seen in fig. (2.16(b)) this can be confirmed for light quarks from experimental observations. Similar to the energy profile a good agreement between MC simulations and data can be seen. Only unbiased jets found in the event hemisphere were used. This is defined as the plane perpendicular to the event axis [40]. The ratio of the expected mean multiplicity of gluon and quark jets is of significant interest. As mentioned in the introduction in the asymptotic limit of $E_{\text{particle}} \ll E_{\text{jet}}$ this ratio should follow the ratio of C_A/C_F for soft particles.

The fragmentation functions mentioned in eq. (2.51) also are expected to differ between quarks and gluons. This is because the scale dependence of the fragmentation functions is dominated by the splitting of $P_{g\to gg} \sim C_A$ for gluons and $P_{q\to gq} \sim C_F$ for quarks [39]. As seen in [39] this could be observed.

As mentioned above, the discriminant used for this thesis is the track width of the jets. It is defined as

$$w_{\text{trk}} = \frac{\sum_{\text{trk}\in\text{ jet}} p_{\text{T, trk}} \Delta R_{\text{trk, jet}}}{\sum_{\text{trk}\in\text{ jet}} p_{\text{T, trk}}} . \tag{2.53}$$

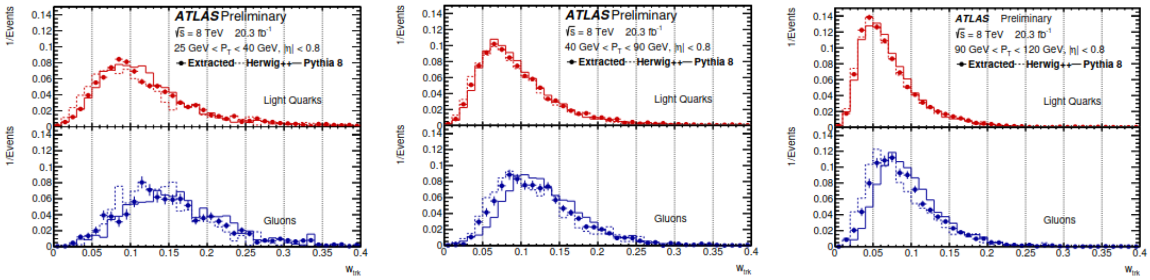
It weights for every track inside the jet the width of that track with its transverse momentum. In fig. (2.17) simulations and data of the track width for different p_{T} and for a fixed detector region $|\eta| < 0.8$ are shown. For the extraction dijet, $\gamma + \text{jet}$ and $Z + \text{jet}$ samples were used [38].



(a) Energy profile for quark and gluon jets. Data is compared with MC simulations. Taken from [39].

(b) Charged partilce multiplicity for 41.8 GeV gluon jets and 45.6 GeV light quark jets. the predicted number from MC simulations are also shown. Taken from [40].

Figure 2.16: Experimental observations of different quark/gluon discriminants and their comparison to MC simulations.



(a) Simulation and data for the track width variable for $25 < p_{\mathrm{T}} < 40$ in GeV. Extracted using dijet and $Z + \text{jet}$ samples.

(b) Simulation and data for the track width variable for $40 < p_{\mathrm{T}} < 90$ in GeV. Extracted using dijet, $\gamma + \text{jet}$ and $Z + \text{jet}$ samples.

(c) Simulation and data for the track width variable for $90 < p_{\mathrm{T}} < 120$ in GeV. Extracted using dijet and $\gamma + \text{jet}$ samples.

Figure 2.17: Simulation and data for the track width variable for different transverse momenta and a fixed $|\eta| < 0.8$. Taken from [38].

2.5 Likelihood Fitting

This section discusses the major statistical method employed in this thesis, the concept of maximum binned likelihood fits for different sources. As major references [41], [42] and [43] have been used.

2.5.1 Formulation of the Problem

Let us assume there is a set of data such that for each element in the dataset there exists a set of values $\{x_i\}$. For an example, this can be a set of hadronic τ decay events with leptons or jets for each which there is a value of $\{p_{\mathrm{T}}^{\ell}, E_{\mathrm{vis}}, ...\}$. These events arise from a number of different sources. In the scope of this thesis we are interested in sources like truth τ decays, QCD jet induced τ fakes or light lepton τ fakes. What we want to determine is the proportion P_j of each source in the data. We do not have an analytic function describing the proportions of these sources. What we use instead are Monte Carlo simulations for high energy particle processes. Therefore, we bin the data into n bins. We denote the number of events in the real data for bin i by d_i. The data is then a set of numbers $\{d_1, ..., d_n\}$. In accordance with [41], we denote the predicted number of events in a given bin i as $f(P_1, .., P_m)$. Here the P_j are the proportions of the sources also called their strength. Denoting the numbers of MC events from source j in bin i with a_{ij} we can write the predicted number of events as

$$f_i = N_D \sum_{j=1}^{m} P_j a_{ij} / N_j \; . \tag{2.54}$$

Here N_D is the total number of events in the data sample and N_j the number of events in the MC sample from source j. As such they are given as

$$N_D = \sum_{i}^{n} d_i \quad \text{and} \quad N_j = \sum_{j}^{m} a_{ji} \; . \tag{2.55}$$

Similar to [41] we rewrite f_i by introducing $p_j = N_D P_j / N_j$. With this f_i becomes

$$f_i = \sum_{j}^{m} p_j a_{ji} \; . \tag{2.56}$$

A possible way to estimate the newly defined p_j and therefore by extension the strength of source j is to minimize the χ^2 assuming a gaussian distribution for the d_i. The χ^2 is given by

$$\chi^2 = \sum_{i} \frac{(d_i - f_i)^2}{d_i} \; . \tag{2.57}$$

This is acceptable for large numbers as the gaussian is a good approximation for a poisson distributed set of data. The d_i are poisson distributed. Due to the high dimensionality of the data, the d_i are often expected to be small. As such, gaussian distribution cannot be assumed. Instead, we rely on the poisson distribution. The probability of observing a particularly d_i is given as

$$\mathcal{P}(d_i|f_i) = e^{-f_i}\frac{f_i^{d_i}}{d_i!} \ . \tag{2.58}$$

The poisson likelihood is then given as

$$\mathcal{L} = \prod_{i=1}^{n} \mathcal{P}(d_i|f_i) \ . \tag{2.59}$$

For practical reasons one often wants to minimze the logarithmic likelihood instead of maximizing the likelihood function. Taking the logarithm of equation (2.59) and neglecting the terms that vanish when calculating the minimum we can write the likelihood function to be

$$\ln(\mathcal{L}) = \sum_{i}^{n} d_i \ln(f_i) - f_i \ . \tag{2.60}$$

This binned likelihood needs to be minimized. This can be done using a dedicated algorithm for minimizing functions such as MINUIT [44]. TRExFitter, the program used to perform the likelihood fits in this thesis, uses MINUIT. TRExFitter is based on Histfactory [43].

2.5.2 Including Nuisances into the likelihood Fit

In order to include systematic uncertainties into the fit, nuisance parameters can be used. Here we restrict ourselves to multiplicative uncertainties. For example, if we fix the sources for muons that fake taus to 1, but still want to estimate the impact of that nuisance on the likelihood estimation we can include this nuisance into eq. (2.59) in the following way

$$\mathcal{L} = \prod_{i=1}^{n} \mathcal{P}(d_i|f_i)\mathcal{G}(F_\mu|\tilde{F_\mu}, \sigma_{F_\mu}) \ . \tag{2.61}$$

Here the function $\mathcal{G}$ is a gaussian[30] with a mean of $\tilde{F_\mu}$ and a width of σ_{F_μ}. Where σ_{F_μ} constraints the nuisance parameter F_μ[31] to its a priori value. Taking the logarithm yields

$$\ln(\mathcal{L}) = \sum_{i}^{n} d_i \ln(f_i) - f_i + \frac{\left(F_\mu - \tilde{F_\mu}\right)^2}{2\sigma_{F_\mu}} \ . \tag{2.62}$$

In principle any uncertainty can be incorporated into a likelihood estimation this way. It is important though to keep the nuisance small otherwise the gaussian is truncated and an appropiate normalization factor is chosen. Other probability densities might be worthy to consider in this situation. [42]

---

[30] Technically other functions than gaussians can be used.

[31] It is important to note that F_μ is used to calculate the f_i and not $\tilde{F_\mu}$.

2.5.3 Statistical Uncertainties from MC

A special kind of nuisance arises from the uncertainties in MC statistics themselves. Taking a step back and summarizing the reasons for possible disagreements between the d_i and f_i in eq. (2.54), we find that those can arise from incorrect p_j, fluctuations in the d_i and fluctuations in the a_{ij}. So far the likelihood fit method reckons with the first two causes but not the third. The proper way of finding the likelihood includes the uncertainty in a_{ji}.[32] The proposed solution by Barlow and Beeston [41] argues that for each source j there is an expected number A_{ji} in each bin that we do not know. This number can be estimated by a poisson distribution in the limit of $A_{ji} << N_j$ with the help of the a_{ji}. As such the total logarithmic likelihood becomes

$$\ln(\mathcal{L}) = \sum_i^n d_i \ln(f_i) - f_i + \sum_i^n \sum_j^m a_{ji} \ln(A_{ji}) - A_{ji} \ . \tag{2.63}$$

Minimizing this likelihood with an algorithm like MINUIT can be quite challenging due to the number of nuisances.[33] Instead it is possible to simplify this problem and find a faster solution.

Solutions to the Problem

The problem at hand is to find the p_j and the A_{ji} that minimize eq. 2.63. This is a problem with $m \times (n + 1)$ unknowns. We can differentiate eq. 2.63 for p_j and A_{ji} and get similar to [41]

$$\sum_{i=1}^n \left(\frac{d_i A_{ji}}{f_i} - A_{ji} \right) = 0 \quad \forall j, \tag{2.64}$$

and

$$\frac{d_i p_j}{f_i} - p_j + \frac{a_{ji}}{A_{ji}} - 1 = 0 \quad \forall i, j \ . \tag{2.65}$$

These are coupled[34] nonlinear equations. We can simplify these equations significantly. The left hand side of eq. (2.65) only depends on the bin index i and can be rewritten as

$$t_i = 1 - \frac{d_i}{f_i} \ . \tag{2.66}$$

With this eq. (2.65) becomes after sorting the A_{ji} and a_{ji} dependence to the right side

$$t_i = 1 - \frac{d_i}{f_i} = \frac{1}{p_j} \left(\frac{a_{ji}}{A_{ji}} - 1 \right) \quad \forall i, j \ . \tag{2.67}$$

This can be solved for A_{ji} to get

$$A_{ji} = \frac{a_{ji}}{1 + p_j t_i} \ . \tag{2.68}$$

---

[32] This is particularly true for relatively small MC statistics.

[33] That are in this case poisson distributed. A binomial distribution would probably be more exact.

[34] Since the f_i are functions of p_j and A_{ji}.

So if p_j is given A_{ji} only depends on t_i. There are n unknown t_i given by eq. (2.66). Remembering that the predicted number of events in bin i is given[35] by eq. (2.56), we can write

$$f_i = \sum_j p_j A_{ji} \overset{(2.68)}{=} \sum_j \frac{p_j a_{ji}}{1 + p_j t_i} \overset{(2.66)}{=} \frac{d_i}{1 - t_i} \,. \tag{2.69}$$

This is only true for bins with $d_i > 0$ as for empty bins $t_i = 1$ and therefore the equation is not reasonable. Keeping this in mind the m eqs. (2.64) are solved by the m p_j values that minimize the binned likelihood. With the A_{ji} given by the n[36] eqs. (2.69).

2.5.4 Discussion of the Solution

So the way to go about finding the solutions to a binned likelihood fit is to find the m unknowns p_j by solving in iterative steps the eqs. (2.64). At every step the A_{ji} have to be found by solving eq. (2.69). There are some interesting points of discussion concerning the solutions as well as how to accomodate for empty bins where the d_i are zero. Those points are discussed in detail in [41]. Fortunately, in this thesis TRExFitter avoids empty bins by setting them to a very small non zero value. Additionally, the binning will be chosen such that empty bins will not arise for any source. When discussing the application of this method in chapter [4] for sources the term templates will often be used. This is also where these templates will be properly defined using truth information.

It is important to note that in practice the binned likelihood in eq. (2.63) with appropiate gaussian[37] nuisance terms will be solved to give the p_j that then can be used to calculate the A_{ji} with which the binned likelihood can be minimized once again.

As discussed in [42], there is a problem that can arise in this approach though. It can be the case that in the regions of the convergence criterion of the minimization procedure the value for A_{ji} becomes discontinous and makes a jump. This can lead to confusion for minimizers like for example MINUIT's MIGRAD.

An improvement to the procedure can be found by treating the statistical uncertainties arising in a bin not separately for all the sources but for all sources combined. This is reasonable especially when the distributions are gaussian or poisson. This reduces the problem to just a single parameter per bin. In [42] it is shown that this leads to a quadratic equation which can be solved for this parameter. Histfactory employs this method [43].

---

[35] When we include the influence of MC statistics in the calculations by replacing $a_{ji} \rightarrow A_{ji}$.

[36] There are only n equations instead of $n \times m$ since for a fixed p_j only the n t_i are unknown.

[37] They are likely to be gaussian.

CHAPTER 3

Experimental Method

This chapter discusses the experimental methods used for data acquisition. Additionally, the basic steps of MC simulations are reviewed building on the already mentioned concept of MC event generators in 2.4. In addition to data acquisition and event simulation, this chapter discusses particle reconstruction with a heavy emphasis of the reconstruction of the τ lepton.

3.1 The LHC and The ATLAS Detector

The data used for this thesis was collected at the Large Hadron Collider[1] [45] at CERN[2] near Geneva in Switzerland. This section discusses the LHC and the ATLAS detector. It mainly relies on [45] and [46].

3.1.1 The Large Hadron Collider

The LHC is a proton-proton collider designed to achieve a center of mass energies of up to $\sqrt{s} = 13$ TeV [45]. In the figure (3.1) the LHC accelerator complex is shown. In order to achieve colliding beams of energies around 6.5 TeV an intricate acceleration chain is employed. The protons are extracted from hydrogen gas by stripping the hydrogen atoms of their electrons via an electric field. These protons then enter the Liner accelerator[3]. Here radiofrequency cavities accelerate the protons which are then injected to the Proton Synchroton Booster. The protons are further accelerated and enter the Proton Synchroton[4]. In the PS the protons are accelerated to a beam of 25 GeV. From here onward they are injected into the SPS which is the Super Proton Synchroton and reach energies of about 450 GeV. They then enter the beam pipes of the LHC. The protons enter the beam pipes in two beams. One beam circulates clockwise inside the beam pipes the other one counterclockwise in order to let the beams collide. In the beam pipe they are accelerated to their maximum beam energies of up to 6.5 TeV and collided. In order to reach those energies and high luminosities the beam pipes are vacuumized. Additionally an intricate magnetic system is installed around those beam pipes to keep the beam circulating and to make certain that protons with deviating energies are seeded out. This is

---

[1] Shortly called LHC.

[2] Conseil Européen pour la Recherche Nucléaire, (European Organization for Nuclear Physics)

[3] Also called Linac.

[4] Often abbreviated PS.

achieved through multipole magnetic fields. Because of the nature of the Collider, two beam pipes are necessary. Both have to be cooled to allow the magnet system to be in a superconducting state. The strength of the magnetic field reaches up to 8.3 T. Along the beam pipes they are several intersection points where the beams can collide. One of them is designated for the ATLAS experiment. This is where the ATLAS detector is installed. Other intersection points include the CMS ALICE and LHCb experiments. ATLAS and the CMS experiment are both general purpose experiments designed to test the SM and physics beyond. The ALICE experiment studies the properties of quark-gluon plasma and the LHCb experiment the physics of b quarks and CP violation. In addition there are some smaller experiments employed such as the LHCf, TOTEM and MoEDAL experiments. This thesis uses data from the ATLAS experiment as such the ATLAS detector will be discussed in more detail in the next subsection 3.1.2.

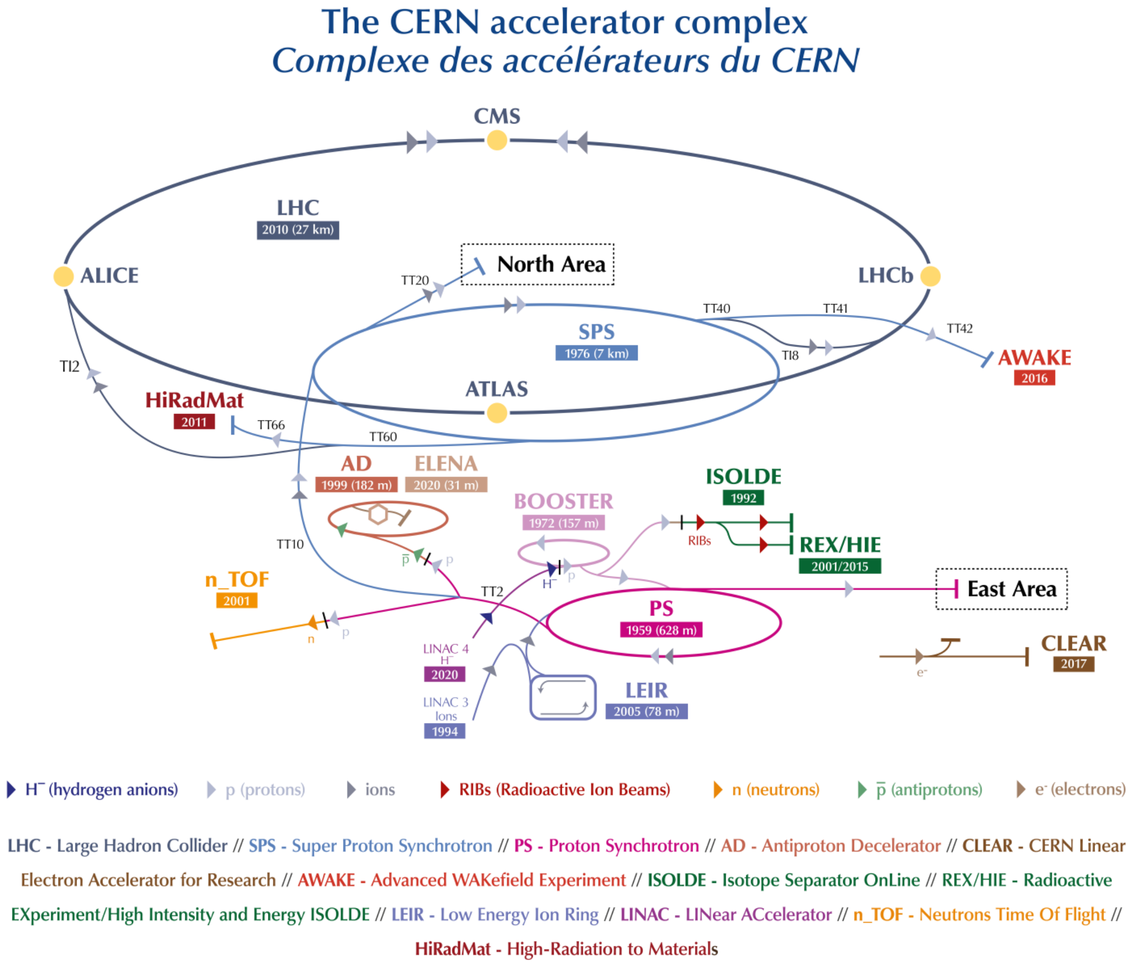


Figure 3.1: The Accelerator complex of the LHC at CERN. Taken from [47].

3.1.2 The ATLAS Detector

The ATLAS[5] detector is seen in fig. (3.2). It has a cylindrical structure. Its length is 44 m and its height is 25 m. It weights around 7000 t. For the ATLAS detector a right handed coordinate system is used. The z-axis follows along the longitudinal direction through the detector such that the plane spanned by the x and y direction is perpendicular to the beam axis. Due to the cylindrical structure we can define the azimuthal angle as the angle along the $x - y$ plane and the polar angle θ with respect to the beam axis. The rapidity

$$y = \frac{1}{2}\ln[(E + p_z)/(E - p_z)] \tag{3.1}$$

is commonly used as a measure for relativistic velocity. The reason for this is that differences between rapidities are invariant under a Lorentz boost along the beam axis. When a particle is highly relativistic the rapidity can be approximated through the pseudo rapidity

$$\eta = -\ln[\tan(\theta/2)] \; . \tag{3.2}$$

Employing the definition of the pseudorapidity and the azimuthal angle we can define regions in the detector coordinate system as well as distances inside the detector similar to eq. (2.52). The detector consists of an inner detector and an outer detector part. The inner detector is the system closest to the beam pipe. It provides position and momentum information as well as information of charged particles. Additionaly, it is used in vertex finding and particle identification. The outer detector consists of calorimeters like the electromagnetic calorimeter and the hadronic calorimeter. The most outer layer

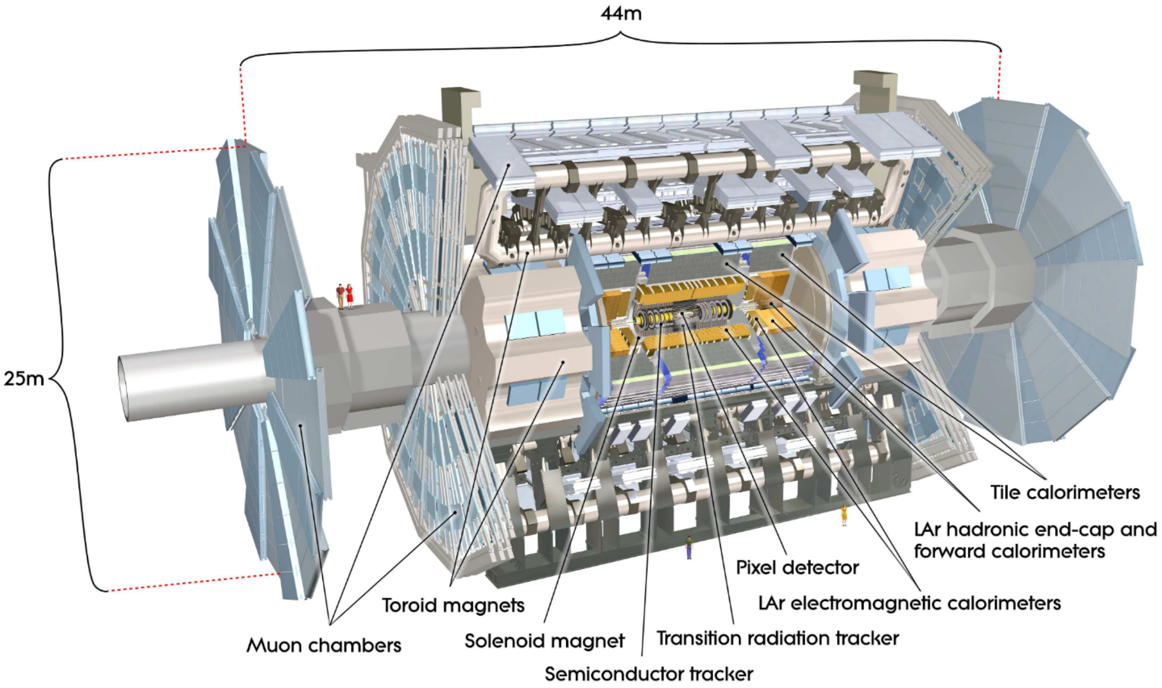


Figure 3.2: A schematic illustration of the ATLAS detector. Taken from [46].

---

[5] ATLAS is short for A Toroidal LHC ApparatuS.

contains the muon spectrometer. The calorimeters measure the energy of the electromagnetic showers as well as the hadronic showers respectively. The muon spectrometer measures the momentum and the trajectory of muons. In addition to these parts, there is a superconducting magnet system inside the detector to acquire momentum and charge information. The system is made up of toroids and provides 2 T for the inner detector and 0.5 T and 1 T for the muon detector in the central and endcap regions. A schematic illustration of the ATLAS detector is given in fig. (3.2).

The Inner Detector

The inner detector covers a range of $|\eta| < 2.5$. It has a cylindrical form and a length of 7 m and a radius of 1.15 m. An illustration of the inner detector can be seen in fig. (3.3). The inner detector consists of three subsystems. The Pixel detector whose main job is the reconstruction of the primary and secondary vertices[6], the semiconductor tracker[7] and the transition radiation tracker[8].

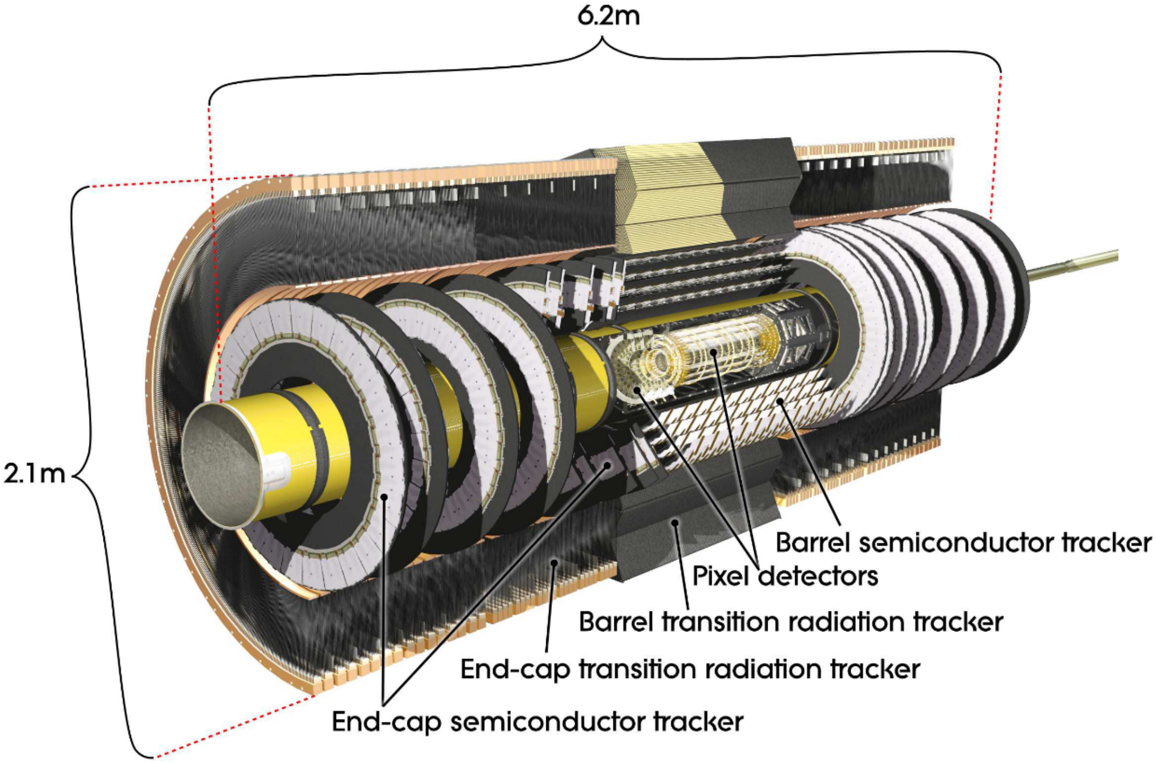


Figure 3.3: The inner detector of ATLAS. Taken from [46].

The **Pixel detector** is a silicon detector with four layers of silicon pixel sensors. It is placed closest to the beam axis of the along the detector with a minimal distance of 33 mm from the beamline [46]. A pixel inside the detector has the dimension of 50×400 µm^2. Overall there are 80 million pixel. For

---

[6] The primary vertex is the vertex of hard scattering while the secondary vertices are the vertices for the decay of the initially produced particles.

[7] In short called SCT.

[8] Shortly called TRT.

the readout it has 46000 channels per module with an abundance of 1744 modules in total. Between the beam pipe and the innermost pixel layer a new layer was inserted, the inserted beam layer. It has approximately 6 million pixels and 224 modules to ensure good performances on vertexing and b-tagging.

The **SCT** surrounds the pixel detector. It is cheaper than the pixel detector per unit area and used for precise tracking. Its distance from the beam pipe is 300 mm. The barrel region contains 4 concentric layers which yields a total number of 2122 silicon strip modules. The dimension of the strip pitch is set to 80 µm in order to achieve precise tracking. This yields a spatial resolution of 16 µm in the $R\phi$ plane and 580 µm along the z direction. In the endcap region the strips are aligned radially with constant ϕ.

The **TRT** is the outer most detector of the three inner subdetectors. Its main purpose is to provide accurate measurements of position and momentum as well as discrimination between electrons and charged pions. This is due to the difference in their transition radiation inside the TRT. The barrel region possesses 52544 straws in parallel direction to the beam pipe. Both endcap regions have 122880 straws radially aligned to the beam axis. Each straw is filled with gas and functions as an individual drift chamber. They are 4 mm wide and 144 cm long. Inside each straw is a thin gold plated tungsten wire. The straws are put on a high voltage such that when particles enter it can function as a drift chamber. The high voltages in combination with the ionization from the particles passing through creates a cascade of charged particles travelling to the anode. Each straw has a resolution of 130 µm in the $R\phi$ direction.

The Calorimeters

Adjacent to the inner detector are the calorimeters. They cover a range of $|\eta| < 4.9$. Their main purpose is the energy measurement of electrons, photons and jets in addition to the identification of those particles. The overlapping detector regions of the calorimeters with the inner detector are grained finer than the rest. They can be separated into the electromagnetic calorimeter[9] that measures the energy of electrons and photons and the hadronic calorimeter[10] which extracts the energy information from hadronically induced parton showers.

Electrons and Photons inside the **ECAL** deposit their energies through particle showers. Via different interactions between the photons and electrons with the detector material like for example pair production and bremsstrahlung a chain of high energy particles are produced. These cascade further due to electromagnetic processes into particle showers. In order to keep the ECAL more compact materials with high atomic number Z are employed. As the mean distance that particles travel until their energy is at $1/e$ of the original value is proportional to the inverse of Z [48] it is advantageous to use materials with higher atomic numbers. Closest to the beam axis in regions where a lot of support structure material is placed a pre-sampler is added to avoid information loss of the particle energies before entering the ECAL. The presampler consists of thin layers filled with liquid argon. The make up of the ECAL is that of alternating layers. One layer consists of an active material to induce particle showers whose energy deposits are recorded and the other of an absorbing material. The design of these two layers is an accordion like shape. As mentioned before the ECAL is longitudinally subdivided into three regions of $|\eta|$ and ϕ. Those regions have different granularities.

---

[9] Abbreviated with ECAL.

[10] Abbreviated with HCAL.

They consist of a barrel part for $|\eta| < 1.475$ and two end-cap parts $1.375 < |\eta| < 3.2$. This can be seen in a schematic illustration of the calorimeter system in fig. (3.4) where the electromagnetic end-caps and barrel are visible.

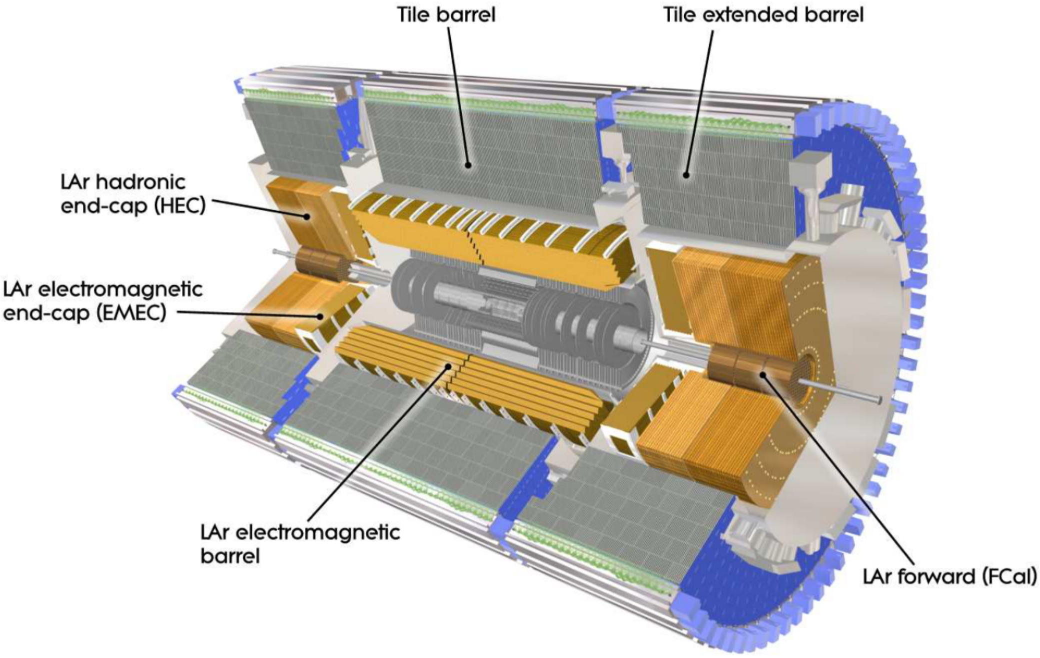


Figure 3.4: Schematic illustration of the ATLAS calorimeter system. Taken from [46].

The purpose of the **HCAL** is to measure the appearance and energy deposit of strongly interacting particles and jets. It consists of a scintillator tile calorimeter in the barrel part that covers $|\eta| < 1.7$. As an absorber the tile calorimeter uses iron and as an active material plastic scintillator plates. In addition to the identification of jets and their energy deposits it also provides a shielding for the muon spectrometer. In a range of $1.5 < |\eta| < 4.9$ are the hadronic endcap calorimeters[11] and a forward calorimeter[12] of high density. The material needs to be resistant to high levels of radiation in these parts of the calorimeter. For this reason a copper-liquid argon mix is used in the HEC and the first part of the FCal. For the two other parts of the FCal a tungsten-liquid argon mix is used. The $|\eta|$ coverage of the calorimeter system is sufficient enough to allow a good missing energy measurement.

The Muon Spectrometer

Because of their high mass and non strongly interacting nature, muons compared to electrons do not usually initiate particle showers. As such they leave the outer detector system described so far without much detection. For the purpose of detecting the muons the muon spectrometer was built. As described before it has its own magnetic field which allows the measurement of the muonic transverse momentum independent of the inner detector. A schematic illustration of the muon spectrometer

---

[11] Also called HEC.

[12] Abbreviated with FCal.

can be seen in figure (3.5). It uses four different chamber systems the monitored drift tubes[13], the cathode-strip chambers[14], resistive plate chambers[15] and thin gap chambers[16]. The MDT and CSC are designed to measure track coordinates while the RPC and TGC are used for triggering due to their fast drift times.

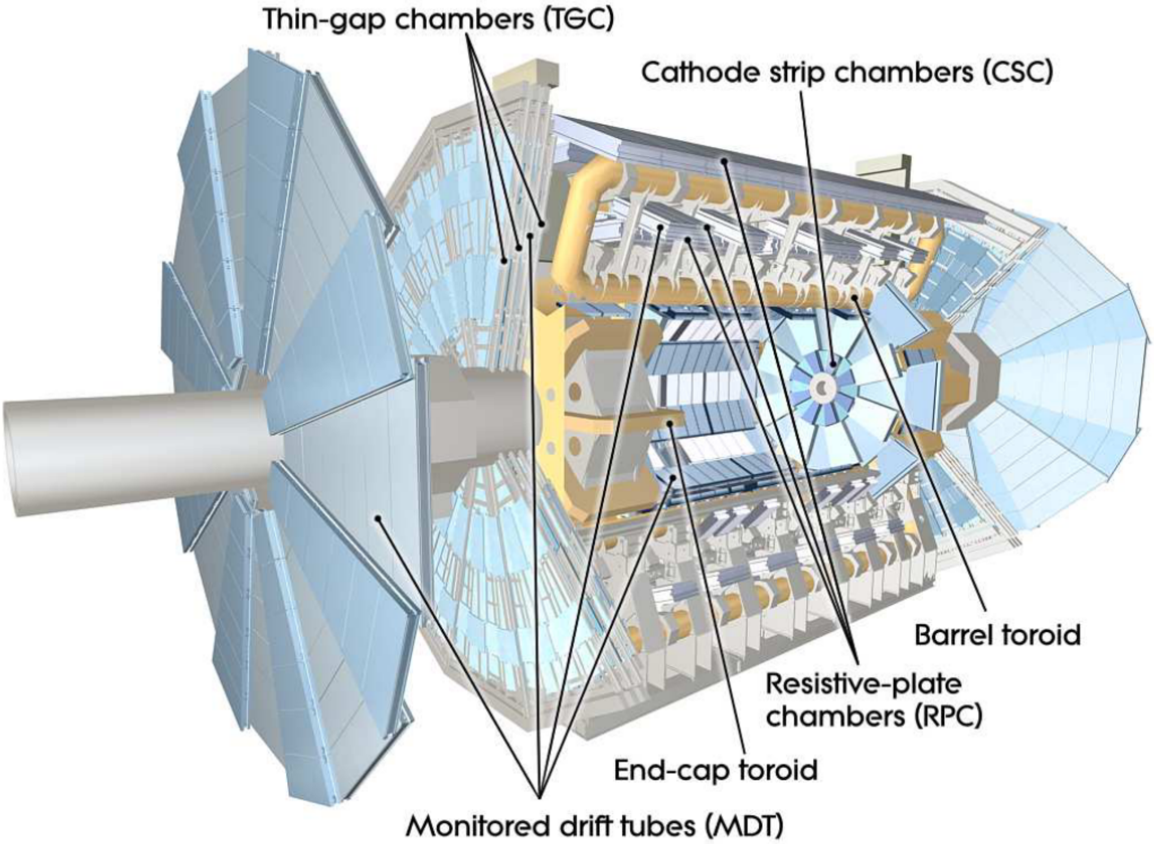


Figure 3.5: Schematic illustration of the ATLAS muon spectrometer. Taken from [46].

As the LHC has an extremely high interaction rate of up to 10^9 Hz an efficient **trigger system** is necessary. The trigger system reduces these interaction rates to include those events that are interesting for the physics analysis. This is done in three consecutive steps. The first step is called Level 1 and uses limited detector information from the muon system and the calorimeters to define a region of interest[17]. These RoI are used for the second step Level 2. It has full access to all the calorimeter information from the RoI and lowers the event rate to 3.5 kHz. The last step is the event filter. It reduces the event rate to 200 Hz. The event filter takes information from magnetic field, alignement and calibration constants into account. This steps happens offline. Afterwards the data will be saved in a mass storage system.

---

[13] Also called MDT.

[14] Also called CSC.

[15] Also called RPC.

[16] Also called TGC.

[17] Abbreviated with RoI

3.2 Reconstruction and Simulation of Events

This section explores the necessary steps of event simulation and reconstruction in a very brief manner. It starts with a discussion of MC simulations as an addendum to the already mentioned event generators in section 2.4. The full set of steps from hard deep inelastic scattering calculations to digitization are lined out. Afterwards the analysis of data from low level analysis methods to particles reconstruction techniques are discussed. For the reconstruction of τ leptons special attention will be given.

3.2.1 Event Simulation

Monte Carlo techniques are used to calculate the multidimensional phase space integrals necessary to simulate particle events at high energy proton-proton collisions. The first step is done by MC event generators which are briefly mentioned in section 2.4. They mimic the initial proton-proton collisions and simulate the fragmentation and hadronization phases. The underlying physics for jets in these phases have been discussed in section 2.4 as well. For these calculations MC integration techniques are used [49]. The most important steps will briefly be summarized in the following again. The first step in the event generation is the hard process calculation which is described in section 2.4. After that parton showers are simulated. These form color neutral states and hadronize in the third step. Those hadrons can then decay and this decay is simulated as well as the corresponding showers and interactions by the decay products including QED radiation. The most used event generators are HERWIG [50], PYTHIA [51] and SHERPA [52]. In the next step the detectors response to the events generated has to be simulated. These detector simulations depend on the detector at hand. In the case of this thesis, on the geometry and material properties, of the ATLAS detector. The detector simulation for the ATLAS detector is done using a software called GEANT4 [53]. The detector simulation can either be done in a fast simulation[18] or a full simulation[19]. The fast simulation uses parametrized parton showers. Similar to data these responses have to be digitized. The digitization process hereby converts the currents and voltages into a response in the read-out system of the detector. The last step is the reconstruction of the simulated data. The MC simulation and samples used in this thesis will be discussed in chapter 4.

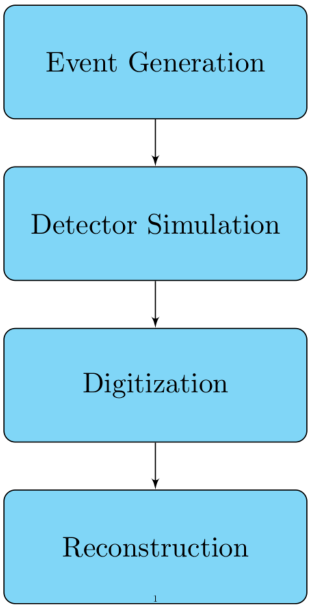


Figure 3.6: ATLAS Monte Carlo Event Simulation

---

[18] Often called AFII.

[19] Often called FS.

3.2.2 Particle Reconstruction

In order to use the acquired data and simulated processes for analysis the physical objets need to be reconstructed. In this subsection the methods and concepts used for this in the ATLAS experiment are briefly discussed.

Object reconstruction begins with low-level analysis where tracks, energy deposition and vertices are used to define low-level objects. The tracks are reconstructed using the Kalman filter algorithm. Here track candidates are determined which need to fulfill certain criteria depending on pixel hits, p_{T} threshold and potential track properties [54]. Track candidates passing these criteria are then extended with TRT information. Vertices are reconstructed using vertex algorithms [55]. Information for this is acquired in the innermost detector layers. In addition to primary vertices secondary vertices usually resulting from particle decays of initial particles are reconstructed as well [55]. Energy deposition of jets are reconstructed using topological clusters. These are a group of cells around a seed whose energy is above a signal to noise ratio threshold. As described in section 2.4 soft radiation from partons is collimated. As such colored partons form jets inside the detector. Their energy deposition in the detector can be seen in these topoclusters from which jet algorithms can be employed to reconstruct the particle object. These algorithms have been discussed in section 2.4.

A special kind of jet are those that stem from b quarks. Due to the relatively long lifetime for b flavored hadrons and their relatively high mass b jets possess unique characteristics that can be employed for efficient tagging. Their high mass leads to wider jets and higher multiplicities as well as a higher invariant mass. In combination with a lifetime that is too short for b flavored hadrons to leave the detector but long enough to be properly tracked these properties allow the tagging of b jets to be relatively efficient [56].

The reconstruction of light leptons[20] is based on track and energy deposit information as well. Electrons produce easily identifiable tracks in the inner detector and energy deposition in the ECAL. Inner detector tracks from primary vertices are matched with the energy deposition in the ECAL where supercluster algorithms are used [57]. Those electron candidates are then discriminated against electron fakes by using multivariate techniques. Employing a likelihood scheme using relevant detector information yields a likelihood value. Based on this value working points that classify electrons into loose, medium or tight identifications can be defined. In order to differentiate between prompt[21] electrons and non-prompt electrons isolation criteria are introduced [57]. Muons do not deposit energies in the calorimeters due to their larger mass compared to electrons. They are reconstructed independently in the inner detector and the muon spectrometer. Using the reconstruction of both sub-detectors the muon candidates are matched and then combined [58]. For these combined muons criteria using a set of variables are used to assess the quality of the combined muon. Similar to the electron identification case, working points can be defined. In the case of muons as well one needs to differentiate between prompt and non-prompt muons. Similar isolation criteria to electrons are applied [58].

Due to the good coverage of the calorimeters in the ATLAS detector it is possible to reconstruct the existence of neutrinos in a reaction by looking at the missing energy since they are not detected in the ATLAS detector. The missing energy is determined by looking at all calibrated hard objects in the calorimeter as well as all soft objects from soft gluon radiation.

---

[20] Light leptons in this thesis refers to electrons and muons compared to the much heavier τ lepton.

[21] Prompt electrons refer to electrons from primary vertex processes whereas non promt electrons originate from photon conversions or hadron decays.

3.2.3 Reconstruction and Identification of τ leptons

As this thesis is interested in the abundance of falsely reconstructed τ leptons, the reconstruction and identification of τ leptons is discussed separately. As such, this subsection focuses on the reconstruction of hadronically decaying τ leptons. The properties of τ leptons have been discussed in section 2.3.

Hadronically decaying τ leptons,decay into 1-prong or 3-prongs. These consists usually of charged pions and additional neutral pions. Decays into 5-prongs are possible but rare compared to 1-prong and 3-prong decays [17]. Due to the boost of the tau lepton the decay products are collimated and form a jet. These jets can be confounded with QCD induced jets which form the main background of the hadronically decaying tau signal. The phenomology of QCD induced jets differs from that of tau induced jets. They are usually wider and their particle multiplicity increases with higher energy.[22] This behavior is illustrated in figure (3.7).

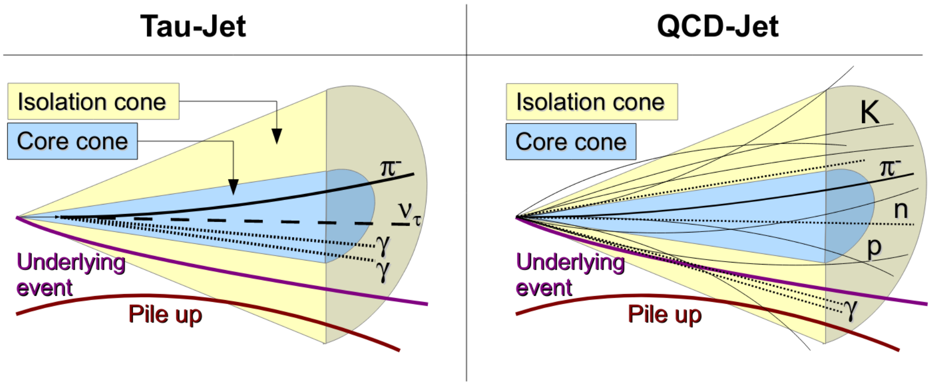


Figure 3.7: Illustration of a QCD induced jet and a hadronically decaying tau lepton. Taken from [59].

The reconstruction and identification of tau leptons can be split into the following steps. First tau candidates are reconstructed using seed jets. In a cone around the central jet axis the jet specific information is extracted and discriminative variables are calculated by using this information. The τ candidates are then identified against QCD induced jets, potential electron and muon fakes by using a variety of methods. These steps will be discussed in more detail in the following. As major references [60] and [59] have been used.

Tau Candidate Reconstruction

Because of the similarity of the detector signature of hadronically decaying taus to the signatures of jets and leptons a dedicated reconstruction algorithm is used to reconstruct hadronically decaying τ leptons.

Jet objects are reconstructed as described in the section 2.4 by using the anti-k_t algorithm with a distance parameter of $R = 0.4$.[23] These jets are then used as potential **seeds** for the reconstruction

---

[22] This is one of the reasons why its easier to differentiate between tau induced jets and QCD induced jets for high transverse momenta. Therefore a better MC/data ratio would be expected for higher p_{T}.

[23] Here R is given by $R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$ as the distance from the jet axis.

algorithm. Additional topological clusters are used as an input for the algorithm. Jets fulfilling the requirements of $p_{\mathrm{T}} \geq 10$ GeV and $|\eta| < 2.5$ are used as seeds for the algorithm [60].

For these seeds the primary vertex is matched using the **Tau Jet Vertex Association**[24] algorithm [60]. This vertex is then later used for track association as well as for the direction of cell and cluster based variables. These variables are calculated, having this vertex as the origin of their coordinate system.

In order to calculate the τ candidate's **detector axis**, a barycenter is formed consisting of the four vector sum of the constituent topological clusters. In all these calculations the τ candidate as well as the constituent information are assumed to be massless. As such, the transverse momentum corresponds directly to the transverse energy. From these barycenters the τ candidates detector axis is calculated in a cone of $R = 0.2$ around the barycenter using clusters. Using the tau vertex coordinate system the four vectors of those clusters are calculated again and summed up. This yields the intermediate tau axis. Using this intermediate tau axis the tau direction in η and ϕ can be defined. Using clusters within $\Delta R < 0.2$ the energy of the τ candidate is calculated. The energy calibration is simulation based and adjusted independently of the jet energy scale [61]. This is because the constituents of τ decays are made up primarily of neutral and charged pions. After the energy calibration and all basic hadronic tau variables have been calculated the final η position is calculated again [60].

Criterias are set for **tracks** inside the radius of $\Delta R < 0.2$ of the intermediate axis such that only tracks satisfying these criteria are accepted. According to [60] those criteria are

- $p_{\mathrm{T}} \geq 1$ GeV,
- number of pixel hits ≥ 2,
- Number of pixel hits + number of SCT hits ≥ 7,
- distance of the closest approach of the track to the tau vertex $d_0 \leq 1$ mm,
- longitudinal distance z_0 of closest approach $|z_0 \sin\theta| \leq 1.5$ millimeter.

Tracks in the range of $0.2 \leq \Delta R \leq 0.4$ used for variable calculation also need to satisfy these criteria. Depending on the number of tracks counted in the core cone[25] τ candidates are classified as being 1-prong or 3-prongs.

As mentioned above, in order to reject τ candidates against jets from the QCD background an additional identification step is necessary. For this a set of **discriminating variables** are calculated in the reconstruction. Those include variables describing the shower shape as QCD induced jets are usually wider than hadronic τ decays for a given p_{T} as well as variables based on the number of tracks reconstructed around the candidate. There are also calorimeter based variables used like the tau energy fraction in the centermost cone.[26] A list of important variables can be found in [60]. Calorimeter based variables are effected by pile up as this pile up leads to extra energy deposition in th detector which leads to extra clusters. To mitigate this the area used for the calculation of calorimeter based variables is reduced to $\Delta R \leq 0.2$. Whereas the discriminative power of the $0.2 \leq \Delta R \leq 0.4$ is still contained in the track based variables [60].

---

[24] Shortly called TJVA

[25] Defined as the cone with distance $\Delta R \leq 0.2$.

[26] Defined by $\Delta R \leq 0.1$.

Discrimination against QCD, Electrons and Muons

QCD induced jets are the main source of background for the hadronically decaying tau signal. Their respective detector signatures are very similar. To reject QCD jets against hadronic tau decays a dedicated tau identification algorithm is used. This is done either with a BDT[27] or a neural network. In the BDT or neural network training one distinguishes between 1-prong and 3-prongs. Only tau candidates satisfying $p_T \geq 15$ GeV and $|\eta| \leq 2.5$ are considered for training. These candidates have to lie within $\Delta R < 0.2$ of a truth matched hadronic tau decay. Based on the output of the BDT or respectively neural network three working points are defined which give a fixed signal efficiency. For 1-prong these are 70% for *tight*, 60% for *medium* and 40% for *loose*. For 3-prongs these are 65% for *tight*, 55% for *medium* and 35% for *loose*. The signal efficiency is defined as the number of hadronic tau candidates passing identification divided by the number of true hadronic taus. A performance review can be viewed in [60].

The nature of 1-prong hadronic tau decays can be mimicked by **electrons**. Properties to discriminate between electron induced jets and hadronic tau decays are transition radiation information which differs for both cases and the shower shape which is longer and wider for hadronically decaying τ leptons. Similar, to the case of discrimination of QCD jets against hadronically decaying taus, a BDT is used. The BDT makes use of variables related to the aforementioned features of electron and tau induced jets. Additionally to that, candidates overlapping with reconstructed electrons are removed and based on the BDT score, three signal efficiency working points are defined. Studies suggest that also 3-prong hadronic tau decays can be confounded with electron induced jets [60].

In the case of **Muons**, they are unlikely to deposit energy into the calorimeters. It is possible though for a muon track to be associated with a sufficiently energetic cluster inside the calorimeter and reconstructed as a tau candidate. To distinguish between those muons and hadronically decaying taus a cut based method is employed [60]. The first cut is to remove tau candidates that overlap geometrically with reconstructed muons. The second cut is based on low efficiency region in the MS. Here the muons are not reconstructed and can therefore not overlap with the tau candidate. Those regions exist for very small $|\eta|$ for this cut the low-efficiency regions are defined to be $|\eta| < 0.1$ and $1.15 < |\eta| < 1.3$. A second cut focuses on the case that the muon deposits enough energy in the calorimeter such that its track in the MS is too skewed and it is not reconstructed as a muon. In this case the energy deposition is mainly done in the hadronic calorimeter. By looking for cases of very few energy deposits in the ECAL and significant ones in the HCAL we can cut this case. The last cut is based on low energy muons that are stopped in the calorimeter and therefore not seen in the MS. These muons must overlap with other energy deposits in the calorimeter to be reconstructed as tau candidates. Because of this their electromagnetic fraction is expected to be large compared to a low track momentum fraction. The resulting efficiency of these cuts is 96% [60].

---

[27] Meaning Boosted Decision Tree. It is a multivariate method exploiting correlations between variables to increase separation power.

CHAPTER 4

Method

This chapter gives an overview of the methodology for estimates of tau fakes. The analysis goal will be defined. The channels of the *tHq* process will be investigated for their assumed process topology and signal and control regions will be defined. Additional the setup will be explained in detail, including the samples and framework used. Lastly the method will be presented and explained in detail. Two methods are used and compared in this thesis. Both have their advantages and disadvantages and those will be investigated.

4.1 Analysis Goal

The goal of this analysis is to estimate the abundances of falsely reconstructed τ leptons. This needs to be done in terms of the sources responsible for faking a τ lepton and the dependencies of the abundance of fakes on the detector region $|\eta|$, the transverse momentum p_{T} and the number of prongs in the hadronic τ decay. The estimation will result in a scale factor[1] that corrects the Monte Carlo prediction to improve the data to MC ratio. This will hopefully improve the possibility of observing the *tHq* process in the ML channel. For this purpose two methods are employed. Both are based on Monte Carlo template binned likelihood fits using the TRExFitter [62] framework.

4.2 The tHq Analysis

The importance of observing the *tHq* process has already been discussed in subsection 2.2.2. A major interest lies in determining the sign of the top Higgs Yukawa coupling Y_t. As already mentioned in this thesis we focus on the ML decay channels. The Higgs decay of interest is the $H \longrightarrow \tau\tau$ decay. Depending on the way the τ leptons decay three distinctive channels can be defined the lep-lep channel $H \longrightarrow \tau_{\mathrm{l}}\tau_{\mathrm{l}}$, where both τ leptons decay leptonically and the two channels of interest for this thesis. Those are the had-had channel $H \longrightarrow \tau_{\mathrm{had}}\tau_{\mathrm{had}}$ where both τ leptons decay hadronically and the lep-had channel $H \longrightarrow \tau_{\mathrm{had}}\tau_{\mathrm{l}}$ where one τ decays leptonically and the other decays hadronically. In addition to the decaying τ lepton, the top quark in the process as seen for example in the Feynman diagrams in fig. (2.5) and fig. (2.6) can also decay into a light lepton. Hence the channels of interest for this thesis are the $2\ell 1\tau_{\mathrm{had}}$ lep-had channel as seen in fig. (2.6) and the $1\ell 2\tau_{\mathrm{had}}$ had-had channel. There is a

---

[1] Which will be abbreviated as SF in the following.

disambiguity where the light leptons originate from in the ML channel. This can be investigated by looking at the sign of the light leptons and their flavor due to charge conservation and the properties of the exchanging gauge bosons.[2] In this thesis we are not interested in this ambiguity and focus on the lep-had and had-had channels as a whole. A primary focus will be placed on the lep-had channel for the purpose of this analysis. This is because estimating the abundances of falsely reconstructed τ leptons is far easier when only including one tau compared to two taus. In this case correlations between the τ leptons and the nature of leading and sub-leading decays needs to respected. Additionally, due to only having one light lepton that can trigger, the statistics for the had-had channel is far smaller than the statistics for the lep-had channel.[3]

4.2.1 Major Background Processes in the lep-had Channel

In order to define reasonable regions to determine the tau fakes, the topology of the lep-had region needs to be investigated. For this purpose the MC simulations are used. The expected background processes have already been discussed in section 2.3. Detailed information about the samples used will be given in the next section. A list of samples can be found in the appendix **??**. The selection cuts required for the samples are discussed in section 4.4.

A table with the yields predicted from MC in the ML lep-had channel can be seen in table (4.1). The table has been split into regions of different working points for the tau reconstruction and amounts of b jets. For example the region Medium b denotes the region for a medium tightness working point for tau reconstruction and a fixed number of $1b$ quark jets. The reason for splitting the lep-had channel beforehand is to better analyze the effects of different selection cuts to the channel. Additionally, since the current SR of the tHq analysis uses a selection for different numbers of b jets, it is useful to investigate the effect of separating for different b jets. For the same reason it is necessary to look at a SR that is topological quite similar to that of the tHq analysis. For the purpose of this section, the Medium tightness $1b$ region will be regarded as the SR and the regions of non-Medium, meaning *loose* tightness, will be used as control regions.[4] From table (4.1) we can see that the signal tHq is expectedly small. The biggest background processes are $t\bar{t}$ for non zero numbers of b jets and Z + jets. The Z + jets process especially dominates for $0b$ jets. In this region the W + jets is also not negligible. Other important processes, concern the Diboson process as well as the ttV process.[5] A visual representation of the process distribution in the lep-had channel can be seen in the piechart, where the miniscule tHq process is left out, in fig. (4.3). In fig. (4.1) the MC prediction of the process yields can be seen in a logarithmic scale for the number of events on the y axis. In both figures the processes have been splitted into the discussed regions. An additional split into regions of different numbers of prongs has been done as well. The yields in the table are scaled to luminosity.

---

[2] Depending for example on a W boson decay, a Z + jets interaction or the signal process which is a Higgs decay.

[3] The terms had-had and $1\ell 2\tau_{\text{had}}$ and lep-had and $2\ell 1\tau_{\text{had}}$ will be used interchangeably in this thesis.

[4] Often abbreviated as CR.

[5] Here the V represents any vector boson.

	Medium $1b$	non-Medium $0b$	non-Medium $1b$	non-Medium $2b$
tHq	3.946	0.659376	0.917081	0.267291
$t\bar{t}$	13 506.2	3 176.18	6 477.85	3 090.41
W+jets	160.07	2 301.83	214.555	9.01619
Z+jets	6 086.67	276 959	9 305.82	640.191
Diboson	314.832	4 036.68	227.387	19.2837
Triboson	1.30047	7.14411	0.389827	0.02127
tZq	63.3718	8.1602	12.0567	3.11947
ttV	242.744	30.7141	69.6031	39.869
tWZ	23.9396	3.41454	5.43931	1.205
ttH	80.4227	6.99326	17.7485	12.6087
t-channel	62.3139	118.459	74.308	4.45802
tW	3.89373	0.680516	0.89862	0.183809
Higgs	22.7982	98.9141	11.7994	3.43347
four tops	0.610436	0.0254858	0.0931098	0.10511
Total	20 573.1	286 749	16 418.9	3 824.17
Data	25 102	295 413	21 377	4 544

Table 4.1: Yields scaled to luminosity of the lep-had channel

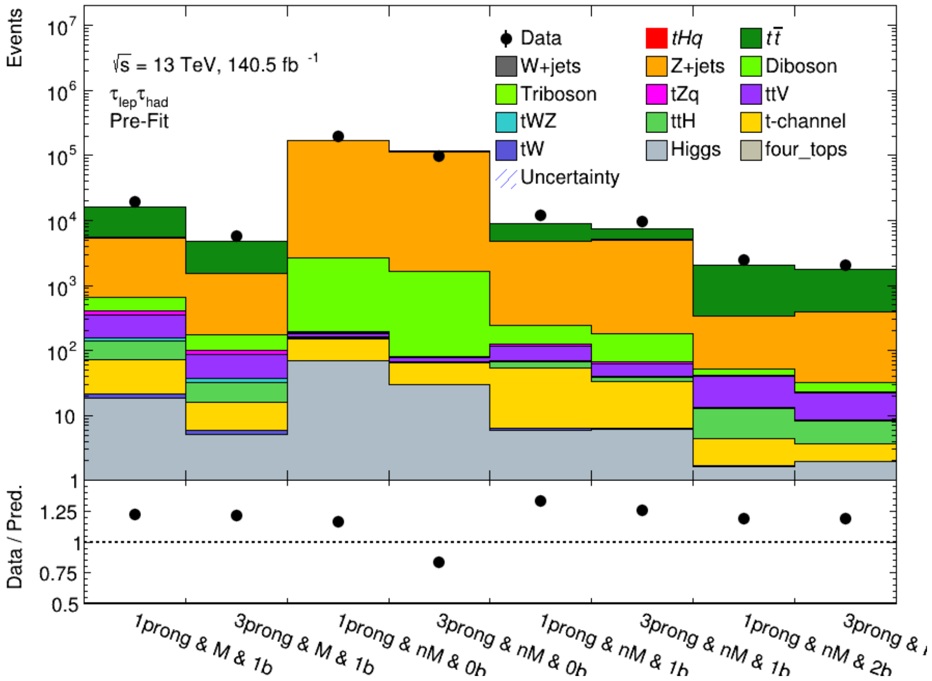


Figure 4.1: Comparison of data and simulated events of various processes contributing to different selection regions related to the lep-had channel.

4.2.2 Major Background Processes in the had-had Channel

In this subsection we take a closer look at the had-had channel of the tHq ML final states. Similarly to the lep-had channel the yields can be seen in the table (4.2). The region splitting is the same as for the lep-had channel. The first thing to notice is that the had-had channel contains far fewer events than the lep-had channel. This can be explained by the replacement of one light lepton with a hadronically decaying tau. Light leptons are good triggers for processes. By having fewer light leptons than in the lep-had channel the effectiveness of the process triggering is effectively reduced. This leads to an overall reduced yields of events. The signal event tHq is similarly rare as in the lep-had channel. For non zero b jets the $t\bar{t}$ process dominates the background again. For the $0b$ region the $Z+$jets and $W+$jets processes dominate. Other processes of interest are the Diboson process and ttV. Similarly to the lep-had channel, a visual representation of the process distribution can be seen in the piechart in fig. (4.5). A breakdown of the processes in terms of regions in a logarithmically scaled plot can be seen in fig. (4.2). In both figures the regions were additionally splitted into regions of different prongs again.

The yields presented in this subsection are supposed to be an estimate for the process distribution in the lep-had and had-had channel respectively. Statistical errors in these estimations have been left out. In addition, we can see that the ad-hoc MC to data ratio in the had-had channel is inferior to that of the lep-had channel and that in both cases the MC prediction mostly undershoots data.

	Medium $1b$	non-Medium $0b$	non-Medium $1b$	non-Medium $2b$
tHq	2.09477	0.223061	0.331998	0.087674
$t\bar{t}$	1 787.9	257.779	539.655	271.642
W+jets	255.302	9 577.46	254.235	11.0626
Z+jets	254.871	3 510.47	159.456	6.99742
Diboson	45.5866	448.155	21.6575	0.889753
Triboson	0.201968	0.854569	0.0314747	1×10^{-6}
tZq	14.1772	1.43949	2.03501	0.523629
ttV	36.5413	3.49224	7.67841	4.38195
tWZ	2.44578	0.237614	0.369067	0.0850032
ttH	25.4773	1.55802	4.03296	2.85946
t-channel	21.6565	20.0067	23.1958	6.34835
tW	1.4737	0.121473	0.210627	0.035243
Higgs	1.15124	44.2201	0.983124	0.446611
four tops	0.0602974	0.00163885	0.007697	0.0100704
Total	2 448.94	13 866	1 013.88	305.37
Data	3 403	14 313	2 042	684

Table 4.2: Yields scaled to luminosity of the had-had channel

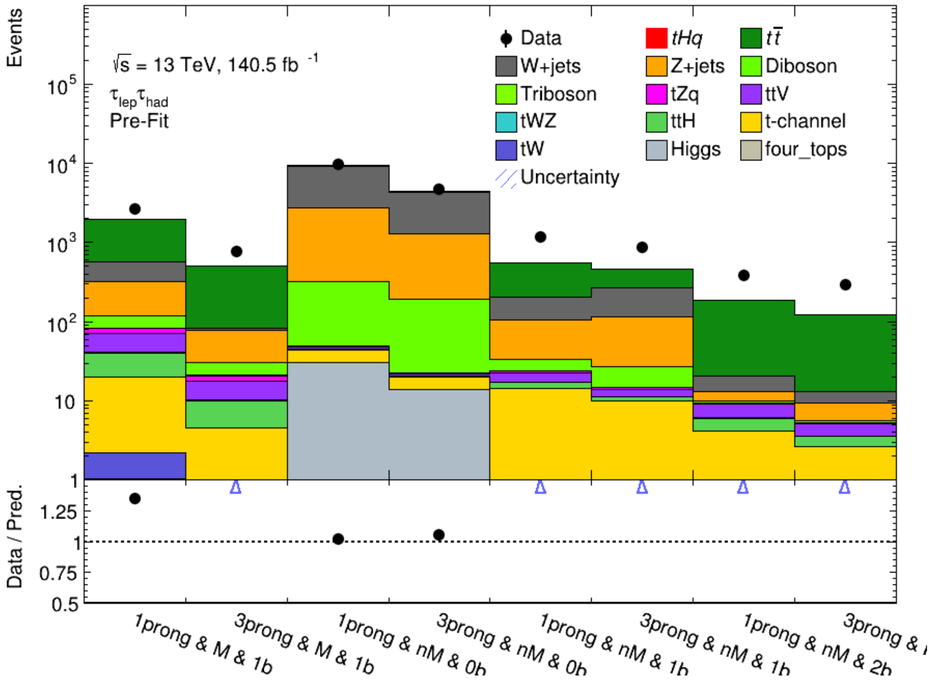


Figure 4.2: Comparison of data and simulated events of various processes contributing to different selection regions related to the had-had channel.

4.3 Region Definitions

In the following section the regions used in this analysis will be defined based on the estimated yields from the MC predictions in section 4.2.

4.3.1 The Signal Region

The SR for this thesis is the region in which an application of the SFs should yield a significant improvement in the MC to data ratio. It has the same process topology as that of the *tHq* analysis.[6] It is required for the SR that the statistics is sufficiently large enough. This allows further cuts to be done in the *tHq* analysis which leaves enough MC and data events for the *tHq* signal fit. As can be seen in the yields it is necessary to work at least with a *medium* tau identification working point. Otherwise the yields for the *tHq* process is far too small and it gets vastly overshadowed by the background. For this purpose the SR in this thesis will demand a *medium* tau identification working point. Furthermore, the number of *b* jets is required to be one as this yields the best probability to observe the *tHq* signal. Therefore, the SR for this analysis will defined as the $1b$ jet *medium* tightness region. It is worth mentioning that additional cuts and requirements for the region definitions are placed on the *tHq* SR.[7] The SR chosen can be seen marked by a red rectangle in fig. (4.3). Here we splitted the regions into 1-prong and 3-prong parts and differentiated the channel in terms of processes. As we can see the $t\bar{t}$

---

[6] The only difference are additional cuts in some variables like neural network outputs.

[7] Those include for example cuts based on NN output variables to separate processes from *tHq*.

process and the $Z+$jets dominate this region.

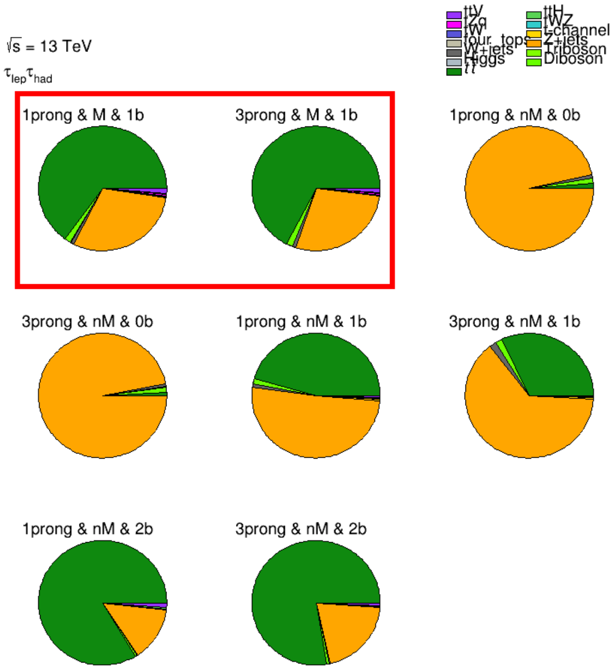


Figure 4.3: Piecharts of the different processes in the lep-had channel for different regions.

4.3.2 The Control Regions

The CRs are used to determine the SFs to correct the MC to data ratio in the SR. For this purpose we need to look at CRs that fulfill a set of requirements based on the level of the analysis. Those requirements are

- enough statistics in order to decrease the statistical uncertainty,
- topological similarity to the SR in terms of processes,
- well defined behavior under MC simulations.

Having enough statistics is necessary especially when splitting the CR into subregions for different variables. In this thesis we will split the CRs into subregions of p_{T}, the number of prongs and potentially $|\eta|$. This multi dimensionality drastically decreases the amount of MC events in a given subregion.[8] Fewer MC events lead to greater statistical uncertainties originating from the simulation. As such high statistics in the chosen CRs is necessary.

In order to calibrate the abundances of fakes and correct the normalization of MC simulations in the best possible way, it is necessary for the CRs to be similar to the SR. There are two major

---

[8] This is also a reason why we employ a likelihood fit and needed to take a proper look at the Barlow Beeston approach to incorporate MC statistical uncertainties.

philosophies to go about insuring that the application of corrections to the fake abundances is sensible. The first one is to concentrate on the details of fake origin and to try to determine a fake factor[9] or SF for all physically different attributes related to fake production. For example one could derive a FF for different p_{T}, $|\eta|$ and number of decay modes[10] for different contributions of quark initiated jets and gluon initiated jets that fake hadronic τ decays. The idea would then be that these factors are universally applicable as long as the method of extraction is done properly.[11] For example, the ATLAS internal Fake Tau Task Force[12] employs this kind of strategy to develope an universal tool of fake tau determination. The benefits of this approach lies in possible high statistics in the specific events that are investigated. The disadvantages are the requirement for clean samples in which the contributions from, in this example quark or gluon fakes, can be safely determined. In such an approach the process topology of the measurement is not relevant for its application. Another possible approach, which is the approach undertaken in this thesis, is to determine the fake abundances in a topology similar to the SR in which the analysis[13] takes place. In such an approach, even though advantageous, it is not necessary to determine the SFs for every possible physically different τ faking process. Additionally, this can be quite difficult to accomplish in a process topology of a certain complexity. Instead, this negligence can be absorbed through the choosing of a decent set of CRs. An important aspect of this thesis is to investigate the duality of both approaches and understand if a possible combination of both can yield significantly good results. This is why the process topology needs to be chosen wisely. So what we want to do is to constrict the τ fakes for the major background processes in our SR. Additionally, the tHq analysis needs to constrict important background processes as well and therefore fits over a series of regions.

The last requirement for a good CR mentioned in the list above is the compability of that region to the method used, its modeling in MC and the understanding of the behavior of the physics of that region under the methods of investigation. For example, the $0b$ region showed strange, not well understood behavior for our method of tau fake estimation. Therefore, we chose to avoid it. This will be discussed in greater detail in chapter [5]. Instead, in order to constrain the fakes originating from the dominant processes in this region which are mainly $Z+$jets and $W+$jets we chose a $1b$ region with an artificially enriched contribution from $Z+$jets events. This is done by loosening up the b tagging efficiency working point from 70% to 85% in the sample production. In this case a higher number refers to a lower light flavored jet rejection abundance. A piechart of the process distribution in this region can be seen in fig. (4.4). The piecharts in the red rectangle show the processes in the region with an enriched abundance of $Z+$jets events. This *alternative* $1b$ *region* will be used as the CR to constrain fakes for $Z+$jets and for $W+$jets events in the lep-had channel. The table of yields scaled to luminosity for this region can be found in appendix A. For the constraint of fakes for $t\bar{t}$ and Diboson events we will use the *loose*[14] regions for $1b$ and $2b$ flavored jets. The same regions are used for our current state had-had analysis. A piechart for the processes in the had-had region can be seen in fig. (4.5).

---

[9] A fake factor, often abbreviated as FF, is a number that is constructed that if applied to a CR it yields the abundances of fakes in a given SR. A proper definition will be given in chapter [5].

[10] This could include number of prongs but also decays into a given number of charged and neutral pions.

[11] This includes the cleanliness of samples for fakes only originating from either quarks or gluons, statistics enabled and other factors as well.

[12] Shortly called FTTF.

[13] In our case for the observation of the tHq process.

[14] Which is called non-Medium in the plots throughout this thesis.

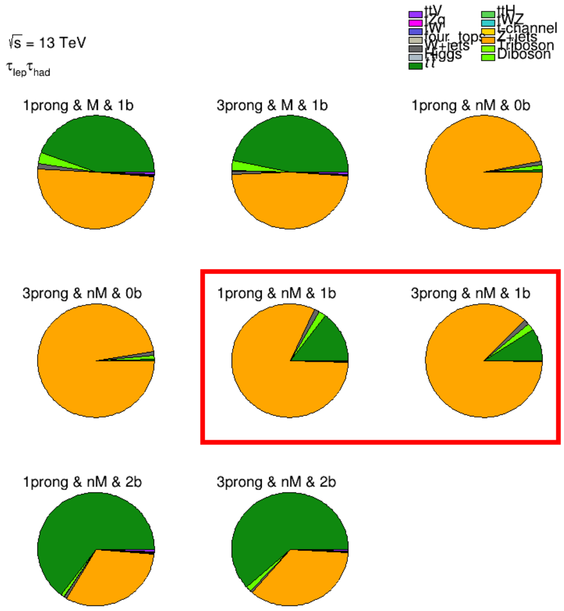


Figure 4.4: Piecharts of the different processes in the lep-had channel with enriched $Z+\text{jets}$ events. Shown for different regions.

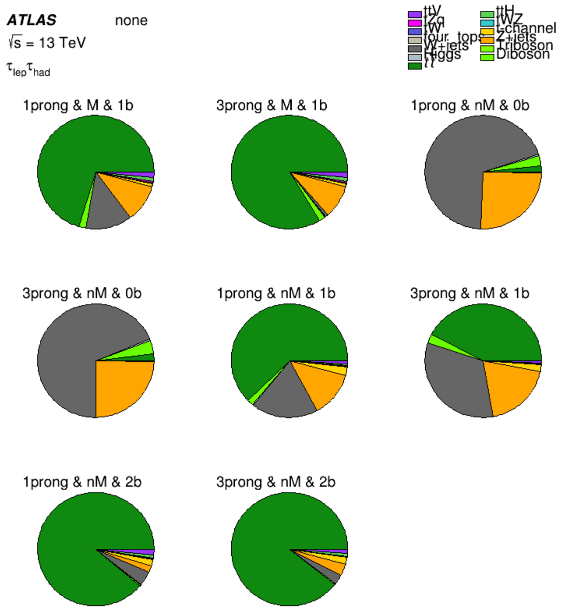


Figure 4.5: Piecharts of the different processes in the had-had channel for different regions.

4.4 Setup

In this section we will explain the setup used for our analysis. This includes the samples and the selection cuts used as well as the framework and additional tools to perform the likelihood fits.

4.4.1 Samples and Data

The details of MC event generation and sample processing has been described in chapters [2] and [3]. The MC samples used in this thesis are based on the Physics Modelling Group Recommendations.[15]. The samples have been produced with AnalysisTop-21.2.163 [63] using the SingleTopAnalysis package[64]. Different MC16 campaigns have been used to create the full set of samples used for this thesis. Here MC16a, MC16d, MC16e correspond to the different PU conditions in the periods of 2016, 2017 and 2018 when the data was taken. For this reason a rescaling is necessary such that the MC simulations integrated luminosity matches that of the data in the given period. Therefore, re-weighting of the MC simulation is necessary. Full simulations are denoted with FS and fast simulations with AFII they mainly differ in the way the parton showers are modeled. The version of these samples used for this thesis corresponds to the internal group name version 32. the list of samples has been chosen to cover the full range of processes in the lep-had or had-had channel respectively without double counting. More information about the samples can be seen in [65]. The SingleTopAnalysis level samples are then post processed using a group internal topLoop derivate tHqLoop to cut the simulation down to the preferred ML analysis channel. In our case these are the lep-had and had-had channels.

The data used in this thesis has been collected by the ATLAS detector at the LHC in the period of 2015 until 2018 with a center of mass energy of $\sqrt{s} = 16$ TeV in Run 2. The total integrated luminosity received from the data sums up to 139.0 fb^{-1}.[16] A more detailed breakdown can be seen in table (4.3).

Year	Integrated Luminosity in fb^{-1}
2015	3.2
2016	33.0
2017	44.3
2018	58.5
Total	139.0

Table 4.3: Table of data and its integrated luminosity taken at the LHC from 2015 up to 2018.

4.4.2 Framework

The framework used in this thesis consists out of TRExFitter [62]. It is a framework for binned likelihood fits. It can be downloaded from github through [62] where more information can be

---

[15] Shortly called PMG.

[16] Uncertainties in the data are not mentioned but can be seen for example mentioned here [65]

found. It uses the format from Histfactory [43] to build statistical models. The framework in which it expresses these models is Roofit with additional classes provided by Roostats. In this thesis TRExFitter was mainly used for fitting and plotting the samples and data.

4.4.3 Selection Cuts

The selection cuts applied to the samples after processing tHqLoop can be found in table (4.4). With additional requirements for jets that are

lep 1	**lep 2**	**lep 3**
$p_{\mathrm{T}} > 27$ GeV	$p_{\mathrm{T}} > 20$ GeV	$p_{\mathrm{T}} > 10$ GeV
Electron	**Muon**	**Tau**
$p_{\mathrm{T}} > 10$ GeV	$p_{\mathrm{T}} > 10$ GeV	$p_{\mathrm{T}} > 20$ GeV
$\|\eta\| < 2.5$	$\|\eta\| < 2.5$	$\|\eta\| < 2.5$
Loose	*Loose*	*Medium (SR)/Loose (CR)*

Table 4.4: Selection cuts for the post tHqLoop processed samples.

- $p_{\mathrm{T}} < 35$ GeV,
- $|\eta| < 4.5$,
- $2 \leq$ number of jets ≤ 6,
- $0 \leq$ number of b jets ≤ 3
- $|\eta| < 2.5$ for b jets.

Additionally we veto the crack detector region $1.37 < |\eta| < 1.52$ and the muon detector crack region discussed in 3.2.

4.5 Methodology

We now focus on the methodology employed in this thesis to estimate the fake taus. The basic idea is presented which will lead to the basic definitions of the templates used for the analysis. Afterwards the method is discussed in detail. This thesis employs two different methods to estimate the falsely reconstructed τ lepton abundance. The first method only considers one parameter of interest, whereas the second method is a direct complexification of the first one. Both methods have their respective advantages and disadvantages which will be discussed. We will mainly focus on the lep-had channel due to it only having one τ lepton that can potentially be faked. This significantly simplifies the analysis. Thoughts about the had-had channel and the projection of SFs to from lep-had to had-had will also be given in this thesis.

4.5.1 The General Idea

The basic idea to determine the fake abundances is the following. We assume MC simulations give a decent prediction of the shape of kinematic distributions concerning tau fakes. We also observe a significant difference between the normalization predictions from MC and the measured data. Therefore, we want to determine the SF for the tau fakes from a likelihood fit of MC to data. This means that we use the shape predicted from MC and the normalization from data. As such this method will be a semi-data driven method relying on MC simulations.

This will be done in the following way. We restrict ourselves to events in the lep-had channel. Every event in the lep-had channel has per definition one reconstructed hadronically decaying tau lepton. Therefore, in principle by counting the hadronically decaying τ lepton events we have the events in the lep-had channel. From MC truth information we know if a reconstructed hadronically decaying tau lepton was actually a tau lepton or if it was another object in the detector that was falsely reconstructed. Using this truth information we split all the lep-had events into those events where a truth tau has been reconstructed as a tau lepton and events where the tau lepton has been falsely reconstructed. These falsely reconstructed events are split again into the different processes of faking a hadronically decaying τ event. By this rule we split our MC simulation into templates. We now perform a binned likelihood fit from MC to data for different sources j as discussed in section 2.5. The different sources in this case are the templates created from truth information. They consist of events where the reconstructed τ lepton was either a truth tau or another object that faked a hadronically decaying tau. The fit allows us to determine the p_j of these templates which correspond to their new normalizations. The difference between the old and new normalization of a given source j is its SF. Therefore, by employing a likelihood fit we improve the MC prediction of the proportion of tau fake events in our CR. This will be done in the defined CRs. These CRs are subdivided into regions of different $|\eta|$, transverse momenta and number of prongs. Both methods differ in the split of fake contributions. This leads to different templates and henceforth to different demands for the fit. This will be discussed in their respective subsections.

4.5.2 General Template Definitions

The basic template definitions are discussed in this subsection. The lep-had events can be split based on truth information into the following templates

- events where the reconstructed τ lepton was a truth tau,
- events where the reconstructed τ lepton was faked by a QCD jet,[17]
- events where the reconstructed τ lepton was faked by an electron,[18]
- events where the reconstructed τ lepton was faked by a muon.[19]

The template concerning the events where the jets faked the tau can further be split into contributions from quark and gluon jets. This split will present the difference between the first and the second method. To understand the MC prediction in these templates we can take a look at fig. (4.6) for the

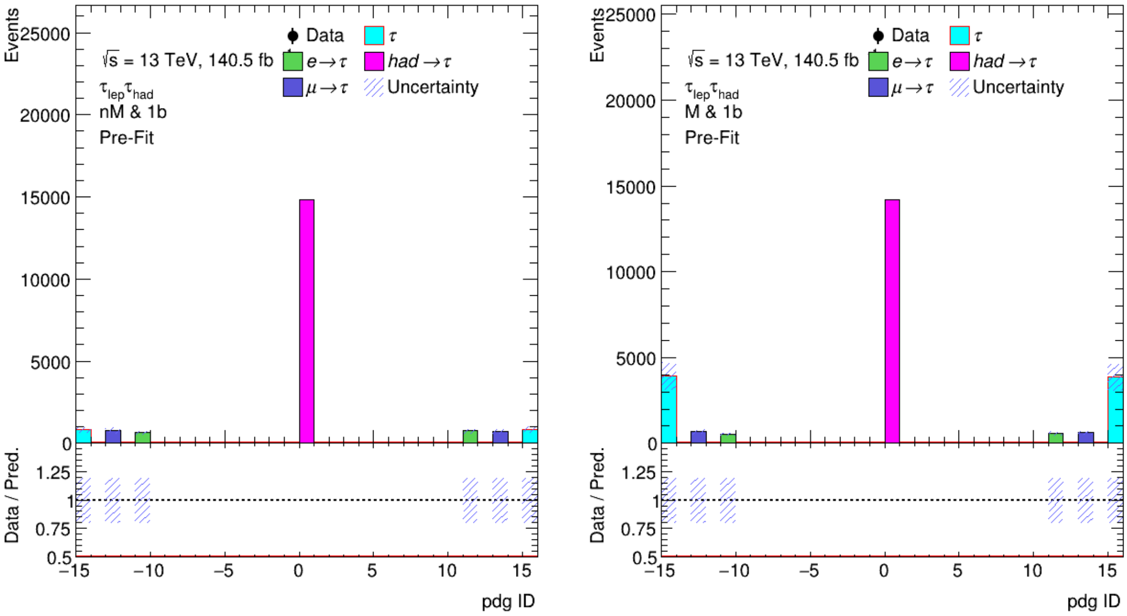
---

[17] Informally called jet faking tau from now on.

[18] Informally called electron faking tau from now on.

[19] Informally called muon faking tau from now on.

lep-had channel. Here the templates defined above are given the color cyan for truth taus that are reconstructed as taus, violet for jet that fake taus, green for electrons that fake taus and blue for muons that fake taus. Additionally to that the plots show the CR[20] for 1b jet and 1-prong and the SR[21] for 1b jet and 1-prong. We can see that the vast number of events are predicted to be QCD jets that are falsely reconstructed as τ leptons. As expected the number of truth tau events increase but we still have a large background in fake contributions. By far the biggest contributions to tau fakes are estimated to be the jet faking taus. For this reason we will mainly focus on the QCD jet contribution to tau fakes.



(a) CR: for non-Medium and 1b jet and 1-prong

(b) SR: for Medium and 1b jet and 1-prong

Figure 4.6: Plots for the events in the lep-had channel split into templates of truth taus (cyan), jet faking taus (violet), electron faking taus (green) and muons faking taus (blue).

4.5.3 The 1-Bin Method

The first method we employed will be called the 1-Bin Method. Because the major contributions to the tau fake abundance are expected to originate from QCD jets that fake the hadronically decaying τ we will restrict ourselves to only fitting the source of this contribution. This is due to the following reason. In order to fit several MC sources in the same fit to the data it is advantageous to have discriminative power between these sources. This will be explained in more depth when discussing the second method. It is not always simple though to find discriminative variables for the different sources. The uncertainty of the prediction is strongly correlated to the discriminative power of these variables. In some cases it might therefore be more advantageous, if the contribution of these sources is expected to be small, that one treats them as nuisances rather than sources for the fits themselves. The treatment

---

[20] Meaning a *loose* (non-Medium) identification working point.

[21] Meaning a *medium* identification working point.

as nuisances then yields a systematic uncertainty originating from these sources. Additionally to, that there is always a correlation between these sources which is the higher the less discriminative the variable is. For these reasons and to reduce our reliance for the shape of the MC prediction we estimate the τ fake abundance by only fitting the source of the QCD jets that fake the τ lepton.

Assuming this we can take a look at the likelihood equations discussed in section 2.5. For simplicity reasons we ignore the contributions that originate from the uncertainty in the MC statistics. This just means that we treat the A_{ji} as given constants rather than parameters of interests themselves. This is totally fine since we showed using the Barlow Beeston method that their determination is independent anyway. Because we only have one degree of freedom, meaning one parameter of interest we do not need to look at a full distribution of the MC simulation for a variable. Instead, it is enough to only look at the total proportion that the source contributes to the MC simulation. This is done using only 1-bin in TRExFitter. Hence arises the name 1-Bin Method. By doing this we avoid the possibility of bad modeling of variable shapes by MC. Unfortunately, we also ignore the different contributions to the fake tau abundance originating from quark and gluon sources. Especially when applying the SFs to a given SR this can be problematic since the proportions of gluon initiated fakes and quark initiated fakes can differ compared to the SR where we determined the SFs. There is no way for us to determine if these proportions are the same other than looking at the predicted process topology of these regions. This is the motivation from complexifying this method by differentiating between quark and gluon initiated fakes as will be seen in the next subsection.

We now take a closer look at the likelihood equation (2.63). As mentioned above, we treat the MC predictions as fixed. The solutions for the SFs are given by the differentiating eq. (2.63) with respect to p_j. This yields the eqs. (2.64) for the different sources j. Because we fix all the sources that are not jet faking taus we end up with only one equation. Additionally because we restrict ourselves to 1 bin $i = 1 = n$ is fixed. The eq. (2.64) becomes

$$\frac{d_i A_{ji}}{f_i} - A_{ji} = 0\,. \tag{4.1}$$

The sum dropped out because we restricted ourselves to only one bin and the different sources and there respective derivatives are zero. Remembering that f_i is given by eq. (2.54) we can write

$$d_i = f_i = \sum_{j=1}^{m} p_j A_{ji} = \sum_{j=1}^{m} N_D P_j A_{ij}/N_j\,. \tag{4.2}$$

Our SFs are then given by $\alpha_j = N_D P_j / N_j$ and all contributions except those from the jets faking taus are fixed to unity. We can drop the subscript i for d_i since awe only have one bin. Calling the sources by their names $j \in \{\mu, \mathrm{el}, \mathrm{had}, \tau\}$ we then have the simple linear equation[22]

$$d = \alpha_\mu A_\mu + \alpha_{\mathrm{el}} A_{\mathrm{el}} + \alpha_\tau A_\tau + \alpha_{\mathrm{had}} A_{\mathrm{had}}\,. \tag{4.3}$$

With

$$\alpha_\mu = 1, \quad \alpha_\tau = 1, \quad \alpha_{\mathrm{el}} = 1 \quad \text{and} \quad \alpha_{\mathrm{had}} = \text{to be determined}\,. \tag{4.4}$$

The simple equation (4.3) can be solved for α_{had} and we determined our SF. Naturally nuisance terms

---

[22] Here had denotes jet faking taus, el denotes electron faking taus, μ denotes muon faking taus and τ denotes truth taus.

are included for the sources that have been fixed in the actual likely hood fit. For those a simple MINOS error is calculated for the nuisance.

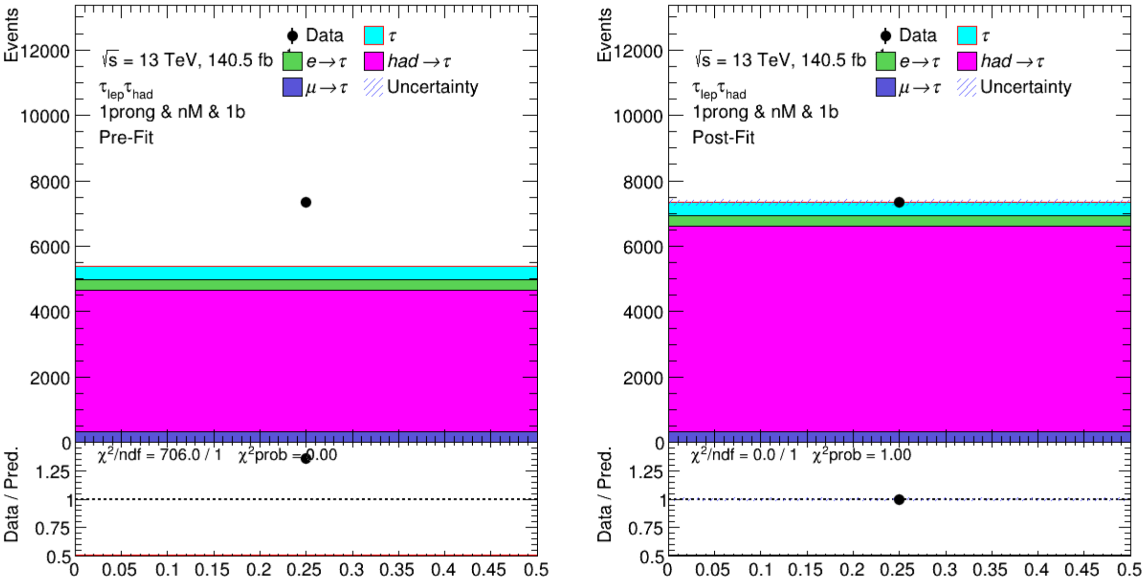
It is prudent to properly define the templates from a technical level. For this truth information is used. The basic definition goes as follows. The variable `had_tau_1_true_pdg` from tHqLoop denotes the truth origin of the reconstructed hadronically decaying τ lepton. Using this and the pdgIds [17] for the different contributions the templates are constructed. Those pdgIds are plotted in fig. (4.6).

A visual representation of a fit in the 1-Bin Method can be seen in the fig. (4.7). In the first row the fit for in the CR region is shown for non-Medium, $1b$ jet, 1-prong and $20 < p_\mathrm{T} < 30$. The second row shows the application of the SF to the SR for Medium, $1b$ jet, 1-prong and $20 < p_\mathrm{T} < 30$. For better visualization in the application to the SR the variable for trackwidth of the jet was used. This variable was explained in eq. (2.53). We can see that the MC to data improvement compared to the pre-fit plot in the SR exists.

In the following an account of the things measured in the 1-Bin method will be given. This includes

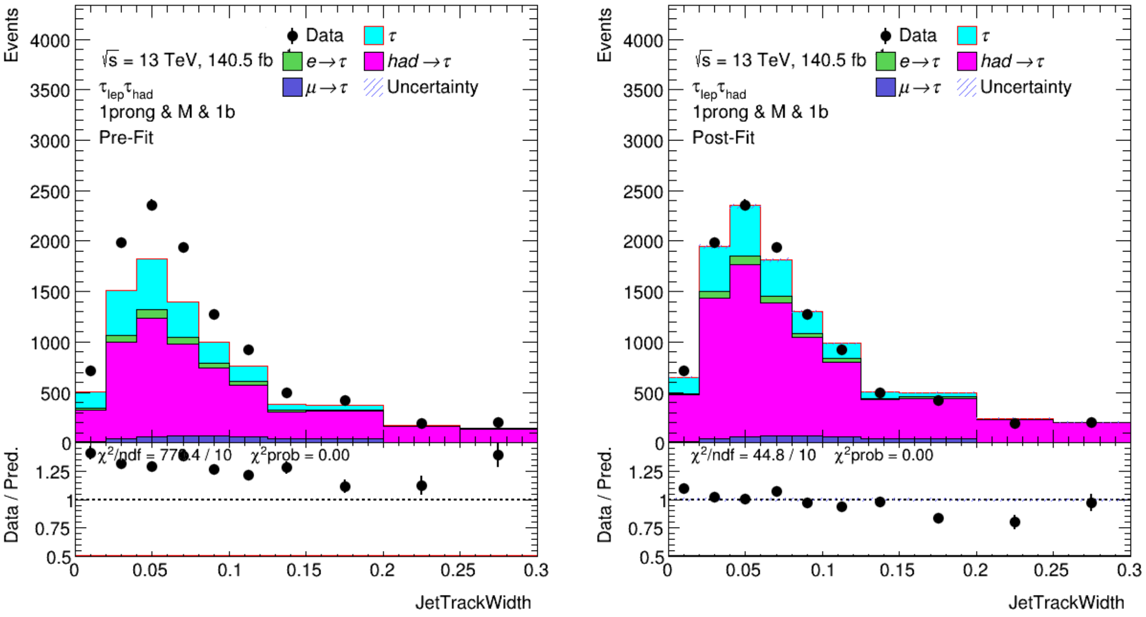
- the dependence of $|\eta|$ for 1-prong and 3-prongs in bins of $[0, 0.8, 1.37]$ and $[1.52, 2.0, 2.5]$,
- the dependence of p_T for 1-prong and 3-prongs in bins of $[20, 30, > 40]$ GeV.

The choice for these bins was based on the requirement for enough statistics per fit and template and for reasons of comparison with the other method.



(a) CR: pre-fit for non-Medium and 1b jet and 1-prong and $20 < p_T < 30$ GeV

(b) CR: pos-tfit plot for non-Medium and 1b jet and 1-prong and $20 < p_T < 30$ GeV



(c) SR: pre-fit for Medium and 1b jet and 1-prong and $20 < p_T < 30$ GeV

(d) SR: post-fit plot for Medium and 1b jet and 1-prong and $20 < p_T < 30$ GeV

Figure 4.7: Example plots for the 1-Bin Method in the CR for pre-fit and post-fit and the SR for pre-fit and post-fit. The plot in the SR is just the application of the SF on the variable *JetTrackWidth*.

4.5.4 The Quark/Gluon Fit Method

The second method is called the Quark/Gluon Fit Method. It is the straightforward complexification of the aforementioned 1-Bin Method. In this method we split the jet faking taus template into contributions from gluon jets and quark jets that fake the hadronically decaying τ lepton. The advantages and disadvantages of doing so have already been discussed. The main advantage is that by looking at the contributions from gluons and quarks to tau fakes separately we respect the different kinematics of quark and gluon jets. A detailed discussion of this can be seen in section 2.4. The disadvantage is that we rely on the shape modeling from MC for a discriminative quark/gluon variable. This variable for reasons discussed in section 2.4 as well is chosen to be the *JetTrackWidth* variable. It is formally defined in eq. (2.53).

The necessity of a discriminative variable follows once again from the likelihood equations. In order to understand this we argue as follows. Let us assume that we only have two sources. The generalization to n sources is straightforward but for simplicity reasons we restrict ourselves to two. Since we have two sources we cannot assume only one bin. As we have seen before this leads to the eq. (4.1) only in this case this equation exists for sources A_{1_i} and A_{2_i}. If both templates have absolutely no discrimination in the variable then we have

$$A_{1_i} = A_{2_i}, \quad \forall i \, . \tag{4.5}$$

This would mean that we have two unknown but only one linear independent equation. This cannot be solved. In fact TRExFitter, when given this case, just puts one source arbitrarily to the minimum amount allowed and fits the other one.

So we need discriminative power in the variable that we choose. This is why for the discrimination between quark and gluon jets the track width was chosen. Additionally we can look at the opposite extreme case to understand that the SFs from both sources are always correlated no matter how good the discrimination of the variable is. this is important since for the gluon and quark discrimination the *JetTrackWidth* variable possesses limited discriminative power. To see that a minimum correlation is always the case we assume that the two templates from the sources 1 and 2 are maximally discriminated in the variable chosen. This means that without loss of generality up to some bin $\tilde{n}$ only contributions $A_1 i$ from source 1 exist and that from $\tilde{n}$ forward only $A_2 i$ contributes. Mathematically spoken this means

$$A_{2i} = 0 \quad \text{for} \quad i \leq \tilde{n} \quad \text{and} \quad A_{1_i} = 0 \quad \text{for} \quad i > \tilde{n} \, . \tag{4.6}$$

In this case we can split the sum in eq. (2.64) and it becomes

$$\sum_{i=1}^{\tilde{n}} \left(\frac{d_i A_{ji}}{f_i} - A_{ji} \right) + \sum_{i=\tilde{n}+1}^{n} \left(\frac{d_i A_{ji}}{f_i} - A_{ji} \right) = 0 \quad \forall j \, . \tag{4.7}$$

When including eq. (4.6) we get the two equations

$$\sum_{i=1}^{\tilde{n}} \left(\frac{d_i A_{1i}}{f_i} - A_{1_i} \right) = 0 \quad \text{for} \quad j = 1 \quad \text{and} \tag{4.8}$$

$$\sum_{i=\tilde{n}+1}^{n} \left(\frac{d_i A_{2i}}{f_i} - A_{2i} \right) = 0 \quad \text{for} \quad j = 2 \, . \tag{4.9}$$

Now, because f_i as seen in eq. (2.54) contains a sum over all sources j depending on the p_j these equations are still correlated.

In reality a prefect discriminative variable does not exist. In addition to this, the likelihood procedure in TRExFitter uses the MINUIT algorithm to minimize the likelihood directly. Still the behavior of these minima are contained in the equations discussed above.

The templates used in this method are the same ones as defined in the 1-Bin Method. The only difference is the split of the jet faking tau template into the quark jet faking a hadronically decaying tau template and the gluon jet faking a hadronically decaying tau template. On a technical level this is done by demanding for those events where the `had_tau_1_true_pdg` variable is 0, which denotes a jet, that an additional variable called `had_tau_1_true_partonTruthLabelID` is either $1 - 5$ for quarks and $1 - 5$ for gluons. There are additional events that are considered jets but are neither in the range $1 - 5$ nor 21. Those events will be merged with the small background template of muons faking taus. Those events are really few in between which means that this can lead to empty bins in this template to mitigate that we merge the two smallest background. Especially since we expect the muon fakes to contribute the least.

The sources that are fitted over are the quarks and gluon jets faking taus. The other sources are fixed to unity and again treated as nuisance parameters in the fit in order to estimate the uncertainty from the contamination.

For both methods the overall variance of the nuisances in the fits will be discussed in chapter [5]. An additional challenge that arises in this method is to choose the correct binning for the variable of the fit. The binning has been carefully adjusted to have enough statistics per bin and to have well defined templates. This includes non-zero bins for every template.

A visual representation of this method can be seen in fig. (4.8). The regions for the fit are the same as for the 1-Bin Method. We can see an improvement in the SR after application of the SF to it the post-fit plot compared to the pre-fit plot. The gluon jets faking tau templates is yellow and the quark jets faking the tau template is red.

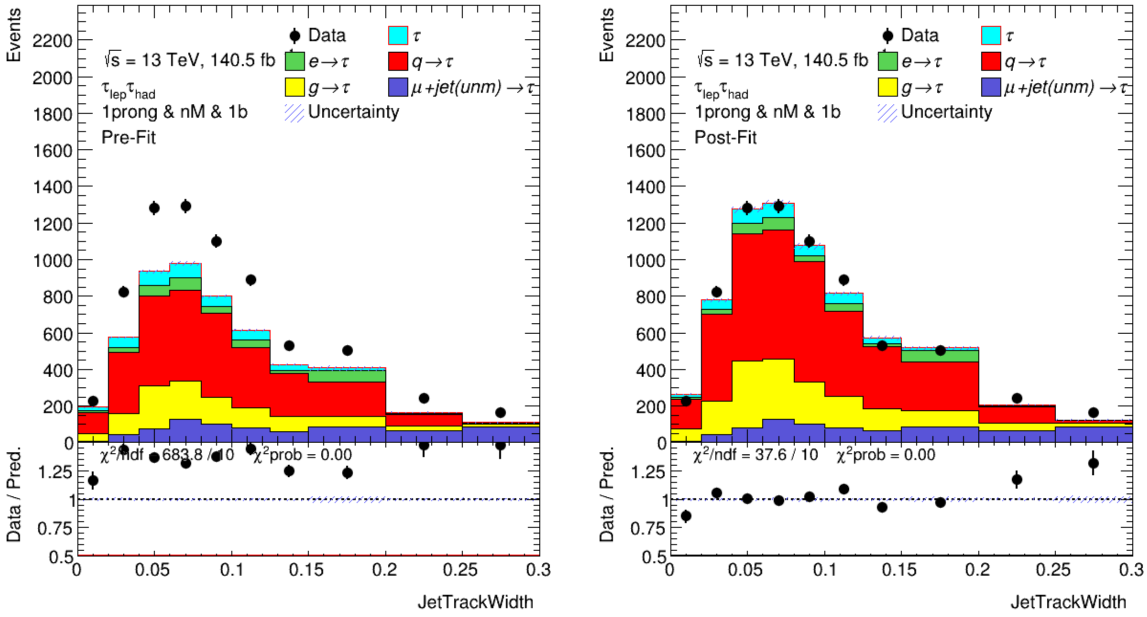
For both methods we did not ask the light leptons to be tight. In fact we did not put any requirement on their working point other than the initial cut. We did not expect the impact of light leptons to be big. Several plots have tended to support that. In order to increase statistics a light lepton working point was therefore not set. This might be a possible oversight in our analysis.[23]

The bins for the estimation in this method in terms of transverse momentum $|\eta|$ are the same as for the 1-Bin Method. The had-had channel has not been used in either method to determine the SFs. This is due to the extra complexity and the significantly lower statistics. For this reason the Quark/Gluon Method is unlikely to be successfully used. Instead of this 1-Bin Method fits for the regions are currently tested. For this the differentiation between leading and sub-leading jets[24] is necessary. Nevertheless, an application of the SFs derived in th lep-had channel to the had-had channel for both methods will be tested for its improvement of the MC to data ratio.

---

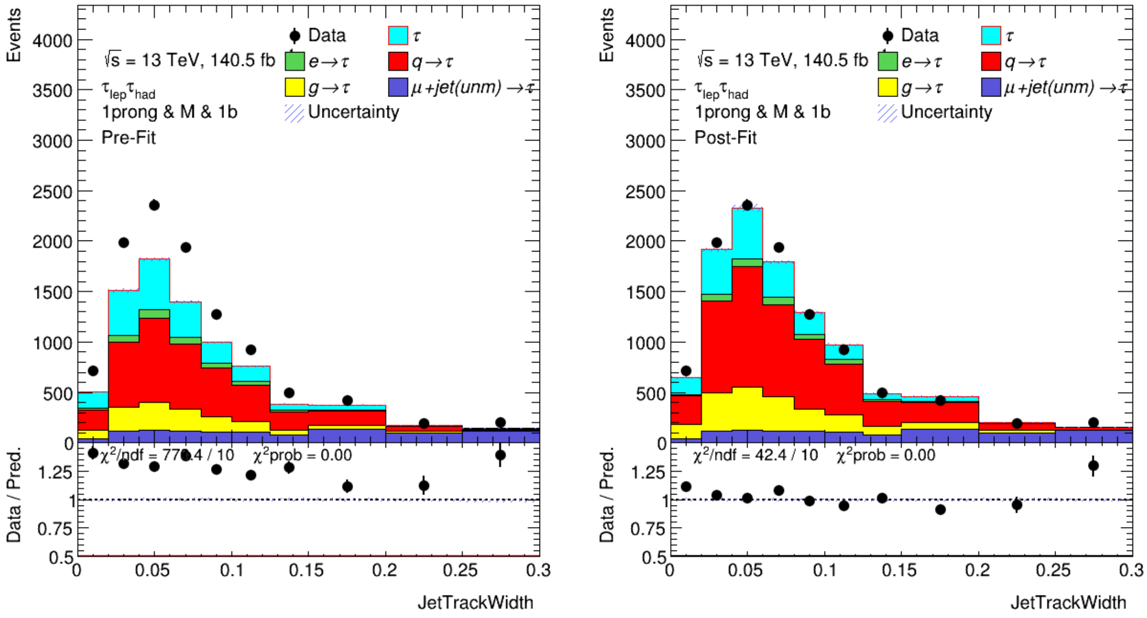
[23] Additionally setting the light leptons to tight proves unexpectedly convoluted in the analysis setup that has been used for this thesis.

[24] Originating from the hadronically decaying tau leptons.



(a) CR: pre-fit for non-Medium and $1b$ jet and 1-prong and $20 < p_{\mathrm{T}} < 30$ GeV

(b) CR: pos-tfit plot for non-Medium and $1b$ jet and 1-prong and $20 < p_{\mathrm{T}} < 30$ GeV



(c) SR: pre-fit for Medium and $1b$ jet and 1-prong and $20 < p_{\mathrm{T}} < 30$ GeV

(d) SR: post-fit plot for Medium and $1b$ jet and 1-prong and $20 < p_{\mathrm{T}} < 30$ GeV

Figure 4.8: Example plots for the Quark/Gluonn Method in the CR for pre-fit and post-fit and the SR for pre-fit and post-fit. The plots in the CR and SR have been done for the variable *JetTrackWidth*.

CHAPTER 5

Results

In the following chapter the results obtained from the methods in chapter [4] will be discussed. First, the potential dependence on $|\eta|$ is investigated and then the SFs determined in the CRs and the alternative region for both methods are presented and discussed. This leads to a comparison between the methods and also between the fake abundances predicted by the Fake Tau Task Force fake factor method. This method will also be briefly presented. In the end the overall application of the SFs to the had-had region, as well as, to a full p_T plot will be presented. The necessary information for every plot will be given either in this chapter or in the appendix.

5.1 $|\eta|$ Dependence

This section investigates the $|\eta|$ dependence of the SFs for both methods. This is done by performing the binned likelihood fit in both methods. Several fits are executed for this purpose in dependence of the number of prongs and different regions of $|\eta|$. The binning in $|\eta|$ was chosen to achieve enough statistics in each bin and to make full use of the tracking coverage. As mentioned before the gap between $1.37 < |\eta| < 1.52$ and the muon gap are excluded. The binning was chosen to be $|\eta| \in [0 : 0.8 : 1.37;\ 1.52 : 2.0 : 2.5]$. The number of prongs are separated into 1-prong and 3-prongs. The SFs were determined with the setup discussed in chapter [4] for both methods respectively.

5.1.1 $|\eta|$ Dependence: 1-Bin Method

As mentioned before, for this fit only the template of jets faking the hadronically decaying tau lepton has been varied, while the other sources are fixed. The range allowed for the jet faking tau normalization to vary is from 0.1 to 10. A list of the results of these fits can be seen in table (5.1). The fits have been performed without the use of systematics including the influence from the normalization of the contamination.[1] Even without these contributions to the total uncertainty a dependence of the SFs on $|\eta|$ can be excluded from the 1-Bin Method estimation. A visual representation of this behavior can be seen in fig. (5.1). The variable for which the SFs have been applied in the SR is the *JetTrackWidth* variable. The uncertainties mentioned are statistical uncertainties only. A detailed view of MC to data

---

[1] This is due to a limitation of time in the submission of this thesis rather than a preference in the procedure. A better analysis would include the systematic uncertainties in the estimation of this dependence. It is clear to see though that even without the systematic uncertainties from the other sources included the SFs do not depend on $|\eta|$.

improvement of every plot in this $|\eta|$ analysis for the 1-Bin Method can be seen in section B. The SR is defined as *medium* and 1b jet and the CR in which the fit was done is defined as *loose* and 1b.

	1-prong	3-prongs
$\|\eta\|$	SF (had)	SF (had)
0 : 0.8	1.43 ± 0.02	1.40 ± 0.03
0.8 : 1.37	1.38 ± 0.03	1.36 ± 0.03
1.52 : 2.0	1.33 ± 0.03	1.29 ± 0.04
2.0 : 2.5	1.34 ± 0.03	1.32 ± 0.04

Table 5.1: SFs determined in the 1-Bin Method in dependence of $|\eta|$ for 1-prong and 3-prongs.

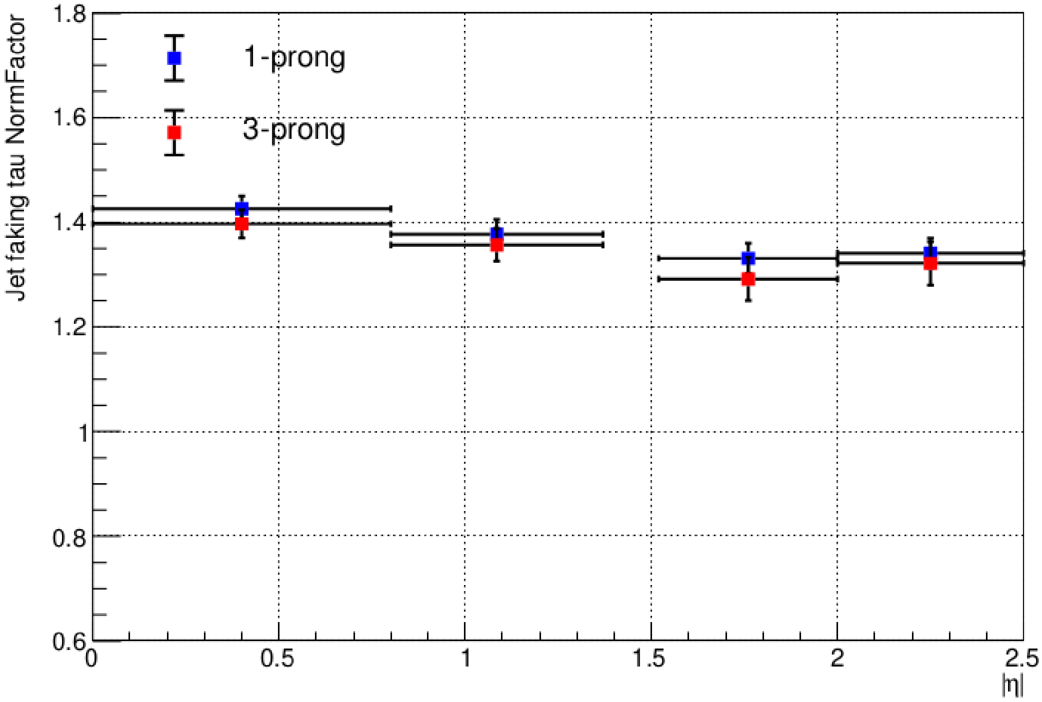


Figure 5.1: Dependence of the SFs for different $|\eta|$ for 1-prong and 3-prongs estimated in the 1-Bin Method.

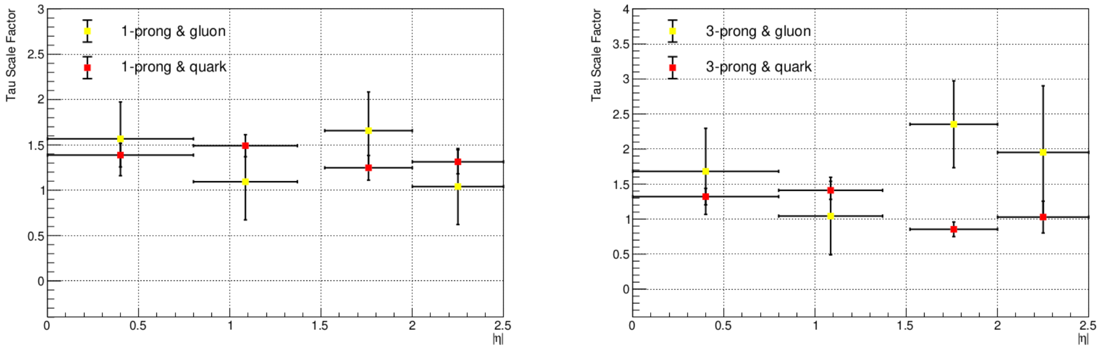
5.1.2 $|\eta|$ Dependence: Quark/Gluon Method

In a similar manner, the $|\eta|$ dependence was estimated for the Quark/Gluon Fit Method. The region cuts in $|\eta|$ and the number of prongs are the same as in the 1-Bin Method. The SR and CRs are also the same. The fits have been performed with respect to the varying templates of quarks faking taus as well as gluons faking taus. The range in which these normalizations have been allowed to vary is from 0.1 to 10. The other sources have been fixed to unity. Again, the contribution of the contamination to the systematic uncertainty has been neglected.[2] Some additional complications arise in the Quark/Gluon Method compared to the 1-Bin Method. The most prominent of these is the choosing of the right binning, such that the templates are always well defined and have zero empty bins. similar to the 1-Bin Method we can exclude a dependence of the SFs on $|\eta|$ within the given uncertainties. The uncertainties are statistical uncertainties only.

The results can be seen in table (5.2). The SFs are visualized in the figures (5.2). The detailed view of the pre- and post-fit plots in the SR can be seen in B.

	1-prong		3-prongs	
$\|\eta\|$	SF (gluons)	SF (quarks)	SF (gluons)	SF (quarks)
0 : 0.8	1.57 ± 0.41	1.39 ± 0.13	1.68 ± 0.61	1.32 ± 0.12
0.8 : 1.37	1.09 ± 0.42	1.49 ± 0.12	1.04 ± 0.55	1.41 ± 0.13
1.52 : 2.0	1.66 ± 0.43	1.25 ± 0.14	2.35 ± 0.62	0.85 ± 0.10
2.0 : 2.5	1.04 ± 0.42	1.31 ± 0.13	1.95 ± 0.95	1.03 ± 0.23

Table 5.2: SFs determined in the Quark/Gluon Method in dependence of $|\eta|$ for 1-prong and 3-prongs.



(a) 1-prong

(b) 3-prong

Figure 5.2: Dependence of the SFs for different $|\eta|$ for 1-prong and 3-prongs estimated in the Quark/Gluon Fit Method.

---

[2] For the exact same reasons as in the 1-Bin Method. The arguments are the same as the ones mentioned before.

5.1.3 $|\eta|$ Dependence: Conclusion

From the figures (5.1) and (5.2) or alternatively from the tables (5.1) and (5.2) we can see that there is no dependence of the SF on $|\eta|$ within the statistical uncertainties. We can also see in appendix B that for each fit there is a good data to MC improvement when applying the SFs to the SR. So both methods agree in their predictions and for both methods the fits look proper. There is, however, the aforementioned caveat of having left out the influence that the contamination by other possible contributing sources might have. Especially for muons it can be expected that their SFs are dependent on $|\eta|$ since their identification and reconstruction is dependent on the η coverage of the Muon Spectrometer and its ability to identify muons. We can however reasonably assume due to high signal efficiencies in muon identification [58] that these contributions do not change this trend. In fact, because we only directly determine the SFs from jets faking taus or their split into respective gluon and quark related parts, it is a reasonable enough argument that leaving out the $|\eta|$ dependence is justified. The indication from both methods for this is strong. In addition to that, an estimation of the SFs in a multidimensional binning including $|\eta|$, p_T and the number of prongs is not feasible due to statistical limitations. For this reason alone the $|\eta|$ dependence would have to be dropped. In this sense the aforementioned discussion based on the estimation in both methods is just an additional argument for a step that is necessary due to statistical circumstances anyway. In the following analysis the influence of the contamination of the contribution from the other sources will not be ignored.

5.2 Scale Factors for the 1-Bin Method

We now present the procedure and results from the SF estimation for the 1-Bin Method. The first CR was chosen to be *loose* and $1b$ jet, while the SR was chosen to be *medium* and $1b$ jet. We only vary one source.[3] The other sources are fixed to unity. The range for the fit of the jets faking taus source is from 0.1 to 10. This allows enough freedom for the likelihood fit to find its minimum. The contamination of the other sources, with special attention given to the electrons faking taus source are treated as nuisances in the fit. Their variance range has been determined by looking at the signal efficiencies of their identification as seen in for example [57], [58] and [60]. They have been chosen as 0.02 for the muon and truth tau templates and 0.2 for the electron template. The 1-Bin Method was employed.[4] The fits were produced in regions cut for 1-prong and 3-prongs and a binning of $p_T \in [20 : 30 : 40 :> 40]$. For every fit an asimov fit was performed to test the influence of the systematics.[5] Every fit was performed and then tested for the systematics. The results of the fit can be found in table (5.3). A visual representation is given in fig. (5.3). In the figures (5.4) and figs. (5.7) the pre-fit and post-fit[6] plots of the SR can be seen. In addition to this, the application of the SFs to the SR for a full range p_T plot can be seen in fig. (5.10) for 1-prong and in fig. (5.11) for 3-prongs. To improve visuality this region was binned and the variable used was the *JetTrackWidth* variable. A good improvement from MC to data ratio is visible. Because we only vary one component the correlation matrix will not be discussed. Instead, the deviation of the nuisance parameters can be seen in appendix D. Their deviation is expectedly small.

---

[3] The source of jets faking taus, denoted with had.

[4] Which was explained in chapter [4].

[5] Consisting of the normalization from the other sources.

[6] Meaning after applying the SF to the region.

	1-prong	3-prongs
p_T in GeV	SF (had)	SF (had)
20 : 30	1.45 ± 0.03	1.36 ± 0.03
30 : 40	1.36 ± 0.04	1.39 ± 0.05
> 40	1.24 ± 0.05	1.16 ± 0.05

Table 5.3: SFs determined in the 1-Bin Method in dependence of p_T for 1-prong and 3-prongs.

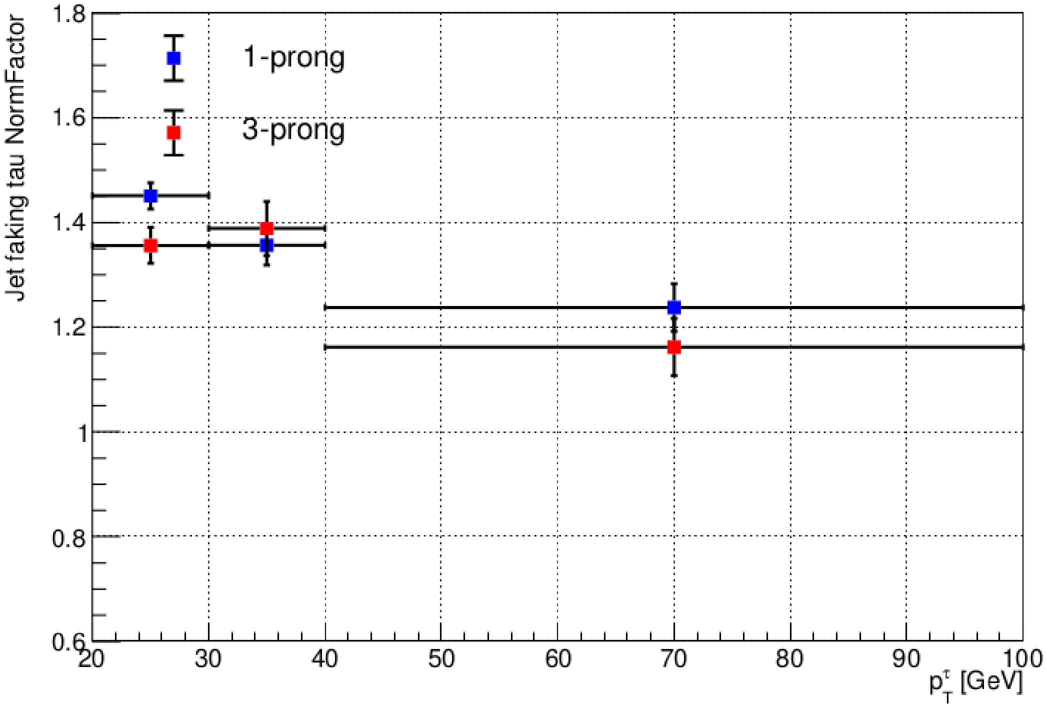
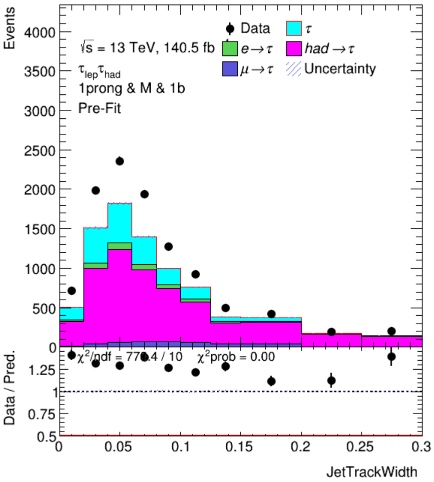
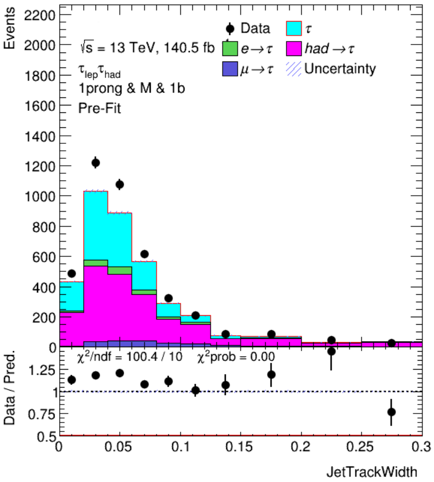


Figure 5.3: Dependence of the SFs for different p_T for 1-prong and 3-prongs estimated in the 1-Bin Method.

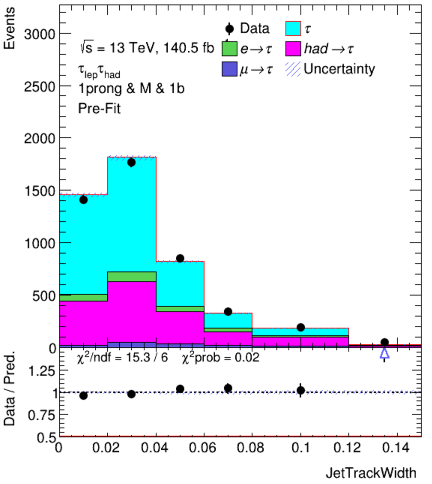
Figure 5.4: Pre- and Post-fit plots for the SR in the 1-Bin Method estimation of the p_{T} dependence of the SFs. For 1-prong.



(a) $p_{\mathrm{T}} \in [20:30]$ in GeV

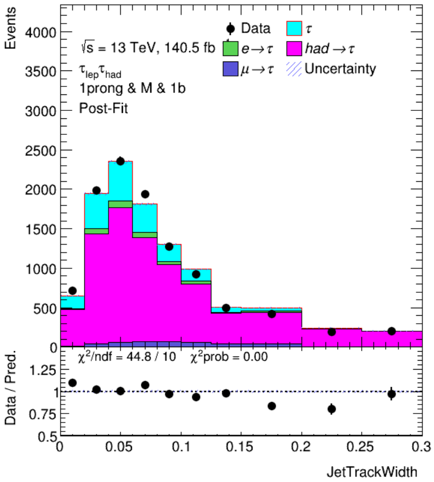


(b) $p_{\mathrm{T}} \in [30:40]$ in GeV

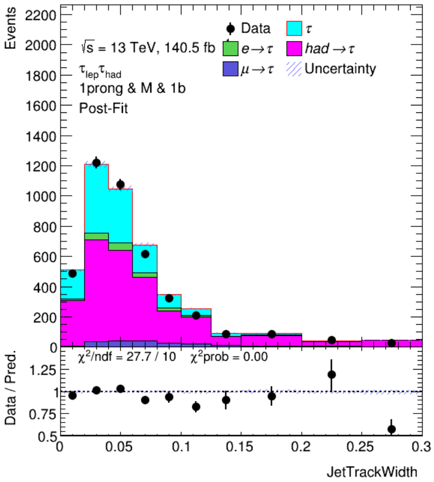


(c) $p_{\mathrm{T}} > 40$ in GeV

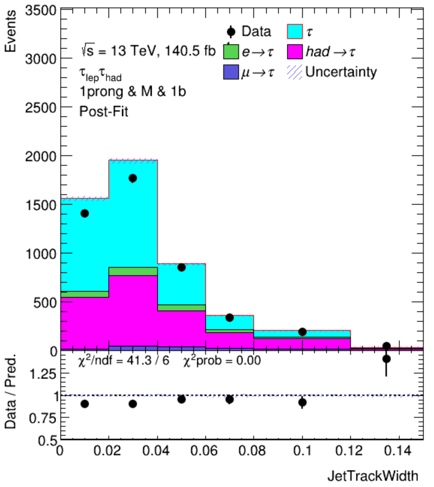
Figure 5.5: Pre-fit Plots



(a) $p_{\mathrm{T}} \in [20:30]$ in GeV



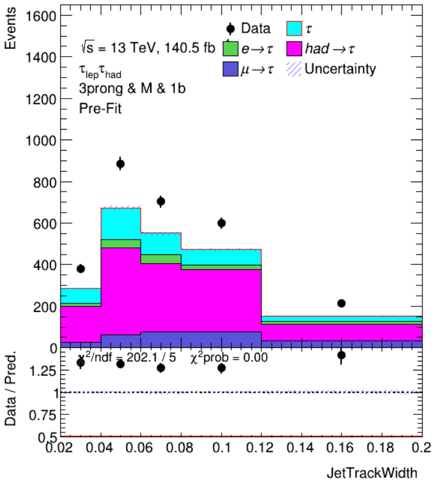
(b) $p_{\mathrm{T}} \in [30:40]$ in GeV



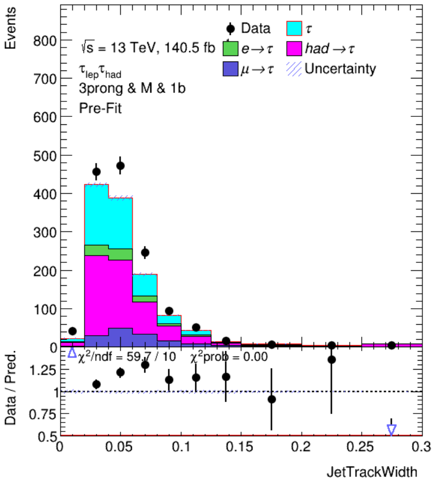
(c) $p_{\mathrm{T}} > 40$ in GeV

Figure 5.6: Post-fit Plots

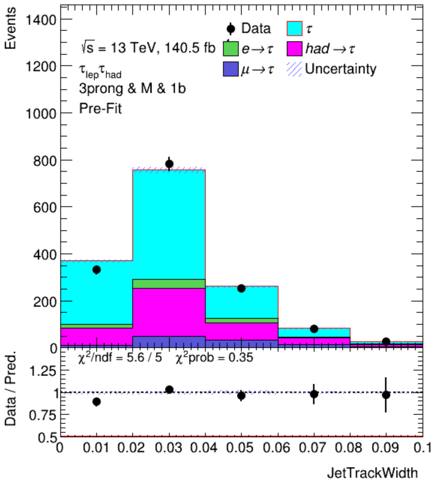
Figure 5.7: Pre- and Post-fit plots for the SR in the 1-Bin Method estimation of the p_T dependence of the SFs. For 3-prong.



(a) $p_T \in [20:30]$ in GeV

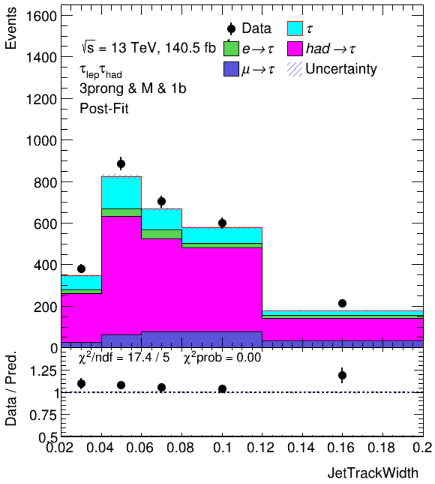


(b) $p_T \in [30:40]$ in GeV

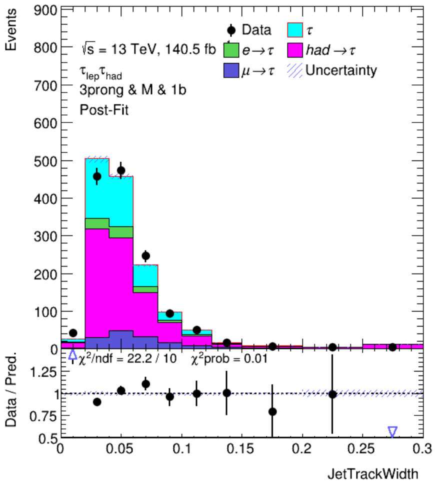


(c) $p_T > 40$ in GeV

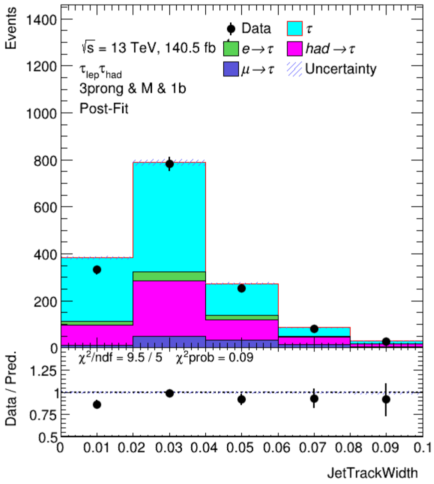
Figure 5.8: Pre-fit Plots



(a) $p_T \in [20:30]$ in GeV

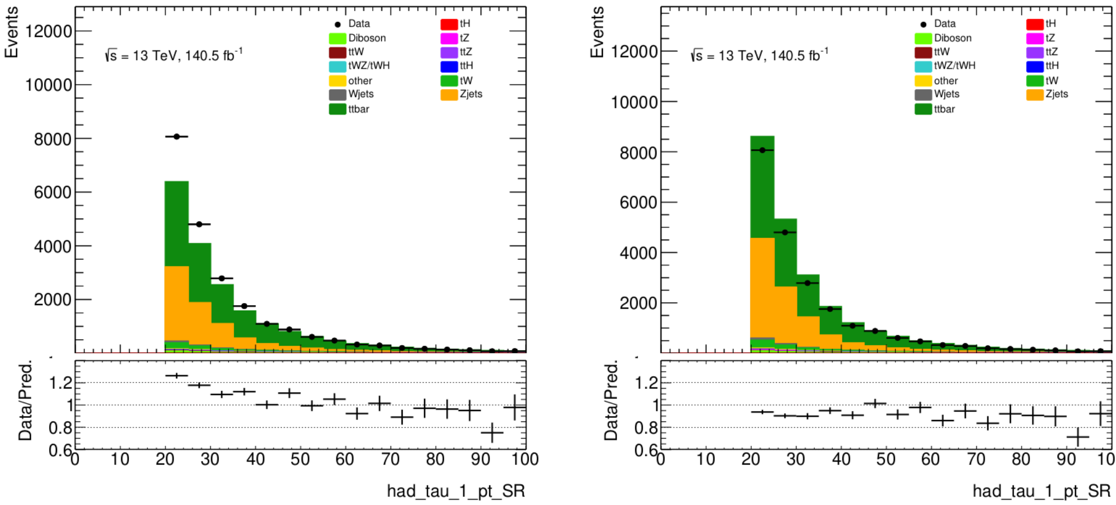


(b) $p_T \in [30:40]$ in GeV



(c) $p_T > 40$ in GeV

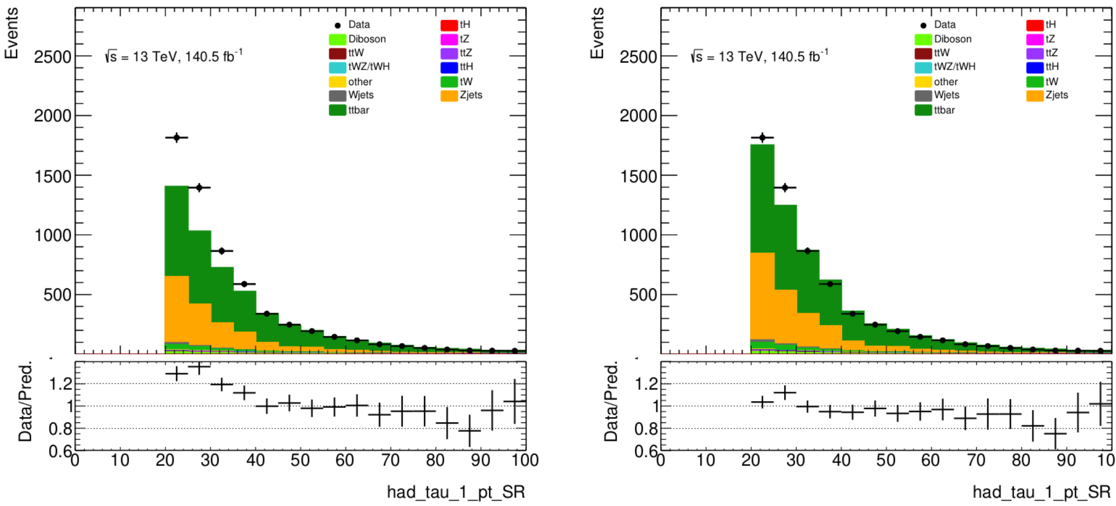
Figure 5.9: Post-fit Plots



(a) Pre-Fit

(b) Post-Fit: SFs from the 1-Bin Method Method.

Figure 5.10: Full p_{T} plot in the SR for 1-prong.



(a) Pre-Fit

(b) Post-Fit: SFs from the 1-Bin Method Method.

Figure 5.11: Full p_{T} plot in the SR for 3-prong.

5.3 Scale Factors for the Quark/Gluon Method

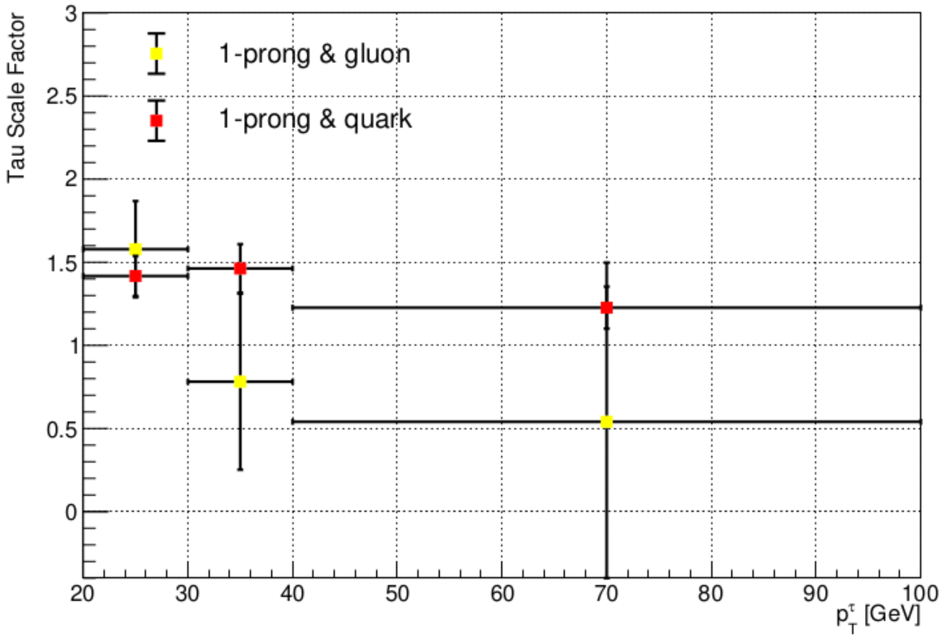
We now present the procedure and results from the SF estimation for the Quark/Gluon Fit Method. The first CR was chosen to be *loose* and $1b$ jet while the SR was chosen to be *medium* and $1b$ jet. We vary two sources, those being the quark jets faking taus and gluon jets faking taus. To discriminate between these templates, we fit over the *JetTrackWidth* variable. The contamination of the other sources with special attention given to the electrons faking taus source are treated as nuisances in the fit. Their variance range has been determined by looking at the signal efficiencies of their identification as seen in for example [57], [58] and [60]. They have been chosen as 0.02 for the muon and truth tau templates and 0.2 for the electron template. We employed the Quark/Gluon Fit Method. In addition to fitting over the $1b$ jet CR we simultaneously fitted over the $2b$ jet *loose* region as well. As discussed in chapter [4] this is reasonable. The addition of this region allows us to reduce the statistical uncertainties. For each fit an asimov fit was produced to ascertain the effect from the nuisance parameters. Then the fit was performed in a carefully chosen binning to maintain the well defined templates and avoid negative or zero bins. After each fit the correlation matrix was checked as well as the influence of the systematics. This includes the comparison between the nominal and their $+1\sigma$ and -1σ variations as well as their correlations to themselves and the parameter of interests.[7] In addition to that the MC to data improvement in the CRs and SR were investigated as well as the influence of pruning the systematics. Based on these criteria, fits were either accepted as good or improved. The SFs from the Quark/Gluon estimation can be seen in table (5.4). A visualization of the SFs can be seen in the figures (5.12). In addition to this the application of the SFs to the SR for a full range p_{T} plot can be seen in fig. (5.19) for 1-prong and in fig. (5.20) for 3-prongs. The uncertainties are the total uncertainties that includes the systematic uncertainties from the nuisance parameters. The pre- and post-fit plots can be seen in figs (5.13) and figs. (5.16). The correlation matrices and nuisance parameter plots can be seen in appendix C.

	1-prong		3-prongs	
p_{T} in GeV	SF (gluons)	SF (quarks)	SF (gluons)	SF (quarks)
20 : 30	1.58 ± 0.29	1.42 ± 0.12	1.11 ± 0.47	1.27 ± 0.13
30 : 40	0.78 ± 0.53	1.46 ± 0.15	2.58 ± 0.51	1.16 ± 0.11
> 40	0.54 ± 0.96	1.23 ± 0.13	0.52 ± 0.43	1.18 ± 0.08

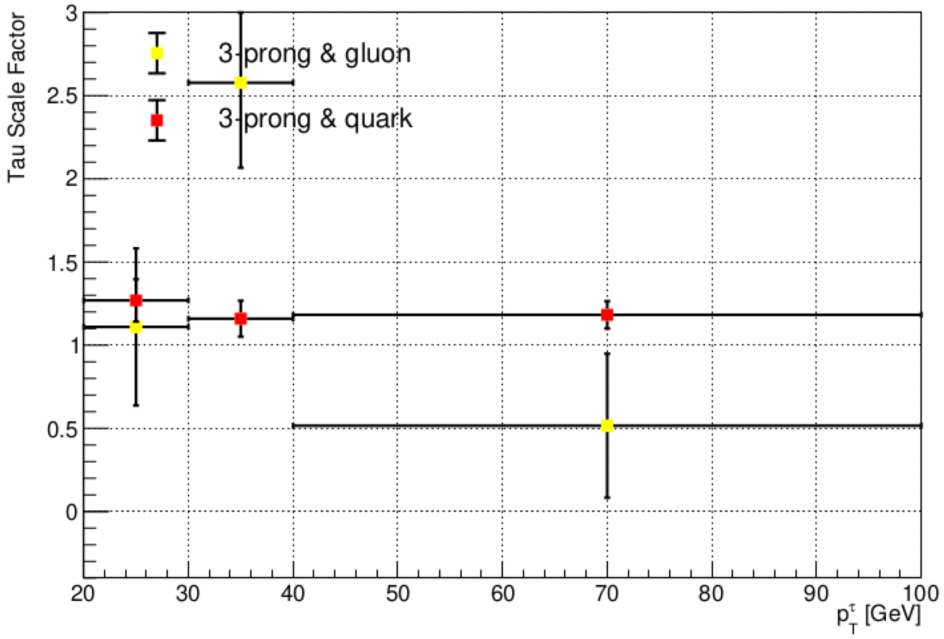
Table 5.4: SFs determined in the Quark/Gluon Method in dependence of p_{T} for 1-prong and 3-prongs.

---

[7] Abbreviated as POI.



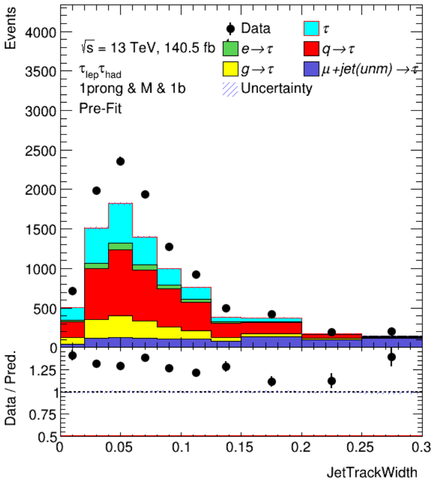
(a) 1-prong



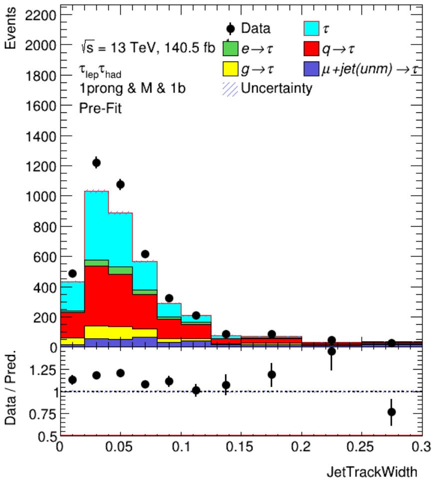
(b) 3-prong

Figure 5.12: Dependence of the SFs for different p_T for 1-prong and 3-prongs estimated in the Quark/Gluon Fit Method.

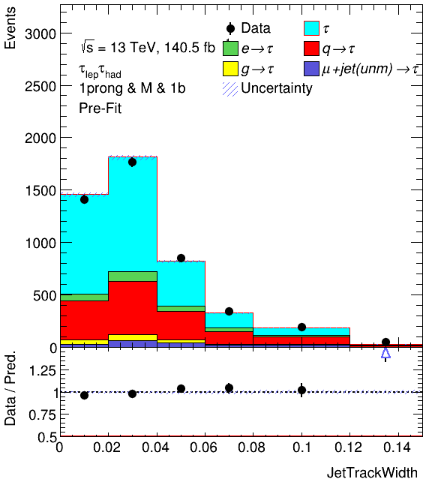
Figure 5.13: Pre- and Post-fit plots for the SR in the Quark/Gluon Method estimation of the p_T dependence of the SFs. For 1-prong.



(a) $p_T \in [20:30]$ in GeV

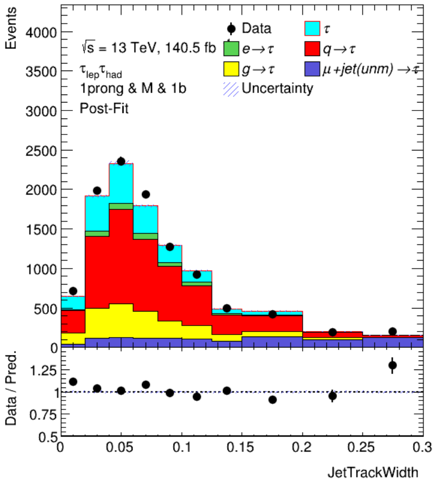


(b) $p_T \in [30:40]$ in GeV

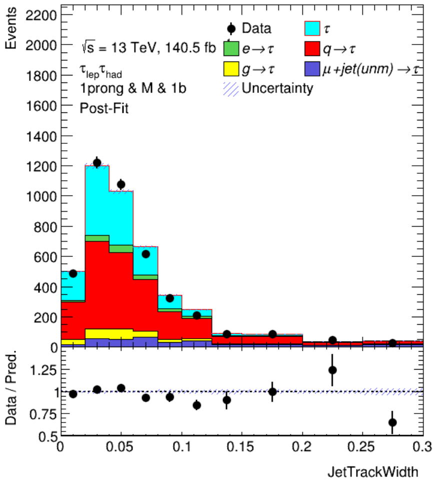


(c) $p_T > 40$ in GeV

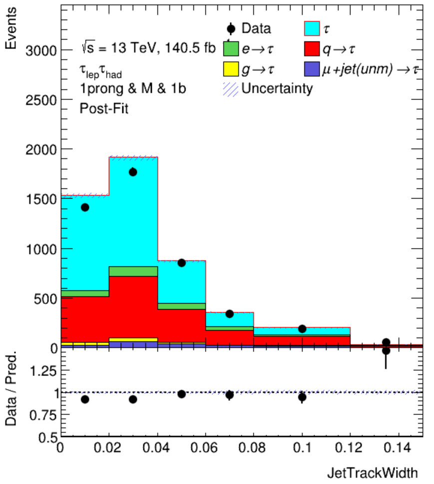
Figure 5.14: Pre-fit Plots



(a) $p_T \in [20:30]$ in GeV



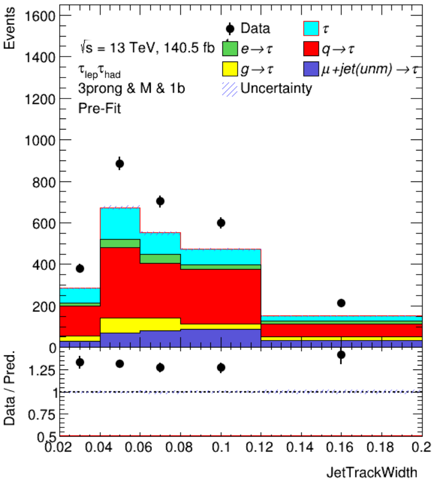
(b) $p_T \in [30:40]$ in GeV



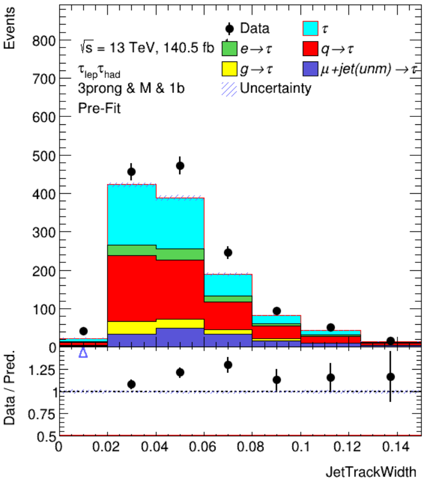
(c) $p_T > 40$ in GeV

Figure 5.15: Post-fit Plots

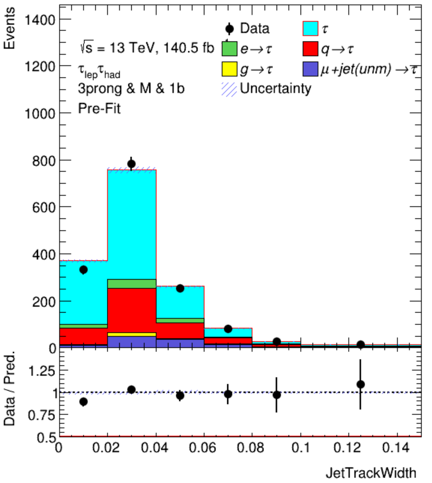
Figure 5.16: Pre- and Post-fit plots for the SR in the Quark/Gluon Method estimation of the p_T dependence of the SFs. For 3-prong.



(a) $p_T \in [20:30]$ in GeV

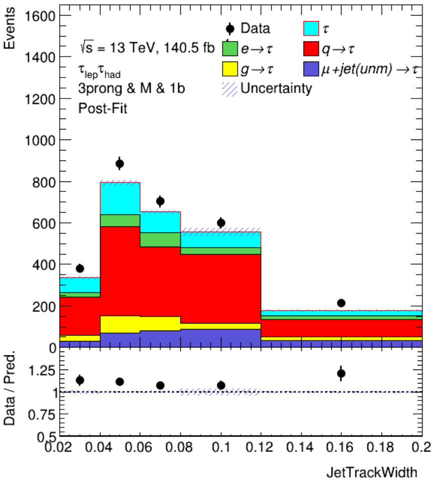


(b) $p_T \in [30:40]$ in GeV

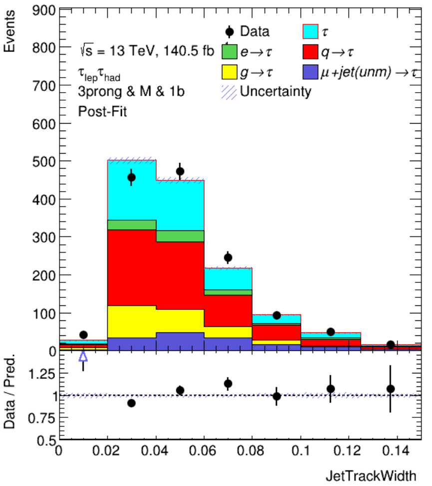


(c) $p_T > 40$ in GeV

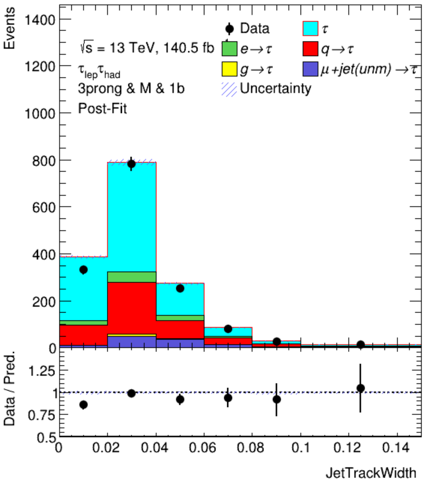
Figure 5.17: Pre-fit Plots



(a) $p_T \in [20:30]$ in GeV

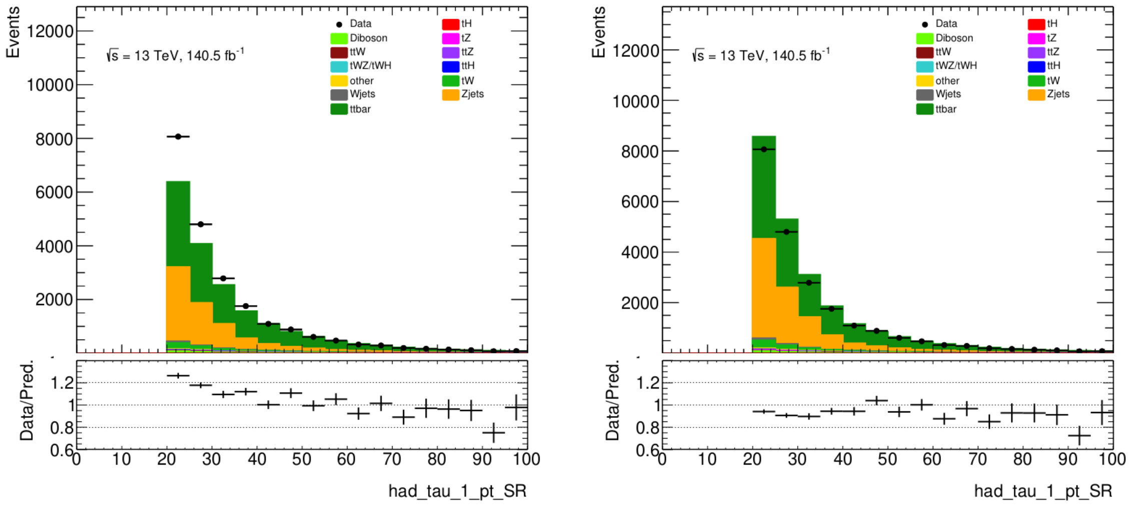


(b) $p_T \in [30:40]$ in GeV



(c) $p_T > 40$ in GeV

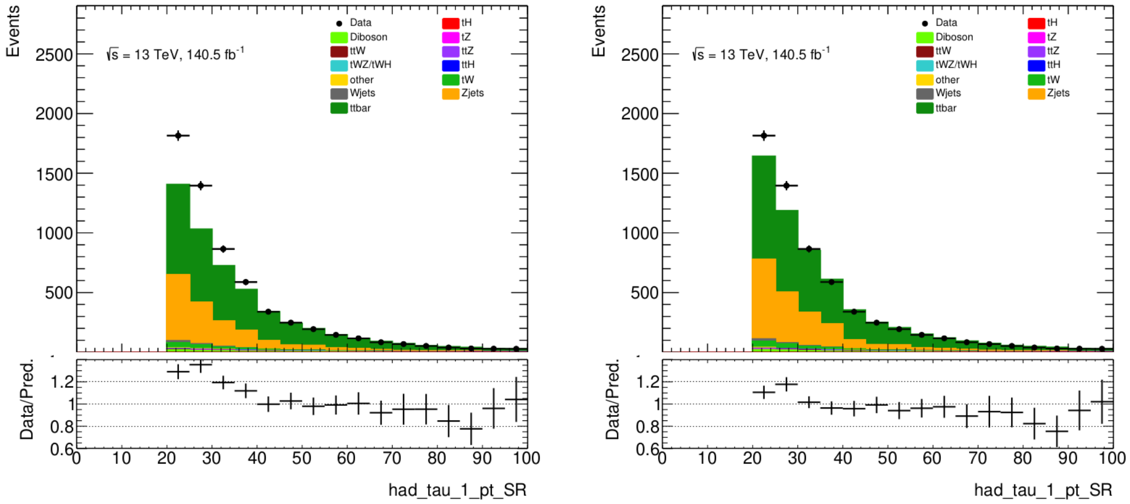
Figure 5.18: Post-fit Plots



(a) Pre-Fit

(b) Post-Fit: SFs from the Quark/Gluon Method.

Figure 5.19: Full p_{T} plot in the SR for 1-prong.



(a) Pre-Fit

(b) Post-Fit: SFs from the Quark/Gluon Method.

Figure 5.20: Full p_{T} plot in the SR for 3-prong.

5.4 Scale Factors in the Alternative Region

In this section, the SFs for the alternative region discussed in chapter [4] are presented. This region is added to determine SFs specifically to constrict the $Z+\text{jets}$ process. As discussed in chapter [4] the $0b$ jet region was deemed ill suited for this. The fits did not converge properly and did not seem stable. Additionally, the SFs for gluons were determined to be unphysically high. In comparison, the now presented alternative region is devoid of these problems. For brevity, we restrict ourselves to present only the SFs. They can be seen in table (5.5) for the 1-Bin method. A visual representation can be seen in fig. (5.21). The SFs for the Quark/Gluon Method can be seen in table (5.6). A visualization can be seen in fig. (5.22). Since we did not observe an $|\eta|$ dependence in the *normal* $1b$ region, we do not expect a dependence in this region. The same arguments as before apply. The $|\eta|$ dimensionality needs to be dropped to acquire enough statistics for the fits. The errors are total uncertainties for the same aforementioned setup.

	1-prong	3-prongs
p_T in GeV	SF (had)	SF (had)
20 : 30	1.24 ± 0.01	1.13 ± 0.01
30 : 40	1.27 ± 0.02	1.12 ± 0.02
> 40	1.11 ± 0.02	1.03 ± 0.02

Table 5.5: SFs determined in the 1-Bin Method in dependence of p_T for 1-prong and 3-prongs. For the *alternative* region.

	1-prong		3-prongs	
p_T in GeV	SF (gluons)	SF (quarks)	SF (gluons)	SF (quarks)
20 : 30	2.51 ± 0.28	0.92 ± 0.07	0.98 ± 0.26	1.16 ± 0.05
30 : 40	1.57 ± 0.39	1.24 ± 0.04	0.57 ± 0.36	1.16 ± 0.11
> 40	2.77 ± 0.47	0.94 ± 0.05	2.20 ± 0.68	0.94 ± 0.07

Table 5.6: SFs determined in the Quark/Gluon Method in dependence of p_T for 1-prong and 3-prongs. For the *alternative* region.

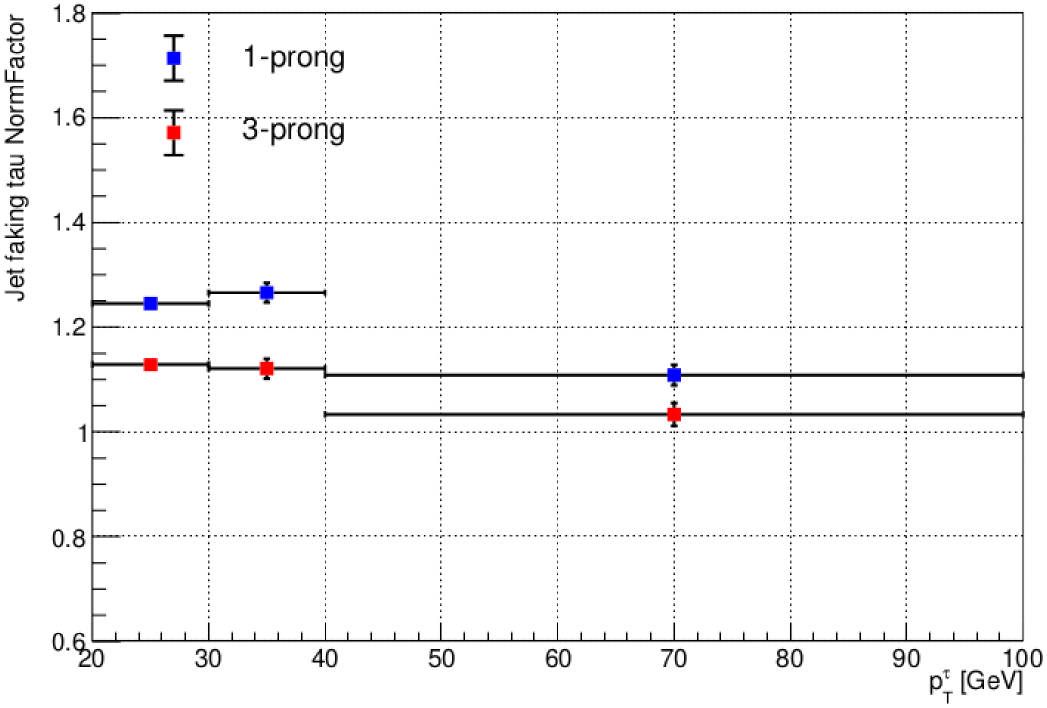
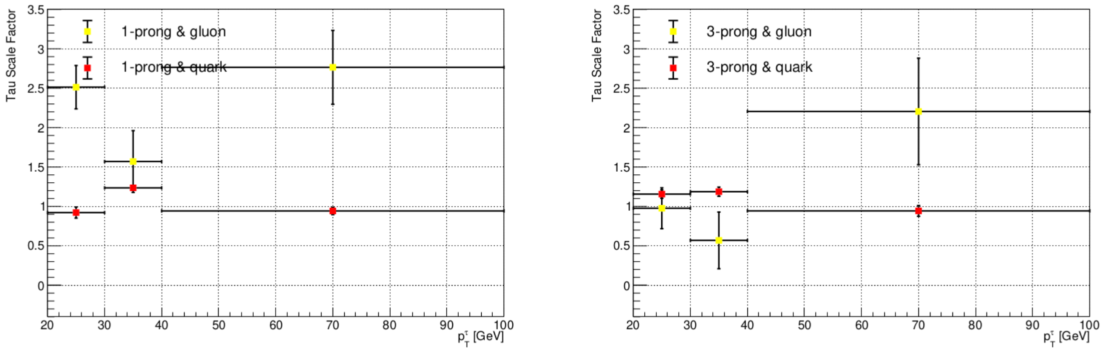


Figure 5.21: Dependence of the SFs for different p_T for 1-prong and 3-prongs estimated in the 1-Bin Method for the *alternative* region.



(a) 1-prong

(b) 3-prong

Figure 5.22: Dependence of the SFs for different p_T for 1-prong and 3-prongs estimated in the Quark/Gluon Fit Method for the *alternative* region.

5.5 Discussion of Results and Comparison between the Methods

In this section we discuss the results for the SFs obtained and presented so far. A comparison between these methods and preliminary predictions from the FTTF will be presented in the next section.

The first thing to notice is a general improvement in the MC to data ratio in the SR for both methods. The SFs become smaller for higher transverse momenta. Correspondingly, the pre-fit MC to data ratio in the highest $p_\mathrm{T} > 40$ GeV bin is similar in quality to that of the post-fit. The general trend of smaller SFs for higher transverse momenta is expected. One reason is that the differences in the profiles for QCD jets and hadronically decaying taus becomes bigger for higher p_T. As we saw in fig. (3.7) QCD jets are wider than tau jets. This difference in width increases for higher transverse momenta. Therefore, the MC to data ratio is already good at these bins. For the Quark/Gluon Method a similar trend can be observed. It is important to note that for very low gluon contributions, a relatively high SF does not automatically mean high gluon contributions after the corrections. Rather this implies that the modeling of these contributions would need significant corrections. There are isolated instances where this can be seen.

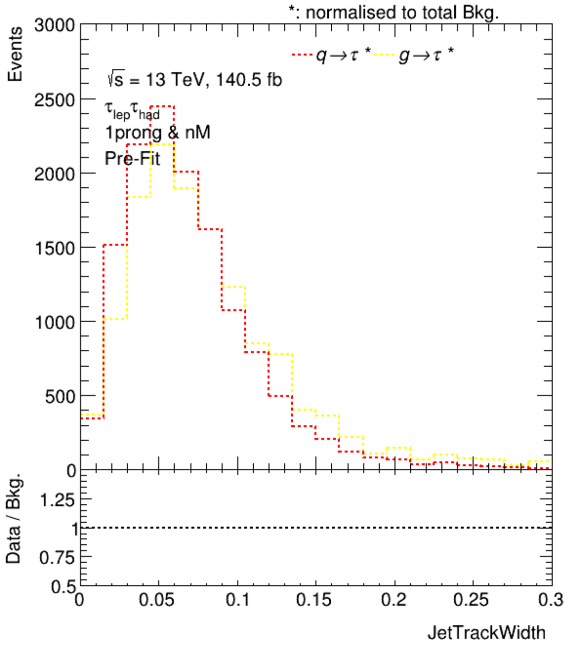


Figure 5.23: Normalized templates for the gluon faking taus (yellow) and quarks faking taus (red) contributions.

In general the SFs for the gluon and quark templates are highly correlated. This can be seen explicitly in the correlation matrices figs. (C.4). The correlation[8] between both SFs is usually about 90%. This is because the discriminative power of the variable used is limited. An example of the shape of the variable predicted by our MC simulation can be seen in fig. (5.23). This is a plot for $20 < p_\mathrm{T} < 30$ and 1-prong in the $1b$ jet *loose* region. We expect the shape distributions to look like fig. (2.17). Comparing the expected shapes and the shape from our MC simulation in that given cut of p_T and prongs, we can see that the difference in the quark and gluon templates is quite small. This varies for different cuts of p_T and number of prongs. We discussed already in chapter [4] that there is always going to be a correlation between the templates of the sources, even when assuming a scenario of maximum discrimination. In our case the templates are quite similar. We can, however, see that the gluon template is slightly shifted to higher values of track width. This is what we would expect due to the higher value of C_A compared to C_F. In addition to this, the separation between the simulated templates in fig. (2.17) is not much bigger. This is true especially for higher p_T. Furthermore, the simulation shows the variable for light quark and gluon jets. We include all quarks up to b quarks in the quark template. Because the biggest difference between quarks and gluons is expected to be between light quarks and gluons, the inclusion of heavier charm and bottom quarks smears out this

---

[8] Technically it is an anticorrelation between the SFs.

separation. Since the likelihood fit is able to differentiate the template enough to find a convergence for the fit, it is reasonable to assume that this problem manifests itself in the resulting uncertainties[9] and their correlation. Meaning as long as the shapes of the templates behave properly, the results should represent the physical reality. A significant MC to data improvement supports this statement. The correlation of the uncertainties for both templates needs to be decorrelated though. This will be discussed in the following subsection. It is also worthwhile to emphasize that correlations in this method arise naturally from the use of only two sources to fit. Intuitevely, it is sensible to assume that if one template goes up the other one has to go down in order to maximize the likelihood.[10] Additionally, the possibilities to differentiate the two sources are limited.[11] Other variables, like the energy profile, led to less promising results. At the heart of this method lies the confidence in the modeling of the template shapes. Unfortunately, it is not possible to confirm absolutely, whether they are modeled well or not. We can, however, look at the behavior expected from physics for the different variables and an improvement in the MC to data ratio. The uncertainty from this knowledge is incorporated into the uncertainties of our SFs due to the mathematical machine that is the binned likelihood fit. We can also compare the predictions from this method with the predictions from other methods like the 1-Bin Method or the FTTF and make conclusions about the quality of the shapes this way. It is, however, necessary to take a closer look into the contributions of quarks and gluons for the jets that fake taus since we cannot assume that the ratio of quark to gluon jets in the CR and SR is the same. We try to mitigate this by looking at similar topologies, but ultimately a discrimination between quark and gluon contributions is desirable due to their different jet properties which result in different tau faking behavior. For this reason a dual approach of investigating both methods has been adapted. It is worthwhile to mention, that in addition to the above, the uncertainties from the Quark/Gluon Method are rather large. This is a consequence of limiting statistics and the relative bad discrimination between the templates.

Furthermore, it is interesting to take a closer look at some examples for the nuisance parameters and the correlation matrices. An example that exhibits behavior worth mentioning, is the fit for $p_T > 40$ GeV and 3-prongs. The correlation matrix and nuisance parameters can be seen in fig. (5.24(a)) and fig. (5.24(b)) respectively. In this fit the correlation between the electron nuisance and the quark fit POI is quite high. In fact for 3-prongs this correlation is usually higher than for 1-prong. This results in a relatively big deviation of the value of the electron nuisance from its expected value. Another fit in which this is the case, is the $20 < p_T < 30$ GeV and 3-prongs fit. A possible way to explain this higher correlation between electron nuisances and quark templates is to think of those electrons as possibly originating from b quark decays. In this case they will have an automatic correlation to the quark jets. It also could explain why this happens more often for 3-prongs since additional emissions from soft gluons may create tracks[12] around the electron, such that these fakes are reproduced as 3-prong decays. Another reason is that the modeling of 3-prong decays is more difficult, than the modeling of 1-prong events. This is also a reason why for 3-prongs the MC to data ratio is not necessarily better than for 1-prong events, even though more information is available. Ultimately this argument is speculative but nevertheless a trend is hinted at.

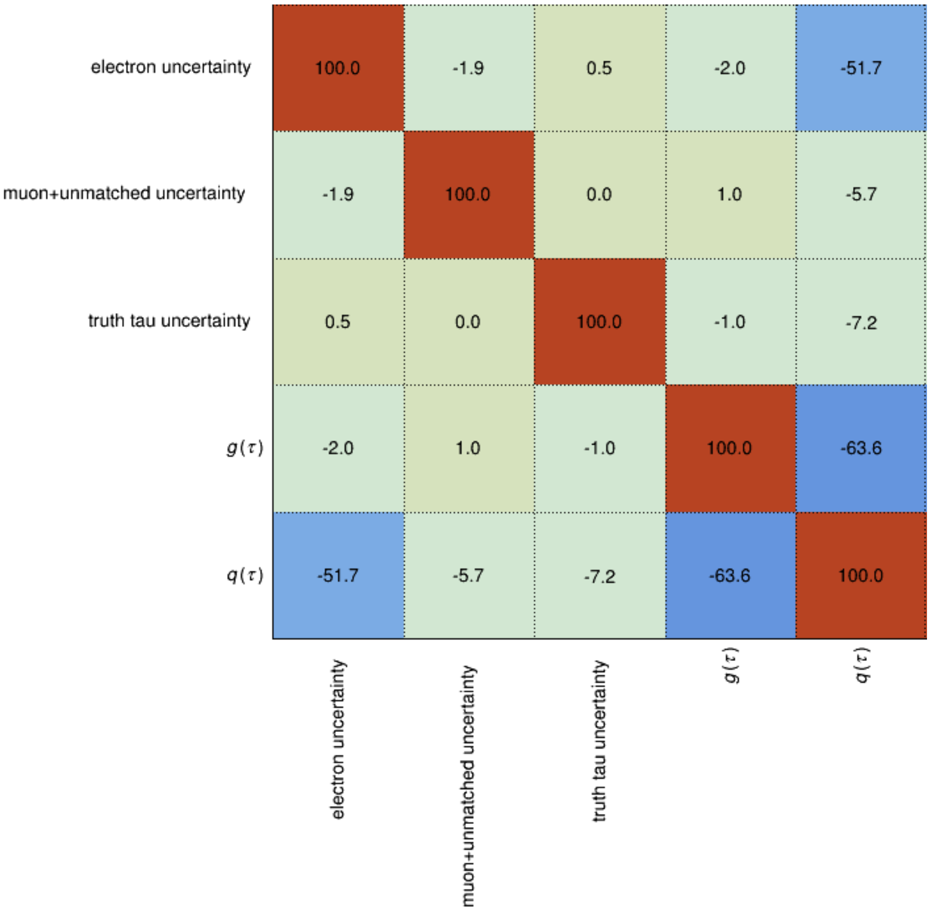
---

[9] Which are fairly high.

[10] Or minimize the logarithmic likelihood.

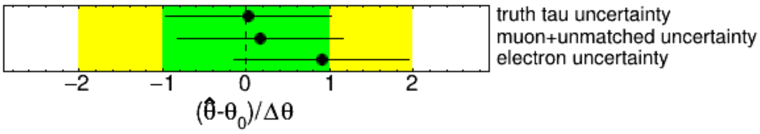
[11] In terms of technical access to other discriminative variables and a total amount of those variables.

[12] Which is likely due to the behavior of b jets.



	electron uncertainty	muon+unmatched uncertainty	truth tau uncertainty	$g(\tau)$	$q(\tau)$
electron uncertainty	100.0	-1.9	0.5	-2.0	-51.7
muon+unmatched uncertainty	-1.9	100.0	0.0	1.0	-5.7
truth tau uncertainty	0.5	0.0	100.0	-1.0	-7.2
$g(\tau)$	-2.0	1.0	-1.0	100.0	-63.6
$q(\tau)$	-51.7	-5.7	-7.2	-63.6	100.0

(a) Correlation matrix for Quark/Gluon Method Fit in the cut of $p_{\mathrm{T}} > 40$ GeV and 3-prongs.



(b) Nuisance parameters for Quark/Gluon Method Fit in the cut of $p_{\mathrm{T}} > 40$ GeV and 3-prongs.

5.5.1 Decorrelation of Errors

It is necessary to discuss the decorrelation of the uncertainties for both SFs. This will be done in this subsection. A sketch of an idea is presented to decorrelate the uncertainties.

Let us assume that the two SFs span a two dimensional space. In this space the possible configurations of results for the SFs within the bounds of their uncertainties are defined by an ellipse. We assume that the correlation matrix that describes the correlation between the SFs also describes the correlation between the uncertainties of those. The idea is to calculate the eigenvalues of the correlation matrix and apply them accordingly. The correlation matrix is a real symmetric matrix and can therefore always be diagonalized. For the case of two SFs this will quickly be done. We have a general correlation matrix

$$M = \begin{pmatrix} 1 & b \\ b & 1 \end{pmatrix} \tag{5.1}$$

and eigenvectors v_1 and v_2. We can diagonalize M by finding the solutions for its characteristic equation

$$\lambda^2 - 2\lambda + 1 - b^2 = 0 \, . \tag{5.2}$$

With λ being the eigenvalues and b the correlation. This can be solved for λ and we get

$$\lambda_{1,2} = 1 \pm \sqrt{\left(\frac{1}{4} - (1 - b^2)\right)} \, . \tag{5.3}$$

Using this we can calculate the eigenvectors that diagonalize the matrix M. They fulfill the eq.

$$(M - \lambda \mathbf{1}) \cdot \tilde{v}_i = 0 \quad \text{for} \quad i = 1, 2 \, . \tag{5.4}$$

This can be solved and we acquire the eigenvectors of the problem

$$\tilde{v}_i = \begin{pmatrix} 1 \\ \frac{\lambda_i - 1}{b} \end{pmatrix} \quad \text{for} \quad i = 1, 2 \, . \tag{5.5}$$

The uncertainty of $\tilde{v}_i$ can then be written down from the correlation b and the uncertainty from the original SFs that form the former basis. Re-expressing those original SFs in terms of the $\tilde{v}_i$ and their uncertainties should yield decorrelated uncertainties. This subsection should be regarded as a proposed solution for dealing with the correlation.

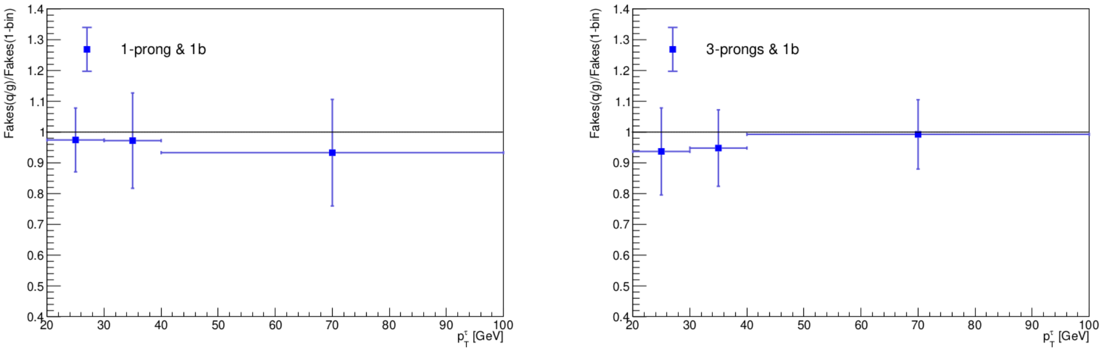
5.5.2 Comparison between the Methods

We now compare the methods and their yields. There are two issues that complicate a direct comparison between methods. The first thing is a direct consequence of the nature of the SFs themselves. They do not directly offer insight into the abundance of a given fake contribution in MC, rather they indicate deviations of MC predictions to a normalization that is closer to data and therefore[13] to the physical reality. We can therefore not compare the SFs for one source directly with the SFs for two sources. In order to allow a direct comparison we decided to calculate the number of fakes predicted by both methods that belong to the source of jets faking taus. In technical terms this means all events

---

[13] If the fitting machine was set up properly.

in MC that fulfill `had_tau_1_true_pdg=0`. This comparison approach leads to the second issue. The second issue is that the Quark/Gluon Method incorporates a combined template of muon and unmatched jet contributions. These are not properly understood and treated as a minor background in order to avoid negative behavior of this template in relation to the binning. In order to deal with this we take the muon template prediction of the 1-Bin Method and substract that from the Quark/Gluon Method muon and unmatched template. This will then provide the true unmatched contributions. Those unmatched contributions and the contributions from quarks and gluons faking taus are then added up and compared with the jet faking tau abundance. The results can be seen in fig. (5.25). A table of the ratios can be seen in table (5.7). The ratio is always close to one and one is always included in the uncertainty. The uncertainties stem from the uncertainties for the fakes which come from the MC uncertainties as well as the total uncertainties of the SFs. They have been propagated to the ratio using error propagation.The plots are results from the application of the SFs to the $1b$ SR.



(a) 1-prong

(b) 3-prongs

Figure 5.25: Ratio Fakes(Quark/Gluon Method) over Fakes(1-Bin Method) in the SR for different p_T and number of prongs.

We can clearly see that both methods predict similar amounts of fakes. The value of one is always inside the uncertainty of the ratio. This can be seen as an indication that the templates for the quarks and gluons are reasonable well modeled. In fact the MC to data improvement of both methods are similar. This can also be understood as a consequence of fitting over the same process topology in both methods. It also indicates that the ratio of quark to gluon jets in the SR is close to that of the CR. Otherwise differences between the methods would have been observed. It is necessary to point out that the similarity between the predictions for fake abundances in the SR for both methods can only be assumed in terms of the bounds set by the uncertainty of the result. The possibility for differences are there but they may be smaller then our predictive power which is limited by our uncertainty.[14] It is necessary to mention that both methods are not completely independent from each other. They both use the same CR[15] and therefore are not fully independent. Nevertheless the results from table (5.7) indicate to well modeled quark and gluon templates.

---

[14] Which mainly comes from the Quark/Gluon uncertainty.

[15] Even though for the Quark/Gluon Method the $2b$ region is included.

	1-prong	3-prongs
p_T in GeV	Fakes(Quark/Gluon Method)/Fakes(1-Bin Method)	Fakes(Quark/Gluon Method)/Fakes(1-Bin Method)
20 : 30	0.97 ± 0.10	0.94 ± 0.14
30 : 40	0.97 ± 0.15	0.95 ± 0.12
> 40	0.93 ± 0.17	0.99 ± 0.11

Table 5.7: Ratio Fakes(Quark/Gluon Method) over Fakes(1-Bin Method) in the SR for different p_T and number of prongs.

5.6 Fake Abundances from Fake Tau Task Force Fake Factors

Unfortunately, we are limited in the comparisons as a crosscheck for our method. So far, both methods agree for the total fake abundances within the uncertainty and both methods improve MC to data rather well. There is, however, another crosscheck that is possible to take a look at. In this subsection we are going to compare our results with the total fake abundance predictions from the FTTF for our SR.[16] For comparison with our analysis, we received the fake factors[17] from the Fake Tau Task Force[18].[19] The FF of the FTTF is defined as follows

$$FF = \frac{N_{\text{pass ID}} - N_{\text{pass ID, true } \tau}}{N_{\text{fail ID}} - N_{\text{fail ID, true } \tau}} \,. \tag{5.6}$$

Here $N_{\text{fail ID}}$ are the number in some region where the identification has been failed. Therefore $N_{\text{pass ID}}$ is defined accordingly. In our case this will be the SR and CR with different working points for tau identification. The fake factors have been measured for quark and gluon jets faking taus for different numbers of prongs, decay modes and p_T. This was done using either $Z+$jets events for quarks or multijets events for gluons. The overall idea was mentioned briefly in chapter [4]. Instead of looking for FF's for the different sources in an analysis topology they look for the fake contributions.[20] Due to this approach, high statistics can be achieved, which lead to small uncertainties in the estimation. Another advantage is that other when substracting the truth tau contributions from MC, which can be expected to be rather well modeled, this method is data driven and does not rely on the fake modeling of MC simulations. A downside is that the topology of the SR in which these FF's are applied is left out completely. If the initial assumption of process independents should prove wrong[21] or the samples of data used in determining the FF's is not clean enough in its quark and gluon purities, the method could lead to bad modeling of tau fakes. Problems could also arise when the processes, in the region where the FFs are applied, are much different than multijets and $Z+$jets events.

---

[16] Or rather the fake abundances acquired after applying the FTTF FF to our region.

[17] In short FF.

[18] In short FTTF.

[19] This was at the end of May 2021.

[20] similar to the template definitions in this thesis.

[21] Even though this assumption seems sensible at first.

In order to calculate the estimated tau fakes from this method we can use

$$N_{\text{SR pass ID}} = N_{\text{SR fail ID}} \cdot FF \; . \tag{5.7}$$

This can be done for the quark and gluon fakes for different decay modes. A list of the FF's can be seen in appendix E. The decay modes are defined as

- $dm_0 = 1p0n$,
- $dm_1 = 1p1n$,
- $dm_2 = 1pXn$,
- $dm_3 = 3p0n$,
- $dm_4 = 3pXn$.

Where $1p$ denotes 1-prong and $0n$ denotes one neutral pion in the decay. In order to acquire the fake estimation from the FF's we apply the FF on the data in the CR similar to eq. (5.7). In order to correct for the truth tau contamination[22] we substract the MC truth tau value from the data yields. Since the FTTF splits the transverse momenta for $p_\text{T} > 40$ GeV into additional bins, we sum these contributions up to allow a fair comparison. The same is done for the different decay modes.

5.6.1 Comparison to the Quark/Gluon Fit Method

We now compare the estimated fake abundances from the Quark/Gluon Fit Method and the FTTF Fake Factor method. The results can be seen in table (5.8) for 1-prong and table (5.9) for 3-prongs. The predictions of both methods do not agree. A few trends can be seen. The first thing to mention is that the Quark/Gluon Method predicts more quarks and less gluons compared to the FTTF method in the 1-prong case. The total number of fakes are not similar, but, due to the large uncertainties of the Quark/Gluon Method, they are sometimes within the uncertainties of each other. Both methods predict a lesser amount of fakes for declining transverse momenta, which is also related to to the decrease in events for smaller p_T. For the 3-prongs estimation, the number of fakes are vastly more different between both methods. What is especially striking is the small number of total fakes[23] predicted by the FTTF results. This is in stark contrast to larger amounts of fakes in our analysis. In addition to that, the uncertainties in our analysis are fairly large, while the uncertainties for the FTTF predictions are rather small. This is due to larger possible statistics that are made possible by their approach.

There are a number of things worth mentioning. The first one is that the FTTF results could indicate vast mismodeling of our MC simulation. In the case of 3-prongs the FTTF results would not improve but rather make the MC to data ratio worse. Our analysis improves this ratio. It is necessary to mention though that an improvement between MC and data comes more naturally from the way this method functions. Even though the results of the FTTF method[24] shows vast discrepancies between data and their predictions, it is left to be asked which process[25] should be extended to fill this discrepancy. The

---

[22] In the case for the fake tau abundance estimation the truth taus can be considered a contamination.

[23] Gluon fakes added to quark fakes.

[24] As presented in this thesis.

[25] In the sense of our templates, meaning truth taus other partons faking taus and similar.

assumption of well modeled truth tau abundances should not be forfeited. In addition to that the truth tau modeling was used to correct the results according to eq. (5.7) in the FTTF method itself. By questioning the results from the Quark/Gluon Fit Method one also questions in extension the results from the 1-Bin Fit Method due to their agreement within uncertainties as seen in fig. (5.25). Even though both methods are not fully independent[26] this leads one to assume that, if the results of this thesis are questioned, rather than in the procedure, the problem would lie in the MC simulations[27] themselves. The similarities in their predictions indicates that the modeling of the shapes in the Quark/Gluon method is not bad.

Another possibility for the differences is of course an error prone application of the FTTF methods to our region.[28] Of course it is not a given that the preliminary results from the FTTF are applicable to our region in the first place. The SR we investigate is dominated by $t\bar{t}$ events. We have a large amounts of b jets in the SR. This is quite different from the processes the FTTF used to measure their fake factors. This difference in topology could be a possible reason for why their FFs could not succesfully be applied to our SR. It might be worthwhile to investigate this point further.

The vast discrepancy of the estimations, using the FTTF fakes factors, to the estimations, using our methods, remains an issue that should be investigated. For now a definitive reason for this difference cannot be given. One major point of uncertainty both, deviating results, rely on, is whether the MC simulation and data for this region as a whole[29] is reasonably well modeled. Or to a lesser degree how badly modeled we can assume the MC simulation to be.[30] The FTTF methods predictions would mean that the MC simulation is very badly modeled. Additionally, if we assume that the modeling of the truth tau template is good, than there needs to be some other influence that explains the difference between data and the expected fakes from the FTTF and the truth tau abundance. The other major point of uncertainty is, whether the application of the preliminary FFs to our SR is reasonable to begin with.[31]

	FTTF		Quark/Gluon Fit Method	
p_T in GeV	Number of quark Fakes	Number of gluon Fakes	Number of quark Fakes	Number of quark Fakes
20 : 30	3 461 ± 23	2 681 ± 73	5 032 ± 428	1 889 ± 346
30 : 40	1 169 ± 16	945 ± 48	2 132 ± 213	256 ± 173
> 40	768 ± 24	668 ± 66	1 669 ± 172	83 ± 147

Table 5.8: Yields estimated for quark and gluon fakes in the FTTF method and the Quark/Gluon Fit Method for 1-prong.

---

[26] As discussed before. They rely on the same CRs and MC simulation.

[27] As used in form of tHqLoop postprocessed samples.

[28] Even though, as of writing this, no errors have been found.

[29] Not just the shape modeling.

[30] Since the point of our analysis is to correct mismodeling in MC.

[31] Because of the large number of b jets and $t\bar{t}$ events in our SR compared to the lack thereof for the FF measurement of the FTTF.

	FTTF		Quark/Gluon Fit Method	
p_T in GeV	Number of quark Fakes	Number of gluon Fakes	Number of quark Fakes	Number of quark Fakes
20 : 30	298 ± 3	247 ± 11	1 360 ± 136	228 ± 97
30 : 40	107 ± 2	82 ± 7	532 ± 50	199 ± 39
> 40	56 ± 2	48 ± 6	424 ± 30	17 ± 14

Table 5.9: Yields estimated for quark and gluon fakes in the FTTF method and the Quark/Gluon Fit Method for 3-prong.

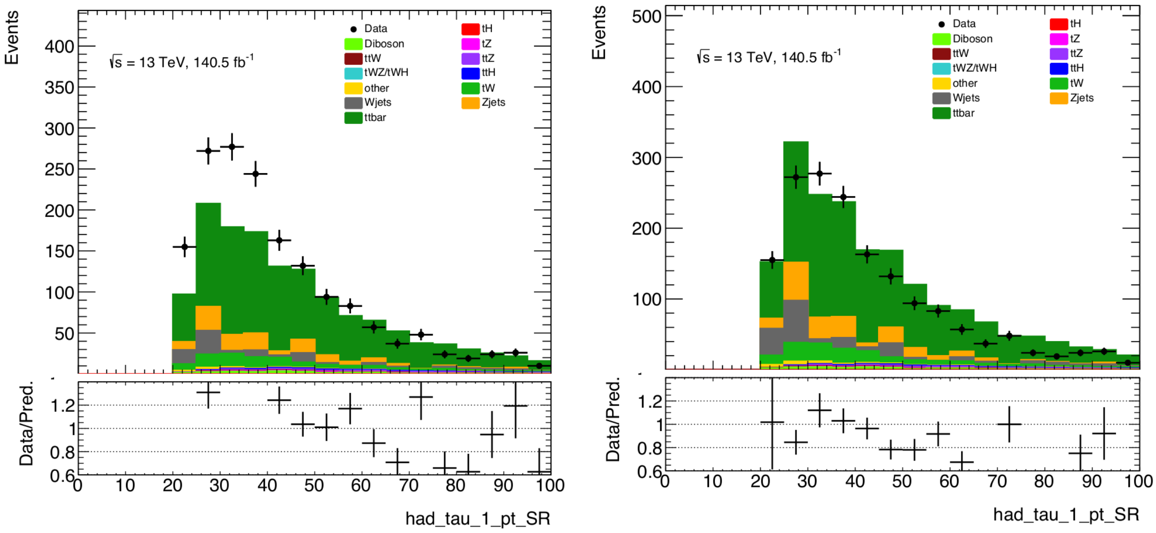
5.7 Application of the Scale Factors to the Had-Had Channel

In this section we apply the SFs from the lep-had channel for both methods to the had-had channel and look at the possible improvement of the MC to data ratio. Determining tau fakes in the had-had channel comes with a set of additional challenges. For once the number of τ leptons per event are doubled. So we need to differentiate between two tau leptons and estimate the fakes for both. We can sort these hadronically decaying taus into a leading tau and a sub-leading tau based on their transverse momentum. Here the tau lepton with the larger momentum is the leading tau. Different circumstances can arise. For instance, it can be the case that the leading tau is a fake and the sub-leading tau is not. The opposite is also possible as well as any combinations of fakes and not fakes for both taus. This gives four distinct cases which need to be considered. In addition to this, the statistics in the had-had region is much smaller than in the lep-had region. A possible reason for this is that there are fewer light leptons which can trigger an event. This makes an approach for the split between quarks and gluons very difficult. Additionally, there might be a correlation between the SFs from both taus as well.

In order to deal with these issues, an approach is investigated that at the current state of analysis does not differentiate between the number of prongs and tries to fit for all four combinations the distributions of leading and sub-leading taus in the 1-Bin Fit Method.

In this thesis, we want to determine if an improvement in MC to data ratio can be observed when we apply our SFs to the had-had region for both taus. In fig. (5.26) the pre-fit distribution is compared with the distribution after applying the SFs from the 1-Bin Method to the leading tau of the had-had SR. The same is done for the sub-leading tau. This result can be seen in fig. (5.27). We also applied the SFs from the Quark/Gluon Fit method to the aforementioned disctributions. The results can be seen in fig. (5.28) for the leading tau and in fig. (5.29) for the sub-leading tau respectively

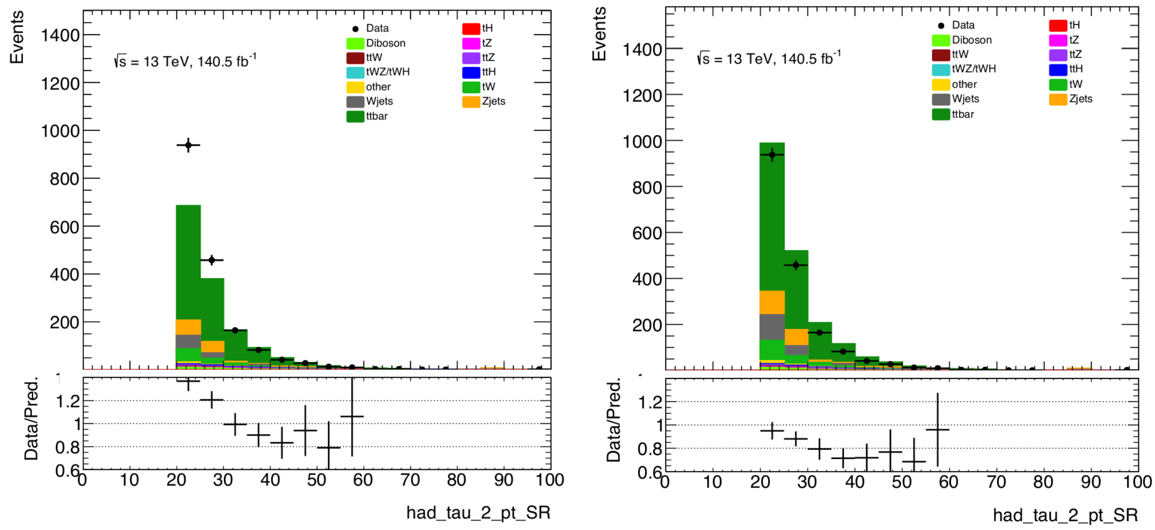
We observe a MC to data improvement for both methods in the had-had SR. This is true for the leading and the sub-leading tau as well. Further investigation into this area has to be performed but these preliminary applications hint at a similarity in the SFs for the lep-had channel to those of the had-had channel.



(a) Pre-Fit

(b) Post-Fit: SFs from the 1-Bin Method.

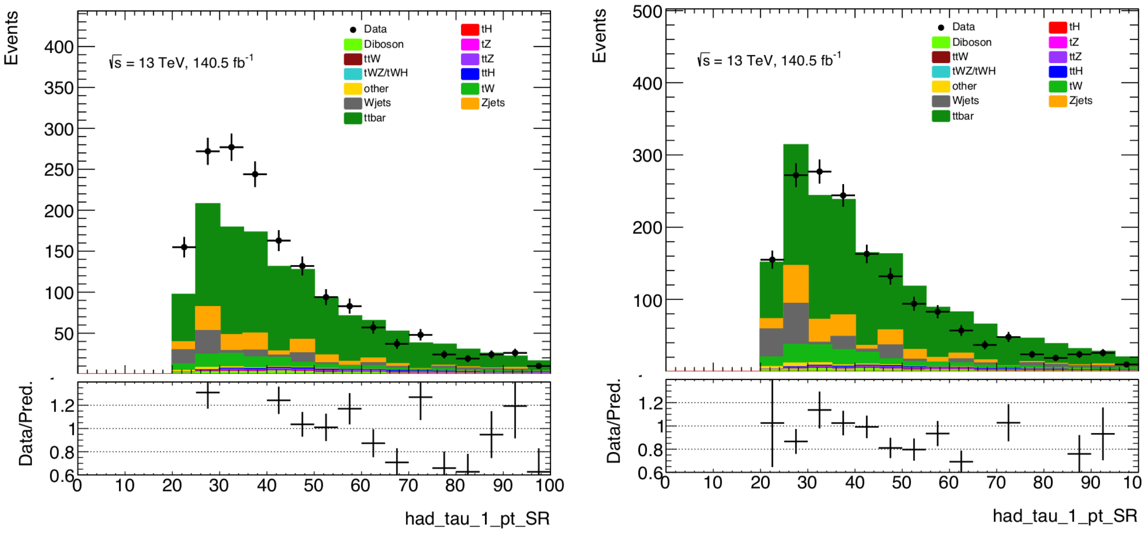
Figure 5.26: Full p_{T} plot in the SR for the leading tau lepton.



(a) Pre-Fit

(b) Post-Fit: SFs from the 1-Bin Method.

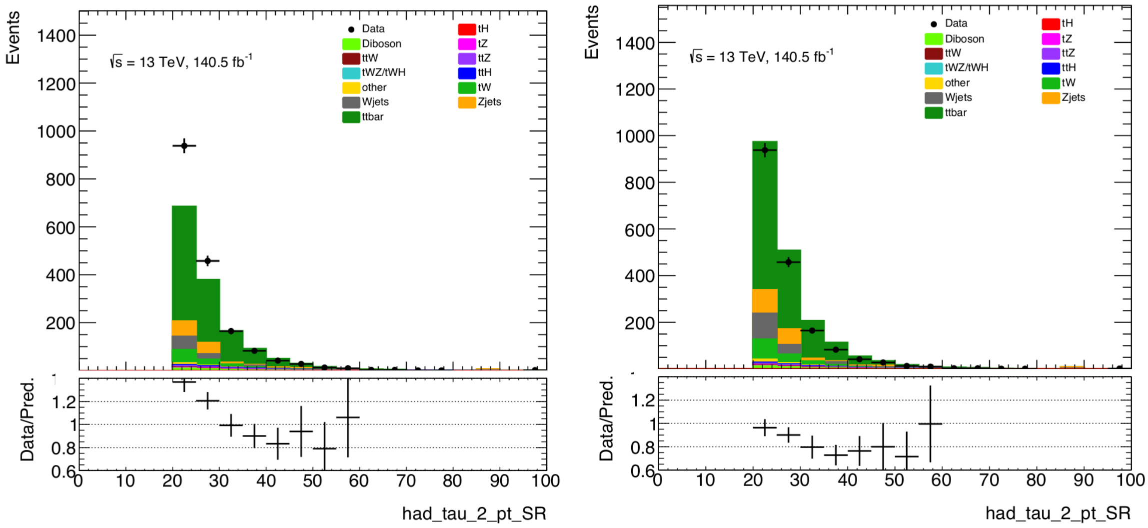
Figure 5.27: Full p_{T} plot in the SR for the sub-leading tau lepton.



(a) Pre-Fit

(b) Post-Fit: SFs from the Quark/Gluon Fit Method.

Figure 5.28: Full p_T plot in the SR for the leading tau lepton.



(a) Pre-Fit

(b) Post-Fit: SFs from the Quark/Gluon Fit Method.

Figure 5.29: Full p_T plot in the SR for the sub-leading tau lepton.

CHAPTER 6

Conclusion

In this thesis two methods based on the principle of binned likelihood fits were presented. The estimation of the fake tau abundances from both methods have been discussed and compared. In addition to this, a comparison with the fake factor method from the Fake Tau Task Force was done. Furthermore, the scale factors for both methods were applied to the had-had channel as well. For both methods the scale factors were determined in dependence of the number of prongs and their transverse momentum. A dependence on $|\eta|$ was excluded. For the estimation the control and signal regions were investigated for their process topologies.

The scale factors determined in the 1-Bin Method improve the Monte Carlo to data discrepancy for the signal region in the lep-had channel. An improvement in Monte Carlo to data ratio can also be observed in the had-had channel. The scale factors determined in the Quark/Gluon Method improve the data to Monte Carlo ratio in the signal region as well. An improvement in the had-had channel using these scale factors can also be seen. The total uncertainties in the Quark/Gluon Fit Method are larger than those of the 1-Bin Method. This lies within expectations and is the consequence of binning the Monte Carlo simulation and therefore reducing the statistics per bin. Additionally the Quark/Gluon Fit Method performs a fit over two sources. These sources need to discriminated against in the fit, in order to reduce the correlation of their scale factors. For this, a quark gluon jet discriminant, based on the p_{T} weighted track width, was used. The discriminative power of this variable is limited. This leads to large correlations between the uncertainties. An idea to decorrelate between these uncertainties was proposed. Due to the large correlation, the goodness of the modeling of the quark and gluon templates concerning their jet track width needs to investigated. For physical predictions when performing likelihood estimations a good modeling of the shape from the Monte Carlo simulation is necessary. In order to ascertain the validity of the Quark/Gluon Methods estimation, a variety of benchmarks have been discussed. For this an improvement in Monte Carlo to data ratio is vital. A comparison with the 1-Bin Method of the total number of estimated fakes has been performed as well. This resulted in both methods agreeing within the given uncertainties in the signal region. This leads one to conclude that the modeling of the jet track widths of gluons and quarks is well and that the ratio of quarks to gluons between control region and signal region is similar within the uncertainties. Otherwise both methods would not agree in the signal region. The comparison with the estimates from application of the Fake Tau Task Force fake factors to our data resulted in large differences. There is a vast disagreement between the Quark/Gluon Method predictions and the predictions from the application of the fake factors. There is a possibility that the application of the fake factors to our signal region is too naive, since their predictions worsen the observed Monte Carlo to data ratio. Therefore we have to conclude that either our Monte Carlo simulation is vastly wrong or that the application of the fake factors to our analysis does not work. This can be regarded as an important issue worthy of further investigation.

Bibliography

[1] Particle Data Group, K. Nakamura et al., *Review of Particle Physics*, J. Phys. G **37** (2010) 075021, URL: `http://pdg.lbl.gov`.

[2] I. C. Brock, *Users Guide to Writing a Thesis in a Physics/Astronomy Institute of the Universität Bonn*, URL: `http://pi.physik.uni-bonn.de/pi_only/thesis.php`.

[3] F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*, Wiley, 1984, ISBN: 9780471887416 (cit. on pp. 11, 14).

[4] ATLAS Collaboration, *Expected Performance of the ATLAS Experiment – Detector, Trigger and Physics*, 2009, arXiv: `0901.0512`.

[5] E. M. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, 1st ed., CRC Press, 2018 (cit. on pp. 3–6, 8–12, 14–17, 25).

[6] D. M. Schwartz, *Quantum Field Theory and the Standard Model*, 1st ed., Cambridge University Press, 2014 (cit. on pp. 3, 6, 8–10, 12, 13, 15).

[7] L. Landau and E. Lifschitz, *Mechanics*, 3rd ed., Pergamon Press, 1976 (cit. on pp. 4, 5).

[8] L. Landau and E. Lifschitz, *The Classical Theory of Fields*, 4th ed., Butterworth-Heinemann, 1975 (cit. on p. 5).

[9] F. Wilson, *Fermi's theory of beta decay*, American Journal of Physics **36** (1968) (cit. on p. 7).

[10] A. Salam, “Weak and Electromagnetic Interactions”, *Elementary particle theory. Relativistic groups and analyticity. Proceedings of the Eighth Nobel Symposium*, ed. by N. Svartholm, Almquist & Wiksell, 1968 367 (cit. on p. 7).

[11] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264 (cit. on p. 7).

[12] S. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579 (cit. on p. 7).

[13] G. Arnison et al., *Experimental observation of isolated large transverse energy electrons with associated missing energy at s=540 GeV*, Physics Letters B **122** (1983) 103, ISSN: 0370-2693, URL: `https://www.sciencedirect.com/science/article/pii/0370269383911772` (cit. on p. 7).

[14] P. Bagnaia et al., *Evidence for* $Z^0 \to e^+e^-$ *at the CERN* $\bar{p}p$ *Collider*, Phys. Lett. B **129** (1983) 130 (cit. on p. 7).

[15] J. Goldstone, *Field theories with "superconductor" solutions*, Nuovo Cimento **19** (1960) 154, URL: `http://cds.cern.ch/record/343400` (cit. on p. 9).

[16] Y. Nambu, *Quasi-Particles and Gauge Invariance in the Theory of Superconductivity*, Phys. Rev. **117** (3 1960) 648, URL: `https://link.aps.org/doi/10.1103/PhysRev.117.648` (cit. on p. 9).

[17] P. Zyla et al., *Review of Particle Physics*, PTEP **2020** (2020) 083C01 (cit. on pp. 11, 18, 19, 21, 23, 46, 62).

[18] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (12 1963) 531, URL: `https://link.aps.org/doi/10.1103/PhysRevLett.10.531` (cit. on p. 13).

[19] Y. Fukuda et al., *Evidence for Oscillation of Atmospheric Neutrinos*, Physical Review Letters **81** (1998) 1562, ISSN: 1079-7114, URL: `http://dx.doi.org/10.1103/PhysRevLett.81.1562` (cit. on p. 13).

[20] T.-M. Yan and S. D. Drell, *The parton model and its applications*, International Journal of Modern Physics A **29** (2014) 1430071, ISSN: 1793-656X, URL: `http://dx.doi.org/10.1142/S0217751X14300713` (cit. on p. 14).

[21] K. R. J. Husto and G. Zanderighi, *Quantum Chromodynamics pdg Review*, URL: `https://pdg.lbl.gov/2020/reviews/rpp2020-rev-qcd.pdf` (cit. on p. 14).

[22] M. Gell-Mann, *Symmetries of Baryons and Mesons*, Phys. Rev. **125** (3 1962) 1067, URL: `https://link.aps.org/doi/10.1103/PhysRev.125.1067` (cit. on p. 15).

[23] H. Lehmann, K. Symanzik and W. Zimmermann, *Zur Formulierung quantisierter Feldtheorien*, Nuovo Cimento Serie **1** (1955) 205 (cit. on p. 16).

[24] F. Abe et al., *Observation of Top Quark Production in $\overline{p}p$ Collisions with the Collider Detector at Fermilab*, Phys. Rev. Lett. **74** (14 1995) 2626, URL: `https://link.aps.org/doi/10.1103/PhysRevLett.74.2626` (cit. on p. 18).

[25] *Updated coupling measurements of the Higgs boson with the ATLAS detector using up to 25 fb^{-1} of proton-proton collision data*, tech. rep., All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2014-009: CERN, 2014, URL: `https://cds.cern.ch/record/1670012` (cit. on p. 19).

[26] M. Farina et al., *Lifting degeneracies in Higgs couplings using single top production in association with a Higgs boson*, Journal of High Energy Physics **2013** (2013), ISSN: 1029-8479, URL: `http://dx.doi.org/10.1007/JHEP05(2013)022` (cit. on p. 19).

[27] H. N. We, *Associated Production of a Top Quark and a Higgs Boson in pp Collisions at 13 TeV Using the ATLAS Detector*, MA thesis: University of Bonn, 2020 (cit. on p. 20).

[28] T. A. Collaboration et al., *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, 2009, arXiv: `0901.0512 [hep-ex]` (cit. on p. 22).

[29] A. Tropiano, *Measurement of the ttbar production cross section in the fully hadronic decay channel in pp collisions at 7 TeV*, EPJ Web of Conferences **28** (2012) 12042, ISSN: 2100-014X, URL: `http://dx.doi.org/10.1051/epjconf/20122812042` (cit. on p. 23).

[30] K. Rabbertz, *Jet Physics at the LHC*, 1st ed., Springer International Publishing, 2017 (cit. on pp. 24–28).

[31] R. Kogler et al., *Jet substructure at the Large Hadron Collider*, Reviews of Modern Physics **91** (2019), ISSN: 1539-0756, URL: `http://dx.doi.org/10.1103/RevModPhys.91.045003` (cit. on pp. 24, 25, 27).

[32] D. B. Renner, *Status of Average-x from Lattice QCD*, (2011), URL: `http://dx.doi.org/10.1063/1.3631515` (cit. on p. 24).

[33] J. Collins, *Foundations of Perturbative QCD*, 1st ed., Cambridge University Press, 2011 (cit. on p. 25).

[34] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*, Nuclear Physics B **126** (1977) 298, ISSN: 0550-3213, URL: `https://www.sciencedirect.com/scie%C2%B4nce/article/pii/0550321377903844` (cit. on p. 25).

[35] I. Knowles, *Spin correlations in parton-parton scattering*, Nuclear Physics B **310** (1988) 571, ISSN: 0550-3213, URL: `https://www.sciencedirect.com/science/article/pii/0550321388900922` (cit. on p. 26).

[36] B. Andersson et al., *Parton fragmentation and string dynamics*, Physics Reports **97** (1983) 31, ISSN: 0370-1573, URL: `https://www.sciencedirect.com/science/article/pii/0370157383900807` (cit. on p. 26).

[37] P. Gras et al., *Systematics of quark/gluon tagging*, Journal of High Energy Physics **2017** (2017), ISSN: 1029-8479, URL: `http://dx.doi.org/10.1007/JHEP07(2017)091` (cit. on pp. 29, 30).

[38] *Discrimination of Light Quark and Gluon Jets in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector*, tech. rep., All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2016-034: CERN, 2016, URL: `https://cds.cern.ch/record/2200202` (cit. on pp. 29, 31).

[39] M. Tasevsky, “Differences between quark and gluon jets as seen at LEP”, *International Conference on New Trends in High-Energy Physics: Experiment, Phenomenology, Theory*, 2001 116, arXiv: `hep-ex/0110084` (cit. on pp. 29–31).

[40] K. Ackerstaff et al., *Multiplicity distributions of gluon and quark jets and tests of QCD analytic predictions*, Eur. Phys. J. C **1** (1998) 479, arXiv: `hep-ex/9708029` (cit. on pp. 30, 31).

[41] R. J. Barlow and C. Beeston, *Fitting using finite Monte Carlo samples*, Comput. Phys. Commun. **77** (1993) 219 (cit. on pp. 32, 34, 35).

[42] J. S. Conway, *Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra*, 2011, arXiv: `1103.0354 [physics.data-an]` (cit. on pp. 32, 33, 35).

[43] K. Cranmer et al., *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, tech. rep., New York U., 2012, URL: `https://cds.cern.ch/record/1456844` (cit. on pp. 32, 33, 35, 58).

[44] F. a. James, *MINUIT: Function Minimization and Error Analysis Reference Manual*, (1998), CERN Program Library Long Writeups, URL: `https://cds.cern.ch/record/2296388` (cit. on p. 33).

[45] L. Evans and P. Bryant, *LHC Machine*, Journal of Instrumentation **3** (2008) S08001, URL: `https://doi.org/10.1088/1748-0221/3/08/s08001` (cit. on p. 37).

[46] T. A. Collaboration et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3** (2008) S08003, URL: `https://doi.org/10.1088/1748-0221/3/08/s08003` (cit. on pp. 37, 39, 40, 42, 43).

[47] E. Mobs, *The CERN accelerator complex - 2019. Complexe des accélérateurs du CERN - 2019*, (2019), General Photo, URL: `https://cds.cern.ch/record/2684277` (cit. on p. 38).

[48] C. W. Fabjan and F. Gianotti, *Calorimetry for particle physics*, Rev. Mod. Phys. **75** (2003) 1243 (cit. on p. 41).

[49] S. Weinzierl, *Introduction to Monte Carlo Methods*, ArXiv High Energy Physics - Phenomenology e-prints (2000) (cit. on p. 44).

[50] M. Bähr et al., *Herwig++ physics and manual*, The European Physical Journal C **58** (2008) 639, ISSN: 1434-6052, URL: `http://dx.doi.org/10.1140/epjc/s10052-008-0798-9` (cit. on p. 44).

[51] T. Sjöstrand, *The Pythia event generator: Past, present and future*, Computer Physics Communications **246** (2020) 106910, ISSN: 0010-4655, URL: `http://dx.doi.org/10.1016/j.cpc.2019.106910` (cit. on p. 44).

[52] E. Bothmann et al., *Event generation with Sherpa 2.2*, SciPost Physics **7** (2019) (cit. on p. 44).

[53] S. Agostinelli et al., *Geant4—a simulation toolkit*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **506** (2003) 250, ISSN: 0168-9002, URL: `https://www.sciencedirect.com/science/article/pii/S0168900203013688` (cit. on p. 44).

[54] M. Aaboud et al., *Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2*, The European Physical Journal C **77** (2017), ISSN: 1434-6052, URL: `http://dx.doi.org/10.1140/epjc/s10052-017-5225-7` (cit. on p. 45).

[55] F. Meloni, *Primary vertex reconstruction with the ATLAS detector*, Journal of Instrumentation **11** (2016) C12060, URL: `https://doi.org/10.1088/1748-0221/11/12/c12060` (cit. on p. 45).

[56] M. G. Gándara and the LHCb Collaboration, *Flavour tagging performance in LHCb*, Journal of Physics: Conference Series **171** (2009) 012103, URL: `https://doi.org/10.1088/1742-6596/171/1/012103` (cit. on p. 45).

[57] M. Aaboud et al., *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at*

$$\sqrt{s} = 13$$

TeV

, The European Physical Journal C **79** (2019), ISSN: 1434-6052,
URL: `http://dx.doi.org/10.1140/epjc/s10052-019-7140-6` (cit. on pp. 45, 70, 75).

[58] G. Aad et al., *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV*,
Eur. Phys. J., C **81** (2020) 578. 44 p, 60 pages in total, author list starting page 44, 23 figures, 3 tables. Submitted to journal EPJC. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/MUON-2018-03/,
arXiv: `2012.00578`, URL: `http://cds.cern.ch/record/2746302` (cit. on pp. 45, 70, 75).

[59] C. Limbach, *Reconstruction and Identification of Tau Leptons in ATLAS*, tech. rep., CERN, 2014, URL: `https://cds.cern.ch/record/1954897` (cit. on p. 46).

[60] *Identification of the Hadronic Decays of Tau Leptons in 2012 Data with the ATLAS Detector*, tech. rep., All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2013-064: CERN, 2013, URL: `https://cds.cern.ch/record/1562839`
(cit. on pp. 46–48, 70, 75).

[61] *Determination of the tau energy scale and the associated systematic uncertainty in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector at the LHC in 2012*, tech. rep., All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2013-044: CERN, 2013, URL: `http://cds.cern.ch/record/1544036`
(cit. on p. 47).

[62] A. H. Michele Pinamonti Loic Valery and T. Dado, *TRExFitter documentation*,
URL: `https://trexfitter-docs.web.cern.ch/trexfitter-docs/` (cit. on pp. 49, 57).

[63] *AnalysisTop21*, URL: `https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/AnalysisTop21#21_2_163_built_on_2021_03_10` (cit. on p. 57).

[64] *SingleTopAnalysis*,
URL: `https://gitlab.cern.ch/atlasphys-top/singletop/SingleTopAnalysis`
(cit. on p. 57).

[65] C. Escobar,
Single-top ntuple production (v32) for Run-2 samples (multi-lepton channel with taus),
URL: `https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/SingleTopProductionV32ll#ttbar_samples` (cit. on p. 57).

APPENDIX A

Yields of the alternative lep-had Region

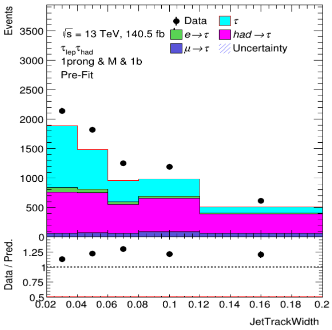
	Medium 1b	non-Medium 0b	non-Medium 1b	non-Medium 2b
tHq	4.16924	0.391964	0.863587	0.512816
$t\bar{t}$	14 937.4	1 288.77	5 357.37	5 634.45
W+jets	440.513	3 119	617.774	65.0116
Z+jets	16 337.2	254 778	37 041.4	2 970.55
Diboson	1 002.91	3 327.87	858.344	109.522
Triboson	4.60753	5.89858	1.74778	0.125159
tZq	67.1899	4.51938	12.0693	5.76302
ttV	205.753	12.8684	52.4303	59.9167
tWZ	23.7834	1.62658	4.98919	2.83496
ttH	62.1476	2.59599	12.1368	16.4251
t-channel	77.1257	83.0952	98.211	15.0136
tW	3.99128	0.424876	0.829554	0.440545
Higgs	43.7306	86.4941	27.681	7.49667
four tops	0.305685	0.00694417	0.0401039	0.104206
Total	33 210.8	262 712	44 085.9	8 888.17
Data	40 560	258 608	51 056	10 772

Table A.1: Yields scaled to luminosity of the alternative lep-had region

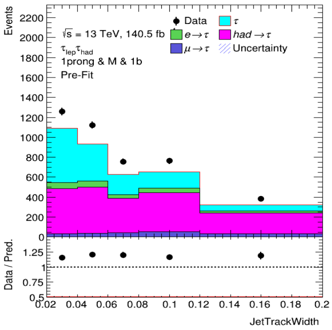
APPENDIX B

Pre- and PostFit Plots for the Estimation of the $|\eta|$ Dependence in the 1-Bin and Quark/Gluon Fit Method

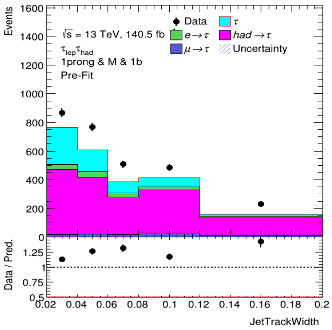
Figure B.1: Pre- and Post-fit plots for the SR in the 1-Bin Method estimation of the $|\eta|$ dependence of the SF's. For 1-prong.



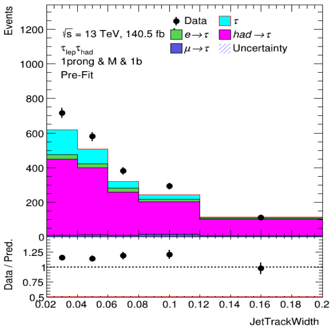
(a) $|\eta| \in [0 : 0.8]$



(b) $|\eta| \in [0.8 : 1.37]$

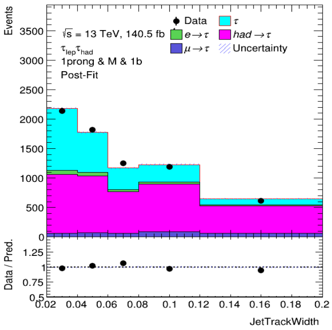


(c) $|\eta| \in [1.52 : 2.0]$

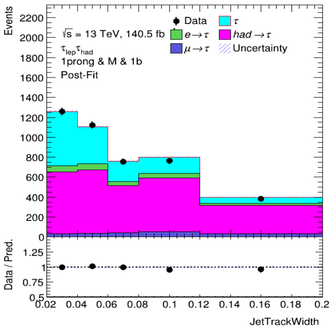


(d) $|\eta| \in [2.0 : 2.5]$

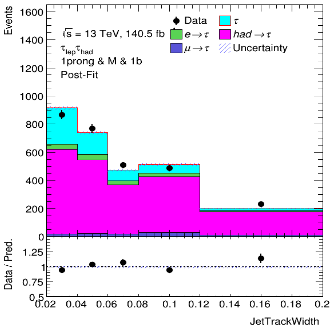
Figure B.2: Pre-fit Plots



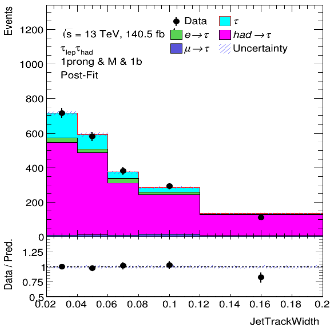
(a) $|\eta| \in [0 : 0.8]$



(b) $|\eta| \in [0.8 : 1.37]$



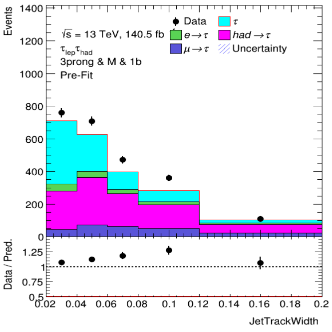
(c) $|\eta| \in [1.52 : 2.0]$



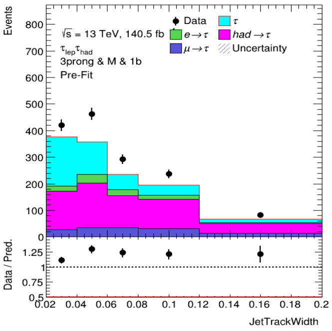
(d) $|\eta| \in [2.0 : 2.5]$

Figure B.3: Post-fit Plots

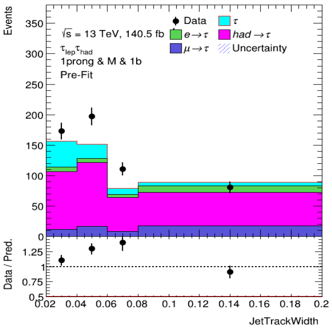
Figure B.4: Pre- and Post-fit plots for the SR in the 1-Bin Method estimation of the $|\eta|$ dependence of the SF's. For 3-prong.



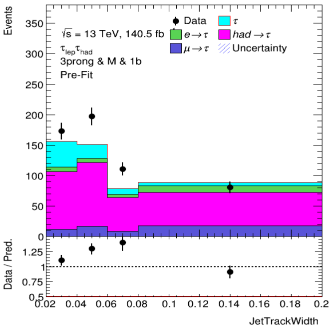
(a) $|\eta| \in [0 : 0.8]$



(b) $|\eta| \in [0.8 : 1.37]$

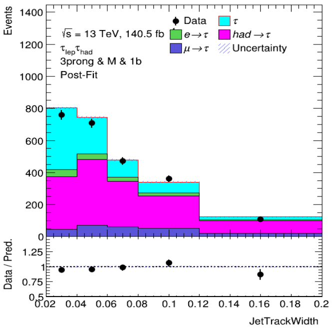


(c) $|\eta| \in [1.52 : 2.0]$

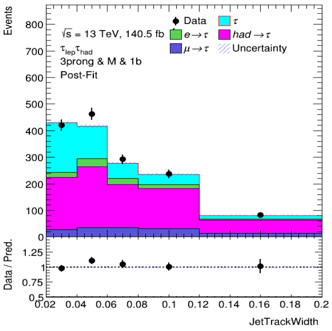


(d) $|\eta| \in [2.0 : 2.5]$

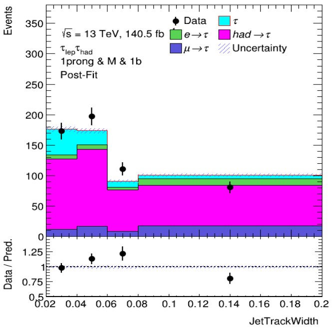
Figure B.5: Pre-fit Plots



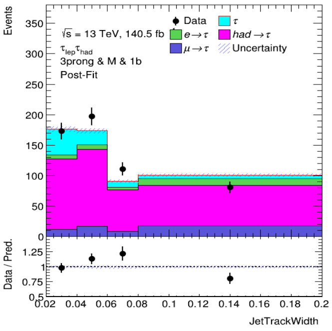
(a) $|\eta| \in [0 : 0.8]$



(b) $|\eta| \in [0.8 : 1.37]$



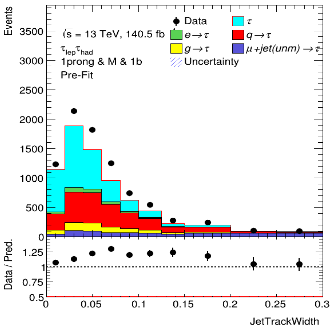
(c) $|\eta| \in [1.52 : 2.0]$



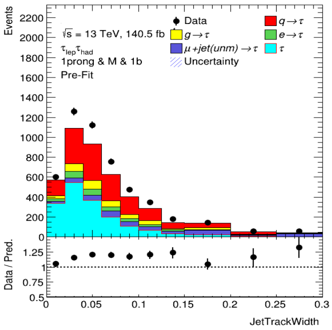
(d) $|\eta| \in [2.0 : 2.5]$

Figure B.6: Post-fit Plots

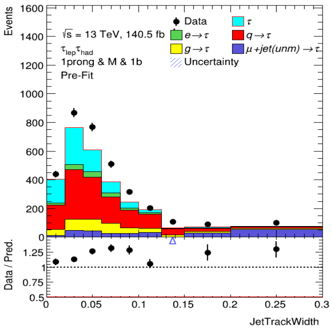
Figure B.7: Pre- and Post-fit plots for the SR in the Quark/Gluon Fit Method estimation of the $|\eta|$ dependence of the SF's. For 1-prong.



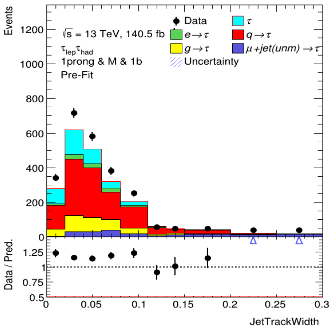
(a) $|\eta| \in [0 : 0.8]$



(b) $|\eta| \in [0.8 : 1.37]$

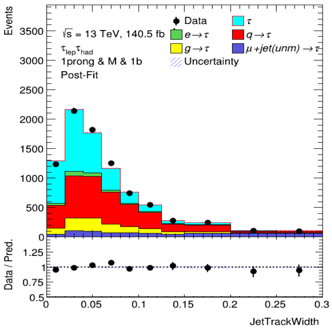


(c) $|\eta| \in [1.52 : 2.0]$

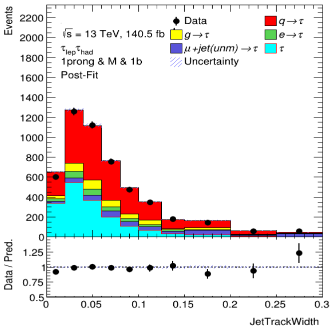


(d) $|\eta| \in [2.0 : 2.5]$

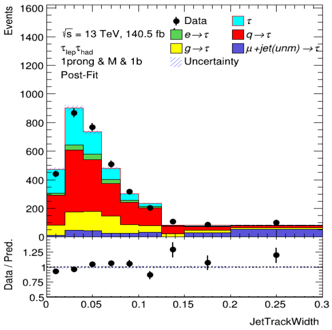
Figure B.8: Pre-fit Plots



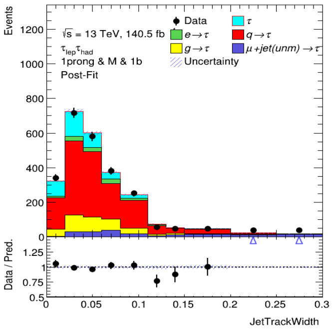
(a) $|\eta| \in [0 : 0.8]$



(b) $|\eta| \in [0.8 : 1.37]$



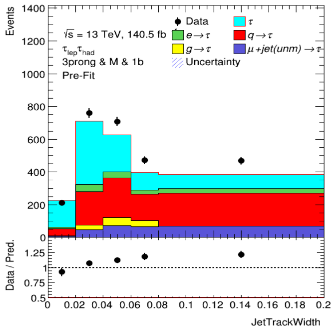
(c) $|\eta| \in [1.52 : 2.0]$



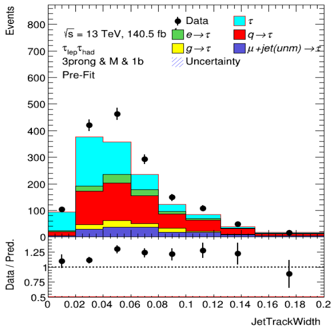
(d) $|\eta| \in [2.0 : 2.5]$

Figure B.9: Post-fit Plots

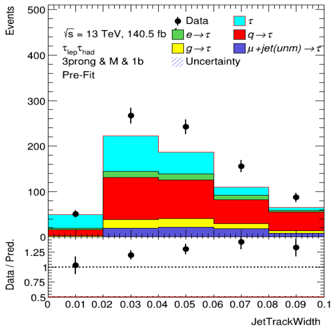
Figure B.10: Pre- and Post-fit plots for the SR in the Quark/Gluon Fit Method estimation of the $|\eta|$ dependence of the SF's. For 3-prong.



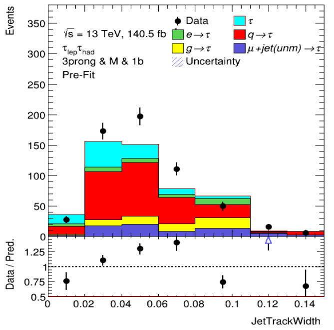
(a) $|\eta| \in [0 : 0.8]$



(b) $|\eta| \in [0.8 : 1.37]$

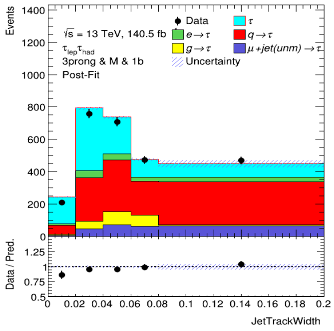


(c) $|\eta| \in [1.52 : 2.0]$

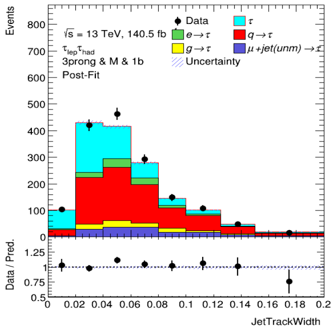


(d) $|\eta| \in [2.0 : 2.5]$

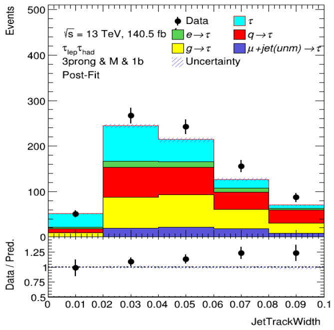
Figure B.11: Pre-fit Plots



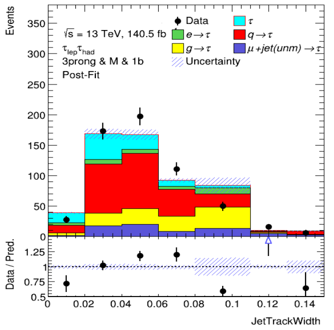
(a) $|\eta| \in [0 : 0.8]$



(b) $|\eta| \in [0.8 : 1.37]$



(c) $|\eta| \in [1.52 : 2.0]$



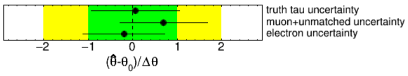
(d) $|\eta| \in [2.0 : 2.5]$

Figure B.12: Post-fit Plots

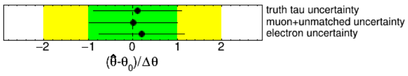
APPENDIX C

Correlation Matrices and Nuisance Parameters in the Quark/Gluon Method

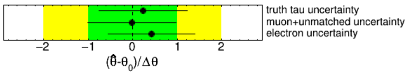
Figure C.1: Nuisance Parameters of the Quark/Gluon Method fits for different p_T for 1-prong and 3-prongs.



(a) $p_\mathrm{T} \in [20 : 30]$ in GeV

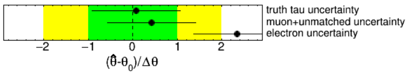


(b) $p_\mathrm{T} \in [30 : 40]$ in GeV

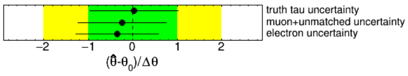


(c) $p_\mathrm{T} > 40$ in GeV

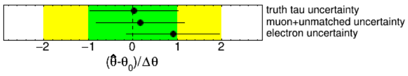
Figure C.2: 1-prong



(a) $p_\mathrm{T} \in [20 : 30]$ in GeV



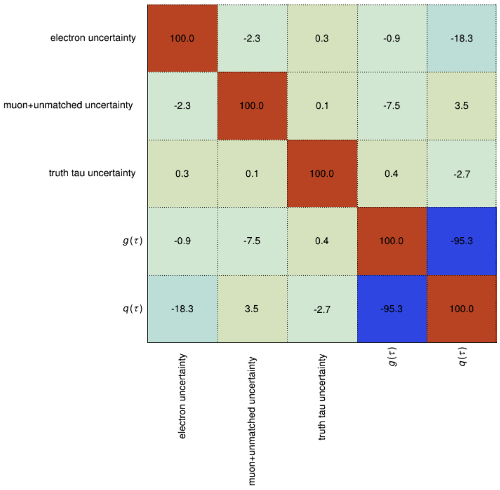
(b) $p_\mathrm{T} \in [30 : 40]$ in GeV



(c) $p_\mathrm{T} > 40$ in GeV

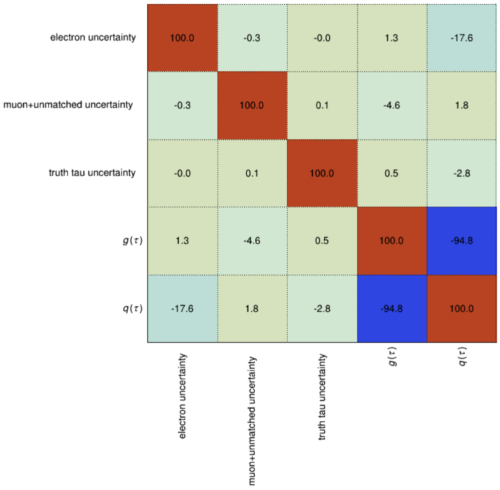
Figure C.3: 3-prongs

Figure C.4: Correlation matrices of Quark/Gluon Method fits for different p_T for 1-prong and 3-prongs.



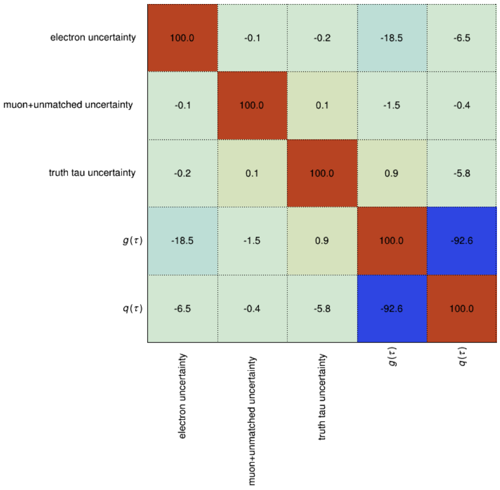
	electron uncertainty	muon+unmatched uncertainty	truth tau uncertainty	$g(\tau)$	$q(\tau)$
electron uncertainty	100.0	-2.3	0.3	-0.9	-18.3
muon+unmatched uncertainty	-2.3	100.0	0.1	-7.5	3.5
truth tau uncertainty	0.3	0.1	100.0	0.4	-2.7
$g(\tau)$	-0.9	-7.5	0.4	100.0	-95.3
$q(\tau)$	-18.3	3.5	-2.7	-95.3	100.0

(a) $p_T \in [20:30]$ in GeV



	electron uncertainty	muon+unmatched uncertainty	truth tau uncertainty	$g(\tau)$	$q(\tau)$
electron uncertainty	100.0	-0.3	-0.0	1.3	-17.6
muon+unmatched uncertainty	-0.3	100.0	0.1	-4.6	1.8
truth tau uncertainty	-0.0	0.1	100.0	0.5	-2.8
$g(\tau)$	1.3	-4.6	0.5	100.0	-94.8
$q(\tau)$	-17.6	1.8	-2.8	-94.8	100.0

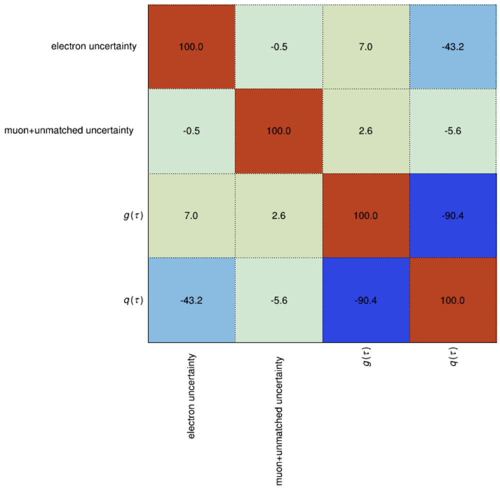
(b) $p_T \in [30:40]$ in GeV



	electron uncertainty	muon+unmatched uncertainty	truth tau uncertainty	$g(\tau)$	$q(\tau)$
electron uncertainty	100.0	-0.1	-0.2	-18.5	-6.5
muon+unmatched uncertainty	-0.1	100.0	0.1	-1.5	-0.4
truth tau uncertainty	-0.2	0.1	100.0	0.9	-5.8
$g(\tau)$	-18.5	-1.5	0.9	100.0	-92.6
$q(\tau)$	-6.5	-0.4	-5.8	-92.6	100.0

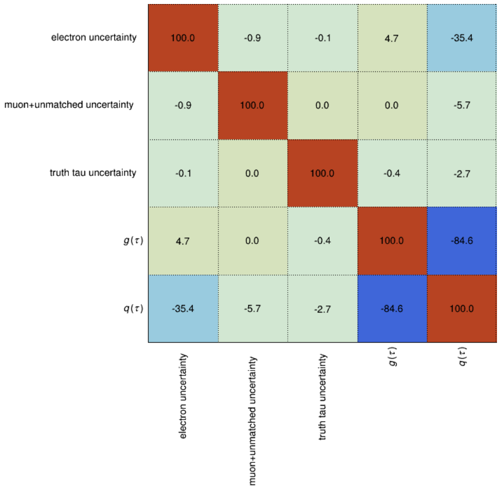
(c) $p_T > 40$ in GeV

Figure C.5: 1-prong



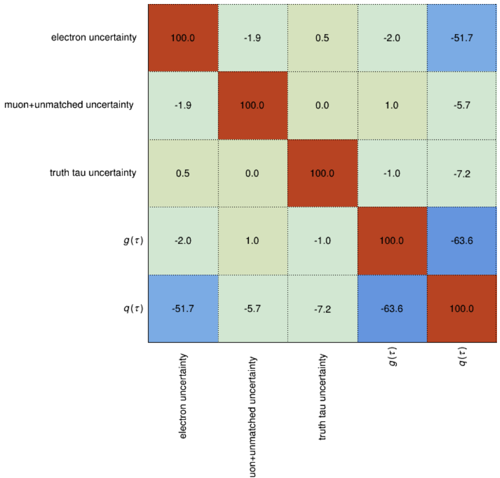
	electron uncertainty	muon+unmatched uncertainty	$g(\tau)$	$q(\tau)$
electron uncertainty	100.0	-0.5	7.0	-43.2
muon+unmatched uncertainty	-0.5	100.0	2.6	-5.6
$g(\tau)$	7.0	2.6	100.0	-90.4
$q(\tau)$	-43.2	-5.6	-90.4	100.0

(a) $p_T \in [20:30]$ in GeV



	electron uncertainty	muon+unmatched uncertainty	truth tau uncertainty	$g(\tau)$	$q(\tau)$
electron uncertainty	100.0	-0.9	-0.1	4.7	-35.4
muon+unmatched uncertainty	-0.9	100.0	0.0	0.0	-5.7
truth tau uncertainty	-0.1	0.0	100.0	-0.4	-2.7
$g(\tau)$	4.7	0.0	-0.4	100.0	-84.6
$q(\tau)$	-35.4	-5.7	-2.7	-84.6	100.0

(b) $p_T \in [30:40]$ in GeV



	electron uncertainty	muon+unmatched uncertainty	truth tau uncertainty	$g(\tau)$	$q(\tau)$
electron uncertainty	100.0	-1.9	0.5	-2.0	-51.7
muon+unmatched uncertainty	-1.9	100.0	0.0	1.0	-5.7
truth tau uncertainty	0.5	0.0	100.0	-1.0	-7.2
$g(\tau)$	-2.0	1.0	-1.0	100.0	-63.6
$q(\tau)$	-51.7	-5.7	-7.2	-63.6	100.0

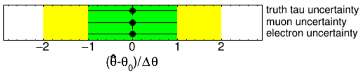
(c) $p_T > 40$ in GeV

Figure C.6: 3-prongs

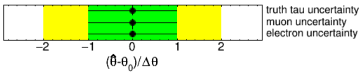
APPENDIX D

Nuisance Parameters in the 1-Bin Method

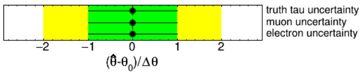
Figure D.1: Nuisance Parameters of the 1-Bin Method fits for different p_{T} for 1-prong and 3-prongs.



(a) $p_{\mathrm{T}} \in [20:30]$ in GeV

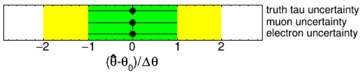


(b) $p_{\mathrm{T}} \in [30:40]$ in GeV

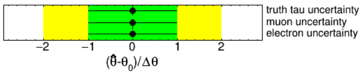


(c) $p_{\mathrm{T}} > 40$ in GeV

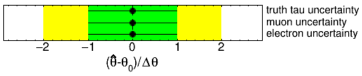
Figure D.2: 1-prong



(a) $p_{\mathrm{T}} \in [20:30]$ in GeV



(b) $p_{\mathrm{T}} \in [30:40]$ in GeV



(c) $p_{\mathrm{T}} > 40$ in GeV

Figure D.3: 3-prongs

APPENDIX E

Fake Tau Task Force Fake Factors

Listing E.1: Preliminary FF’s from the FTTF

```
ff_zmm_dm0
pt bin:      FF      unc.
20.0 30.0: ff 0.253325283527 +- 0.00169621143803
30.0 40.0: ff 0.257479429245 +- 0.00376764222229
40.0 60.0: ff 0.221458137035 +- 0.00490428669353
60.0 90.0: ff 0.174620777369 +- 0.00830860077086
90.0 150.0: ff 0.173809349537 +- 0.0166588708751
150.0 300.0: ff 0.16221742332 +- 0.044134134054

ff_mj_dm0
20.0 30.0: ff 0.20155146718 +- 0.00537638030535
30.0 40.0: ff 0.211563959718 +- 0.0114874508614
40.0 60.0: ff 0.187011927366 +- 0.0154164228422
60.0 90.0: ff 0.126406744123 +- 0.0182074788898
90.0 150.0: ff 0.250402063131 +- 0.0770053506657
150.0 300.0: ff 0.155837848783 +- 0.0193740066182

ff_zmm_dm1
20.0 30.0: ff 0.171441078186 +- 0.000810111182907
30.0 40.0: ff 0.157793179154 +- 0.0014725463309
40.0 60.0: ff 0.128320530057 +- 0.00175758787901
60.0 90.0: ff 0.0995610356331 +- 0.00275227186233
90.0 150.0: ff 0.0845352709293 +- 0.00476513010411
150.0 300.0: ff 0.0736702904105 +- 0.0112606735878

ff_mj_dm1
20.0 30.0: ff 0.127953588963 +- 0.0025189435067
30.0 40.0: ff 0.12859107554 +- 0.00461675861566
40.0 60.0: ff 0.116752833128 +- 0.00583639354893
60.0 90.0: ff 0.089846804738 +- 0.00894314572101
90.0 150.0: ff 0.0820041894913 +- 0.0127525622668
150.0 300.0: ff 0.0432809889317 +- 0.00536797694893
```

```
ff_zmm_dm2
20.0 30.0: ff 0.0948242098093 +- 0.000950555485877
30.0 40.0: ff 0.0964095816016 +- 0.00157076832732
40.0 60.0: ff 0.0834181755781 +- 0.00182472914218
60.0 90.0: ff 0.0680741667747 +- 0.00272614272727
90.0 150.0: ff 0.0667965561152 +- 0.00475872324535
150.0 300.0: ff 0.0433608926833 +- 0.00849081170214

ff_mj_dm2
20.0 30.0: ff 0.0729707479477 +- 0.00299263994537
30.0 40.0: ff 0.0733744204044 +- 0.00474199915123
40.0 60.0: ff 0.0684221833944 +- 0.00572927603479
60.0 90.0: ff 0.0514135546982 +- 0.00627690272282
90.0 150.0: ff 0.0532114468515 +- 0.00518503274142
150.0 300.0: ff 0.0488734431565 +- 0.00433779983595

ff_zmm_dm3
20.0 30.0: ff 0.0437965616584 +- 0.000409851815819
30.0 40.0: ff 0.0375614725053 +- 0.000668721762715
40.0 60.0: ff 0.0298070274293 +- 0.000804607764137
60.0 90.0: ff 0.0228478927165 +- 0.00129734269601
90.0 150.0: ff 0.0209054201841 +- 0.0022072238163
150.0 300.0: ff 0.0136962579563 +- 0.00432440582808

ff_mj_dm3
20.0 30.0: ff 0.0352525264025 +- 0.00133853120605
30.0 40.0: ff 0.0278108157218 +- 0.00195185657182
40.0 60.0: ff 0.0245603304356 +- 0.00251914561685
60.0 90.0: ff 0.01613676548 +- 0.00336347618299
90.0 150.0: ff 0.0126737877727 +- 0.00120751634589
150.0 300.0: ff 0.00902333576232 +- 0.000932742779456

ff_zmm_dm4
20.0 30.0: ff 0.0181961357594 +- 0.000297583731305
30.0 40.0: ff 0.0182524900883 +- 0.000485610923596
40.0 60.0: ff 0.0147620504722 +- 0.000554094807782
60.0 90.0: ff 0.00966475531459 +- 0.000798403554302
90.0 150.0: ff 0.00598668958992 +- 0.00119142184333
150.0 300.0: ff 0.00490593025461 +- 0.00261820633447

ff_mj_dm4
20.0 30.0: ff 0.0160327907652 +- 0.00102452022012
30.0 40.0: ff 0.0151956258342 +- 0.00153387220619
40.0 60.0: ff 0.0152069870383 +- 0.00202908666084
60.0 90.0: ff 0.00680216122419 +- 0.00101740889641
90.0 150.0: ff 0.0061655016616 +- 0.000819974031046
150.0 300.0: ff 0.00707851815969 +- 0.00106345037057
```

List of Figures

List of Tables