Search for new physics in the tau plus missing energy final states at CMS

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften
der RWTH Aachen University zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften genehmigte Dissertation
vorgelegt von

Master of Science
Klaas Ole Padeken
aus Neunkirchen-Seelscheid

Berichter: Prof. Dr. Thomas Hebbeker
Prof. Dr. Lutz Feld

Tag der mündlichen Prüfung: 18.4.2017

Abstract

In this thesis the first search for a heavy charged vector boson W' in the final state with a hadronicaly decaying tau lepton and two neutrinos is described. The analysed data include $20\,\mathrm{fb}^{-1}$ of LHC proton proton collisions at $\sqrt{s} = 8\,\mathrm{TeV}$ as well as $2.3\,\mathrm{fb}^{-1}$ at $\sqrt{s} = 13\,\mathrm{TeV}$ recorded with the CMS detector. A signal would appear as an excess of events with high transverse mass where the standard model background is low.

The algorithm used to reconstruct taus from the detector information was extended to provide a high efficiency and correct energy reconstruction for boosted hadronic tau decays at the TeV scale.

No significant excess is observed in the recorded data. Limits are set on models in which the W' decays preferentially to fermions of the third generation. These results substantially extend previous constraints on this model. W'-boson masses below 2.0 to 3.3 TeV are excluded, depending on the model parameters. In addition, the existence of a W' boson with universal fermion couplings is excluded at 95% credibility level for W' masses below 2.7 TeV with the $\sqrt{s} = 8\,\mathrm{TeV}$ data and for masses below 3.3 TeV with $\sqrt{s} = 13\,\mathrm{TeV}$ data.

For further reinterpretation, a model-independent limit on potential signals is also presented for various transverse mass thresholds.

Finally a generalized interpretation of non resonant signals in the e/μ plus missing transverse energy channel is proposed.

Zusammenfassung

In dieser Arbeit wird die erste Suche nach einem schweren geladenen Vektorboson W' im Endzustand mit einem hadronisch zerfallenden Tau-Lepton und zwei Neutrinos beschrieben. Die analysierten Daten umfassen $20\,\mathrm{fb}^{-1}$ der LHC Proton Proton Kollisionen bei $\sqrt{s} = 8\,\mathrm{TeV}$ sowie $2.3\,\mathrm{fb}^{-1}$ bei $\sqrt{s} = 13\,\mathrm{TeV}$, welche mit dem CMS Detektor ausgenommen wurden. Ein Signal würde als ein Überschuss von Ereignissen mit hoher transversaler Masse erscheinen, wobei der Standardmodelluntergrund für diese Art von Ereignissen niedrig ist.

Um diese Suche durchzuführen wurde die Tau-Rekonstruktion im CMS-Detektor für geboostete Tau Zerfälle erweitert, um eine hohe Effizienz und eine korrekte Energierekonstruktion bei der TeV Skala zu gewährleisten.

Es wird kein signifikanter Überschuss in den Daten beobachtet. Es werden Ausschlussgrenzen auf Modelle gesetzt, in dem das W' bevorzugt in Fermionen der dritten Generation zerfällt. Diese Ergebnisse erweitern die vorherigen Ausschlussgrenzen auf dieses Modell erheblich. Eine untere Grenze auf erlaubte W' Boson-Massen kann gesetzt werden. Diese liegt zwischen 2.0 und 3.3 TeV, je nach Modellparameter. Darüber hinaus wird die Existenz eines W'-Bosons mit universellen Fermionkopplungen mit einem Konfidenzniveau von 95% für W' Massen unter 2.7 TeV für die $\sqrt{s} = 8\,\mathrm{TeV}$-Daten und für Massen unter 3.3 TeV für $\sqrt{s} = 13\,\mathrm{TeV}$-Daten ausgeschlossen.

Um weitere Interpretationen der Ergebnisse zu ermöglichen wird auch ein modellunabhängiger Grenzwert auf potenzielle Signale für verschiedene transversale Massenschwellen errechnet.

Zusätzlich wird eine verallgemeinerte Interpretation von nicht resonanten Signalen vorgestellt, welche anhand des Endzustandes in e/μ plus fehlender transversaler Energie eingeführt wird.

Contents

1. Introduction

The standard model of particle physics was very successful in the last five decades, but new processes of physics beyond the standard model (BSM) are expected to appear at the TeV scale. With the CMS detector at the LHC new, previously untested energy regimes are accessible to searches for new physics. The main aspect BSM theories studied in this thesis are events with one tau and missing transverse energy (E_T^{miss}), which are analysed for signs of a new heavy gauge boson. Additionally a new model independent limit is proposed for non resonant signals, where the e/μ and E_T^{miss} final state is used.

This is the first published search for BSM physics in the tau and E_T^{miss} final state. It faces two major challenges: To reconstruct hadronic tau decays at the TeV scale with a high efficiency and to describe the background in the final state, for which it is not possible to derive it completely from data. These points are addressed by updating the tau reconstruction to include highly boosted taus and an extensive study of the dominating W-boson background.

As signal a generic model of a heavy charged gauge boson is used. Several extensions to this model, which have phenomenological consequences, are tested. These modifications are a generalization of the coupling structure and a dedicated model, motivating the large masses of the third fermion family with a non universal coupling structure. To allow an easy reinterpretation of the final result with any model, limits are also provided in a model unspecific way.

The distribution of the main discriminating variable in tau plus E_T^{miss} final states is drastically different, compared to previous mono-lepton searches. The idea of a generalized limit is also transferred to the e/μ plus E_T^{miss} final state, where a new parametrisation for non resonant limits is proposed.

The analysed data is recorded in two data taking periods of the CMS detector. The first data set is collected in 2012, using $20\,\mathrm{fb}^{-1}$ of integrated luminosity of the proton proton collision provided by the LHC at a centre of mass energy of $\sqrt{s} = 8\,\mathrm{TeV}$. The second data set is recorded in 2015 at a centre of mass energy of $\sqrt{s} = 13\,\mathrm{TeV}$. It amounts to $2.3\,\mathrm{fb}^{-1}$. These are analysed separately due to comprehensive changes in the recording conditions.

This thesis is divided into thirteen chapters. After the introduction the theoretical and experimental foundations are explained, needed for the analysis of the investigated models. Building on this the analysis of the 8 TeV and 13 TeV data is described. The results of these analyses are first presented separately and combined afterwards. Finally a new generalized limit for non resonant mono-lepton searches is proposed.

During the preparation of this thesis several parts of my analysis have been published. The initial publication of the analysis is a public analysis summary (PAS) [1] followed by the paper [2]. The analysis of the $\sqrt{s} = 13\,\mathrm{TeV}$ data is published as a PAS in [3]. In addition to the analysis also the tau reconstruction is published in [4]. I have also contributed extensively

to the mono-lepton channel ($e/\mu + E_T^{miss}$), which is not documented in detail in this thesis. My contribution was to the publication of [5, 6, 7]. During this work I also supervised following theses [8, 9, 10, 11, 12, 13], detailing the search for dark matter in the mono-lepton channel. In order to provide a systematic method for the quickly evolving dark matter model landscape, the generalized limit for non resonant signals is proposed in this thesis. In the following all figures not provided by me to these publications, will be referenced in the figure captions.

Leading up to this analysis was among others the master thesis [14], which started the search in the tau plus E_T^{miss} channel.

2. Theory

In order to perform a search for new physics, the foundations of the new physics have to be understood. Therefore the needed known physics concepts will be introduced in this chapter. The focus of this section will be on the aspects that are needed to understand the investigated models for new physics. These will be presented in the second part of this chapter.

2.1. The Standard Model

The standard model of particle physics (SM) describes the known elementary particles and their interactions in terms of a field theory based on gauge symmetries. The foundation of the SM was laid in the 60s and the current representation, which is our understanding of the fundamental concepts in nature, was developed on this basis. Elementary particles are fundamental particles that cannot be described as a combination of other particles. Each particle has its unique quantum numbers, such as charge or spin. There are two kinds of elementary particles: fermions with half integer spin and bosons with an integer spin. The interaction between them can be categorized in three fundamental forces: electroweak (QED[1], QFD[2]), strong (QCD[3]) and gravitational interaction. Within the SM one can describe very precisely the transition from one particle to another, particle decays, annihilations and the production in vacuum. The energy range this model can describe starts at a few eV and is tested up to a few hundred GeV. The SM gives a unified description of the electromagnetic, weak and strong interactions in terms of the gauge group $\mathsf{U}(1)_Y \times \mathsf{SU}(2)_L \times \mathsf{SU}(3)_C$. However, the SM does not contain a description of gravitation. This is described by the theory of general relativity. No unification of the standard model in particle physics and gravitation was achieved so far. Therefore a hard validity limit of the SM is given by the Planck scale of 10^{19} GeV where quantum-gravity effects are expected to become relevant.

2.1.1. Particles

The fermions are divided into leptons and quarks, where the former participate only in the electroweak interaction, while the latter also interact strongly. For every type of fermion a particle and an antiparticle exist with opposite electric charge. Unless stated explicitly in this thesis the particle also refers to its antiparticle. All fermions are listed in Table 2.1. The fermions are arranged in three families, which differ by the mass of the fermions. The increasing mass pattern of the fermion families is unexplained. The third family's fermions are much heavier than the others and they can decay according to a complex structure. The top

[1] quantum electrodynamics
[2] quantum flavourdynamics
[3] quantum chromodynamics

		particle	name	mass
lepton	charged l.	e	electron	0.51 MeV
		μ	muon	105 MeV
		τ	tau	1776 MeV
	neutrinos	ν_e	electron neutrino	<2 eV
		ν_μ	muon neutrino	<2 eV
		ν_τ	tau neutrino	<2 eV
quarks	up-type	u	up	2.3 MeV
		c	charm	1275 MeV
		t	top	174 GeV
	down-type	d	down	4.8 MeV
		s	strange	95 MeV
		b	bottom	4.18 GeV

Table 2.1.: Table of all fermions in the SM [15, 16].

is the only quark that decays electroweakly before hadronizing, and the tau is the only lepton that decays hadronically. The lightest stable charged fermions the u and d quarks and the e leptons form bound states and are observable in our universe as matter.

The bosons correspond to the different interactions in the SM. The electroweak interaction is mediated by the γ, W- and Z-bosons, while the gluons mediate the strong force. The recently discovered H-boson [17] is needed to allow mass terms for particles in the Lagrangian, which otherwise would violate gauge invariance. All bosons are listed in Table 2.2.

particle	name	mass	force
g	gluon	0 GeV	QCD
γ	photon	0 GeV	QED
Z	Z	$91.1876^{+0.0021}_{-0.0021}$ GeV	QFD
$W^\pm$	W	$80.385^{+0.015}_{-0.015}$ GeV	QFD
h	Higgs	$125.09^{+0.24}_{-0.24}$ GeV	electroweak symmetry breaking

Table 2.2.: Table of all bosons in the SM [15].

2.1.2. Electroweak Interactions

The electroweak theory [18, 19, 20] is the most important model to describe rare processes at the LHC. The main background for this analysis will be the W-boson production and decay, therefore a short overview of the electroweak theory will be given, namely the unified description of the electromagnetic and the weak interaction. The corresponding gauge group is therefore a combination of the abelian, electromagnetic group $U(1)_Y$ and the non-abelian, weak group $SU(2)_L$ to a $SU(2)_L \times U(1)_Y$. From this group structure we can deduce that that there are four bosons: three generators of the $SU(2)_L$, W^i_μ ($i = 1, 2, 3$) and the generator of the $U(1)_Y$, B_μ. The fermions can be written in a bases that shows their chirality. For the first

quark family, this is:

$$\psi_1 = \begin{pmatrix} u \\ d \end{pmatrix}_L, \quad \psi_2 = u_R, \quad \psi_3 = d_R. \tag{2.1}$$

This way we can write a compact form of the electroweak interaction Langrangian density as:

$$\mathscr{L}_{int} = g\bar{\psi}_1\gamma^\mu \frac{\sigma_i}{2} W^i_\mu \psi_1 + g' \sum_{j=1}^{3} y_j \bar{\psi}_j \gamma^\mu B_\mu \psi_j h.c., \tag{2.2}$$

where σ_i are the three Pauli matrices and the couplings of the weak interaction are g and g', connected via the weak mixing angle θ_W with $g \cdot \sin\theta_W = g' \cdot \cos\theta_W = e$. The γ^μ are known as the Dirac matrices. They are defined via the anticommutator relation $\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\nu\mu}$, with the Minkowski metric $\eta^{\nu\mu}$. The bosons W^i_μ and B_μ can be mapped to the observed bosons. While the charged bosons W^1_μ and W^2_μ are connected to the observed W-boson via:

$$W^\pm_\mu = (W^1_\mu \mp iW^2_\mu)/\sqrt{2}, \tag{2.3}$$

the neutral bosons W^3_μ and B_μ are connected to the observed bosons via the electroweak mixing:

$$\begin{pmatrix} W^3_\mu \\ B_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} \tag{2.4}$$

For the quarks the flavour eigenstates q' are not the same as the mass eigenstates q. The two are connected by the CKM mixing matrix:

$$\begin{pmatrix} d' \\ c' \\ b' \end{pmatrix} = V_{qq'} \begin{pmatrix} d \\ c \\ b \end{pmatrix} \tag{2.5}$$

The leading order interaction of a W-boson with two fermions ($W \to f_i \bar{f}_j$) can be written as:

$$\mathscr{L}_{int} = \frac{V_{ij}}{2\sqrt{2}} g \bar{f^i} \gamma^\mu (1-\gamma^5) W_\mu f^j + \text{h.c.}, \tag{2.6}$$

where γ^5 is $i\prod_\mu \gamma^\mu$. In this thesis we will test various modifications to this interaction. Because on the one hand, the W-boson decay into $l\nu$ is the main background for the search for physics beyond the SM, on the other hand, models with properties similar to the W boson will be investigated. From the Langrangian one can see that the W boson couples only to left handed fermions and right handed antifermions, which is formally given by the chirality operator $P_L = \frac{1}{2}(1-\gamma^5)$.

The leading order cross section for the process $f_i \bar{f}_j \to W \to f_{i'} \bar{f}_{j'}$ can be written as [15]:

$$\frac{d\sigma}{d\Omega} = \frac{N_c^f}{N_c^i} \frac{1}{256 \cdot 4 \cdot \pi^2} \cdot V^2_{ij} \cdot V^2_{i'j'} \cdot \frac{s \cdot g^4}{(s - m^2_W)^2 + s\Gamma^2} (1 + \cos^2\theta + 2\cos\theta), \tag{2.7}$$

where N_c is the number of colours for the initial or final fermion f, θ is the angle between the fermions and $\sqrt{s}$ is the centre of mass energy for this process. One may note several things

about this process. The mass of the particle in the propagator M_W separates the process into three regions ($s \ll m_W^2$), ($s \approx m_W^2$) and ($s \gg m_W^2$). When the transferred energy $\sqrt{s}$ is smaller than the mass of the W boson, the term $(s - m_W^2)$ is dominated by the W-boson mass and for large values of s the W-boson mass can be neglected. These two regions are called off-shell regions, which means that the momentum transfer of the process is not in the region of the W mass. For $s \ll M_W^2$ the process can be described as a contact interaction where the heavy mediator cannot be resolved. For $s \gg M_W^2$ the W boson can be produced with a heavier mass. The process is called on shell, when $s \approx M_W^2$, then the $(s - m_W^2)$ term vanishes and the width of the propagator determines the cross section.

For two general fermions the width is given by [15]:

$$\Gamma(W \to f_i \bar{f}_j) = N_c \frac{g^2 \cdot M_W \cdot |V_{ij}|^2}{24\pi} \tag{2.8}$$

The total width is then the sum over all possible decay modes $\Gamma = \sum \Gamma_{ij}$. For the cross section a few simplifications were made, to write these short versions. The masses of the fermions are small in comparison with the W-boson mass, therefore they are neglected here. This is of course incorrect, which can be directly seen for the decay of the top quark where the fermion is more than two times heavier than the boson, see Tables 2.1 and 2.2. Only the leading order processes are considered here. Higher order corrections are crucial for this analysis. Therefore the effect and basic principles of next to leading order (NLO) calculations are explained in Section 2.1.5. Remark that the cross section scales with g^4. On the mass pole, the quadratic width also scales with g^4, which implies that for a universally coupling boson the on-shell cross section is independent of the coupling strength.

2.1.3. Electroweak Symmetry Breaking

The QCD description is non trivial, but the theory is in itself consistent. This is not the case for the electroweak description without the Higgs mechanism. A mass scale has to be introduced externally which defines the boson mass scales and therefore the weak coupling strength. For a long time the outstanding problem was, that in the SM all particles were massless. The introduced masses would lead to gauge invariance violating Lagrangian. The solution is that the boson masses are generated dynamically by the interaction with an additional field, the Higgs field, which spontaneously breaks the gauge invariance.

A short outline of this mechanism will be provided here, in order to motivate some models beyond the SM. This summary is based on [21, 22, 23, 24]. For a detailed discussion consult e.g [24, 25].

The idea of the Higgs mechanism is that a complex scalar field could couple to the SM particles. By spontaneously breaking the electroweak gauge symmetry a non vanishing ground state would generate mass terms in the Lagrangian.

The Lagrangian of a complex scalar field ϕ can be written as[4]:

$$\mathcal{L}_S = \partial_\mu \phi^\dagger \partial^\mu \phi - V(\phi) \text{ with } V(\phi) = \mu^2 \phi^\dagger \phi + \frac{1}{2}(\phi^\dagger \phi)^2 \tag{2.9}$$

[4]We will see at the end of this section, that this is not the most general form. One can also add fermion interactions.

This Lagrange density is gauge invariant, but is also can have a non vanishing ground state.

$$< 0|\phi(x)|0 >= \begin{pmatrix} 0 \\ v \end{pmatrix} = \phi_G \neq 0 \tag{2.10}$$

In order to determine the interactions of the scalar field with the SM the partial derivative ∂_μ in equation 2.9 has to be replaced by the covariant derivative D_μ, which is determined by the gauge transformations of the field:

$$D_\mu \phi = [\partial_\mu + ig\frac{\sigma^i}{2}W_\mu^i - i\frac{Y_H}{2}g' B_\mu]\phi, \tag{2.11}$$

where $Y_H = 1/2$ is the hypercharge for the Higgs field. Using this equations one can work out the mass terms for the bosons by comparing:

$$(D_\mu \phi_G)^\dagger D^\mu \phi_G = \frac{v^2}{4}g^2 W_\mu^+ W^{-\mu} + \frac{v^2}{8}(g'^2+g^2)\left(\frac{g'}{\sqrt{g'^2+g^2}}B_\mu - \frac{g}{\sqrt{g'^2+g^2}}W_\mu^3\right)^2, \tag{2.12}$$

to the observable boson from Equation 2.4. One can see the masses of the bosons as e.g. $M_W = (gv/2)^2$. Keeping all terms, also the interactions with a scalar boson, the Higgs boson appear.

As we have seen the masses of the bosons are generated dynamically by the introduction of a non vanishing ground state of a scalar field. But the masses of the fermions are still not accounted for. They can be introduced by the same scalar field with a coupling proportional to their masses. For e.g. the up- and down-quarks defined as defined Equation 2.1, the interaction is:

$$\mathscr{L}_{Y,int} = \bar{Q} d_R \phi \overset{\phi \to \phi_G}{=} -Y_d \frac{v}{\sqrt{2}} \bar{d_L} d_R \tag{2.13}$$

As one can see the masses as constructed here only work for the charged leptons and the down-type quarks. In order to have mass terms for the up-type quarks more detailed discussion of the Higgs field is needed, see e.g.[24].

2.1.4. Quantum Chromodynamics

As this thesis will analyse events from a hadron collider the strong interactions are an important model to understand the background and the production of the events.

QCD [26] is described by the non-abelian gauge group SU(3) and the characterizing charge is the so called colour charge (q_c with $c = 1, 2, 3$). The interaction is mediated by eight gluons, which carry colour charge. The strong coupling α_s has a non-negligible dependence on the energy scale. Therefore QCD can not be described with a perturbative approach at low energies, which is very efficient for high energies ($\alpha_s \approx 0.1$ at energies of 100 GeV-1 TeV [15]). The running of α_s is shown in Figure 2.1. As feature of the theory there are no free quarks or gluons observable, which is often referred to as confinement. High energetic light quarks or gluons first hadronise and then form a parton shower, which is observable as a jet of final state particles in modern particle detectors.

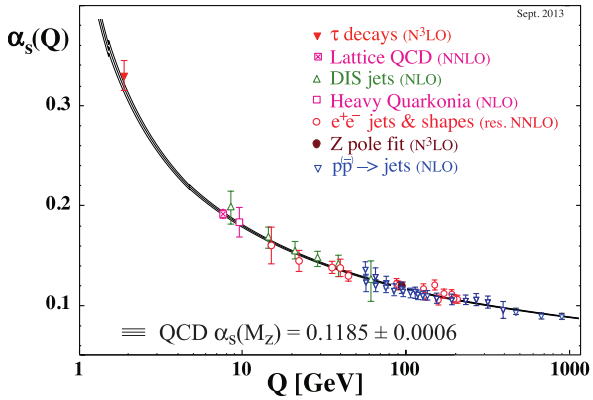


Figure 2.1.: Running of the strong coupling constant α_s as a function of the energy transfer Q [15].

Many of the used results from QCD are either perturbative calculations, thus only valid at higher energies, or parametrizations for low energy effects[5]. The diverging coupling is especially problematic for numerical calculations. There are many applications at the LHC, where the interplay of these two regimes is important. One example is the description of jets, which either are calculated with the perturbative approach or parametrized. For the parametrization of the fragmentation the most commonly used framework in CMS simulations is PYTHIA [27, 28]. The use of numerical calculations for high energy QCD and parametrizations for low energy QCD effects is a subject of constant improvements as discussed in [29].

2.1.5. Higher Order Effects for $pp \to W \to l\nu$

A short summary of the higher order effects will be provided, in order to show the recent developments and the difficulties for the next-to-leading-order (NLO) calculations[6]. This will be used for the correct description of the background.

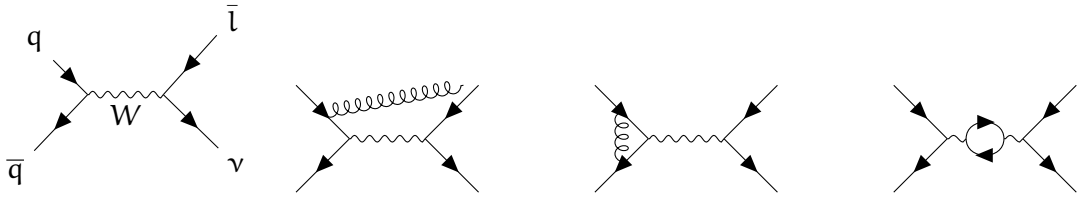
For the correct calculation of the cross sections higher orders have to be calculated, which in this context means the evaluation of diagrams with additional vertices of the strength[7] α_s or α. The number of diagrams that have to be evaluated grows exponentially with the number of allowed vertices. An analytical evaluation of the higher order diagrams is therefore complicated. In the past years the automatic numerical evaluation of higher order diagrams and calculation of generic diagrams had a huge success. An example for the automatic calculation of NLO QCD diagrams is the OPENLOOPS library [32]. In the last years also the automatic calculation of NLO electroweak (NLO EW) diagrams was added [33]. In order to understand the difference between leading-order (LO) and NLO four exemplary Feynman diagrams are shown in Figure 2.2. For the numerical NLO calculations several divergencies arise, which will be discussed using these diagrams. One has to note here that an analytical evaluation of

[5] In principle the low energy QCD regime is also accessible with lattice calculations, but these calculations have not reached the needed precision.

[6] As the differential NLO calculations are still under development, the following paragraph is based on [30] and private discussions with Alexander Mück and Lennart Oymanns. More details for the $pp \to W \to l\nu$ process can be found in [31].

[7] α is proportional to g^2. To be exact $\alpha_i = \frac{g_i^2}{4\pi}$.

higher orders show no not handable divergences in the calculation. Many of the divergences arise from the numerical calculation and the evaluation of non inclusive quantities.



(a) LO (b) NLO initial state radiation (QCD or electroweak) (c) NLO vertex loop (QCD or electroweak) (d) NLO one loop (electroweak)

Figure 2.2.: Feynman diagrams for a LO s-channel process and some examples for NLO corrections.

The Figure 2.2a shows the LO s-channel diagram. If we assume that the initial state particles are quarks and the final state particles are leptons, the NLO QCD corrections mainly affect the initial state, while the NLO EW corrections can affect all parts of the diagrams. In light of the values for $\alpha_s(M_Z) = 0.118$ and $\alpha(M_Z) = 7.297 \times 10^{-3}$ [15], it is clear that the QCD corrections from higher order processes are much stronger than the electroweak processes. For modern Monte Carlo (MC) event generators such as MadGraph [30, 34] the LO process can be defined as $pp \to W \to l\nu$+n jets, which makes the counting of the order ambiguous. Figure 2.2b and the crossed graph with the gluon in the initial state can be viewed as LO W+ 1 jet process or NLO graph to diagram 2.2a. Most modern generators therefore define NLO processes as α_s^{n+1} or α^{m+1}, where n and m are the orders of the defined process. This separation is ambiguous due to the definition of "the defined process". The answer is that Figure 2.2b is both NLO and LO. For the numerical calculation of LO n-jet diagrams divergences appear for collinear and soft particle radiations, which would be cancelled by virtual corrections at the full NLO or higher order calculations. LO generators typically circumvent this, by using thresholds for the momenta of the radiated particles, requiring e.g. for a jet a typical minimal p_T of $10 - 20\,\text{GeV}$. The low momentum jets are parametrized. Full NLO calculations also provide the correct virtual corrections and the correct parametrisation of the low momentum jets have to be adjusted to only describe the missing NNLO effects. Another problem arises from the radiation of photons and heavy gauge bosons, because they are often generated in a separate process like $pp \to W$ and $pp \to WZ$. The virtual graphs for the NLO W-boson process, radiating a Z boson, would not be simulated in the explicit WZ process. A NLO EW calculation of the W-boson process would therefore only contain the virtual corrections, and thus most likely have a negative contribution to the cross section.

Another aspect of higher order corrections can be seen in Figure 2.2c. The diagram effectively changes the coupling constant at the vertex. It can be renomalized analytically. It is then absorbed in the running of the coupling constant, as shown in Figure 2.1 for α_s. Numerical NLO calculations can calculate a correction to the vertex due to a phase space constraint or additional initial state radiation. Therefore the NLO calculation has to agree with the LO calculation on the renormalizaiton scheme. They have to use the same parametrization of the running coupling constant.The last diagram 2.2d changes the width of the propagator and an be completely renormalized by analytical methods.

One of the reasons for the so called NLO revolution however was the use of on-shell calculations

with so called tree diagrams instead of feynman diagrams, where much more of the calculations can be reused [35, 36]. More recently it was shown that the loop integrations can be solved generally in a D-dimensional approach which evolved in a series of publications, of which [37] is the most recent.

The first goal of the automatisation was the NLO QCD calculations for the strong coupling. For the NLO EW corrections in certain parts of the phasespace such as the high mass M_W boson so called Sudakov logarithms arise at the TeV scale. These logarithms have the form of $g^2 \cdot \log^2(s/m^2)$, where m is the mass of e.g. an external lepton and s is the usual Mandelstam variable of the process. Therefore recent attempts were made to include the electroweak corrections in the calculations e.g. with HORACE [38], mcSANC [39] and most recently OPENLOOPS within SHERPA [40][8]. The resulting effects of the higher order calculations to the W-boson background are studied in detail in Section 7.2.3 and 8.2.1.

2.1.6. Parton Distribution Functions

At the Large Hadron Collider (LHC) protons are accelerated and collided. This brings certain advantages with respect to electron colliders, like the essential high centre of mass energy, but also has some disadvantages. One of them is that protons are not elementary particles. Protons have a substructure and the particles taking part in the interactions are the partons (quarks and gluons) constituting the protons. The propability to find a specific parton at a specific energy inside the proton can be described in terms of parton distributions functions (PDF). The idea is that the partons in the proton have a probability to carry a fraction of the proton momentum. This probability for each parton depends on the energy transfer Q of the interaction and the fraction of the energy with respect to the proton x ($f_i(x, Q^2)$). The cross section of a partonic (hard) interaction P can then be written as [41]:

$$\sigma(pp \to X_P) \propto \sum_i \sum_j C_{ij}^P(x_1, x_2, \alpha_s(Q^2)) \otimes f_i(x_1, Q^2) \otimes f_j(x_2, Q^2), \tag{2.14}$$

where $C_{ij}^P(x_1, x_2, \alpha_s(Q^2))$ describes the hard process for the parton types i and j at an energy Q with the coupling constant α_s depending on Q^2. For the PDF the dependence on Q^2 can be calculated using the DGLAP equations [42, 43, 44]. The dependency of the PDFs on x at the current understanding of QCD cannot be determined from first principles. They describe the dynamics in the proton, which can not be calculated. In order to determine these proton form functions data from various collision experiments were analysed e.g. $ep, en, p\bar{p}, pp$[9]. While the DGLAP equations are known since the mid to end 70s and dedicated ep colliders such as HERA measured the PDFs in great detail, the NLO QCD calculations of the energy evolution could only be solved 20 years later [45]. Since then the combination of various data sources and the estimation of the involved uncertainties has grown. The current PDFs are at next-to-next-to-leading-order (NNLO) QCD, include also photon contributions and have largely decreased the uncertainty. But there are still discrepancies in the description of the PDFs, depending on the combination and fit method used for the data. An example PDF is shown in Figure 2.3 and 2.4. One can see that for higher Q^2 the uncertainties are smaller,

[8] The code is not yet available publicly
[9] For a full list see [15] Chap.19

because the energy evolution can be described perturbatively. The so called valence quarks, which constitute a proton (uud), have a probability peak at $x \approx 1/3$, while the other quarks and the gluon are least probable at high x and steeply rise towards low x, which means that it is more probable to get a low x $\bar{u}$ than a high x u. For this analysis we will have $\bar{q}q'$ combinations in the initial state with high Q^2 and high x. For the heavier quark flavours (s, c, b), the probability to find a quark or an antiquark in the proton is the same, since these quarks can only be produced dynamically in pairs. The difference between various PDF sets is much larger than the uncertainty of the individual PDF sets. This will be refrected in the uncertainties, which will be used later.

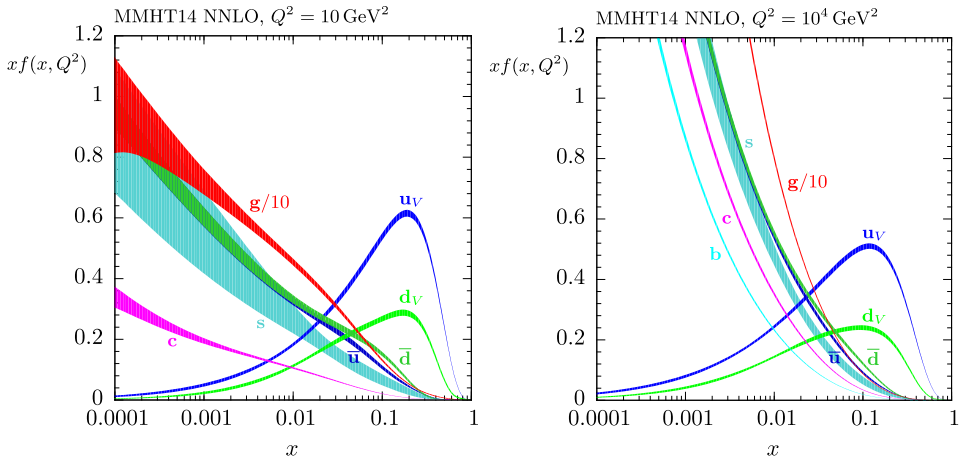


Figure 2.3.: An example PDF (MMHT2014) at NNLO at $Q^2 = 10\,\mathrm{GeV}^2$ and $Q^2 = 10^4\,\mathrm{GeV}^2$, with associated 68% confidence-level uncertainty bands, taken from [46].

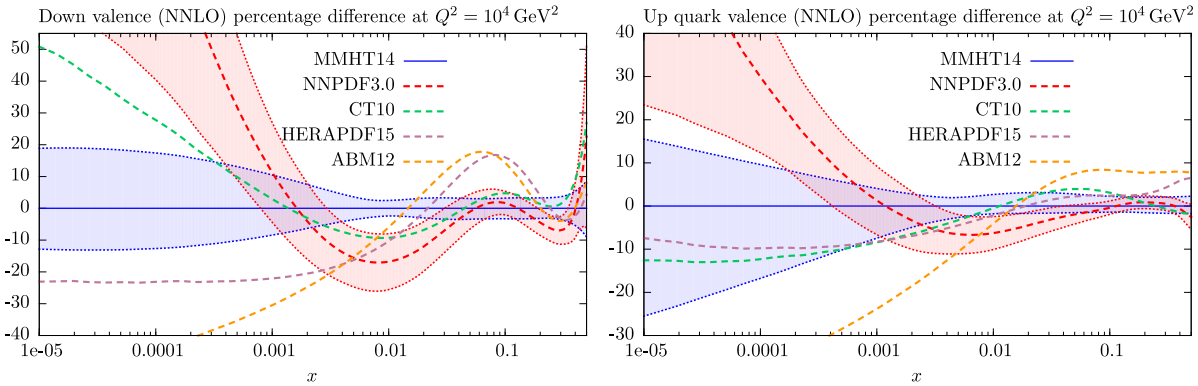


Figure 2.4.: The difference of various PDF sets for $Q^2 = 10^4\,\mathrm{GeV}^2$, taken from [46].

The PDFs build the foundation of every cross section calculation and, as we will see in Section 7.2.3, have a huge influence on the result and thus effect the background prediction.

2.2. Limitations of the Standard Model

Despite the huge success of the SM there are phenomena that cannot be explained within this theory. Therefore many extensions of the current standard model are aiming at solving one or more of the shortcomings of the SM. Different theories aim for different shortcomings of the SM. To understand the motivations a short selection of problems within the SM is given. A more detailed review can be found in the introductions of [47, 48, 49]:

- There are two quantities observed in astrophysical and cosmological measurements, which cannot be explained within the SM: Dark Matter and Dark Energy.
- A generalized description of gravity as quantum field theory, or a description of the particle fields as geometric theory - probably a unified description of both in a new framework - is still missing.
- The origin of the weak scale is unexplained. The electroweak scale is not protected by any symmetry, and therefore is unnaturally low, compared to the Planck scale.
- The recently discovered Higgs boson is relatively heavy, which requires a high fine-tuning in loops that have a large amplitude and cancel each other.
- The origin of the neutrino masses is still debated. It is unclear why they are that low, compared with the rest of the fermions.
- While the masses of the fermions in the SM can be introduced in the theory, no mechanism in the SM explains the difference of masses between the fermion families.
- No gauge unification is possible within the SM.

As one can see many of these shortcomings are a lack of motivation for specific observations (Dark Matter) or scale arguments. But the underlying assumption, when discussing the shortcomings of the SM, is that the theory is only an approximation. It describes the status quo of the known measurements. With a new measurement e.g. neutrino masses, the existing theoretical framework is updated. Therefore some new theories try to motivate their approach by a simplification of the SM, or an introduction of a new symmetry, while others propose new phenomena that would have been missed so far.

One of the most prominent extensions of the SM is Supersymmetry (SUSY) [47], which introduces a new symmetry between bosons and fermions. As a result many properties of the SM can be calculated and scales are constrained better. But so far no evidence was found and the phase space left to explore SUSY can not explain several of the problems of the SM.

One of the prominent simplifications is the introduction of a Grand Unified Theory (GUT) [50]. The running of the coupling constants suggests that the electroweak and QCD scale could be unified at high energies. But no successful theory was developed yet.

As all these new model ideas cannot be checked at the same time, for this thesis models are considered that can be probed at the LHC with a single lepton in the final state. In particular, in the next section models will be discussed which can produce a heavy charged gauge boson, decaying into τ+E_T^{miss}. A phenomenological model will be introduced, which serves as a benchmark model for this analysis. In addition one specific model will be discussed, which can explain the mass differences of the third fermion family to the first two generations. Last but

not least various approaches to interpret the events as results from Dark Matter produciton will be discussed, which will finally be investigated in the $e/\mu + E_T^{miss}$ final state.

2.3. A Heavy Gauge Boson W' in Theories Beyond the SM

One of the most common features of BSM theories is the prediction of new heavy particles. There are several known methods to include new particles in a consistent way into the SM. The main focus of this thesis is on a heavy charged vector boson, generally referred to as W'. In most extensions also a neutral vector boson is predicted, here referred to as Z'. In order to position the search in the model landscape that the LHC is searching for a brief summary of different W' model types will be given and then only the models used in the following analysis will be described.

As mentioned before **SUSY** is one of the most popular extensions of the SM. It also generates additional particles, but the new particles are shifted w.r.t. the SM particles by spin $\frac{1}{2}$. New charged heavy bosons would therefore be scalars. One of the assumptions for SUSY is R-parity conversation, which as a consequence allows SUSY particles to be produced only in pairs. Nonetheless in addition to the dedicated searches [51], also the lepton+E_T^{miss} signature can play a role, as shown in [52]. If R-parity is conserved a W' candidate could be the super partner of the charged leptons. But strong limits on the R-parity violating couplings for these particles exist.

A similar approach can be summarized under the label **extra dimension models**. The idea is that additional extra dimensions could extend our known four dimensions. These additional dimensions are compactified and therefore not accessible at low energies. Depending on the model not all particles are able to travel in these extra dimensions. A common consequence of this compactification is the creation of so called Kaluza-Klein modes, which are excitations of the SM particles in the extra dimensions. Here a heavy W-boson excitation can be predicted [53, 54].

In light of the recent Higgs discovery models with a **composite** structure of the **Higgs** got more attention. These models mainly solve the problem that the Higgs mass, and therefore the scale of the weak interaction has to be constrained to the specific measured value, to avoid fine tuning. Therefore new strongly coupled particles are proposed, which have a natural energy scale. These models either directly predict a heavy charged spin-1 boson or a heavy scalar charged boson [55].

Other models are directly motivated by the group structure that has been observed and for example propose a left-right symmetry, which adds to the $SU(2)_L$ a $SU(2)_R$ or directly a larger group such as $SU(5)$ or higher[10]. The structure of these groups require additional gauge bosons and therefore predict a W' boson.

In order to have a reference model that provides a generic W' the sequential standard model [56] was chosen that will be described in the next section.

[10]The original $SU(5)$ [50] grand unified theory was proven to be incorrect, because it allows proton decays.

2.3.1. The Sequential Standard Model

The Sequential Standard Model (SSM) [56] assumes that a carbon copy of the SM W- and Z-boson with higher masses exists. Because not all details are specified with this assumption, several additional assumptions and properties will be used in the following:

The branching fraction and width of the W'-boson has to be modified with respect to the SM. Since the W' is heavier than the top and bottom quarks, the decay into tb is allowed. The same could happen for the WZ decay, but the coupling of WZ is very model dependent, therefore assumed to vanish. With these assumptions the decay width is [57, 58]:

$$\Gamma_{W'\to\bar{f}f'} = M_{W'} \cdot \frac{g^2 \cdot C_{\bar{f}f'}}{2 \cdot 48\pi} \cdot F(\frac{M_{\bar{f}}}{M_{W'}}, \frac{M_{f'}}{M_{W'}}) \tag{2.15}$$

The factor $C_{\bar{f}f'}$ denotes the colour factor of the fermions and is 1 for leptons and 3 for quarks. The function $F(x_1, x_2)$ is a correction to the leading order width from NNLO calculations and has the following structure [57, 58]:

$$F(x_1, x_2) = (2 - x_1^2 - x_2^2 - (x_1^2 - x_2^2)^2) \cdot \sqrt{(1 - (x_1 + x_2)^2) \cdot (1 - (x_1 - x_2)^2)} \tag{2.16}$$

The total width of the W'-boson is

$$\Gamma_{W'\to\bar{f}f'} = M_{W'} \cdot \frac{g^2}{2 \cdot 48\pi} \cdot (18 + 3 \cdot F(\frac{M_t}{M_{W'}}, \frac{M_b}{M_{W'}})) \tag{2.17}$$

For high masses the influence of the fermion masses can be neglected and expression 2.17 can be approximated as:

$$\Gamma_{W'} = \frac{4}{3}\frac{M_{W'}}{M_W}\Gamma_W \tag{2.18}$$

If the same coupling structure as in the SM is present there will be interference between the W- and W'-boson [59, 60]. The differential cross section is:

$$\frac{d\sigma}{d\Omega} \cdot s \propto \left(\frac{g_{SM}^2}{s - M_W^2}\right)^2 + \left(\frac{g_{W'}^2}{s - M_{W'}^2}\right)^2 \pm 2\left(\frac{g_{SM}^2}{s - M_W^2} \cdot \frac{g_{W'}^2}{s - M_{W'}^2}\right), \tag{2.19}$$

where the decay width is neglected. One can see that the sign of the interference depends on the last term and in general the term will be added if the coupling to the initial and final particles is the same. If the coupling is different, e.g. an opposite sign between leptons and quarks, the last term will be subtracted. The sign of the last term is also depending on the value of s with respect to M_W^2 and $M_{W'}^2$. The effect was studied in [57, 6] and was found to be less than 10%. As we will see for the τ+E_T^{miss} final state, one can not resolve the regions of constructive and destructive interference. Therefore the effect is reduced to a cross section modification and will be ignored in the following.

For the τ+E_T^{miss} final state different variables can be used to separate the signal from data. The best signal to background separation can be achieved using the transverse mass of the event. It is defined as:

$$M_T = \sqrt{2 \cdot p_T^\tau \cdot E_T^{miss} \cdot (1 - \cos\Delta\phi_{\tau, E_T^{miss}})} \tag{2.20}$$

The momentum p_T^τ is the measurable transverse momentum of the hadronic tau decay τ^{had}. The missing transverse energy E_T^{miss} is comprised of all particles, which can not be detected with a LHC detector. For a hadronically decaying tau, E_T^{miss} is formed from the two neutrinos. One neutrino from the W' decay and one neutrino from the hadronic tau decay. The distribution of a W' signal for different decay stages of the W' is shown in Figure 2.5. Two W' masses are shown 1 TeV (dark red) and 3 TeV (blue). The decay stages illustrate the loss of information due to the neutrinos. The solid line shows the invariant mass of the undecayed tau and the neutrino. The dashed lines show the transverse mass of the undecayed tau and the neutrino, which uses in contrast to the invariant mass only transverse information. The dotted lines show the M_T distribution as it is measurable in the detector with the two neutrinos combined to E_T^{miss} and p_T^τ from the τ^{had}.

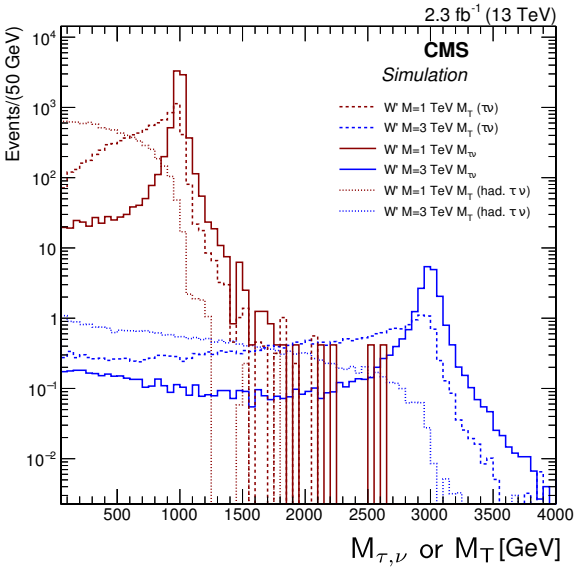


Figure 2.5.: The signal shape of the $W' \to \tau\nu$ final state with different reconstruction stages.

To analyse the regions that are discussed previously in Equation 2.7 for the W boson, these distributions will be analysed in more detail. For the solid line, which shows the invariant mass of the neutrino and the tau, a clear Breit-Wigner resonance is visible. For the 3 TeV mass it is already visible that the boson has a significant contribution below the peak, but masses below the resonance also contribute. This is due to the available centre of mass energy, which is the partonic centre of mass energy. For processes with a low energy transfer, the W' signal will appear like a contact interaction. The dashed line is the transverse mass, of the undecayed tau and the neutrino from the W decay ($M_T(\tau, \nu)$). The shape changed to a Jacobian peak structure. A peak is still visible at the W' mass, but due to missing information of the z-direction, the values are spread to lower energies. For the 3 TeV signal due to its already high off-shell production, an almost rectangular shape is visible. For the dotted line the M_T distribution of the full hadronicly decayed tau is shown ($M_T(\tau^{had}, \nu\nu)$). The signal has no visible peak strucute left. Instead the signal resembles a rectangular shape with values between the W' mass and 0 GeV. At the W' mass the rectangular shape is dropping off, only barely reaching the W' mass. For the 3 TeV W' the off-shell production distorts the signal even further, giving a falling distribution, until shortly before the W' mass the signal dies of completely.

The cross section of the process $f_i \bar{f}_j \to W' \to f_{i'} \bar{f}_{j'}$ is analogous to the W-boson production and decay shown in Equation 2.7. The W-boson mass and width just have to be substituted for the W' mass and width.

After the discussion of the experimental setup and the reconstruction, we will see how such a signal can be detected with the CMS detector.

One may note that the coupling is set arbitrarily to the same strength as the electroweak coupling. If this is changed, but the universality is kept, the cross section at the W' peak is still the same, as can be seen from Equation 2.7. The main influence is then given by the width of the W'. This changes, when the W' couples non universally to the fermions, because the peak cross section is no longer independent from the coupling constant.

2.3.2. Non Universal W' Model

As discussed, one issue for the SM is the mass difference of the third fermion generation to the first two generations. Or phrased differently: Why has the top quark a mass of 173 GeV and the up quark a few MeV?

A possible way to explain the difference is the introduction of a Higgs doublet. The Two Higgs Doublet Models (2HDM) can have many manifestations. In contrast to the common 2HDM, here the second Higgs doublet is not used to generate masses for the up-type quarks and the neutrinos, but it only couples to the third fermion generation. In order to allow this division in the Higgs sector, a second $SU(2)_L$ group has to be introduced. This would not only solve the mass hierarchy of the fermions, but also give us new heavy (charged) bosons. The structure of the electroweak sector therefore is $SU(2)_l \times SU(2)_h \times U(1)_Y$, where the $SU(2)_l$ provides the interactions for the first two (light) generations and the $SU(2)_h$ only for the third (heavy) generation. Due to this structure the model is generally referred to as non-universal model[11]. Since the non universality is an important aspect for this thesis, it will be referred to as non-universal gauge interaction model (NUGIM). In order to understand the differences to the SM, it will be reviewed here shortly. For a full description of the model see [61]. The phenomenology of the W' in this model is summarized in [62].

Similar to the electroweak theory described in Section 2.1.2, the observable bosons are a mixture of the gauge bosons, where the symmetry is broken at an energy scale u. The mixing angle of the extended gauge group will be referred to as θ_E. The couplings are given by $g_l = \frac{g}{\cos\theta_E}$ and $g_h = \frac{g}{\sin\theta_E}$. For the Higgs sector two scalar particles emerge, ϕ_l and ϕ_h. They have two vacuum expectation values v_h and v_l and mix as in other 2HDM or SUSY, with the parameter $\tan\beta = \frac{v_h}{v_l}$. A large value of $\tan\beta$ would explain the heavy third generation and will be assumed in the following.

The mass of the gauge bosons can be expressed as a series of $\epsilon = \frac{v}{u}$, where v is the SM vacuum

[11] In literature these models are also called $G(221)$ - following the group structure - or topflavor model.

expectation value and u is the scale of the new group mixing [61]:

$$M_W^2 = \frac{1}{2}g^2v^2 - \frac{1}{2}g^2v^2(\sin^2\beta - \sin^2\theta_E)^2\epsilon^2 + \mathcal{O}(\epsilon^4) \tag{2.21}$$

$$M_{W'}^2 = \frac{1}{2}g^2u^2\frac{1}{\sin^2\theta_E^2\cos^2\theta_E^2}[1 + (\sin^2\beta - 2\sin^2\beta\sin^2\theta_E + \sin^4\theta_E)\epsilon] + \epsilon^2 + \mathcal{O}(\epsilon^4) \tag{2.22}$$

$$M_Z^2 = \frac{1}{2}g^2v^2\frac{1}{\cos^2\theta_W} - \frac{1}{2}g^2v^2\frac{1}{\cos^2\theta_W}(\sin^2\beta - \sin^2\theta_E)^2\epsilon^2 + \mathcal{O}(\epsilon^4) \tag{2.23}$$

$$M_{Z'}^2 = \frac{1}{2}g^2u^2\frac{1}{\sin^2\theta_E^2\cos^2\theta_E^2}[1 + (\sin^2\beta - 2\sin^2\beta\sin^2\theta_E + \sin^4\theta_E)\epsilon] + \epsilon^2 + \mathcal{O}(\epsilon^4) \tag{2.24}$$

$$\tag{2.25}$$

One can see that at the zeroth order in ϵ the SM masses are reproduced, the W' and Z' masses are the same and the masses are independent of $\tan\beta$.

Since the Higgs sector has no significant impact on the phenomenology of the W', it will be ignored in the following (high $\tan\beta$ region). The vertex Feynman rules for the W' processes are shown in Table 2.3.

The resulting total width of the W' will then be [62]:

$$\Gamma_{W'} = \Gamma_{Z'} = \Gamma_{W'}^{SSM} \times \frac{(4 + \frac{1}{4})\cot^2\theta_E + 8\tan^2\theta_E}{12 + \frac{1}{4}} \tag{2.26}$$

vertex	Feynman rule
$\bar{\tau}\nu W'^-, \tau\bar{\nu}W'^+$	$-\frac{ig}{\sqrt{2}}\cot\theta_E\gamma^\mu P_L$
$\bar{u}dW'^+, u\bar{d}W'^+$	$\frac{ig}{\sqrt{2}}\tan\theta_E V_{ij}\gamma^\mu P_L$
$\bar{d}uW'^-, d\bar{u}W'^-$	$\frac{ig}{\sqrt{2}}\tan\theta_E V_{ij}^\dagger\gamma^\mu P_L$

Table 2.3.: Table of the relevant vertex Feynman rules for the NUGIM W' [62].

For the third generation the branching fraction (B) behaves like $B = B_{SSM}\cdot\cot^2\theta_E\cdot\frac{\Gamma_{SSM}}{\Gamma_{W'}}$ and for the light fermions the behaviour is $B = B_{SSM}\cdot\tan^2\theta_E\cdot\frac{\Gamma_{SSM}}{\Gamma_{W'}}$. The total cross section times branching fraction for the third generation fermions at the LHC is then $\sigma\times B \propto B_{SSM}\cdot\frac{\Gamma_{W'}^{SSM}}{\Gamma_{W'}}$, while for the light fermions the cross section times branching fraction is $\sigma\times B \propto B_{SSM}\cdot\tan^2\theta_E\cdot\frac{\Gamma_{W'}^{SSM}}{\Gamma_{W'}}$. The total width and branching fractions are summarized in Figure 2.6. One can see that for values of $\cot\theta_E < 1$ the decay into light quarks is dominant, while for larger $\cot\theta_E$ the third family is the preferred decay of the W'. For large $\cot\theta_E$ the width of the W' also increases drastically.

As one can see, the width of the W' strongly depends on $\cot\theta_E$. In order to study the signal the effect of the width on the observable distributions has to be studied. The effect on the invariant and transverse mass is shown in Figure 2.7.

Due to the different dependence of the width for the production and decay the cross section of the NUGIM W' is different even at the W' mass peak.

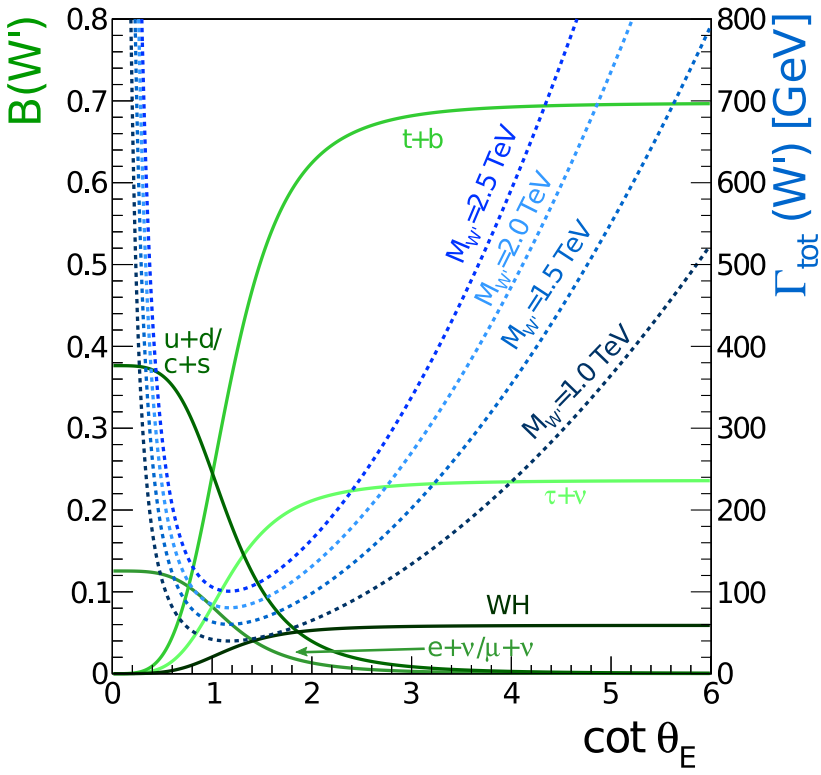


Figure 2.6.: Branching fractions (left-hand scale and solid lines) and total width (right-hand scale and dotted lines) for W' decays in the NUGIM. For $\cot\theta_E = 1$ the values are the same as those in the SSM, rescaled to accommodate the $W' \to Wh$ decay channel.

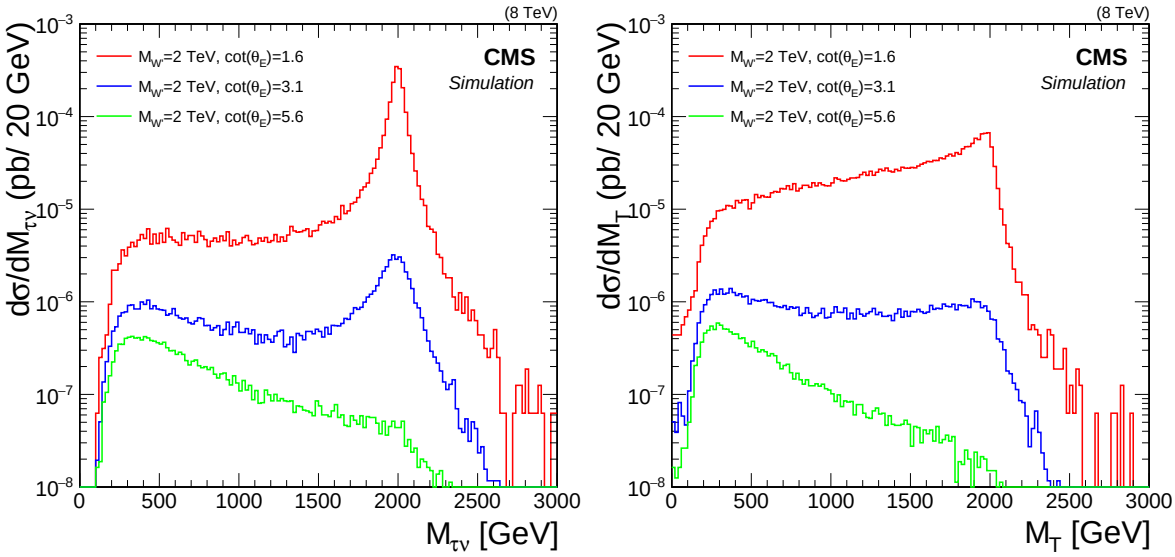


Figure 2.7.: For a W' mass of 2 TeV the dependence of the width with respect to various $\cot\theta_E$ values is shown. In the left plot the invariant mass of the undecayed tau and the neutrino is shown, where in the right plot the transverse mass of the undecayed tau and the neutrino is shown.

2.4. Dark Matter as Non Resonant Signal

One of the keystones of the cosmological standard model ΛCDM [48] is Dark Matter (DM). It was first discovered in the 1930s in spiral galaxies [63] where the rotation velocity of the outer stars could only be explained by introducing non luminous "dark" matter. After decades of investigating rotation curves, in the last two decades other observations confirmed the existence of DM in our universe. To the number of observations one can count the infrared and gravitational lensing measurements [64], the detailed measurement of the cosmic microwave background in our universe [65] or structure formation simulations for galaxies. All these measurements formed a more or less clear picture of the cosmological properties of DM. But there are still many open questions, like the DM density distributions in galaxies or the interaction properties. For a detailed discussion of the status of DM measurements, see [48]. The particle description of DM is unclear up to date. There are a few properties that have to be fulfilled for a particle to be a good DM candidate, that can be summarized as follows [66]:

- Does it match the appropriate dark matter relic density, measured in the cosmic microwave background?
- Is it cold?
- Is it neutral?
- Is it consistent with Big Bang nucleosynthesis?
- Does it leave stellar evolution unchanged?
- Is it compatible with constraints on self-interactions?
- Is it consistent with direct DM searches?
- Is it compatible with gamma-ray constraints?
- Is it compatible with other astrophysical bounds?
- Can it be probed experimentally?

While this list is also a list of the known properties for DM candidates, it does not constrain the models used for particle DM searches much. One of the standard DM candidates is the lightest stable supersymmetric particle and a lot of phenomenology studies have been oriented towards this kind of DM. However, so far no evidence for SUSY has been found. Therefore other possible particle DM candidates are considered. Nonetheless the particle candidate for DM will be denoted χ in the following, using the standard SUSY notation. The searches for DM can be categorized into three search strategies.

The only experiments with the primary purpose to find DM, are the **direct detection** searches [67]. The cosmological DM could hit a nucleus and the recoil energies can be measured. These experiments vary in the material used as target and are dependent on the density and velocity descriptions of the cosmological models. Due to their setup they constrain the mass of a potential DM candidate directly. The assumption is that the $\chi N \leftrightarrow N\chi$ interaction can be described by an effective field theory model where the details of a potential mediator cannot be resolved. A complete list of effective operators[12] can be found in [68]. The most

[12]Of which not all can be transferred into a full theory.

important difference that can be made is if the effective operator is spin dependent (includes a γ^5 operator) or not. Because if the operator is spin dependent the DM couples to the spin of the nucleus whereas a spin independent operator couples to the hole nucleus (heavier nuclei provide a better sensitivity). These kinds of experiments have been very successful in the past and aim to reach the irreducible neutrino background in the next decades.

The **indirect detection** experiments [69] consider the annihilation of DM. The idea is that DM could be gravitationally trapped in the centre of the galaxy or the sun, therefore an increased rate of messenger particles such as ν, γ or p^+ could be observed from these regions in comparison to other regions. While these searches found some interesting candidates in the last years [70, 71], the background to the searches is non trivial and we must learn more about our galaxy before conclusions can be drawn.

The third possibility to search is the **production at colliders**. Aside from the SUSY or extra dimensional searches a new kind of dedicated collider searches have been performed in the last years. One can use a mono-X signature where the DM is produced similar to the direct searches, and an additional initial state radiation particle X is used to tag and identify the event. In the following we will investigate the mono-lepton channel[13] and a short review of the used models will be given.

The first approach to obtain a model which is comparable with the direct and indirect detection experiments, was the use of the same effective field theory in the interaction vertex:

$$\text{Spin-independent vector coupling: (V)} \quad \frac{1}{\Lambda^2} \quad \bar{\chi}\gamma^\mu\chi \quad \xi_i\bar{q}_i\gamma_\mu q_i \tag{2.27}$$

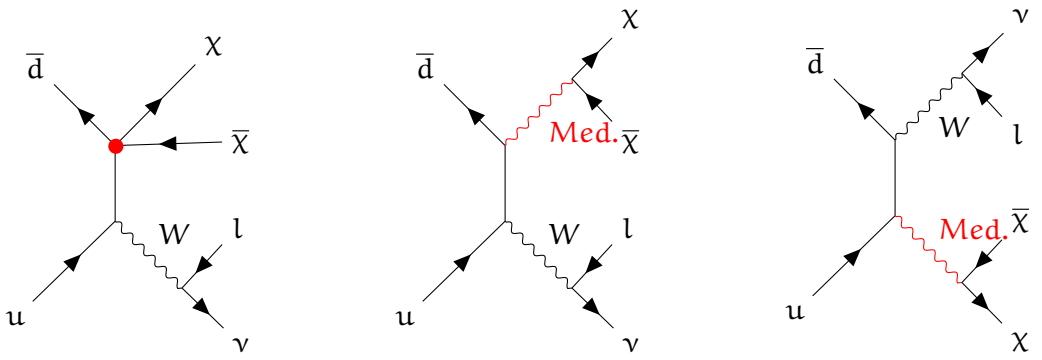
$$\text{Spin-dependent axial-vector coupling: (AV)} \quad \frac{1}{\Lambda^2} \quad \bar{\chi}\gamma^\mu\gamma^5\chi \quad \xi_i\bar{q}_i\gamma_\mu\gamma^5 q_i \tag{2.28}$$

These effective field interactions, as shown in Figure 2.8a, provide a full description of the event and were used in several CMS and ATLAS publications [72, 73, 74, 75, 76, 77, 78]. But the effective scale that can be probed at the LHC is at the order of 1 TeV, and the momentum transfer at the vertex can be higher than the probed effective scale. This means that a potential mediator could be resolved at the LHC. As a result the model depicted in Figure 2.8b, with a resolved mediator was studied. It became clear that in fact the LHC is not sensitive to the DM mass, but to the mass of a potential mediator. Having a specified mediator the properties of the mediator also have to be specified. The first approach would be to have a SSM like mediator, where the Z' would decay into $\chi\bar{\chi}$. This however, would have been detected in the decays of leptons or jets. The mediator properties are discussed in [79, 13]. It was concluded to use a leptophobic boson mediator as reference model.

For the mono-lepton channel another theoretically interesting property arose from the structure of the diagram in Figure 2.8b and 2.8c. While in Figure 2.8b the mediator couples to the down quark, in Figure 2.8c the mediator couples to an up quark. The diagrams have the same initial and final state, therefore they interfere. Due to the evolution of DM signal models in the mono-lepton channel a brief rundown of the analyses will be given here.

The interference effect (absorbed as a relative factor ξ, between the diagrams 2.8b and 2.8c) first was noticed in 2012 by [80] which used the model unspecific limit, as published in the PAS [81].

[13] The correct name would be mono-charged lepton channel, but in the tradition of many experiments the neutrinos are neglected.



(a) effective DM production (b) mediator DM production (down quark coupling) (c) mediator DM production (up quark coupling)

Figure 2.8.: Feynman diagrams for an effective and mediator produced DM.

At the start of this thesis I worked on a dedicated analysis of the proposed model, which was published in 2013 [5], and was consequently turned into a paper published in early 2015 [6]. After this publication the EFT model was abandoned in favour of a simplified model, which has a resolved mediator. This was discussed with others in the "Dark Matter Forum", comprised of members from ATLAS, CMS and theorists, which led to the publication [79] in 2015. Shortly before the conclusion of this forum's report theorists noticed that the diagram 2.8a is not gauge invariant and unitarity violating [82]. In following publications [83, 84], alternative models were proposed.

The interference not only affects the interpretation of the results, but also the search strategies, since the interference affects the steepness of the M_T distribution as shown in Figure 2.9. The

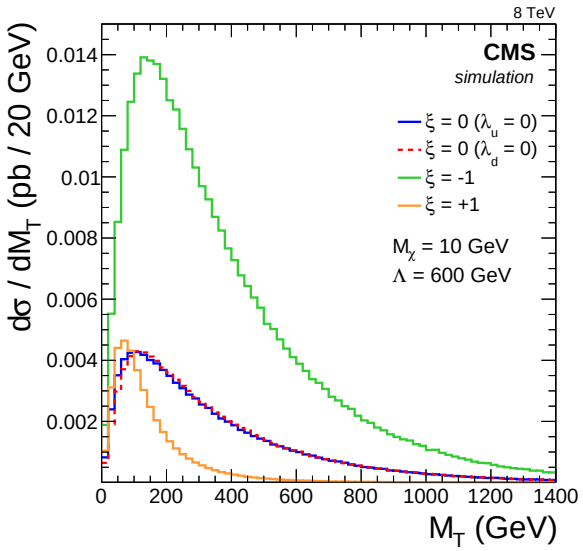


Figure 2.9.: The M_T distributions of different interference scenarios for the EFT model. The parameter ξ, is a relative factor between the up-type and down-type coupling DM. Which of the quark couplings is set to 0 for $\xi = 0$ is irrelevant for the shape.

potential different coupling of a Z' with respect to the quarks is discussed in detail in [84]. The conclusion is that the unitarity violating terms are a sign for the incompleteness of the

theory and can be resolved in three different ways:

- restrict the coupling to $g_u = g_d$ for the Z'
- $g_u \neq g_d$ and additional interactions like for the SM Z
- introduce a unitarity violation scale for the coupling

The divergency is similar to the cross section divergency observed in the SM for the $WW \to WW$ process, which can only be resolved by the introduction of a $WW \to h \to WW$ process, giving strong hints for the Higgs boson, before it's discovery.

Due to the large number of possibilities and the different consequences on the shape, the standard simplified Z' approach will be used in the following thesis. As interference parameter the three cases of $\xi = -1, 0, 1$ will be kept as the extreme cases for slope in M_T. It must be noted here that the cross section predictions for the interference cases $\xi = -1$ and $\xi = 0$ are too high to be expanded to a full theory, but the slope difference can still be valid for a complete theory. In order to provide a model unspecific result, which can be interpreted for any initial state radiated invisible particle, a new signal parametrization will be introduced for the $e/\mu + E_T^{miss}$ final state.

A detailed discussion of the EFT results is published in [6, 5] and [9], while we studied the simplified models in [10, 11, 12, 13].

3. Experimental Setup

This chapter summarizes the experimental setup, which is needed to perform this analysis. First a short overview of the relevant machine parameters and conventions is provided and then a description of the CMS detector is given.

3.1. The Large Hadron Collider

The LHC is built in the tunnel originally constructed for the LEP accelerator[1]. Four large experiments plus a few smaller ones are positioned around the LHC. The two general purpose experiments are ATLAS and CMS, and the specialized experiments are LHCb and ALICE. While ATLAS and CMS are built at new collision points, ALICE and LHCb are built in caverns, which housed the L3 and DELPHI LEP experiments. The LHC [85] is designed to deliver an instantaneous luminosity of $\mathcal{L} = 10^{34}\,\mathrm{cm}^2\mathrm{s}$ at a proton proton centre of mass energy $\sqrt{s} = 14\,\mathrm{TeV}$ in the 27 km long LHC ring. The magnetic field needed to bend the protons is 8.33 T, which can be achieved by superconducting magnets. This has some consequences for the design of the LHC. First, the LHC is to first order a storage ring and to second order an accelerator while the machine injection energy of 450 GeV is ramped to 3.5 TeV. Most of the beam-time is at a constant energy (up to 35 h 23 min [86]). Secondly, the LHC is not easily accessible, for repairs or corrections. Often a complete section has to be warmed up, therefore the operation of such a machine is expensive.

This thesis covers two data sets of CMS data. The first collected in the year 2012 at $\sqrt{s} = 8\,\mathrm{TeV}$ and the second in 2015 at $\sqrt{s} = 13\,\mathrm{TeV}$. The years 2013 and 2014 were used to upgrade the collider and the experiments. For the purpose of this thesis the data collected in 2012 is called Run I and the data collected in 2015 is called Run II. Previous data taking periods in 2010 and 2011 at $\sqrt{s} = 7\,\mathrm{TeV}$ are ignored due to their lower centre of mass energy and less integrated luminosity.

In 2012 the instantaneous luminosity was $\mathcal{L} \gtrsim 0.7 \cdot 10^{34}\,\mathrm{cm}^2\mathrm{s}$ at $\sqrt{s} = 8\,\mathrm{TeV}$ [87], which is very close to the design luminosity. But the bunch spacing - the distance between two proton bunches in the beam - was 50 ns compared to the design bunch spacing of 25 ns. Because the spacing between two bunches was larger the number of protons per bunch could be $1.6 \cdot 10^{11}$. It superceded the design value of $1.15 \cdot 10^{11}$ protons per bunch. The total number of bunches per ring was 1374 (design 2808). This leads to a mean number of interactions per bunch crossing (pileup) events of up to 35 per bunch crossing. The delivered integrated luminosity was $23\,\mathrm{fb}^{-1}$. Of this luminosity $20\,\mathrm{fb}^{-1}$ were recorded with all detector components and could be used for this analysis.

[1] With the LHC as possible extension in mind.

In 2015 the instantaneous luminosity was $\mathcal{L} = 0.5 \cdot 10^{34}\,\mathrm{cm}^2\mathrm{s}$ at $\sqrt{s} = 13\,\mathrm{TeV}$ [87]. During the first part of the data taking period the LHC had a bunch spacing of 50 ns, but this was switched to 25 ns, when machine behaviour was understood. The number of protons per bunch was reduced to $1.5 \cdot 10^{11}$, therefore the number of bunches per ring could be raised to 2244. Due to the smaller bunch spacing and the lower number of protons per bunch the mean number of interactions per bunch crossing was about 15. The delivered integrated luminosity at 25 ns bunch spacing was $4\,\mathrm{fb}^{-1}$. Of this luminosity $2.3\,\mathrm{fb}^{-1}$ could be used for this analysis, because all detector components worked properly.

3.2. The CMS Detector

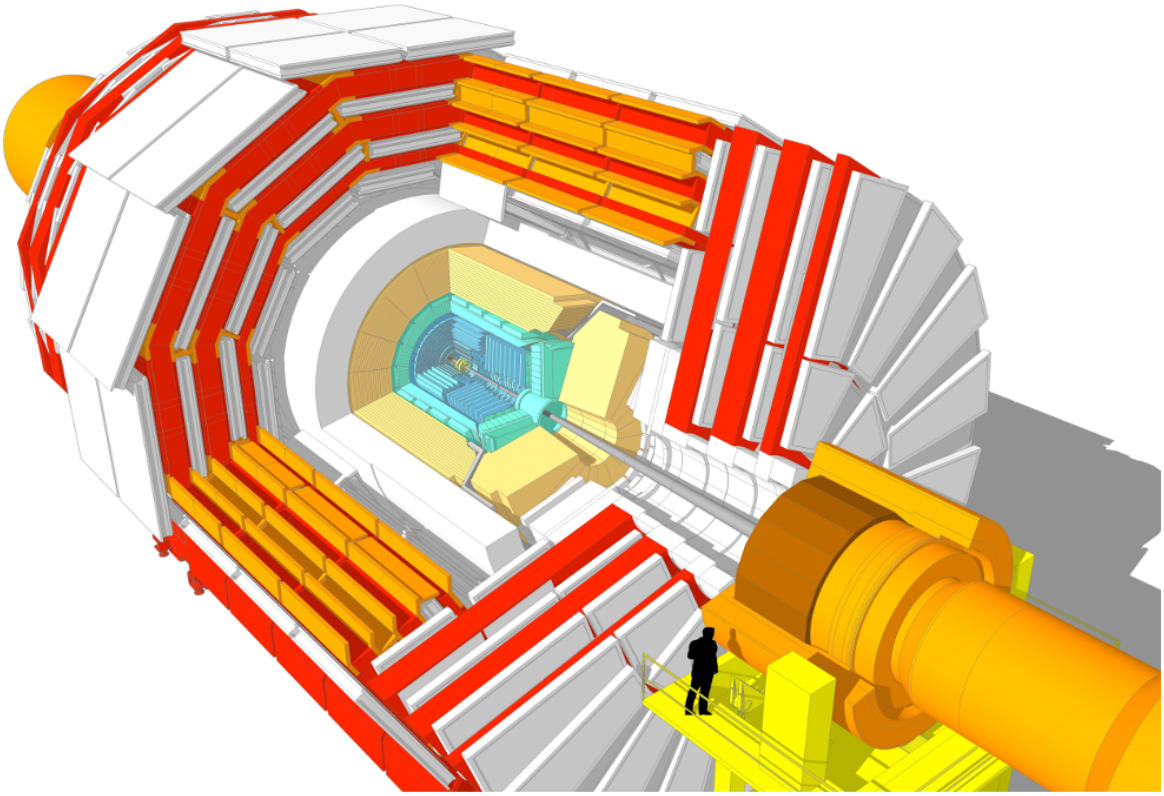


Figure 3.1.: The CMS Detector [88]

The Compact Muon Solenoid (CMS) detector, as show in Figure 3.1 is located 100 m underground at point 5 of the LHC. It is 28.7 m long and 15 m in diameter. This is small for a general purpose experiment at the LHC. The compactness of the detector has several advantages. First, most of the active detector material can be within the solenoid. Second, the underground space could be split into two caverns, one for the detector and one for the support infrastructure. This allows maintenance on most of the vital systems during operations. The third advantage of the compact design is that it allows a wheel structure of the detector. The CMS detector consists of five wheels and two endcaps, which can be moved if access is needed. The 14 000 ton detector is built to cover as much of the 4π surface around the interaction as possible. The compactness is possible because of the 3.8 T superconducting solenoid magnet, which enfolds over $70\,\mathrm{m}^3$

detector material. Only the muon system is built outside of the magnet in the return yoke of the magnetic field.

As a general purpose detector the design of CMS has to perform well for two detector aspects:

- good charged particle reconstruction
- good neutral particle reconstruction

While the first can be achieved with a high magnetic field and a good tracker, the second can be achieved with a good calorimeter. One of the challenges at the CMS construction was to combine both. The CMS detector has a high resolution tracking system with a low material budget and a fast calorimeter with a good energy resolution, which is not shielded by material of the magnet. Finally the muon system provides a good muon resolution up to high energies.

In the following the individual detector components are described briefly to allow the reader to connect the abstract physics results to the actual measurements.

3.2.1. Coordinate System and Detector Quantities

The CMS detector is a cylinder build around the interaction point. Therefore cylindrical coordinates are used, with the interaction point as point of origin. For every event the primary vertex is defined as the reconstructed vertex with the highest $\sum p_T$, but not all particles originate from the primary vertex, therefore coordinate system ϕ or θ are defined in can differ for different particles. For detector components and neutral particles the nominal detector centre is used as coordinate center. The z-axis is positive in the direction of the Jura mountains. In order to allow boost independent angle distances, θ is translated into the pseudo rapidity $\eta = -\ln(\tan\frac{\theta}{2})$. Also an angular distance is often used which is defined as $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$. Due to the unknown momentum in the z-direction, transverse quantities are used like E_T or p_T. These are defined corresponding to $p_T^2 = p_x^2 + p_y^2$. The direction of an energy measurement in CMS is defined by the position of the energy measurement in the detector and the CMS centre.

3.3. The Solenoid

The CMS superconducting solenoid [89] is a key component of the detector. The 3.8 T magnetic field encloses a cylinder of 3 m radius and 12.5 m length. This surrounds all detector components except a tail catcher of the hadronic calorimeter and the muon system. The magnet is operated at a temperature of 4.5 K, with liquid helium. The huge magnetic field is returned on the outside of the coil with 10 000 t of iron, in which the muon system is embedded.

3.4. The Inner Tracker

The innermost detector part of CMS is the tracking system [90]. It is a silicon based tracker, with a high resolution pixel core and a strip design on the outside. The pixel detector ranges from 4.4 cm to 10.2 cm in radius, where the strip detector extends from the pixel detector to a radius of 1.2 m. The layout is shown in Figure 3.2 (left).

The Pixel cell size is $100 \times 150\,\mu m^2$ and the analogue pulse height readout and the combination of the three pixel layers allows a similar resolution in $r-\phi$ and z directions of 15-20 µm [91].

Due to the larger volume and lower occupancy the strip detector is built of $6 \times 12\,cm^2$ to $10 \times 9\,cm^2$ rectangular modules, this are slightly tilted to each other, which allows a measurement of the θ angle.

The tracker endcaps (TEC) are constructed from petals arranged to form the forward cylinder of the tracker. The tracker endcap ranges from $|z| = 1.2\,m$ to $|z| = 2.8\,m$ in the forward direction. The inner diameter allows the exchange of the Pixel detector and provides coverage up to $|\eta| = 2.5$.

The performance of the tracker can for example be determined width cosmic muons, since they can be measured twice in the detector. Figure 3.2 (right) shows the tracker and muon system resolution[2] for muons up to the TeV scale [92]. For an updated (not public) version of the plot, see [93]. The low momentum uncertainty of muons can be aproximated by:

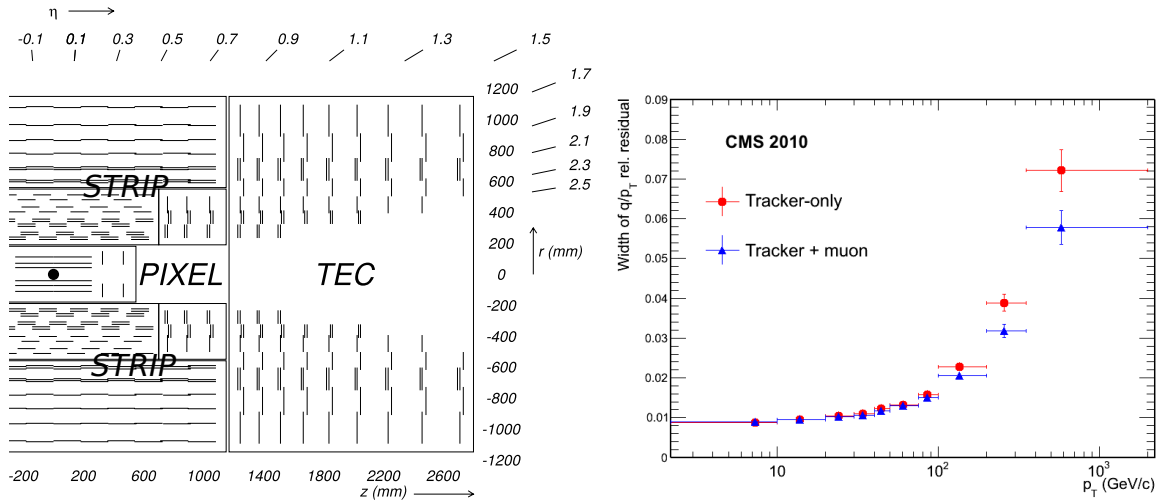


Figure 3.2.: On the left side the layout of the inner tracking system is shown, modified from [88]. The muon resolution with tracker and muon system from cosmic muons is shown on the right [92].

$$\left(\frac{\sigma_{Track}}{p_T}\right)^2 = \left(7\% \cdot \frac{p_T}{1\,TeV}\right)^2 + (0.9\%)^2 \tag{3.1}$$

One can see that the tracker has a resolution, which is at the 1% level for low p_T.

[2]The plot 3.2 (right) shows the width of the $1/p_T$ distribution. By construction of the quantity the charge q is given in the nominator

3.5. Calorimeters

Most particles that are produced in the hard interaction can be stopped in the detector[3]. In order to stop a particle the material budget of the detectors has to be increased with respect to the tracker and the deposited energy has to be measured. This is done in the calorimeters. The calorimeter is for the main part inside the solenoid and is subdivided into two systems, the inner electromagnetic calorimeter (ECAL) and the hadron calorimeter (HCAL). The arrangement of the calorimeters can be seen in Figure 3.3.

3.5.1. Electromagnetic Calorimeter

The ECAL [94] is comprised of homogeneous lead-tungstate ($PbWO_4$) crystals. It is divided into barrel (EB) and endcap region (EE). The crystals have a radiation length of 0.89 cm and are arranged to point to the detector centre. The ECAL covers a $|\eta|$ range up to 3. The crystals have a granularity which corresponds to the Molière radius of 2.2 cm. This allows the measurement of an electron with a high precision in a single crystal. The transition region ($1.4442 < |\eta| < 1.566$) between EB and EB is often excluded for objects that rely on a good ECAL resolution.

The resolution of the ECAL could be determined in test beams to be:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E/\,\mathrm{GeV}}}\right)^2 + \left(\frac{12\%}{E/\,\mathrm{GeV}}\right)^2 + (0.3\%)^2, \tag{3.2}$$

where 2.8% is the stochastic term from the electromagnetic shower, 12% is the noise from the readout and 0.3% is the systematic uncertainty due to the calibration setup. For high energies the ECAL resolution is dominated by the systematic uncertainty, while for low energies the stochastic and noise terms are dominant.

3.5.2. Hadronic Calorimeter

The HCAL [96] is built to provide the best possible 4π coverage of the detector and measure all hardonic showers before the muon system. Therefore it increases the number of hadronic interaction length λ_I of the ECAL in front of the HCAL by $10 - 15\lambda_I$ (including the magnet and the HO). The HCAL is divided into four subsystems: hadron barrel (HB), endcap (HE), outer (HO) and forward (HF) calorimeters. In total the HCAL covers a range of $|\eta| < 5$ to the very forward direction, see Figure 3.3. The HB and HE are sampling calorimeters, constructed from brass, steel and plastic scintillators. They enclose the interaction point up to $|\eta| = 3$. The HO is a tail catcher in the barrel region, which is positioned outside the solenoid. The central detector wheel has two HO layers of scintillators, before and after the steel of the innermost ring of the return yoke, where the other wheels have one layer directly behind the magnet. The HF is a very forward calorimeter. The forward region has the harshest radiation conditions in the detector. Therefore quartz fibres where embedded into a steel absorber, using the Cherenkov light of the passing particles to measure the energy.

[3]From the detectable particles only the muons leave the detector.

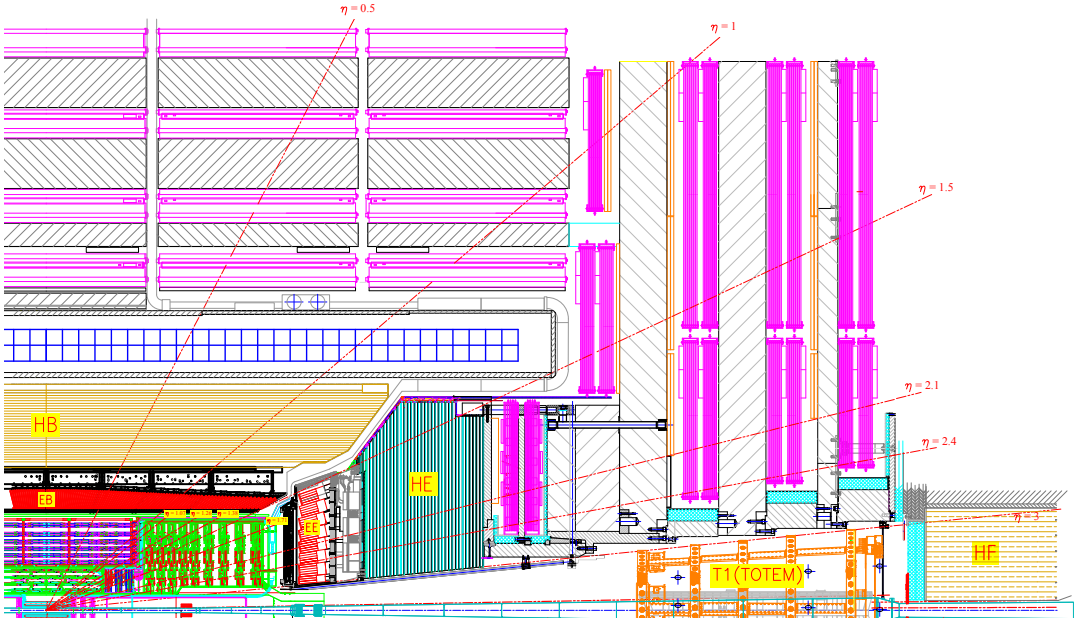


Figure 3.3.: The longitudinal layout of the CMS detector, modified from [95].

Due to the sampling structure and the nature of hadronic showers the stand alone energy resolution for $\pi^{\pm}$ is worse than the ECAL resolution [88]:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{138\%}{\sqrt{E/\,\mathrm{GeV}}}\right)^2 + (13\%)^2. \tag{3.3}$$

But if the measurement is combined with the ECAL, the resolution for hadronic showers improves to:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{80\%}{\sqrt{E/\,\mathrm{GeV}}}\right)^2 + (8\%)^2. \tag{3.4}$$

The detailed resolution depends on the position in the detector and the measured particle. As we will see in Section 4.6.2, combining all detector components a jet energy resolution of better than 3% can be achieved for high p_T jets.

3.6. Muon System

The muon system [97] is the outermost detector of CMS and is embedded in the iron return yoke of the magnet. The system uses different detectors in the barrel and endcap region. In the barrel region drift tubes (DT) are employed, whereas in the forward region cathode strip chambers (CSC) are used. Muons are not stopped in the detector[4]. Resistive plate chambers (RPC) are installed in both, the barrel and endcap region, which allow a fast tagging of the muon hits in addition to the DT and CSC time tags.

[4]For muons with p_T <5 GeV the strong magnetic field causes them to curl in the tracker, so that they do not reach the muon system.

Due to the barrel-endcap design of CMS the two muon systems overlap. While the DTs are covering $|\eta| < 1.2$, the CSCs cover the region $0.9 < |\eta| < 2.4$. The muon system is marked magenta in Figure 3.3.

The DT system consists of four stations, of which each have up to three so called super layers, which consists of four layers of drift tubes each. A drift tube has a single wire resolution of 250 µm, which provides for 8 hits a resolution of better than 100 µm. This resolution is comparable to the effect of the multiple scattering on the resolution for a $p_T = 200\,\text{GeV}$ muon.

The CSCs consist also of stations where the outermost station, which is furthest from the beampipe was installed between Run I and Run II. The CSC system consists of trapezoidal petals. A spatial resolution of $\approx 75\,\mu\text{m}$ is achieved for the innermost stations and $\approx 150\,\mu\text{m}$ for the others.

The muon resulting resolution can be seen in Figure 3.2 (right). This shows that the muon resolution is about 5% at for TeV-scale muons. The limiting factor for this resolution is the showering of the muons in the iron of the return yoke.

3.7. Trigger

The interaction rate at CMS can be up to 40 MHz. Not every event can be stored or even fully reconstructed. Therefore the rate is reduced by the CMS trigger system. It is comprised of two subsystems, the Level 1 (L1) trigger [98] and the High-Level Trigger (HLT)[99]. Where the L1 system is implemented in programmable hardware and can filter for basic detector responses. The HLT is a software based system, which provides many aspects of the full reconstruction. The rate is typically reduced by the L1 to $\approx 100\,\text{kHz}$ and by the HLT to $\approx 1\,\text{kHz}$. One of the benefits of this two level trigger system is the flexibility of the HLT. New triggers can be implemented and updated without change of hardware. The use of cross triggers[5] has a huge success in CMS. Due to the almost full reconstruction of the events in the HLT, complex objects such as jets and E_T^{miss} can be used in one trigger simultaneously. For the future the L1 trigger is updated to also combine multiple detector components.

More details on the CMS detector can be found in [88].

[5] Triggers that consider more than one object in the final state.

4. Tau Reconstruction

The tau reconstruction in CMS uses the decay modes of the taus hadronic tau decays (τ^{had}). These decay modes are substantially different from quark or gluon decays. The τ^{had} jets are narrower and consist of fewer constituents the quark and gluon jets will be called QCD jets in the following. In order to reconstruct the tau decay modes, information from all subcomponents of the detector have to be combined. To achieve this a particle flow technique was formed which is now one of the standard reconstruction algorithms in CMS [100, 101]. This differs significantly from previous tau reconstruction algorithms, which were mainly calorimeter driven[1]. A discrimination against light charged leptons is also possible. The particle specific detector responses can be used to build these discriminators.

This chapter will mainly describe the status of the reconstruction at the time of the 2012 data taking and my developments for a high p_T extension entering the final default tau reconstruction at the end of Run I. In Section 4.5 the changes for Run II are summarized. The performance of the high p_T tau reconstruction will be discussed in Section 4.6.

4.1. The Tau Lepton

The tau lepton is the charged lepton of the third fermion family and the heaviest lepton. With a mass of 1776 MeV it is heavier than most of the light quark mesons. This allows the hadronic decay of the tau lepton, which sets it apart from all other leptons.

Due to the lepton number conserving weak interactions, the tau final states are always accompanied by one neutrino in the hadronic final states ($\tau^- \to \nu_\tau + \tau^{had}$) and two neutrinos in the leptonic final states ($\tau^- \to \nu_\tau l^- \bar{\nu}_l$). The neutrinos are not detectable without a huge volume of instrumented material (e.g. IceCube or Kamiokande), therefore the full tau energy cannot be measured.

The leptonic decays of the tau can not be distinguished from prompt leptons in a $l + E_T^{miss}$ final state as was shown in [103]. Therefore here only the hadronically decaying taus will be discussed. The direction of the neutrino from the $\tau \to \tau^{had}\nu$ decay can not be reconstructed, due to the additional neutrino from the $W'/W \to \tau\nu$ decay.

The transverse momentum of the τ^{had} with respect to the undecayed tau p_T is shown in Figure 4.1. For each tau p_T bin the distributions are normalized to one. Since the region of kinematically allowed τ^{had} p_T values increases with tau p_T, the probability to find a specific τ^{had} p_T decreases. Nonetheless one can see that for a given τ p_T the τ^{had} p_T is not uniformly distributed, but a slight tendency towards higher values is observed, which is due to the boost of the τ^{had} in the τ direction.

[1] For a comparison in the jet reconstruction see [102].

The hadronic decays of the tau mainly occur via light meson resonances. A detailed list can be seen in Table 4.1.

final state	branching fraction in %	resonance	mass [MeV]
$e^-\bar{\nu}_e\nu_\tau$	17.83 ± 0.04		
$\mu^-\bar{\nu}_\mu\nu_\tau$	17.41 ± 0.04		
$\pi^-\nu_\tau$	10.83 ± 0.06		
$\pi^-\pi^0\nu_\tau$	25.52 ± 0.09	ρ	770
$\pi^-\ 2\pi^0\nu_\tau$ (ex. K^0)	9.30 ± 0.11	a_1	1200
$\pi^-\ 3\pi^0\nu_\tau$ (ex. K^0)	1.05 ± 0.07		
$\pi^-\pi^+\pi^-\nu_\tau$ (ex. K^0 , ω)	8.99 ± 0.06	a_1	1200
$\pi^-\pi^+\pi^-\pi^0\nu_\tau$ (ex. K^0 , ω)	2.70 ± 0.08	a_1	1200
$h^-\omega\nu_\tau$	2.00 ± 0.08		
$\sum$ other decays with $B < 1\%$	4.36 ± 0.110		

Table 4.1.: All tau decay modes with a branching fraction above 1% and the main associated resonance, if known [15].

This shows that 65% of the tau decays are hadronic. Of this hadonic tau decays about 80% have only one charged hadron in the decay. A reconstruction of these tau decays should therefore concentrate on the correct identification of these decay modes.

Another property that sets the tau apart is the lifetime. The tau has a mean decay length of $c\tau_{tau} = 87.03\,\mu m$, which means that the tau decays at some distance from its production point. This has consequences for the reconstruction in the detector. From the space contraction $l_{lab} = \frac{l_0}{\gamma}$ one can estimate that a tau with 900 GeV decays in the average at a distance of 4.4 cm which is the distance from the innermost pixel layer in CMS to the beampipe centre and therefore the simulation of the decays has to include this effect.

The effect of the boost on the opening angle is shown in Figure 4.1 (right). The opening angle is defined as the maximum ΔR between the centre of the hadronic τ^{had} and its final state particles. While for low p_T taus the ΔR is relatively broad, for taus with $p_T > 500\,GeV$ the majority of the hadronic opening angles are smaller than $\Delta R = 0.06$. This is much narrower than a quark or gluon jet.

For the tau reconstruction in CMS these three characteristics will be used: the decay resonances, the narrow jet and the lifetime.

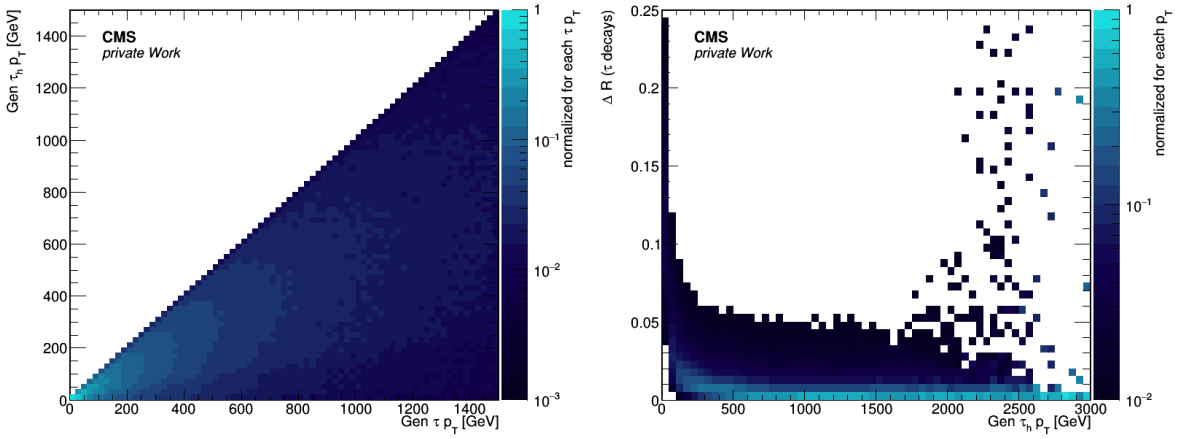


Figure 4.1.: The left plots shows the p_T of the τ^{had} with respect to the p_T of the undecayed tau. For each τp_T the evnets are nomalized to show the relative distribution. On the right, the opening angle of the τ^{had} $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta^2)}$ is shown with respect to the p_T of the tau evaluated with PYTHIA 8.

4.2. Particle Flow Algorithm

The particle flow (pf) algorithm is used to identify and sort all particles in the detector, creating a generator like object classification. The particle flow algorithm links measurements of all CMS-subdetectors and combines them according to their uncertainties. One design goal of the pfAlgorithm is to reconstruct all objects and allow the building of complex objects like jets. The so called pfCandidates are sorted into seven different types, ordered in the algorithm, they are:

- μ pfMuons are built from tracks in the tracker and/or the muon system with no ECAL cluster
- e pfElectrons are built from GSF-tracks[2] and tracks with compatible ECAL clusters
- $h^{\pm}$ pfChargedHadrons are built from tracks linked with calorimeter entries
- γ pfGammas are built from ECAL clusters that are not linked to tracks
- h^0 pfNeutralHadrons are built from clusters in the HCAL
- e^{HF}/γ^{HF} are electromagnetic particles in the hadron forward calorimeter
- h^{HF} are hadronic particles in the hadron forward calorimeter

The aim is not a perfect particle identification, but a categorisation. This means that in contrast to dedicated object identifications a high efficiency is more important than a correct particle identification, assigning all detector responses to one of the objects listed above. The creation of these candidates is done in three steps. First hits are combined in all detector components separately into pfElements, which means that tracks are fitted and calorimeter entries are clustered. Secondly these pfElements are linked between detector

[2]Gaussian Sum Filter, details see [104].

components to pfBlocks. The objects reconstructed with the standard muon, electron and gamma reconstruction, see Section 5, are removed from the pfBlocks collection and used directly as pfCandidates [105]. The remaining pfBlocks are linked as described in [101]. The linking is based mainly on geometrical aspects, but also the energy compatibility is checked. In the last step these linked objects are sorted iteratively into particle candidates where the ordering is determined by the list above. While some properties of the physical particles are used, e.g. the mass of the muon or the mass of the pion for the $h^{\pm}$, this sorting is not a complete particle identification. A pfGamma is an energy cluster in the ECAL that could not be linked to a muon, electron or charged hadron, it may still be a light meson component of a jet, or detector noise.

Using the particle flow reconstruction a better resolution can be achieved compared to other algorithms, due to measurement of the same particle in different components of the detector, and the categorization of all detector hits into one standardized format. This format allows for example the detection and removal of pileup in the event, just by removing all particles that have some distance in z from the primary vertex.

The effect of the particle flow reconstruction on the detector resolution can best be seen in the missing energy. In Figure 4.2 the E_T^{miss} resolution is shown for calorimeter based E_T^{miss}, calorimeter based and track corrected E_T^{miss} (tcE_T^{miss}) and pfE_T^{miss}. The comparison was done for the first data in 2011, but is still valid. The calorimeter only based E_T^{miss} has a resolution which is two times worse than pfE_T^{miss}. The track corrected calorimeter E_T^{miss} has a resolution in the same order of magnitude as the pfE_T^{miss} but it is worse for all measured points then the pfE_T^{miss}.

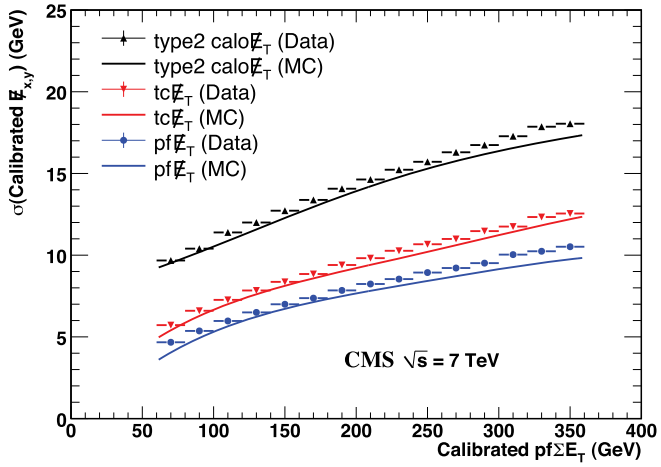


Figure 4.2.: Width of the missing transverse energy resolution as a function of the $\sum E_T$ of the particle flow reconstruction. Plot from [106]

4.3. The Tau Reconstruction as Hadrons Plus Strips

The Hadron Plus Strip (HPS) reconstruction builds on top of the pfCandidates. The initial aim of the HPS was a highly efficient tau identification at energies from 20 GeV to 100 GeV.

As can be seen in Table 4.1 the hadronic tau decays have final states of charged and neutral hadrons (mainly $\pi^{\pm}$ and π^0). While the $\pi^{\pm}$ are reconstructed as $h^{\pm}$ (called "Hadrons"), the π^0 form clusters of pfGammas (called "Strips").

The HPS algorithm starts with a pfJet. The components of each pfJet are separated into charged hadrons ($h^{\pm}$) and strips (π^0). The π^0s form strips in the ECAL since the π^0 can decay into two photons, which in turn can convert into two electrons. These electrons are effected by the magnetic field, which bends them in the ϕ direction. The resulting strips are therefore wider in ϕ as in η. In a $\eta \times \phi = 0.05 \times 0.2$ window all contributions form $\pi^0 \to \gamma\gamma$ decays are combined. The threshold for pfGammas is 0.5 GeV, and only strips with a total energy above 2.5 GeV are kept. The charged hadrons are built from $h^{\pm}$ with $p_T > 0.5$ GeV and the possible secondary vertex distance to the primary vertex has to be less than 0.4 cm in the z direction and less than 0.03 cm in the transverse plane[3]. One has to note here that h^0 pfCandidates are ignored by the reconstruction, because the neutral pions are expected to interact mainly in the ECAL. Leakage to the HCAL would be combined to a pfGamma candidate. Also the forward pfCandidates e^{HF}/γ^{HF} and h^{HF} are not used, because no track can be recorded for this detector region.

Since the tau decays mostly consist of charged and neutral pions, they will be used synonymous to charged hadrons and neutral hadrons in the following. In the jargon of the tau reconstruction the number of charged particles is often referred to as prongs. An one $h^{\pm}$ one π^0 decay is therefore called an one prong with one π^0.

From these candidates the τ^{had} constituents are built. These strips and charged hadrons are then combined into potential decay modes where for each decay mode a mass window corresponding to the decay resonance is allowed. The possible mass windows are listed in Table 4.2. A schematic representation of the algorithm can be see in Figure 4.3. Since not all decay modes can be associated to a resonance the mass windows are kept very loose and mainly constrain the mass to a value below the tau mass.

decay mode	mass window
$1h^{\pm}\ 0\pi^0$	$M_{\pi^{\pm}}$
$1h^{\pm}\ 1\pi^0$	0.3 MeV-1.3 MeV
$1h^{\pm}\ 2\pi^0$	0.4 MeV-1.2 MeV
$3h^{\pm}\ 0\pi^0$	0.8 MeV-1.5 MeV

Table 4.2.: All decay modes with their corresponding mass window in the reconstruction.

Via this algorithm all jets are sorted into decay modes, visualized in Figure 4.4, if more than one combination of the decay modes is possible the highest energetic decay mode combination is chosen as τ^{had}. Other hadrons are not counted as part of the τ and will enter the isolation cone, as we will see later. The decay mode finding therefore is highly efficient (99% for τ^{had} for taus from Z-boson events [107]). In addition to this decay mode the τ^{had} is required to have a ΔR cone size of less than $3\,\text{GeV}/p_T$. In order to accommodate the detector resolution this relative definition is truncated at $\Delta R = 0.05$ for high p_T and $\Delta R = 0.1$ for low p_T τ^{had}.

[3]This assumes that the primary vertex is linked to the tau vertex. For this analysis the requirement is not used, because the primary vertex can not alwasy be reconstructed correctly for a $W' \to \tau^{had}\nu\nu$ event.

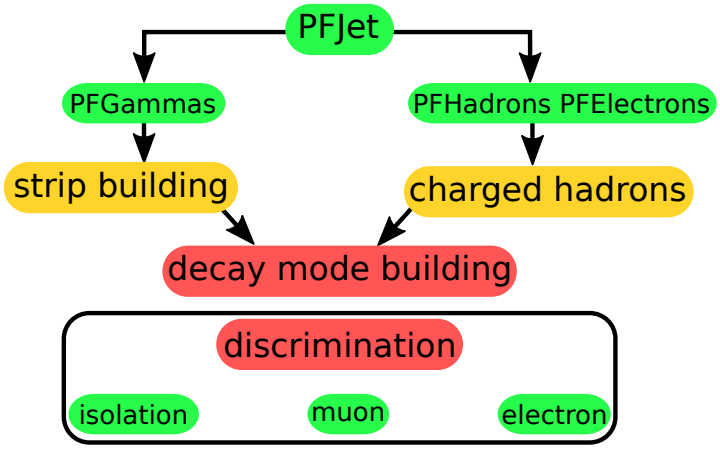


Figure 4.3.: A schematic view of the HPS reconstruction algorithm.

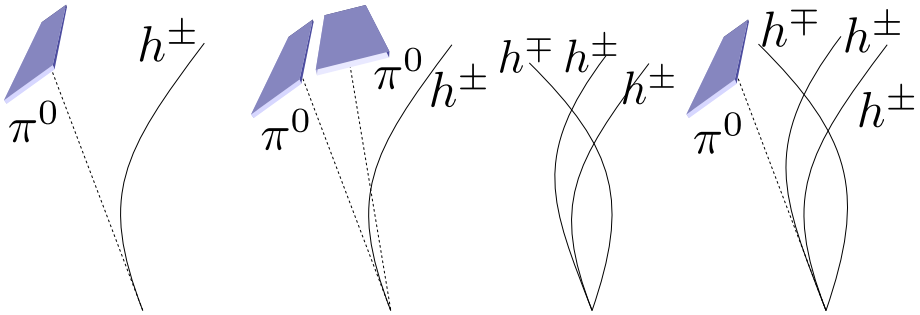


Figure 4.4.: The graphs illustrate the reconstruction by the HPS algorithm, showing the charged hadrons as lines and the strips from the neutral pions as blue boxes. Considered tau decays: the two left graphs illustrate the decays $\tau \to h^{\pm}\pi^0$(B ~ 26%) and $\tau \to h^{\pm}\pi^0\pi^0$(B ~ 9.5%) would both lead to a "one-prong" signature. The two graphs on the right lead to an observed "three-prong" signature from the decays $\tau \to h^{\pm}h^{\pm}h^{\mp}$(B ~ 9.8%4) and $\tau \to h^{\pm}h^{\pm}h^{\mp}\pi^0$(B ~ 4.8%). Not shown is the one prong decay without a π^0(B ~ 11.6%).

4.3.1. Tau Discriminators

The sorting of the decay mode into signal and isolation candidates has consequences on the isolation. If the correct decay mode is not found, the isolation may fail. For the isolation there are two approaches: a cut-based and a multivariate analysis (MVA)[4] [108] discrimination. The input variables are the same for both methods, the multivariate discriminator has slightly higher efficiencies and lower fake probabilities, but the decisions of the MVA are less comprehensive.

For this analysis both approaches will be used to provide results. For the Run I analysis the main result will be given with the cut-based discriminator and only a few key distributions will also be shown with the MVA isolation. This is done, because this analysis is the first analysis using τ^{had} candidates with transverse momenta this high and the cut-based can provide a more reproducible result, even if the MVA discriminator has a higher efficiency and a lower fake probability. For the Run II analysis, we will see that the fake probability of the cut-based

[4]Uses a boosted decision tree.

discriminator is too high to provide reasonable results, therefore the MVA discriminator will be used, which was then already validated with the Run I results of this analysis.

The input is given by all charged particles and photons with $p_T > 0.5\,\text{GeV}$ with a cone of $\Delta R = 0.5$ around the τ^{had}, excluding the tau signal candidates. The contribution from additional interactions in the same bunchcrossing (pileup) is taken into account, by requiring a distance $\Delta z < 0.2\,\text{cm}$ from the primary vertex. The isolation is then computed as:

$$I = \sum_{h^{\pm}+e+\mu} (p_T) + \max\left(0, \sum_{\gamma} p_T - \Delta\beta\right) \tag{4.1}$$

The $\Delta\beta$ correction should mimic the effect of pileup on the neutral pfGamma candidates, which cannot be identified by their vertex. The assumption is that the contribution from neutral and charged particles due to pileup is correlated. Therefore $\Delta\beta$ is the sum of all charged particles p_T value with $\Delta z > 0.2\,\text{cm}$ in a $\Delta R < 0.8$ region around the τ^{had} candidate times 0.46. The value is obtained from simulation.

For the cut-based isolation three working points are chosen to be $I < 2\,\text{GeV}$ (loose), $I < 1\,\text{GeV}$ (medium) and $I < 0.8\,\text{GeV}$ (tight).

The MVA-based isolation discriminator uses a variety of input parameters:

- p_T and η of the τ^{had}
- charged and neutral isolation sums and $\Delta\beta$ correction as for the cut-based discriminator, see Equation 4.1.
- reconstructed decay mode
- transverse impact parameter and its significance
- distance and significance of the distance between primary and secondary vertex, if the secondary vertex is identified.

The MVA discriminator is trained with simulated events, where the reconstructed tau candidates can be matched to the generator information. If a generator level tau is found, the tau is treated as a signal, if not it is treated as a background for the MVA. The idea of the MVA discriminator is that p_T and η dependent effects are taken into account and the additional information of the tau lifetime is used. In order to not effect the training of the MVA by different p_T and η distributions in background and signal, all events are weighted to provide the same distribution in the $p_T - \eta$ plane. As for the cut-based discriminator several working points (loose, medium, tight and very tight) have been chosen to reflect different combinations of efficiency (60% for loose to 48% for very tight on a $Z \to \tau\tau$ sample) and fake probability.

For the discrimination against muons also a cut-based and a MVA based discriminator was developed. Since for the identification of muons with $p_T > 7\,\text{GeV}$ the muon system can be used, the cut-based discrimination is very efficient and the MVA discriminator can not improve the performance further. Hence only the cut-based discriminator will be described here. In general a τ^{had} is tagged as a muon, if signals are found in the muon system close to the tau candidate. There are two working points (loose and tight).

For the loose working point of the muon discriminator the tau has to have no track segments that are reconstructed in two stations within a $\Delta R = 0.3$ cone around the τ^{had} candidate, and the sum of ECAL and HCAL energies should be smaller than $0.2 \cdot p_T$ of the leading track of the τ^{had} candidate.

For the tight working point the same requirements have to be fulfilled and in addition, no hits should be present in a $\Delta R = 0.3$ cone in the CSC, DT and RPC systems. The expected efficiency on a $Z/\gamma^* \to \tau\tau$ event is 99.3% and 99.1% for the loose and tight discriminator respectively.

For the electron discrimination no additional detector system can be used, therefore the discrimination is more complex than the muon discrimination and only a MVA based discrimination is used. The input variables are [4]:

- p_T and η of the τ^{had}
- p_T, η, σ_{p_T} of the Gaussian Sum Filter (GSF) track.
- $N_{hits}^{GSF} - N_{hits}^{KF} / N_{hits}^{GSF} + N_{hits}^{KF}$ where N_{hits}^{GSF} are the number of hits in the tracker reconstructed with the Gaussian Sum Filter algorithm and N_{hits}^{KF} the number of hits reconstructed with the Kalman-Filter algorithm [109, 110].
- χ^2/N_{dof} of the track fit
- $M_{\tau^{had}}$
- $E_{ECAL}/(E_{ECAL} + E_{HCAL})$ of all pfCandidates associated to the τ^{had} candidate.
- E_{ECAL}/p and E_{HCAL}/p where p is the momentum of the leading charged particle from the τ^{had} candidate
- $\sqrt{\sum (\Delta\eta)^2 p_T^\gamma}$ and $\sqrt{\sum (\Delta\phi)^2 p_T^\gamma}$ the respective p_T-weighted (in GeV) root-mean-square distances in η and ϕ between the photons in any strip and the leading charged particle.
- $\sum (E_\gamma)/E_\tau$
- $F_{brem} = (p_{in} - p_{out})/p_{in}$ where p_{in} is the curvature of the GSF track at the innermost position in the detector and p_{out} at the outermost position
- $(E_e + \sum E_\gamma)/p_{in}$ where E_e is the ECAL energy associated to the electron itself and E_γ is the energy associated to the bremsstrahlung photons
- $\sum E_\gamma/(p_{in} - p_{out})$

As for the isolation discriminator p_T and η are used to parametrise the dependence of the other input variables. One additional input information is the distance of the GSF track to the next ECAL module boundary, since electrons entering the ECAL close to the boundaries of a module can have a deformed electromagnetic shower or a lager signal in the HCAL. The MVA output is used to define different working points in terms of efficiency and fake probability (loose, medium and tight).

Further details can be found in [107, 4, 111, 112].

4.4. Changes for High p_T Taus

The tau reconstruction algorithm was originally designed for low energy taus. For highly boosted taus there are several effects that change the detector response and therefore the identification of taus. The consequences of a non optimal tau identification are shown in Figure 4.5 (left), which shows the energy response of taus as was the status at the beginning of this thesis. This variable is most sensitive to the reconstruction of the tau, because it shows that not all pfParticles from the τ^{had} are picked up by the tau reconstruction. In Figure 4.5 (right) the mean response is shown w.r.t. the generator level τ^{had} p_T. For high p_T taus only 40% of the tau p_T is reconstructed. One question is, if the full τ^{had} is even recorded by the detector. This can be answered by comparing the corresponding energy response of all pfCandidates in a $\Delta R < 0.12$ cone around a tau candidate to the generated taus, as done in Figure 4.5 (right). Since the energy response is flat with respect to the generated tau, it shows that the reconstruction of the decay mode and the τ^{had} energy has to be adjusted.

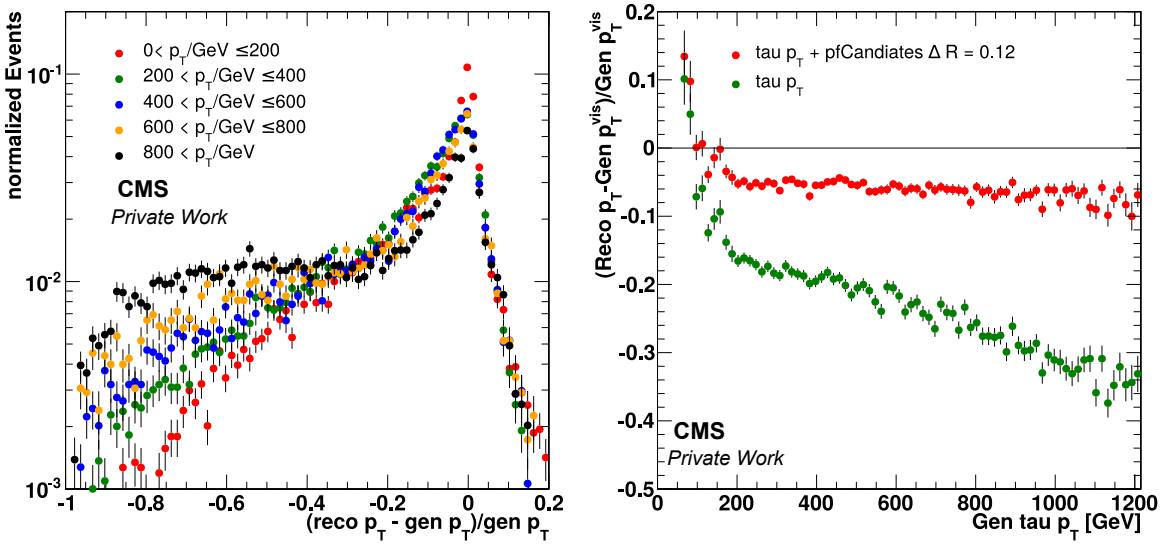


Figure 4.5.: Tau energy response evaluated with respect to generator level information for different p_T ranges (left) and as a mean (right).

The suspicion after this observation was that it could affect the decay modes in different ways. This was found to be true as is illustrated in Figure 4.6. Shown is the mean energy fraction of the pfCandidates with respect to the generated τ^{had}, that are missed by the reconstruction. For the one prong decays no significant amount of energy is lost by missed charged hadrons, as expected. The charged hadrons that are picked up by the cone are due to pileup interactions. For the one prong without π^0 the only significant unassociated energy is a result of unassociated h^0 that are created, but ignored by the algorithm. For the one prong with one or two π^0, the unassociated energy is also in form of γ candidates. The total mean fraction of lost energy for one prong taus is 20%. In the three prong decay modes the amount of energy from not picked up energy is larger. There is a substantial amount of energy lost due to charged hadrons, but also the amount of h^0 and some γ is increased.

The problem of the lost charged hadrons will be addressed in the next section. For the other

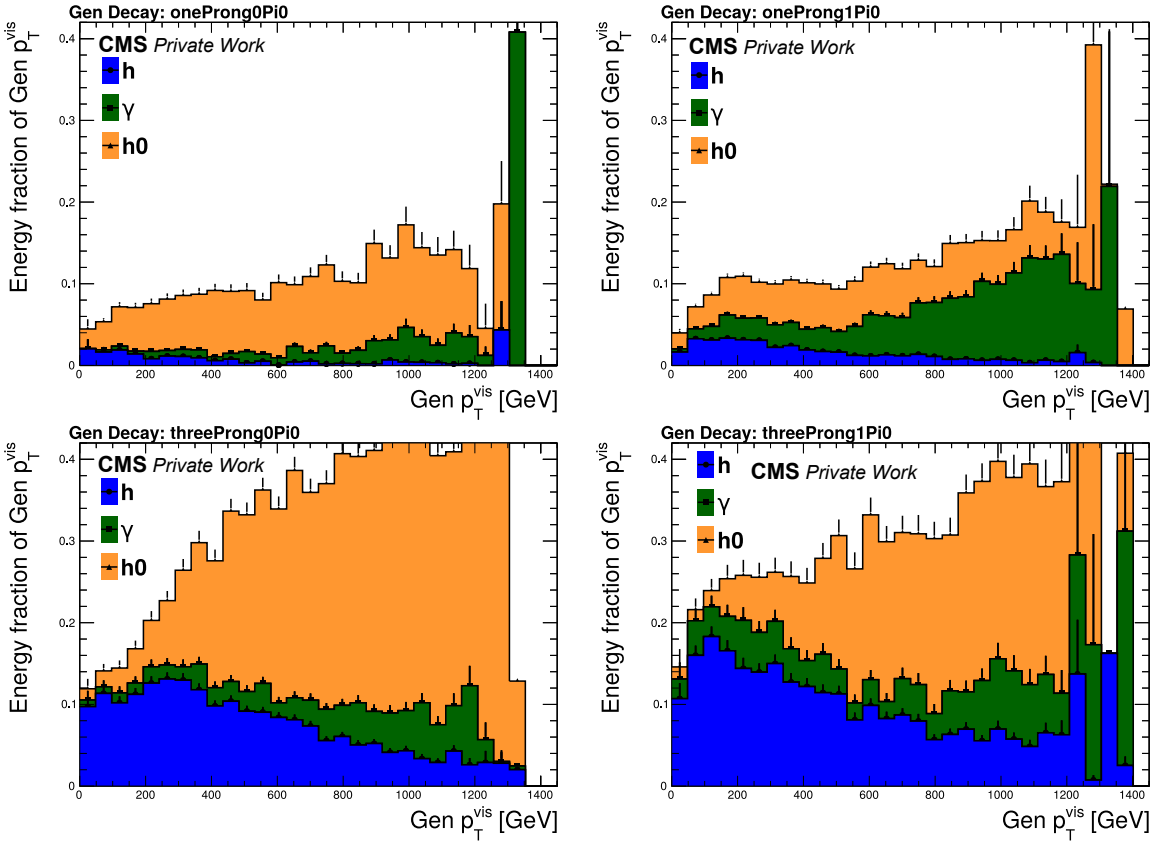


Figure 4.6.: The plots show the stacked fraction of energy w.r.t. the not picked up pfCandidates in a $\Delta R < 0.12$ cone around the tau. The plots are separated into generator level decay mode, which is denoted at the top of the plots.

pfCandidates we can investigate the geometrical distance a little further. This is summarized in Figure 4.7. The energy response is shown for different types of recombining the missed pfCandidates. While in red circles the standard tau reconstruction is shown, in blue crosses, the energy response for all pfCandidates in a $\Delta R < 0.12$ cone around the tau is shown. The violet squares show only h^0 and γ candidates added to the reconstructed taus. As one can see this curve coincides reasonably well with the cone. This tells us that no other pfCandidates are needed for the correct reconstruction. The green circles show the energy response of the pfCandidates that were picked with the algorithm and all h^0 and γ candidates that have a $\Delta R < 0.005$ to a $h^\pm$ candidate[5]. One can see that the mean energy response is decreasing with higher tau energies. If one adds also h^0 and γ candidates that have a $\Delta R < 0.005$ to the gamma candidates in the tau (yellow rectangles), the energy response is flat. From this several conclusions can be drawn. There are additional h^0 and γ pfCandidates that are in close proximity to the tau pfCandidates. But there are not only h^0 candidates that are close to $h^\pm$ from the tau cone, but also h^0 candidates that are close to the γ candidates from the

[5]As this is an obvious quantity to check it was previously missed in the tau reconstruction, because the ϕ and η variables for neutral pfCandidates are defined with respect to the detector centre, while the charged pfCandidates use their measured vertex as coordinate centre. The distance is only small if the η and ϕ variables of the charged hadrons at the entrance to the calorimeter are used with respect to the detector centre.

strip.

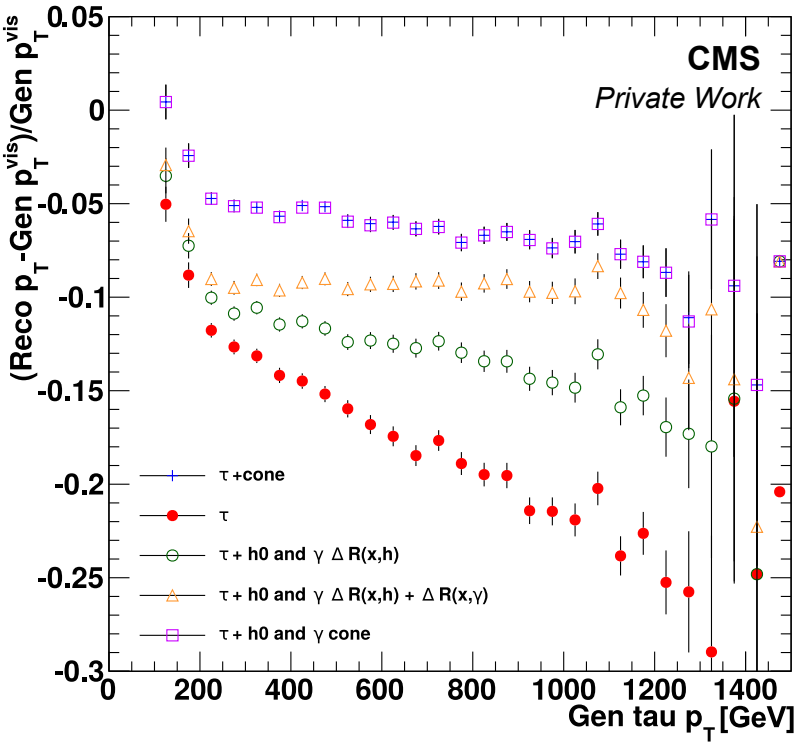


Figure 4.7.: Energy response of the reconstructed τ^{had} candidate. The various lines show the energy response for different additions to the tau reconstruction. As a cone around the tau, a $\Delta R < 0.12$ criteria was chosen. Different pfCandidates within this cone where added to the reconstructed tau.

These phenomena can be explained by three main effects in the reconstruction of high p_T taus:

- track merging
- track momentum underestimation
- mass window constrain

As will be shown in the following it is possible to recover these effects and still have a unified tau reconstruction for the CMS detector.

4.4.1. Track Merging

For high p_T taus the tracks in the detector can be close to each other. Due to a finite resolution they can merge in the detector. Two effects have an influence on the tracks of the highly boosted τ^{had}. The distance between two tracks in the pixel detector can be below the resolution of the detector and the calorimeter clusters cannot be split to have two charged hadrons. This

effects are only relevant for three prong tau decays, which account only for about 10% of the tau decays.

Though the tracking is done interactively as described in [113], the track seeding may not be able to pick up two individual tracks from the same pixel seeds. The distance between two tracks is dependent on the opening angle of the tau decay and the charge. In contrast to e.g. photon conversions two of the $h^{\pm}$ can have the same charge, therefore the distance will not increase due to the magnetic field. The effect of the boost can be seen in Figure 4.8 where the distance of the equally charged hadrons to each other is plotted with respect to the p_T of the τ^{had}. The distance is gaussian distributed at low p_T. For p_T larger than 300 GeV the gaussian distribution of the ΔR distribution is truncated at the lower side. The left plot is evaluated on generator level, while the right plot is evaluated with the reconstructed three prong decays. For the right plot all changes which will be discussed in this chapter are applied, but still one can see that above $p_T = 2\,\text{TeV}$ not many three prong decays are reconstructed with three prongs.

The second effect is due to the interplay between the calorimeter clusters and the track measurement. Tracks can be seeded by calorimeter clusters and clusters are only linked to tracks if the clusters can be split. The calorimeter cluster of charged pions have a varying size in the ECAL and HCAL. The shower of charged pions can be described by the hadronic shower model [114]. This can be summarized for high p_T pions in the CMS detector as follows: pions radiate energy due to ionisation as they transverse the detector and at some point start a shower. The starting point of the shower in the detector depends on the pion energy and the material, but is different for each event. The transverse shower size in general depends on the energy and the material. However, in the CMS detector the size of the shower is mainly given by the detector granularity. Figure 4.9 (left) shows the distance of the equally charged pions from a τ^{had} with 1 TeV at the ECAL entrance. The distance depends on the energy fraction the charged pions carry with respect to the τ^{had}. As black line the distance of three calorimeter cells is stated, which is the expected size of a high p_T pion cluster. As the energy of the τ^{had} has to be shared between at least three decay products, the most probable value for the carried energy is 1/3. Which would not allow to separate the calorimeter clusters of the two equally charged pions in the calorimeter.

To summarize, tracks are lost due to finite detector resolution. In order to recover the efficiency of the tau reconstruction two things were changed. As the identification of charged hadrons is not as effective as the track reconstruction, because some quality criteria are applied in the pfAlgorithm, the prong counting is done on a combination of tracks (with very little quality selections) and charged hadrons. The potential double-counting of tracks is corrected for. The second change is, that the unphysical two prong decay mode is allowed. This means, if from a three prong decay one prong is lost due to the detector resolution, the decay mode can still be built. As a consequence the charge information of these events is lost for most events[6]. The failure of linking between track and calorimeter cluster is also taken into account, by adding all h^0 clusters at the end of the found tracks.

[6]The charge could still be used if two equally charged pions are associated to one τ^{had}, but in the majority of the τ^{had} candidates one of the equally charged pions is lost, because these tracks are not separated as much by the magnetic field as opposite charged pions are.

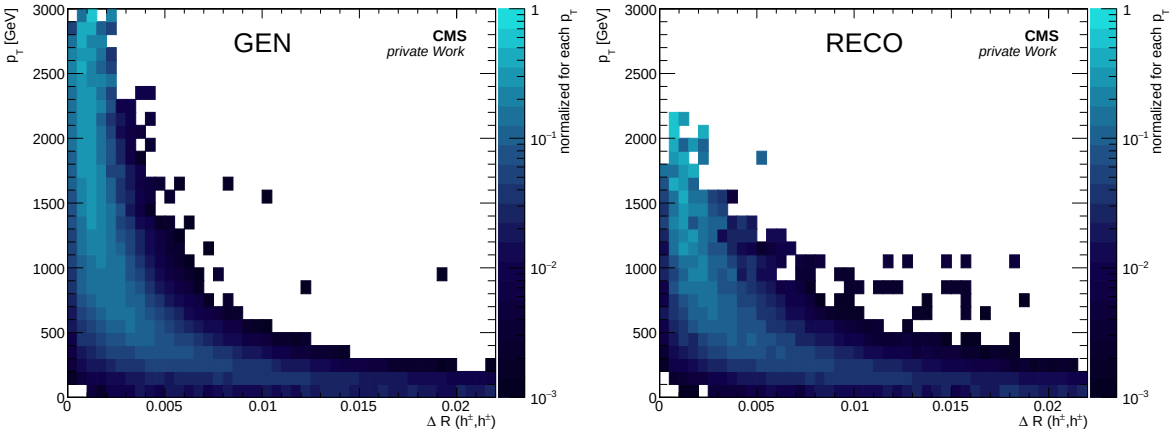


Figure 4.8.: The figure shows the distance ΔR of the two equally charged $h^{\pm}$ from the τ^{had} with respect to the p_T of the tau where for each p_T the distributions are normalized to one. For this figure all W' 13 TeV signal samples were combined. The plotted ΔR is evaluated on generator level (left) and after reconstruction (right) for all three prong decays. For the reconstruction all changes to the tau reconstruction, which are discussed in this chapter, are applied.

4.4.2. Track Momentum Underestimation

One of the premises of the pfAlgorithm is "tracker first", which is built on the assumption that the best energy resolution for charged hadrons is achieved by the track. As a consequence the linking of the calorimeter clusters with the tracks "trusts" the track measurement more. If the track and the calorimeter measurement are not compatible, they are not combined. A large cluster is assumed to originate from a neutral γ or h^0 and a $h^{\pm}$, and is therefore split into these two components. This means the calorimeter measurement of the $h^{\pm}$ is not used, but the size of the cluster is constrained to match exactly the pion hypothesis of the track. This approach was highly successful for low p_T taus and jets at all energies in the past [102, 107]. Whereas all other high level physics objects use either a cluster of pfCandidates (e.g. jets, E_T^{miss}) or have a dedicated high p_T identification, for high energy taus it poses a problem, since the HPS algorithm relies on the particle flow categorisation . The type of the pfCandidates is not used outside the tau reconstruction. The HPS algorithm assigns the artificially created γ candidates to the strip part of the tau. The h^0 candidate is ignored completely. The problem of these artificial strips is that the decay mode is changed, and that the reconstructed resonance mass of the $h^{\pm}$ and the strip does not have to be in the corresponding mass windows.

This behaviour is expected if one compares the resolution of the different detector components as shown in equation 3.1 and 3.2-3.4. The energy dependence of the tracker and calorimeter resolution is $\sigma_{Track}/p_T = A \cdot p_T \oplus C$ and $\sigma_{Calo}/E = S/\sqrt{E} \oplus N/E \oplus C$. It is clear that the tracker resolution at some point gets worse than the energy resolution and therefore the initial particle flow paradigm no longer holds. For the CMS detector the mean resolution is equal for tracker and calorimeter at roughly 500 GeV, as can be seen in Figure 4.9 (right). The exact point varies however on the actual track quality and the calorimeter measurement. For pions from the hard interaction the calorimeter measurements can be much more precise than the stated resolution and the track quality can be worse due to additional pions from the three prong decays or pileup tracks. The shown ECAL resolution is only valid for electromagnetic

showers fully contained in the ECAL (only e or γ). The values are shown for $\eta = 0$ were $E_T = E$.

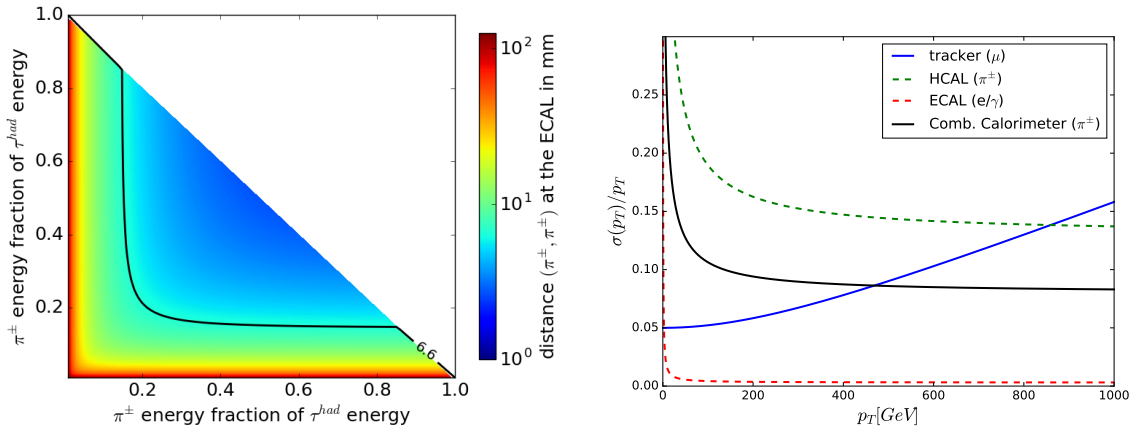


Figure 4.9.: For a 1 TeV τ^{had}, the distance of two pions with the same charge at the ECAL entrance is plotted w.r.t. the fraction of the energy they carry, is plotted on the left. In the right plot the relative resolution for tracker, ECAL and HCAL and the combined calorimeter is plotted, as shown in equation 3.1 and 3.2-3.4. Source [90, 94, 96]

In summary the "tracker first" premises does not work well for all high p_T $h^{\pm}$ pfCandidates and the calorimeter information should not be neglected. As a consequence for the tau reconstruction, the h^0 pfCandidates that are around the tau candidate are added to the tau. It was shown that the artificially created neutral pfCandidates have a distance of $\Delta R = 0$ with respect to existing pfCandidates in the tau, see Figure 4.7. To save computing time and because the result is the same, the energy is calculated from the existing τ^{had} candidate plus all neutral pfCandidates in a $\Delta R < 0.12$ cone around the tau. In order not to pick up noise the h^0 are required to have $p_T > 50\,GeV$ and this "post decay mode energy estimation" is only done for τ^{had} candidates with $p_T > 200\,GeV$. The artificially created γ candidates are not added by this procedure, because they are integral to the strip reconstruction of the decay modes and enter the isolation discrimination. A majority of the created γ candidates is picked up by the changed mass windows.

4.4.3. Mass Window Constraint

The mass window for the construction of the decay modes are set to reflect the decay modes as shown in Table 4.2. For the highly boosted taus the mass resolution degrades, therefore the mass windows are too tight for high p_T taus. The mass windows were modified to the values shown in Table 4.3.

The values were obtained evaluating multiple variations of the mass windows on a $Z' \to \tau\tau$ sample. The result can be seen in Figure 4.10. For this mass plot all pfCandidates in a $\Delta R = 0.3$ cone around a generated τ^{had} were combined to a tau candidate[7]. The mass of this tau candidate is plotted w.r.t. the generated p_T. For each p_T bin the mass distribution is

[7] In order to be as close as possible to the tau reconstruction h^0 and h^{HF} candidates were ignored.

decay mode	mass window
1h $0\pi^0$	set to $\pi^\pm$ mass
1h $1\pi^0$	$0.3 - \max(1.3, \min(1.3 \cdot \sqrt{\frac{p_T}{100\,\text{GeV}}}), 4.2))\,\text{MeV}$
1h $2\pi^0$	$0.4 - \max(1.2, \min(1.2 \cdot \sqrt{\frac{p_T}{100\,\text{GeV}}}, 4.0))\,\text{MeV}$
2h $0\pi^0$	$0 - 1.2\,\text{MeV}$
2h $1\pi^0$	$0 - \max(1.2, \min(1.2 \cdot \sqrt{\frac{p_T}{100\,\text{GeV}}}, 4.0))\,\text{MeV}$
3h $0\pi^0$	$0.8 - 1.5\,\text{MeV}$

Table 4.3.: All decay modes with their corresponding mass window in the reconstruction with the modifications for the high p_T taus.

normalized to one. The plots are divided into the decay modes as reconstructed with a modified reconstruction software, the two prong decay mode is allowed. The blue lines indicate the allowed mass for the reconstruction before the mass windows were adjusted. The red line shows the adjusted used maximum mass and the green line shows the allowed minimum mass.

One can see that for the one prong no π^0 decay mode the majority of the reconstructed tau mass has only one charged hadron, which mass is set to $M_{\pi^\pm} = 139.6\,\text{MeV}$. The events with diverging reconstructed mass have additional pfCandidates in the cone that are either from one prong $1\pi^0$ decays or detector noise, such as pileup or readout noise. The amount of events, which will probably not pass the isolation criteria is at the order of 3%. For the one prong $1\pi^0$ reconstructed decay mode the improvement of the algorithm is visible. The distribution begin to spread for higher tau p_T towards higher tau masses. And the decay mode can now associate these higher mass events with the charged hadrons. This effect is not present for the charged hadron only decay modes "two prong no π^0" and "three prong no π^0", due to the high spacial resolution of the tracker as shown in the appendix, Figure A.1. For the "two prong one π^0" decay mode most of the events have a mass within the mass window, but due to the loss of one charged hadron the distribution is much broader in general. For the three prong one π^0 decay mode one can see that the mass reconstruction is mainly not in the mass window. This is because the three prong one π^0 has the highest fake probability from jets and therefore has harsher requirements. Due to the low expected event rate and the high fake probability, this decay mode is reconstructed by the HPS algorithm, but not used.

4.4.4. Summary of the Changes for the High p_T Tau Reconstruction

For high p_T taus several changes for the HPS algorithm were made. While at the start of this investigation it was unclear if the HPS algorithm could be used for high p_T taus at all, it was shown that all effects could be understood and explained. The main issue of the high p_T objects was the strict reliance on the pfAlgorithm, when it comes to object identification. Since the aim of the pfAlgorithm is more a categorisation of physics objects than an identification, high p_T objects tend not to be reconstructed as a single pfCandidate e.g. $h^\pm$, but are split into multiple objects.

Several changes have been made that take the pfCandidate type splitting into account. For the counting of prongs not only $h^\pm$ and e candidates are used, but also tracks. Artificially

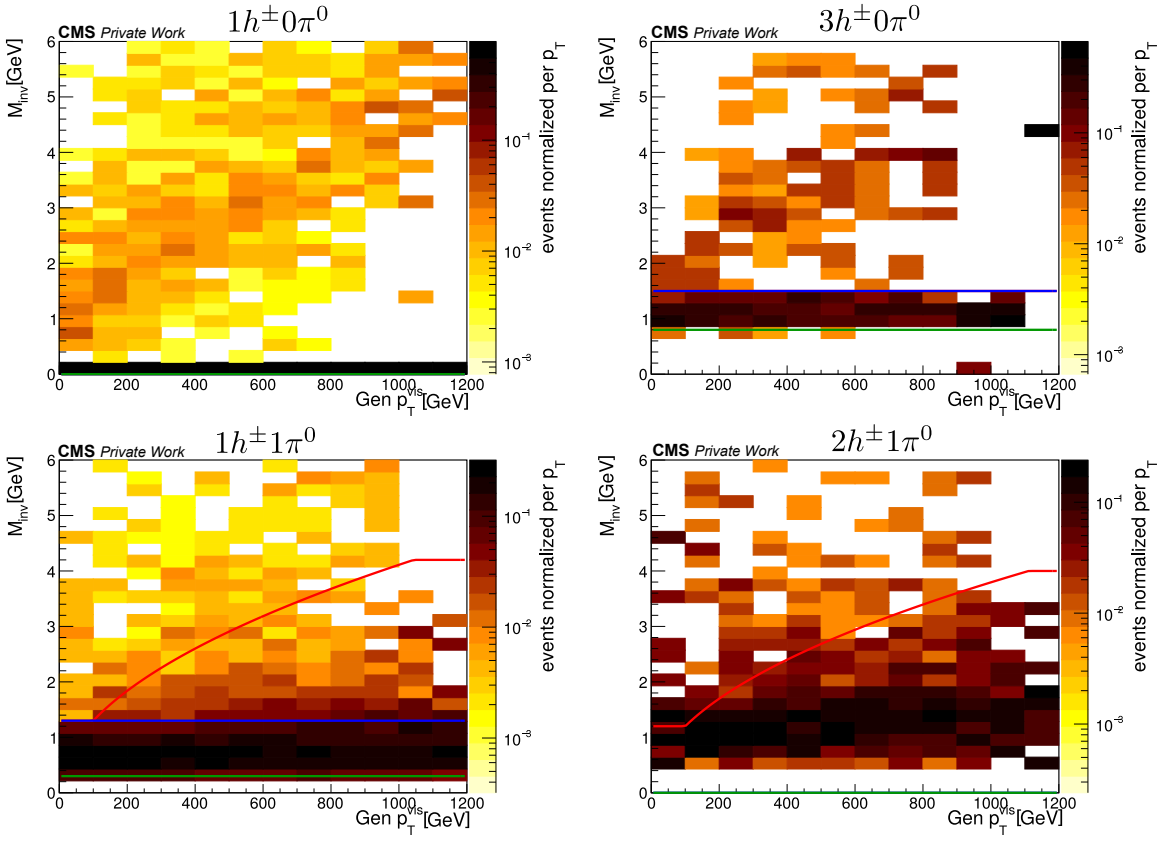


Figure 4.10.: The invariant mass of all pfCandidates in a $\Delta R < 0.3$ cone w.r.t. the visible tau p_T at generator level. The distributions are normalized to one for each p_T bin. The lines indicate the mass window constraints where green indicates the minimum mass, blue the maximum mass before the high p_T changes, and red the maximum mass after the high p_T changes.

created h^0 candidates are added to the tau after the decay mode finding and reconstruction. Also the two prong decay mode is allowed for high p_T taus. Finally the mass windows are less stringent for high p_T taus. A sketch of the updated HPS algorithm can be seen in Figure 4.11.

As can be seen in Figure 4.12, the decay mode reconstruction still has a strong dependence on the tau p_T. For high p_T taus most of the taus are reconstructed as one prong plus $1\pi^0$.

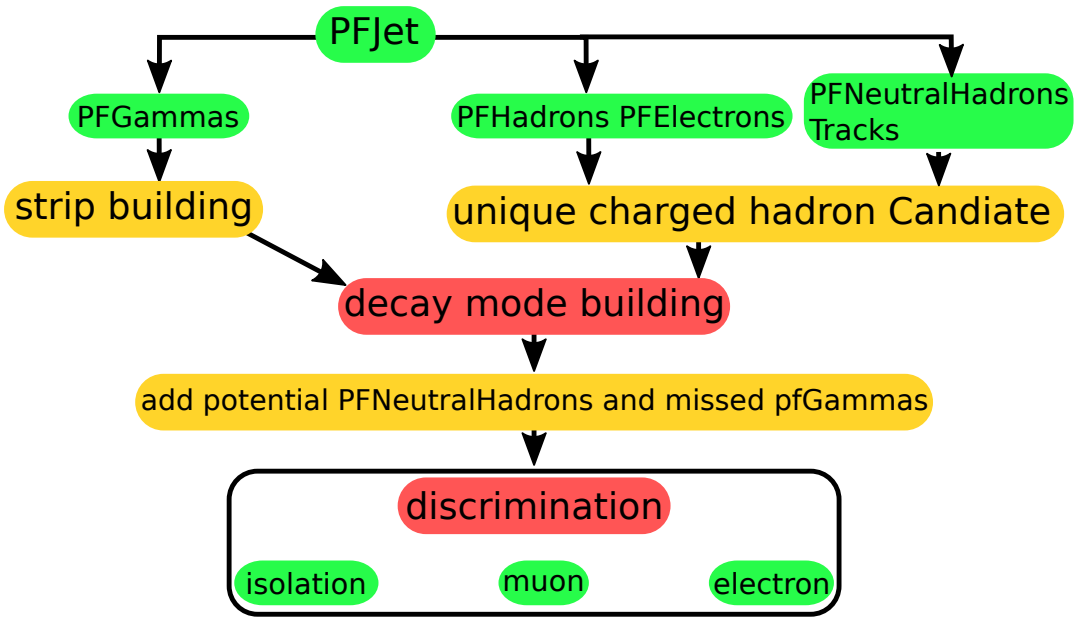


Figure 4.11.: A schematic view of the modified HPS reconstruction algorithm.

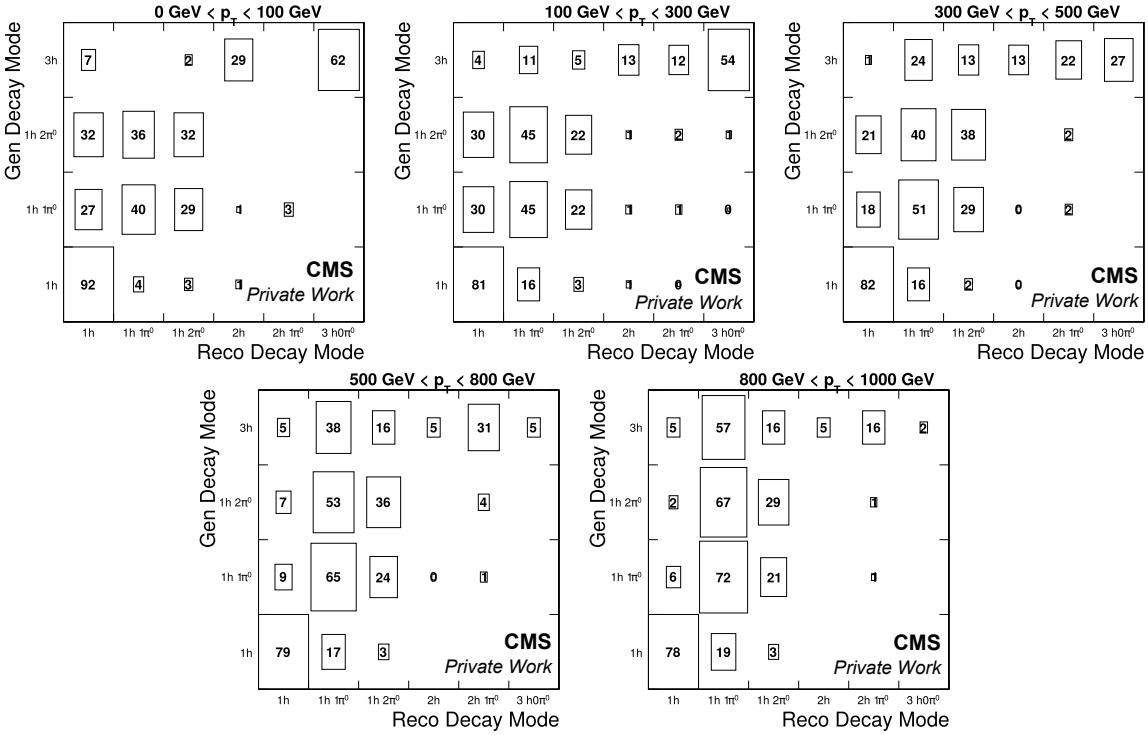


Figure 4.12.: The decay mode reconstruction for different energies. The shown numbers give the percentage of the events that for a given generated decay mode, are reconstructed with the shown reconstructed decay mode.

4.5. Tau Reconstruction Changes in Run II

For the Run II tau reconstruction some modifications of the input parameters were made. The tau reconstruction as a high level object is effected by these changes. A full description of the tau reconstruction in Run II can be found in [111]. For Run II the anti-k_T pfJet reconstruction distance parameter R changed from $R = 0.5$ (Run I) to $R = 0.4$ (Run II). Since the pfJets are only used to seed taus, which are expected to be well within $R = 0.3$, this change had no consequences for the tau reconstruction.

For Run II other changes are at the level of the pfCandidate reconstruction. The particle flow algorithm added the linking of tracks to converted photons. This helps in identifying the strips, but in a boosted object such as the tau, clusters are easily linked to the wrong cluster or the track quality was not good enough to build a $h^{\pm}$ candidate. As a consequence the tau decay mode building treats all photons with tracks as potential $h^{\pm}$ candidates, and an iterative approach chooses the tau decay mode with the highest p_T, as before.

Also concerning the photon conversion the fixed strip size of 0.05×0.2 in $\eta - \phi$, is relaxed to allow, if one conversion electron enters the strip, an additional electron outside the strip is allowed to be combined into a γ. This recombination allows to modify the strip window as a function of $p_T(e/\gamma)$:

$$\Delta\eta(p_T) = 0.20 \cdot \left(\frac{p_T}{\text{GeV}}\right)^{-0.66} \tag{4.2}$$

$$\Delta\phi(p_T) = 0.35 \cdot \left(\frac{p_T}{\text{GeV}}\right)^{-0.71} \tag{4.3}$$

With these changes the MVA discriminators were retrained.

4.6. Performance of the Tau Reconstruction Algorithm at High p_T

The performance of the tau algorithm can be viewed from several sides. On the one hand the efficiency and fake probability determine signal and non signal events that can be observed. On the other hand the efficiency and the fake probability are indicators for some misreconstruction effects. Another important quantity is the energy scale of the tau reconstruction. As can be seen in the previous Section 4.4 the reconstructed energy can be affected by the choice of pfCandidates for the τ^{had} candidate.

4.6.1. Tau Reconstruction Efficiency

The performance of the high p_T taus can be evaluated with the simulation up to very high tau p_T. For a comparison with data two final states are available: $Z \to \tau\tau$ which allow the probing of taus up to 100 GeV, and $t\bar{t}$ events which allow the probing of the tau efficiency up to 200 GeV [111, 4]. These studies show that the tau identification is working well and all the effects are modelled in the detector, as far as they are known.

As an example for the comparison between data and simulation the efficiency results from the tag and probe measurement of $Z \to \tau\tau$ events are given in Figure 4.13. As one can see, the

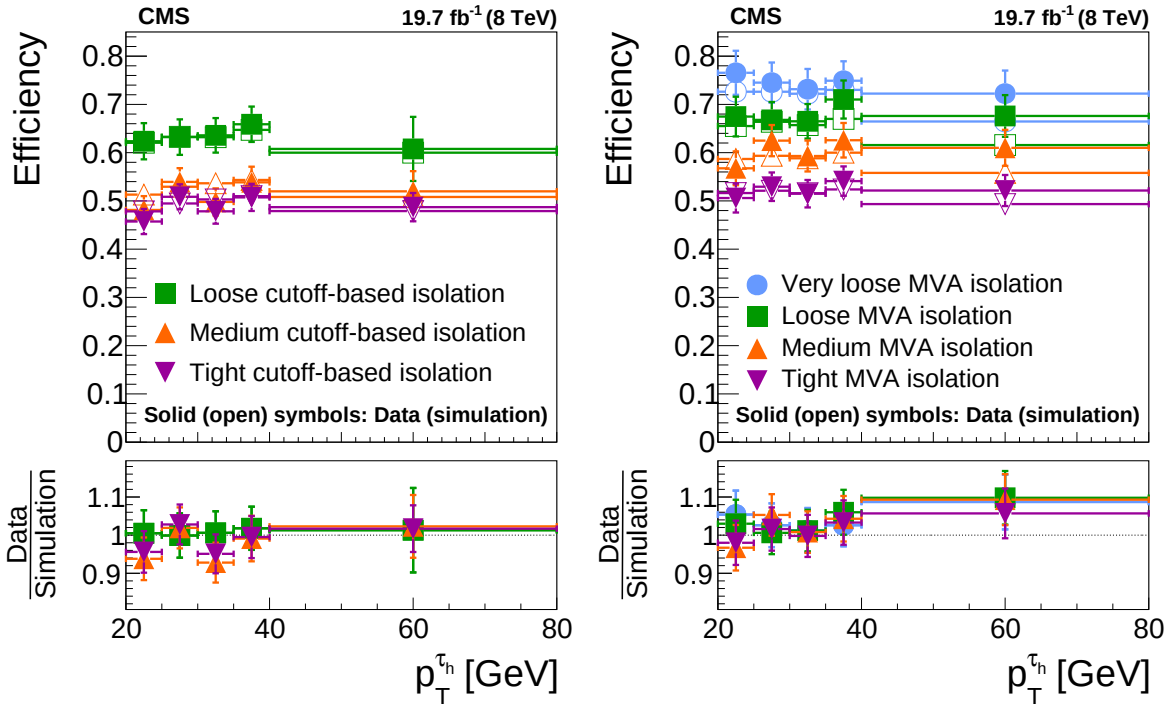


Figure 4.13.: The efficiency of the tau lepton using $Z \to \tau\tau$ events evaluated on data and simulation [4]. The left plot shows the cut-based isolation efficiency, while the right plot shows the MVA-based isolation efficiency.

reach of the tag and probe method is limited. But no significant deviation between data and simulation can be seen.

The events with the highest tau p_T at CMS are recorded in the tau+E_T^{miss} channel, which is the final state of this analysis. In order to understand the behaviour of the tau identification efficiency and fake probability simulation studies were made. As a reference sample here a Z' sample with $M_{Z'} = 2.5\,\text{TeV}$ is used. The efficiency includes the decay mode reconstruction and the isolation.

The result can be seen in the left plots of Figures 4.14-4.16 where the efficiency is shown as a function of various quantities. One can see that the efficiency w.r.t. the tau p_T (Figure 4.14) is constant. This contains a lot of information. If the decay mode reconstruction did not pick up all pfCandidates created in the detector, one would see a drop of efficiency for higher p_T, because the left over pfCandidates would enter the isolation cone and therefore the isolation would fail. If the decay mode picks up too many pfCandidates, the fake probability, would rise for larger p_T. For the evaluation of the fake probability a QCD sample is used. The results can be seen in the right plots of Figures 4.14-4.16. The efficiency and fake probability for the MVA discriminator is shown in Figure A.2-A.4 in the appendix.

One can also see in Figure 4.14 (right) that the fake probability is steeply falling for $p_T < 200\,\text{GeV}$ and then reaches a plateau. This is the expected behaviour since for high p_T jets the number of particles and therefore the number of charged hadrons rises with the QCD jet p_T. But above roughly 200 GeV the probability to get a jet with a low number of charged hadrons

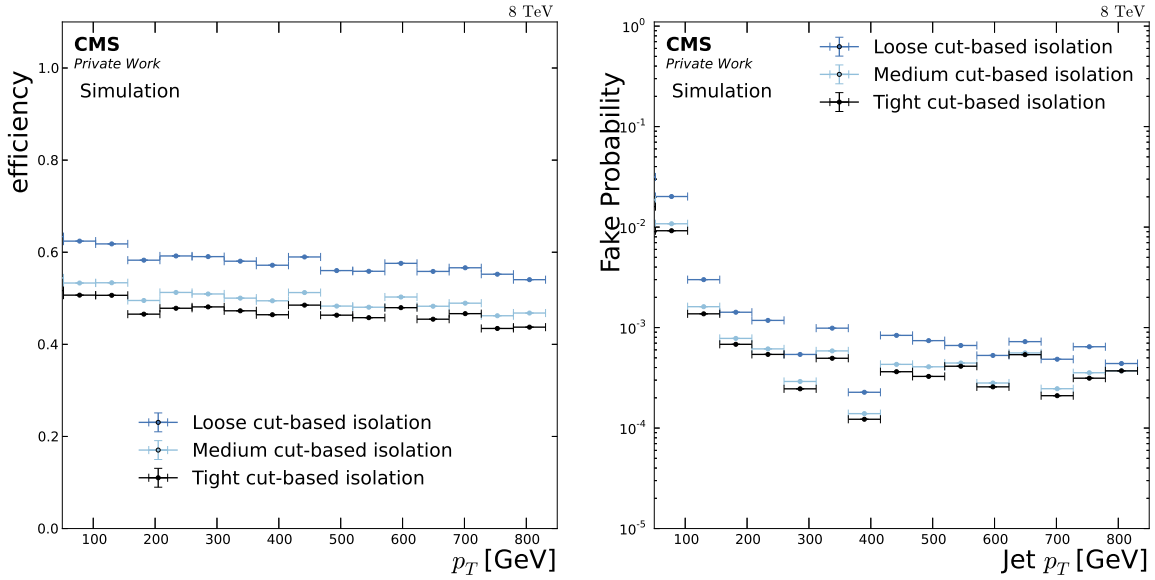


Figure 4.14.: The tau identification efficiencies (left) and fake probability (right) for high p_T taus as a function of p_T.

stays the same, thus the fake probability does not change by orders of magnitude anymore.

In Figure 4.15 one can see the dependence of the efficiency on the number of vertices, which is a measure of the pileup in the detector. A detailed description of pileup simulation will be given in Section 7.2.1. For the cut-based isolation one can see a slight increase of the efficiency towards a higher number of vertices. This can be explained, if the pileup adds energy to the strips of the tau decay the strips can pass the decay mode whereas they could not pass the decay mode without the pileup. The effect is small and the fake probability is relatively flat w.r.t. the number of vertices, compared to the p_T dependence. Therefore no additional changes are made.

In Figure 4.16 the η dependence is shown. For the cut-based isolation one can clearly see the structure of the ECAL in the efficiency plot. The fake probability is reasonable flat.

For 13 TeV the MVA tau identification performance is similar to the 8 TeV performance. The performance of the isolation is shown in Figure 4.17 for the p_T dependence. One can see that the identification efficiency for the cut-based isolation has a p_T dependence and the MVA discriminators show only a slight dependence. The increase of the efficiency for the cut-based discriminator is not visible if the two prong decay mode is not allowed in the isolated tau. The increase is therefore expected to be the effect of the more efficient decay mode finding. For a detailed discussion of this see [115]. The fake rate for the MVA based isolation discriminator is about one order of magnitude lower as for the cut-based discriminator at high p_T.

In summary the tau identification could be extended to high p_T taus. The efficiency is flat with respect to the p_T of the τ^{had}. The dependence on other key variables like η or the number or the number of vertices is understood. For the 8 TeV analysis the choice between the cut- and MVA-based isolation discriminator can not be based on the performance comparison of simulation with data. Therefore cut-based isolation will be used in the following for all results.

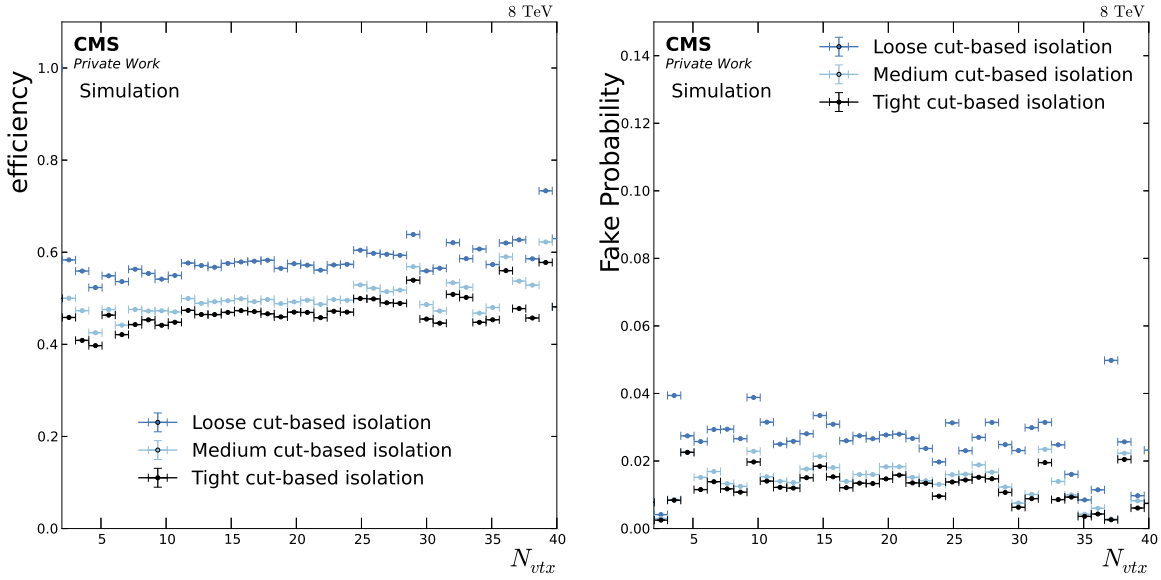


Figure 4.15.: The tau identification efficiencies (left) and fake probability (right) for high p_T taus as a function of the number of vertices.

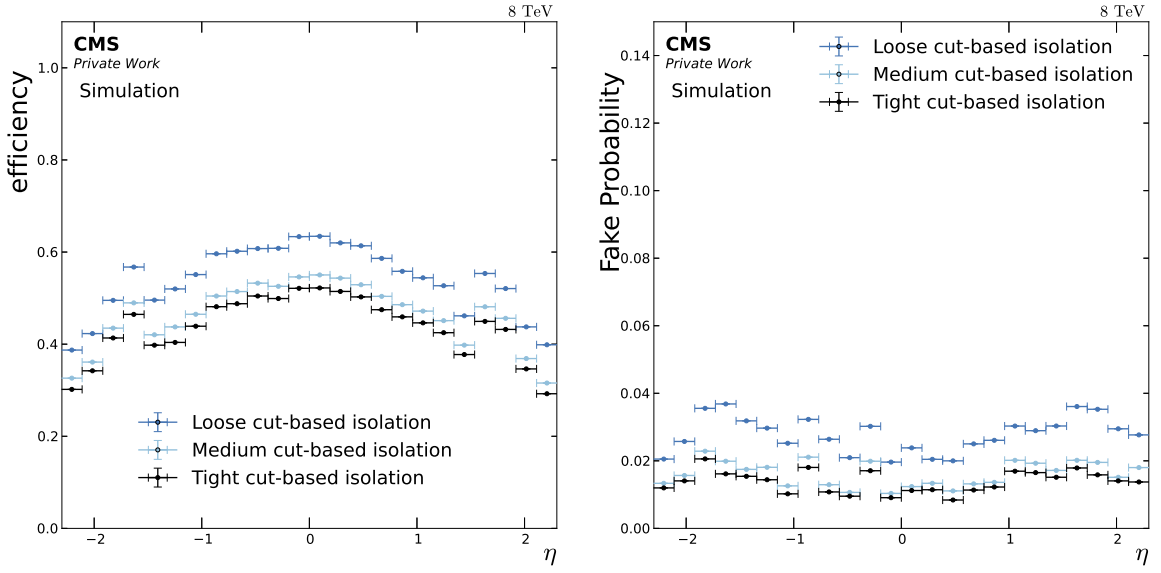


Figure 4.16.: The tau identification efficiencies (left) and fake probability (right) for high p_T taus as a function of η.

In order to check the behaviour of the MVA isolation discriminator key results will also be provided with this discriminator.

For the 13 TeV analysis the contribution from fakes in the signal region is much higher. Therefore the medium MVA isolation discriminator is used, which checked with the 8 TeV analysis to provide similar results.

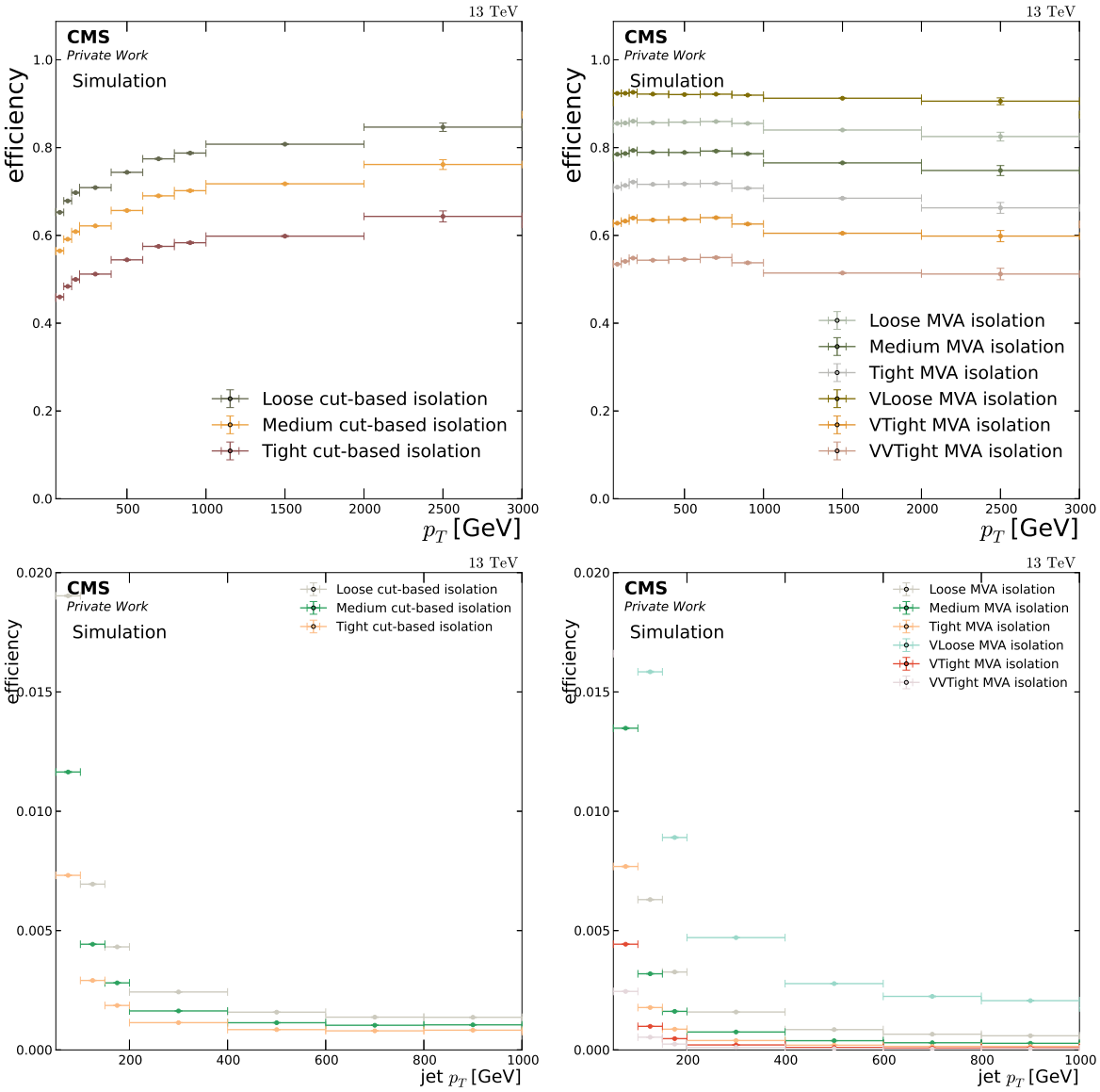


Figure 4.17.: Efficiency for the used isolations and the fake probability. The efficiency and fake rate is estimated with W events. The left side shows the cut-based isolation discriminators and the right side shows the MVA-based isolation. The efficiency of the cut-based isolation shows an increase w.r.t. the τ^{had} p_T. This is due to the increase of the decay mode efficiency.

4.6.2. Tau Energy Scale

The tau energy scale, meaning the detector response to the τ^{had}, is an important quantity. The intrinsic nature of the decaying tau reduces the information one can gain about the undecayed tau, especially the energy of the tau can not be reconstructed without further information from the underlying physics process. A precise energy measurement of the hadronic decay products is important in order to have a good description of the τ^{had} from simulation and therefore a good agreement between data and simulation is crucial. Two sources should be considered to determine the energy scale uncertainty.

First the uncertainty of the τ^{had} components, which are measured with different parts of the detector and thus have a different scale uncertainty. By components, here the detector measurements of tracks and the calorimeter clusters are meant. Since the tau decays into multiple particles with lower energy, e.g. one prong with two π^0, the scale of the subcomponents is understood better than that for a single particle with the same energy e.g. an e or μ, which is understood well. A τ^{had} with $p_T = 1\,\text{TeV}$ for example can have components with not more than 500 GeV energy. Additionally, the charged pions can be measured in several detector parts. If one assumes roughly the known uncertainties for the detector components one can combine these uncertainties to an uncertainty on the τ^{had} p_T. The result is that the sale uncertainty from the measurements of the subcomponents is in the order of 0.1% of the τ^{had} p_T.

This leaves the second uncertainty for the τ^{had} scale. If not all the decay products of the τ^{had} are included in the τ^{had} candidate, a systematic offset for the τ^{had} reconstruction would be observed. The nature of such an offset can have different effects. If the energy response has an offset w.r.t. the generator, this would not be problematic, since the offset can be corrected for. In light of the tau neutrino in the decay this correction does not even have to occur, since the full decay energy can never be reconstructed. Only the correct description of the tau response is crucial. An effect that cannot be corrected for is for example a non linear energy response, as seen in Figure 4.5 (right) for the old τ^{had} reconstruction as described in Section 4.3. For an energy response that is 15% off at 200 GeV and 20% off at 400 GeV, the same reconstructed p_T can be associated with different generated p_T. Therefore, a flat energy response is important. Studies so far showed, that the energy response shows no dependence on p_T at high p_T. This can be checked with data up to a $p_T \approx 200\,\text{GeV}$. The energy response was checked with $Z \to \tau\tau$ and $t\bar{t}$ events, as detailed in [4, 111]. The scale in the $Z \to \tau\tau$ events was measured comparing various templates of the reconstructed invariant mass of all visible $Z \to \tau^{had}\mu + E_T^{miss}$ decays and using the visible mass of the hadronic tau decays $M_{\tau^{had}}$. The resulting corrections showed that the τ^{had} energy response is well simulated. For a W' like signal the energy response of taus with high p_T is important. In order to test the energy response in data the jet reconstruction can be used, from which we know the energy scale at a 1-3% level to the highest energies. The ratio of the τ^{had} energy to the jet energy, normalized to the τ^{had} energy can provide a good measure if a systematic divergence occurs. This value is shown in Figure 4.18. One can see that the jet p_T is by construction almost always above the τ^{had} p_T. For the lower p_T region the jet reconstruction can be up to 100% higher than the τ^{had} p_T due to additional activity in the events. For higher p_T taus this difference gets smaller. But one can also see, that the τ^{had} energy and jet energy differ for individual τ^{had} candidates and the τ^{had} reconstruction provides the more precise energy reconstruction.

In Figure 4.19 (left) the mean of the distributions from Figure 4.18 is shown. On the left

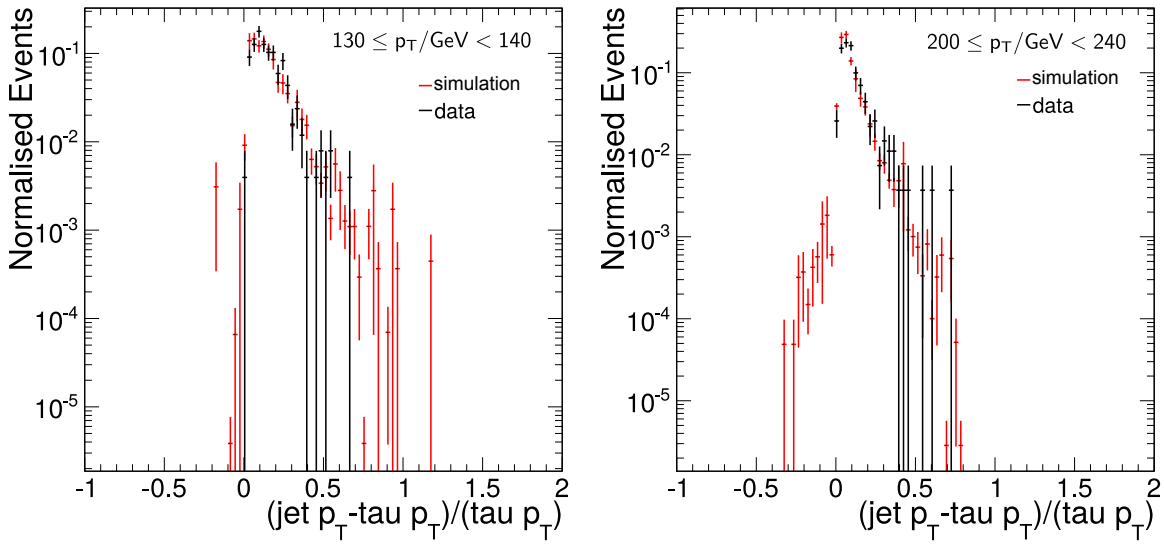


Figure 4.18.: The plots show the ratio of the jet and τ^{had} p_T difference to the τ^{had} p_T ($130 \leqslant p_T/\,\text{GeV} < 140$ and $200 \leqslant p_T/\,\text{GeV} < 240$).

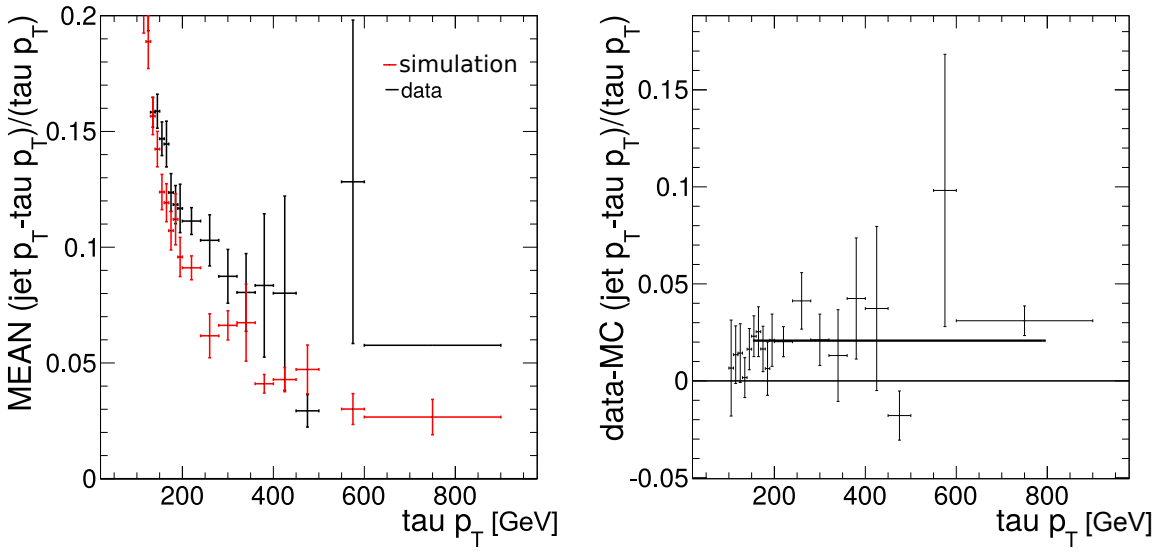


Figure 4.19.: The lower plots show the mean of the distributions in Figure 4.19 in the left plot and on the right the difference of data and simulation of the value in the left plot.

the mean shows that for low p_T τ^{had} s the jet p_T largely over-estimates the τ^{had} p_T. This is expected, since constant offsets in the jet energy from pileup and other activity in the events lead to a larger relative offset at low p_T. The error bars correspond to the RMS of the distribution, which is not defined for a single event and therefore the one event error is set to zero. In Figure 4.19 (right) the difference between simulation and data of the left plot is shown. This difference suffers from low event numbers at high p_T, but the difference can be found to be of the order of 2-3%, confirming the low p_T energy measurements [4, 111].

Since on data the τ^{had} energy scale can only be probed up to roughly 800 GeV the flatness of the energy response has to be shown with the simulation. This is shown in the appendix in Figure A.5 for the SM *W* boson background and various signal mass points. This shows that the τ^{had} reconstruction has an offset of 5% that is flat up to 1.5 TeV. The highest data event is around 600 GeV. This offset is not corrected for since it does not depend on p_T and the absolute energy scale of the τ^{had} does not allow more conclusions about the τ p_T.

For the Run II reconstruction, the background is simulated with more events and more detailed study of the τ^{had} energy response can be done. In Figure 4.20 various energy response differences are shown. In the left plot the difference between the reconstructed and generated τ^{had} is displayed in blue. As can be seen, the τ^{had} response is reasonably flat at high p_T, but shows a slightly rising structure from τ^{had} $p_T = 80$ GeV where the difference is roughly 10% to nearly no difference for a τ^{had} p_T at the TeV scale. For a τ^{had} p_T below 80 GeV the energy response has a large energy dependence. This is due to the trigger threshold, which requires a reconstructed τ^{had} with more than 80 GeV. In red the same is shown for the jet response. One can see, that the jet largely over estimates the τ^{had} p_T at low energies up to a τ^{had} p_T of roughly 200 GeV where the reconstruction agrees with the simulated τ^{had} energy. Above 200 GeV the jet energy response raises again w.r.t. the generated p_T and has a 2% offset at 2 TeV. In the right plot finally the difference between jet reconstruction and τ^{had} reconstruction is shown. For low p_T one can see a large, expected energy difference between jet and τ^{had} p_T. For higher τ^{had} p_T the energy measurements of the jet and τ^{had} agree on the 1-2% level.

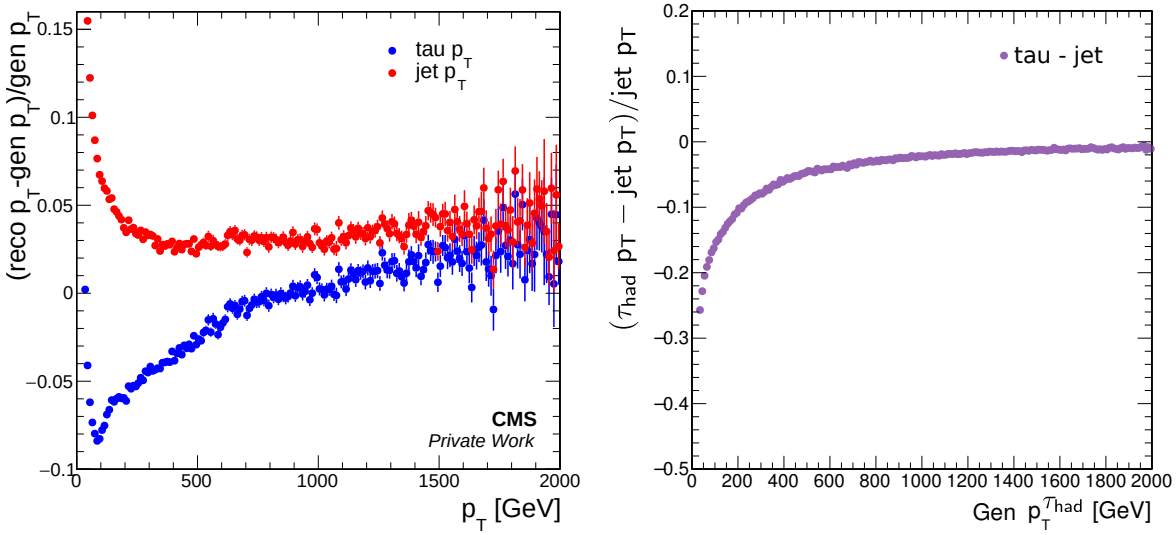


Figure 4.20.: The plots show the mean ratio of the generator matched taus with respect to the tau reconstruction and the jet reconstruction (left), and the ratio of jet to tau reconstruction (right).

On data the difference between jet energy and τ^{had} energy reconstruction is shown in Figure 4.21. While in the left plot the full selection described in Section 8.4 is applied, on the right plot only the preselection[8] is applied. This means no kinematic signal requirements are used. One can see that for most energy regions there is no difference between data and simulation. For the preselection plot a structure in the data energy response can be seen at 300-500 GeV. This is due to jets faking taus where the energy response is different. These events also lead to a large spread in the energy difference distributions, represented by the growing RMS of the data distribution.

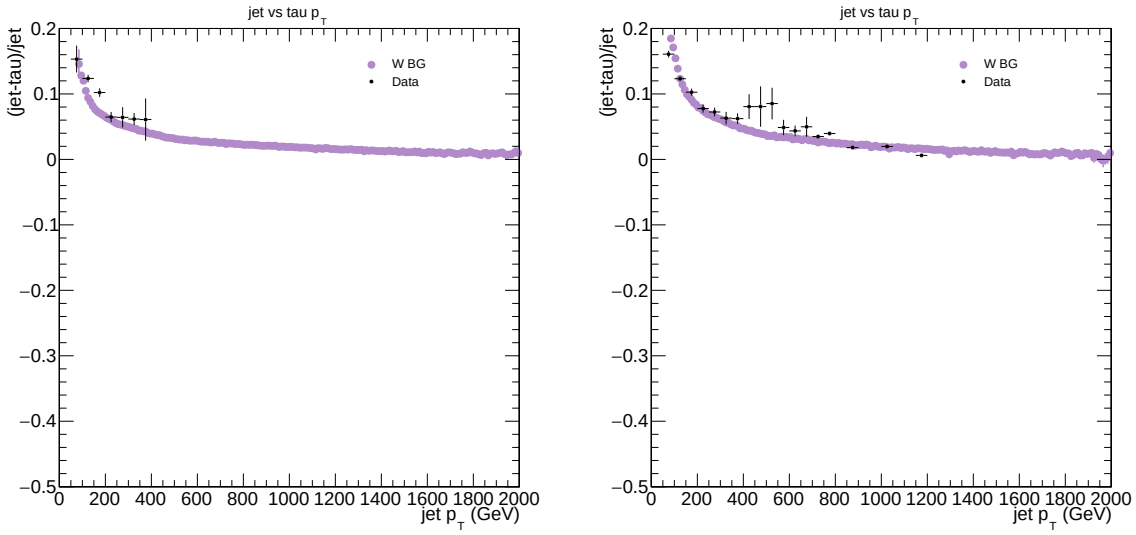


Figure 4.21.: Tau response shown in comparison with the jet response in data and simulation. This comparison is done for the events selected for the analysis as described in Section 8.4. In the left plot the full signal selection is applied and in the right plot only the one tau passing the trigger and identification is required.

Two effects can influence the results of this method. First it is possible, that one component is not reconstructed at all. This is suppressed by the fact that we can measure the one pion in different detector components. The simulation of jets could also be not accurate. This is anticipated in the jet reconstruction and the data to simulation differences are estimated with several physics processes. These differences between data and simulation for the jet reconstructions are small. The second effect is the influence of fake taus on the measurement. This can be suppressed by using a tighter working point of the isolation. Tighter working points have been checked and no difference to the τ^{had}- jet difference has been observed. This difference is 2%, thus compatible with the τ^{had} scale uncertainty estimated by the tau POG, which is 3% [111, 112]. An inefficiency in the τ^{had} reconstruction would also show up in the τ^{had} isolation efficiency and fake probability. Since the τ^{had} isolation has low thresholds for additional particles next to the τ^{had}, the taus would fail the isolation requirement.

[8] Required are trigger, and one isolated τ^{had} within the p_T and η acceptance.

5. Reconstruction of Other Objects

The main work during this thesis was focused to provide a good tau reconstruction. But in the analysis other physics objects are needed as well. In this chapter, the relevant objects will be introduced in the following sections.

5.1. Electron

For this thesis electrons are relevant as a discrimination against taus and in the context of the mono-lepton searches. As it is not the main focus of this thesis only an overview of the most important ideas is provided.

The electron reconstruction in the CMS detector [116] is driven by two aspects. The CMS tracker has a very high accuracy and large tracking volume and the ECAL has a very good energy resolution. For the identification algorithm this results in an approach from two sides. The tracks are reconstructed using a Kalman Filter and a Gaussian Sum Filter (GSF) method. The Kalman Filter is the standard track reconstruction algorithm in CMS where the GSF is a modification of the Kalman Filter, which takes into account the radiation of bremsstrahlung in the high magnetic field and relatively dense tracker. The tracks are an important input for the discrimination of electrons against photons, but also help with the cluster assignment in the ECAL. The other approach for the electron reconstruction is the reconstruction of cluster shape in the ECAL.

In the ECAL an electron shower is contained to 97% in a 5×5 crystal array, but due to the bending of the electron in the magnetic field and the continuous photon radiation a spread of the ECAL clusters in the ϕ direction is expected. These so called superclusters (SC) are built differently, depending on the ECAL shape in the barrel and endcap sections.

As a complementary approach to this clustering also a pfCluster ansatz is used where, regardless of the cluster shape ECAL clusters are built around a seed cells. Only cells that are above the noise level are used. With this approach energy deposits in a single crystal can be shared between two clusters, if the cluster shape or matched tracks indicate this.

The tracks and clusters are then combined using geometrical and resolution properties of the tracks and the clusters. All GSF tracks are linked with SCs, and for electrons, that are only seeded by a track, the pfCluster is used. The seeding of electrons in the ECAL can be challenging for instance in the transition region between barrel and endcap ($1.442 < |\eta_{SC}| < 1.56$), therefore this region is treated differently by some identification algorithms.

For the discrimination of electrons against other particles two different sets of cuts were used in this thesis. The cut-based identification, which aims to provide a general identifications

Quantity	cut value	
	($\lvert\eta\rvert < 1.442$)	($1.56 < \lvert\eta\rvert < 2.5$)
SC $E_T >$	35 GeV	35 GeV
$\lvert\Delta\eta_{in}\rvert$	0.005	0.007
$\lvert\Delta\phi_{in}\rvert$	0.06	0.06
$\sigma_{i\eta i\eta}$	-	0.03
E^{2x5}/E^{5x5}	0.94 or $E^{1x5}/E^{5x5} > 0.83$	-
EM + Had iso.	$2\,\text{GeV} + \rho \cdot 0.28 + 0.03 \cdot E_T$	$2.5\,\text{GeV} + \rho \cdot 0.28 + 0.03 \cdot (E_T - 50\,\text{GeV})$ ($E_T > 50$ GeV) or $2.5\,\text{GeV} + \rho \cdot 0.28$ ($E_T < 50$)
H/E	0.05	0.05
tracker iso.	5 GeV	5 GeV
exp. lost hits	$\leqslant 1$	$\leqslant 1$
$\lvert d_0\rvert$	0.02 cm	0.05 cm

Table 5.1.: The table summarizes the Selection criteria for the HEEP identification, as they were optimized for Run I (HEEP 4.1).

of all electrons, but it is mainly tuned for electrons with $p_T < 100\,\text{GeV}$, and the HEEP[1] identification, which was developed to provide an efficient identification of electrons with a $p_T > 100\,\text{GeV}$. While the cut-based identification is used in the τ+E_T^{miss} analysis to veto additional electrons in the event, the HEEP identification is used for electrons with $p_T > 100\,\text{GeV}$ in the context of the non resonant search. While for the HEEP identification only one working point exists, the cut-based isolation has several working points, which correspond to different identification efficiencies. For the τ+E_T^{miss} analysis the "veto" working point is used, while for the non resonant search, the "loose" working point has been used.

The values for the two discriminators are listed in Table 5.1 and 5.2 for Run I and Run II. The values for the two identifications are very similar and can be summarized in three categories.

The shape variables are geometric criteria, where the width of the signal in the ECAL crystals in η and ϕ is evaluated ($|\Delta\eta_{in}|$ and $|\Delta\phi_{in}|$), as well as the spread in the ECAL ($\sigma_{i\eta i\eta}$). The ratio of the core energy to the surrounding energy (E^{2x5}/E^{5x5} and E^{1x5}/E^{5x5}) also helps to identify electrons.

For the isolation two approaches are used. Where the HEEP identification uses ECAL and HCAL measurements directly, the cut-based identification uses pfCandidates to compute the isolation. For both isolations a $\Delta R < 0.4$ cone is used (for Run II the cone was reduced to 0.3 for the cut-based isolation) where the electron itself is excluded from the isolation quantity. The pfIsolation is corrected for pileup:

$$pfIso_{rel} = \left(\sum_{\Delta R<0.4} p_T(h) + \sum_{\Delta R<0.4} p_T(\gamma) + max(0, \sum_{\Delta R<0.4} p_T(h^0) - \rho \cdot A_{eff}) \right) / E_T. \quad (5.1)$$

The ρ correction is similar to the $\Delta\beta$ correction for the taus. Here ρ is the mean energy per area of the event and A_{eff} is the effective area of an electron. The area definition originates from the jet pileup subtraction and is described in [117]. The effective area is η dependent

[1] High Energy Electron Positron

and represents the usual size of an electron in this region of the detector. For the HEEP identification a single effective area (0.28) is used. One other isolation variable is the ratio of HCAL energy over ECAL energy (H/E), where high values indicate that the object is a charged hadron.

The photon variables like the expected, but lost, hits in the Pixel or the vertex probability are meant to discriminate electrons from photons. This is important, since the probability for photons to convert in the CMS tracker is 70% [118]. The expected lost hits variable checks, if a hit is expected in the innermost pixel layers, but not recorded, which rejects most of the photons. In addition for Run II a conversion veto was implemented, which checks if a converted photon was reconstructed, using this specific electron. The vertex cuts of the distance to the primary vertex in z direction ($|d_z|$), as well as in the transverse plane ($|d_0|$) is larger for converted photons than for electrons. But this also rejects electrons from pileup interactions. In Run I also the association to a vertex was checked separately. A reconstruction quality cut that is only used in the cut-based identification is the comparability of the track momentum measurement with the energy measurement. In order to not be affected by the worsening momentum resolution with higher p_T from bremsstrahlung and track reconstruction the difference of the inverse quantities is used ($|1/E - 1/p|$).

The electron reconstruction is highly efficient ($> 95\%$) and a little better in the barrel region ($|\eta| < 1.479$) than in the endcap region ($1.479 < |\eta| < 2.5$). The transition region[2] $1.442 < |\eta_{SC}| < 1.56$ is not used in the HEEP identification, to suppress misidentification of jets as electrons [119].

[2]Since for the HEEP reconstruction the direction of the SC is more important than the direction measured by the combined object, the η definitions of the endcap and barrel region is slightly different.

8 TeV cut values				
	($\lvert\eta\rvert < 1.479$)		($1.479 < \lvert\eta\rvert < 2.5$)	
working point	Veto	Loose	Veto	Loose
$\lvert\Delta\eta_{in}\rvert <$	0.007	0.007	0.01	0.009
$\lvert\Delta\phi_{in}\rvert <$	0.8	0.15	0.7	0.1
$\sigma_{i\eta i\eta} <$	0.01	0.01	0.03	0.03
H/E <	0.15	0.12	-	0.1
$pfIso_{rel} <$	0.15	0.15	0.15	0.15
$\lvert 1/E - 1/p\rvert$ GeV <	-	0.05	-	0.05
exp. lost hits <=	-	1	-	1
vertex fit prob <	-	0.000001	-	0.000001
$\lvert d_0\rvert$/ cm <	0.04	0.02	0.04	0.02
$\lvert d_z\rvert$/ cm <	0.2	0.2	0.2	0.2

13 TeV cut values		
	($\lvert\eta\rvert < 1.479$)	($1.479 < \lvert\eta\rvert < 2.5$)
working point	Veto	Veto
$\lvert\Delta\eta_{in}\rvert <$	0.0152	0.0113
$\lvert\Delta\phi_{in}\rvert <$	0.216	0.237
$\sigma_{i\eta i\eta} <$	0.0114	0.0352
H/E <	0.181	0.116
$pfIso_{rel} <$	0.126	0.144
$\lvert 1/E - 1/p\rvert$ GeV <	0.207	0.174
exp. lost hits <=	2	3
conversion veto	yes	yes
$\lvert d_0\rvert$/ cm <	0.0564	0.222
$\lvert d_z\rvert$/ cm <	0.472	0.921

Table 5.2.: Cut-based identification values as used for the Run I and Run II analysis. The "veto" working point is used for the τ+E_T^{miss} analysis and the "loose" working point is used for the generalized limits of the $e + E_T^{miss}$ final state.

5.2. Muon

The muon reconstruction in CMS is driven by the high magnetic field. Since muons are minimum ionizing particles for a large part of the energies covered by the CMS detector, muons leave the detector without fatal interactions[3]. This allows the measurement of muons only as tracks. In general the bending of the muon in the magnetic field is described by:

$$p_T = q \cdot r \times B \tag{5.2}$$

For a good muon resolution a large radius and a high magnetic field is needed. For CMS the magnetic field of 3.8 T provides a high curvature, even for high p_T muons. The muon identification therefore mainly requires quality criteria on the track fit. The fit uses a Kalman Filter and is done independently in the tracker and in the muon system. The used tight muon identification has the following requirements. The muon has to have a p_T >20 GeV and be within the muon systems acceptance of ($|\eta| < 2.4$). For the muon trigger the range is $|\eta| < 2.1$ for Run I. This is extended for Run II to $|\eta| < 2.4$. The muon has to be reconstructed in the muon system and in the tracker, with at least one hit in the pixel system and two matched muon stations and one muon segment. To guarantee a good resolution at least five hits in the tracker are required. In order to not include cosmic or pileup muons, the transverse impact parameter with respect to the beamspot has to be less than 0.2 cm, and the longitudinal distance of the tracker track w.r.t. the primary vertex has to be $dz < 5$ mm. In order to avoid ambiguities the muon is also required to be used by the particle flow algorithm as muon.

For non high p_T muons the $\chi^2/ndof$ of the track fit is required to be smaller than ten. For high p_T muons (only used for p_T >200 GeV) the track fit is modified. Three dedicated algorithms are used additionally for high p_T muons [120]. For high p_T muons there is a non negligible chance of showering in the iron of the magnetic return yoke. Therefore the tracks are refitted with the *tracker plus first muon station* (TPFMS) algorithm, which only uses the information from the tracker and the innermost muon station for the p_T measurement. The second algorithm is called *picky*. The idea is to find muon chambers with high occupancy and tag them as electromagnetic showers. For all hits in this chamber only those are picked for the muon, that have a low $\chi^2/ndof$. The third algorithm is the *dynamically truncated* algorithm and is only used for Run II. It either stops the Kalman-Filter algorithm in the first station where the energy loss is above a threshold, or skips the station if the next station is within the expectation.

Between these three algorithms and the standard muon reconstruction (tracker, tracker + muon system) the fit with the highest track probability is chosen. Fits where the error on the track $\Delta p_T/p_T$ is larger than 25% are ignored. If all fits have a worse resolution, the threshold is raised by 15%. For a detailed implementation see [121]. For the high p_T muon identification the selection is applied on the chosen track, and the resolution is required to have a $\Delta p_T/p_T < 30\%$.

For the muon isolation there are two choices, either a relative tracker based isolation $iso_{tk} = (\sum_{\text{tracks with } \Delta R<0.3} p_T(track))/p_T$, has to be smaller than 10%. This is used only for high p_T

[3]Muons can hit a nucleus of the detector material, and collide inelastically.

muons where ionisation deposits or showers from the muon could change the calorimeter measurement. The second choice uses the particle flow information, $iso_{pf} = (\sum_{\Delta R<0.4} p_T(h^{\pm}_{no\ pu}) + max(0, \sum_{\Delta R<0.4} p_T(h^0) + \sum_{\Delta R<0.4} p_T(\gamma) - 0.5 \sum_{\Delta R<0.4} p_T(h^{\pm}_{pu})))/p_T$. Here the charged hadrons are separated into hardrons from pileup interactions and hard interactions with the association to the primary vertex.

5.3. Jet

Hadronically decaying particles form jets in the detector. For the analyses presented here jets are not a direct observable, but many aspects like tau reconstruction and E_T^{miss} depend on the building of jets. Since the LHC is a hadron collider a good description of the hadronic activity in the detector is crucial. There are various jet algorithms that cluster the measurements in the detector to represent these hadronic decays [122]. In CMS the particle flow candidates are clustered with various algorithms [123, 124]. The standard algorithm is the anti-k_T algorithm where the key component is the distance definition between two candidates:

$$d_{ij} = min(k_{Ti}^{-2}, k_{Tj}^{-2}) \frac{\Delta R_{ij}^2}{R}. \tag{5.3}$$

The k_T is the transverse momentum of the protojet or the candidate, that is considered to be joint to the protojet. R is the distance parameter. It is set to 0.5 in Run I and to 0.4 in Run II. The algorithm starts with a seed and clusters objects i and j, which are closest to this seed and not too far away to the seed, combines these objects and calculates the next distance. The distance ΔR_{ij} is defined as before. The algorithm stops, when all objects have been clustered. If the distance to the beam defined as $d_{i,B} = k_{Ti}^{-2}$ is larger than d_{ij}, the object i is discarded.

The advantage of the anti-k_T algorithm with respect to e.g. the k_T algorithm - where not the inverse of the squared k_T is used, but k_T^2 - is that the clustered jets are collinear and infrared save. Collinear save means that the algorithm is not affected by collinear splitting of the hard partons, and infrared save means that soft QCD radiation is inside the defined distance cone.

One benefit of a modern detector and computing framework is the ability to calibrate the energy response of the raw data after the data was already reconstructed. In CMS there are different momentum correction levels that are obtained using different types of measurements. They are separated in offset corrections (L1), corrections depending on η and p_T (L2 and L3).

The L1 offset has evolved at CMS to reduce the effect of "in time" and out "of time" pileup, for details on pileup see Section 7.2.1. The offset correction is an additive correction, which is dependent on the mean energy deposit per area ρ in the event, η and the area of the jet in the detector. Further details can be found in [124].

The L2 and L3 corrections are obtained from simulation. Observed values are corrected to correspond to the true p_T as expected in the simulated events as best as possible. In addition to these corrections, dijet, $Z \to \mu\mu/ee$ and γ+jets events are employed to correct for data and simulation differences. These correction factors aim to remove differences in the balance of one jet to another (dijet), or the response of the ECAL or muon system with respect to the

rest of the detector ($Z \to \mu\mu/ee$ and γ+jets events). These differences are below 4% in p_T for $|\eta| < 3$.

This jet energy calibration decreases the jet energy scale uncertainty as well as the jet energy resolution.

The calibration and reconstruction of jets in the CMS detector has reached a high precision where the uncertainty on the jet p_T is at the 1% level for central jets, as shown in Figure 5.1. For forward jets the uncertainty increases slightly to $2-3\%$.

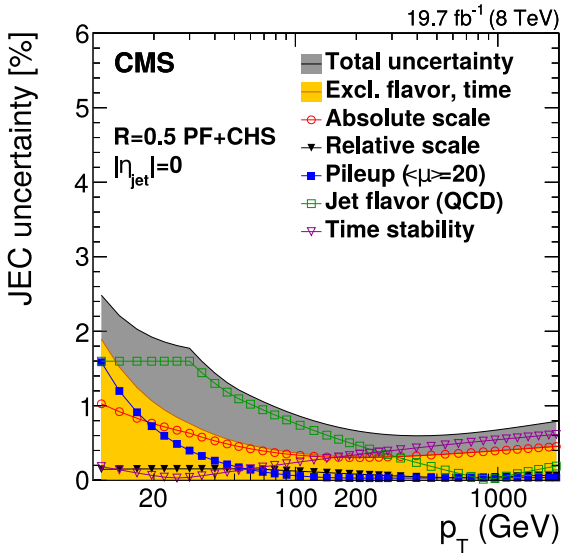


Figure 5.1.: The uncertainties of the jet energy calibration (JEC) for 8 TeV [124]

5.4. Missing Transverse Energy

In proton-proton collisions the longitudinal momentum of the partons is unknown, but the initial transverse momentum is negligible. Therefore the transverse momenta are conserved in the hard interaction, and the missing transverse momentum (E_T^{miss}) can be calculated. It is a measure for the particles that leave the detector undetected. In the standard model only the neutrinos contribute to this E_T^{miss}.

Technically the E_T^{miss} is defined as the modulus of the negative vectorial sum of the transverse momenta of all pfCandidates:

$$E_T^{miss} = |-\sum \vec{p}_T|. \tag{5.4}$$

Historically a calorimeter based $\vec{E}_T$ was used, therefore the abbreviation E_T^{miss} was kept. Per definition the azimuthal angle (θ or η) of the missing energy is not relevant, but the polar angle of the missing energy with respect to other particles in the detector is well defined and provides information that can be used. For that purpose $\vec{p}_T^{miss}$ will be used, defined as $\vec{p}_T^{miss} = -\sum \vec{p}_T$. The jet energy corrections are also transferred to E_T^{miss}, which is usually referred to as type-I corrected E_T^{miss}:

$$\vec{p}_T^{\,\text{miss}} = \vec{p}_{T,\,\text{raw}}^{\,\text{miss}} + \sum_i \vec{p}_{T,i}^{\,raw} - \sum_i \vec{p}_{T,i}^{\,corr} - \sum_i \vec{O}_i^{\,RC}, \tag{5.5}$$

where $\vec{p}_{T,\,\text{raw}}^{\,\text{miss}}$ is the uncorrected $\vec{p}_T^{\,\text{miss}}$, $\vec{p}_T^{\,raw}$ is the uncorrected jet p_T, $\vec{p}_T^{\,corr}$ is the fully corrected jet p_T, and $\vec{O}_i^{\,RC}$ is the average offset due to pileup, as obtained with a Random Cone method. For further details see [124]. The sum runs over all jets with $p_{T,corr} > 10\,\text{GeV}$ in the event.

It should be noted here that since the analyses considered in this thesis have multiple particles that produce missing energy in an event. For all events only one lepton is in the final state. Therefore the E_T^{miss} will largely only reflect the p_T of the one lepton. Hence, E_T^{miss} is also a measure for the additional activity in the event for this analysis.

6. Analysis Framework

The reconstruction and analysis of the recorded CMS data needs a development framework. The work flow of such a framework is more or less always the same. The huge amount of data collected by CMS is broken down in several steps to allow an analysis. For this thesis two different frameworks were used and largely developed during the course of this thesis.

The 8 TeV tau analysis will be described first, and then the 8 TeV dark matter and 13 TeV tau analysis framework will be outlined.

6.1. 8 TeV Analysis Framework

For the 8 TeV tau analysis the existing n-Tupleizer framework [125] was used on top of the AOD (Analysis Object Data) output of the CMS SoftWare (CMSSW). The main CMSSW version was 5.3.15. The previously existing n-Tupleizer was modified and rewritten for my master thesis and this analysis. The output format is a native ROOT [126] tree. The event size could be reduced from AOD with 400 KB/event to an n-tuple with 2 KB/event[1]. The speed[2] of the analysis is 1.6 ms/event. While this produces very small files and a fast analysis, the coding structure is not very flexible. All information is stored as a tuple of strings, integers, floats or doubles. While no other data type is needed, the cross reference is not easy. The information is ordered by indices, and one additional particle or a change in the ordering poses problems for versioning the analysis. In addition to this a C++ class file had to be updated for each version, which proofed to be error-prune. The analysis then used ROOT to fill histograms, smaller trees and other ROOT data formats.

6.2. 13 TeV Analysis Framework

In parallel to the ROOT n-Tupleizer, the MUSiC [127] group has used a similar tool, which used a more flexible implementation of the PXL library [128] output. Within this library it is possible to store particles and add arbitrary datatypes. String are used as keys to these objects. The more elegant solution had drawbacks in terms of size (15 KB/event) and speed[3] (8.8 ms/event). Nevertheless, this more flexible solution was chosen to be the starting point of a unified analysis framework "Three A Physics Analysis Software" (TAPAS) [129]. One of the many advantages is the object identifications or handling of systematic uncertainties, which are implemented once and can be used by various analysis groups. The file size could be reduced

[1] The reference numbers in this chapter are calculated using the $t\bar{t}$ events from POWHEG.
[2] All computing speed measurements is evaluated on an AMD FX-6100 Processor.
[3] The speed was evaluated on 8 TeV data with the latest version of the analyser.

to 12 KB/event at a speed of 4.2 ms/event. This has to be compared with the miniAOD format, which was introduced in CMSSW 6.X and contains all information which analyses needs, at 35 KB/event. The final 13 TeV analysis used CMSSW 7.6.3 with the miniAOD version `RunIIFall15MiniAODv2`. One of the other advantages of the TAPAS framework is that nearly the same code for the 8 TeV and the 13 TeV analysis can be used. The time to convert the CMS data into an analysable format was reduced from roughly one month to the order of 1-2 days. In a parallel effort the tools that accompany the analysis could be unified and shared within the work group.

As the documentation and description of algorithms can often not transport the full extend of the used programs the best description is often found in the source code of the program itself. For CMSSW this is available at [130] with the previously stated versions available as tags. For the analysis the source code is also published. The Run I analysis at [131] and for the Run II analysis at [132].

7. Tau Analysis at 8 TeV

For this thesis two data sets were analysed. They can be separated into Run I and Run II data. As the analyses are very similar for both data sets, the general analysis strategy will be introduced in this chapter with the Run I data. For the Run II analysis the analysis will be described in Chapter 8. In order to reduce the repetition, the focus of the Run II presentation will be the improvements and differences with respect to the Run I analysis. One of the aspects that will be described in detail for both analysis is the data driven estimation of the QCD jet background, since substantial changes were made for the Run II analysis. For both Chapters the structure will be similar: First the data recording conditions, simulations and selections will be introduced and then the data driven estimate will be described. At the end of each chapter the final distribution of the $\tau + E_T^{miss}$ final state in M_T will be provided with the full background prediction.

7.1. Dataset of pp Collisions Recorded in 2012

The Run I analysis is done with 19.71 fb^{-1} of $\sqrt{s} = 8$ TeV proton-proton collision data collected with the CMS detector in the year 2012. As described in Section 7.3, the jet and E_T^{miss} trigger is used[1]. The data has to pass a certification process to insure that all detector components have worked properly during the recording of the data. The latest certified dataset is used. It is the so called ReReco from 22nd of January 2013[2], which contains the latest calibrations, reconstructions and detector alignments known at that time.

7.2. Process Simulation

Many different samples, produced with MC generators, are used in this analysis to simulate the SM and signal processes.

Since the samples are produced with a number of events which does not match the integrated luminosity of the used collision data, the samples have to be weighted by a factor

$$w = \frac{\sigma \cdot \mathcal{L}}{N_{MC}} \tag{7.1}$$

[1] `HLT_MonoCentralPFJet80_PFMETnoMu95_NHEF0p95_v*` and `HLT_MonoCentralPFJet80_PFMETnoMu105_NHEF0p95_v*`

[2] With the certified data sets as provided in the Cert_190456-208686_8TeV_22Jan2013ReReco_Collisions12_JSON

where σ is the full cross section for the process, $\mathcal{L}$ is the integrated luminosity of the used collision data and N_{MC} the number of generated events.

At the time of the analysis most generators operate at LO, but for all used processes the inclusive (N)NLO cross sections are known. For the main W-boson background, these higher order correction were evaluated differentially as a function of the boson mass. This will be described in Section 7.2.3. As described before, the hard interactions have to be matched with a parton shower simulation. For the samples, PYTHIA 6.426 [27] was used. The parton shower and underlying event parametrizations were tuned with a portion of the 7 TeV data. The used tune is Z2*[133]. It is modifies the description of the underlying event response and the multiparton iterations to match the data. The shower parameters were not modified with respect to previous tunes determined at the Tevatron. The hadronized events are then processed by a detector simulation in order to simulate the response of individual detector components. For the simulation of the detector response GEANT4 [134] was used.

Since PYTHIA 6 has only a rudimentary implementation of the tau decay, the dedicated library TAUOLA [135] is used to simulate the decay. As noted in Section 4.1, the tau lepton has a long enough lifetime that it even may not decay before reaching the instrumented part of the detector. Therefore an interplay between TAUOLA, PYTHIA and GEANT is used. Taus with a flight distance of 10 mm or more are processed by GEANT, which simulates the effect of the magnetic field and possible interactions with the material. If the tau has no hard interaction, the decay of the tau is simulated by TAUOLA and hadronized with PYTHIA is then positioned at the end of the flight route, and processed further with GEANT. If GEANT simulates an inelastic scattering of the tau with the detector material, the decay and further detector responses are handled by GEANT. So far the simulation has reached an enormous precision and in particular all high p_T tau effects are described by it.

7.2.1. Pileup Simulation

At the high instantaneous luminosity of the LHC more than one proton-proton interaction occurs at each bunch crossing resulting in multiple interactions per event. These additional interactions are called pileup. The characterising variable for pileup is the number of vertices per event since each interaction per event should lead to at least one additional vertex. The number of interactions (and therefore number of vertices) per event depends on the instantaneous luminosity of the LHC at the interaction point and changes with time. The average number of pileup interactions per event is $N_{PU} \approx 20$. Most of these interactions are QCD processes with a low momentum transfer. These have to be considered in the simulation since many event and object features like the missing transverse energy and the tau isolation could be affected by pileup. The pileup is simulated by adding minimum bias[3] events to all simulated processes where the distribution of the number of interactions per event represents the distribution expected to occur in real data for a specific run period. Minimum bias events are generated to represent the average interaction at the LHC and are dominated by soft QCD processes. The pileup scenario used for all samples in this analysis is S10. Since the number of

[3]The term minimum bias originates from the trigger logic of colliders. Minimum bias triggers are usually random triggers that record an event with no requirements, except an expected bunch crossing. To simulate minimum bias events all SM processes enter the same generator sample. Due to the different expected cross sections minimum bias events are dominated by QCD processes.

pileup events changes continuously with rising luminosity, but the samples are generated for a whole run period, a procedure called pileup reweighting is used to adapt the distribution of the number of pileup interactions per event to the current value [136]. The difference between the abundance of a specific number of interactions per event in simulation and in the current data distribution is used to generate weights for each event depending on its expected interactions. These weights are used to reweight all simulated distributions. The number of pileup events in data could also be measured, using the number of reconstructed vertices. This would introduce a potential systematic, if the vertex reconstruction efficiency is different in data and simulation. The number of pileup events is instead calculated from the instantaneous luminosity and the total inelastic proton-proton cross section.

7.2.2. Background Samples

All SM processes which can lead to an experimental signature similar to the one of a W' boson have to be considered as a possible source of background events. The main source of background is the off-shell, high mass region of the SM W-boson leading to the same diagram as a W'-boson decay. Other backgrounds are due to QCD multijet processes, $t\bar{t}$ and single top production, dibosons and Drell-Yan (DY) events. One disadvantage of a final state with only one measurable particle in the detector is the limited number of selections one can apply to purify the signal. This also limits the possibilities to estimate the background from control regions. The W-boson process as the leading background cannot be estimated from data, because no signal free control region is constructable. Therefore, this analysis relies heavily on the simulation of this background and much emphasis was directed to get the best possible simulation of W-boson events. The background estimation of the multijet and $Z \to \nu\nu$ + jets processes will be done with a data driven approach and is described in Section 7.5.

The following sources of background are included (The full overview of sample names, cross section and event numbers can be seen in Table A.1 in the appendix):

- $W \to \ell\nu$ ($\ell = e, \mu, \tau$) The simulation of the main W background is split into two regions. The low energy region which is described with events generated with MadGraph at LO and the high mass samples, which are generated with pythia. The low energy region describes the events with $M_W < 200\,\text{GeV}$, the samples are divided in bins of $p_T(W)$. The p_T bins are 0-50 GeV, 50-70 GeV, 70-100 GeV, 100-180 GeV and 180 GeV to infinity. Here MadGraph was chosen, because additional jets are simulated at tree level, hence a better description is given. For the high mass events the description of additional jets is not as important, hence the simpler pythia simulation can be used. The high mass samples are binned in $\hat{p}_T$. $\hat{p}_T$ is a quantity used in the core of the pythia simulation, which is defined as the transverse momentum in the rest frame of the hard interaction [27][4]. The used bins are $\hat{p}_T$ from 100 GeV to 500 GeV and 500 GeV to infinity, which can be roughly translated to the W-mass with $m \approx 2\hat{p}_T$.
- Top pair and single top production is another potential source of high p_T leptons and E_T^{miss}, and has to be considered as a background process. All the samples have been produced with powheg.

[4] While it is a well defined quantity in the pythia generator, the choice this quantity was not wise. A generator independent quantity like the invariant mass of the process would have been a better choice.

- DY decays to taus (generated with PYTHIA) are a source for background events if one tau is not identified and the energy enough to fake a sufficient amount of E_T^{miss}. The other DY decay channels only contribute to the signal with a high p_T recoil jet faking a tau. Therefore MADGRAPH is used to describe these events. Here p_T (Z) bins are used to divide the samples. They are 0-50 GeV, 50-70 GeV, 70-100 GeV, 100-180 GeV and 180 GeV to infinity. The PYTHIA tau sample is used to describe the high mass region in Z-mass bins of 100-200 GeV, 200-400 GeV, 400-800 GeV, 800-1000 GeV and 1000 GeV to infinity.

- Diboson processes (WW, WZ, ZZ) with W/Z decaying to anything are generated with PYTHIA. To improve the statistics at high mass, samples with $\hat{p}_T > 500$ GeV are produced additionally.

- QCD multijet processes are an important source of background since they have the highest production cross section at the LHC and a jet can fake a HPS tau. The samples are generated with PYTHIA.

- $Z \to \nu\nu$ +jet events also provide a background, if the jet fakes a τ^{had}. The samples are generated with MADGRAPH and are divided by the p_T (Z) at generator level. The used bins are 70-100 GeV and 100 GeV to infinity. The lower momentum events would not be seen in the detector.

All samples have been generated in leading order except for the top samples which have been generated in next-to leading order. The DY and top samples are scaled to next-to-next-to leading order QCD corrected cross sections by an energy independent correction factor and the Diboson samples are scaled to next-to leading order QCD cross sections. The QCD samples cannot be scaled to higher order corrections since the theoretical and statistical uncertainties are too large to make this approach sensible. The QCD jet processes (QCD multijet and $Z \to \nu\nu$ +jet) are not used for the background determination from QCD processes which is done by a data driven method instead (see Section 7.5). They are used only for cross checks of this method. All inclusive higher order cross sections are taken from [137].

7.2.3. Higher Order Corrections for SM W

The higher order corrections of the $pp \to W \to \tau\nu$ +jets process can have a significant impact on the differential distributions. Not only the QCD corrections from higher order calculations can change the differential distributions but also the electroweak corrections can become large due to their large Sudakov logarithms. Up to now, no method was found which allows the matrix element level combination of QCD and electroweak corrections with reasonable expense. Therefore, the corrections are calculated individually and are then combined. As the differential parameter the invariant mass of the W-boson process was used, because it is a generator independent, well defined quantity. Since the combination of the QCD and electroweak corrections should be combined on matrix element level, which is not done so far, a method has been developed to combine both corrections [138, 139]. The corrections are

combined additively ($\left[\frac{d\sigma}{dM}\right]_{QCD \oplus EW}$) or factorized ($\left[\frac{d\sigma}{dM}\right]_{QCD \otimes EW}$):

$$\left[\frac{d\sigma}{dM}\right]_{QCD \oplus EW} = \left[\frac{d\sigma}{dM}\right]_{QCD} + \left[\frac{d\sigma}{dM}\right]_{EW} - \left[\frac{d\sigma}{dM}\right]_{LO} \tag{7.2}$$

$$\left[\frac{d\sigma}{dM}\right]_{QCD \otimes EW} = \left(\frac{\left[\frac{d\sigma}{dM}\right]_{QCD}}{\left[\frac{d\sigma}{dM}\right]_{LO}}\right) \times \left[\frac{d\sigma}{dM}\right]_{EW} \tag{7.3}$$

The additive approach assumes that the electroweak corrections except final state radiations have an additive nature and the same correction needs to be added for all orders of QCD corrections, while the factorized approach assumes that the higher order electroweak corrections are the same for all orders of QCD. Following the recommendation from [139], the additive approach is used for combining the corrections, and the difference to the factorial approach is used as an uncertainty.

The NNLO QCD differential cross section for the SM W can be calculated with different programs such as FEWZ [140] and DYNNLO [141]. While in the latest FEWZ version (3.1b2), the NLO EW correction for the neutral Drell-Yan process is implemented, these corrections are not implemented for the W-Boson production.

The NLO EW differential cross section can be calculated with separate programs such as HORACE [142] and MCSANC [39]. The NLO prediction of the recent HORRACE version 3.1 and MCSANC 1.01 disagree for high invariant masses, but this is a wrong implementation of HORRACE and could be resolved by a correction to the HORRACE code[5]. A comparison is shown in Figure 7.1(left). For taus (and also for electrons) the reconstruction of the other objects in the detector is also of interest, because, for example, the radiated photons will be reconstructed as part of the tau (or electron). Hence photons that are in a $\Delta R < 0.1$ cone around the tau are recombined. The impact of these recombinations is shown in Figure 7.1 (right). The tendency is that the about 2% higher NLO EW cross section is due to the infrared safe tau measurement in the detector. A wider cone had no significant effect on the cross section.

For the neutral Drell-Yan process ($pp \to Z/\gamma \to ll$) an interesting higher order electroweak effect occurs due to an additional initial state. In the past only the QCD contributions where taken into account when PDFs where studied. Since the PDF calculations have reached a high accuracy also the electroweak interactions in the proton became of interest. For the ll final state a t-channel diagram with two photons in the initial state is possible. This diagram is not negligible for high η leptons with a high invariant ll-mass, and depends heavily on the QED modelling of the PDFs. For the charged Drell-Yan process the contribution is expected to be small since no tree diagram $l\nu$ production is possible.

In Figure 7.2 the ratio of the (N)NLO to the LO differential cross section is shown. For the high invariant mass region the two figures visualize the effect of the PDF on this region. For the left Figure the LO cross section is calculated with the LO PDF set "CTEQ6L1" as it is also used for the PYTHIA simulations. In the right plot of Figure 7.2 the LO cross section is calculated with a NNLO PDF set "CT10NNLO". As one can see the curved structure at high masses is manly due to a difference in the PDF description of the events at high $M_{\tau\nu}$. The corrections are also large around the W-boson peak, which is due to the difference of LO to

[5]The correction was obtained after a private discussion with the authors.

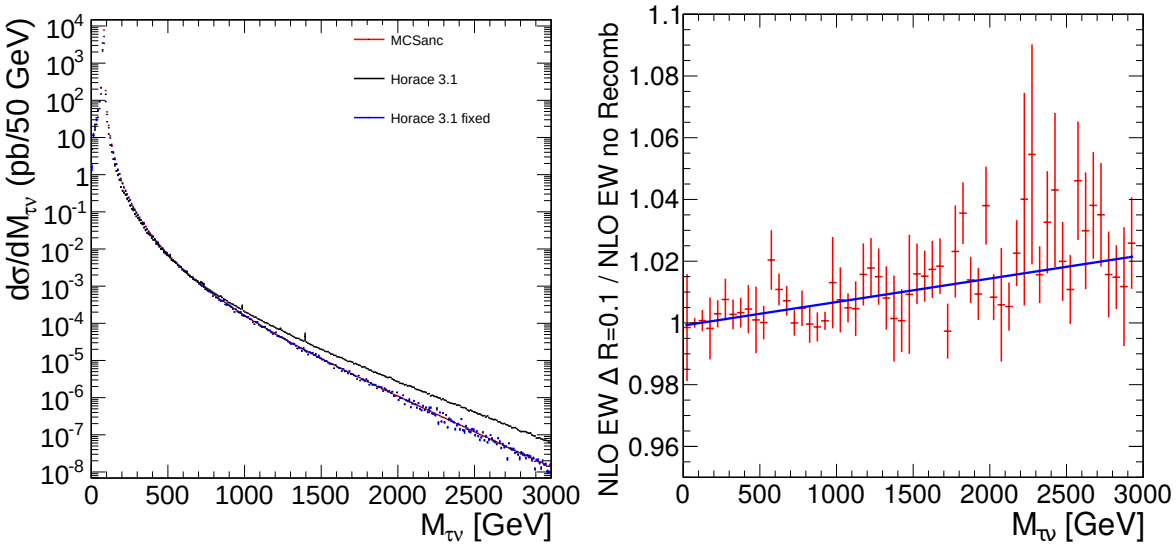


Figure 7.1.: A comparison of different NLO EW cross sections for different NLO programs (left) and the impact of tau recombination in the calorimeter (right). The dots represent the numerical difference and the line shows a linear fit to the difference.

NLO diagrams, where more production channels and initial and final state radiation can shift the position of the W-boson peak in the distribution. However, this shift does not have a large effect for the used high mass regions. For the NLO electroweak corrections the photon induced processes where also considered using the NNPDF set "NNLPDF 2.3 NNLO α_s QED".

As we have seen in Figure 2.4 the prediction from different PDF groups for high x is not consistent at the moment. This can have a large influence on the background and signal prediction. A large improvement for this final state is the evolution of the DGLAP equations at NNLO, because high Q^2 events can be described in detail.

7.2.4. Signal Samples

In order to determine the selection efficiency of the analysis for the various W' models considered in this analysis, many signal samples have been produced. The SSM samples are generated with PYTHIA in a mass range between 300 GeV and 4 TeV and the NUGIM samples are generated with MadGraph in the mass range between 1 TeV and 3.4 TeV for the mixing angle parameters $\sin^2\theta_E = 0.031, 0.04, 0.05, 0.1, 0.2, 0.3$. The cross sections for the different samples can be seen in Table A.2. All samples are produced in LO with the PDF set "CTEQ6L1" [143], and for the SSM samples cross sections in NNLO QCD have been calculated [144].

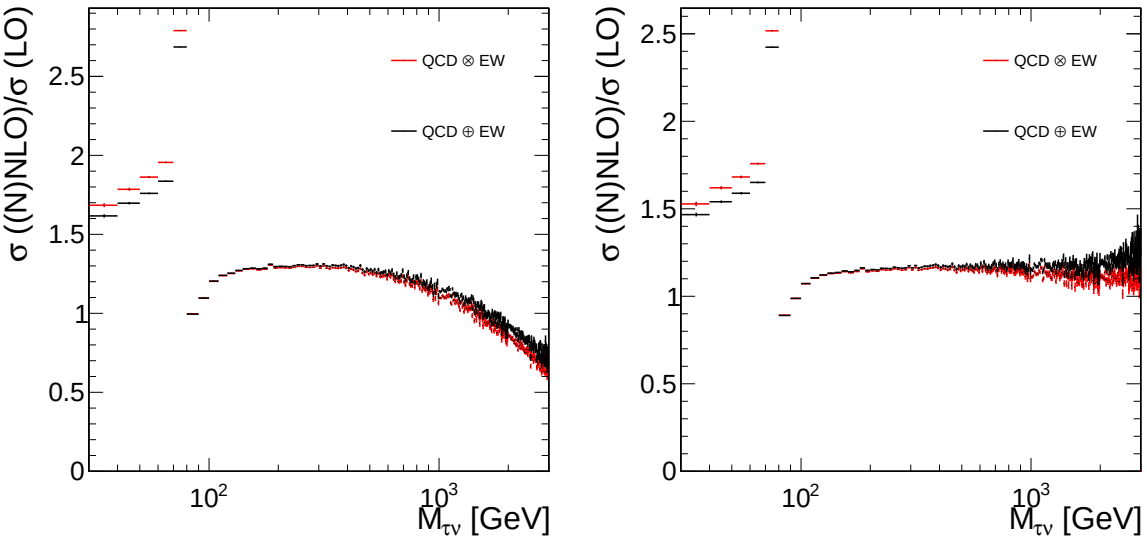


Figure 7.2.: k-factors for the additive and factorial combination of NNLO QCD and NLO EW corrections. In the left plot normalized to LO PDF (CTEQ6L1) and in the right plot normalized to NNLO PDF (CT10NNLO).

7.3. The Trigger

There are several triggers that can be used to search for a τ+E_T^{miss} final state. The signal can be separated into two physics objects in the detector, E_T^{miss} and τ. For the 2012 operations there are two possible approaches to trigger the signal: Trigger the individual objects or use a combination of both objects to trigger. For triggering the individual objects in the events only a E_T^{miss} trigger is usable, since a single tau trigger would have a too high trigger rate. The alternative is a cross trigger, which combines τ^{had} and E_T^{miss} in the HLT. This was developed before the start of this thesis and ran for the 2012 data taking period. But the trigger has had an inefficiency for high p_T taus ($p_T^{had} > 100\,GeV$), which is documented in [14]. Therefore three triggers will be tested for the use in this analysis. Fist the situation of the tau+E_T^{miss} trigger will be stated, then the E_T^{miss} trigger will be tested, and finally the a jet+E_T^{miss} trigger will be presented. The jet+E_T^{miss} trigger is the best choice for this analysis, given the not working tau+E_T^{miss} trigger.

7.3.1. The Tau + E_T^{miss} Trigger

A dedicated trigger for the tau + E_T^{miss} final state was developed for the charged Higgs search, where the tau lepton has a much lower p_T^{had} and is not balanced with E_T^{miss}. Due to a mistake in the configuration, the track of the tau had to have a momentum uncertainty 0.5 times smaller than the uncertainty measured in the calorimeter. As a consequence, the trigger efficiency above 100 GeV dropped to zero. The trigger was studied in detail in [14].

7.3.2. The E_T^{miss} Trigger

There are two main E_T^{miss} triggers in CMS. The calorimeter based E_T^{miss} trigger with a threshold of 120 GeV on HLT level and the pfE_T^{miss} trigger with a threshold of 150 GeV on HLT level. To validate the trigger turn on the working cross trigger was used. The results can be seen in Figure 7.3. The trigger efficiency is defined as the number of evens in a specific E_T^{miss} bin passing the E_T^{miss} trigger, divided by the number of events passing the tau-E_T^{miss} cross trigger. The errors and efficiencies shown for the data points are estimated using Clopper-Pearson confidence intervals. For the simulation this is not possible, because weighted events are used. The shown errors are estimated using binomial statistics as implemented in ROOT [126]. In order to get a correction function for the simulation a Gauss error function was fitted:

$$\frac{a}{2}(1 + \mathrm{Erf}(\frac{x-b}{c\sqrt{2}})) \tag{7.4}$$

The fit depends on the error estimated with the statistical approach. For weighted events the binomial approach is correct for high efficiencies, while the gaussian error is correct for low efficiencies. Therefore the fit was performed using both methods. The fit to the gaussian efficiency estimation has a better agreement at high E_T^{miss} to the data fit and reflecting the statistical uncertainties at high E_T^{miss}. For these distributions the minimal requirement was one well reconstructed tau, see 7.4. The E_T^{miss} is offline pfE_T^{miss} with type I correction. The fit results are in the appendix Table A.3. The plots show that even with a higher threshold of 150 GeV the pfE_T^{miss}, the turn on is steeper and flat above ~200 GeV.

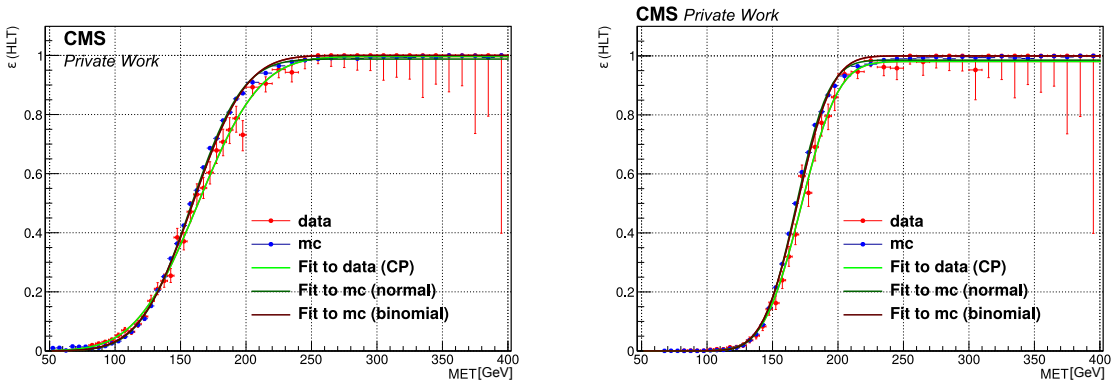


Figure 7.3.: The E_T^{miss} trigger turn on curve for calorimeter based E_T^{miss} (left) and pfE_T^{miss} (right). The turn on was measured with the tau E_T^{miss} trigger (tau $p_T >$ 35 GeV and pf$E_T^{miss} >$ 70 GeV), which has a much lower threshold in E_T^{miss} and therefore can be seen as independent.

7.3.3. The Jet + E_T^{miss} Trigger

An interesting trigger for this analysis is the `HLT_MonoCentralPFJet80_PFMETnoMu95_NHEF0p95_v*` and the later version `HLT_MonoCentralPFJet80_PFMETnoMu105_NHEF0p95_v*`. The trigger components can be broken down to at least one jet with $p_T >$ 80 GeV and $E_T^{miss} >$ 95 GeV (105 GeV in the later runs). The L1 seed is `L1_ETM40`. Both the jet and the E_T^{miss} are

reconstructed in the HLT, using the pfCandidates. The jet has to be central ($|\eta| < 2.6$) and with a p_T above 80 GeV. In order to avoid beam halo effects or other unphysical effects the neutral energy fraction in the jet is constrained to be less than 0.95. The E_T^{miss} is reconstructed without taking muon candidates into account.

The trigger can be used for triggering taus, since none of the constrains on the jet would constrain a tau. Every tau is constructed from a jet the jet could be used as a reference object.

This trigger has several advantages with respect to the E_T^{miss} triggered events. The E_T^{miss} threshold is much lower, and the additional requirement on a central jet ($|\eta| < 3$) does not bias the selection, since the τ^{had} can also be reconstructed as a jet and a fixed ratio between τ^{had} p_T and E_T^{miss} is expected in the signals, and therefore is required as a selection criteria, see Section 7.4. The disadvantage is that an independent validation of the trigger is difficult. One can see the different turn on thresholds for the three described triggers for p_T and E_T^{miss} in Figure 7.4, which is evaluated on simulated events. The jet-E_T^{miss} cross trigger has the lowest threshold in tau p_T^{had}. The turn on in the tau p_T^{had} and E_T^{miss} is due to the correlation with the jet cut off of 80 GeV. In the E_T^{miss} distribution the tau-E_T^{miss} cross trigger has the lowest threshold. In order to avoid turn on effects the jet-E_T^{miss} cross trigger is used in combination with a jet $p_T > 100$ GeV and $E_T^{miss} > 140$ GeV.

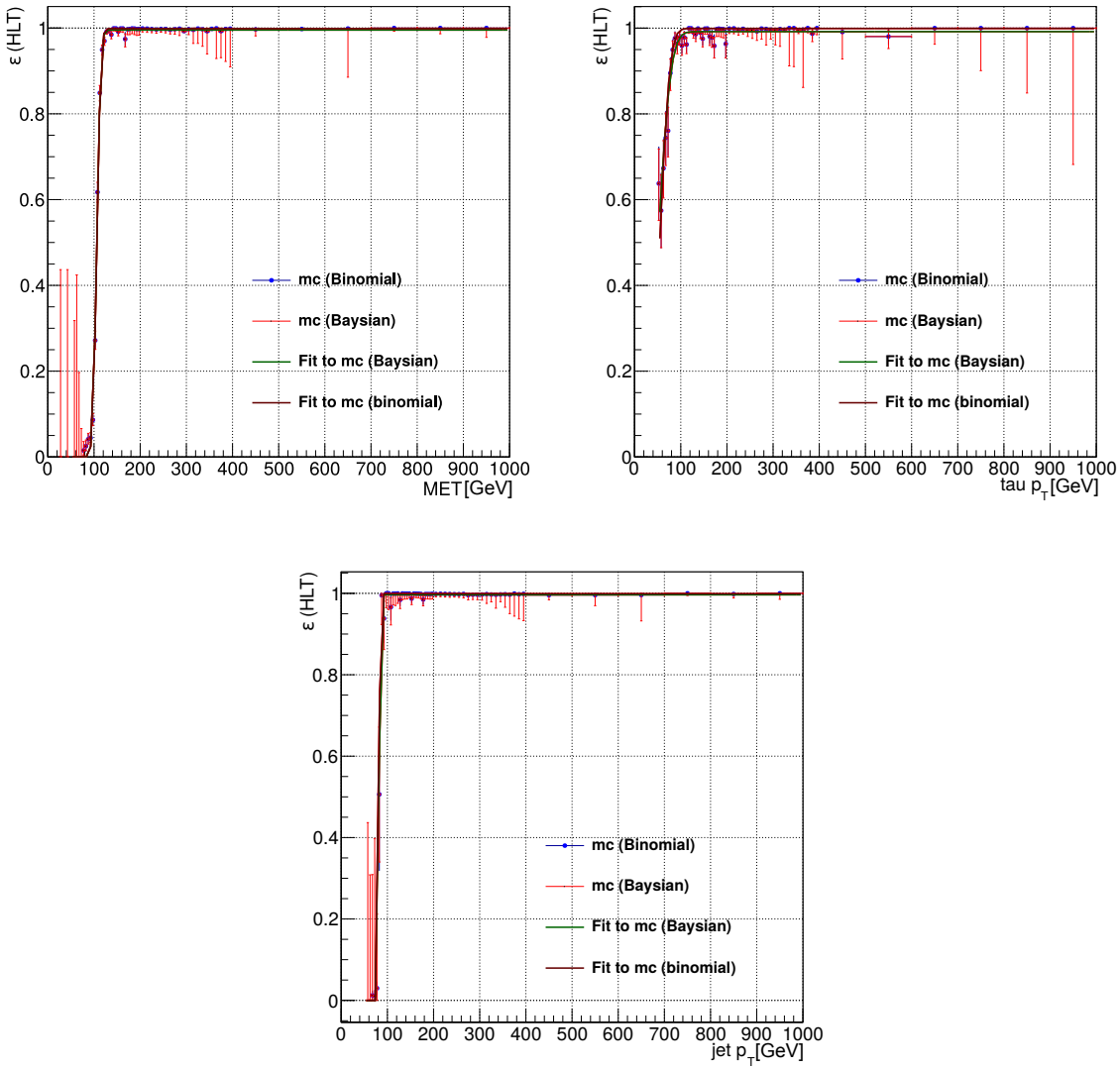


Figure 7.4.: Jet-E_T^{miss} trigger turn on is shown for simulated events. For the simulation W events were considered and the $E_T^{miss} > 140\,GeV$ cut was applied for all events in the jet and tau p_T^{had} plot, vice versa for the jet $p_T > 100\,GeV$ cut. The plots show the efficiency for E_T^{miss} tau p_T^{had}, jet p_T from left to right.

7.4. Event Selection

7.4.1. General Event Selection

The recorded events are required to fulfil a few general event criteria, which ensure the correct reconstruction of the events. As discussed before the trigger requires at HLT at least one jet with $p_T > 80\,\text{GeV}$ and $E_T^{\text{miss}} > 105\,\text{GeV}$. This sets the offline thresholds for the analysis to at least one jet with $p_T > 100\,\text{GeV}$ and $E_T^{\text{miss}} > 140\,\text{GeV}$.

To connect the jet with the expected tau in the event, exactly one tau with a τ^{had} $p_T > 80\,\text{GeV}$ and $|\eta| < 2.3$ is required to be in the event. The tau p_T^{had} can be reduced with respect to the jet p_T, because hadronically decayed taus are reconstructed as a subset of the jet constituents. The selection will not be 100% efficient, but the already contained phase space is not further unnecessarily reduced. For additional, taus used as veto, the p_T^{had} can be reduced to 50 GeV, where an the identification efficiency and miss identification rate are acceptable. The taus have to fulfil several identification requirements: The decay mode reconstruction has to have provided a valid decay mode, including the two prong decay mode. The electron rejection has to pass the loose MVA based identification working point[6]. The muon rejection has to pass the loose cut-based working point[7] and as isolation requirement the loose cut-based identification[8] was used. In order to have a comparison to the MVA based isolation, some results are also given for the medium MVA-based working point[9]. But since this is the first search for a W' signal in this channel and the isolation is an important discriminator against background, the cut-based isolation is preferred to the MVA-based isolation. This is done for the cut-based isolation all input parameters are well controlled and a MVA output for high p_T taus could have an unforeseen behaviour.

To further reduce electroweak background, a veto on events containing muons, that fulfil the (non high p_T, tight) identification and are isolated with $\text{iso}_{\text{pf}} < 12\%$. Also events with electrons that fulfil the veto working point are rejected. For both leptons the minimal allowed p_T is 25 GeV and the η region is according to the detector coverage $|\eta| < 2.4$ for muons and $|\eta| < 2.5$ for electrons.

This selection should be enough to guarantee a good reconstruction of all events. But since the detector does not behave like a MC generator some additional requirements are recommended that ensure no electronic noise or reconstruction effect enters in the analysed events. These filters where mainly developed to reduce events with large E_T^{miss}. These filters are [145]:

- The **CSC tight beam halo** filter which reduces events, where collisions, which occur before the detector produce a beam halo in the detector. They can be detected with the CSCs.
- The **HB HE noise** filter reduces events, where unusual noise rates in the HCAL were detected without any additional activity in the events.

[6] *againstElectronLooseMVA5*
[7] *againstMuonLoose3*
[8] *byLooseCombinedIsolationDeltaBetaCorr3Hits*
[9] *byMediumIsolationMVA3newDMwLT*

- The **HCAL laser** filter reduces events, where the HCAL calibration laser fired during the data recording.
- The **ECAL dead cell trigger primitive** filter reduces events, where ECAL crystals, that are removed from the reconstruction due to problems in the readout, are hit.
- The **Tracking failure** filter reduces events, where no tracks but large calorimeter deposits where observed.
- The **Bad EE Supercrystal** filter reduces, where some 5×5 crystal regions give anomalously high energies.
- The **ECAL laser correction** filter reduces events, where for some crystals unusually large transparency corrections were observed.
- The **Additional tracking** filter reduces events, where aborted track reconstruction and coherent tracker noise is caught.

While the list is quite long, the final state of tau+E_T^{miss} is not affected by these kinds of events. One additional requirement is kept from the past which has no effect in the conditions of the 2012 data taking: Events are required to have at least one vertex. For events which have on average 20 vertices, this is no constraint.

7.4.2. Kinematic Event Selection

In order to select W' boson like events, the topology of the signal has to be evaluated. The kinematic distribution of W'-boson events, is indistinguishable from a W-boson event produced off shell with a high mass. For the lepton channels the only objects visible in the detector are the leptons. All neutrinos contribute to the E_T^{miss} in the event. For a hadronic tau decay this leaves basically a mono-jet signature with the jet tagged as tau. The event is still balanced in tau and E_T^{miss} allowing to exploit the two body kinematics. A distribution of the transverse mass on generator level and for reconstructed events is shown in Figure 2.5 and 7.5. The transverse mass is defined as in equation 2.20. One can see in Figure 7.5 that the visible part of the tau combined with E_T^{miss} does not have a peak structure. In contrast to previous generator plots Figure 7.5, the effect of the trigger threshold can be seen.

Two quantities can be used to further exploit the back-to-back, two body kinematics of the final state: The azimuthal angle between E_T^{miss} and the reconstructed τ^{had} ($\Delta\phi(\tau^{had}, E_T^{miss})$) and the ratio of the energies p_T^{had}/E_T^{miss}. The two body kinematics can be used even with two neutrinos in the final state since the only detectable object in a signal event should be the τ^{had}. In principal, these quantities are heavily correlated with M_T and are not needed to achieve clear signal separation at high M_T. But they provide a reduction of two types of events: The background contamination from di-jet events is largely reduced with the p_T^{had}/E_T^{miss} selection. This ensures that the object with the largest p_T in the event is the τ^{had}. For di-jet and Drell-Yan events an increase with rising p_T^{had}/E_T^{miss} is expected. On the other hand, if the p_T^{had} is small and the E_T^{miss} is large, the source of the E_T^{miss} is not the τ^{had}. The angle between E_T^{miss} and the lepton is heavily effected by the boost of the initial heavy particle. The W-boson events with high W-p_T are affected more severely by a correct QCD description. These events can have a high M_T if the E_T^{miss} is shifted by the recoil jet, wrongly identified as τ^{had}.

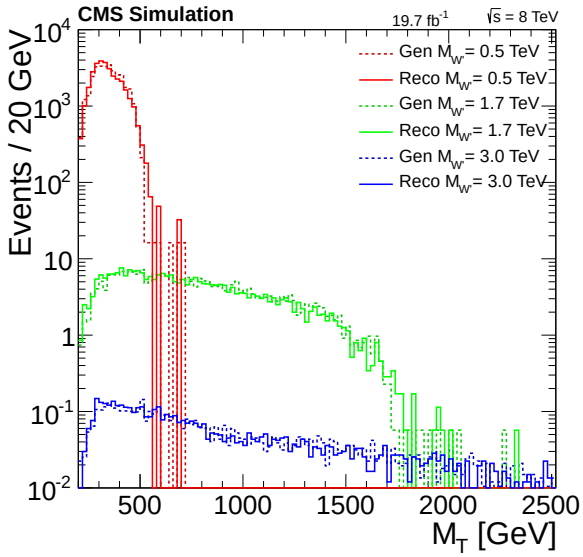


Figure 7.5.: The plot shows that the generated and reconstructed W' signals agree up to the highest energies. No Jacobian peak is visible and the signal shape is a flat, almost rectangular distribution, between zero and the W' mass. The low M_T side is reduced due to the trigger threshold.

As signal selection the following selections are chosen:

- $0.7 < p_T^{had}/E_T^{miss} < 1.5$
- $\Delta\phi(\tau^{had}, E_T^{miss}) > 2.4$

These selections where chosen on the basis of the selection of the previously existing $W' \to e/\mu\nu$ analysis [6]. As the tau + E_T^{miss} final state suffers a higher background from miss reconstructed objects the lower threshold of p_T^{had}/E_T^{miss} and $\Delta\phi(\tau^{had}, E_T^{miss})$ is raised. The effect of this selection on the signal and background was checked and found to reduce the background efficiently, while not too constraining on the specific signal.

Additional studies using additional selections with quantities involving the leading jet (not identified as tau) showed no significant improvement.

The kinematic distributions for p_T^{had}/E_T^{miss} and $\Delta\phi(\tau^{had}, E_T^{miss})$ are shown in Figure 7.6. The distributions show that the background describes the data well. The systematic uncertainties are evaluated bin to bin and are described in Section 9.1. For the shown distribution the cut on the other shown quantity is applied. One can see the signal has clear peaks at $p_T^{had}/E_T^{miss} \approx 1$ and $\Delta\phi(\tau^{had}, E_T^{miss}) \approx \pi$. In the low p_T^{had}/E_T^{miss} region one can see that the contribution from missidentified taus is high. For large p_T^{had}/E_T^{miss} the events are completely described by QCD jets, which are evaluated from data as described in Section 7.5. For the $\Delta\phi(\tau^{had}, E_T^{miss})$ distribution a surprisingly good description of the data by the background is given, as the angle between τ^{had} and E_T^{miss} is influenced by the modelling of the additional activity in the event. For the high W mass PYTHIA sample this means that no jets are calculated on matrix element level, but the description is parametrised by the tune of the generator.

The signal efficiencies for SSM W' are shown in Figure 7.7 for different W' masses and various stages in the analysis. Where in the left plot all efficiencies are calculated with respect to the

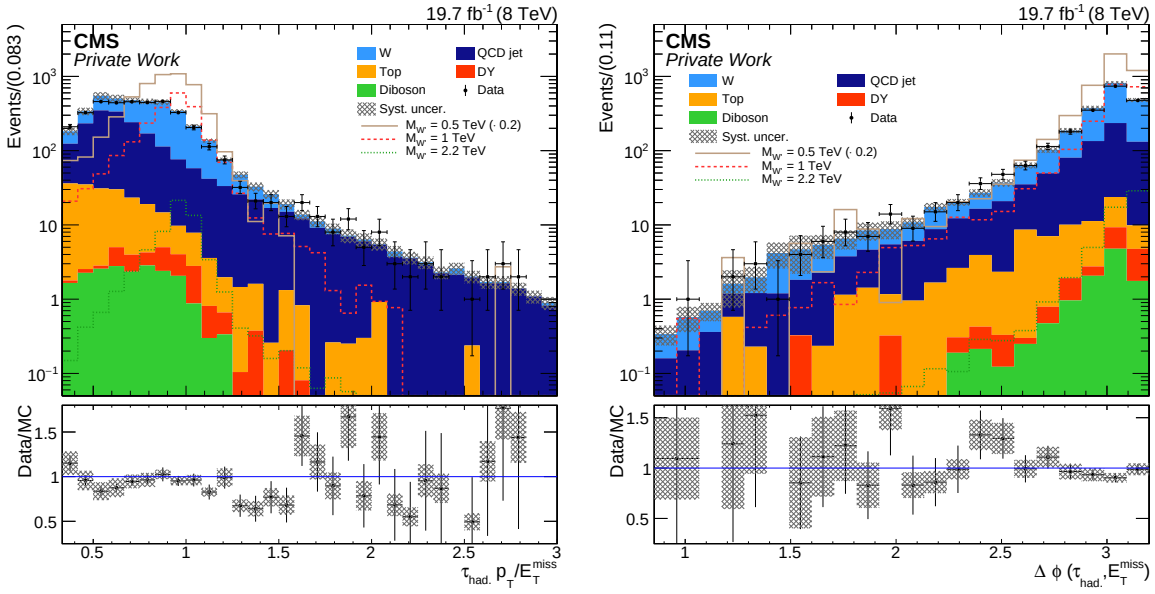


Figure 7.6.: In the plots the simulated backgrounds are shown for the kinematic distributions for p_T^{had}/E_T^{miss} (left) and $\Delta\phi(\tau^{had}, E_T^{miss})$ (right). The QCD multijet background is estimated from data, see Section 7.5. For the plots all events with the full signal selection except the shown value is applied. The bin width in the ratio plot is increased with respect to the bin width in the upper plot in order to show the ratio only for bins with at least one expected events.

number of simulated events, in the right plot the efficiencies are calculated with respect to various stages.

- **trigger and offline threshold**: events, which have at least one jet with $p_T > 100\,\text{GeV}$ and $E_T^{miss} > 140\,\text{GeV}$
- **isolation**: events, which have one isolated tau with $p_T > 80\,\text{GeV}$ and $|\eta| < 2.3$. Events with additional taus are vetoed if they have a τ^{had} $p_T > 50\,\text{GeV}$
- **kinematics**: events, which passed the kinematic cuts.

The trigger efficiency shows a rising flank up to 2 TeV, due to the signal peak moving towards higher M_T. Above 2 TeV the signal starts to be produced off shell, shifting the spectrum towards lower M_T. The isolation shows a large drop in efficiency to about 20%, while a M_T threshold of 200 GeV does not effect the signal. The kinematic selection is about 80% efficient w.r.t. events passing the isolation criteria. The behaviour is similar to the electron and muon channel, where the geometric acceptance (p_T and η) rises with the W' mass up to 2 TeV (pink curve) and then off-shell production of W' shifts the events towards lower energies, thus reducing the geometric acceptance.

The main dependence of the efficiency is the tau p_T. Therefore, the efficiency is shown in Figure 7.8, as a function of generator level tau p_T.

Even when the most events where the taus are faked by a jet are described with a data-driven method, not all backgrounds can be described using this method. The kinematics of e.g. a $t\bar{t}$ is too different from a multijet background to be described by the same method. For all

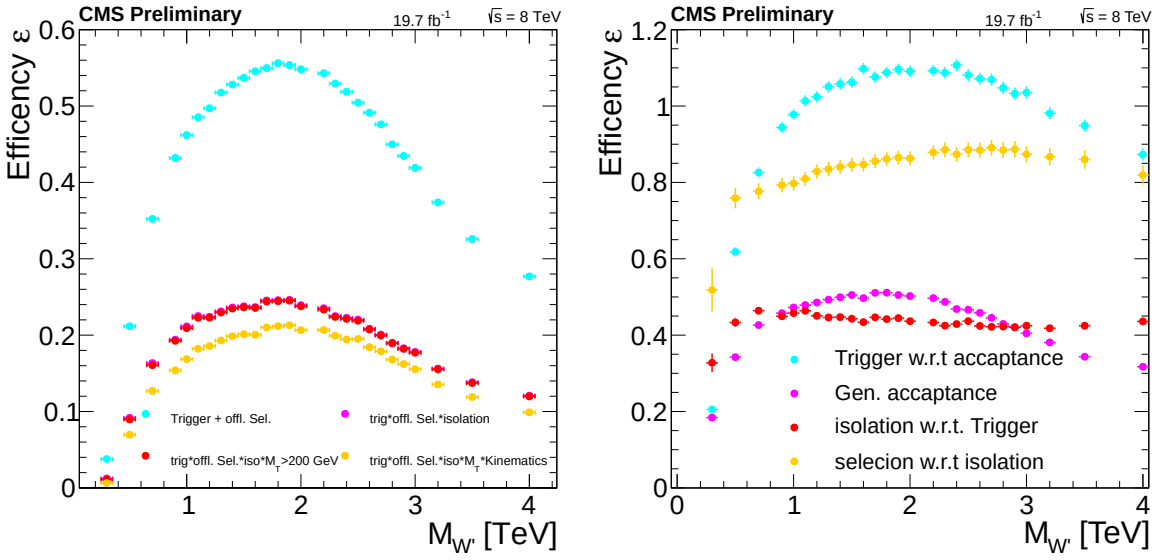


Figure 7.7.: The SSM signal efficiency for various stages in the analysis, binned in W' mass. The left plot shows the absolute efficiency including the hadronic branching faction. The right plot shows the efficiencies relative to different selections. The efficiency for the trigger, offline selection and tau isolation (red) is similar to the efficiency for the trigger, offline selection, tau isolation and $M_T > 200\,\text{GeV}$ (magenta) in the left plot.

electroweak backgrounds the description from simulation is used, except for the $Z \to \nu\nu$ + jets background, because the data driven method can describe it. Therefore Figure 7.9 shows the taus matched to their originating fermions obtained from the generator information. The figure shows the used cut-based (left) and the MVA based isolation (right). For the used cut-based isolation the wrongly reconstructed taus contribute up to 40% at low M_T, but their contribution falls steeply with increasing M_T down to 1 TeV where the fake contribution is at the 1% level. For the evaluation of the M_T distribution we therefore have to keep in mind that wrongly identified taus are not only described by the QCD jet contribution, but also by the simulation. For high M_T the expected contribution from jets faking taus is only given by the QCD jet background.

Additional kinematic distributions are shown in the appendix A.3.

A summary of all selection criteria can be found in Table 7.1 to provide a unified collection of all selections in a compact from.

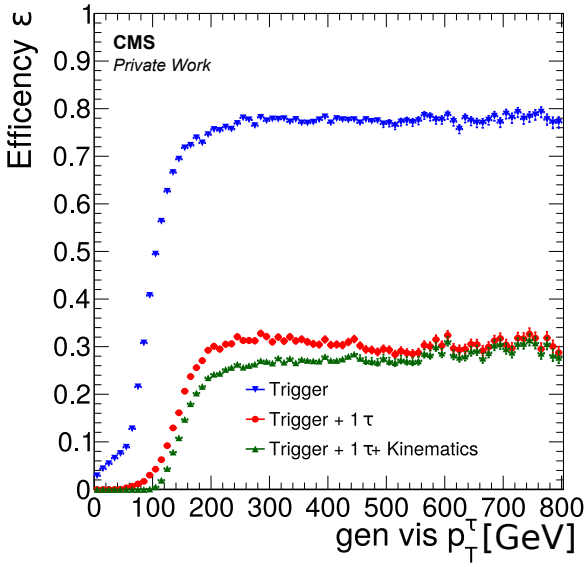


Figure 7.8.: The signal efficiencies depending on the visible generator level tau p_T.

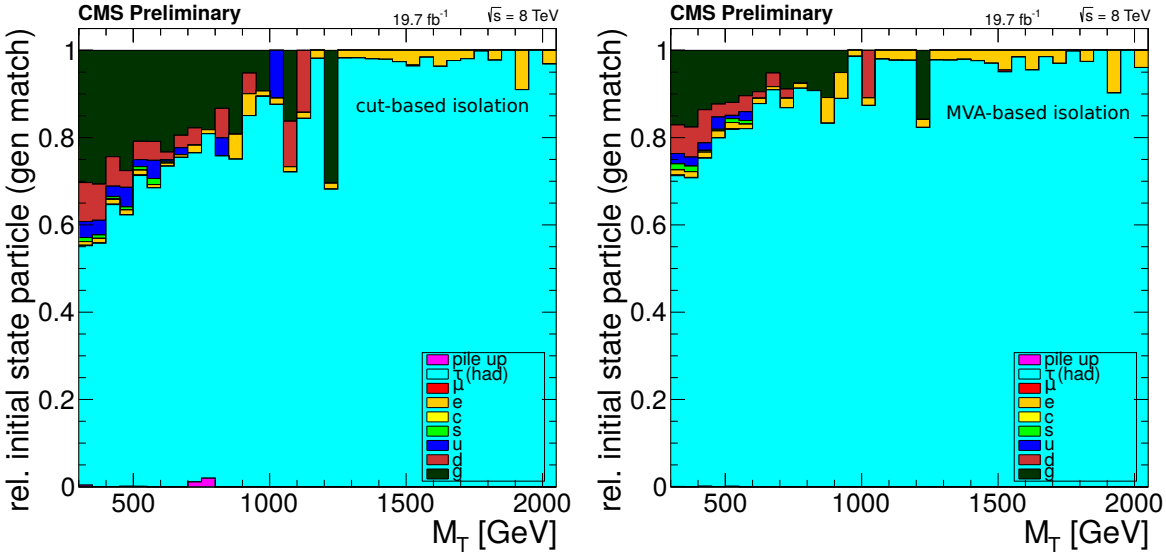


Figure 7.9.: For the background that is used from simulation the origin of the tau was evaluated, depending on M_T. Each M_T bin is normalized to represent the tau origin as a fraction. The left plot shows the loose cut-based, while the right plot shows the MVA isolation as control distribution.

Name	Value
General event properties	
acceptance for τ veto events with additional taus with $p_T > 50\,\text{GeV}$	$\lvert\eta\rvert < 2.3$ and $p_T > 80\,\text{GeV}$,
good vertex	$Ndof_{vtx} > 4$, $z_{vtx} < 24\,\text{cm}$, $(x^2_{vtx} + y^2_{vtx}) < 4\,\text{cm}^2$
event filter	E_T^{miss} and general event filters
Trigger	
Trigger	`HLT_MonoCentralPFJet80_PFMETnoMu105_NHEF0p95_v*` `HLT_MonoCentralPFJet80_PFMETnoMu95_NHEF0p95_v*`
offline thresholds	$E_T^{miss} > 140\,\text{GeV}$ and lead jet $p_T > 100\,\text{GeV}$
tau-ID	
decay mode	found one decay mode compatible with a τ^{had} (new incl. two pro
QCD discriminant	loose cut-based isolation or medium MVA-based isolation for control plots
electron discriminant	loose MVA based
muon discriminant	loose cut based
Kinematic	
just one good τ	true
$\frac{p_T^\tau}{E_T^{miss}}$	$0.7 < \frac{p_T^\tau}{E_T^{miss}} < 1.5$
$\Delta\phi(p_T^{had}, E_T^{miss})$	$\Delta\phi(p_T^{had}, E_T^{miss}) > 2.4$
Lepton Veto	No well identified muons or electrons

Table 7.1.: All selections for this analysis.

7.5. Determination of the Multijet Background from Data

Since processes where QCD jets are wrongly identified as tau jets are an important source of background for this analysis, a reliable method to estimate this contribution is needed. As was said before, the simulation prediction cannot be used, due to important, but unknown higher order corrections and insufficient statistics. Therefore, a data-driven approach is used instead. A so called "ABCD method" is used which yields a QCD jet sample for the final distributions where the shape and the normalisation is derived from data. An overview of the method is shown in Figure 7.10 which is explained in detail in the following two sections. The basic idea is that a tight-to-loose ratio can be measured from control regions C and D. The ratio uses tightly isolated taus and loosely isolated taus to provide an estimate of the fake probability. It is then used to scale the non isolated events to the isolated ones. All selection criteria listed in Table 7.1 are applied on the events used in this method except p_T^{had}/E_T^{miss}, $\Delta\phi(p_T^{had}, E_T^{miss})$ and tau isolation. This method is used to describe QCD multijet events as well as Z + jets events with $Z \to \nu\nu$ since the fake probability only depends on the recoiling jet. In the regions A, C and D contributions from processes not derived from data, but described by simulation, mainly W +jets, $t\bar{t}$, DY and diboson processes, are subtracted from the data to avoid double counting. Since the tau isolation is evaluated with the cut-based and MVA-based approach, this method is performed with both. The MVA approach is shown to provide a comparison and validate this isolation method for high p_T taus. The numbers are only quoted for the cut-based isolation, because the difference between the approaches is small.

In the following we will first discuss the template sample A and then the tight-to-loose ratio in regions C and D.

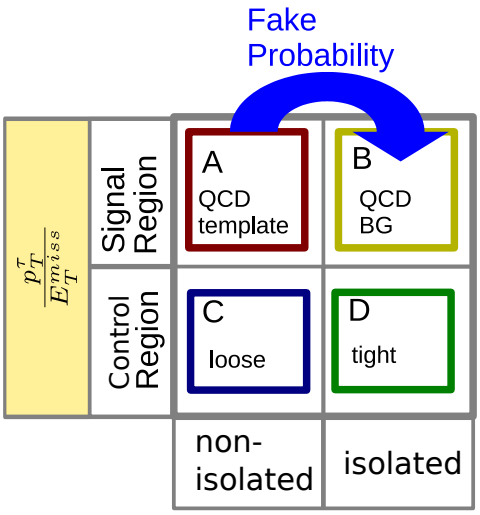


Figure 7.10.: Sketch of the ABCD method. Region A is used to obtain a QCD template sample, region C and D is used to determine the normalisation and region B is the signal region. The variable to distinguish the control and signal region is $\frac{p_T^{had}}{E_T^{miss}}$. The region with $\frac{p_T^{had}}{E_T^{miss}} \geqslant 1.5$ is the control region, while the region with $\frac{p_T^{had}}{E_T^{miss}} < 1.5$ is the signal region.

7.5.1. The QCD Jet Template Sample (Shape)

The first step of the ABCD method is to generate a template sample A from data which is dominated by QCD jet processes but includes as many selection steps as possible to minimise the effect of missing selections on the normalisation and shape of this sample. In this implementation of the method all selections are applied except for the tau isolation which is inverted. While in the regular event selection an isolated tau is required, this template sample should not contain such isolated tau leptons but instead at least one non-isolated tau candidate. Since the isolation selection is the main suppression against QCD jet processes, this inversion of the selection enhances the QCD jet contribution to this template sample. The remaining contamination of the other background processes (mainly W +jets, $t\bar{t}$) to this sample is removed by subtracting their contribution as determined from simulation. The distribution of non-isolated events is shown in Figure 7.11 (left) for M_T and p_T^{had}/E_T^{miss} in the region A (for higher p_T^{had}/E_T^{miss} also for region C). The simulated background is only shown for the weak processes. The difference between these backgrounds and the data is therefore an estimate of the QCD jet contribution. The resulting distribution is a sample dominated by QCD jet processes, gained from data. Figure 7.11 (right) shows the M_T distribution from region A. The subtracted contamination amounts in region A to approximately 30%. The largest contamination is from boosted $W \to l\nu$ events. The second largest contribution to the contamination are the processes with top production. Due to the different kinematics one cannot use this method to evaluate the shape of these contributions.

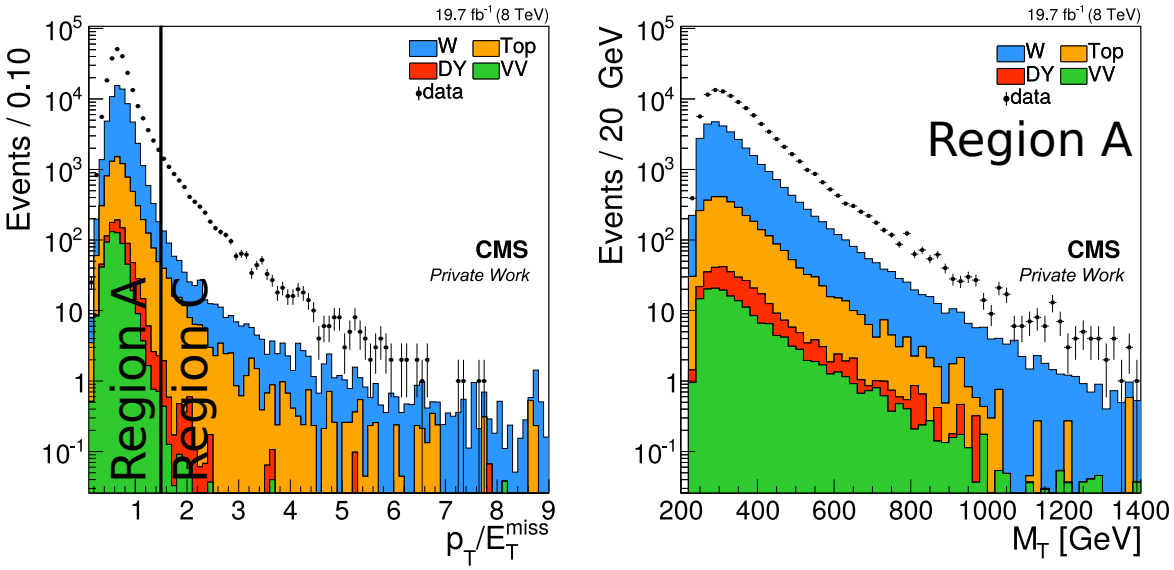


Figure 7.11.: The p_T^{had}/E_T^{miss} distribution in the signal region with inverted isolation for the taus (region A and C). On the right the resulting M_T distribution in A is shown. The QCD jet simulation (multijet and $Z \to \nu\nu$) is not plotted to only show the subtracted background (coloured) and the data.

7.5.2. Tight-To-Loose Ratio (Normalization)

In order to scale the non isolated region A to the isolated signal region B, a tight-to-loose ratio approach is used. The tight-to-loose ratio can be interpreted as a probability for a QCD jet to pass the isolation.

This probability is derived from a disjoint, signal free regions C and D, which are separated from the signal events by using the events with $\frac{p_T^{had}}{E_T^{miss}} > 1.5$. The set of all taus passing these selections and failing the isolation is called the "loose" tau sample while the sample which is yield from all taus passing the isolation criterion is called the "tight" sample.

The tight-to-loose ratio is derived out of this sample which is defined as

$$R_{TL} = \frac{\text{isolated taus}}{\text{non isolated taus}} = \frac{D}{C}$$

The contribution from electroweak processes (W, Top, DY and VV) are subtracted, using the description from simulation. The number of events are (cut-based isolation):

$$\text{C: } \overbrace{15642}^{\text{data}} - \overbrace{1803.0}^{\text{simulation}} = 13839.0 \tag{7.5}$$

$$\text{D: } \overbrace{120}^{\text{data}} - \overbrace{29.0}^{\text{simulation}} = 91.0 \tag{7.6}$$

$$\Rightarrow R_{TL} = 0.0066(1 \pm 12\% \text{ (stat.)}) \tag{7.7}$$

The ratio can be further evaluated in different tau components such as $1h1\gamma$ or $3h0\gamma$. This is done to correct for a potential bias in the chosen phase space. Especially the fake probability could be different for gluon and for quark jets. This difference can be compensated with the use of the decay mode as a final state. Other parametrizations (e.g. η, p_T^{had} and E_T^{miss} dependent) have been evaluated and did not show a behaviour that was non trivial. This does not mean the fake probability does not depend on p_T^{had}, but a flat parametrisation was reasonable for the signal region, since only high p_T taus are used. This will be discussed in the next Section 7.5.3.

The tau components show a strong influence on the tight-to-loose ratio as shown in Figure 7.12. The largest ratio is in the one hadron and one photon bin ($1h^{\pm}1\gamma$). Many QCD jets have this final state. The tau reconstruction has an energy ordering for the components. Therefore a $1h^{\pm}1\gamma$ jet will have a high energetic hadron and photon, which is very tau like. For additional tau components the information to compute the tau properties increases thus the tight-to-loose rate drops. The fake probability for the tau with the components of $3h^{\pm}1\gamma$ is much higher than the other τ^{had} final states. This on the one had due to the difference between the tau components and the decay mode. The decay mode in to $3h^{\pm}1\pi^0$ is not used in the reconstruction, hence taus with the constituents $3h^{\pm}1\gamma$ are the rare cases where the 1γ is added to one of the charged hadrons.

The grey bands show the estimation of the systematic effect from the p_T^{had}/E_T^{miss} threshold. The ratio is evaluated several times with changed p_T^{had}/E_T^{miss} thresholds and upper bounds. The lower thresholds were changed from 1.5 to 1.55, 1.6 and 1.7, where the upper bound is

loosened from 4 to $p_T^{had}/E_T^{miss} < 8, 10$ and 20. Since a dependence on this threshold should not effect the result it reflects the systematic influence on the ratio. Shown are the minimal and maximal values from these variations. The point in each bin represents the mean.

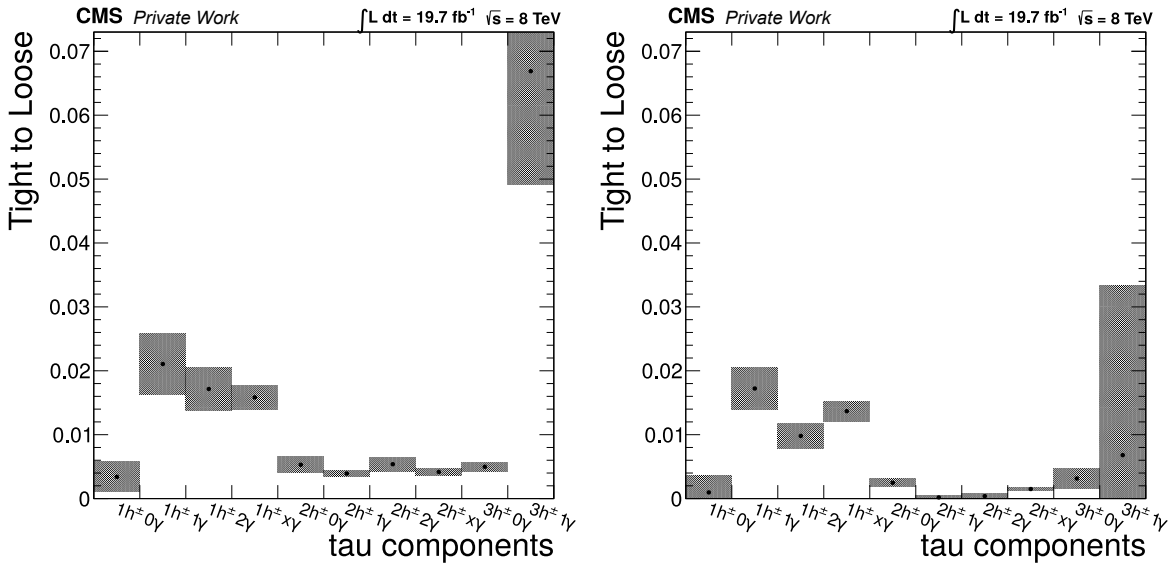


Figure 7.12.: The ratio of tight to loose taus w.r.t. the tau pfCandiates is shown for several selections. In the left plot the cut-based isolation is shown. In the right plot the MVA-based is shown for comparison. The shaded areas correspond to the shift due to the variation of the p_T^{had}/E_T^{miss} threshold.

The obtained ratio represents the probability for one jet to fake one tau. To take into account multiple jets the ratio has to be applied for each jet according to the following formula:

$$R_{TL}^{event} = \sum_k \left(P(jet_k \text{ faking } \tau) * \prod_{i!=k} (1 - P(jet_i \text{ faking } \tau)) \right) \tag{7.8}$$

where $P(jet_k \text{ faking } \tau)$ depends on the decay mode of the jet.

The dependence of R_{TL}^{event} on other variables is also checked and is listed in the following section. The expected p_T^{had} dependence is small, because the trigger threshold is relatively high and the fake probability has reached its minimum.

7.5.3. Crosschecks of Fake Probability

This QCD background estimation is crosschecked with several changes to the setup of the method. The first method is an evaluation of the tight-to-loose ratio using the QCD simulation, another uses a dataset collected with the E_T^{miss} trigger. Other parametrisation of the fake probability are also checked.

The result from simulation is shown in Figure 7.13. The black line represents the used data driven estimate described in the previous two sections. The coloured areas represent the prediction from isolated simulation, split in $Z \to \nu\nu$ and QCD multijet. One can clearly see that the QCD multijet background does not have enough events to give an accurate description.

The grey line corresponds to the non isolated simulated events ($Z \to \nu\nu$ and QCD multijet) scaled with a tight-to-loose ratio evaluated like on data, but with the simulation as input. The tight-to-loose ratio on simulation has a large statistical uncertainty of $\approx 70\%$ due to an insufficient number of isolated events.

The shape of the distribution can be confirmed. Within the simulation uncertainty the correct order of magnitude can also be confirmed.

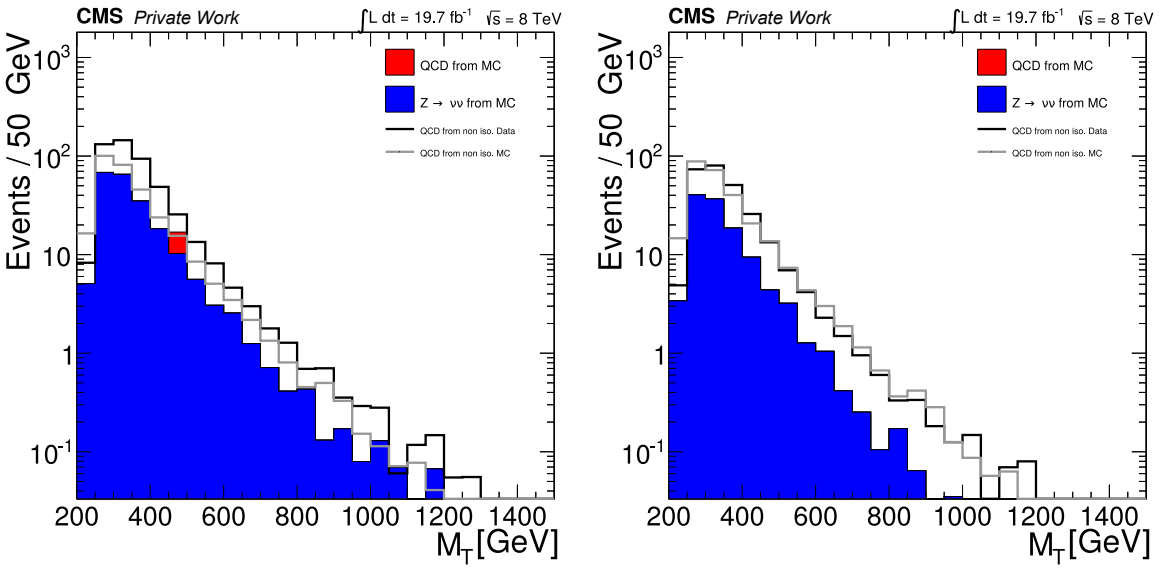


Figure 7.13.: The resulting QCD contributing in the signal region for two isolations, loose cut-based left and medium MVA right. As can be seen the direct contribution from QCD simulation is only from a few events. The distributions from scaled QCD agree reasonably with the data driven estimation.

The fake probability was also evaluated using the E_T^{miss} triggered dataset. The tight-to-loose ratio is $R_{TL}^{E_T^{miss}} = 0.0052(1 \pm 34\%\ (stat.))$ for loose cut-based isolation and $R_{TL}^{E_T^{miss}} = 0.001(1 \pm 103\%\ (stat.))$ for medium MVA isolation, which can be compared to $R_{TL}^{singlebin} = 0.0066(1 \pm 12\%\ (stat.))$. The large error of the cut-based isolation $R_{TL}^{E_T^{miss}}$ is due to $8(data) - 5.16(simulation)$ events in the isolated region D. The calculated ratios can confirm the order of magnitude even if they have a large uncertainty.

In Figure 7.14 a comparison between the two triggers is shown for the M_T distribution weighted with R_{TL} . For the E_T^{miss} trigger the statistical evaluation in several decay mode bins was not possible, due to low event rates in the isolated events with $p_T^{had}/E_T^{miss} > 1.5$.

Another cross check is to use a single bin to evaluate the tight-to-loose ratio and not bins in tau decay modes as shown in Figure 7.15 (left). The result (red line) agrees almost with the QCD jet prediction evaluated in bins of tau components. The same is true for the tight-to-loose ratio evaluated in terms of tau p_T^{had}. Both tau p_T^{had} binned and single binned tight-to-loose ratio have similar errors as estimated for the reference QCD jet prediction.

Other isolation working points were also calculated using this method. For all working points a reasonable QCD jet prediction could be calculated.

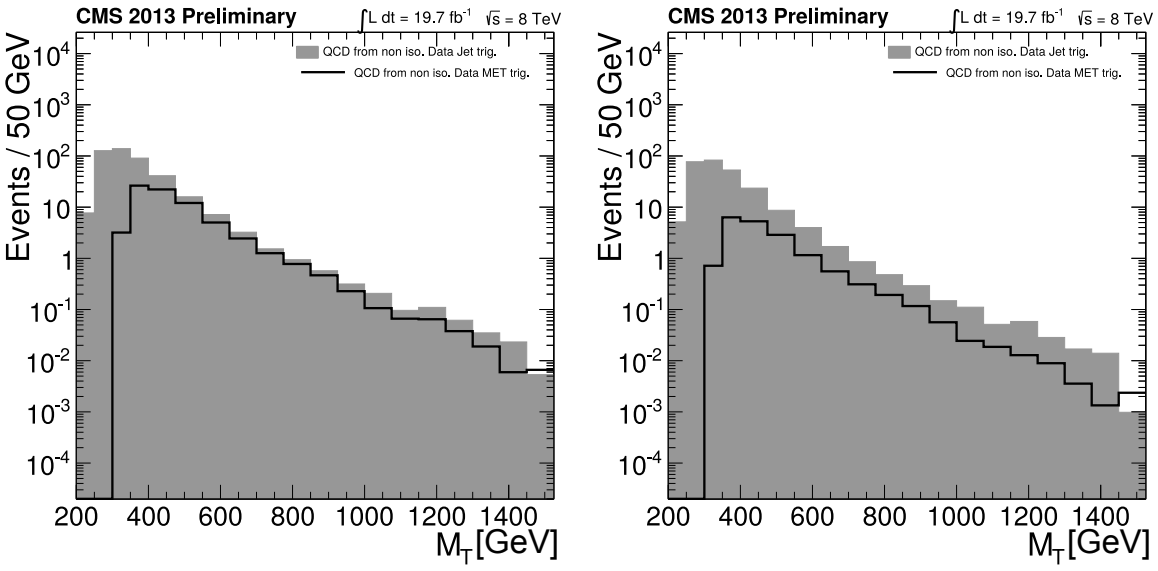


Figure 7.14.: Tight-to-loose result with E_T^{miss} triggered events. For loose cut-based isolation (left) and medium MVA (right)

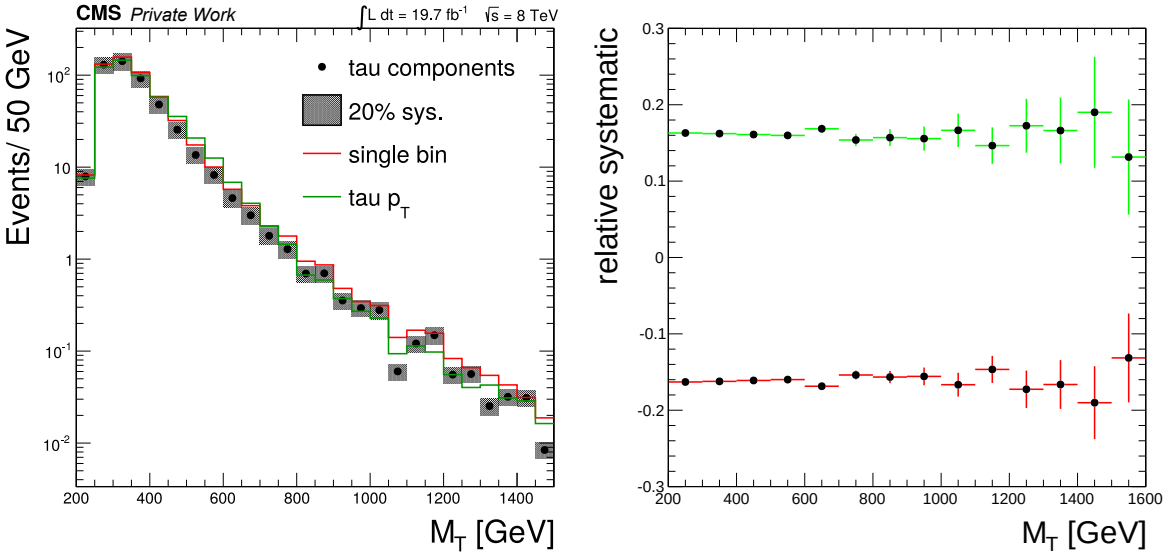


Figure 7.15.: If the fake probability is evaluated in a single bin (red curve) and as a function of tau p_T^{had} (green curve), the prediction agrees with the prediction from the decay mode binned fake probability almost within 20% uncertainty (left). Relative systematic shift in the M_T distribution from the shifts in Figure 7.12 for the loose isolation (right).

The systematic variation shown in Figure 7.12 could in principle have an effect on the M_T distribution, if the jet composition would change with M_T, but this is not expected. Shown in Figure 7.15 is the systematic shift for the upper and lower point of the grey area in Figure 7.12. The systematic shift is about 17% and flat with M_T. In addition to this the statistical error for the tight-to-loose ratio is about 10%.

Combining these two results and summarising the various tests with the tight-to-loose ratio a systematic uncertainty of 20% on the decay mode dependent R_{TL} is used. In order to incorporate a potentially wrong fake probability in the simulation, the subtracted simulation is varied by 20% and the bin by bin result is used as a systematic uncertainty.

7.6. Full Background Prediction and Final Distributions

The number of QCD jet events after the full selection derived by the previously presented ABCD method is $N_{QCD} = 620$ for loose cut-based isolation.

Adding this contribution of QCD jet processes to the contribution of background events derived from simulation results in the final distributions of this search. The spectrum as well as the cumulative distribution of the transverse mass which is used for the statistical analysis in Section 10.2 can be seen in Figure 7.16. The Figure shows the loose cut-based isolation (left) as well as the medium MVA-based isolation (right) for comparison. Both isolation discriminators provide similar results. For the MVA-based isolation the tighter working point (medium instead of loose) has a similar signal efficiency, but a lower background prediction. The total number of all background events which includes the contribution from simulation as well as the contribution from the QCD data-driven method above a certain M_T threshold can be seen in Table 7.3 and 7.2 together with their uncertainty for cut-based and MVA isolation, respectively. The QCD jets include jets from invisible Z decays. No excess of data above the Standard Model prediction can be seen which exceeds the overall uncertainty. For both isolations the largest difference is at 600-800 GeV where less events are observed than predicted by the background. For higher M_T the number of expected and observed events is compatible again. In the tau+E_T^{miss} channel no signal is expected that could explain a local underfluctuation of events. As can be seen in the statistical analysis the fluctuation is less than 2σ. For a plot, where above the shown M_T all events are integrated, see Figure 7.17. The reasonable agreement between data and background for both isolations can be interpreted as a confirmation that the MVA-based isolation has no unexpected dependencies.

The calculation of the systematic uncertainties on the number of predicted background events will be described in Section 9.1. The distributions for p_T^{had} and E_T^{miss} are shown in Figure A.6. The charge separated distributions for M_T and tau p_T^{had} are shown in Figure A.7.

The statistical interpretation derived from these distributions can be found in Chapter 10 in combination with the Run II limits. The Run II analysis will be discussed first in the following Chapter 8. This will highlight the difference to the 8 TeV analysis and based on the knowledge acquired by this first implementation of the analysis.

M_T^{min} / GeV	Diboson	DY	Top	QCD jet	W	Background	Data	$M_{W'}$	
								2.2 TeV	1 TeV
200	10	10	54	620	1400	2080 ± 34	2000	55	2000
400	2.3	2.9	7.6	150	230	397.9 ± 5.5	360	50	1500
800	0.064	0.072	0	3.6	7.4	11.15 ± 0.49	10	32	230
1000	0.0091	0.027	0	1.1	1.9	3.06 ± 0.19	4	25	25
1200	0.0031	0.016	0	0.32	0.61	0.944 ± 0.095	1	19	1.1
1400	0.0011	0.0076	0	0.13	0.18	0.323 ± 0.046	0	12	0.45

Table 7.2.: Summary of simulated background samples with data driven QCD for loose cut-based isolation for the Run I analysis. The events are integrated above the M_T^{min} threshold in the row. The errors include the statistical and systematic uncertainties as defined in Section 9.1.

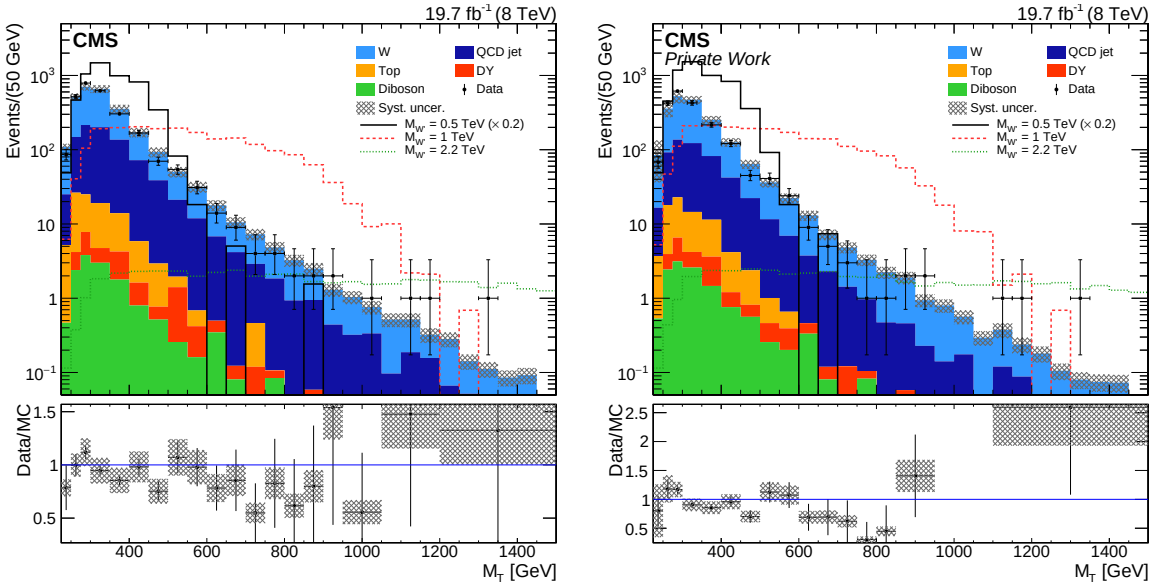


Figure 7.16.: Final M_T distribution with full background prediction. The left plot shows the M_T spectrum for the loose cut-based isolation, while the right is for the medium MVA isolation. The QCD background contribution is taken from the ABCD method described previously. Good agreement between data and simulation can be seen over the full spectrum. The background is described well and no evidence for a W' boson occurs in the data.

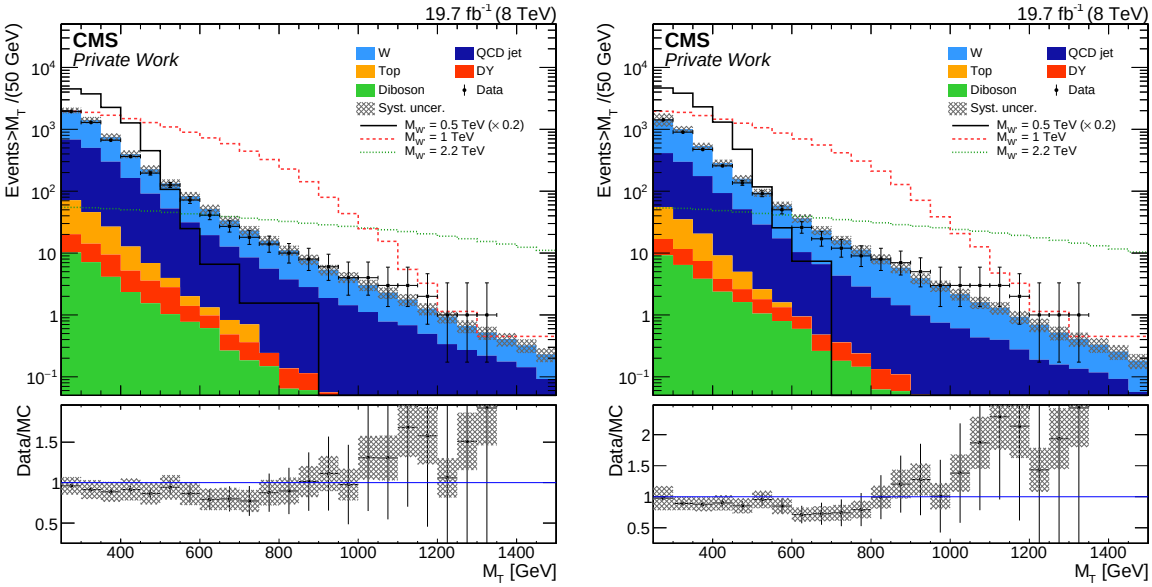


Figure 7.17.: Final M_T distribution with full background prediction. The events are integrated above the M_T threshold on the x-axis. The left plot shows the M_T spectrum for the loose cut-based isolation, while the right is for the medium MVA isolation. Good agreement within the uncertainties between data and simulation can be seen over the full spectrum. The background is described well and no evidence for a W' boson occurs in the data.

M_T^{min} / GeV	Diboson	DY	Top	QCD jet	W	Background	Data	$M_{W'}$	
								2.2 TeV	1 TeV
200	9.5	7.7	40	360	1100	1496 ± 32	1500	54	2000
400	2.4	1.4	5.3	84	190	285.5 ± 4.4	260	49	1500
600	0.60	0.34	0	9.4	26	36.4 ± 1.0	26	39	700
800	0.063	0.071	0	1.8	6.1	8.05 ± 0.39	8	31	210
1000	0.0076	0.023	0	0.58	1.6	2.17 ± 0.12	3	24	21
1200	0.0025	0.014	0	0.17	0.51	0.697 ± 0.062	1	18	1.1
1400	0.0011	0.0084	0	0.083	0.16	0.257 ± 0.036	0	12	0.45

Table 7.3.: Summary of simulated background samples with data driven QCD for medium MVA isolation for the Run I analysis. The events are integrated above the M_T^{min} threshold in the row. The errors include the statistical and systematic uncertainties as defined in Section 9.1.

8. Tau Analysis at 13 TeV

Before the statistical interpretation of the Run I results will be given, the analysis of the Run II data is described. This analysis is by nature very similar to the Run I analysis. Due to the nature of the data taking period the Run II analysis could be build on the experiences gained from the fist τ+E_T^{miss} analysis. Therefore some aspects of the analysis could be extended. The following description will be on the differences to the Run I analysis or describe another aspect of a specific problem to provide a deeper understanding of the analysis. One part, which is changed drastically is the data driven estimate, hence a full description of this aspect will be provided in Section 8.5.

8.1. Data Set of pp Collisions Recorded in 2015

This analysis is done with the Run II data, with 2.3 fb^{-1} of $\sqrt{s} = 13$ TeV proton-proton collision data collected with the CMS detector in the year 2015 with 25 ns bunch spacing. As trigger the dedicated single tau plus E_T^{miss} trigger[1] is used. The last certified dataset is used from the 16th December 2015[2].

8.2. Monte Carlo Simulation

The processes used as background for this search are the same as for the Run I analysis. The used samples are all in the latest production campaign for the 2015 simulation[3]. One of the main differences is the use of PYTHIA 8 instead of PYTHIA 6 for the hadronisation and simulation of the underlying event. Also during the first long shutdown of the LHC the NLO calculation which was previously only available in POWHEG was also implemented in MADGRAPH in a generic way. In addition to the use of MINIAOD the simulation of pileup has also changed[4]. PYTHIA 8 is used to simulate the minimum bias events for the pileup mixing. The distribution of additional interactions was adjusted to fit the recorded data, hence only a small modification due to the used subset of the data was needed, compared to Run I. The average number of pileup interactions per event in the 25 ns bunch spacing scenario is $N_{PU} \approx 10$. The influence of previous and subsequent collisions is larger, due to the higher frequency. This effect could be averted by adjusting the object identifications as described previously in Section 5.

[1] `HLT_LooseIsoPFTau50_Trk30_eta2p1_MET80_JetIdCleaned_v*`
[2] Cert_13TeV_16Dec2015ReReco_Collisions15_25ns_JSON
[3] The production is tagged as `RunIIFall15MiniAODv2-PU25nsData2015v1_76X_mcRun2_asymptotic_v12`
[4] The version of the pileup scenario is `2015_25ns_Startup_PoissonOOTPU`

8.2.1. Background Samples

For the division into physics backgrounds the same processes were used as for the Run I analysis. For the simulation the developments of the generators were included. The PYTHIA high mass tails were produced in PYTHIA 8, which made it possible to divide these samples not in $\hat{p}_T$ but in the invariant mass of the W-boson. This has the advantage that the transition between different generators is well defined. Also the mass separations were only produced with a lower mass threshold, and no upper threshold. While potentially some generated events are not used it is possible to check the behaviour of the individual samples. For the non W background the number of processes generated at NLO also increased, such that the high mass DY, all top and diboson processes are at NLO.

The full overview of sample names, used cross sections and event numbers can be seen in Table A.4 in the appendix. The used generator can be read from the name.

For the NNLO QCD and NLO electroweak differential cross section correction of the $pp \to W \to \tau\nu$ process the same approach was used as for the 8 TeV analysis. In Figure 8.1 for both combinations (additive and multiplicative) are shown at the 13 TeV centre of mass energy. For the 13 TeV k-factor the effect of different particles on the NLO EW cross section can be seen in Figure 8.2. One can clearly see that the electroweak corrections for muons are the strongest, while for electrons and taus the effect is at the 1-2% level. This is due to the reconstruction of the objects. While electrons and taus are reconstructed in the calorimeters, where additional photons radiated as final state radiation would be combined with the hard particle, muons are reconstructed as tracks and radiated photons change the invariant mass of the W-boson under investigation here.

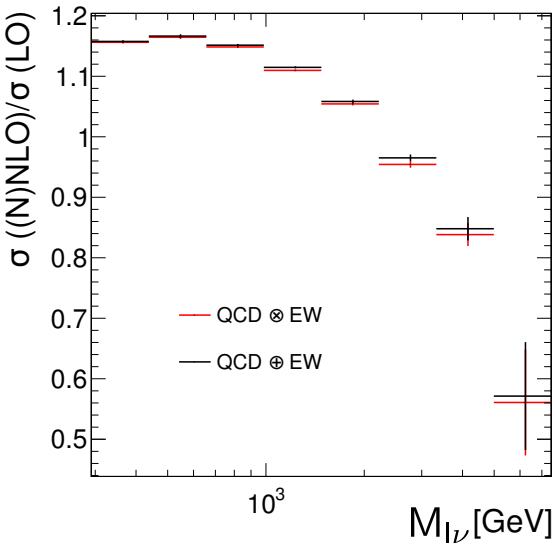


Figure 8.1.: k-factors for the additive and factorial combination of NNLO QCD and NLO EW corrections for the $pp \to W \to \tau\nu$ process.

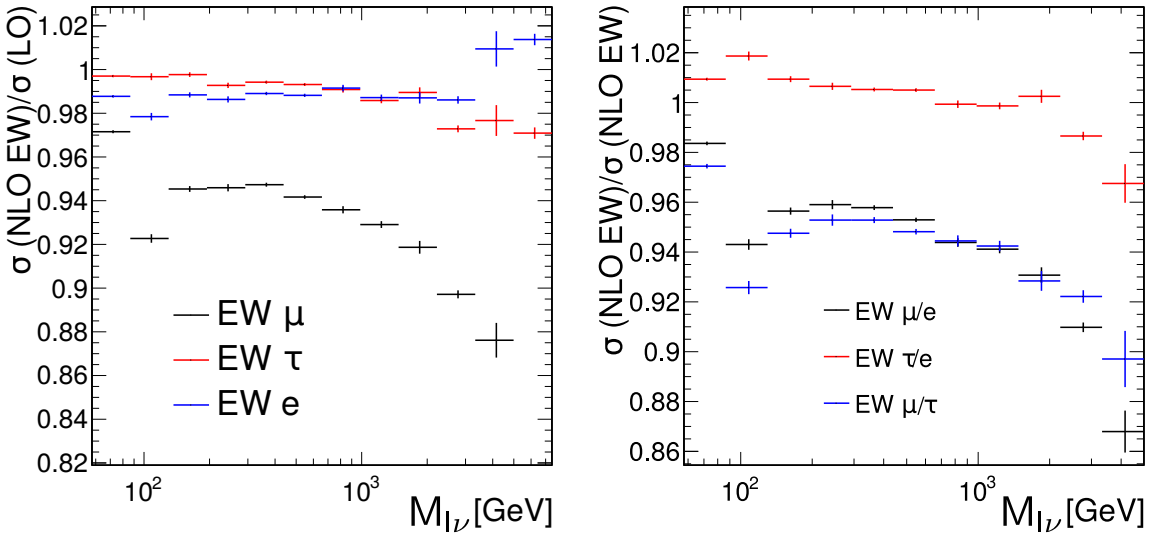


Figure 8.2.: The difference of the electroweak corrections for the $pp \to W \to \tau\nu$ process with respect to different final state leptons. On the left the ratio of NLO to LO cross section can be seen and on the right the ratios of different final state leptons as indicated in the legend.

8.2.2. Signal Samples

In order to determine the selection efficiency of the analysis for the various W' models considered in this analysis, many signal samples have been produced. The SSM samples are generated with PYTHIA in a mass range between 0.5 TeV and 5.8 TeV. The cross sections for the different samples can be seen in Table A.2 in the appendix. All samples are produced in LO with the PDF set CTEQ6L1 [143], and for the SSM samples cross sections in NNLO QCD have been calculated [144].

8.2.3. Reweighting of Signal Samples

Since the production of signal samples needs time and resources, not all samples were generated and processed with the full detector simulation. In order to reduce the computing time the samples with a different total width of the W' were studied on generator level. The samples for different width were simulated with MADGRAPH for various coupling strength. The connection between the coupling strength and the width of a W' boson is shown in Equation 2.17. The effect on the invariant mass distribution is a sharpening ($g_{W'} < 1$) or broadening ($g_{W'} > 1$) of the peak of the distribution, see Figure 8.3. The peak position at the invariant mass as shown in Figure 8.3 (left) is independent of the coupling. The narrow coupling $g_{W'}/g = 0.01$ appears to have a smaller cross section, because the bin width of 10 GeV is broader than the peak.

In order to evaluate the shape changes of these different widths, the transverse mass of the undecayed tau and neutrino is used instead of the transverse mass calculated from the tau decay, denoted $M_T^{\tau\nu}$. The tau decay is not included, since MADGRAPH does not take the polarization of the tau into account and further processing with PYTHIA would produce decayed taus, but also cause dependencies on effects like underlying event and initial state

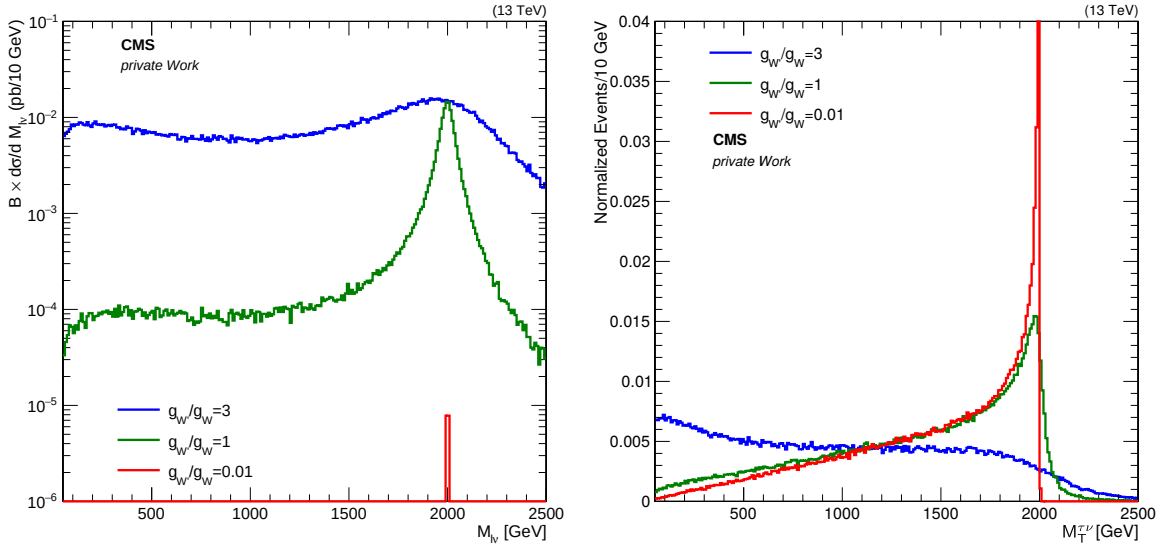


Figure 8.3.: The invariant mass distribution for tau and neutrino on generator level for various coupling constants, resulting in different widths is shown on the left. The transverse mass distribution for different width is shown on the right.

radiation description. This $M_T^{\tau\nu}$ can then be used to weight the reconstructed events in the final distribution to obtain a reconstructed M_T distribution for a different W' width. The weight and the resulting M_T distribution can be seen in Figure 8.4. Since each event in the final distribution is weighted, the detector response is conserved. The used weights are shown in Figure 8.4(left). The effect on the reconstructed M_T is shown in Figure 8.4(right). One can see that a broadening of the W' width decreases the signal strength, because the M_T distribution is shifted to lower values, which are not triggered. The sharpening of the signal does not affect the final distribution as much as it would affect an event from a $W' \to e/\mu\nu$ decay, since the resonant peak is smeared over all lower M_T values due to the additional neutrino from the τ decay.

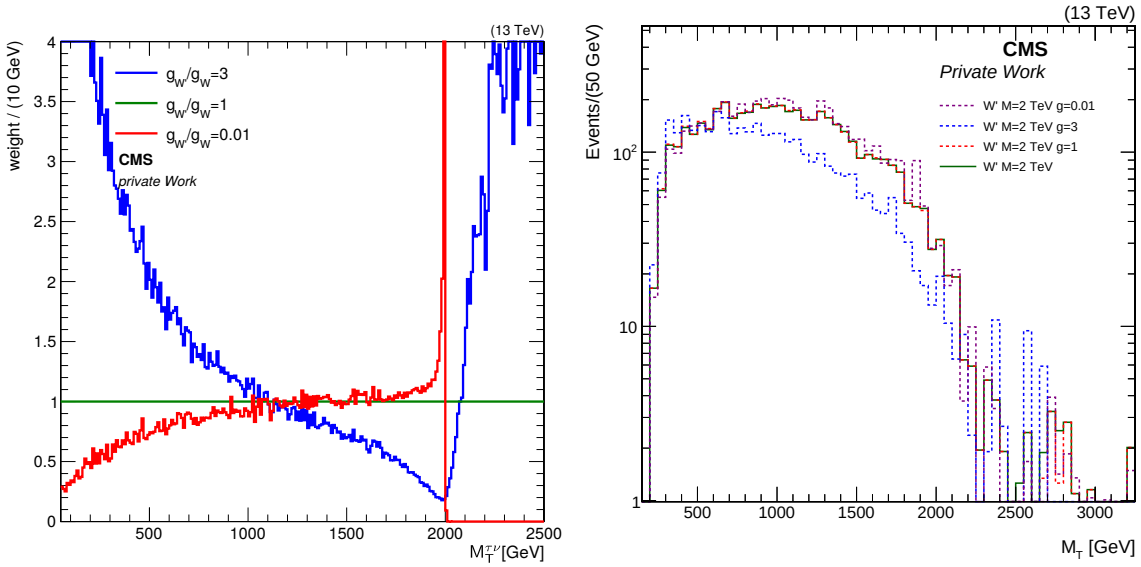


Figure 8.4.: The weight of different coupling constants with respect to M_T is shown on the left. The reweighted reconstructed signal M_T distribution is shown on the right. The right plot is not normalized to the number of generated events.

8.3. Tau + E_T^{miss} Trigger

In contrast to the Run I analysis the tau + E_T^{miss} trigger[5] could be used in the full dataset. It was unprescaled during the complete data taking in 2015.

In contrast to the Run I analysis an evaluation of the trigger efficiency of data and simulation was performed. The distributions are shown in Figure 8.5. To obtain this data and simulation efficiency the trigger efficiency was evaluated individually for the two trigger legs, tau and E_T^{miss}. As preselection for the trigger evaluation is one isolated tau with a leading charged hadron p_T^{had} lager than 50 GeV has to be in each event.

For the tau leg[6] an additional selection is that the offline E_T^{miss} is larger than 170 GeV, the tau has $|\eta| < 2.1$. The data also has to be triggered by another trigger, because it would not have been recorded. For this the pfE_T^{miss} trigger[7] was used. The resulting efficiency is shown in Figure 8.5(left). The obvious, but important thing to note here is that in contrast to the Run I data the tau trigger is fully efficient for high p_T τ^{had} events. The trigger efficiency does not reach 100%, because the E_T^{miss} leg of the trigger is not fully efficient at 120 GeV, but if this would be raised to 200 GeV, not enough tau events would have been triggered to provide a meaningful efficiency.

For the E_T^{miss} leg the tau is required to have a $p_T > 80$ GeV. The tau trigger[8] was used and the results are in Figure 8.5(right). This single tau trigger is prescaled during data taking. This means that not every time the trigger fired the event is recorded, but every n-th time with $n > 1$. The prescale factor n changes during the data recording and is adjusted to the

[5] `HLT_LooseIsoPFTau50_Trk30_eta2p1_MET80_JetIdCleaned_v*`
[6] `LooseIsoPFTau50_Trk30_eta2p1`
[7] `HLT_PFMET_*` the threshold changed over run
[8] `LooseIsoPFTau50_Trk30_eta2p1_v*`

instantaneous luminosity. The tun on of the distribution is very broad, but no inefficiency at high E_T^{miss} could be observed. The broadness of the E_T^{miss} is not desirable, but can not be changed. If the offline threshold would be raised to only use the fully efficient part of this trigger e.g. $E_T^{miss} > 200\,GeV$ only roughly 10 events would have entered the final distribution. As the description of the data by the simulation is reasonable well, the modelling of the simulation is used. The offline E_T^{miss} threshold is set to 120 GeV. As offline threshold for the tau 80 GeV is used.

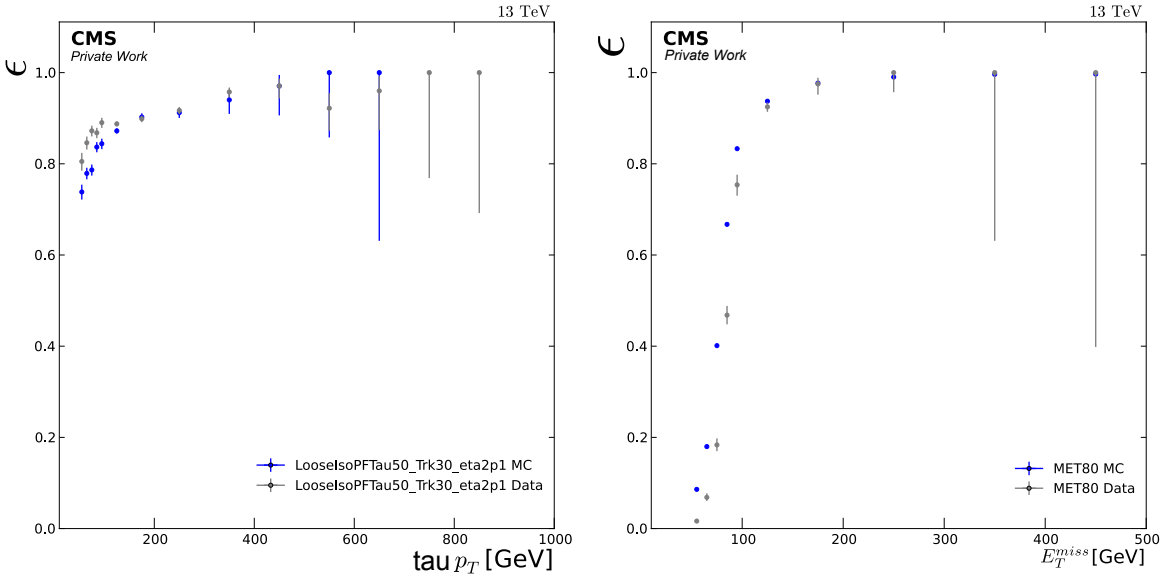


Figure 8.5.: The plots show the trigger efficiency for the tau (left), triggered by a E_T^{miss} trigger with $pfE_T^{miss} > 170\,GeV$. The right plot shows the efficiency of the E_T^{miss} leg of the tau + E_T^{miss}, evaluated with the prescaled single tau trigger.

8.4. Event Selection

The Run II selection follows the same lines as the Run I analysis. The main difference with respect to 8 TeV is the use of the MVA discriminator. The p_T^{had}/E_T^{miss} selection was symmetrized around one.

The threshold for the τ^{had} p_T and E_T^{miss} can be kept or even lowered. The leading tau has to have a $p_T > 80$ GeV, where the E_T^{miss} has to be larger than 120 GeV. Due to the trigger requirement, the tau has to be in a range of $|\eta| < 2.1$. For the tau identification the HPS tau is required to pass the decay mode reconstruction including the two prong decays, the retrained loose electron discriminator[9], the loose muon discriminator[10] and the medium MVA-based discriminator [11].

The used kinematic selection is with the symmetric p_T^{had}/E_T^{miss} boundaries:

$$0.7 < p_T^{had}/E_T^{miss} < 1.3 \tag{8.1}$$

$$\Delta\phi(\tau^{had}, E_T^{miss}) > 2.4 \tag{8.2}$$

As veto against light leptons the same identification working points are used as in the Run I analysis. These identification criteria are updated to compensate for the Run II conditions. The p_T^{had} and η constrains are kept[12]. For details see Section 5.1 and 5.2. From the event filters the same reconstruction effects are caught to suppress artificially created E_T^{miss}. The individual codes where updated and details can be found at [146].

The distributions for p_T^{had}/E_T^{miss} and $\Delta\phi(\tau^{had}, E_T^{miss})$ are shown in Figure 8.6. The distributions show that the background estimation describes the data distributions well. The systematic uncertainties are evaluated bin by bin and are described in Section 9.1.

In contrast to the Run I distribution the background of the p_T^{had}/E_T^{miss} is shifted to lower values. This is due to the different trigger, while it is more efficient for the tau leg, the E_T^{miss} leg is not as efficient. As can be seen by comparing Figure 8.6 to the Run I Figure 7.6, the signal is not effected by this. For the $\Delta\phi(\tau^{had}, E_T^{miss})$ distribution the lower trigger thresholds can also be seen. For low $\Delta\phi(\tau^{had}, E_T^{miss})$ a rise in events is observed, which are dominated by boosted W events, where the τ^{had} and E_T^{miss} is expected to point in the same direction. Another change with respect to the Run I analysis is that the data driven fake probability is no longer estimated in the hight p_T^{had}/E_T^{miss} region, as will be described in Section 8.5.

The signal efficiencies for SSM W' are shown in Figure 8.7. The efficiencies are shown for different W' masses and various stages in the analysis. In the both plots all efficiencies are calculated with respect to the number of simulated events. While in the right plot the efficiency is plotted as a funciton of the W' mass, int the right plot the efficiencies are calculated as a function of to M_T. The used stages are:

- **One reconstructed tau**: events which fired the trigger and the event has at least one tau with $p_T > 80$ GeV and $E_T^{miss} > 120$ GeV.

[9] againstElectronLooseMVA6

[10] againstMuonLoose3

[11] byMediumIsolationMVArun2v1DBnewDMwLT

[12] For both leptons the minimal allowed p_T is 25 GeV and the η region is according to the detector coverage $|\eta| < 2.4$ for muons and $|\eta| < 2.5$ for electrons.

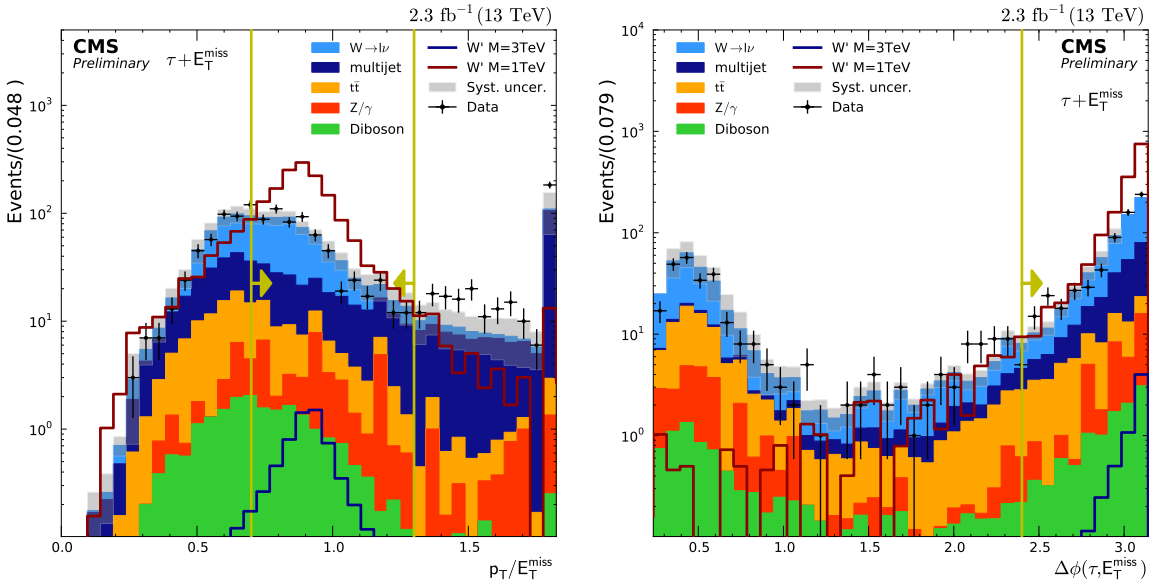


Figure 8.6.: In the plots the simulated backgrounds are shown for the kinematic distributions for p_T^{had}/E_T^{miss} (left) and $\Delta\phi(\tau^{had}, E_T^{miss})$ (right). The QCD multijet background is estimated from data, see Section 8.5. For the plots all events with the full signal selection except the shown value is applied.

- **preselection**: events which have exactly one isolated tau with $|\eta| < 2.3$ (veto for additional taus with $p_T > 50\,\text{GeV}$)
- **full selection**: events which passed the kinematic cuts as outlined in Equation 8.1 and 8.2.

For Figure 8.7 (left) the efficiency is also provided for events that have a $M_T^{min} > 400$ at generator level. As expected from Figure 8.7 (right) the efficiency is almost independent from the W' mass for these events.

As for the 8 TeV analysis not all fake taus are described with the data driven approach. Therefore Figure 8.8 shows the origin of the taus, matched at generator level. Similar to the cut-based isolation, used for the Run I analysis, the MVA-based isolation has a contribution of up to 40% from jets faking taus at low M_T, but this contribution is reduced to 10% for high M_T events.

The evaluation of the fake contribution from an independent sample allows the systematic evaluation of the distributions of many variables without a self predicting QCD multijet background. For the kinematic selections, the distributions for p_T^{had}, E_T^{miss}, p_T^{had}/E_T^{miss} and $\Delta\phi(\tau^{had}, E_T^{miss})$ are shown without the selections and with the inverted selections in the appendix A.5. A comparison of various modifications to the signal selection is also shown in the appendix A.5.

All selections are collected in Table 8.1 to provide a unified collection of all selections in a compact from.

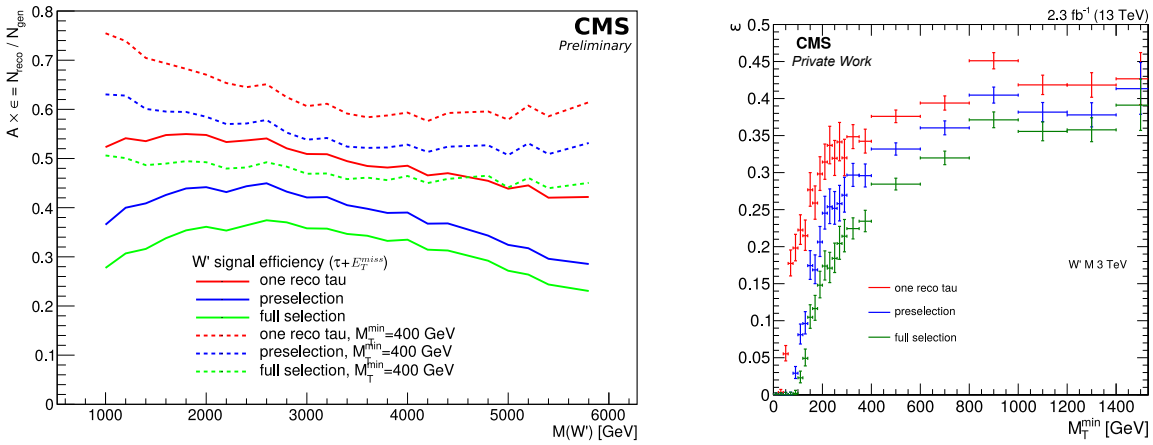


Figure 8.7.: The SSM signal efficiency for various stages in the analysis binned in W' mass (left) and as a function of the M_T threshold M_T^{min} (right). For the right plot all events above M_T^{min} are integrated.

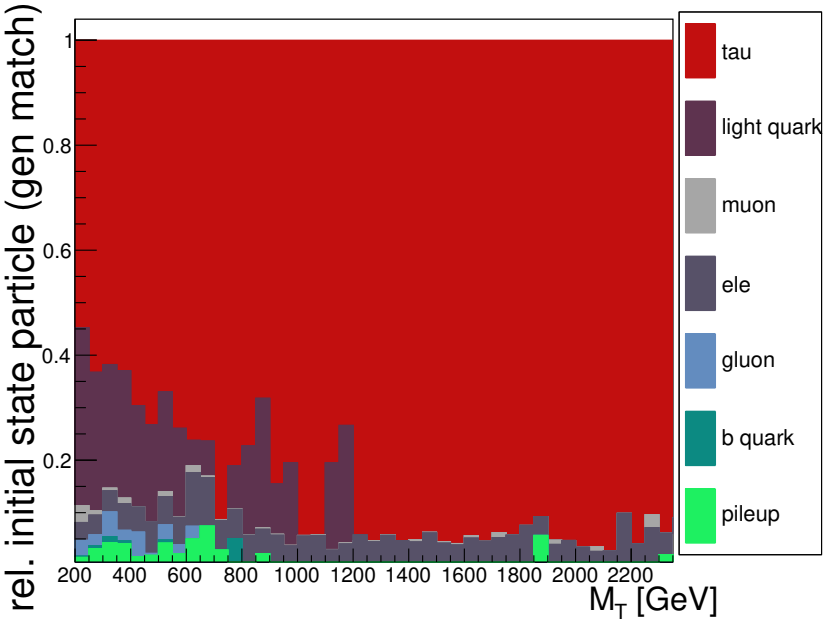


Figure 8.8.: For the background that is used from simulation, the origin of the tau was evaluated, depending on M_T. Each M_T bin is normalized to represent the tau origin as a fraction.

Name	Value
General event properties	
acceptance for τ leading tau τ	$\lvert\eta\rvert < 2.3$ and $p_T^{had} > 50\,\text{GeV}$, $\lvert\eta\rvert < 2.1$ and $p_T^{had} > 80\,\text{GeV}$
good vertex	$Ndof_{vtx} > 4$, $z_{vtx} < 24\,cm$, $(x_{vtx}^2 + y_{vtx}^2) < 4\,cm^2$
event filter	E_T^{miss} and general event filters
Trigger	
Trigger	at least one loosely isolated τ^{had} with p_T^{had} >50 GeV and $\lvert\eta\rvert < 2.1$ and $E_T^{miss} > 80\,\text{GeV}$
offline thresholds	$E_T^{miss} > 120\,\text{GeV}$ tau $p_T^{had} > 80\,\text{GeV}$
tau-ID	
decay mode	found one decay mode compatible with a τ^{had} (new incl. two prong)
QCD discriminant	medium MVA-based isolation
electron discriminant	loose MVA based
muon discriminant	loose cut based
Kinematic	
just one good τ	true
$\frac{p_T^{had}}{E_T^{miss}}$	$0.7 < \frac{p_T^{had}}{E_T^{miss}} < 1.3$
$\Delta\phi(p_T^{had}, E_T^{miss})$	$\Delta\phi(p_T^{had}, E_T^{miss}) > 2.4$
Lepton Veto	No well identified muons or electrons

Table 8.1.: All selections for the analysis performed in Run II.

8.5. Determination of the Multijet Background from Data

As for the Run I analysis the background from QCD jets faking taus is important and can not be described well by simulation. The production cross section for the QCD background has risen with respect to the 8 TeV centre of mass energy. This increase is larger in the processes with gluon initial states than in initial states with quarks, as one would expect from Figure 2.3. The increase in QCD multijet background is especially noticeable for the low p_T^{had} tau fakes, where more jets lead to a higher fake contribution and in consequence a larger p_T^{had} dependence. The method for the estimation of multijet background is changed w.r.t. the previous analysis, where a kinematic side band was used to evaluate the fake probability. In the 13 TeV analysis, events with light leptons in the final state (dominantly W-boson events) are used to evaluate the fake probability. Two reasons led to the change in the method: The p_T^{had}/E_T^{miss} sideband with $2.3\,fb^{-1}$ of data has not enough events to calculate a fake probability and the advanced analysis framework allows an easy evaluation of electron or muon triggered data.

For the prediction four regions are used which are depicted in Figure 8.9. The control regions (C and D), used to evaluate the fake probability are required to have one isolated electron or muon and a tau candidate, while the signal regions are required to have one tau candidate (A and B). The regions are divided in isolated tau candidates which pass the QCD jet isolation criteria[13] (regions B and D) and non-isolated tau candidates regions (A and C).

The control region is dominated by boosted W-boson events with an isolated electron or muon. The e/μ vetos are inverted for this purpose and the p_T thresholds raised from 25 GeV to match the single lepton triggers [14] to 35 GeV and 53 GeV for electrons and muons, respectively. Also the allowed η region is constrained accordingly.

The signal region with isolated taus (B) contains processes that have fake taus. Since the simulation description of the isolated taus is used, these backgrounds are subtracted in the non isolated signal region (A). Figure 8.10 shows the non isolated tau distribution for the distributions previously shown in Figures 8.6 and the final M_T distribution as shown in Figure 8.16 for the isolated taus. In Figure 8.10 the contribution from $Z \to \nu\nu$+ jets and QCD events is also plotted from the simulation. These processes will later be predicted by data. Because the signal region is also defined to contain exactly one isolated tau, the probability for multiple jets faking a single tau has to be calculated as depicted in Equation 7.8.

As can be seen in Figure 8.10, the description of non isolated events is not perfectly modelled by the simulation. The largest discrepancy is visible for the non signal region of $p_T^{had}/E_T^{miss} > 1.3$ where the LO QCD simulation is not able to model events with a large p_T^{had} for the tau candidate and a small E_T^{miss}. To correctly describe this region a more precise higher order simulation would be needed. The signal region $0.7 < p_T^{had}/E_T^{miss} < 1.3$ is described reasonably well, even by the simulation. The distribution of the $\Delta\phi(\tau^{had}, E_T^{miss})$, where only events with $0.7 < p_T^{had}/E_T^{miss} < 1.3$ are plotted, is described better by the simulation, but the insufficient number of generated QCD events is clearly visible. These plots confirm that a data driven approach to model the jet-tau fakes is important. The last distribution in Figure 8.10 shows the M_T distribution after the kinematic selection for non isolated tau candidates. In contrast

[13] byMediumIsolationMVArun2v1DBnewDMwLT
[14] HLT_Ele27_eta2p1_WPLoose_Gsf_v*, HLT_Mu50_v*

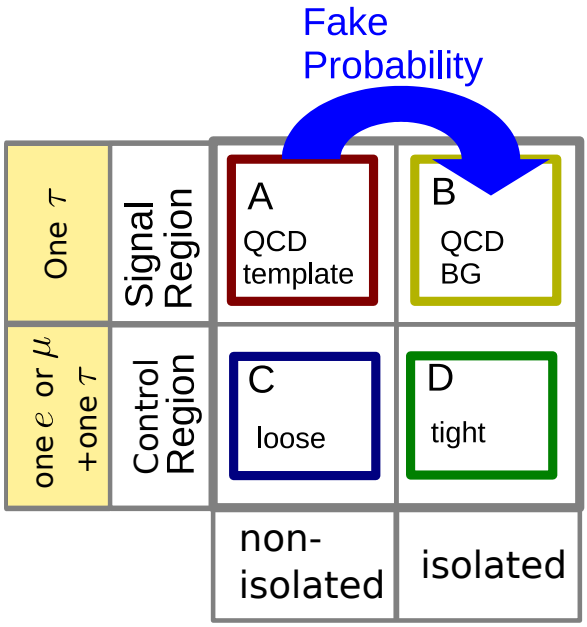


Figure 8.9.: The regions that are used to define the fake probability (C and D) and the regions that are used to apply the fake probability (A to predict B).

to the Run I analysis a reasonable description by the QCD simulation is possible, but the number of generated events is insufficient. One can see that the composition of the non-isolated background depends on M_T. While at low M_T the QCD background is almost equal to the boosted W background, the importance of the QCD background is reduced at high M_T and the background fraction described by $Z \to \nu\nu+$ jets is increased.

8.5.1. Evaluation of the Jet-Tau Fake Probability

The distributions in the e/μ final state with isolated (D) and non-isolated (C) events can be seen in Figure 8.11. The grey hashed coloured background is the fraction of events, where the tau is matched to a "true" tau at generator level. As expected the fraction of true taus is larger in the isolated region (D). The W-boson background is treated as the dominating fake background in this final state, since the fake probability from e/μ is negligible in comparison with the jet fake probability. The same plots shown separately for events with one e or μ in Figure A.9 and A.10 in the appendix.

For both the isolated and non-isolated region, one can see that the low and high p_T^{had} region is not well described by the simulation. The low p_T^{had} region is in both regions dominated by QCD-jet events, which are only simulated at LO. For the high p_T^{had} region the W process is dominant, which is also only simulated at LO. The W process is corrected for NNLO QCD and NLO electroweak effects, but this correction is done as a function of the invariant mass $M_{l\nu}$. Most of the shown events in Figure 8.11 have an invariant mass close to the W peak. A correct higher order correction would have to be evaluated as a function of the transverse momentum of the W. In summary the description of the isolated and non-isolated taus in

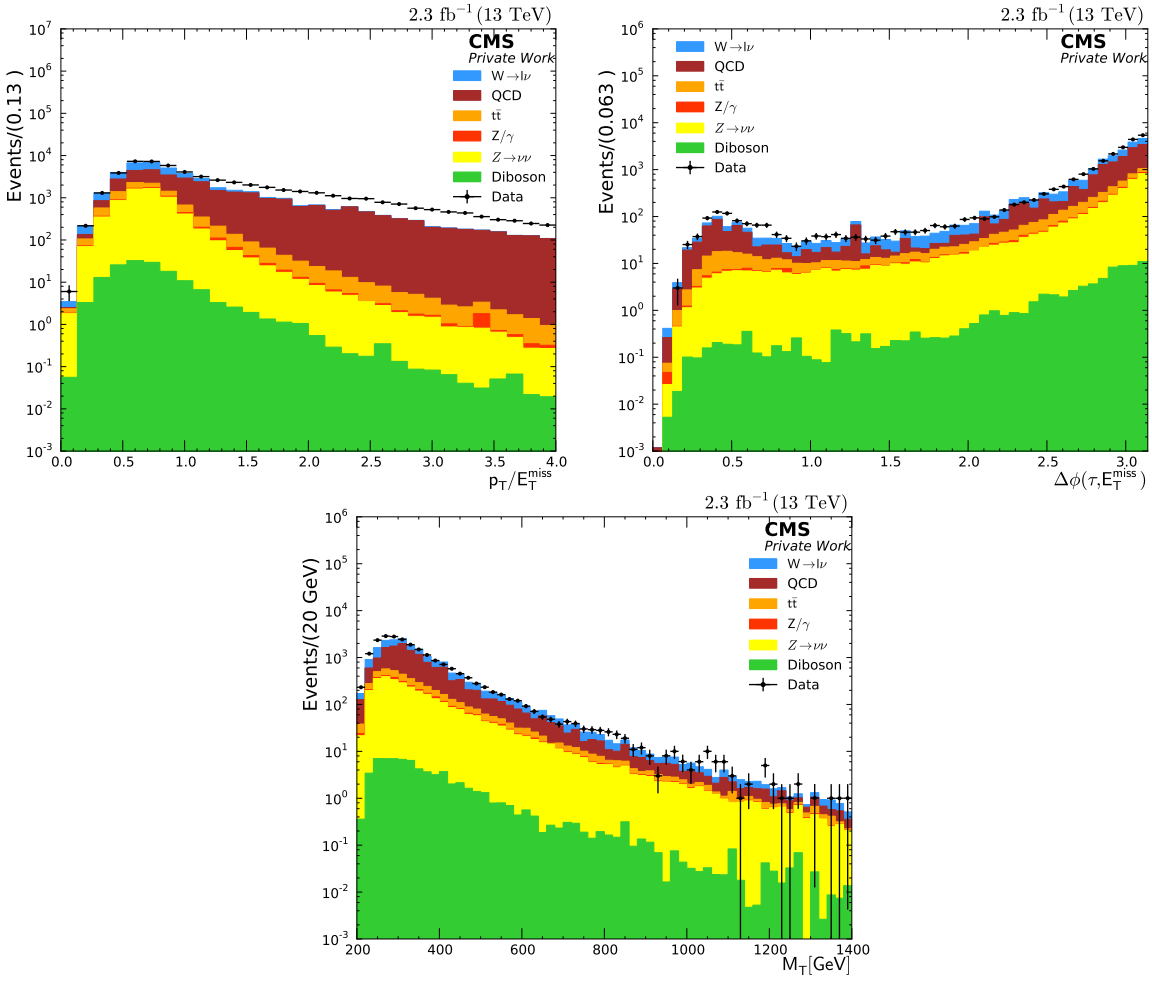


Figure 8.10.: The plots show the non-isolated tau distribution (region A) for the Figures 8.6 and 8.16. The contribution from fake jets is filled with $Z \rightarrow \nu\nu$ and QCD simulation. For the upper plots the signal selection of $\Delta\phi(\tau^{had}, E_T^{miss}) > 2.4$ is applied for the p_T^{had}/E_T^{miss} plot and the $0.7 < p_T^{had}/E_T^{miss} < 1.3$ selection is applied for the $\Delta\phi(\tau^{had}, E_T^{miss})$. For the M_T distribution both selections are applied.

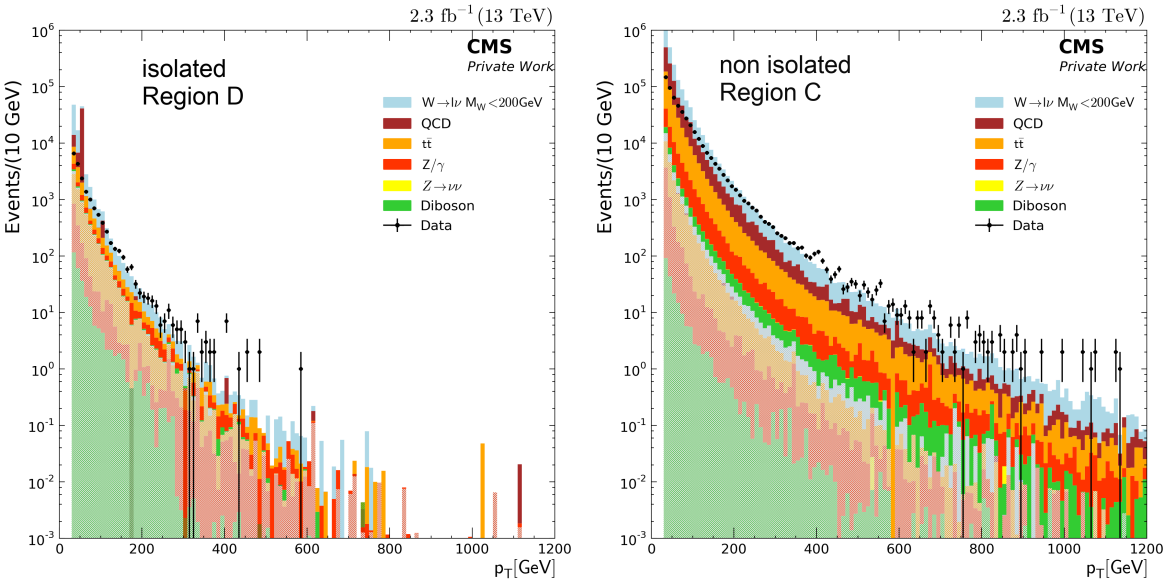


Figure 8.11.: The plots show the tau candidate p_T^{had} distribution of isolated -region D- (left) and non-isolated events -region C- (right) in the combined $e\mu$ control region. The grey hashed coloured backgrounds are matched to a tau on generator level.

the e/μ plus τ^{had} final state is expected to be rudimentary from the simulation, underling the importance of a data driven approach.

The tau p_T^{had} distribution with the subtracted events containing a "true" tau can be seen in Figure 8.12 (left). The plots show the distributions from Figure 8.11 for data and simulation where the grey hashed area is subtracted. The non-isolated region C has one order of magnitude more events as the isolated region A, as expected. In the right plot the ratio of both regions are shown. While the data estimation has a smooth description of the fake probability, simulation suffers from the statistical fluctuations. A fit to the fake probability with the function $'a/(p_T * b + c) + d'$ provides a smooth description, even at high p_T. The fake probability evaluated in data is used. The jet to tau fake probability for taus with p_T^{had} larger than 300 GeV is below 2%. The description of the QCD simulation at low energies is non perturbative and relies on many aspects such as the parametrization of the Pythia 8 tune. Therefore a non perfect description is expected.

Since the tau fake probability can have many influences, its dependence on other variables is checked by binning the fake probability in η and in terms of the decay mode. The various fake probabilities can be seen in Figure 8.13. The decay mode has the largest impact on the fake probability. The one prong zero π^0 decay mode has a fake probability in the order of 10%. This seems very high, but it concerns only a very small faction of the background since the probability for a QCD jet to have only one track and no associated ECAL clusters is very low. If there are ECAL clusters, the decay mode would have picked them up and the reconstructed decay mode would be "one prong one π^0". In terms of η, as expected, the central regions have less fakes than the forward region.

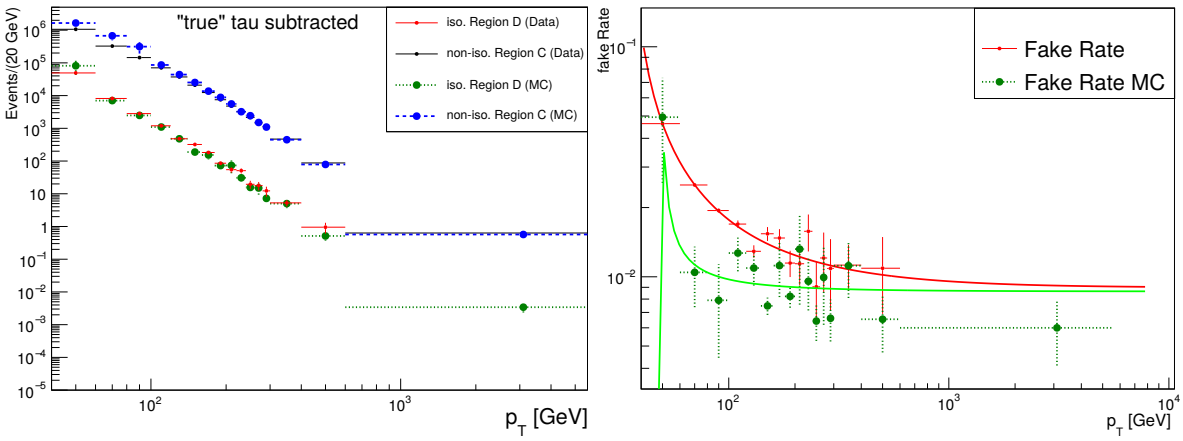


Figure 8.12.: The plots show the distributions from Figure 8.11 for data and simulation, both with events containing a "true" tau subtracted. On the right the ratio of both regions are shown, and fitted with the function given in the text.

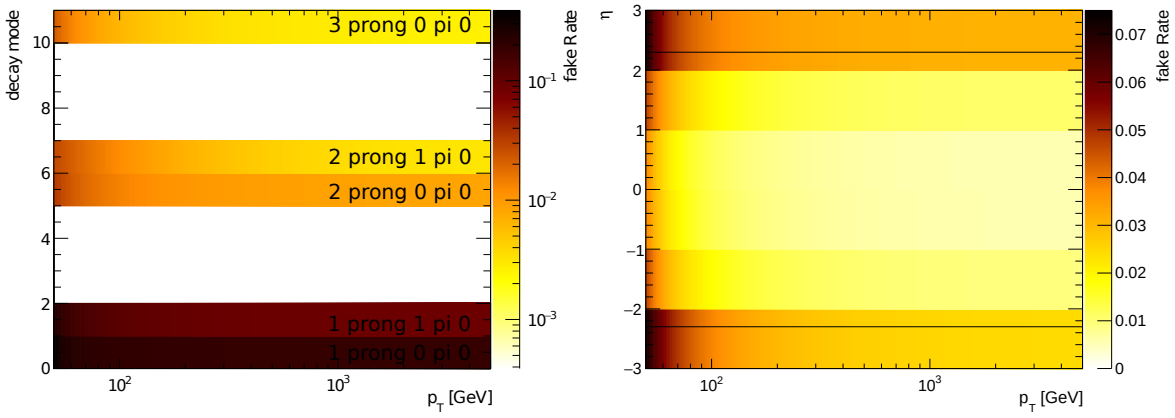


Figure 8.13.: The plots show the tau candidate fake probability based on the fit described previously, on the left p_T^{had} and decay mode dependent, and on the right p_T^{had} and η dependent. Both plots are evaluated on data events, with truth matched taus subtracted. The left plot is used as a control distribution.

8.5.2. Comparison and Check of the Fake Probability

In order to see if the prediction from the data driven estimate is able to reproduce the desired background contributions the method is applied on simulation only. In Figure 8.14 the simulation for the isolated taus (region B, solid histograms) is plotted with the scaled simulation from the non isolated taus (region A, black line). To scale the simulation the function shown as green line in Figure 8.12 (right) is used. The shown values are the control plots p_T^{had}/E_T^{miss}, $\Delta\phi(\tau^{had}, E_T^{miss})$ and the distribution of the τ^{had} candidates p_T. For the p_T^{had}/E_T^{miss} distribution the $\Delta\phi(\tau^{had}, E_T^{miss})$ selection is applied, vice versa. For the p_T^{had} distribution both selections are required. One can see that the QCD description from simulation suffers from an insufficient number of generated events. The events in the high p_T^{had}/E_T^{miss} region can not be described smoothly by simulation. For the $\Delta\phi(\tau^{had}, E_T^{miss})$ and M_T distribution the shape of the backgrounds can be reproduced astoundingly well, even with the requirement that

the simulation should contain exactly one non isolated tau, which is a non trivial quantity to model. The comparison of isolated events from simulation and non-isolated, weighted events from simulation shows that the method works and the shape of the distributions can be reproduced.

Using the fake probability estimated from data the various parametrisations can be compared in Figure 8.15 the parametrisation in terms of p_T^{had} (Figure 8.12 right), p_T^{had} and decay mode (Figure 8.13 left) and p_T^{had} and η (Figure 8.13 right) are plotted as lines. The same distributions as in Figure 8.14 are shown for region B, with the same requirements. The prediction from simulation as seen in Figure 8.14 is shown as well, to allow an easy comparison. Also shown is the distribution of the measured data, where the non fake background is subtracted[15]. The ratio to the p_T^{had} and η dependent fake probability prediction is shown below the plots. One can see that for high p_T^{had}/E_T^{miss} the simulation predicts not as many events as one would expect. For low p_T^{had}/E_T^{miss} the agreement is very good. In the $\Delta\phi(\tau^{had}, E_T^{miss})$ distribution, the region with low event numbers ($\Delta\phi(\tau^{had}, E_T^{miss}) < 2.4$) suffers from huge fluctuations. The expected shape of the control region data can be reproduced nonetheless. The fake probability estimated as a function of the decay modes can not provide a smooth description for this low event number region. The fake probability depending on p_T^{had} and depending on p_T^{had} and η provide almost the same smooth prediction. In the p_T^{had} distribution one can see that they are slightly disagreeing at high p_T^{had}. Since the description for both is smooth, the $\eta - p_T^{had}$ dependent fake probability will be used as it takes into account one additional dependency.

The uncertainty on the p_T^{had}-η dependent fake probability is estimated to be 50% this contains most of the other parametrisations as indicated in the grey bands in the ratios of Figure 8.15. The value is chosen conservatively and no additional uncertainty from the variation of the subtracted background is used.

[15] In order to have a meaningful uncertainty the uncertainty of the background is set to $\sqrt{n}$.

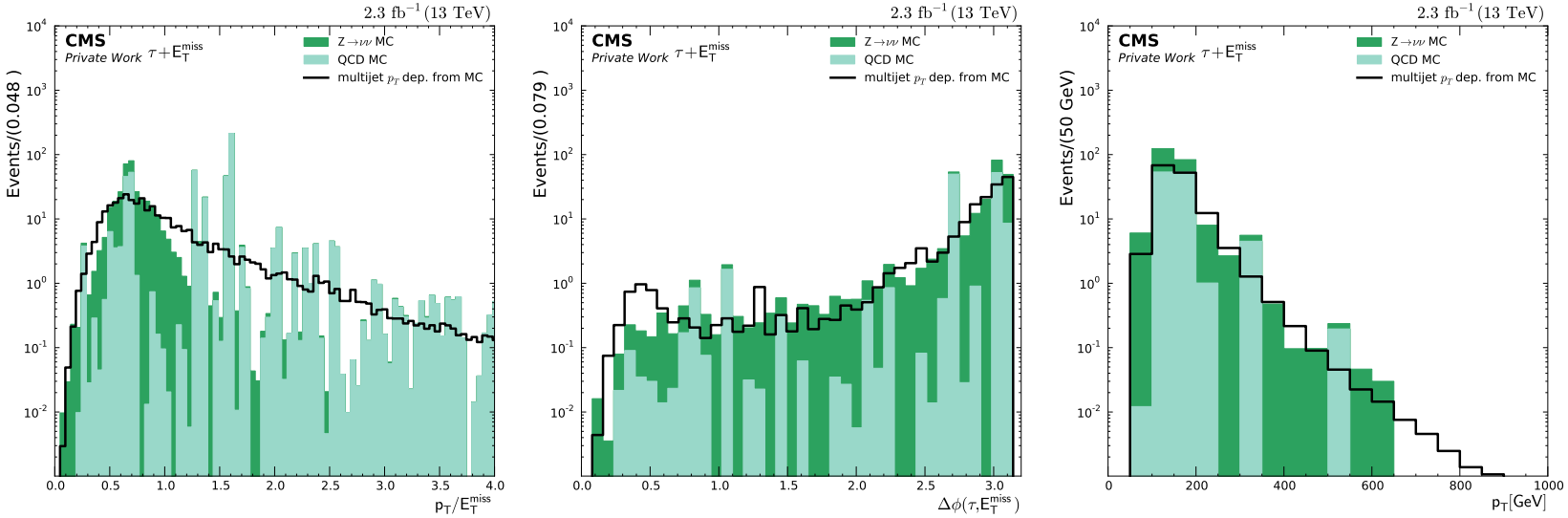


Figure 8.14.: The histogram shows the direct prediction from simulation in region B. The line shows the prediction from region A scaled with the fake probability evaluated in simulation, from Figure 8.12. The shown distributions are p_T^{had}/E_T^{miss}, $\Delta\phi(\tau^{had}, E_T^{miss})$ and p_T^{had}. For the p_T^{had}/E_T^{miss} distribution the $\Delta\phi(\tau^{had}, E_T^{miss})$ selection is applied, vice versa. For the p_T^{had} distribution both selections are applied.

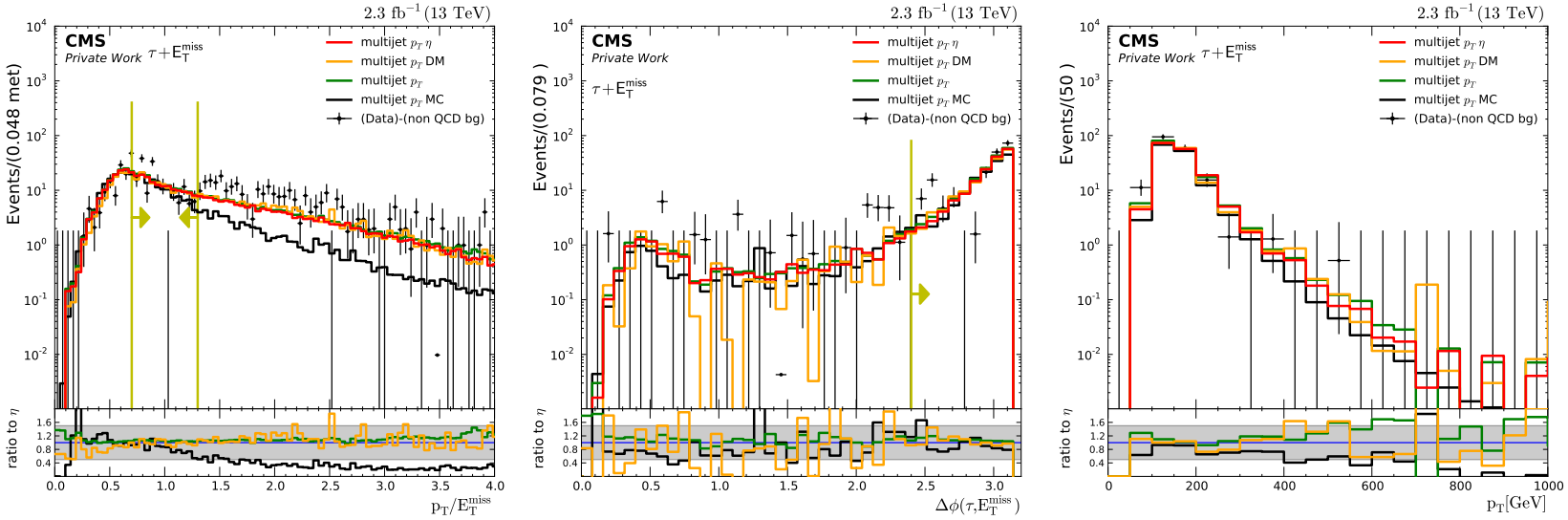


Figure 8.15.: The same plots as in Figure 8.14, with the data from region A scaled to region B. For this scaling several parametrisations of the fake probability are used. The p_T^{had}, p_T^{had}-decay mode (DM) and the p_T^{had}-η parametrisation. For the p_T^{had} parametrisation also the prediction from simulation as shown in Figure 8.14 is shown. Also shown is are the data values in region B with the non multijet contributions subtracted.

8.6. Full Background Prediction and Final Distributions

The final multijet contribution is 160 ± 80 events. With respect to the total background the fraction of QCD jet events is roughly 30%. This is the same as for the Run I analysis. This is only possible with the use of the MVA-based isolation discriminator, which has a lower fake probability than the cut-based isolation. The cross section of the QCD jet background increased more than the signal and most other background processes, because with the rise of the centre of mass energy from $\sqrt{s} = 8\,\text{TeV}$ to $\sqrt{s} = 13\,\text{TeV}$ the gluon induced processes are preferred by the PDF[16].

The distribution in M_T is shown in Figure 8.16. The entries are also given in Table 8.2. As can be seen in the ratio of Figure 8.16 (left), the data fluctuates around the predicated background. The biggest discrepancy between background and data is around 650 GeV. This is more pronounced, viewing the right plot of Figure 8.16, where the events above the M_T threshold on the x-axis are integrated. The local significance of the data above the M_T =650 GeV is -2σ. Compared to the Run I analysis, this deficit is roughly at the same position as in Run I. Due to the nature of the tau decay no signal could be found, which would produce a deficit in the M_T distribution. Even a destructive interference as discussed in Section 2.3.1 would not appear as a deficit in data, since events from the W' peak would be reconstructed with lower M_T, predicting more events than the SM background. The localisation of the deficit is interesting, pointing to a correlation between the Run I and Run II background prediction. As the method for the multijet background changed between the two realizations of the analysis, therefore the multijet method as source for this deficit is less likely. The for the main W-boson background the generator for the high mass samples was changed from PYTHIA 6 to PYTHIA 8, but the programs to estimate the (N)NLO differential cross sections is the same (FEWZ and MCSANC). Also the PDF set used for the (N)NLO calculations were not changed. This may be a hint for insufficient theoretical description. It is also possible that the tau identification despite the performed cross checks has an inefficiency, which is not simulated. The changes between the Run I and Run II analysis, such as the different isolation and changes to the strip reconstruction do not support this explanation.

In order to see the difference in the decay modes all kinematic distributions are shown, split into the reconstuced decay modes in the appendix A.5.

M_T^{min} / GeV	diboson	Z/γ	$t\bar{t}$	multijet	W	Background	Data	W'	
								1 TeV	3 TeV
200	11	29	50	160	380	634 ± 98	650	1500	6.1
300	5.4	13	20	87	180	305 ± 53	300	1400	6.1
600	0.33	0.68	0.63	3.1	12	17.0 ± 3.5	11	660	5.3
900	0.024	0.11	0.20	0.46	2.3	3.14 ± 0.80	1	110	4.5
1020	0	011	0.11	0.25	1.3	1.66 ± 0.46	1	28	4.1
1050	0	0.0087	0.11	0.19	1.1	1.44 ± 0.41	0	21	4.0

Table 8.2.: The event numbers for various backgrounds, signal and data above a given M_T^{min} threshold for the Run II analysis. The errors include the statistical and systematic uncertainties as evaluated in 9.2.

[16]This can be seen in Figure 2.3, where with rising Q^2, the gluon probability increases more than the quark probability.

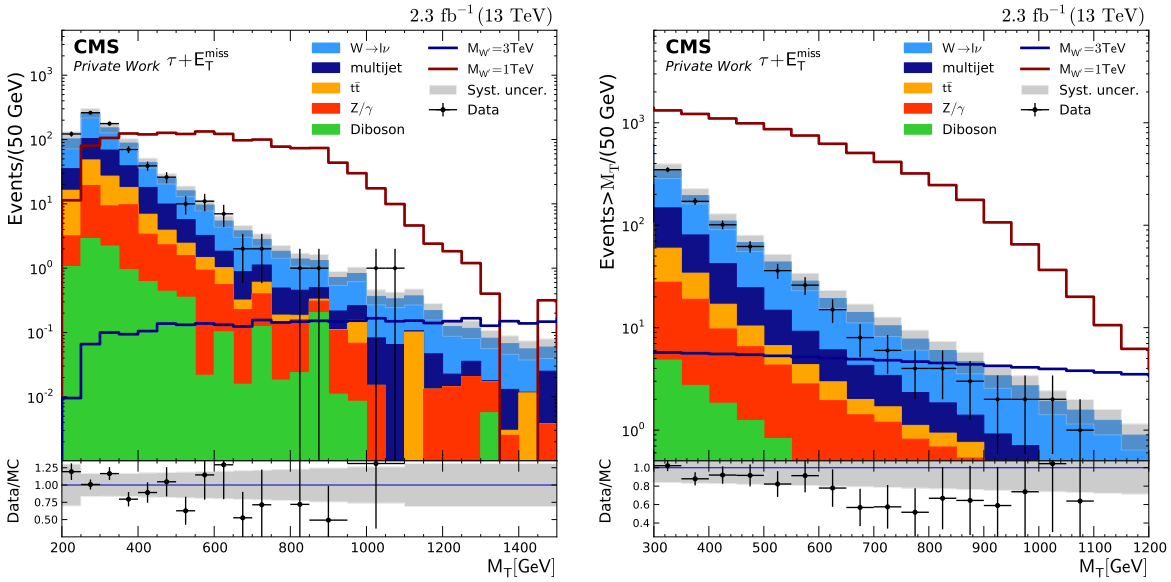


Figure 8.16.: Final M_T plot is shown on the left. In the right plot the events are integrated above the M_T threshold on the x-axis. In the lower ratio the bin width is increased to have a minimum of one event in the background prediction. The figure is the same as published in [3] with a minimal modification of the systematic uncertainty as described in Section 9.2.

9. Systematic Uncertainties

In this chapter the systematic uncertainties that are relevant for the final result will be discussed. The uncertainties are similar in the Run I and Run II analyses. Therefore a detailed description of the systematic uncertainties for the $\sqrt{s} = 8\,\text{TeV}$ data taking period will be given first, and then the differences to the $\sqrt{s} = 13\,\text{TeV}$ data taking period will be given.

9.1. Sources of Systematic Uncertainties at 8 TeV

There are two different types of systematic uncertainties: One type directly changes the number of events in the final M_T distribution by varying scale factors like the uncertainty on the integrated luminosity or on the object identification efficiencies. This does not alter the shape of the M_T spectrum. The second type has an indirect influence like the uncertainties on the energy scales and resolutions of the reconstructed objects which can cause bin-to-bin fluctuations of events in the kinematic distributions. The influences of this second type of uncertainties on the signal efficiencies and background predictions are calculated by repeating the analysis with the shifted or smeared energy values of these objects to obtain distorted final distributions for each uncertainty. The deviation between the original and the distorted distributions is caused by these second type uncertainties. To obtain the effect of each uncertainty on the full signal efficiency, the efficiency is recalculated with each of the distorted distributions.

The following uncertainties are considered in this analysis:

- **Tau identification efficiency uncertainty:** The uncertainty on the efficiency for the tau identification was determined by the tau POG to be 6% [147]. This uncertainty is used to rescale all simulated events in the final distribution. The tau identification efficiency was estimated with Z events, but no indication for a larger efficiency uncertainty at high p_T^{had} was found. The uncertainty could also be confirmed with $t\bar{t}$ events, where events with p_T^{had} >100 GeV were probed, but the energy resolution of this analysis is insufficient at high p_T [147].

- **Tau energy scale uncertainty:** The uncertainty on the scale of the tau energy is not driven by the measurements of the subcomponents of the tau. The main uncertainty on the tau energy scale is due to missing constituents, not included in the tau reconstruction. The difference between data and simulation was evaluated using jet and tau reconstruction algorithms. Details can be found in Section 4. The results where compatible with the estimated uncertainty of 3% for events with lower p_T^{had} as described in [4].

- **Jet energy scale uncertainty:** The uncertainty on the jet energy scale has been studied in [148]. The relevant uncertainty varies between 3% and 5%, depending on p_T and η. All jets that are not identified as taus are shifted accordingly.
- E_T^{miss} **uncertainty:** Since E_T^{miss} is a composite object, all components are evaluated and shifted individually; namely muon and electron energy scale and resolution as well as unclustered energy scale are taken into account. The muon uncertainty is discussed in [92], electron uncertainty in [149] and the unclustered energy, which is determined from energy deposits that do not form a jet of $p_T < 10\,\text{GeV}$ have an uncertainty of 10% on their p_T.
- **Jet energy resolutions uncertainty:** The uncertainty of the jet energy resolution (JER) [148] dependents on jet p_T and η. The energy smearing is done separately for each of the three components of the momentum vector: The true value of each component is used as the mean value of a gaussian distribution with the standard deviation being equal to the initial resolution of this jet increased (or reduced) by the uncertainty. Using these gaussian distributions, new values of the three components are randomly determined. This cannot only change the energy of the jet but also its direction, which is important to model the influence on E_T^{miss}. This approach is very conservative, but as well will see the effect on the final distributions is so small, that it is included in the E_T^{miss} uncertainty.
- **Pileup reweighting uncertainty:** The uncertainty on the pileup weights arises from the uncertainty on the distribution of pileup interactions per event in data. The number of pileup events in data is not identical to the number of vertices in an event and cannot be measured directly. Instead, it is calculated from the overall proton-proton cross section (see Section 7.2). The uncertainty on this cross section is taken into account by recalculating the pileup weights with a second distribution of pileup interactions per event in data, which is obtained from a slightly different assumption for the cross section [136]: The first assumption uses a cross section which is gained by extrapolating the value measured at 7 TeV data to 8 TeV (68 mb). For the second assumption, measurements at 8 TeV are used (69.4 mb). The analysis is repeated with the new weights to obtain the impact on the final distributions.
- **Luminosity uncertainty:** The uncertainty on the determined integrated luminosity is assumed to be 2.6% as stated by [150]. Since the integrated luminosity is used to scale the simulation distributions, this uncertainty has a direct influence of 2.6% on the number of background events. The signal efficiency is independent of the integrated luminosity and therefore this uncertainty has no effect on it, but the limit will be provided as a limit on the cross section, therefore it effects the signal in the same way as the background.
- **QCD data driven method uncertainty:** As described in Section 7.5, the relative uncertainty on the fake probability is estimated to be 20%. The resulting uncertainty on the number of QCD background events is determined by generating two additional QCD background samples where for the first one a fake probability enhanced by 20% and for the second one a fake probability reduced by 20% is used to scale the template sample. In addition the background samples subtracted in this method are varied by ±20% and the resulting distribution is added shape dependent as a systematic uncertainty.
- **k-factor uncertainty:** The W process at hadron colliders can be simulated in all kinematic distributions up to NNLO in α_s, and up to NLO in α. The correction for the

W spectrum is calculated as the additive combination of the higher order effects. The difference relative to the multiplicative combination is used as an uncertainty.

9.1.1. PDF Uncertainties

To estimate the PDF uncertainty a different approach is used. As shown in Section 2.1.6, each PDF set has intrinsic systematic uncertainties, which are based on their inputs and extrapolations. Therefore better extrapolations and more inputs yield in a smaller uncertainty. One of the outstanding problems at the LHC era of PDF estimation is that different groups estimate the PDFs and quote uncertainties for their estimations, but their results do not agree with other groups. An example can be seen in Figure 2.4, where for small and medium range of the PDF parameter x MMHT14 [46] and NNPDF3.0 [151] disagree by more than 2σ. In order to estimate the difference between the PDF groups the total uncertainty envelope on the final state distribution has to be estimated. This is no trivial approach since two different concepts are used by different groups to estimate the PDFs.

The MC based PDF (NNPDF [151]) uses measurements to train a neural network. As an output MC events can be provided, that represent the PDF description of the phase space. This method provides an intrinsic handling of the uncertainties and allows cross checks of the data, method and underlying theory. The uncertainty on a parameter like the cross section can be evaluated, using the spread of the MC outputs and constructing a 68% uncertainty interval.

The hessian based PDFs (CTEQ [152], MMHT/MSTW [46], HERA [153]) use a functional description of the theory, which is optimized to fit the data. These PDFs provide uncertainties in the form of n uncorrelated parameters, which are shifted with respect to their uncertainty. The uncertainty on the cross section can be evaluated from the spread of the fit results, shifted by these n-parameter uncertainties.

A generation of all events with these different uncertainty PDFs would mean an increase of the number of generated events by a factor of $\mathcal{O}(100)$. Since this is not reasonable a reweighting of the events is used, where for each event a weight is calculated, using the initially generated PDF parameters x and Q. The weight then only modifies the initial state of an event. The hadronisation is unchanged. This method works for leading order backgrounds, as was shown in [154]. Since the PDF has a strong α_s dependence, the α_s uncertainty is usually evaluated at the same time as the PDF uncertainty.

The PDF4LHC working group [155] worked out a recommendation on how to combine these uncertainties [156]. Their recommendation is to use PDF sets from three independent groups, including their uncertainties and use the total error envelope as a systematic error on the PDF estimation. For this analysis the PDF sets CT10, MSTW2008 and NNPDF2.3 were used to evaluate this envelope as shown in Figure 9.1 (right). The envelope should then contain one σ of all possible PDF values.

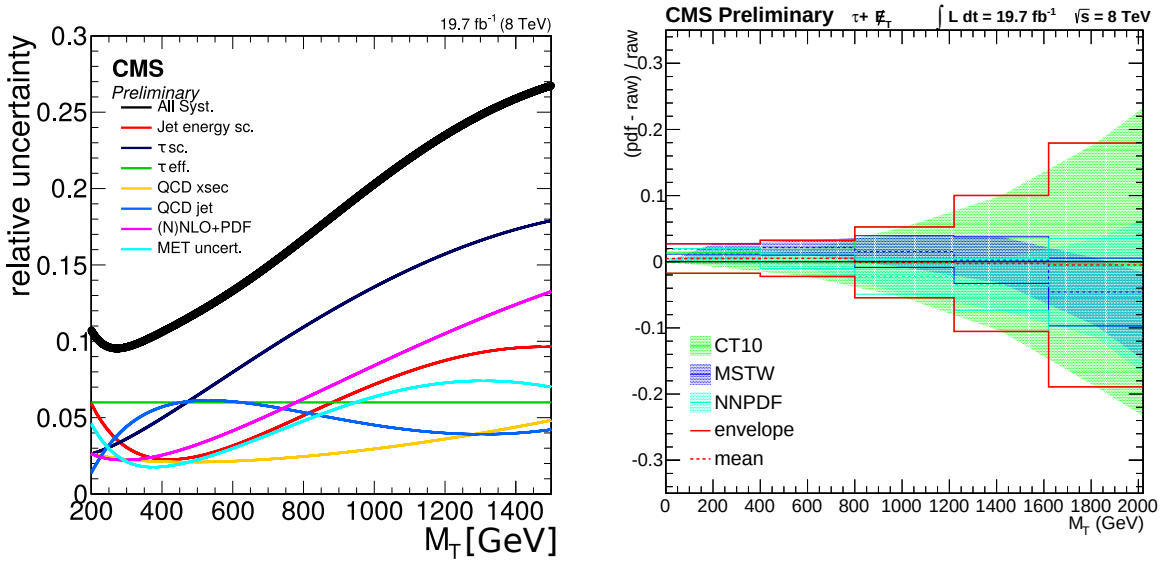


Figure 9.1.: Relative uncertainties on the number of background events in bins of M_T for all uncertainty sources causing a distortion of the spectrum and the uncertainty arising from the uncertainty on the QCD data-driven method. The entry labelled "All Syst." refers to the sum off all uncertainties causing a distortion of the spectrum. The right plot shows the effect of the three different uncertainty PDF sets on the M_T distribution.

9.1.2. Impact of the Uncertainties on Signal Efficiencies and Background Prediction

The impact of the different systematic uncertainties which cause bin-to-bin variations on the M_T spectrum of the background processes can be seen in Figure 9.1 as the relative uncertainty on the number of background events in a certain M_T range. The backgrounds follow a steeply falling distribution with no structure. In order to reduce the effect of insufficient number of simulated events and other statistical non physical effects the background is fitted with an exponential:

$$\exp\left(a + bx + cx^2\right) \cdot x^d \tag{9.1}$$

This fit is used for all shifted distributions listed above and the relative uncertainty is shown in Figure 9.1. One can see that the effects of most uncertainties on the M_T spectrum is small. A significant influence has the tau scale uncertainty, and the tau efficiency uncertainty. The effect of the jet specific uncertainties (jets are cleaned for taus) has an influence of up to 6% on the M_T spectrum. The k-Factor uncertainty is small for low M_T and rises up to 4% at M_T =1.4 TeV. For $M_T \approx 200$ GeV all the uncertainties rise rapidly, because of the trigger turn on in the M_T distribution. The effect of the uncertainty parametrisation on the limit was checked and a cross section limit with and without the parametrisation of the uncertainties was calculated, resulting in the same limit on the W' mass.

The PDF uncertainties are evaluated as described before, reweighting all events after all signal selections.

The impact of the uncertainties on the signal is also evaluated following the same procedure.

Since each signal has a different distribution in M_T a simple visualisation as for the background is not possible. Figure 9.2 shows the shifted distributions for the tau energy scale uncertainty, which is the uncertainty with the largest effect on the signal. One can see that a scale shift has no influence on the event yield, where the distribution is flat, but on the falling edges at low and high M_T the difference to the central not shifted distribution becomes large. This reflects the expected effect of an unknown scale. As a result the exact position of the edge will be masked by high uncertainties. The distributions as shown in Figure 9.2 will enter the statistical evaluation as described in Section 10.1.

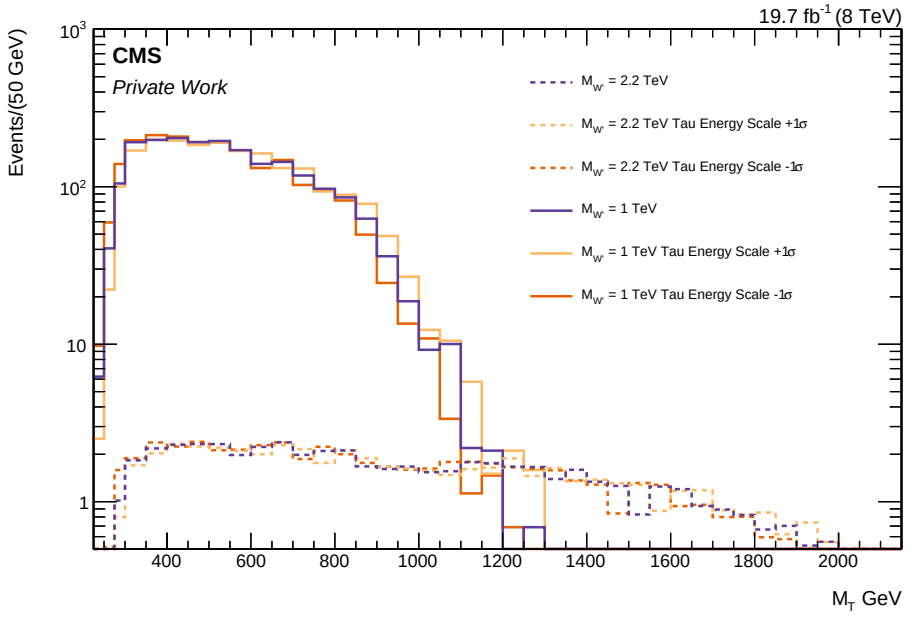


Figure 9.2.: The effects of the tau scale uncertainty for two exemplary W' masses are shown in this Figure. The shifted distributions are shown in yellow and orange for a shift of $\pm 1\sigma$. The unshifted distributions are shown in violet.

9.2. Sources of Systematic Uncertainties at 13 TeV

The systematic uncertainties are to a large extent the same as the 8 TeV uncertainties since no essential parts of the analysis changed. The main changes are the luminosity, which is measured to a precision of 2.7% [157]. The QCD uncertainty is 30% and new jet energy corrections were calculated for the smaller jet cone and changed detector. For the higher centre of mass energy the total inelastic cross section was kept at 69 mb with an uncertainty of 5%. But for the calculation of the pileup events in an event the smaller bunch spacing and consequent smaller number of protons per bunch also is taken into account. Also the PDF uncertainty method changed with respect to Run I procedure. Where previously different PDF sets were evaluated to get an envelope of the current PDF knowledge the PDF4LHC working group produced several meta PDF sets, which already combine the spread of the PDF predictions from various groups [154]. For the Run I analysis the jet energy scale had a larger influence on the background yield since the isolation was less stringent and the trigger used the jet p_T instead of the tau p_T^{had} as threshold.

In contrast to the publication [3] the tau identification uncertainty is not approximated by the formula $6\% + p_T \cdot 20\%/1\,\text{TeV}$. This was the proposal within the tau object working group. The constant 6% uncertainty originate from the Run I study of Z events, where the uncertainty for the highest p_T objects in this channel is estimated to be smaller than 6%. As the discrimination against jets, which has the largest influence on the tau identification efficiency, becomes easier for higher p_T taus[1], the rise in uncertainty can not be justified. As a consequence the 6% tau identification uncertainty is kept, which is motivated CMS by measurements.

The resulting uncertainties for the background yield can be seen in Figure 9.3.

In summary the background uncertainties of the Run I analysis are at the order of 10-25%, depending on the M_T. For the Run II analysis the uncertainties for high M_T is almost the same, but the uncertainty for the low M_T region ($M_T < 500\,\text{GeV}$) is increased due to a larger influence of the QCD uncertainty.

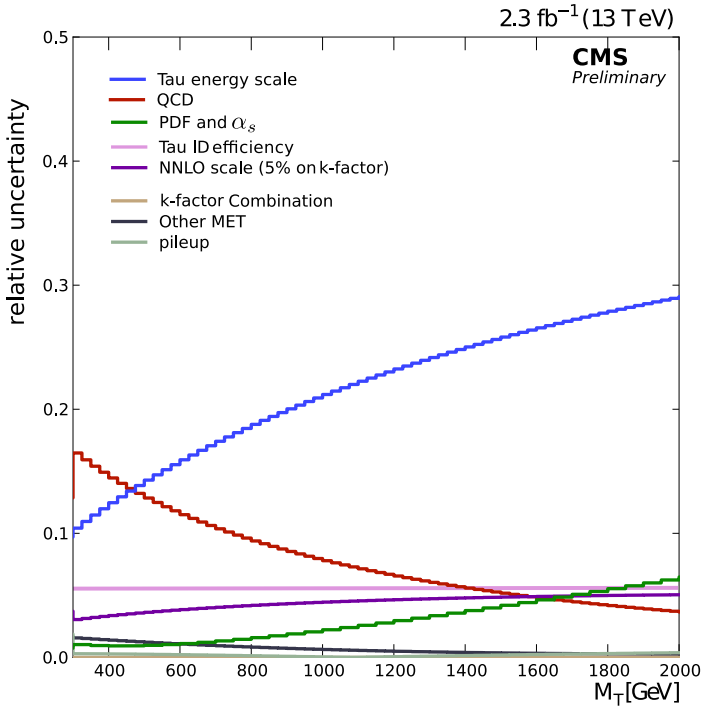


Figure 9.3.: Relative uncertainties on the number of background events in bins of M_T for all uncertainty sources causing a distortion of the spectrum and the uncertainty arising from the uncertainty on the QCD data-driven method.

[1]For QCD jets the number of constituents rises with the energy of the jet. For high p_T jets it is therefore less likely to be misidentified as a tau. This is represented by the almost constant fake probability shown in Figure 4.14.

10. Results of the Tau + E_T^{miss} Analysis

In order to quantify the compatibility of the data to the expectation one has to use a statistical measure, which provides a simple but well defined answer for the agreement between data and expectation. In modern statistics, there are two approaches to quantify the difference between a measurement and a hypothesis, which are called frequentist and bayesian [15]. The results of this statistical analysis depend on the initial purpose of the analysis. Here a limit will be set on a model parameter. For the limit an expected and observed value will be quoted. The difference of these two values is similar to a significance statement if a parameter estimation would have been performed.

10.1. General Concepts of Limit Determination

The frequentist and bayesian approaches have a fundamentally different concept for the definition of probability. For the frequentist statistics probability is interpreted as the frequency of a result for a repeated experiment. Due to the conceptual similarity of a particle physics experiment, which often repeats the same measurement, this approach is used widely for the interpretation of particle physics results. The other approach is the bayesian statistics where probability is more theoretically defined as the "degree of belief" for a specific experimental result to occur. While this seems to be more complicated, it allows a simple mathematical description of the probability for a hypothesis prior to the experiment as a function of the parameters. This is often referred to as a probability density function (p.d.f.). The problematic part is that the result can depend on the shape of this p.d.f., which is assumed prior to the evaluation of the measurements. Nonetheless it allows a natural inclusion of systematic uncertainties. In the frequentist interpretation a similar implementation can be done. The mathematical description of this hypothetical experiments is often taken from the bayesian description and called hybrid calculation, and therefore also a p.d.f. has to be assumed.

Both approaches are used in particle physics and yield similar results. In the past it was agreed between ATLAS and CMS to use the bayesian approach to set limits for a W' result. Therefore it will also be used here. A complete review of the method can be found in [15][1]. Here only the general concept and the used prior functions will be stated. A review of the common statistical parameters and their use can be found in [158] which mainly describes the frequentist approach, but provides a detailed study of the currently used statistical understanding.

The bayesian approach is based on Bayes' theorem which provides a formula for the general setup that all given possibilities are in a set S and A and B are subsets of S. The probability of B being true under the assumption that A is true is the same as A being true under the

[1] Chapter 38

assumption that B times the probability for A being true divided by the probability of B being true:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}. \tag{10.1}$$

Or in an example closer to reality: The probability a hypothesis H is true given the data (P(H|data)) is the same as the probability that the data can occur given H is true (P(data|H)) times your prior believe in the hypothesis $\pi(H)$ (e.g. systematic uncertainties) and normalized to the probability to observe the data (P(data)). The last part can be split up in a sum or integral because $P(B) = \sum_i P(B|A_i)P(A_i)$, or for our hypothesis data: the probability of the data (P(data)) is the sum over all possibilities to have the data given, given all possible hypothesis ($P(data|H_i)$) times the probability of the hypothesis ($\pi(H_i)$).

For models beyond the SM it is common to provide a limit on a model parameter θ. In particle physics it was agreed that exclusion limits should describe the parameter region, which is excluded with 95% probability, usually referred to as 95% confidence level (CL)[2], though this is a terminology used by the frequentist interpretation. This means that $\int_0^{\theta_{95}} P(\theta|\text{data})d\theta = 0.95$ sets the model parameter limits θ_{95} for H. The parameter of interest for most beyond the SM searches is the signal cross section which can be translated to a number of events. The probability to measure a data set, given the hypothesis is usually described by a Poisson likelihood function:

$$0.95 = \int_0^{\theta_{0.95}} d\theta\ \pi(\theta) \cdot L'(\text{data}|\theta) \tag{10.2}$$

$$L'(\text{data}|\theta) = \int d\vec{\nu}\ L_{\text{Poisson}}(\text{data}|\theta, \vec{\nu}) \cdot \pi(\vec{\nu}) \tag{10.3}$$

$$L_{\text{Poisson}}(n, s) = \frac{(s+b)^n}{n!} e^{-(s+b)}, \tag{10.4}$$

where $\vec{\nu}$ are the nuisance parameters, meaning parameters that model the uncertainties of the statistical model. For a counting experiment the parameters of interest are the number of signal events in the search window s and the number of background events in the search window b, which can be translated into cross sections.

For this analysis the parameter of interest is the signal cross section $\sigma = \theta$. This cross section is evaluated for different physics model parameters. For this search the cross section limit is determined as a function of the W' mass. For the limit a multibin approach is chosen which subdivides a distribution in a number of bins. In this analysis the M_T distribution is used. The likelihood function is evaluated for each bin separately considering each bin as a separate channel and combining all bins to derive a limit. This approach takes the shape of the distribution into account and is more sensitive or at least as sensitive as a single-bin method.

The prior functions for s and b are assumed to be log-normal distributions given by the systematic uncertainties (see Section 9.1).

[2]The baysian equivalent is the credibility level, which serves the same purpose and therefore will also be abbreviated CL.

It may be noted here that the quantification of a potential discovery can be calculated in a similar way.

For the calculations the Higgs combination tool [159] on top of the RooStats [160] package is used. A detailed description of the statistical details for this analysis can be found in the thesis by Mark Olschewski [57].

As for the description of the analysis the Run I results will be discussed first, followed by the Run II results. As before the Run I results will establish the basis of the analysis and for the Run II results additional results will be shown.

10.2. Limits of the 8 TeV Analysis

The limits will be set on different models. First the limit on a SSM W' will be set. The model is described in Section 2.3.1. This will be followed by the analysis of the NUGIM model, which is described in Section 2.3.2. As this is a specific model, the result can be compared to several bounds from previous experiments. The final limit before the Run II result will be a model unspecific limit, which can be used for the reinterpretation of the measured M_T spectrum in many models with the $\tau+E_T^{miss}$ final state.

10.2.1. Limits in the Sequential Standard Model

The method described before is used to set limits on an additional cross section beyond the SM contribution arising from processes which lead to similar experimental signatures as a SSM W'-boson decay. The limit on the cross section times $BR(W' \to \tau\nu)$ as a function of the W' mass is shown in Figure 10.1: The solid black line shows the limit observed with $20\,fb^{-1}$ of data while the dotted lines shows the expected limit. The green and the yellow bands indicate the one and two sigma intervals of the expected limit. The observed limit shows only a small deviation from the expected limit and stays within the two sigma band which is due to the fact that no significant deviation from the SM expectation was seen in data. The largest deviation is at 700 GeV in the M_T distribution.

The thin dotted line within a blue uncertainty band is the predicted cross section of a SSM W'. The cross section is calculated in NNLO, taken from [144]. All W' masses leading to a theoretical signal cross section higher than the observed cross section limit are excluded. The intersection between the line of the observed limit and the theoretical prediction indicates the best mass limit on the W'-boson derived from this analysis. The SSM W' is excluded for $0.3\,TeV < M_{W'} < 2.75\,TeV$ at 95% CL in the tau channel.

The low M_T region has an under-fluctuations of data events with respect to the background prediction, leading to a better observed limit w.r.t. the expected for W' masses of about 1 TeV. For high M_T a little less events are expected than observed, leading to a higher observed cross section limit than is expected. The shape of the exclusion limit originates from the shape of the signal distributions in M_T. For low W' masses most of the signal is lost, due to the trigger thresholds, leading to the steep flank of the excluded cross section for $M_{W'} < 500\,GeV$. For higher masses the M_T distributions of standard model and signal are easier to separate, providing a better limit. But for W' masses above 2.5 TeV the low mass off-shell production

of the signal becomes dominant, because the W' peak is PDF suppressed. This leads to a rise in the excluded cross section for very high W' masses.

An overview of the limits for the different mass points, yielding the best expected limit are shown in Table 10.1.

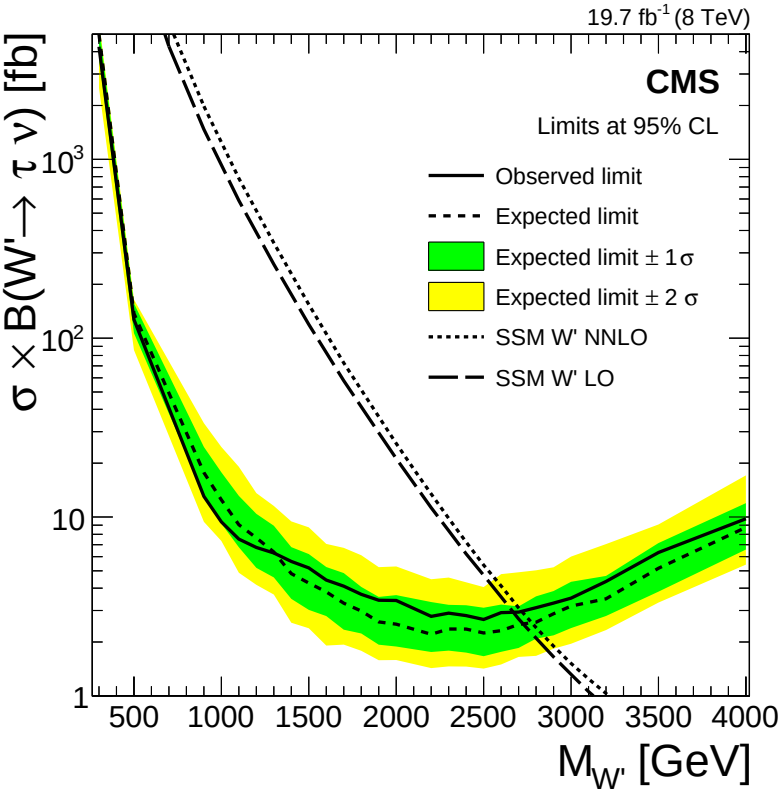


Figure 10.1.: Cross section limit as a function of the mass of the W' boson in the Sequential Standard Model. All W' masses below 2.75 TeV, which corresponds to the intersection of the dotted theoretical NNLO cross section line and the observed cross section limit, in the tau channel are excluded at 95 % CL.

10.2.2. Limits on NUGIM Model

Additional to the limit on the SSM W', limits are set on the parameter space of the NUGIM model, see Section 2.3.2. Contrary to the former one, the limits on the NUGIM model are calculated with leading order signal cross sections since no higher order corrections are calculated, yet. For each value of the parameter $\cot\theta_E$ of the model and each mass value a separate cross section limit is derived since the signal efficiency depends on this parameter (see Figure A.8). From these cross section limits, constraints on the mass of the W' as a function of the mixing angle $\cot\theta_E$ are derived in the same way as described in the previous section for the SSM W'. The cross section limits for the NUGIM model are shown in Figures 10.2, 10.3, 10.4. The mass limit decreases for increasing values of the parameter $\cot\theta_E$. For high values of $\cot\theta_E$ the width of the W' is broad, getting narrower to $\cot\theta_E = 1$, which equals the SSM W'. The

M'_W [TeV]	0.3	0.9	1.1	1.3	1.5
Observed limit [fb]	2000	13	7.3	5.9	5
Expected limit [fb]	1200^{+370}_{-340}	$20^{+9.3}_{-5.1}$	$10^{+4.8}_{-3}$	$6.9^{+2.9}_{-1.9}$	$4.9^{+2.4}_{-1.2}$
M'_W [TeV]	1.6	1.8	2	2.3	2.5
Observed limit [fb]	4.3	3.6	3.5	2.9	2.7
Expected limit [fb]	$3.9^{+1.9}_{-0.95}$	$3.1^{+1.6}_{-0.67}$	$2.8^{+1.4}_{-0.66}$	$2.5^{+1}_{-0.52}$	$2.4^{+1.3}_{-0.51}$
M'_W [TeV]	2.6	2.8	3	3.5	
Observed limit [fb]	2.9	3	3.6	6.7	
Expected limit [fb]	$2.5^{+0.86}_{-0.56}$	$2.8^{+1.2}_{-0.51}$	$3.2^{+1.1}_{-0.69}$	$5.2^{+2}_{-0.99}$	

Table 10.1.: Overview of the cross section limit for the different masses of the SSM W' boson. The table shows the cut-based result.

excluded cross section is lower towards smaller values of $\cot\theta_E$ due to the more pronounced peak at high M_T.

The signal efficiency and cross section times branching ratio of the last two considered parameter points are similar to the SSM W' and therefore the resulting mass limits are similar, too. The limit for $\cot\theta_E = 1$ is $M_{W'} < 2.7$ TeV and falls to $M_{W'} < 2.$ TeV for $\cot\theta_E = 5.6$ at 95% CL.

The resulting constraints from these mass exclusion limits on the parameter space can be seen in Figure 10.5. This search sets significantly better limits than the previous existing constraints from direct and indirect searches for large $\cot\theta_E$ [161, 162, 6] reinterpreted in [62]. These previous constraints are from LHC searches such as $W' \to e\nu$ and $W' \to tb$ which are obtained similar to this analysis from direct searches for a W'. The strongest indirect limits are calculated from the modifications of the SM observations from the mixing with the new group. The decay modes of the muon $\mu^- \to e^+e^-e^-$ and $\mu^- \to \gamma e^-$ would be allowed with a very small branching fraction. Also the CKM matrix would have a non unitary extension.

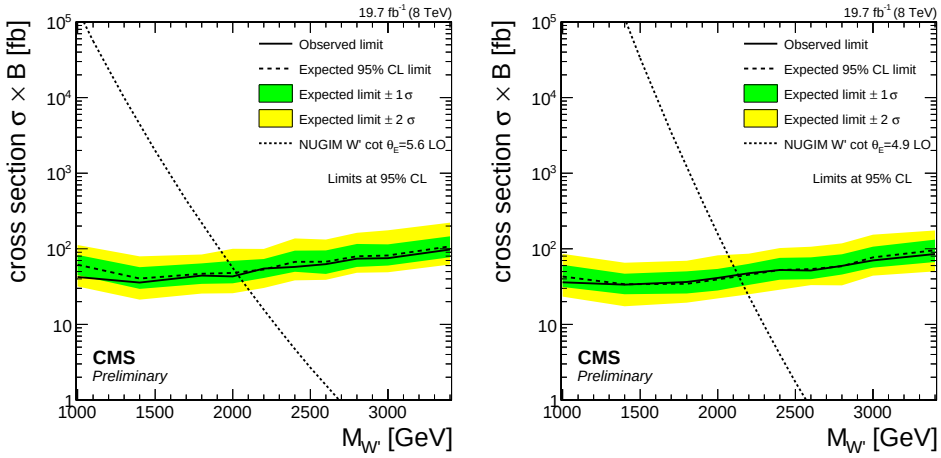


Figure 10.2.: The plots show the limits for $\cot\theta_E = 5.6$ (left) and $\cot\theta_E = 4.9$ (right). For the first parameter point W' masses below 2 TeV are excluded and for the second parameter point W' masses below 2.1 TeV are excluded.

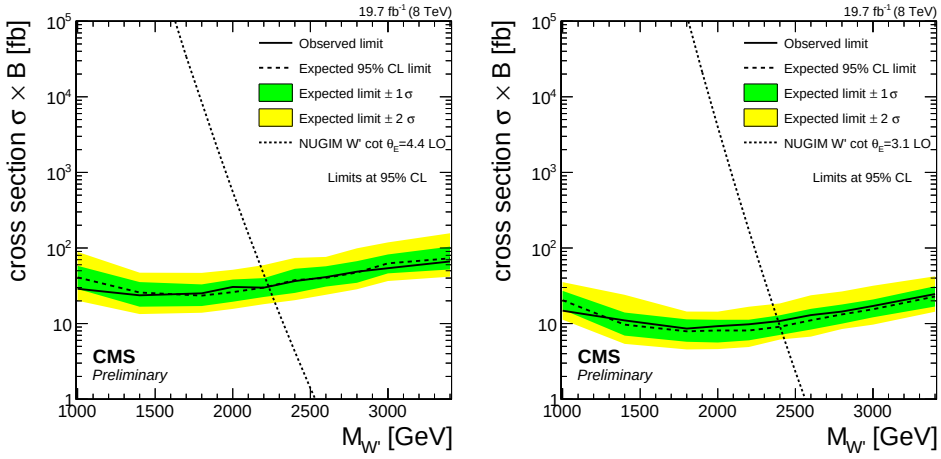


Figure 10.3.: The plots show the limits for $\cot\theta_E = 4.4$ (left) and $\cot\theta_E = 3.1$ (right). A W' mass smaller than 2.2 TeV for the left plot and smaller than 2.4 TeV for the right plot can be excluded at 95% CL.

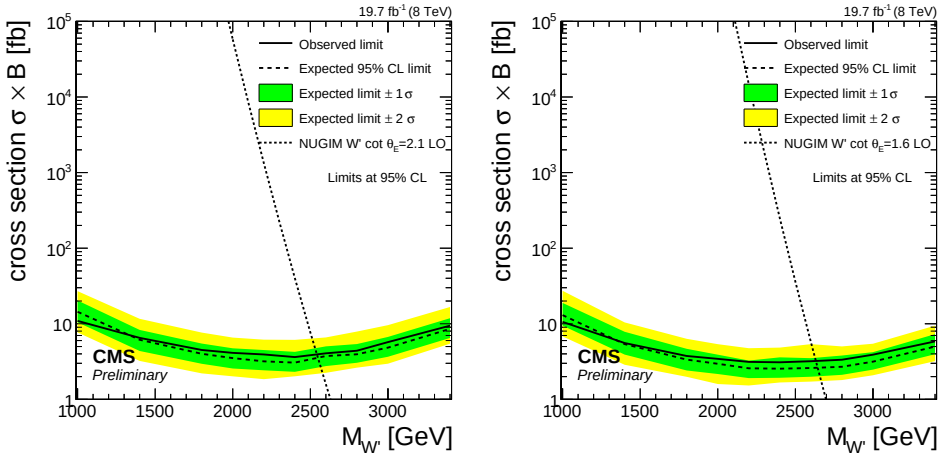


Figure 10.4.: The plots show the limits for $\cot\theta_E = 2.1$ (left) and $\cot\theta_E = 1.6$ (right). For the first parameter, W' masses below 2.5 TeV are excluded and for the second one W' masses below 2.6 TeV are excluded at 95 % CL.

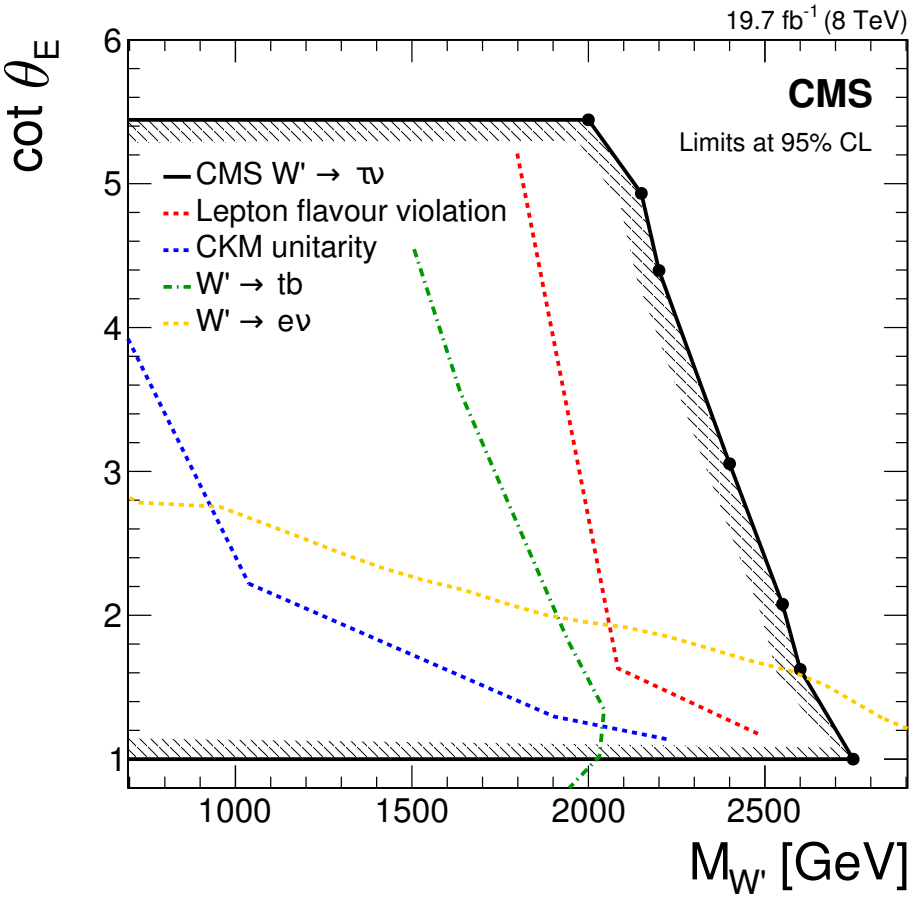


Figure 10.5.: Excluded parameter space of the NUGIM model. The previous limits are taken from [161] while the black line indicates the excluded limit derived from this analysis. Everything inside this line is excluded at 95 % CL.

10.2.3. Model Unspecific Limit

The multibin approach assumes a certain signal shape in M_T. However, new physics processes yielding a tau + E_T^{miss} final state could cause an excess of a different shape. To be independent of models, a single-bin approach compares the number of observed events above a M_T threshold, denoted M_T^{min}, with the SM expectation for this M_T range. The resulting cross section limit as a function of M_T^{min} is shown in Figure 10.6 and Table 10.2. The reconstruction efficiency is estimated to be 42% for W'-boson events satisfying the condition $M_T > M_T^{min}$. It may be noted that the fraction of the signal that satisfies the M_T^{min} requirement depends on the particular model, and is mass-dependent. Since the signal in the detector is only one tau and the loose back-to-back kinematic requirements only ensure that the activity in addition to the tau is not too much, the efficiency for events with $M_T > M_T^{min}$ is very model independent. The systematic uncertainties of a typical W' are uesd in the limit. This allows a reinterpretation in various models by evaluating the signal efficiency, ε_{Signal}, for the M_T^{min} threshold, defined as the number of events in the signal region with $M_T > M_T^{min}$ divided by the total number of generated events: $\varepsilon_{Signal} = N_{M_T > M_T^{min}}/N_{total}$. A new model is excluded if $\sigma_{new\ model}(M_T > M_T^{min}) = \sigma_{total} \cdot \varepsilon_{Signal} > \sigma_{excluded}$.

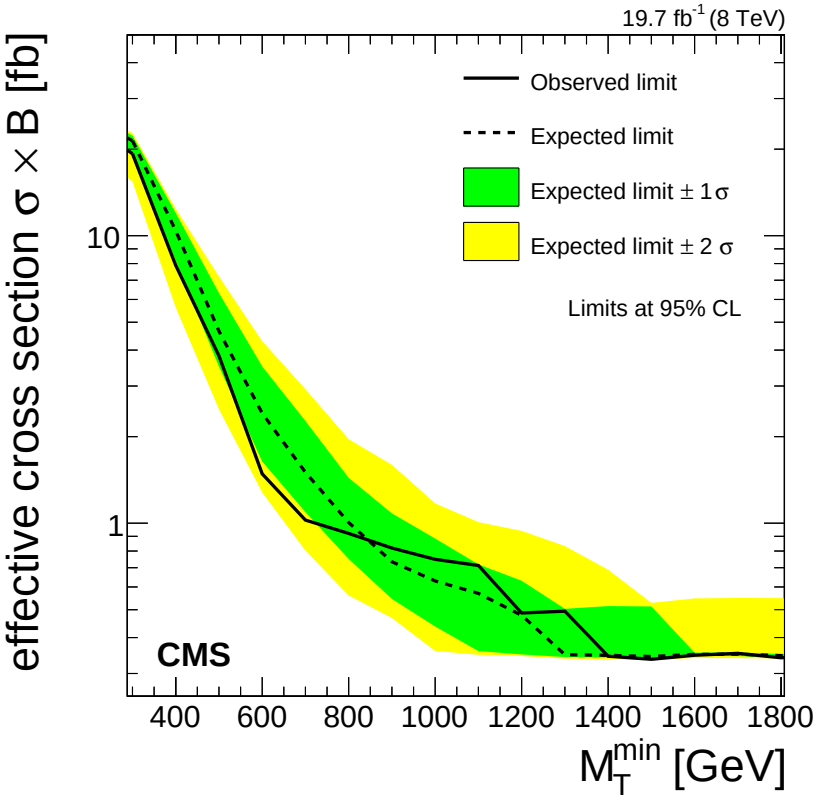


Figure 10.6.: The plot shows the limit for a given fixed efficiency of 42%, for the given M_T^{min} threshold. A signal with an effective cross section in the M_T region above M_T^{min}, that is below the line is excluded at 95 % CL.

M_T [TeV]	0.2	0.3	0.4	0.5	0.6	0.7	0.8
Observed limit [fb]	25	19	7.9	3.8	1.5	1	0.92
Expected limit [fb]	$26^{+1}_{-1.9}$	$21^{+0.89}_{-2.2}$	$10^{+1.6}_{-2.2}$	$4.7^{+1.7}_{-1.2}$	$2.4^{+1.1}_{-0.77}$	$1.5^{+0.76}_{-0.41}$	$1^{+0.43}_{-0.25}$
M_T [TeV]	0.9	1	1.1	1.2	1.3	1.4	
Observed limit [fb]	0.82	0.75	0.71	0.49	0.49	0.34	
Expected limit [fb]	$0.73^{+0.35}_{-0.19}$	$0.63^{+0.25}_{-0.19}$	$0.57^{+0.15}_{-0.21}$	$0.48^{+0.15}_{-0.13}$	$0.35^{+0.15}_{-0.0066}$	$0.35^{+0.17}_{-0.006}$	

Table 10.2.: The Table shows the excluded additional cross section in the $\tau + E_T^{miss}$ final state. The cross section assumes 100% signal above the M_T^{min} threshold and therefore is a limit on $\sigma_{\text{new model}}(M_T > M_T^{min})$.

10.3. Summary of Run I

In summary, the first search for an excess in the transverse mass distribution of the tau+E_T^{miss} channel has been performed. The data sample was collected with the CMS detector in proton-proton collisions at $\sqrt{s} = 8\,\text{TeV}$, and corresponds to an integrated luminosity of $19.7\,\text{fb}^{-1}$. The analysed data contain the events with the highest p_T tau ever recorded at a collider up to now (607 GeV). The reconstruction of the hadronic tau decays has been improved to allow the measurement of these taus.

No significant excess beyond the SM expectation is observed. An SSM W' boson is excluded in the mass range $0.3\,\text{TeV} < M_{W'} < 2.7\,\text{TeV}$ at 95% confidence level. Within the NUGIM the lower limit on the W'-boson mass depends on the coupling constant $\cot\theta_E$ and varies from 2.0 to 2.7 TeV at 95% confidence level. A model unspecific limit was also provided, which allows the reinterpretation of the final state in other models.

10.4. Limits of the 13 TeV Analysis

The limits on the $\tau + E_T^{miss}$ final state of the Run II data will be in a similar fashion as the Run I results. First the SSM W' limits will be calculated, but instead of directly providing the limits on the NUGIM model, a limit is set on a generic coupling parameter $g_{W'}$ as discussed in Section 2.3.1. The coupling parameter changes the width of the W' signal, resulting in a modified M_T distribution. This generic coupling limit can then be used to obtain limits for the NUGIM model. The last Run II result will be model unspecific limit.

The results from Run I and Run II will then be combined and a common limit on the SSM W' mass will be set.

10.4.1. Limits in the Sequential Standard Model

As for the Run I analysis a multibin limit is used to set limits on cross section times branching ratio, depending on the W' mass. The limit is shown in Figure 10.7. In contrast to the Run I limit the signal PDF uncertainty description is shown as a band around the theoretical cross (thin black line with grey band) section, instead of entering the limit calculation. This is done to provide a unified limit approach with the $e/\mu + E_T^{miss}$ analysis in CMS. The NNLO cross section as well as the PDF uncertainties are taken from [163]. All W' masses leading to a theoretical signal cross section higher than the observed cross section limit are excluded. The intersection of the line of the observed limit (thick solid line) and the theoretical prediction indicates the best mass limit on the W' derived from this analysis. The PDF uncertainty on the theoretical signal cross section is small and has a negligible effect on the mass limit. The SSM W' is excluded for W' masses between 0.5 and 3.3 TeV at 95% CL in the tau channel.

As expected from the M_T distribution as shown in Figure 8.16 the observed limit is below the expected limit (dashed line). This means more signal could be excluded than was expected, because less high M_T data events were observed than expected. The largest deviation in the M_T spectrum is about 650 GeV. This translates to the observed limit, where the signal distributions enter, to an offset of all W' masses larger than 700 GeV. For the highest masses the effect of the deficit is mitigated since the deficit is only localized.

An overview of the limits for the different mass points are shown in Table 10.3.

10.4.2. Generalized Couplings

The limit on the cross section depends not only on the position of a potential excess, but can also depend on the width of this excess. In contrast to processes where all final state particles can be recorded (e.g. ll), for the tau+E_T^{miss} signal a change in the width results in a change of the slope of the M_T distribution, as shown in Section 8.2.3. One can evaluate the effect of the width on the resulting limit using reweighted samples. The effect is given with respect to the ratio of the coupling constant $g_{W'}$ to the SM coupling constant g_W, and can be seen in Figure 10.8 for some exemplary W' mass points. One can compare these mass points to one W' mass slice in Figure 10.7. Also my be noted that the cross section limit in Figure 10.7 has a logarithmic cross section scale, while Figure 10.8 has a linear scale.

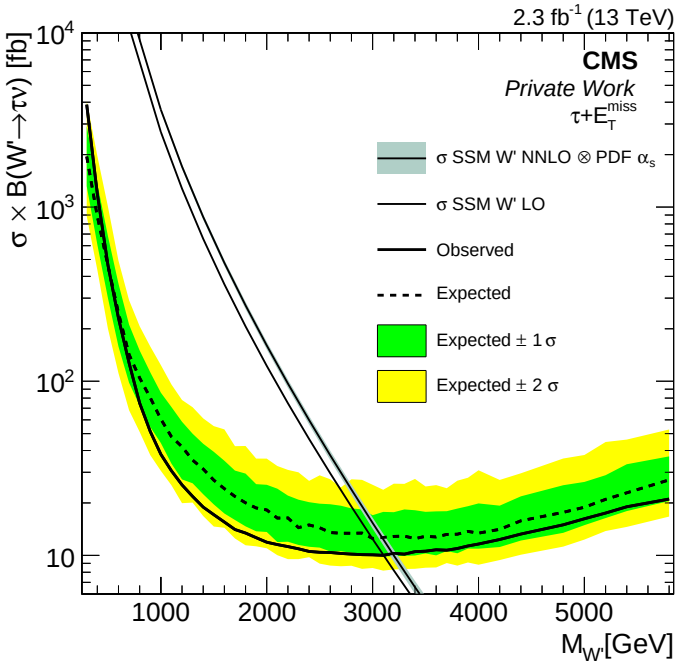


Figure 10.7.: Cross section limit as a function of the mass of the W' boson in the SSM. All W' masses below 3.3 TeV, which corresponds to the intersection of the theoretical NNLO cross section line and the observed cross section limit, in the tau channel are excluded at 95 % CL.

Due to the vast number of signal points that have to be scanned, the limit setting method was reduced from using the full bayesian limit to the use of asymptotic limit, as described in [164]. The resulting limit is worse w.r.t. the bayesian limit by a factor of up to two, as can be seen in Figure 10.7, where the expected cross section limit for a W' with a mass of 3.2 TeV is $12.8^{+4.8}_{-3.0}$ fb and in Figure 10.8 the expected limit for the same mass point is 20^{+15}_{-10} fb. This is expected for low event numbers. Due to the exponential behaviour of most signal cross sections w.r.t. resonance mass, this factor does not change much in the signal parameter space.

In order to evaluate the exclusion limit in a two dimensional plane, the shown limits were parametrized with a second degree polynomial. The theoretical cross section at leading order is used from the generated MadGraph samples and interpolated linearly on a logarithmic scale for $g_W/g_{W'}$. One can also see in Figure 10.8 that for a small width ($g_W/g_{W'} < 1$) no change in the limit is expected, while there is a steep increase in the excluded cross section for high $g_{W'}$ couplings. The excluded cross section as a function of the mass follows the same U-shaped form as the SSM limits, where at high masses the excluded cross section raises, because more signal events are produced off shell at low M_T.

The two dimensional limit as a function of the intersection is shown in Figure 10.9. The result shows that for a large coupling ($g_{W'}/g_W > 4$), a W' boson as heavy as 5.4 TeV can be excluded, while for low couplings of $g_{W'}/g_W < 0.5$ only a W' lighter than 2 TeV can be excluded.

$M_{W'}$ [GeV]	300	500	700	900	1100
Expected limit [fb]	1950^{+960}_{-630}	460^{+220}_{-160}	142^{+69}_{-44}	80^{+32}_{-26}	48^{+23}_{-14}
Observed limit [fb]	3900	470	130	51	31
$M_{W'}$ [GeV]	1200	1400	1600	1800	2000
Expected limit [fb]	42^{+16}_{-15}	$31.3^{+12.4}_{-8.9}$	$24.1^{+10.8}_{-6.8}$	$20.0^{+9.5}_{-5.4}$	$18.2^{+7.2}_{-4.6}$
Observed limit [fb]	25	19	16	13	12
$M_{W'}$ [GeV]	2100	2300	2500	2700	2900
Expected limit [fb]	$16.4^{+7.8}_{-4.4}$	$14.4^{+6.4}_{-3.0}$	$14.4^{+5.4}_{-3.3}$	$13.4^{+5.9}_{-2.8}$	$13.4^{+4.9}_{-3.2}$
Observed limit [fb]	12	11	10	10	10
$M_{W'}$ [GeV]	3000	3200	3400	3600	3800
Expected limit [fb]	$12.5^{+4.5}_{-2.4}$	$12.8^{+4.8}_{-3.0}$	$12.5^{+5.7}_{-2.5}$	$12.8^{+5.1}_{-2.8}$	$13.2^{+5.3}_{-2.9}$
Observed limit [fb]	10	10	11	11	11
$M_{W'}$ [GeV]	3900	4200	4800	5200	5800
Expected limit [fb]	$13.7^{+5.4}_{-2.8}$	$14.1^{+5.2}_{-2.9}$	$17.7^{+7.6}_{-4.2}$	$21.0^{+7.9}_{-4.6}$	$27.1^{+9.9}_{-6.6}$
Observed limit [fb]	11	12	15	17	21

Table 10.3.: Overview of the cross section limit for the different masses of the SSM W' boson.

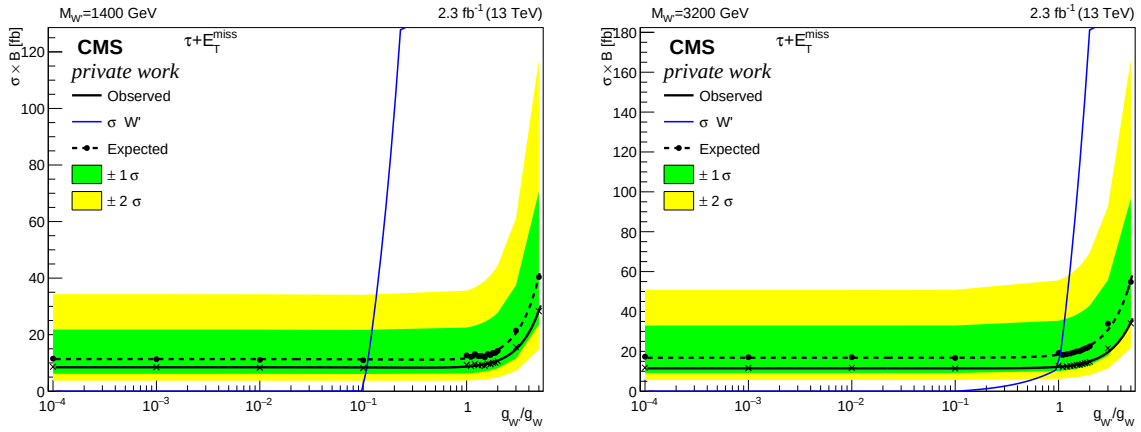


Figure 10.8.: Cross section limits for two exemplary W' masses. The cross section is given as a function of $g_{W'}/g_W$.

10.4.3. NUGIM Limit

The same cross section limits used for the generalized couplings can be used to set limits on the NUGIM model. In order to do this the width of the generalized coupling model has to set in connection with the NUGIM parameter $\cot\theta_E$. The relevant relation is [62]:

$$\Gamma_{W'} = \Gamma^{SSM}_{W'} \times \frac{(4+\frac{1}{4})\cot^2\theta_E + 8\tan^2\theta_E}{12+\frac{1}{4}} = \Gamma^{SSM}_{W'} \cdot \left(\frac{g_{W'}}{g_W}\right)^2. \tag{10.5}$$

Solving this for $\cot\theta_E$ provides the NUGIM parameter for each $\frac{g_{W'}}{g_W}$ value. The functional dependence is plotted in Figure 10.10 (left). The cross section dependence on the width is is $\sigma \times BR \propto \Gamma^{SSM}_{W'}/\Gamma_{W'}$, which in the generalized parameter space means $\sigma_{W'} = \sigma^{SSM}_{W'}/g^2_W$.

The limit is shown in Figure 10.10 (right). The shown parameter space is for $\cot\theta_E \geqslant 1$, because here the tau channel sets the most stringent bounds. The limit is more stringent

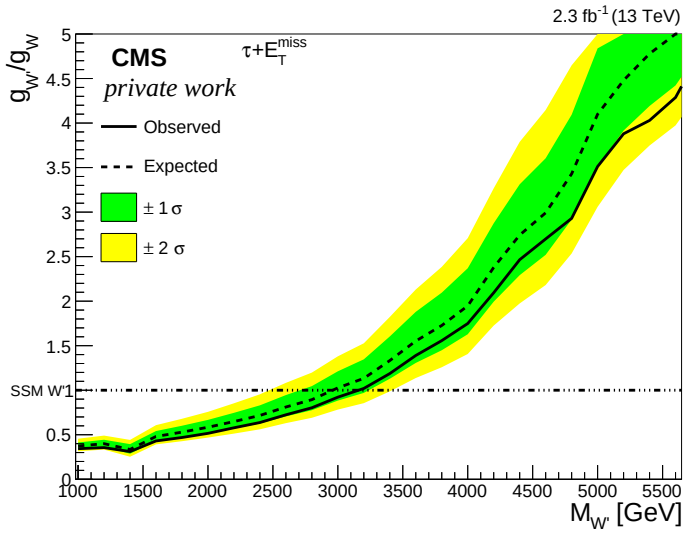


Figure 10.9.: A two dimensional limit of the generalized coupling model is shown. The coupling above the obseved limit, depicted as the solid line, is excluded for the shown W' mass. The used cross section for the W' signal is at leading order.

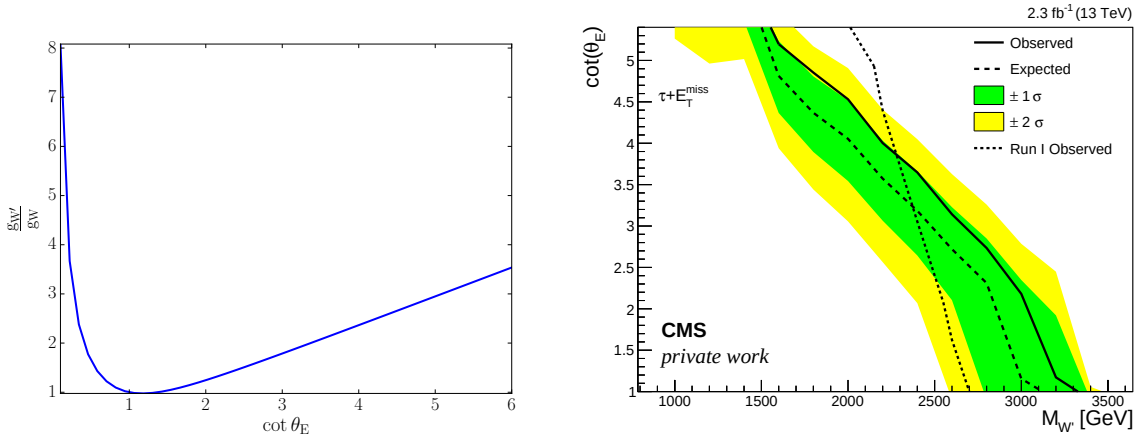


Figure 10.10.: The connection of the generalized coupling constant $g_{W'}$ with the NUGIM parameter $\cot\theta_E$ is shown in the left plot. The limit on $\cot\theta_E$ with respect to the mass is shown on the right. The region left of the solid line is excluded by the data.

for low values of $\cot\theta_E$ than the 8 TeV limit, but for high $\cot\theta_E$ the limit of the 8 TeV data is more stringent due to the higher luminosity and the lower background expectation. This analysis can exclude masses of a new W' boson in the NUGIM model depending on $\cot\theta_E$ from 1.6 TeV to 3.3 TeV. This are the strongest limits set in this model.

10.4.4. Model Unspecific Limit

The multibin approach assumes a certain signal shape in M_T. However, new physics processes yielding a tau+E_T^{miss} final state could cause an excess of a different shape. To be independent

from specific models, a single-bin approach compares the number of observed events above a sliding M_T threshold, denoted M_T^{min}, with the SM expectation for this M_T range. The resulting cross section limit as a function of M_T^{min} is shown in Figure 10.11 and Table 10.4. The reconstruction efficiency is estimated to be about 48% for W' events satisfying the condition $M_T > M_T^{min}$. It may be noted that the fraction of the signal that satisfies the M_T^{min} requirement depends on the particular model, and is mass-dependent. The reconstruction efficiency has an uncertainty corresponding to that of a typical W'-like signal at different M_T^{min} thresholds. For a W' with a mass of 3 TeV the efficiency, including all acceptance and reconstruction factors, is 34% of the total $W' \to \tau\nu$ process, where only events with a hadronic tau decay $\tau \to \tau^{had}\nu$ contribute. For the full tau decay this corresponds to 23%. This mass corresponds to the typical W'-like signal. The difference between 48% the and 34% efficiency is that the trigger turn on is not included in the 48% efficiency since only events with $M_T > 400$ GeV are probed. This allows a reinterpretation in various models by evaluating the signal efficiency, ε_{signal}, for the M_T^{min} threshold, defined as the number of events in the signal region with $M_T > M_T^{min}$ divided by the total number of generated events: $\varepsilon_{signal} = N_{M_T > M_T^{min}}/N_{total}$. If the M_T^{min} threshold is known during generation, it is also possible to use the cross section $\sigma(M_T > M_T^{min})$ directly. The difference between generated and reconstructed M_T is small and should not affect the gained cross section limit more than the provided uncertainties.

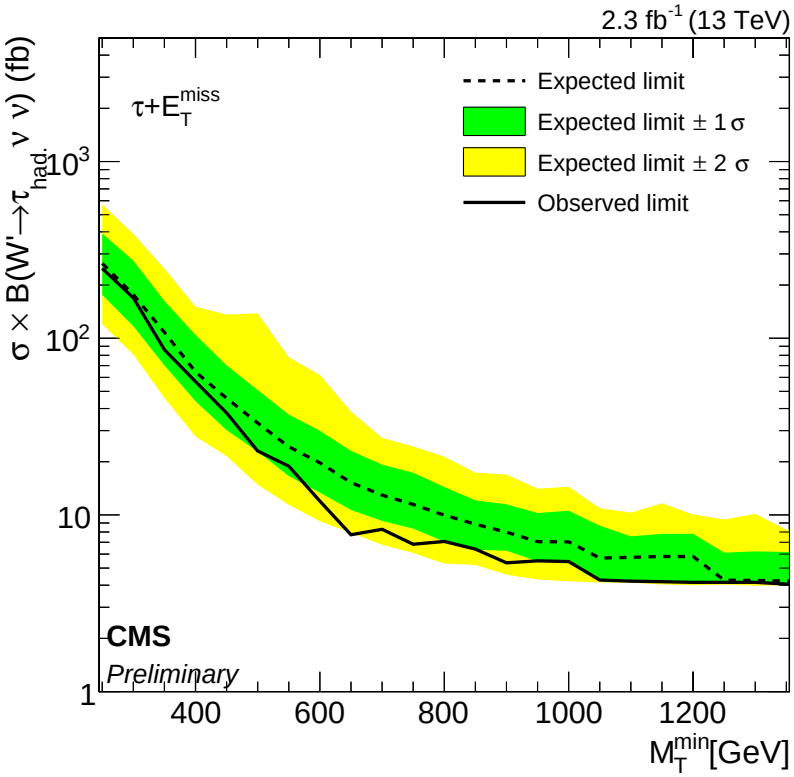


Figure 10.11.: The model unspecific limit shows the limit for a given fixed efficiency of 48%, for the given M_T^{min} threshold. A signal with an effective cross section times its efficiency in the M_T region above M_T^{min}, that is below the line is excluded at 95% CL.

M_T^{min} [GeV]	250	300	350	400	450	500
Expected limit [fb]	265^{+126}_{-89}	176^{+98}_{-59}	108^{+54}_{-38}	64^{+40}_{-20}	46^{+24}_{-16}	33^{+18}_{-10}
Observed limit [fb]	248.2	169.2	86.1	56.8	37.7	23.0
M_T^{min} [GeV]	550	600	650	700	750	800
Expected limit [fb]	$24.3^{+12.7}_{-7.7}$	$19.7^{+10.2}_{-6.3}$	$15.2^{+7.8}_{-4.6}$	$12.9^{+6.3}_{-3.7}$	$11.5^{+5.9}_{-3.1}$	$10.0^{+4.4}_{-2.9}$
Observed limit [fb]	18.9	12.0	7.7	8.3	6.8	7.1
M_T^{min} [GeV]	850	900	950	1000	1050	1100
Expected limit [fb]	$8.9^{+3.2}_{-2.5}$	$8.0^{+3.5}_{-1.7}$	$7.1^{+3.2}_{-1.6}$	$7.0^{+3.5}_{-1.6}$	$5.7^{+3.0}_{-1.4}$	$5.8^{+1.8}_{-1.5}$
Observed limit [fb]	6.4	5.4	5.5	5.4	4.3	4.2

Table 10.4.: The Table shows the excluded additional cross section in the $\tau + E_T^{miss}$ final state. The cross section assumes 100% signal above the M_T^{min} threshold and therefore is a limit on $\sigma_{\text{new model}}(M_T > M_T^{min})$.

10.5. Summary of Run II

A search for an excess in the transverse mass distribution of the $\tau + E_T^{miss}$ channel has been performed. The analysed events are collected with the CMS detector in proton-proton collisions with $\sqrt{s} = 13\,\text{TeV}$, corresponding to $2.3\,\text{fb}^{-1}$. No significant excess beyond the SM could be observed. A SSM W' boson could be excluded for masses between 0.5 TeV and 3.3 TeV with 95% CL. In addition to the SSM, signals with a generalized coupling were also analysed. For a coupling strength $g_{W'}/g_W > 4$ a W' signal with a mass $M_{W'} < 5.4\,\text{TeV}$ can be excluded. This result can also be used to set limits on the Non Universal Gauge Interaction Model. Within the NUGIM the lower limit on the W'-boson mass depends on the coupling constant $\cot\theta_E$ and varies from 1.6 to 3.3 TeV at 95% confidence level. The limit is better w.r.t. the previous 8 TeV limit for $\cot\theta_E$ smaller than 3.8. A model unspecific limit was also provided, which allows the reinterpretation of the final state in other models.

10.6. Combination of the 8 TeV and 13 TeV Results

The 13 TeV results can be combined with the 8 TeV results. But the excluded cross section at 13 TeV being higher, one can exclude more of the signal phase space since the signal cross section also increases. The increase in signal cross section is usually much higher for high W' masses, than the increase of the background cross section. But for low masses the limit obtained from a lower centre of mass energies can be better. In order to set the limit on a meaningful quantity the ratio of the excluded cross section to the expected theoretical signal cross section is used as parameter of interest. This way the different centre of mass energy is taken into account for. The combination can only be done for the same W' mass point. Since slightly different mass points were generated for the 8 TeV search and the 13 TeV search the mass points missing at 13 TeV were reweighted, as described in Section 8.2.3 to provide the missing mass points. The resulting limit is shown in Figure 10.12. One can see that for high masses the difference between the expected limit for the data at 13 TeV (dashed blue line) and the combination with the 8 TeV data is very small, while for the W' masses below 2 TeV the limit from the 8 TeV data (dashed red line) is better than the limit from the 13 TeV data. In the low mass region of $M_{W'} < 500\,\text{GeV}$ the limit is driven by the kinematic threshold of the

analysis, since this is similar at 8 TeV and 13 TeV, the excluded limits are also similar. The limit on the SSM W' mass is at $\sigma_{excluded}/\sigma_{theo.} = 1$ and is 3.3 TeV.

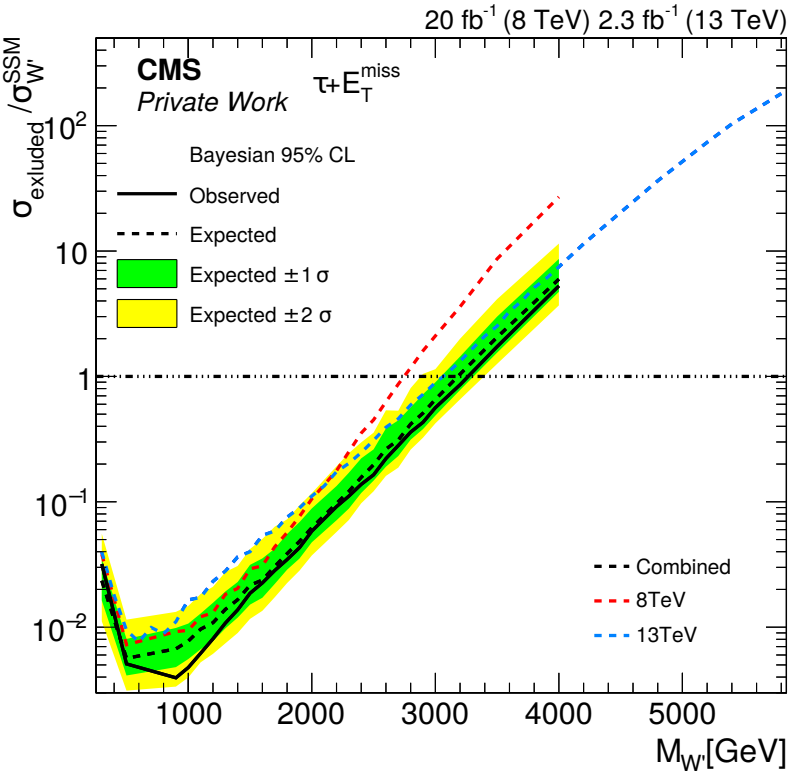


Figure 10.12.: Limit on the mass of a SSM W' using a combination of 8 TeV and 13 TeV results. The red and blue dashed lines show the observed limit from Run I and Run II, respectively. The solid line shows the excluded cross section with respect to the W' cross section.

11. Conclusion of Run I and Run II Tau + E_T^{miss} W' Analysis

The first search for new physics in the tau+E_T^{miss} channel was performed. The analysis provided limits on an additional cross section at the order of 10 fb in the tau+E_T^{miss} channel. The obtained limit on the SSM W' is the first direct limit on the W' mass from the tau final state and only slightly worse than the electron or muon plus E_T^{miss} limit, despite the hadronic final state and the additional neutrino. A comparison to other SSM searches is collected in Table 11.1.

Analysis	obs. mass Limit [TeV]	
	$\sqrt{s} = 8$ TeV	$\sqrt{s} = 13$ TeV
	CMS/ATLAS	CMS/ATLAS
$W' \to \tau\nu$	2.7/-	3.3/-
$W' \to tb$	2.05/1.92	2.38/-
$W' \to e\nu$	3.22/3.08	3.8/4.64
$W' \to \mu\nu$	2.99/2.83	4.0/4.19
$W' \to qq'$	2.2/2.45	2.7/2.9
$Z' \to ee$	2.67/2.79	3.65/3.18
$Z' \to \mu\mu$	2.73/2.53	3.75/2.98
$Z' \to \tau\tau$	-/2.02	2.1/-

Table 11.1.: LHC limits on the SSM from various final states [165, 166, 167, 168, 6, 169, 170, 171, 172, 173, 174, 119, 175, 176, 177, 178, 179].

One can see that while the $W' \to e/\mu + \nu$ final state provides a stronger limit, the $W' \to \tau\nu$ is in the same mass region as the $Z' \to \mu\mu/ee$ results. From the W' results with hadronic final states ($W' \to tb$ and $W' \to qq'$) the $W' \to \tau\nu$ result is by far the strongest limit. The $Z' \to \tau\tau$ limit is also far behind the $W' \to \tau\nu$ limit.

One can also see in the published limits that the tau as a final state is not trivial. While for all other final states ATLAS and CMS calculated limits for $\sqrt{s} = 8$ TeV and $\sqrt{s} = 13$ TeV, the $Z' \to \tau\tau$ was only provided by one of the experiments and for the $W' \to \tau\nu$ no corresponding ATLAS analysis exist yet.

The NUGIM results show that the $W' \to \tau\nu$ search provides the strongest limit for this model. The light lepton and the tb final states are not as sensitive to a new gauge boson as this analysis, despite the higher branching fraction to tb.

The generalized coupling limits show that a relevant signal shape change is only observable for $g_{W'}/g_W > 1$. For lower coupling values the change in the coupling limit is mainly due to the change in the signal cross section.

The result is also provided as a model unspecific cross section limit. This allows the reinterpretation of the limit in all models with this final state.

12. Generalized Model Independent Non Resonant Signal Interpretation

In the past the lepton $+\mathrm{E}_{\mathrm{T}}^{\mathrm{miss}}$ channel was mainly used to investigate limits on resonant signatures such as a W' like signal. In recent years the interest on non resonant signatures became apparent e.g. a DM signal as proposed in [80]. While the experimental signature looks very similar, the region of interest in the M_{T} spectrum is largely different. While a W' resonance is more or less localized - assuming the tail towards low energies is not the first thing to discover - non resonant signals would appear as a slope deviation with respect to the background. For contact interaction and EFT DM signals this was investigate as part of the work for this thesis and published in [6, 57]. Since then the interest in ISR DM searches has grown, but also the models have evolved. The EFT approach was replaced by the simplified models, which were studied in several Bachelor and Master theses for the electron$+\mathrm{E}_{\mathrm{T}}^{\mathrm{miss}}$ channel [9, 8, 10, 11, 12, 13]. The simplified model was in turn discovered to be unitarity violating, as explained in Section 2.4. Taken all the changes on the theory side a model independent interpretation of the lepton $+\mathrm{E}_{\mathrm{T}}^{\mathrm{miss}}$ spectrum would be useful. The main dependence of the obtained limit is the slope of the signal. The proper use of this slope as a model parameter in the limit setting will be shown in this chapter. As the light lepton channel already has three undetected particles it is not sensible to use also the tau channel, where a forth undetected particle would be added. Therefore only the W decays to electron or muon are considered. This is visualized in Figure 12.1. One can see that the events from the decayed tau have about one order of magnitude lower event numbers than the light lepton for high M_{T} (e.g. at 400 GeV 20 events w.r.t. 200 events). The events are not normalized to a specific luminosity here. The decayed tau M_{T} distribution shows an increase in event numbers for low M_{T}, which is not only the kinematic region, where more W-background is expected, but also not triggerd by the CMS detector.

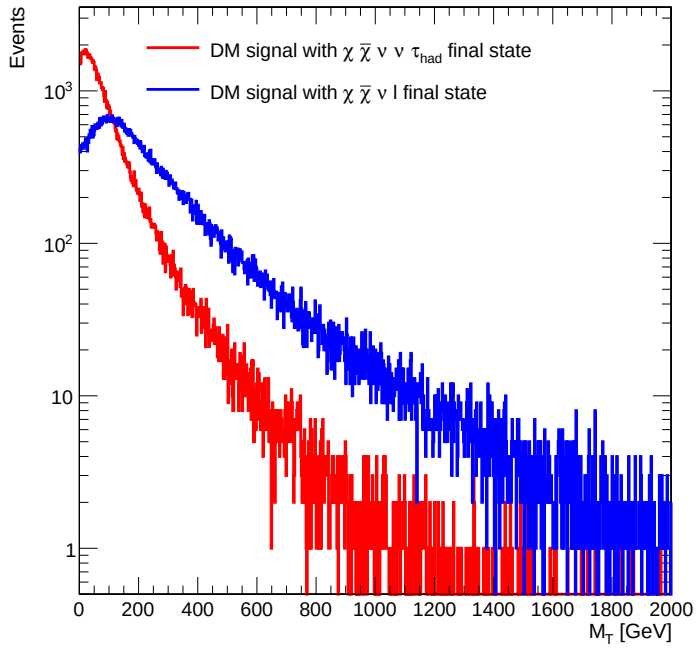


Figure 12.1.: The M_T distribution of a DM signal ($pp \to \chi\chi l\nu$) shows that the tau decay process $pp \to \chi\chi\tau^{had}\nu\nu$ has substantially less events at high M_T than the light lepton final states.

12.1. Analysis Setup

The used analysis data set is the data collected in Run I. The analysis of the Run I data was published in two master theses [9, 8] and in the paper [6]. Since these analyses were performed with different an analysis framework and programs for the electron and the muon channel, the TAPAS framework, as described in Section 6.2, was used with the 8 TeV configuration.

This allowed the lowering of the trigger thresholds, which is most noticeable for the electron channel, where the loosely isolated high p_T trigger has an online threshold of 80 GeV, which can be lowered to 35 GeV with harder isolation criteria. For the muons the non isolated trigger is already at 40 GeV, but this can be lowered to 25 GeV, using the isolated trigger. For both channels this allows the selection of events at the W peak. Since this is not intended to be a complete rework of the 8 TeV analysis, but merely a reinterpretation of the results, the QCD multijet description was not taken from data, but used from the simulation. As shown in Philipp Millet's master thesis [8] the description is acceptable after the final selection. The used samples are the same as for the 8 TeV tau analysis described in Section 7.2. The signal description is based on the $W' \to e/\mu$ and DM EFT samples, details will be described in Section 12.3.

12.2. Event Selection

The event selection follows the W' 8 TeV summary paper [6], where one electron or muon is selected as described in Section 5.1 and 5.2. Also the p_T/E_T^{miss} and $\Delta\phi(l, E_T^{miss})$ selection is used unchanged to the previous papers. The p_T/E_T^{miss} region is $0.4 < p_T/E_T^{miss} < 1.5$ and

the $\Delta\phi(l, E_T^{miss})$ signal region is $\Delta\phi(l, E_T^{miss}) > 2.5$. The distributions of these variables can be seen in Figure 12.2 and 12.3. One can see that the leading order description of the QCD multijet background is not sufficient for high p_T/E_T^{miss}. Especially in the electron channel the QCD multijet is an important background in this region, but since it will have an influence only at low M_T and it is expected to be one order of magnitude below the W background, this discrepancy will be noted as understood leading order effect. This will have no effect on the analysis. One can also see that the simulation can reproduce quantities such as the energy balance and the angle of the E_T^{miss} to the leading lepton sufficiently well, even if most of the simulation is at leading order. At the steeply falling flank of the p_T/E_T^{miss} distribution at low values one can see a slight difference between data and simulation. In this region the events have a p_T at the trigger threshold and very low E_T^{miss}. The simulation of the E_T^{miss} in such events is sensitive to even small detector effects. The description improves if higher E_T^{miss} or a M_T threshold is required. In the $\Delta\phi(l, E_T^{miss})$ distribution as shown in Figure 12.3 one can see that there is a data to simulation disagreement at $\Delta\phi(l, E_T^{miss}) \approx 1$, which is visible in the muon channel, but not visible in the electron channel. This is due to the different trigger threshold of the muon trigger (25 GeV) compared to the electron trigger (35 GeV). The lower threshold also allows for lower values of E_T^{miss}, which are more sensitive to detector effects. This disagreement vanishes for higher p_T thresholds. The description only influences the first bins of the M_T distribution, which are sufficiently well described for this search. For the muon selection it is interesting to note that the Drell-Yan background is much higher than in the electron channel, which is mainly due to the slightly smaller η acceptance, which is $|\eta| < 2.5$ for electrons and $|\eta| < 2.1$ for muons. This reduced acceptance also reduces the effectiveness of the second lepton veto.

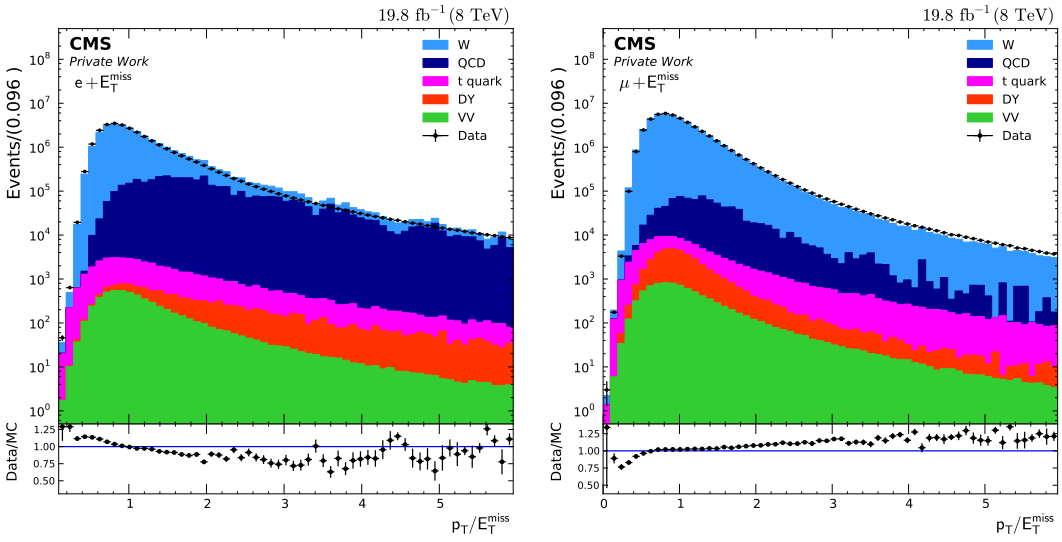


Figure 12.2.: The kinematic distributions for p_T/E_T^{miss} for electrons (left) and muons (right). All selection criteria except p_T/E_T^{miss} are applied.

The p_T and E_T^{miss} plots can be found in the appendix in Figure A.28. The final M_T selection is shown in Figure 12.4. While the electron M_T is binned with a 20 GeV width, the muon channel is binned with a logarithmically increasing bin width, which compensates the muon resolution. No large discrepancies between data and simulation can be seen in the distributions.

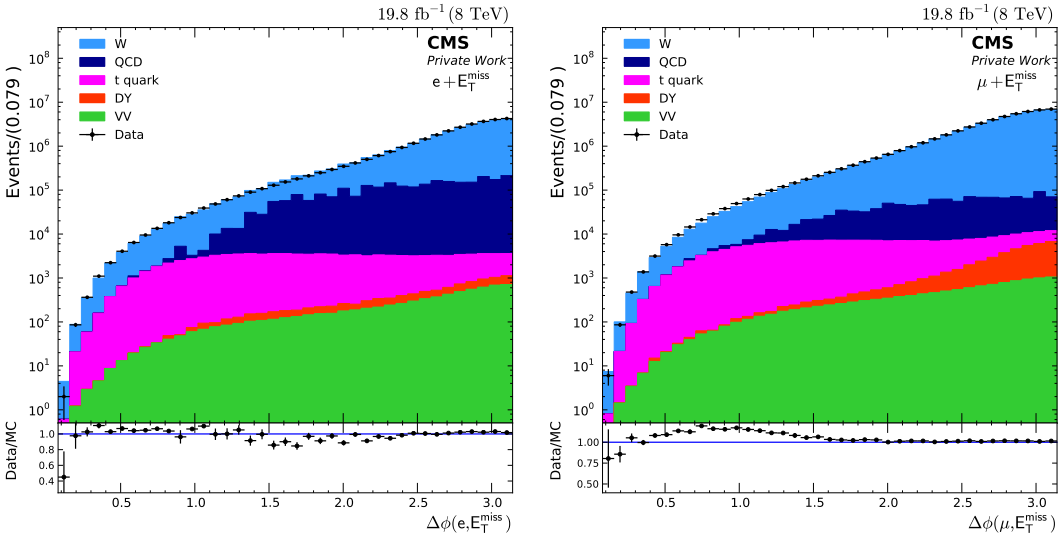


Figure 12.3.: The kinematic distributions for $\Delta\phi(l, E_T^{miss})$ for electrons (left) and muons (right). All selection criteria except $\Delta\phi(l, E_T^{miss})$ are applied.

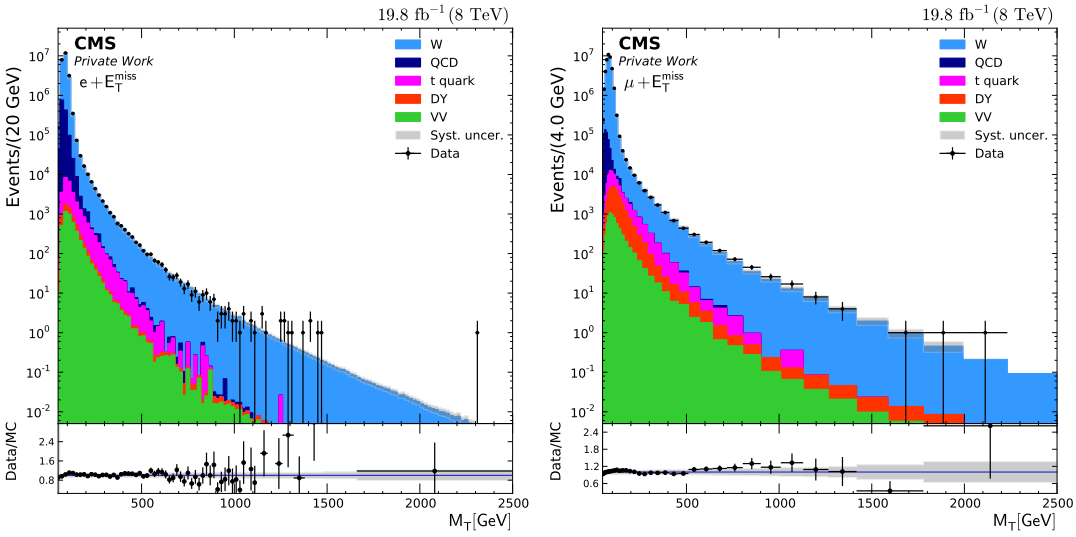


Figure 12.4.: The M_T distribution after all signal selections are applied. On the left for $e + E_T^{miss}$ and on the right for $\mu + E_T^{miss}$.

12.3. Signal Parametrization

The signal events with a non resonant decay in the final state of lepton plus E_T^{miss} are mainly characterized by the energy scale given by M_T. In order to produce a generalized signal distribution, we will investigate the kinematic distributions in more detail.

12.3.1. Signal Shapes with Mono-Lepton Final States

In order to have a handle on the signal shape for non resonant models, the known DM models are investigated. The shape of a non resonant mono-lepton final state is dependent on various model parameters. For the DM EFT model it was shown in the previous Figure 2.9, that a strong shape dependence is due to the interference parameter ξ. This insight also holds for the simplified DM models with a resolved mediator. While for $\xi = +1$, we can see a steeply falling distribution, the distributions for $\xi = 0$ and $\xi = -1$ have a flatter distribution. Other parameters such as vector or axial vector coupling or DM mass do not influence the shape significantly. The only other parameter that has an influence on the shape of the M_T distribution is introduced in the simplified model DM, the mediator mass. This is shown in Figure 12.5.

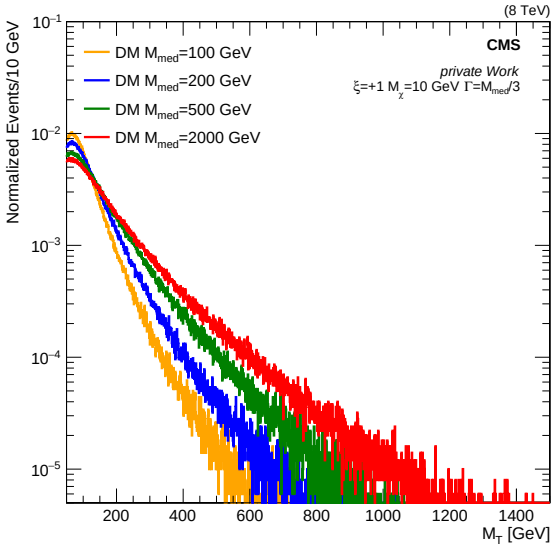


Figure 12.5.: The signal shapes for various mediator masses of the simplified model. The distributions are evaluated with generator information and normalized to their integral.

Since these distributions have no peak structure they can be reduced to a single value, the slope of the distribution. From the DM signal one can also learn more about this slopes extreme values. The distribution for $\xi = -1$ is unitarity violating. A realistic model therefore should not have a much flatter distribution than this.

To use this simple signal shape and reduce it to the only observable parameter the signal M_T

distribution is parametrized with a simple exponential function:

$$f(M_T) = \begin{cases} 0, & \text{for } M_T \leqslant 70\,\text{GeV} \\ A \cdot e^{\,a \cdot M_T}, & \text{for } M_T > 70\,\text{GeV} \end{cases} \tag{12.1}$$

The function is set to zero below 70 GeV to avoid normalization divergences and the normalization is irrelevant, because for all calculations done with the signal only the efficiency is important, not the absolute number. This cutoff is low if compared to the generator distributions of the simplified model, but the turn on will be given by the detector acceptance and the trigger threshold, as can be seen later. The physical observable for these signal shapes is the slope a.

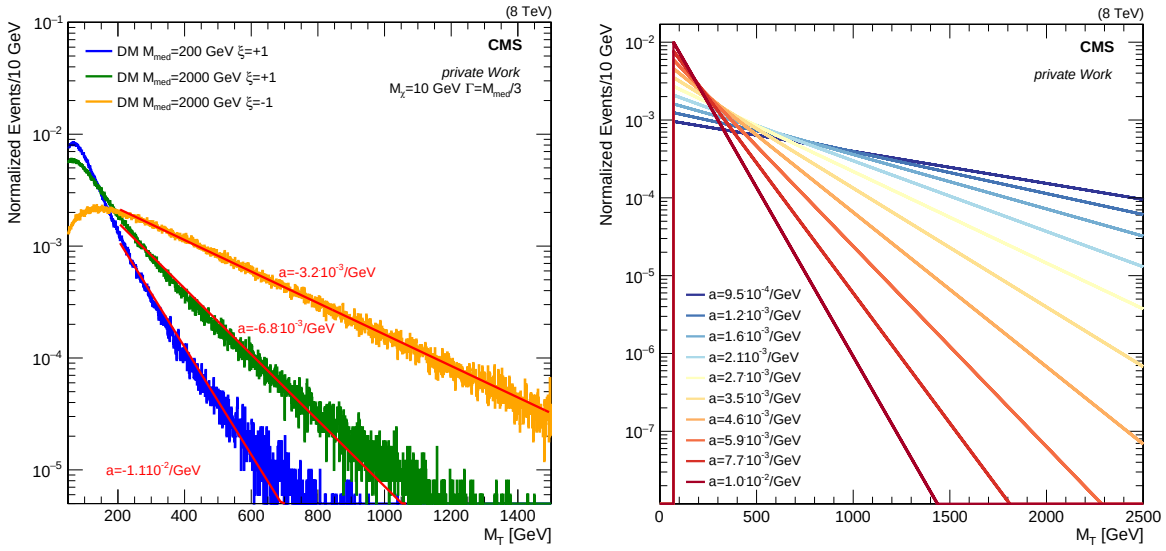


Figure 12.6.: In the left plot the parametrization of the simplified model dark matter with an exponential is shown. The parameter points are chosen such that the maximal and minimal slope can be seen. In the right plot the parametrization is shown for a logarithmic evenly spaced set of slopes.

In Figure 12.6 the maximal and minimal slopes of the DM simplified model are shown. The simple parametrization cannot describe the signal shapes in all detail, but it delivers a good approximation. One can see that the slope range of interest is $(1 \cdot 10^{-2} - 1 \cdot 10^{-4})/\,\text{GeV}$. The peak is not described by the parametrization, but since the low M_T region is dominated by the W-boson peak, this should not be a problem. The low M_T distribution of the background also provides a lower bound for the slope. If the slope is smaller than roughly $1 \cdot 10^{-2}/\,\text{GeV}$ the signal is no longer visible as additional events at hight M_T, but it is completely localized at the W peak. This would appear as a change of the W-boson cross section at the peak and therefore would be out of the scope of this analysis.

12.3.2. Detector Response of a Generic Model

Having a generic model poses the question how the detector response can be simulated. The idea here is to use the detector response of an existing model and calculate event weights that

force the generator level M_T distribution on the described functional shape. The weights can then be applied to the events selected for the signal region. This way all detector effects are simulated by the events and a generic signal shape can be modelled.

As signal input which can be reweighted, there are two kinds of signal available as fully simulated samples, the EFT DM samples and the W' samples. The weights are shown in Figure 12.7 for the EFT DM samples on the left and for the W' samples on the right. One can see that the weights for the DM samples become large for large M_T, which is counter productive for a falling M_T distribution. For the W' signal inputs one can see that the largest weights are in the low M_T region and the Jacobian peaks, at the simulated W' masses, can be seen inversely in the weight. Due to the high weights for the DM sample the W' signal sample is used as input for the reweighting.

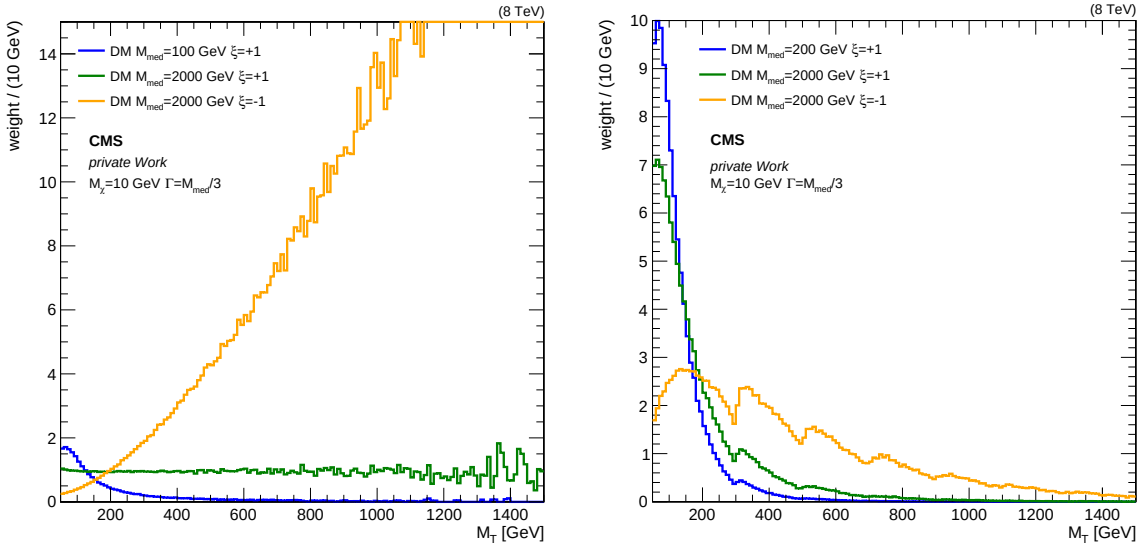


Figure 12.7.: Generator level weights for the DM signal points shown in Figure 12.5 (right). In the left plot the combined EFT DM samples are used as input, while in the right plot the combined W' samples are used as an input. One can see the W' peaks of the available samples in the weight.

In Figure 12.8 (left) the fully reconstructed reweighted M_T can be seen. The EFT DM samples do not have enough events to describe the high M_T region. For the W' reweighted sample the events describe the shape up to 4 TeV. In the right plot of Figure 12.8 the weighted generic samples are compared with the weighted simplified model shapes. One can see that the higher order shapes cannot be described by the simple generic function e.g. a difference at high M_T is visible for $a = 7.1 \cdot 10^{-}3/\,\mathrm{GeV}$, but the general shape can be reproduced.

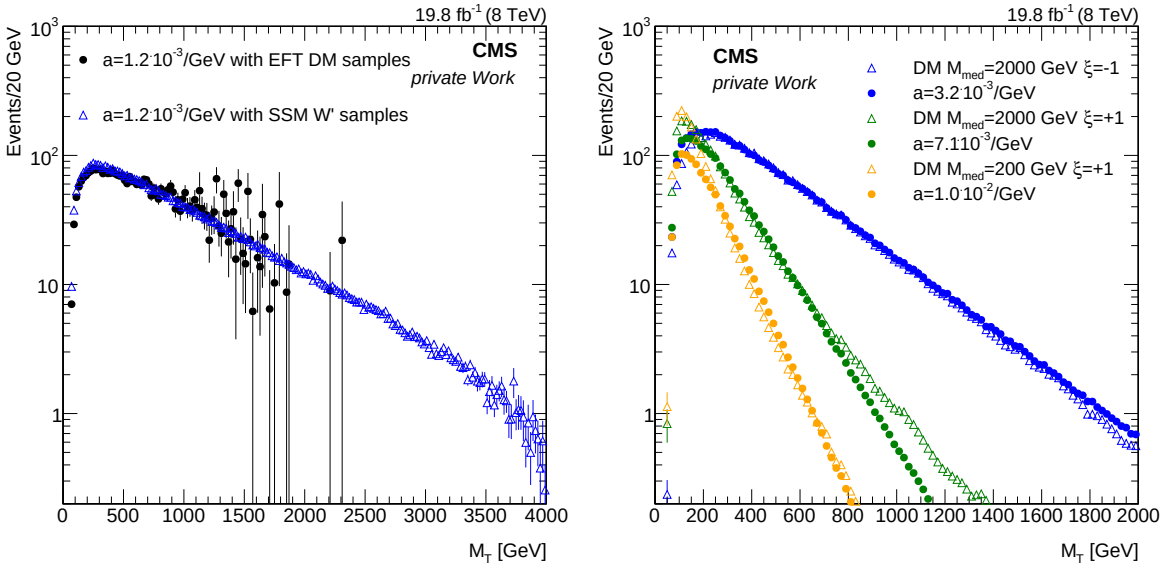


Figure 12.8.: The left plot shows the comparison of the EFT DM reweighted samples w.r.t. the W' reweighted samples. The right plot shows the comparison of the reweighted samples for the parametrization and the simplified model.

12.4. Uncertainties

The systematic uncertainties use the same inputs as in the publication [6]. The same method and numbers are used as described in Section 9.1 for the tau+E_T^{miss} analysis.

For the muon momentum scale a relative uncertainty of $5\% \times p_T/\,\text{TeV}$ is used, whereas the resolution uncertainty is assumed to be 3.2%. Both scale and resolution were measured in the barrel region using cosmic muons [92].

The electron resolution uncertainty is negligible and a scale uncertainty 0.4%(0.8%) is used in the barrel (endcaps) [149].

Compared to the publication the object specific uncertainties have been reevaluated for the extended lower transverse mass distribution and other changes to the analysis. For the W (N)NLO cross section the NNLO in QCD was used and the k-factor is evaluated on the invariant mass of the lepton and neutrino system instead of the transverse mass, which is a better defined quantity on generator level. In addition a 5% scale uncertainty is assumed. For the PDF uncertainty the low $M_T < 200\,\text{GeV}$ region is estimated with 5% uncertainty, because this was not reevaluated. The resulting uncertainties on the background are summarized in Figure 12.9.

For electrons the largest uncertainty at high M_T is the PDF uncertainty, whereas for the muons the momentum scale has a larger uncertainty.

As uncertainty on the luminosity 2.6% is used [150]. For the signal the uncertainties are evaluated with the weighted events. As shown for the $\tau + E_T^{miss}$ analysis the resulting signal uncertainties depend on the signal shape.

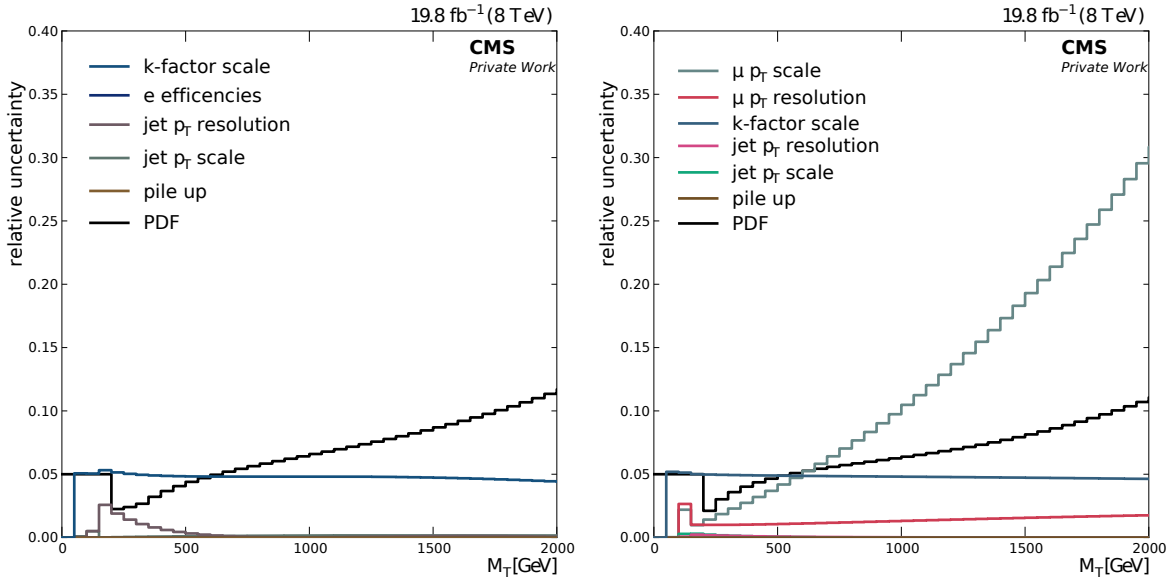


Figure 12.9.: The relative systematic uncertainty on the background event yield as a function of M_T is shown in the left plot for electrons and in the right plot for muons.

12.5. Results

The signal distributions are interpreted using the same multibin approach described in Section 10.2. The cross section limit can be seen in Figure 12.10. One can see a dependence on the slope a. For a steeper signal distribution a less stringent limit can be set. The most stringent limit is obtained for the most shallow signal ($a \approx 10^{-3}/GeV$), but this has to be considered with care. While it is possible to construct a shallow signal this can be ultra violet divergent, which means that at high energies ultimately the probability for an interaction is larger than 1. One has to keep in mind that the PDF suppression of the signal at high M_T is bypassed by the reduction of the signal to a slope a. Nonetheless the distribution could be shallow up to a point, where other new physics appears[1].

Since for many signals the expected signal strength of the e and μ channel is given by the W-boson decays, one can often use a combined limit of both channels. This is shown in Figure 12.11. The expected limit is similar in the e and μ channel. One can see that for signals almost a factor of two in the cross section limit is expected to be gained from the combination.

These limits can be used to interpret a wide range of non resonant models. As an example the simplified model for the DM interpretation is shown in Figure 12.12. The limits split in the electron and muon final states is shown in Figure A.29, in the appendix. It can be compared to the limits set with with the full simplified DM model. The limit is obtained by fitting the DM generator level M_T distribution, obtaining the slope. The limit of the slope value can be obtained by the generalized limit. One can see that the shallow distributions for $\xi = -1$ are described well by this method. The steeply falling distributions for $\xi = +1$ are not as well modelled by the generalized limit, but the limit is within the two sigma expectation of

[1]For resonances we even expect a rise in the cross section up to the peak energy.

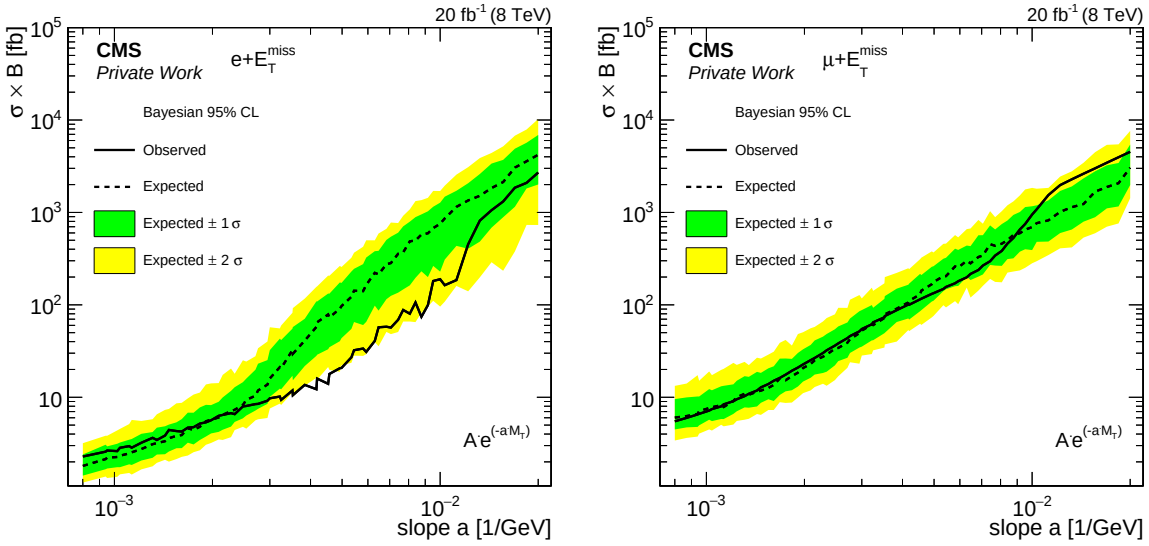


Figure 12.10.: The limit on a generic signal, which is parametrized by the slope of $A \cdot e^{\,a \cdot M_T}$. The signal efficiency is assumed to be similar to the DM models.

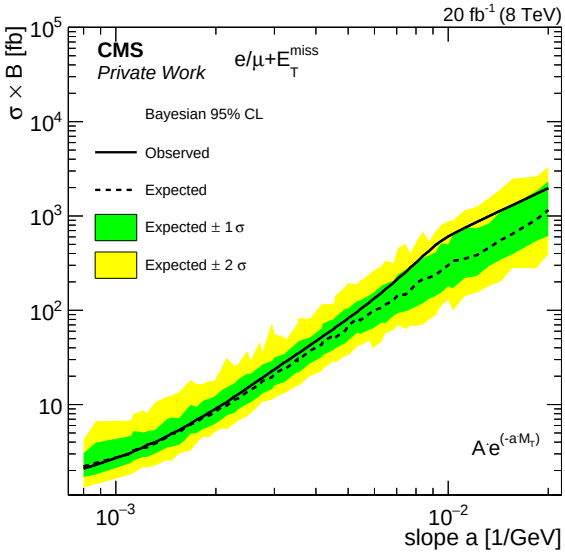


Figure 12.11.: The plot shows the combination of the limits in the electron and muon channel. This can only be used for models, where universality is expected. The limit is set on a generic signal shape, which is parametrized by the slope of $A \cdot e^{\,a \cdot M_T}$. The signal efficiency is assumed to be similar to the DM models.

the expected limit. The obtained limit is the first limit of the simplified model DM for the $\sqrt{s} = 8\,\text{TeV}$ data. The narrow width approach allows to set a limit of 0.5 TeV on the mediator mass (0.4 TeV in μ and 1 TeV in e channel) for the steeply falling $\xi = +1$ signal. For the shallow signal of $\xi = -1$, the limit is 1.8 TeV (1.6 TeV in μ and 2 TeV in e channel). The difference between the electron and muon channel can be seen in Figure 12.10, where the electron limit is below the expectation and the muon limit is at the expectation value or above.

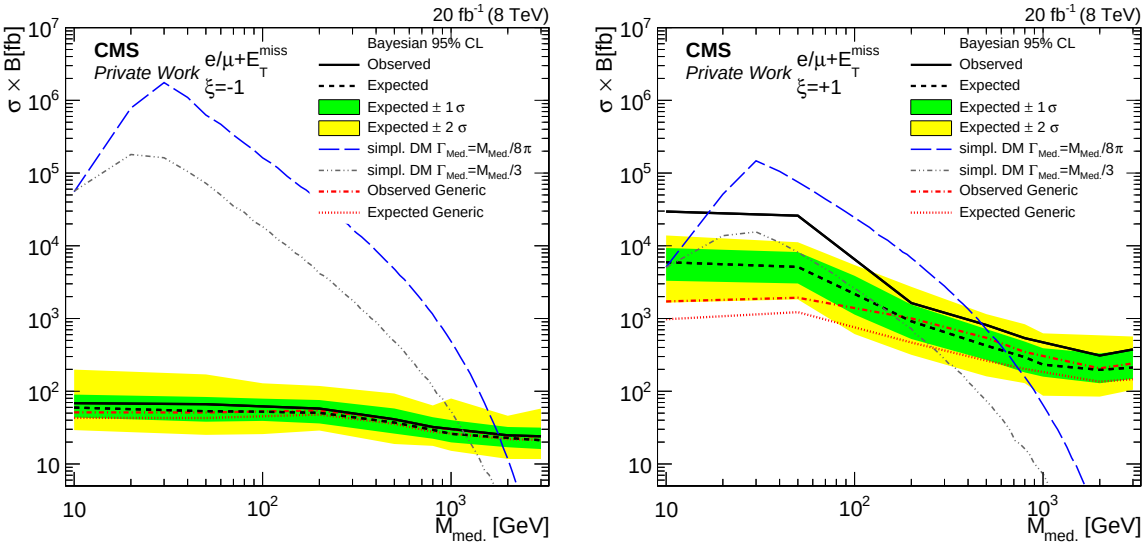


Figure 12.12.: The plots show Dark Matter simplified model limits for the combined e/μ channel. The plots show various ξ parameters $\xi = -1$ (left) and $\xi = 1$ (right). The shown theoretical cross section is plotted for two mediator widths. The narrow width for $\Gamma_{Med.} = \frac{M_{Med.}}{8\pi}$ and the broad width $\Gamma_{Med.} = \frac{M_{Med.}}{3}$.

12.6. Conclusion of the Non Resonant Signal Interpretation

From this test of a generalized non resonant limit in the mono-lepton channel one can conclude that an interpretation in form of a limit on a single parameter is possible. The characterizing parameter of all models where a W is produced in association with invisible particles is the slope of the M_T distributions. This method is better, compared to the single bin limit, used in the model unspecific limit, previously provided in [6] and described in Section 10.2.3, because most of the sensitivity for the limit is gained by the shape of the signal in M_T. This limit can be used for non resonant signals, which have a significant contribution at high M_T.

Using this generalized interpretation the first limit on the simplified model DM at $\sqrt{s} = 8\,\text{TeV}$ can be set. The limit is better than the limit obtained with the $\sqrt{s} = 13\,\text{TeV}$ data in 2015 where the excluded mediator mass is 0.45 TeV [13]. This is due to the higher luminosity of the collected 8 TeV data and the lower background cross sections.

13. Conclusion

To summarize this thesis the fist analysis of the tau+E_T^{miss} final state was performed. The most important premise for this analysis is the good reconstruction of the hadronic tau decays for highly boosted taus. While correct reconstruction of the tau decay modes could not be realized for these highly collimated objects, a good energy reconstruction could be achieved. The decay mode reconstruction is efficient to bring out the difference between hadronic tau jets and QCD jets, allowing a very efficient identification of high p_T taus. The event with the up to date highest reconstructed τ^{had} p_T is observed in this analysis (τ^{had} p_T =608 GeV).

The search for new physics in this final state also requires a good modelling of the expected standard model background. Two approaches were used to attain this. The W-background was modelled with the best theoretical knowledge available. The background from QCD-jets falsely identified as taus is estimated from a control region. This description of the background was found to provide a good estimation with low uncertainties. This description was proven to be sensitive even for non resonant new physics processes. The recoded data show no sign of a new physics, therefore limits on the W' mass can be set. For the sequential standard model W', masses between 0.3 TeV and 3.3 TeV could be excluded. Limits could also be set for a non universal gauge interaction model, where dependent on the mixing parameter of the model $\cot\theta_E$ a limit on the W' mass could be set. The excluded W' masses range from 2 TeV for $\cot\theta_E = 5.5$ to 3.3 TeV for $\cot\theta_E = 1$. These are the strongest limits for this model and the first direct limits in the tau +E_T^{miss} final state. These limits rule out the best fit value from the precision measurements of the standard model, which is $m_{W'} = 2.8$ TeV [61]. However, this best fit estimation was made before the Higgs boson discovery and has a broad minimum. A non universal W' signal may still be in the 3-4 TeV range, but higher values are less favoured by the standard model precision data. The consequences for the standard model precision data is very model dependent and other W' models may have no constrain from these measurements.

In the future of the LHC a total recorded luminosity of 3000 fb^{-1} is planed. For this amount of data a doubling of limit on the W' mass to about 6 TeV is expected at $\sqrt{s} = 13$ TeV. If the centre of mass energy is increased to 14 TeV the expected mass limit is closer to 7 TeV. In contrast to the e/μ+E_T^{miss} channel, the $\tau + E_T^{miss}$ channel gains from the high luminosity, because the signal does not have a peak structure, but is observable as a slope difference in the M_T distribution.

In order to be able to interpret the result in more than one model the model independent limit is calculated. This allows reinterpretation of any signal distribution in the final state of $\tau + E_T^{miss}$. From the ratio of the data to the expectation also a significance can be calculated, which corresponds roughly to the difference of the observed limits to the expected limit in Figures 10.6 and 10.11 in units of σ. This is shown in Figure 13.1. For this plot one has to note the values are correlated, because all events are summed for $M_T > M_T^{min}$. One can see that for the Run I and Run II analysis the most significant deviation of the data from the

standard model expectation is at about 600-750 GeV. The deviation is due to more expected than observed events, which is not expected by any signal in the $\tau + E_T^{miss}$ final state due to the $\tau \to \tau^{had}\nu$ decay. It may be a hint for an incomplete understanding of $W \to \tau\nu$ process, or a statistical fluctuation. Since both deviations are less than 2σ a statistical fluctuation may still be the most likely explanation. A not correctly reconstructed τ^{had} candidate, may also be the cause of this deviation, but this is disfavoured by the Run I result, where the expectation for higher M_T^{min} values agrees with the observed data. This shows that the analysis can also be interpreted as a fist measurement of the $W \to \tau\nu$ distribution for high M_T.

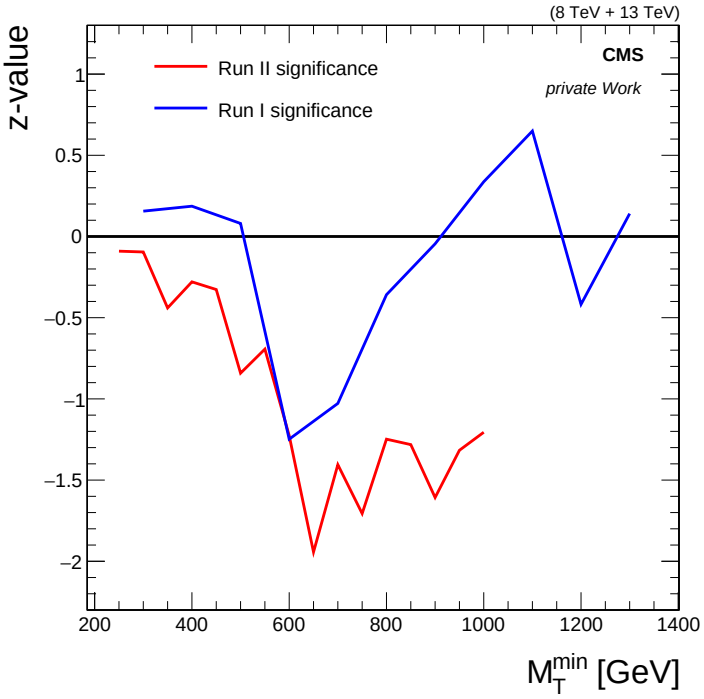


Figure 13.1.: This is a plot of the significance z of the measured data with respect to the expectation. The significance is given as a function of M_T^{min}, where all events with $M_T > M_T^{min}$ are summed. The z value is the difference of the data measurement with respect to the expectation, in units of the standard deviation σ.

The possibility to reinterpret the results in a model independent fashion is also transferred to non resonant excesses in the $e/\mu + E_T^{miss}$ final state. The dynamic development of the dark matter models did not allow the interpretation of the $e/\mu + E_T^{miss}$ final state distributions in terms of a future-proof dark matter model. A limit as a function of the signal slope in the M_T distribution is proven to be comparable with a full description of a dark matter model, and allows a fast reinterpretation of the final state in terms of other models.

In summary a search for signal events were only one of the three final state objects is measurable in the detector as a hadronic jet is possible with previously unaccieved precision and the obtained limits can be the strongest constrains for some new physics models. The standard model still seems to be the best model to describe the $l + E_T^{miss}$ final state, but only about half of the LHC accessible W' mass is excluded so far.

A. Supplementary Material

A.1. Tau ID Performance

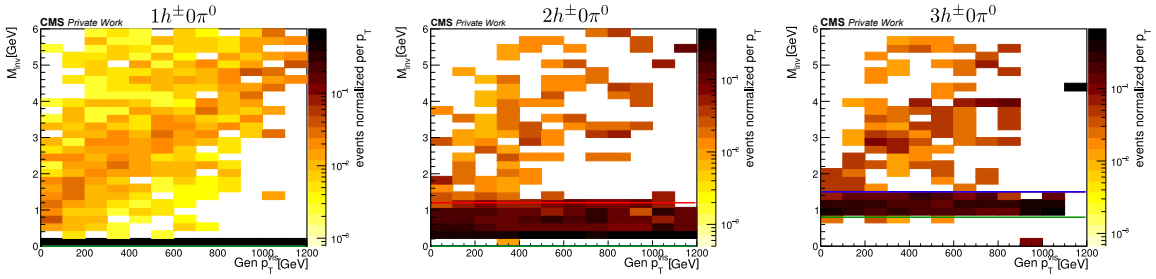


Figure A.1.: The invariant mass of all pfCandidates in a $\Delta R < 0.3$ cone w.r.t. the visible tau p_T at generator level. The event are normalized for each p_T bin. The lines indicate the mass window constrains where green indicates the minimum mass, blue the maximum mass before the high p_T changes, and red the maximum mass after the high p_T changes.

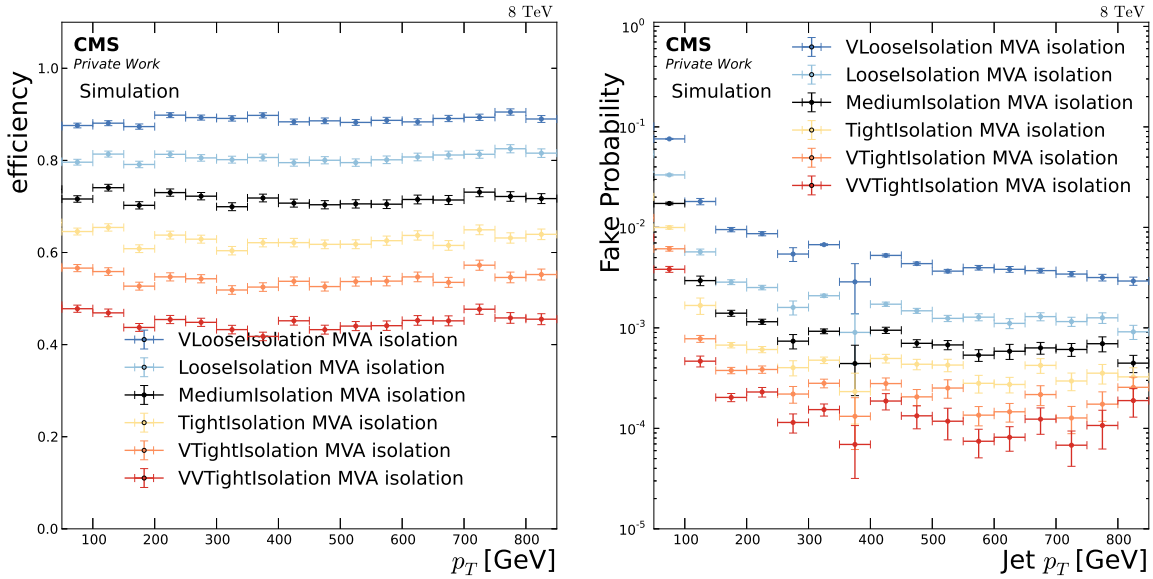


Figure A.2.: The tau identification efficiencies (left) and fake probability (right) for high p_T taus as a function of p_T.

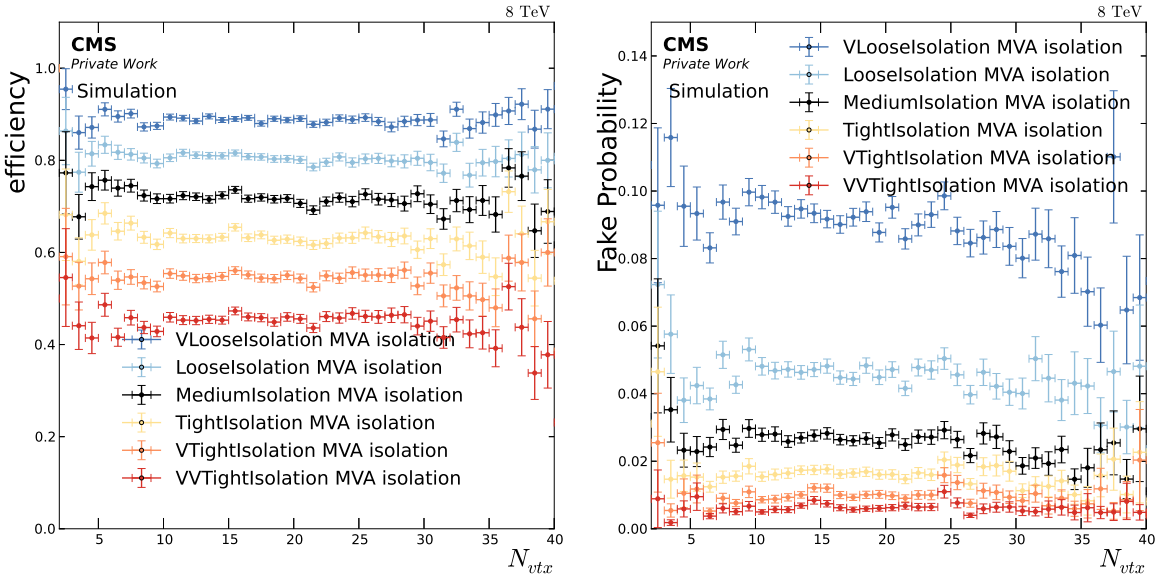


Figure A.3.: The tau identification efficiencies (left) and fake probability (right) for high p_T taus as a function of the number of vertices.

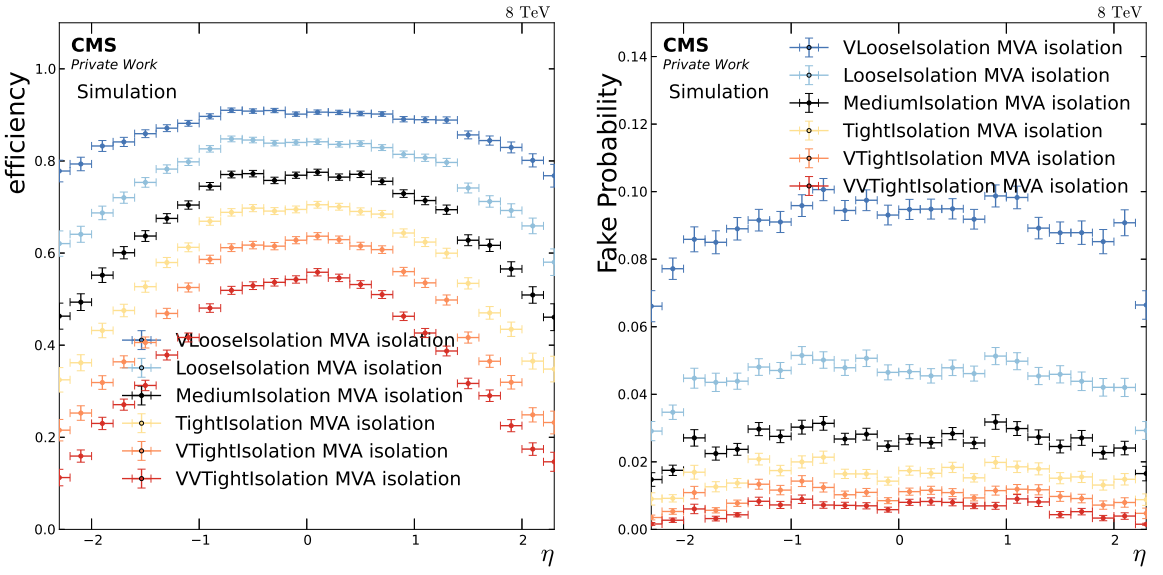


Figure A.4.: The tau identification efficiencies (left) and fake probability (right) for high p_T taus as a function of η.

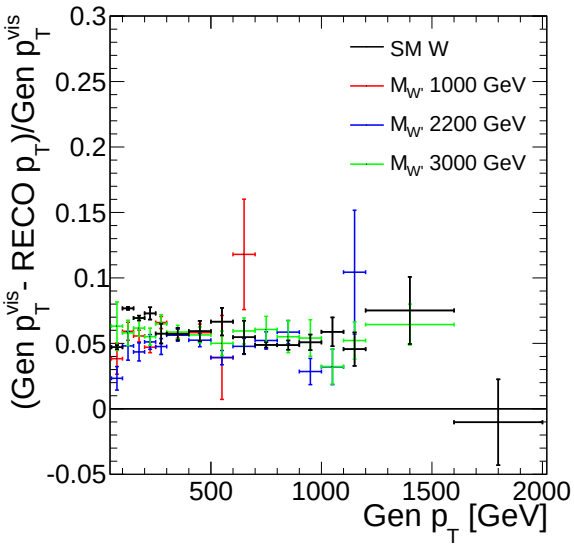


Figure A.5.: Tau response for high p_T. The tau response in simulation is flat w.r.t. p_T.

A.2. Background and Signal Samples for 8 TeV

Process	σ(pb)	Events	PDF set
WW_TuneZ2star_8TeV_pythia6_tauola	57.1 (NLO)	10110431	CTEQ6LL
WWtoAnything_ptmin500_TuneZ2Star_8TeV-pythia6-tauola	0.005235 (NLO)	1000142	CTEQ6LL
WZ_TuneZ2star_8TeV_pythia6_tauola	32.3 (NLO)	9980283	CTEQ6LL
WZtoAnything_ptmin500_TuneZ2Star_8TeV-pythia6-tauola	0.004345 (NLO)	1000035	CTEQ6LL
ZZ_TuneZ2star_8TeV_pythia6_tauola	8.3 (NLO)	9799908	CTEQ6LL
ZZtoAnything_ptmin500_TuneZ2Star_8TeV-pythia6-tauola	0.0017 (NLO)	1000000	CTEQ6LL
DYJetsToLL_PtZ-50To70_TuneZ2star_8TeV-madgraph-tarball	93.8 (NLO)	4930773	CTEQ6LL
DYJetsToLL_PtZ-70To100_TuneZ2star_8TeV-madgraph-tarball	52.31 (NLO)	1413395	CTEQ6LL
DYJetsToLL_PtZ-100_TuneZ2star_8TeV-madgraph	32.9 (NLO)	2662137	CTEQ6LL
DYJetsToLL_PtZ-180_TuneZ2star_8TeV-madgraph-tarball	4.56 (NLO)	1555476	CTEQ6LL
DYToTauTau_M-100to200_TuneZ2Star_8TeV-pythia6-tauola	34.92 (NLO)	200167	CTEQ6LL
DYToTauTau_M-200to400_TuneZ2Star_8TeV-pythia6-tauola	1.181 (NLO)	100083	CTEQ6LL
DYToTauTau_M-400to800_TuneZ2Star_8TeV-pythia6-tauola	0.08699 (NLO)	100455	CTEQ6LL
DYToTauTau_M_1000_TuneZ2star_8TeV_pythia6_tauola	0.00149 (NLO)	50184	CTEQ6LL
ZJetsToNuNu_PtZ-100_8TeV-madgraph	21.4 (LO)	14267721	CTEQ6LL
ZJetsToNuNu_PtZ-70To100_8TeV	32.9 (LO)	22331254	CTEQ6LL
TT_CT10_TuneZ2star_8TeV-powheg-tauola	225.197 (NNLO)	21675970	CTEQ10
Tbar_s-channel_TuneZ2star_8TeV-powheg-tauola	1.76 (NNLL)	139974	CTEQ10
Tbar_t-channel_TuneZ2star_8TeV-powheg-tauola	30.7 (NNLL)	1935072	CTEQ10
Tbar_tW-channel-DR_TuneZ2star_8TeV-powheg-tauola	11.1 (NNLL)	493460	CTEQ10
T_s-channel_TuneZ2star_8TeV-powheg-tauola	3.79 (NNLL)	259961	CTEQ10
T_t-channel_TuneZ2star_8TeV-powheg-tauola	56.4 (NNLL)	3748227	CTEQ10
T_tW-channel-DR_TuneZ2star_8TeV-powheg-tauola	11.1 (NNLL)	497658	CTEQ10
WJetsToLNu_TuneZ2Star_8TeV-madgraph-tarball	36703.2 (NLO)	58554908	CTEQ6LL
WJetsToLNu_PtW-100_TuneZ2star_8TeV-madgraph	228.9 (NLO)	12742382	CTEQ6LL
WJetsToLNu_PtW-180_TuneZ2star_8TeV-madgraph-tarball	23.5 (NLO)	9739464	CTEQ6LL
WJetsToLNu_PtW-50To70_TuneZ2star_8TeV-madgraph	811.2 (NLO)	48426604	CTEQ6LL
WJetsToLNu_PtW-70To100_TuneZ2star_8TeV-madgraph	428.9 (NLO)	3248966	CTEQ6LL
WToENu_ptmin100_ptmax500_TuneZ2Star_8TeV-pythia6	1.457 (NLO)	1000206	CTEQ6LL
WToENu_ptmin500_TuneZ2Star_8TeV-pythia6	0.001525 (NLO)	1000366	CTEQ6LL
WToMuNu_ptmin100_ptmax500_TuneZ2Star_8TeV-pythia6	1.457 (NLO)	1000206	CTEQ6LL
WToMuNu_ptmin500_TuneZ2Star_8TeV-pythia6	0.001525 (NLO)	1000366	CTEQ6LL
WToTauNu_ptmin100_ptmax500_TuneZ2Star_8TeV-pythia6-tauola	1.457 (NLO)	1000206	CTEQ6LL
WToTauNu_ptmin500_TuneZ2Star_8TeV-pythia6-tauola	0.001525 (NLO)	1000366	CTEQ6LL

Table A.1.: Summary of simulated background samples

	SSM		NUGIM					
			$\sin^2\theta_E$					
			0.031	0.04	0.05	0.1	0.2	0.3
$M_{W'}$ (in GeV)	σ_{LO} (pb)	σ_{NNLO} (pb)	σ_{LO} (pb)					
300	113.5	153.2	-	-	-	-	-	-
500	16.48	22.46	-	-	-	-	-	-
700	4.28	5.782	-	-	-	-	-	-
900	1.471	1.981	-	-	-	-	-	-
1000	-	-	2.647	1.833	1.342	0.6435	0.6071	0.7494
1100	0.5881	0.7828	-	-	-	-	-	-
1300	0.2588	0.3408	-	-	-	-	-	-
1400	-	-	0.5810	0.3858	0.2715	0.1181	0.1131	0.142
1500	0.1193	0.1543	-	-	-	-	-	-
1700	0.05781	0.0727	-	-	-	-	-	-
1800	-	-	0.1932	0.1249	0.0854	0.0326	0.0279	0.0342
1900	0.02958	0.03638	-	-	-	-	-	-
2000	0.02123	0.02577	0.1231	0.0789	0.0534	0.0191	0.0148	0.0177
2100	0.01547	0.01855	-	-	-	-	-	-
2200	0.01127	0.01346	0.0822	0.0524	0.0351	0.0119	0.0082	0.0094
2300	0.00839	0.00983	-	-	-	-	-	-
2400	0.00622	0.00724	0.0571	0.0362	0.0241	0.0077	0.0048	0.0052
2500	0.00473	0.00539	-	-	-	-	-	-
2600	0.00357	0.00412	0.0410	0.0259	0.0172	0.0053	0.0029	0.0030
2700	0.00269	0.003104	-	-	-	-	-	-
2800	0.00210	0.00241	0.0301	0.0190	0.0126	0.0037	0.0019	0.0018
2900	0.00165	0.00190	-	-	-	-	-	-
3000	0.00132	0.00152	0.0227	0.0143	0.0094	0.0027	0.0012	0.0011
3100	0.00106	0.00124	-	-	-	-	-	-
3200	0.00087	0.0010	-	-	-	-	-	-
3300	0.00071	0.00085	-	-	-	-	-	-
3400	0.00060	0.00073	0.0136	0.0085	0.0056	0.0016	0.0006	0.0016
3500	0.00051	0.00063	-	-	-	-	-	-
3700	0.00037	0.00047	-	-	-	-	-	-
4000	0.00025	0.00033	-	-	-	-	-	-

Table A.2.: Signal Monte Carlo samples generated in PYTHIA for SSM W' and in MADGRAPH for the NUGIM W'. Cross sections in NNLO are only calculated for the SSM W' samples [144].

	a	b	c	χ^2/N_{dof}
calorimeter E_T^{miss}				
simulation (binomial)	1 ± 1.34e-05	159.4 ± 0.1	45.2 ± 0.2	6.5
simulation (gaussian)	0.988 ± 1.34e-05	158.3 ± 0.1	43.6 ± 0.2	1.6
Data	0.996 ± 1.34e-05	163.3 ± 0.1	52.8 ± 0.2	0.6
particle flow E_T^{miss}				
simulation (binomial)	1 ± 1.37e-05	169.1 ± 0.1	29.5 ± 0.1	11.6
simulation (gaussian)	0.986 ± 1.37e-05	168.1 ± 0.1	28.1 ± 0.1	1.2
Data	0.981 ± 1.37e-05	171.9 ± 0.1	31.7 ± 0.1	0.7

Table A.3.: Results from the trigger fits. For the simulation different errors were used.

A.3. Additional Kinematic Distributions at 8 TeV

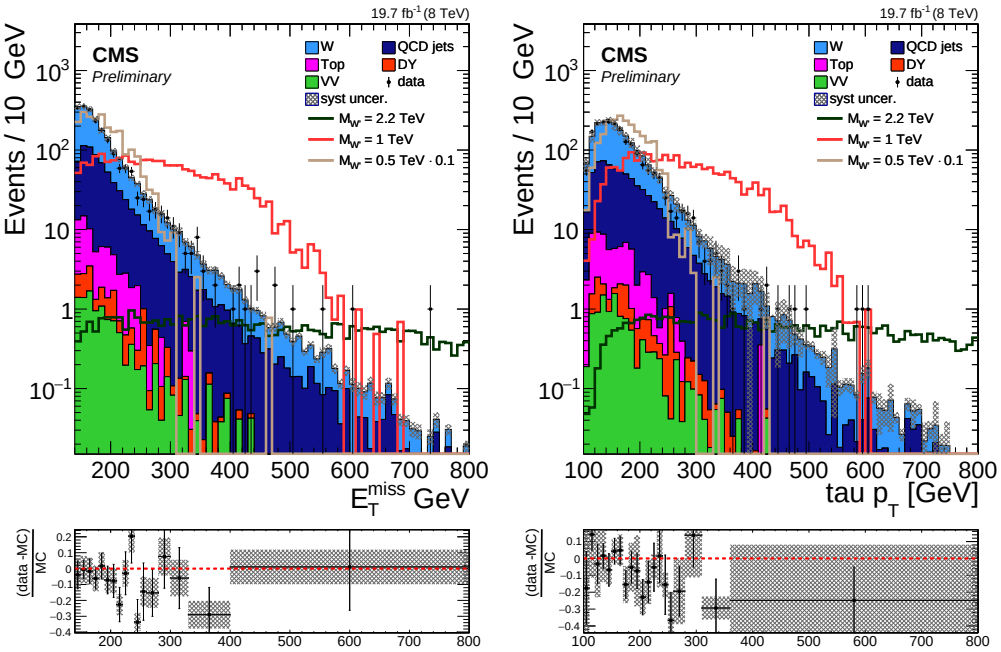


Figure A.6.: The E_T^{miss} and tau p_T distributions for the cut-based isolation. For the final event selection.

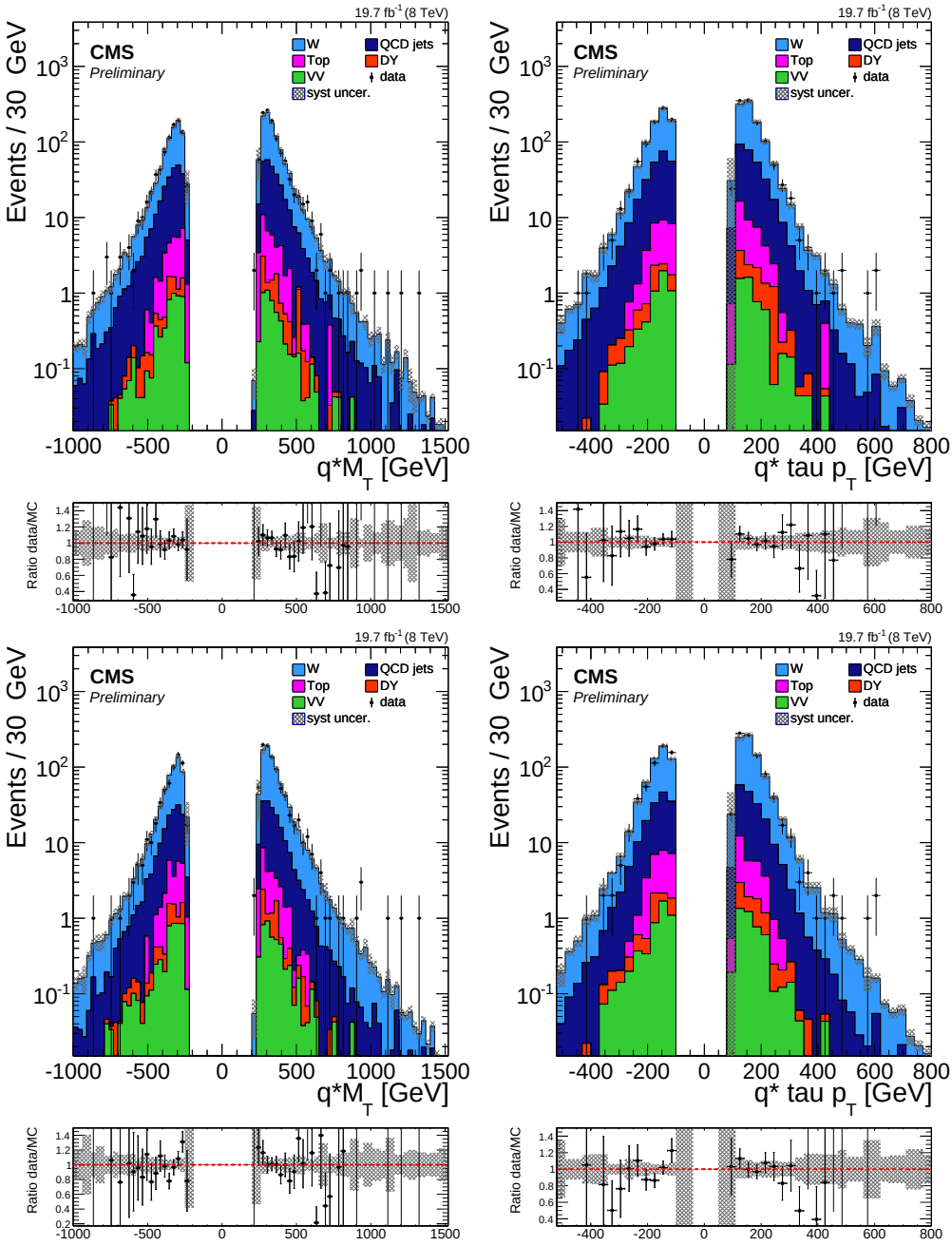


Figure A.7.: The M_T and tau p_T distributions times the charge of the tau for the cut-based isolation (top) and MVA isolation (bottom). For the final event selection.

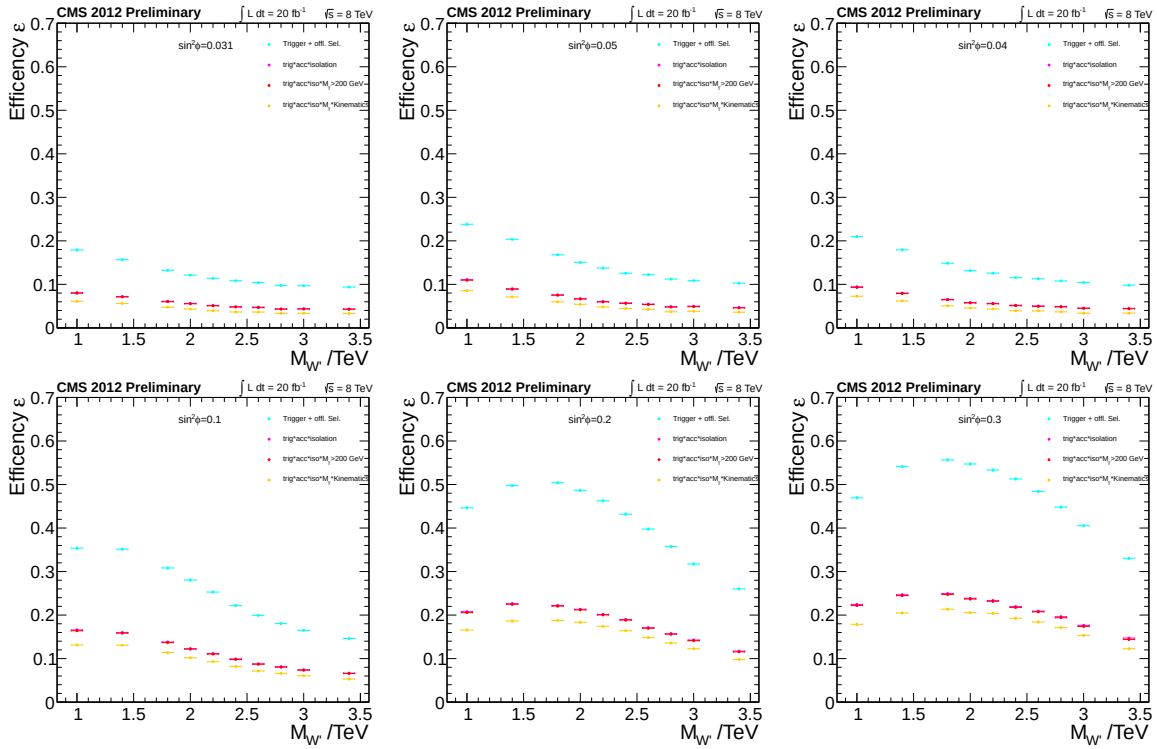


Figure A.8.: The absolute NUGIM signal efficiency for various stages in the analysis. Each plot shows another $\sin^2\theta_E$. The efficiencies reflect the changing signal kinematics.

A.4. Background and Signal Samples for 13 TeV

Process	Order (generator)	σ(pb)	Events
WWTo4Q_13TeV-powheg	NLO	54	2000000
WWToLNuQQ_13TeV-powheg	NLO	53	8573600
WWTo2L2Nu_13TeV-powheg	NLO	12	1979988
WZTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8	NLO	11	14292852
WZTo1L3Nu_13TeV_amcatnloFXFX_madspin_pythia8	NLO	3	1232223
WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	NLO	5.5	25704656
WZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	NLO	0.056	249800
ZZTo2L2Nu_13TeV_powheg_pythia8	NLO	0.75	8556400
ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	NLO	3.4	15029272
ZZTo2Q2Nu_13TeV_amcatnloFXFX_madspin_pythia8	NLO	5	30782664
ZZTo4L_13TeV-amcatnloFXFX-pythia8	NLO	14	10747136
ZZTo4Q_13TeV_amcatnloFXFX_madspin_pythia8	NLO	8.3	30684331
DYJetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	NLO	6000	1719818
DYJetsToLL_M-100to200_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	NLO	230	1054878
DYJetsToLL_M-200to400_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	NLO	7.7	10020
DYJetsToLL_M-400to500_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	NLO	0.42	10745
DYJetsToLL_M-500to700_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	NLO	0.24	10077
DYJetsToLL_M-700to800_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	NLO	0.035	10208
DYJetsToLL_M-800to1000_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	NLO	0.03	9916
DYJetsToLL_M-1000to1500_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	NLO	0.016	9432
DYJetsToLL_M-1500to2000_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	NLO	0.002	9692
DYJetsToLL_M-2000to3000_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	NLO	$5.4 \cdot 10^{-4}$	9061
DYJetsToLL_M-50_HT-100to200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	LO	220	2605822
DYJetsToLL_M-50_HT-200to400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	LO	61	962195
DYJetsToLL_M-50_HT-400to600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	LO	8.2	1069003
DYJetsToLL_M-50_HT-600toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	LO	2.5	1031103
TT_TuneCUETP8M1_13TeV-powheg	NLO	820	284950787
TT_Mtt-700to1000_TuneCUETP8M1_13TeV-powheg	NLO	75	40371857
TT_Mtt-1000toInf_TuneCUETP8M1_13TeV-powheg	NLO	20	26964427
TTTT_TuneCUETP8M1_13TeV-amcatnlo-pythia8	NLO	0.009	1210521
TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	NLO	0.2	250307
TTWJetsToQQ_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	NLO	0.41	831847
ST_t-channel_antitop_4f_leptonDecays_13TeV-powheg-pythia8_TuneCUETP8M1	NLO	81	1630900
ST_t-channel_top_4f_leptonDecays_13TeV-powheg-pythia8_TuneCUETP8M1	NLO	220	3299200
ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	NLO	38	999400
ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	NLO	38	1000000
WJetsToLNu_HT-100To200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	LO	2000	10168757
WJetsToLNu_HT-200To400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	LO	520	4949568
WJetsToLNu_HT-400To600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	LO	75	1943664
WJetsToLNu_HT-600ToInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	LO	1.90E+01	1041358
WJetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	LO	$6.2 \cdot 10^{4}$	47161328
WToENu_M-3000_TuneCUETP8M1_13TeV-pythia8	LO	$2.9 \cdot 10^{-5}$	995350
WToENu_M-1000_TuneCUETP8M1_13TeV-pythia8	LO	1.30E-02	998728
WToENu_M-200_TuneCUETP8M1_13TeV-pythia8	LO	6.20E+00	998887
WToENu_M-500_TuneCUETP8M1_13TeV-pythia8	LO	2.10E-01	951250
WToMuNu_M-200_TuneCUETP8M1_13TeV-pythia8	LO	6.20E+00	993140
WToMuNu_M-500_TuneCUETP8M1_13TeV-pythia8	LO	2.10E-01	997511
WToMuNu_M-1000_TuneCUETP8M1_13TeV-pythia8	LO	1.30E-02	992260
WToMuNu_M-3000_TuneCUETP8M1_13TeV-pythia8	LO	$2.9 \cdot 10^{-5}$	996545
WToTauNu_M-1000_TuneCUETP8M1_13TeV-pythia8	LO	1.30E-02	999316
WToTauNu_M-200_TuneCUETP8M1_13TeV-pythia8	LO	6.20E+00	991772
WToTauNu_M-3000_TuneCUETP8M1_13TeV-pythia8	LO	$2.9 \cdot 10^{-5}$	994288
WToTauNu_M-500_TuneCUETP8M1_13TeV-pythia8	LO	0.21	950080

Table A.4.: Summary of simulated background samples

$M_{W'}$ (in GeV)	σ_{LO} (fb)	σ_{NNLO} (fb)
1000.0	2698.95	3623.49
1200.0	1275.76	1708.94
1400.0	657.281	877.234
1600.0	360.152	478.175
1800.0	206.473	272.375
2000.0	122.55	160.428
2200.0	74.726	96.957
2400.0	46.5844	59.8134
2600.0	29.5602	37.5096
2800.0	19.0347	23.8523
3000.0	12.4122	15.348
3200.0	8.19516	9.98443
3400.0	5.46747	6.56821
3600.0	3.6865	4.37084
3800.0	2.51575	2.9432
4000.0	1.73367	2.00923
4200.0	1.2108	1.39263
4400.0	0.856493	0.981875
4600.0	0.614759	0.70621
4800.0	0.448711	0.518857
5000.0	0.332268	0.38928
5200.0	0.251457	0.298324
5400.0	0.193004	0.233706
5600.0	0.151284	0.18659
5800.0	0.12029	0.151475

Table A.5.: Signal Monte Carlo samples generated in PYTHIA for SSM W'. Cross sections in NNLO are only calculated for the SSM W' samples [163].

A.5. Additional Kinematic Distributions for 13 TeV

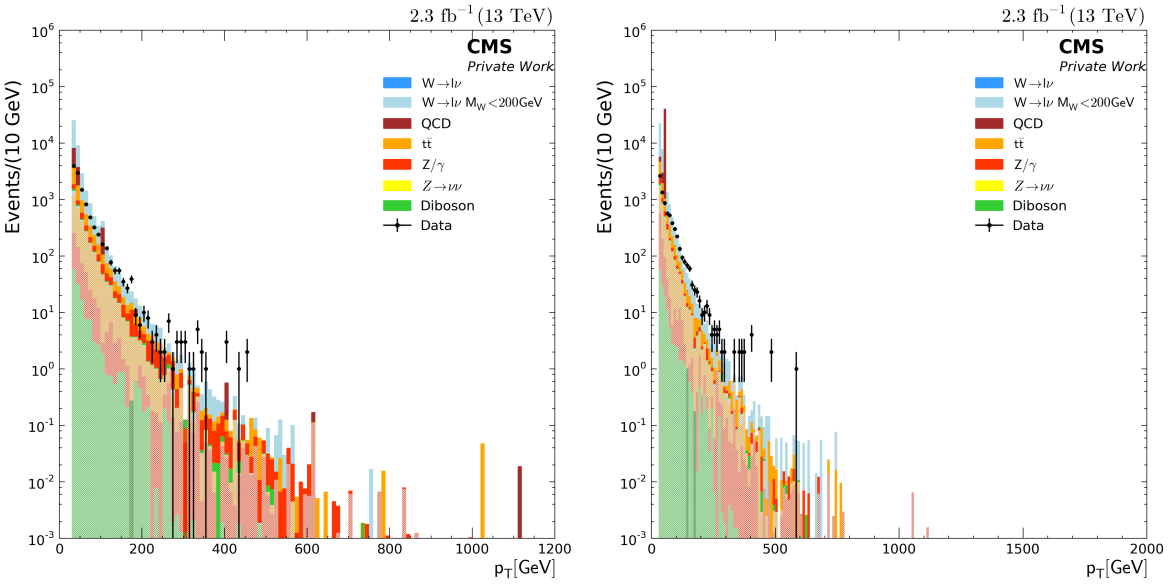


Figure A.9.: Same as in Figure 8.11, separated for events with one electron or one muon. The plots show the tau candidate p_T distribution of the isolated events -region D- in the e control region (left) and μ control region (right). The light coloured backgrounds are matched to a true tau.

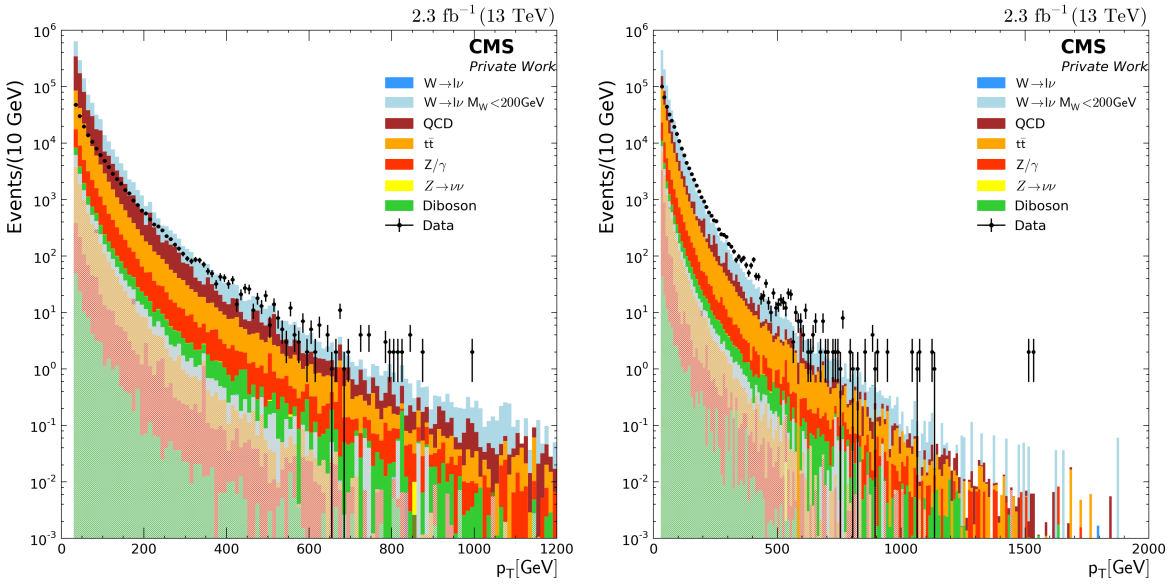


Figure A.10.: Same as in Figure 8.11, separated for events with one electron or one muon. The plots show the tau candidate p_T distribution of the non-isolated events - region C- (right plots) in the e control region (left) and μ control region (right). The light coloured backgrounds are matched to a true tau.

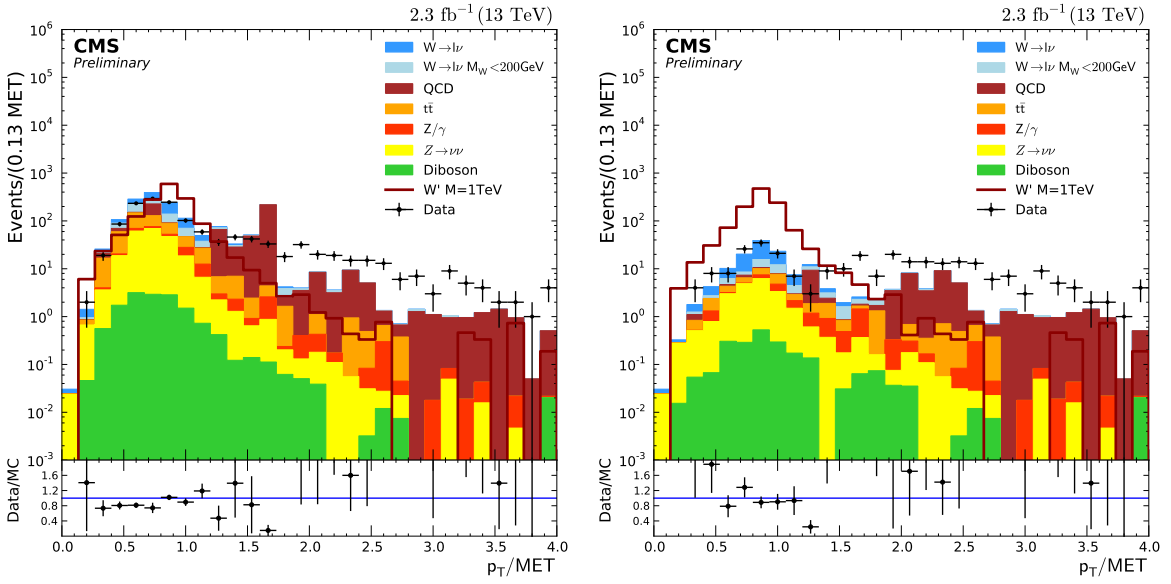


Figure A.11.: In the left plot the p_T^{had}/E_T^{miss} distribution as in figure 8.6 is shown with all backgrounds from simulation. In the right plot the same selection except with a $M_T > 400\,GeV$ selection. This shows that QCD dijet events can fake the signature if one of the jets is only partially reconstructed, and the E_T^{miss} has half of the jets energy.

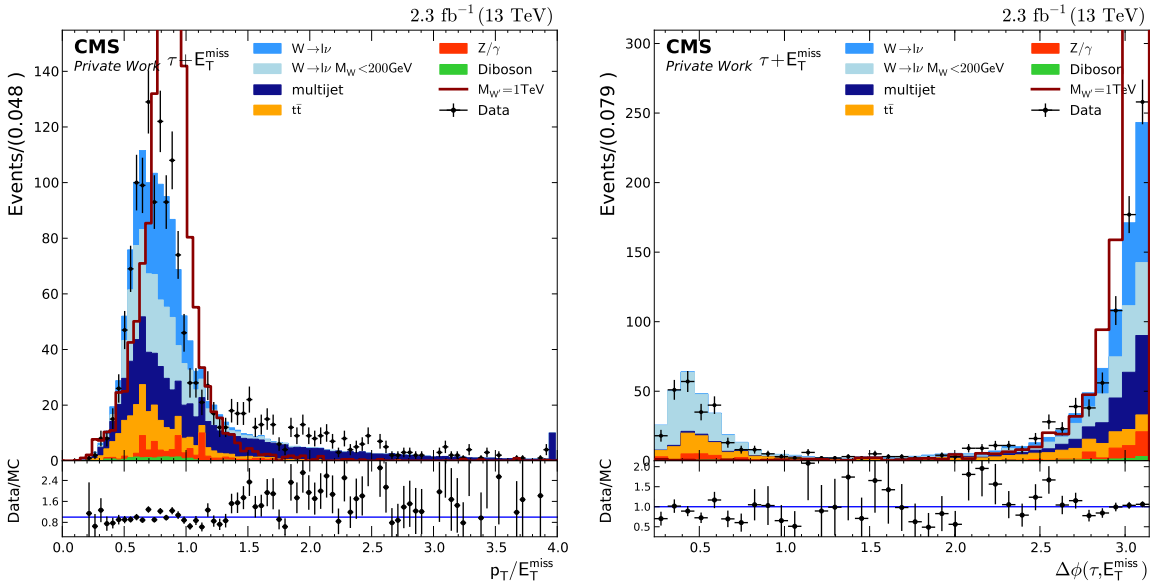


Figure A.12.: In the left plot the p_T^{had}/E_T^{miss} distribution as in figure 8.6 in linear scale. In the right plot the $\Delta\phi(\tau^{had}, E_T^{miss})$ plot in linear scale.

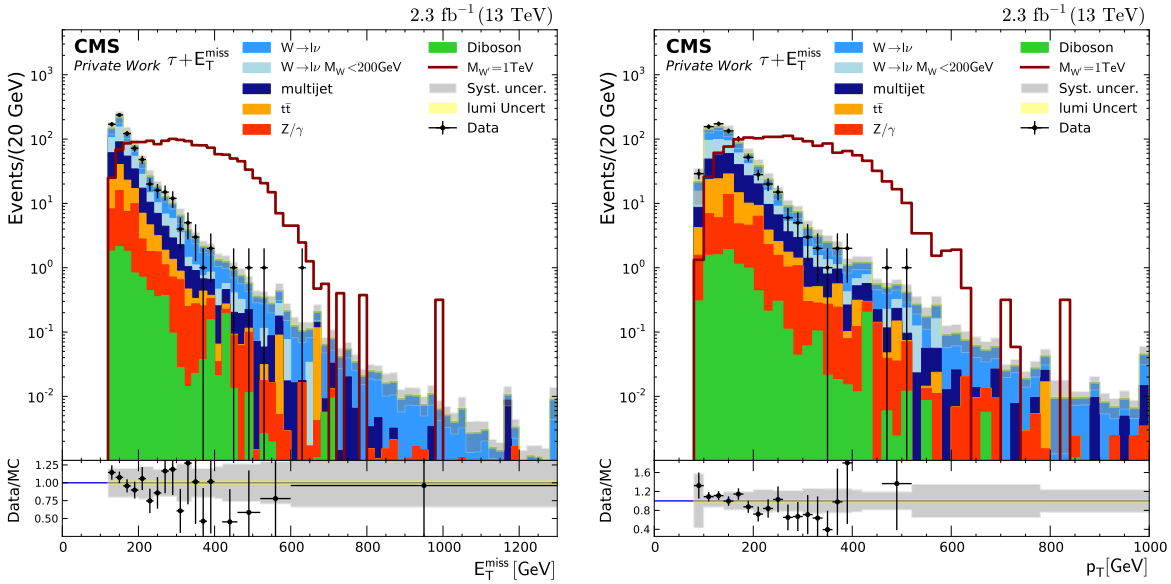


Figure A.13.: In the left plot the E_T^{miss} distribution in the signal region. In the right plot the tau p_T in the signal region.

Control Plots Without the p_T^{had}/E_T^{miss} Requirement

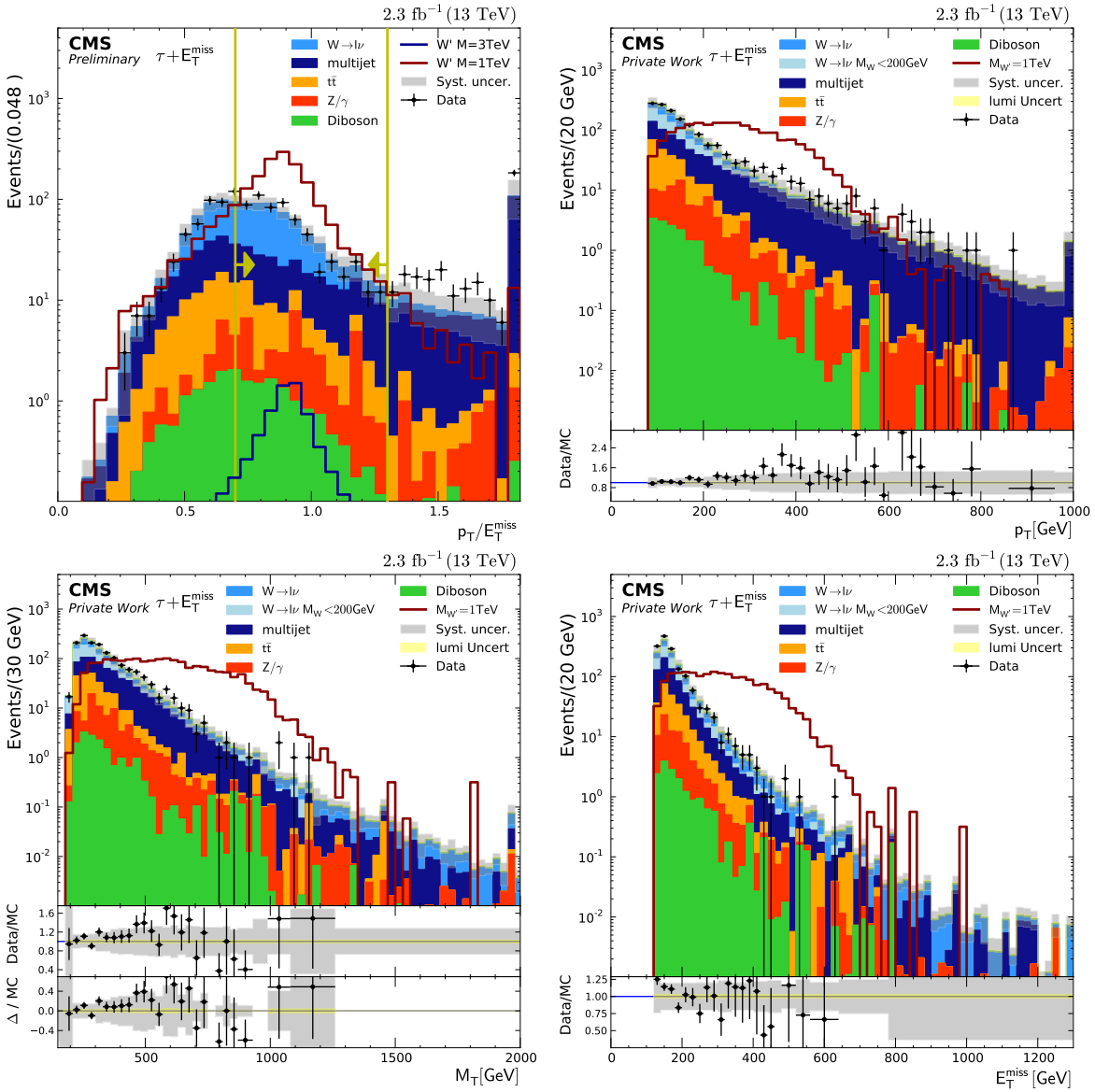


Figure A.14.: Control plots for the full selection without the p_T^{had}/E_T^{miss} Requirement.

Control Plots Without the $\Delta\phi(\tau^{\text{had}}, E_T^{\text{miss}})$ Requirement

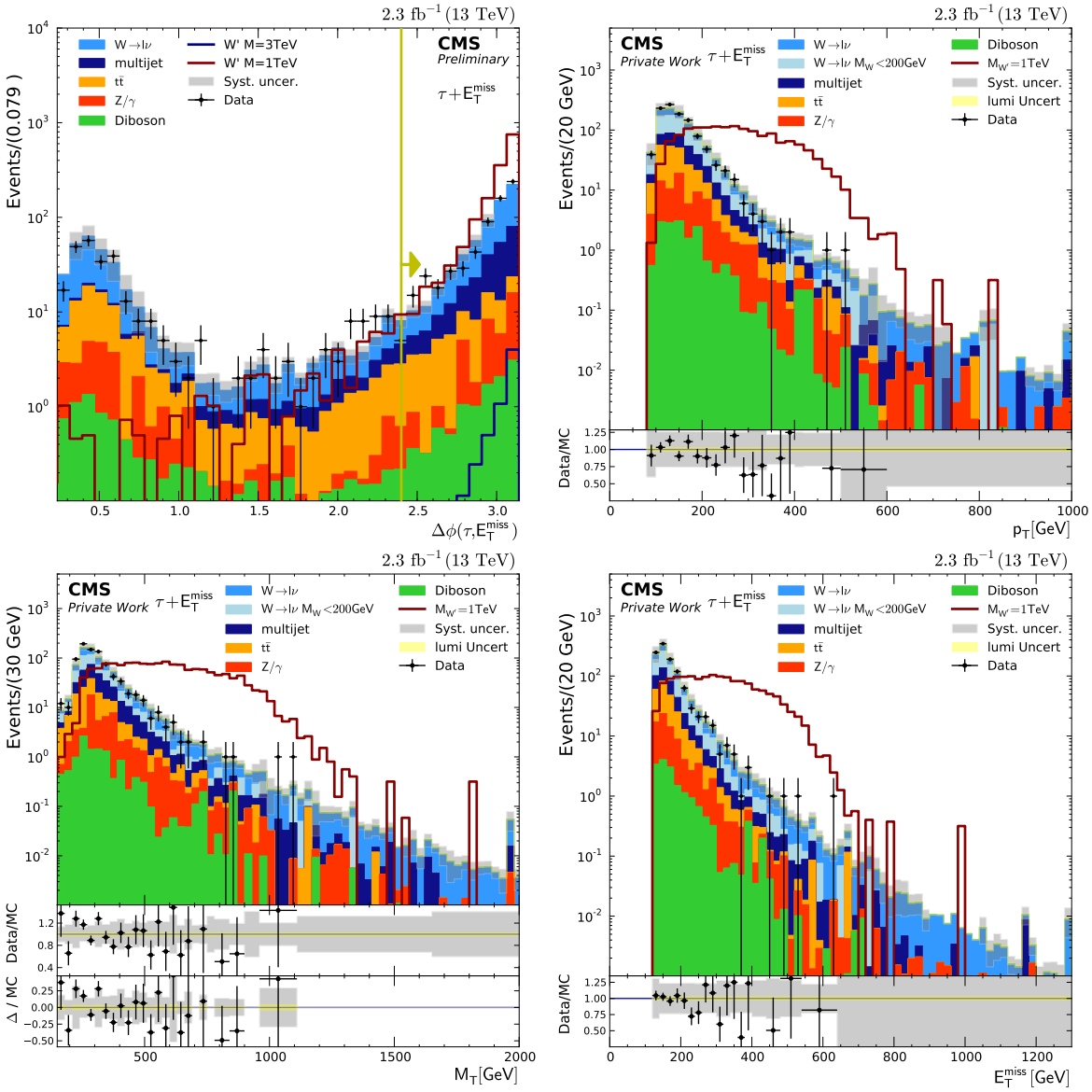


Figure A.15.: Control plots for the full selection without the $\Delta\phi(\tau^{\text{had}}, E_T^{\text{miss}})$ Requirement.

Control Plots with the Inverted p_T^{had}/E_T^{miss} Requirement

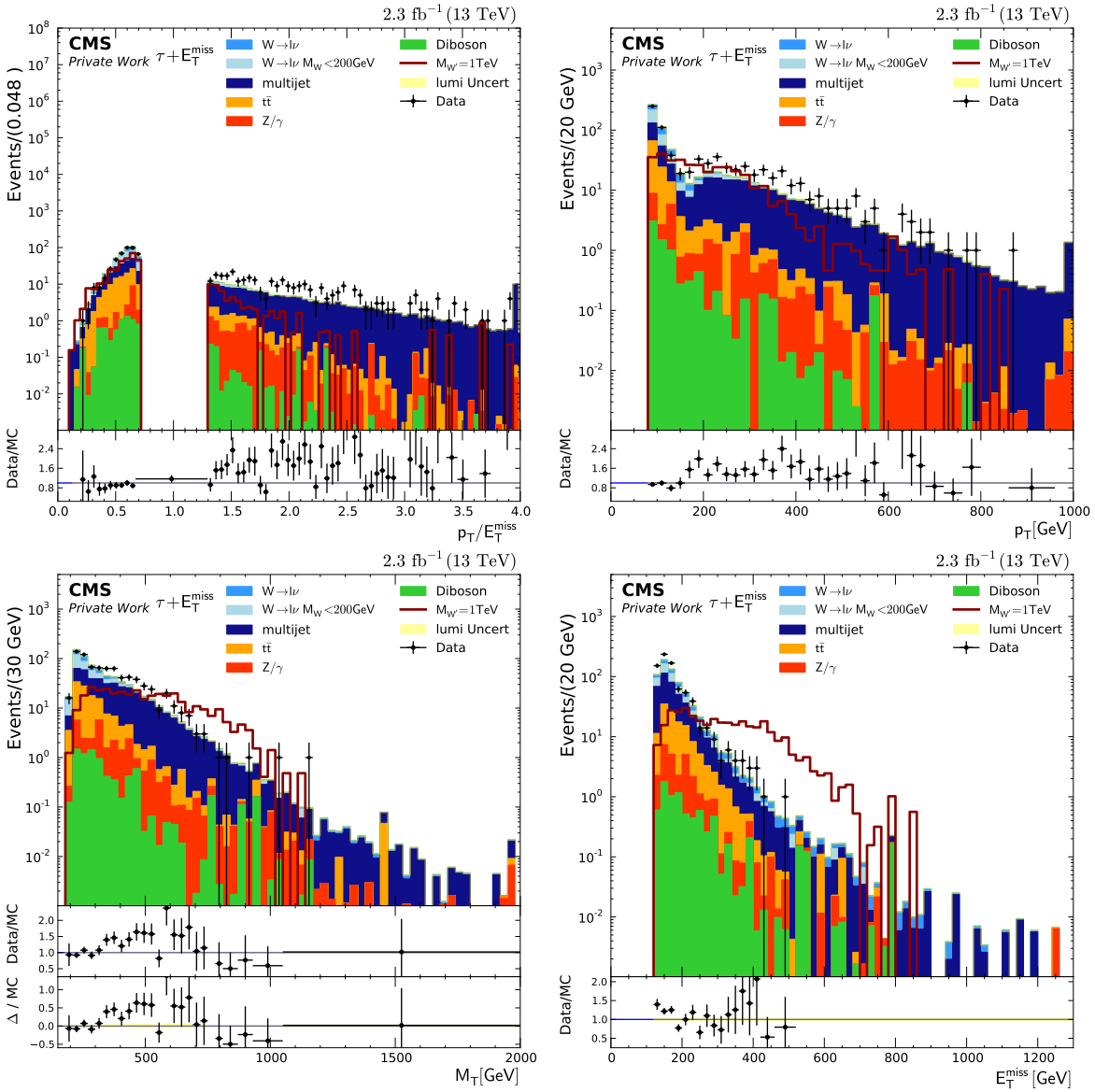


Figure A.16.: Control plots for the full selection with the inverted p_T^{had}/E_T^{miss} Requirement.

Control Plots with the Inverted $\Delta\phi(\tau^{\text{had}}, E_T^{\text{miss}})$ Requirement

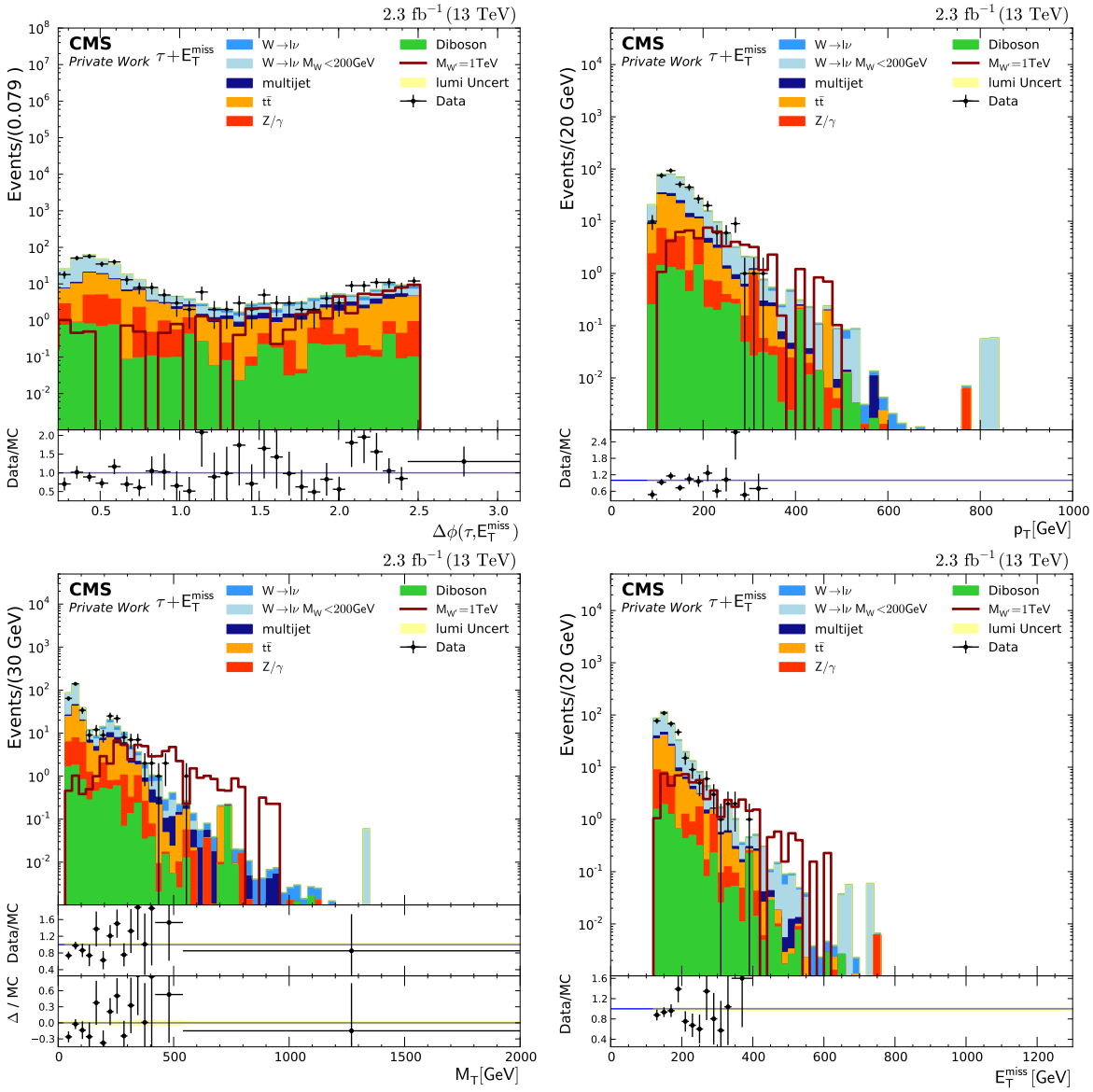


Figure A.17.: Control plots for the full selection with the inverted $\Delta\phi(\tau^{\text{had}}, E_T^{\text{miss}})$ Requirement.

Control Plots in the Signal Region Split into the Decay Mode

The QCD jets are estimated with the decay mode-p_T dependent fake rate, since not considering the decay mode would lead to wrong predictions in the individual decay modes.

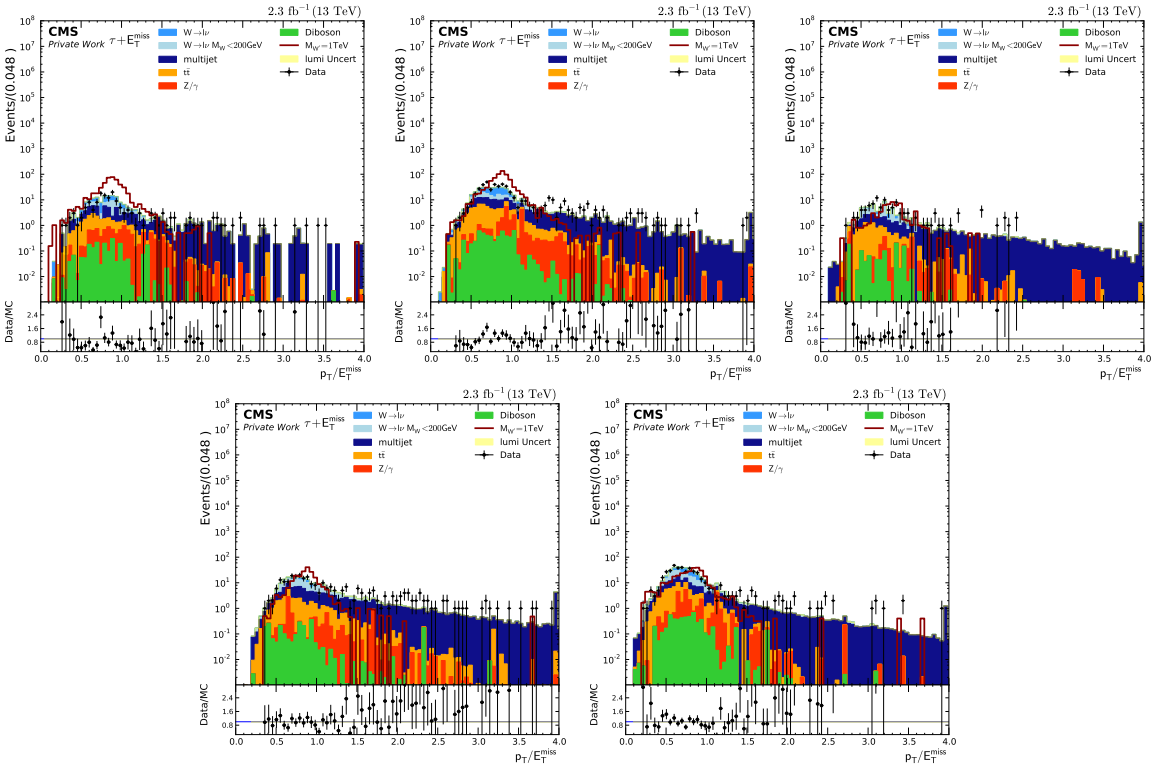


Figure A.18.: Control plots for the full selection, without the p_T^{had}/E_T^{miss} requirement, split into the decay mode. The order is 1Prong $0\pi^0$,1Prong $1\pi^0$ and 2Prong $0\pi^0$ at the top and 2Prong $1\pi^0$ and 3Prong $0\pi^0$ in the bottom.

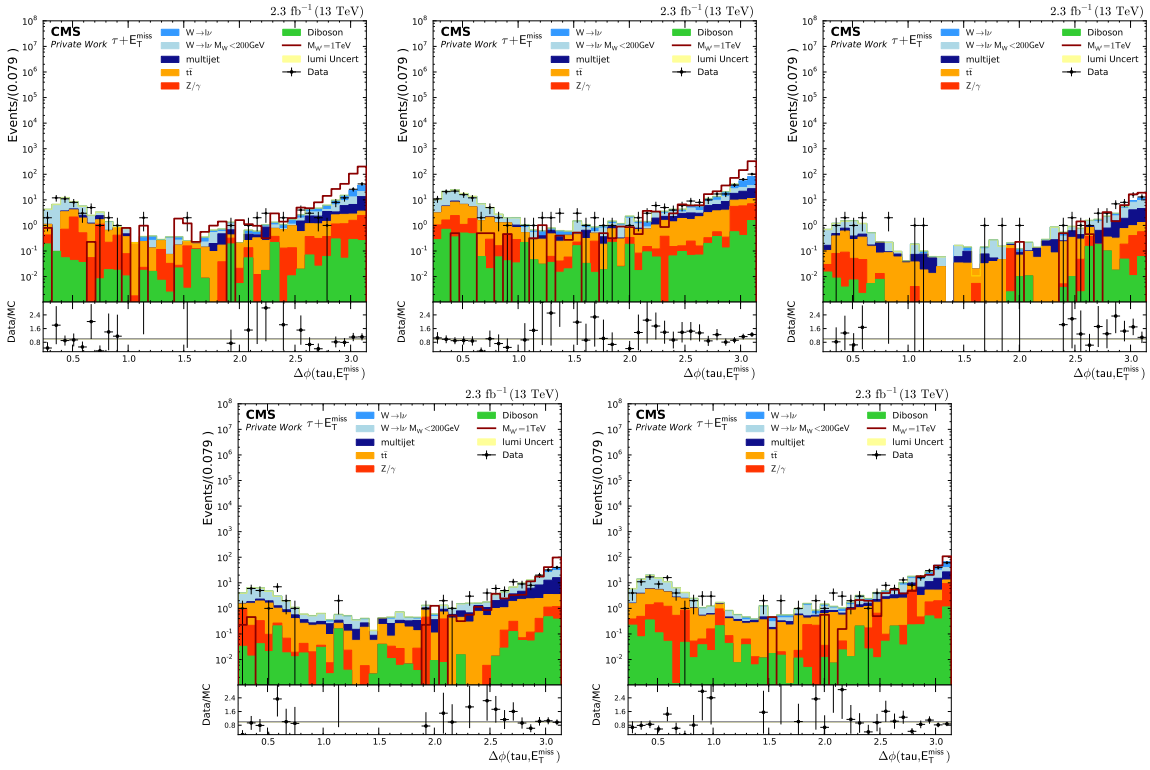


Figure A.19.: Control plots for the full selection, without the $\Delta\phi(\tau^{had}, E_T^{miss})$ requirement, split into the decay mode. The order is 1Prong $0\pi^0$,1Prong $1\pi^0$ and 2Prong $0\pi^0$ at the top and 2Prong $1\pi^0$ and 3Prong $0\pi^0$ in the bottom.

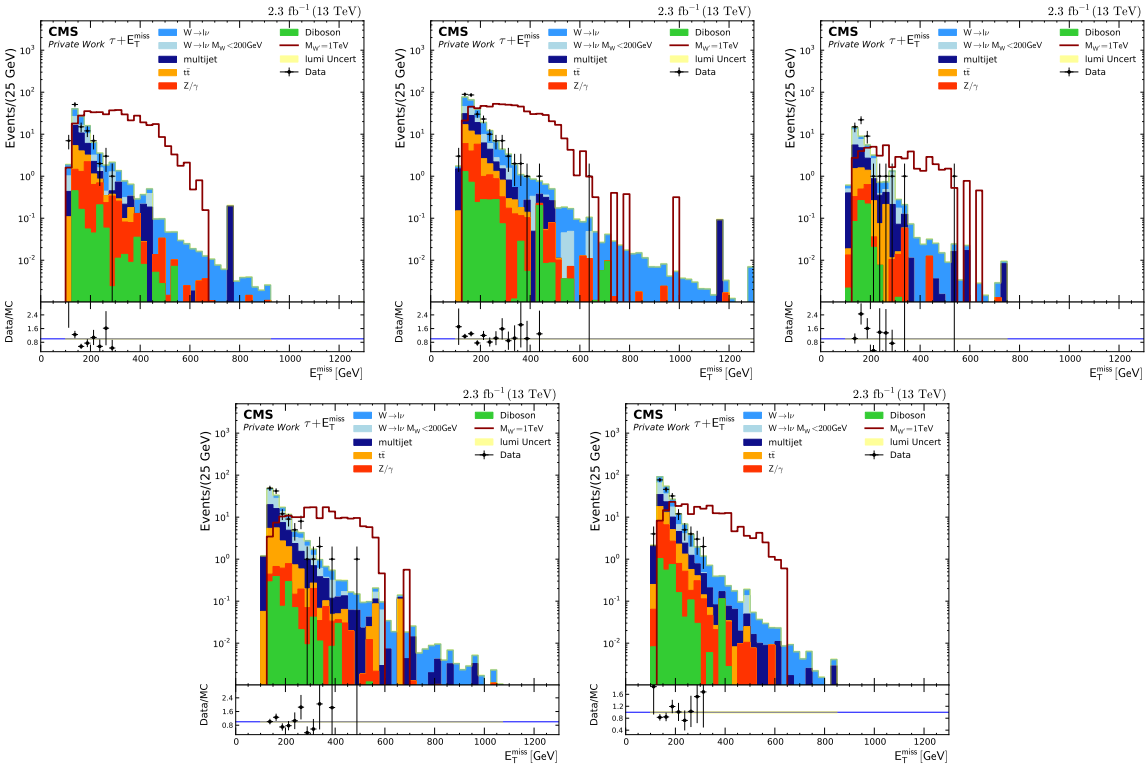


Figure A.20.: Control plots for the full selection split into the decay mode. The order is 1Prong $0\pi^0$,1Prong $1\pi^0$ and 2Prong $0\pi^0$ at the top and 2Prong $1\pi^0$ and 3Prong $0\pi^0$ in the bottom.

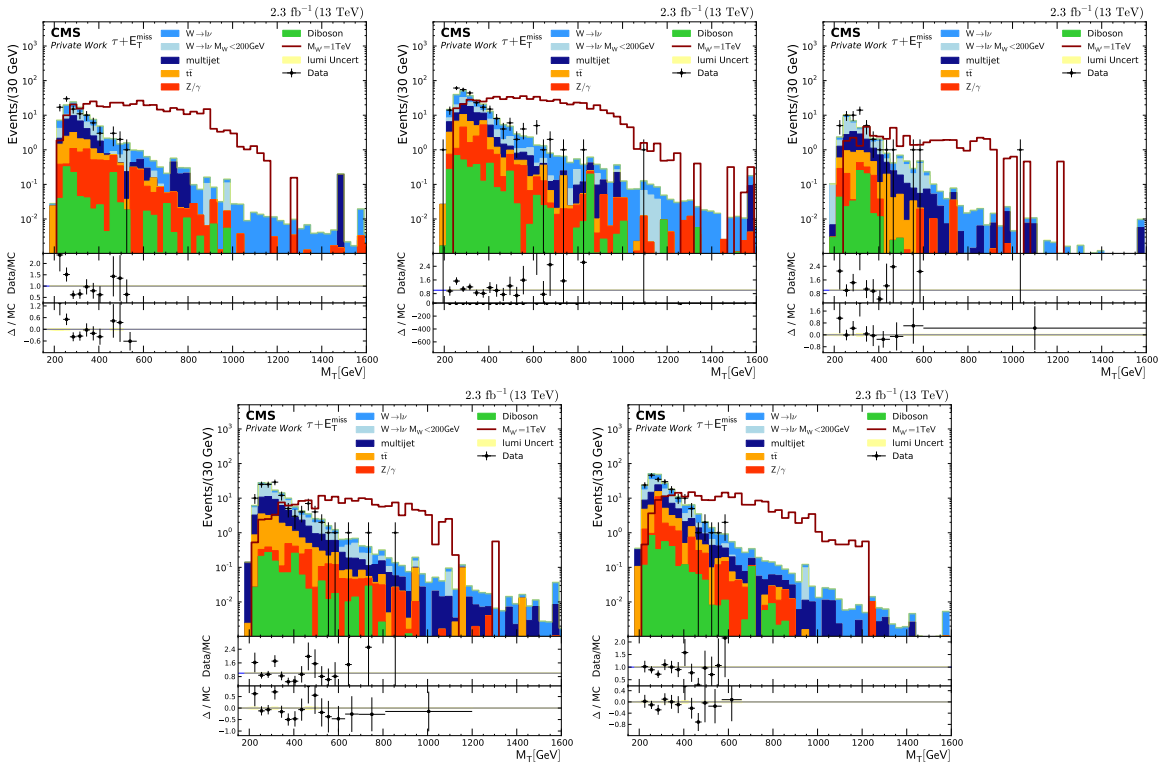


Figure A.21.: Control plots for the full selection split into the decay mode. The order is 1Prong $0\pi^0$,1Prong $1\pi^0$ and 2Prong $0\pi^0$ at the top and 2Prong $1\pi^0$ and 3Prong $0\pi^0$ in the bottom.

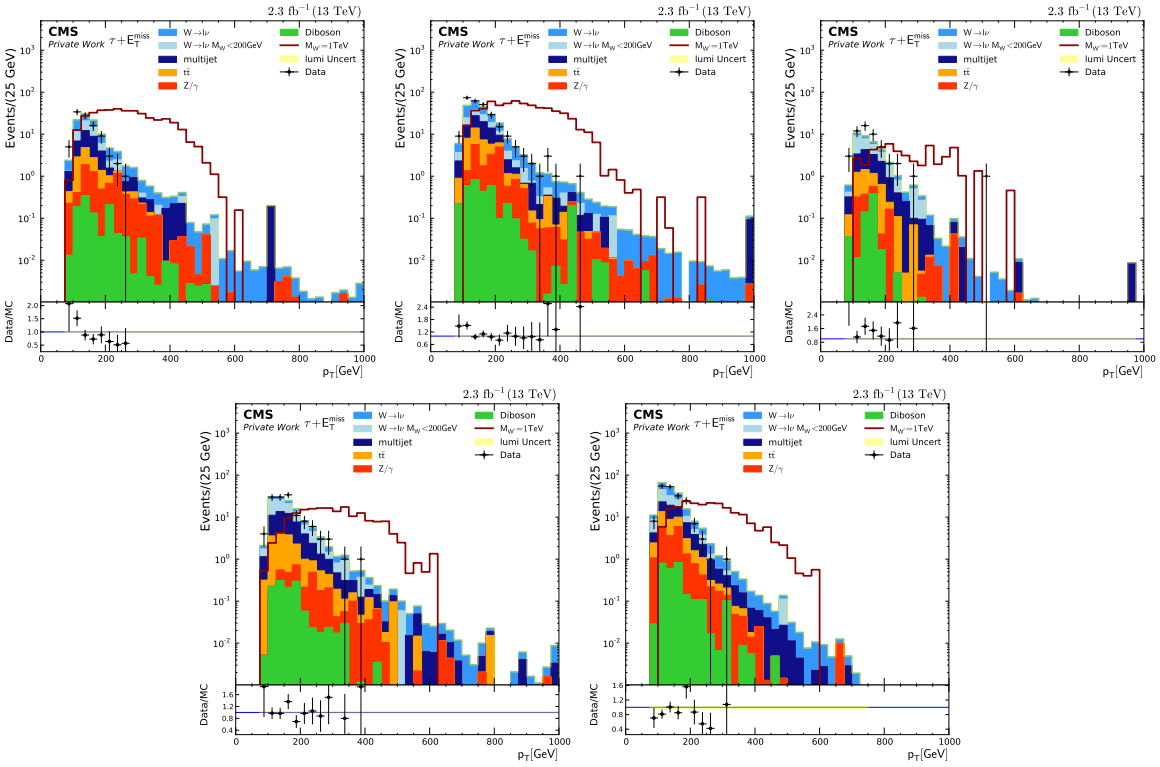


Figure A.22.: Control plots for the full selection split into the decay mode. The order is 1Prong $0\pi^0$,1Prong $1\pi^0$ and 2Prong $0\pi^0$ at the top and 2Prong $1\pi^0$ and 3Prong $0\pi^0$ in the bottom.

Various Selections

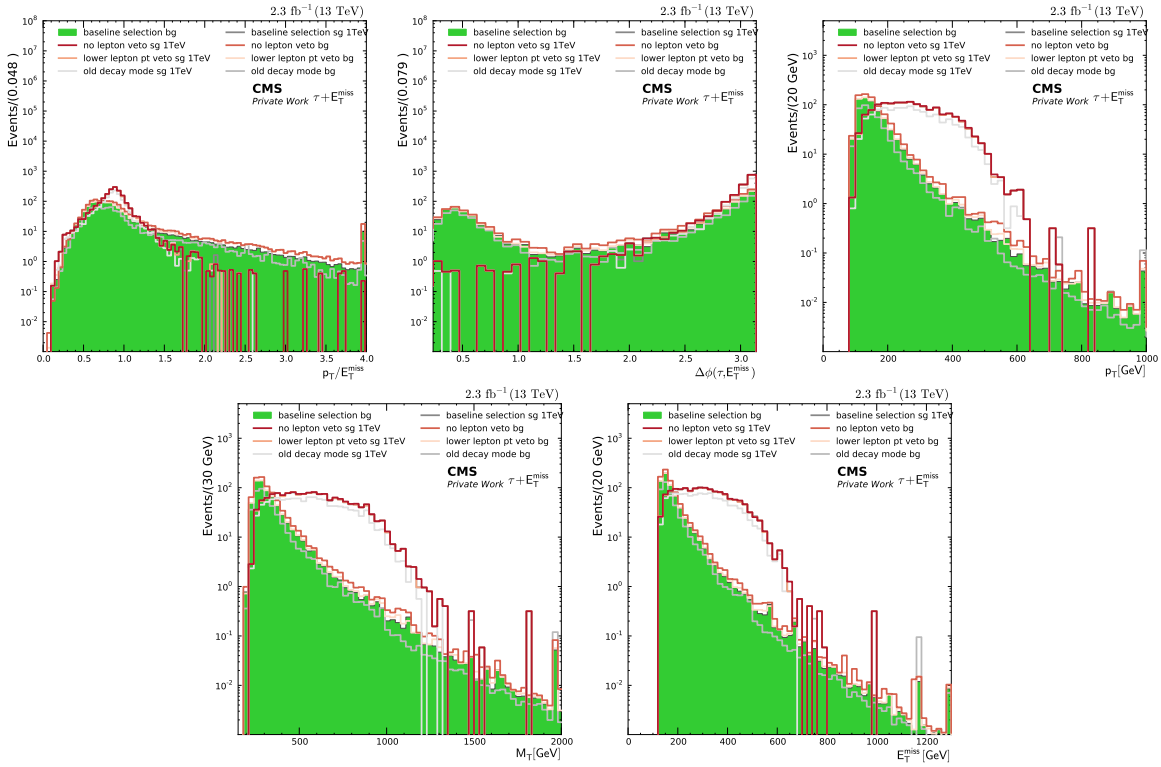


Figure A.23.: Plots that compare the selections listed before with the baseline selection.

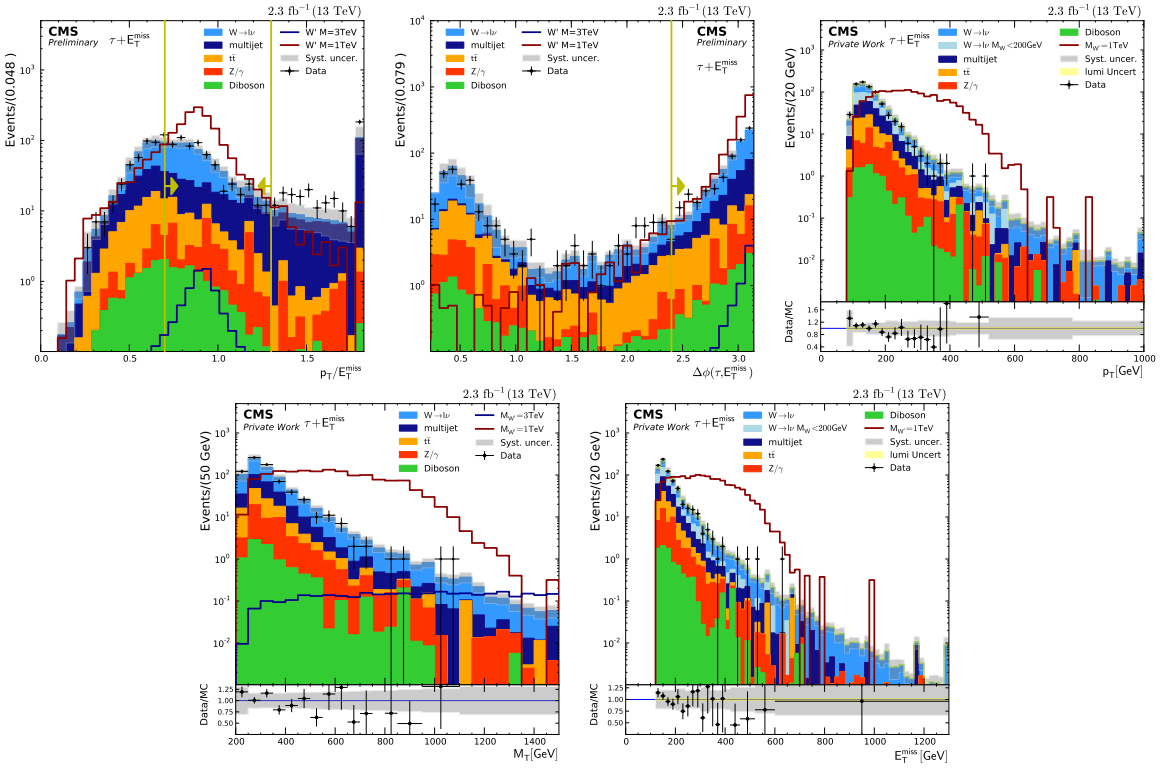


Figure A.24.: All plots for the baseline selection.

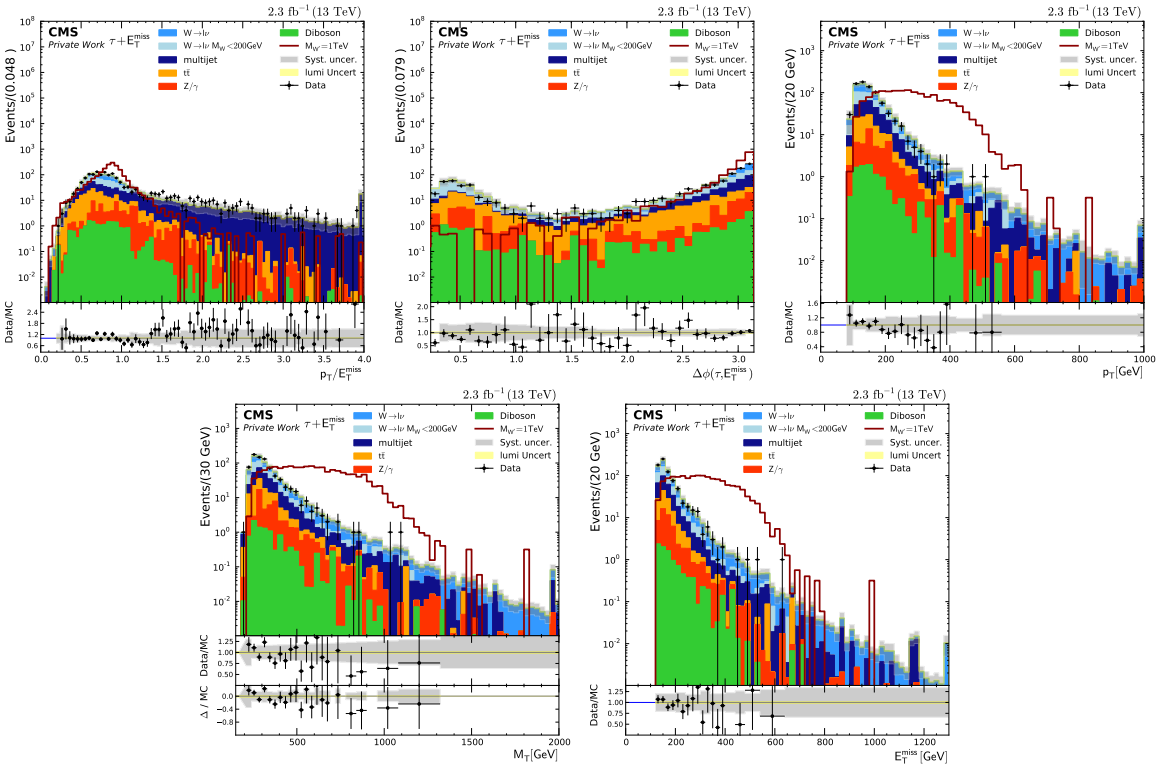


Figure A.25.: Plots for the same selection but with no lepton veto.

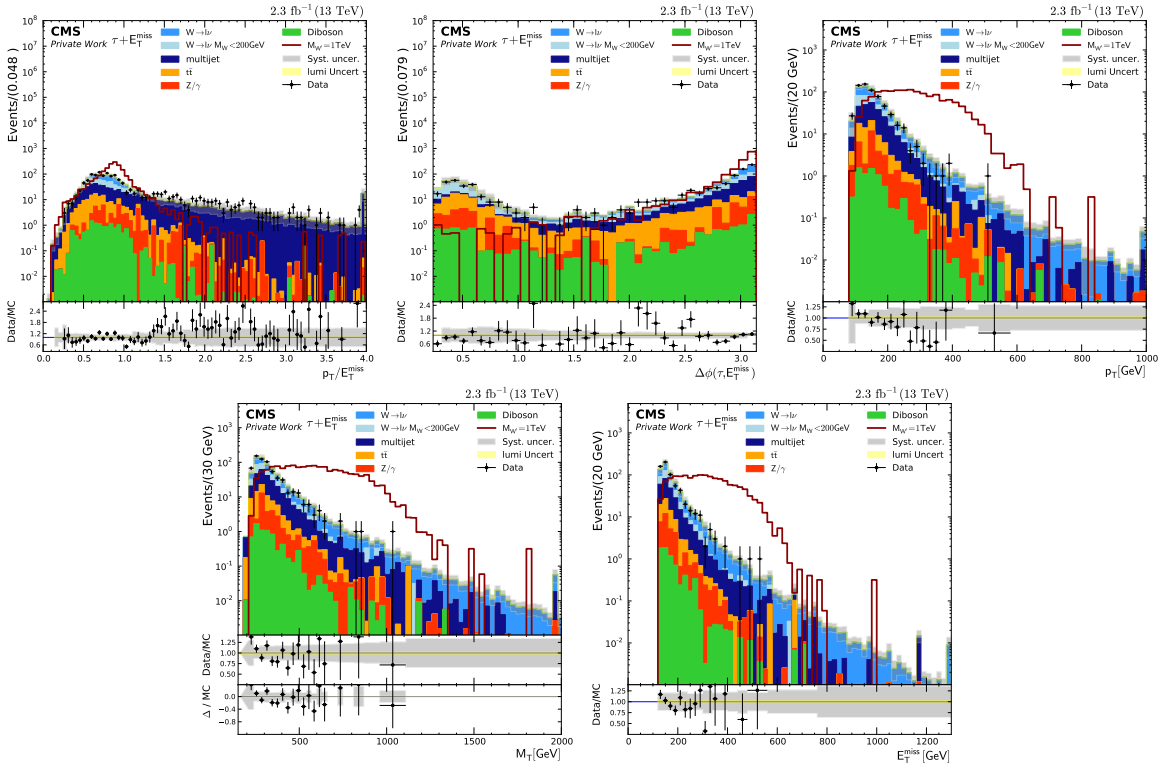


Figure A.26.: Plots for the same selection but with a p_T threshold of 10 GeV for the lepton veto.

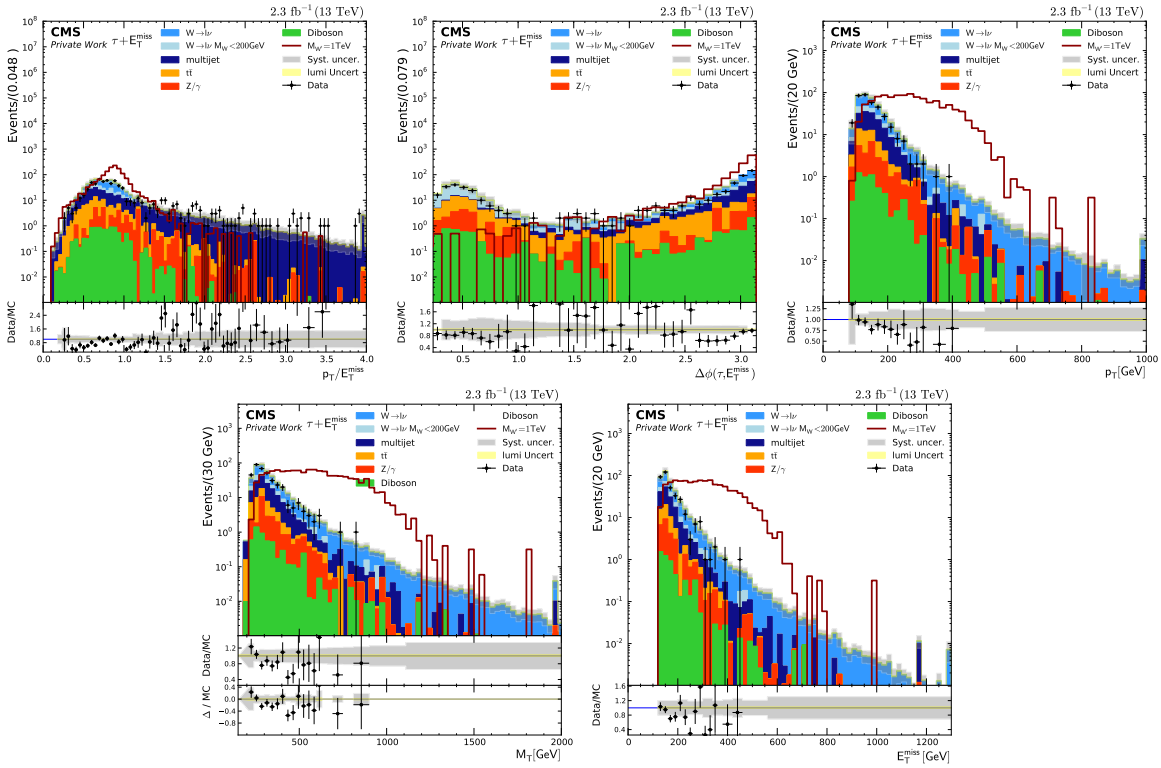


Figure A.27.: Plots for the same selection but with the "old" decay mode, where 2 prongs are not reconstructed.

A.6. Additional Plots for the e/μ-Selection

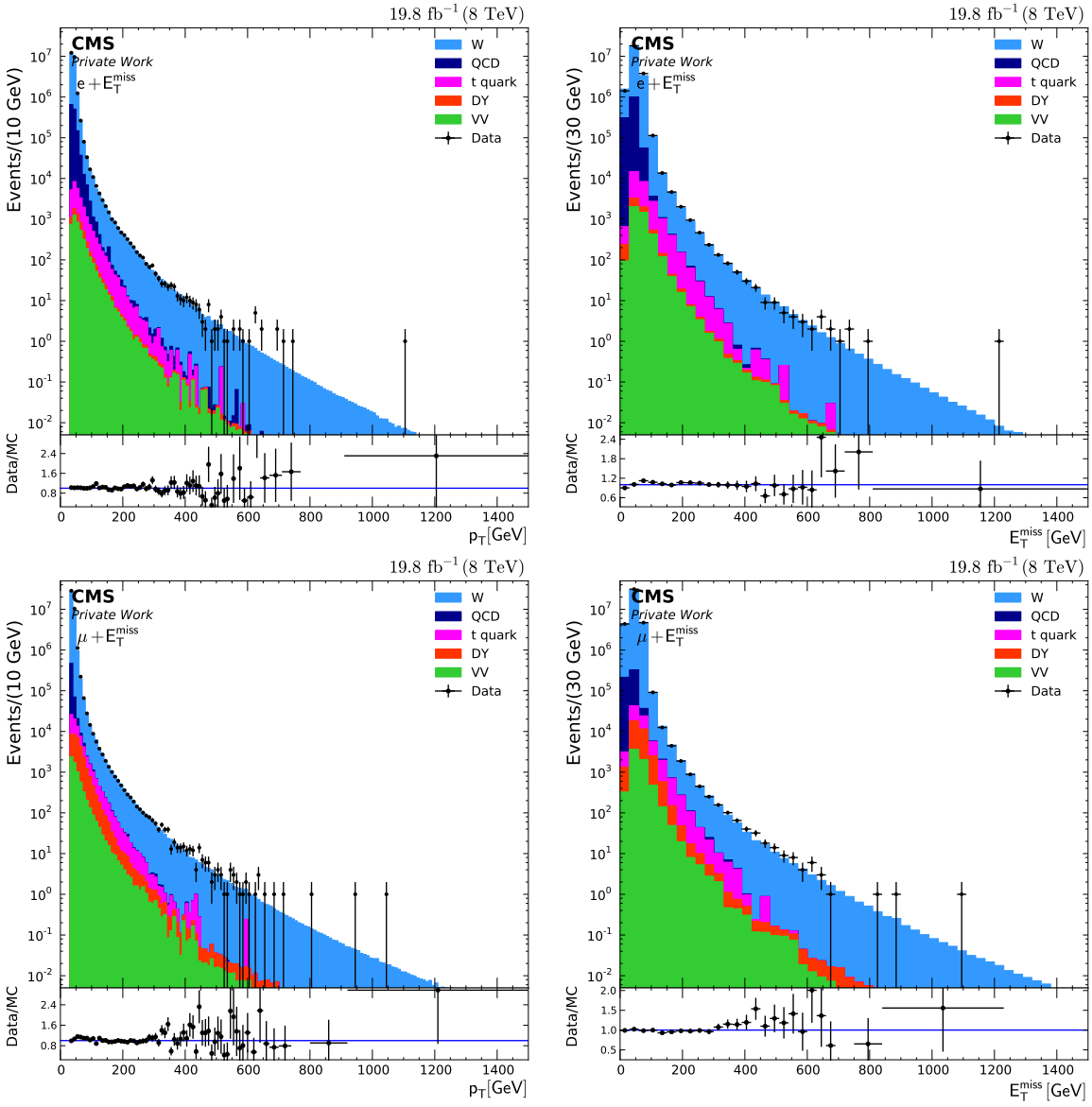


Figure A.28.: p_T (left) and E_T^{miss} (right) for the final selection, where the upper plots are for the electron and the lower plots are for the muon channel.

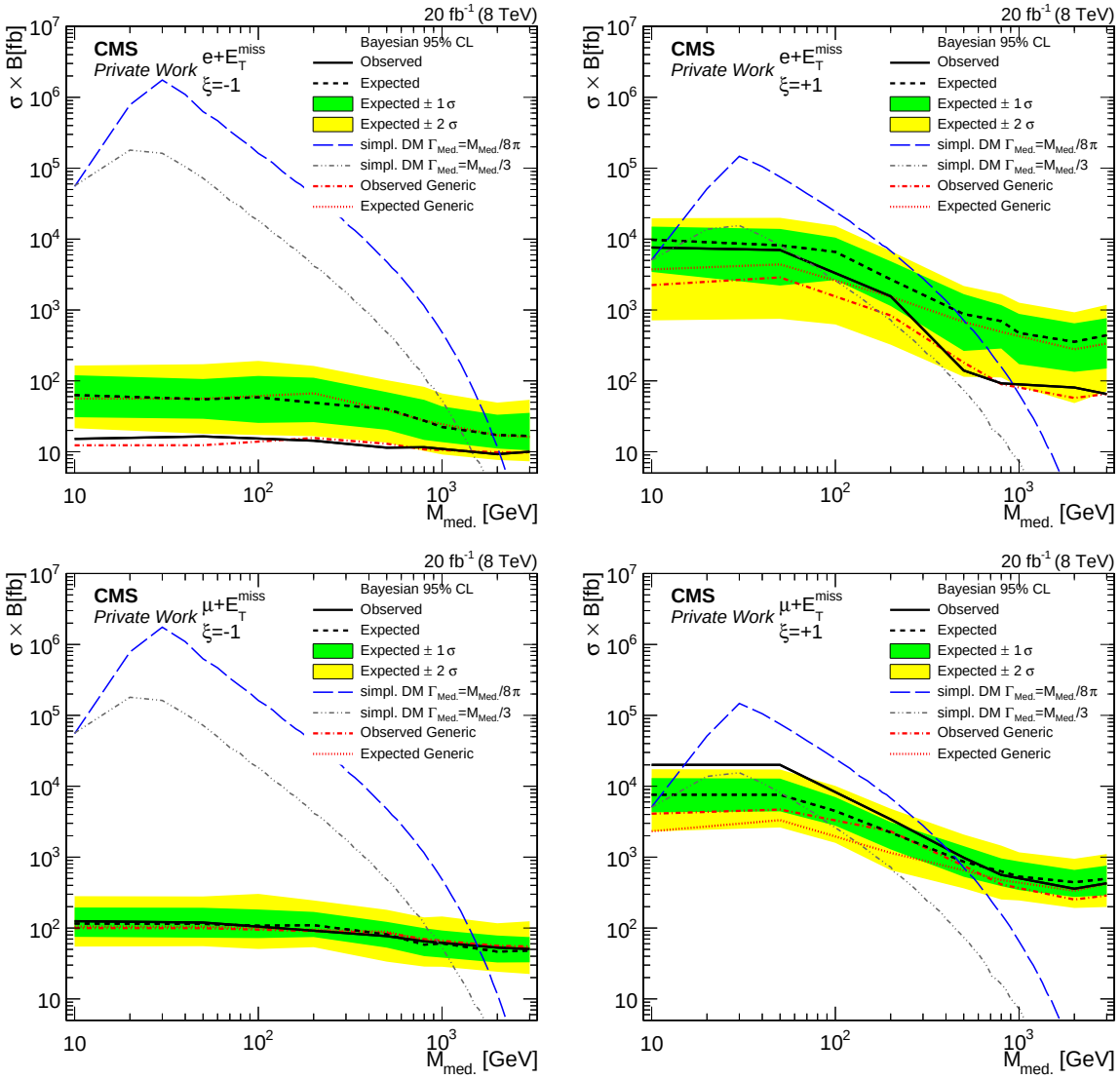


Figure A.29.: The plots show Dark Matter simplified model limits for the combined e/μ channel. The plots show various ξ parameters $\xi = -1$ (left) and $\xi = 1$ (right). The shown theoretical cross section is plotted for two mediator widths. The narrow width for $\Gamma_{Med.} = \frac{M_{Med.}}{8\pi}$ and the broad width $\Gamma_{Med.} = \frac{M_{Med.}}{3}$.

Bibliography

[1] CMS Collaboration Collaboration, "Search for new physics in final states with a tau and missing transverse energy using pp collisions at $\sqrt{s} = 8$ TeV", *CMS Physics Analysis Summary: EXO-12-011* (2015) . http://cds.cern.ch/record/2002036.

[2] CMS Collaboration, "Search for W' decaying to tau lepton and neutrino in proton-proton collisions at $\sqrt{(s)} = 8$ TeV", *Phys. Lett.* B755 (2016) 196–216, arXiv:1508.04308 [hep-ex].

[3] CMS Collaboration Collaboration, "Search for W' decaying to tau lepton and neutrino in proton-proton collisions at sqrt(s) = 13 TeV", *CMS Physics Analysis Summary: EXO-16-006* (2016) . http://cds.cern.ch/record/2140976.

[4] CMS Collaboration, "Reconstruction and identification of lepton decays to hadrons and at CMS", *JINST* 11 no. 01, (2016) P01019, arXiv:1510.07488 [physics.ins-det].

[5] CMS Collaboration, "Search for dark matter in the mono-lepton channel with pp collision events at center-of-mass energy of 8 TeV", *CMS Physics Analysis Summary: EXO-13-004* (2013) . https://cds.cern.ch/record/1563245.

[6] CMS Collaboration, "Search for physics beyond the standard model in final states with a lepton and missing transverse energy in proton-proton collisions at $\sqrt{s} = 8$ TeV", *Accepted by Phys. Rev. D* (2014) , arXiv:1408.2745 [hep-ex].

[7] CMS Collaboration Collaboration, "Search for SSM W' production, in the lepton and missing transverse momentum final state at a center-of-mass energy of 13 TeV", *CMS Physics Analysis Summary: EXO-15-006* (2015) . http://cds.cern.ch/record/2114864.

[8] P. Millet, *Search for new physics in pp collision events with one electron and missing transverse energy using CMS data, Master thesis.* Master thesis, RWTH Aachen, Aachen, 2013. https://web.physik.rwth-aachen.de/~hebbeker/theses/millet_master.pdf.

[9] S. Erdweg, *Search for Dark Matter and W' in the final state with one muon and missing transverse energy with CMS, Master thesis.* Master thesis, RWTH Aachen, Aachen, 2013. https://web.physik.rwth-aachen.de/~hebbeker/theses/erdweg_master.pdf.

[10] D. Noll, *A Study of Dark Matter Models at sqrt s = 13 TeV pp Collisions for Mono-Lepton Searches.* Bachelor thesis, RWTH Aachen, Aachen, 2015. https://web.physik.rwth-aachen.de/~hebbeker/theses/noll_bachelor.pdf.

[11] M. Materok, *Search for dark matter with scalar or pseudoscalar coupling in Run 2 data*. Bachelor thesis, RWTH Aachen, Aachen, 2016. https://web.physik.rwth-aachen.de/~hebbeker/theses/materok_bachelor.pdf.

[12] H. Acaroglu, *A Study of Simplified Models for Dark Matter Searches in pp Collisions*. Bachelor thesis, RWTH Aachen, Aachen, 2015. https://web.physik.rwth-aachen.de/~hebbeker/theses/acaroglu_bachelor.pdf.

[13] V. Kutzner, *Search for dark matter with scalar or pseudoscalar coupling in Run 2 data*. Master thesis, RWTH Aachen, Aachen, 2016. https://web.physik.rwth-aachen.de/~hebbeker/theses/kutzner_master.pdf.

[14] S. Knutzen, *Search for heavy vector bosons W' in the tau decay channel using pp collisions at 8 TeV*. Master thesis, RWTH Aachen, Aachen, 2013. https://web.physik.rwth-aachen.de/~hebbeker/Thesis/knutzen_master.pdf.

[15] Particle Data Group Collaboration, "Review of Particle Physics", *Chin. Phys.* C38 (2014) 090001.

[16] CMS Collaboration, "Measurement of the top quark mass using charged particles in pp collisions at $\sqrt{s} = 8$ TeV", *Phys. Rev.* D93 no. 9, (2016) 092006, arXiv:1603.06536 [hep-ex].

[17] ATLAS, CMS Collaboration, "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV", *JHEP* 08 (2016) 045, arXiv:1606.02266 [hep-ex].

[18] S. L. Glashow, "Partial Symmetries of Weak Interactions", *Nucl. Phys.* 22 (1961) 579–588.

[19] S. L. Glashow, J. Iliopoulos, and L. Maiani, "Weak Interactions with Lepton-Hadron Symmetry", *Phys. Rev.* D2 (1970) 1285–1292.

[20] S. Weinberg, "A Model of Leptons", *Phys. Rev. Lett.* 19 (1967) 1264–1266.

[21] J. L. Rosner, "The Standard model in 2001", in *Heavy flavor physics: Theory and experimental results in heavy quark physics and CP violation. Proceedings, 55th Scottish Universities Summer School in Physics, SUSSP 2001, St. Andrews, UK, August 7-23, 2001*, pp. 1–56. 2001. arXiv:hep-ph/0108195 [hep-ph]. http://alice.cern.ch/format/showfull?sysnb=2270815.

[22] A. Pich, "The Standard model of electroweak interactions", in *2004 European School of High-Energy Physics, Sant Feliu de Guixols, Spain, 30 May - 12 June 2004*, pp. 1–48. 2005. arXiv:hep-ph/0502010 [hep-ph]. http://doc.cern.ch/yellowrep/2006/2006-003/p1.pdf.

[23] R. Wolf, *The Higgs Boson Discovery at the Large Hadron Collider*, vol. 264. Springer, 2015.

[24] L. Edelhäuser and A. Knochel, *Tutorium Quantenfeldtheorie - Was Sie schon immer über QFT wissen wollten, aber bisher nicht zu fragen wagten.* Springer, 2016. ISBN-978-3-642-37675-7.

[25] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory.* Perseus Books, 1995. ISBN-9780201503975.

[26] R. K. Ellis, W. J. Stirling, and B. R. Webber, "QCD and collider physics", *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.* 8 (1996) 1–435.

[27] T. Sjöstrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual", *JHEP* 5 (2006) 26, arXiv:hep-ph/0603175 [hep-ph].

[28] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, "An Introduction to PYTHIA 8.2", *Comput. Phys. Commun.* 191 (2015) 159–177, arXiv:1410.3012 [hep-ph].

[29] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", *Eur. Phys. J.* C76 no. 3, (2016) 155, arXiv:1512.00815 [hep-ex].

[30] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* 07 (2014) 079, arXiv:1405.0301 [hep-ph].

[31] J. H. Kuhn, A. Kulesza, S. Pozzorini, and M. Schulze, "Electroweak corrections to hadronic production of W bosons at large transverse momenta", *Nucl. Phys.* B797 (2008) 27–77, arXiv:0708.0476 [hep-ph].

[32] F. Cascioli, P. Maierhofer, and S. Pozzorini, "Scattering Amplitudes with Open Loops", *Phys. Rev. Lett.* 108 (2012) 111601, arXiv:1111.5206 [hep-ph].

[33] S. Kallweit, J. M. Lindert, P. Maierhofer, S. Pozzorini, and M. Schönherr, "NLO QCD+EW predictions for V + jets including off-shell vector-boson decays and multijet merging", *JHEP* 04 (2016) 021, arXiv:1511.08692 [hep-ph].

[34] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, "MadGraph 5 : Going Beyond", *JHEP* 06 (2011) 128, arXiv:1106.0522 [hep-ph].

[35] Z. Bern, L. J. Dixon, D. C. Dunbar, and D. A. Kosower, "One loop n point gauge theory amplitudes, unitarity and collinear limits", *Nucl. Phys.* B425 (1994) 217–260, arXiv:hep-ph/9403226 [hep-ph].

[36] Z. Bern, L. J. Dixon, and D. A. Kosower, "One loop amplitudes for e+ e- to four partons", *Nucl. Phys.* B513 (1998) 3–86, arXiv:hep-ph/9708239 [hep-ph].

[37] S. D. Badger, "Direct Extraction Of One Loop Rational Terms", *JHEP* 01 (2009) 049, arXiv:0806.4600 [hep-ph].

[38] C. M. Carloni Calame, G. Montagna, O. Nicrosini, and A. Vicini, "Precision electroweak calculation of the production of a high transverse-momentum lepton pair at hadron colliders", *JHEP* 10 (2007) 109, arXiv:0710.1722 [hep-ph].

[39] S. G. Bondarenko and A. A. Sapronov, "NLO EW and QCD proton-proton cross section calculations with mcsanc-v1.01", *Comput. Phys. Commun.* 184 (2013) 2343, arXiv:1301.3687 [hep-ph].

[40] S. Kallweit, J. M. Lindert, P. Maierhofer, S. Pozzorini, and M. Schönherr, "NLO QCD+EW predictions for V + jets including off-shell vector-boson decays and multijet merging", *JHEP* 04 (2016) 021, arXiv:1511.08692 [hep-ph].

[41] R. Thorne, "The Role of uncertainties in parton distribution functions", *PHYSTAT-LHC Workshop on Statistical Issues* (2007) 141–150, arXiv:0711.2986 [hep-ph].

[42] G. Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language", *Nucl. Phys.* B126 (1977) 298–318.

[43] Y. L. Dokshitzer, "Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.", *Sov. Phys. JETP* 46 (1977) 641–653. [Zh. Eksp. Teor. Fiz.73,1216(1977)].

[44] V. N. Gribov and L. N. Lipatov, "Deep inelastic e p scattering in perturbation theory", *Sov. J. Nucl. Phys.* 15 (1972) 438–450. [Yad. Fiz.15,781(1972)].

[45] H. L. Lai, J. Botts, J. Huston, J. G. Morfin, J. F. Owens, J.-w. Qiu, W. K. Tung, and H. Weerts, "Global QCD analysis and the CTEQ parton distributions", *Phys. Rev.* D51 (1995) 4763–4782, arXiv:hep-ph/9410404 [hep-ph].

[46] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, "Parton distributions in the LHC era: MMHT 2014 PDFs", *Eur. Phys. J.* C75 no. 5, (2015) 204, arXiv:1412.3989 [hep-ph].

[47] I. J. R. Aitchison, "Supersymmetry and the MSSM: An Elementary introduction", arXiv:hep-ph/0505105 [hep-ph].

[48] J. Silk *et al.*, *Particle Dark Matter: Observations, Models and Searches*. 2010. http://www.cambridge.org/uk/catalogue/catalogue.asp?isbn=9780521763684.

[49] G. Altarelli, "Collider Physics within the Standard Model: a Primer", arXiv:1303.2842 [hep-ph].

[50] H. Georgi and S. L. Glashow, "Unity of All Elementary Particle Forces", *Phys. Rev. Lett.* 32 (1974) 438–441.

[51] CMS Collaboration, "Phenomenological MSSM interpretation of CMS searches in pp collisions at sqrt(s) = 7 and 8 TeV", *Submitted to: JHEP* (2016) , arXiv:1606.03577 [hep-ex].

[52] C. Arina, M. E. C. Catalan, S. Kraml, S. Kulkarni, and U. Laa, "Constraints on sneutrino dark matter from LHC Run 1", *JHEP* 05 (2015) 142, arXiv:1503.02960 [hep-ph].

[53] C.-R. Chen, M. M. Nojiri, S. C. Park, J. Shu, and M. Takeuchi, "Dark matter and collider phenomenology of split-UED", *JHEP* 09 (2009) 078, arXiv:0903.1971 [hep-ph].

[54] K. Kong, S. C. Park, and T. G. Rizzo, "Collider phenomenology with Split-UED", *JHEP* 04 (2010) 081, arXiv:1002.0602 [hep-ph].

[55] M. Carena, L. Da Rold, and E. Pontón, "Minimal Composite Higgs Models at the LHC", *JHEP* 06 (2014) 159, arXiv:1402.2987 [hep-ph].

[56] G. Altarelli, B. Mele, and M. Ruiz-Altaba, "Searching for new heavy vector bosons in $p\bar{p}$ colliders", *Z. Phys. C* 45 (1989) 109.

[57] M. Olschewski, *Search for new physics in proton-proton collision events with a lepton and missing transverse energy*. Phd, RWTH Aachen, Aachen, 2016. https://publications.rwth-aachen.de/record/572409. Veröffentlicht auf dem Publikationsserver der RWTH Aachen University; Dissertation, RWTH Aachen, 2016.

[58] Z. Sullivan, "Fully differential W' production and decay at next-to-leading order in QCD", *Phys. Rev.* D66 (2002) 075011, arXiv:hep-ph/0207290 [hep-ph].

[59] E. Accomando, D. Becciolini, S. De Curtis, D. Dominici, L. Fedeli, *et al.*, "Interference effects in heavy W'-boson searches at the LHC", *Phys. Rev. D* 85 (2012) 115017, arXiv:1110.0713 [hep-ph].

[60] E. Boos, V. Bunichev, L. Dudko, and M. Perfilov, "Interference between W′ and W in single-top quark production processes", *Phys. Lett. B* 655 (2007) 245, arXiv:hep-ph/0610080 [hep-ph].

[61] C.-W. Chiang, N. Deshpande, X.-G. He, and J. Jiang, "The Family $\mathsf{SU}(2)_\mathfrak{l} \times \mathsf{SU}(2)_\mathfrak{h} \times \mathsf{U}(1)$ Model", *Phys. Rev. D* 81 (2010) 015006, arXiv:0911.1480 [hep-ph].

[62] L. Edelhäuser and A. Knochel, "Observing nonstandard W' and Z' through the third generation and Higgs lens", arXiv:1408.0914 [hep-ph].

[63] F. Zwicky, "Die Rotverschiebung von extragalaktischen Nebeln", *Helv. Phys. Acta* 6 (1933) 110–127.

[64] M. Bradac, D. Clowe, A. H. Gonzalez, P. Marshall, W. Forman, C. Jones, M. Markevitch, S. Randall, T. Schrabback, and D. Zaritsky, "Strong and weak lensing united. 3. Measuring the mass distribution of the merging galaxy cluster 1E0657-56", *Astrophys. J.* 652 (2006) 937–947, arXiv:astro-ph/0608408 [astro-ph].

[65] Planck Collaboration, "Planck 2015 results. XIII. Cosmological parameters", arXiv:1502.01589 [astro-ph.CO].

[66] M. Taoso, G. Bertone, and A. Masiero, "Dark Matter Candidates: A Ten-Point Test", *JCAP* 0803 (2008) 022, arXiv:0711.4996 [astro-ph].

[67] T. Marrodán Undagoitia and L. Rauch, "Dark matter direct-detection experiments", *J. Phys.* G43 no. 1, (2016) 013001, arXiv:1509.08767 [physics.ins-det].

[68] J. Goodman, M. Ibe, A. Rajaraman, W. Shepherd, T. M. P. Tait, and H.-B. Yu, "Constraints on Dark Matter from Colliders", *Phys. Rev.* D82 (2010) 116010, arXiv:1008.1783 [hep-ph].

[69] J. Conrad, "Indirect Detection of WIMP Dark Matter: a compact review", in *Interplay between Particle and Astroparticle physics (IPA2014) London, United Kingdom, August 18-22, 2014*. 2014. arXiv:1411.1925 [hep-ph].

[70] Fermi-LAT Collaboration, "Updated search for spectral lines from Galactic dark matter interactions with pass 8 data from the Fermi Large Area Telescope", *Phys. Rev.* D91 no. 12, (2015) 122002, arXiv:1506.00013 [astro-ph.HE].

[71] AMS Collaboration, "High Statistics Measurement of the Positron Fraction in Primary Cosmic Rays of 0.5500 GeV with the Alpha Magnetic Spectrometer on the International Space Station", *Phys. Rev. Lett.* 113 (2014) 121101.

[72] CMS Collaboration, "Search for dark matter, extra dimensions, and unparticles in monojet events in protonproton collisions at $\sqrt{s} = 8$ TeV", *Eur. Phys. J.* C75 no. 5, (2015) 235, arXiv:1408.3583 [hep-ex].

[73] CMS Collaboration, "Search for the production of dark matter in association with top-quark pairs in the single-lepton final state in proton-proton collisions at sqrt(s) = 8 TeV", *JHEP* 06 (2015) 121, arXiv:1504.03198 [hep-ex].

[74] CMS Collaboration, "Search for dark matter and unparticles produced in association with a Z boson in proton-proton collisions at $\sqrt{s} = 8$TeV", *Phys. Rev.* D93 no. 5, (2016) 052011, arXiv:1511.09375 [hep-ex].

[75] ATLAS Collaboration, "Search for dark matter candidates and large extra dimensions in events with a jet and missing transverse momentum with the ATLAS detector", *JHEP* 04 (2013) 075, arXiv:1210.4491 [hep-ex].

[76] ATLAS Collaboration, "Search for dark matter in events with a hadronically decaying W or Z boson and missing transverse momentum in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector", *Phys. Rev. Lett.* 112 no. 4, (2014) 041802, arXiv:1309.4017 [hep-ex].

[77] ATLAS Collaboration, "Search for dark matter in events with a Z boson and missing transverse momentum in pp collisions at $\sqrt{s}$=8 TeV with the ATLAS detector", *Phys. Rev.* D90 no. 1, (2014) 012004, arXiv:1404.0051 [hep-ex].

[78] ATLAS Collaboration, "Search for dark matter in events with heavy quarks and missing transverse momentum in pp collisions with the ATLAS detector", *Eur. Phys. J.* C75 no. 2, (2015) 92, arXiv:1410.4031 [hep-ex].

[79] DM Forum, "Dark Matter Benchmark Models for Early LHC Run-2 Searches: Report of the ATLAS/CMS Dark Matter Forum", arXiv:1507.00966 [hep-ex].

[80] Y. Bai and T. M. P. Tait, "Searches with Mono-Leptons", *Phys. Lett.* B723 (2013) 384–387, arXiv:1208.4361 [hep-ph].

[81] CMS Collaboration, "Search for leptonic decays of W' bosons in pp collisions at sqrt(s)=8 TeV", *CMS Physics Analysis Summary: EXO-12-060* (2013) . https://cds.cern.ch/record/1522476.

[82] N. F. Bell, Y. Cai, J. B. Dent, R. K. Leane, and T. J. Weiler, "Dark matter at the LHC: Effective field theories and gauge invariance", *Phys. Rev.* D92 no. 5, (2015) 053008, arXiv:1503.07874 [hep-ph].

[83] N. F. Bell, Y. Cai, and R. K. Leane, "Mono-W Dark Matter Signals at the LHC: Simplified Model Analysis", *JCAP* 1601 no. 01, (2016) 051, arXiv:1512.00476 [hep-ph].

[84] U. Haisch, F. Kahlhoefer, and T. M. P. Tait, "On Mono-W Signatures in Spin-1 Simplified Models", *Phys. Lett.* B760 (2016) 207–213, arXiv:1603.01267 [hep-ph].

[85] L. Evans and P. B. (editors), "LHC Machine", *Journal of Instrumentation* 3 no. 08, (2008) S08001. http://stacks.iop.org/1748-0221/3/i=08/a=S08001.

[86] C. Schwick, "CMS Daily Run Meeting with longest fill", 2016. https://hypernews.cern.ch/HyperNews/CMS/get/commissioning/6438.html.

[87] G. Papotti, "Lessons from LHC operation in 2015", 2016. https://indico.cern.ch/event/448109/contributions/1942059/attachments/1216261/1793941/2015overview_paper.pdf.

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

[89] CMS Collaboration, *The CMS magnet project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997. https://cds.cern.ch/record/331056.

[90] CMS Collaboration, *The CMS tracker system project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997. https://cds.cern.ch/record/368412.

[91] CMS Collaboration, A. Starodumov, "Operation of the cms pixel detector", Tech. Rep. CMS-CR-2011-187. CERN-CMS-CR-2011-187, CERN, Geneva, Sep, 2011.

[92] CMS Collaboration, "The performance of the CMS muon detector in proton-proton collisions at sqrt(s) = 7 TeV at the LHC", *JINST* 8 (2013) P11002, arXiv:1306.6905 [physics.ins-det].

[93] CMS Collaboration, Muon POG, "Tests of muon momentum resolution at high pT", 2015. https://indico.cern.ch/event/465390/.

[94] CMS Collaboration, "CMS: The electromagnetic calorimeter. Technical design report", 1997.

[95] CERN EDMS, "Technical drawings of the CMS experiment", 2016. https://edms5.cern.ch.

[96] CMS Collaboration, *The CMS hadron calorimeter project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997. https://cds.cern.ch/record/357153.

[97] CMS Collaboration, *The CMS muon project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997. https://cds.cern.ch/record/343814.

[98] CMS Collaboration, "CMS. The TriDAS project. Technical design report, vol. 1: The trigger systems", 2000.

[99] CMS Collaboration, "CMS: The TriDAS project. Technical design report, Vol. 2: Data acquisition and high-level trigger", 2002.

[100] CMS Collaboration, "CMS Strategies for tau reconstruction and identification using particle-flow techniques", *CMS Physics Analysis Summary: PFT-08-001* (2008) . https://cds.cern.ch/record/1198228.

[101] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET", *CMS Physics Analysis Summary: PFT-09-001* (2009) . https://cds.cern.ch/record/1194487.

[102] CMS Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS", *JINST* 6 (2011) P11002, arXiv:1107.4277 [physics.ins-det].

[103] S. Erdweg, *Study for Sensitivity for W' to tau nu with CMS*. Bachelor thesis, RWTH Aachen, Aachen, 2011. https://web.physik.rwth-aachen.de/~hebbeker/theses/erdweg_bachelor.pdf.

[104] W. Adam, R. Frühwirth, A. Strandlie, and T. Todor, "Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC", *CMS-NOTE-2005-001* (2005) .

[105] CMS Collaboration, "CMSSW configuration for the pfAlgoithm", 2016. https://github.com/cms-sw/cmssw/blob/CMSSW_5_3_X/RecoParticleFlow/Configuration/python/RecoParticleFlow_cff.py.

[106] CMS Collaboration, "Missing transverse energy performance of the CMS detector", *JINST* 6 (2011) P09001, arXiv:1106.5048 [physics.ins-det].

[107] CMS Collaboration, "Performance of tau-lepton reconstruction and identification in CMS", *JINST* 7 (2012) P01001, arXiv:1109.6034 [physics.ins-det].

[108] H.-J. Yang, B. P. Roe, and J. Zhu, "Studies of boosted decision trees for MiniBooNE particle identification", *Nucl. Instrum. Meth.* A555 (2005) 370–385, arXiv:physics/0508045 [physics].

[109] R. Fruhwirth, "Application of Kalman filtering to track and vertex fitting", *Nucl.Instrum.Meth.* A262 (1987) 444–450.

[110] G. Welch and G. Bishop, "An introduction to the kalman filter", Tech. Rep. TR 95-041, UNC-Chapel Hill, 1995. http://www.cs.unc.edu/~welch/kalman/kalmanIntro.html.

[111] CMS Collaboration, "Performance of reconstruction and identification of tau leptons in their decays to hadrons and tau neutrino in LHC Run-2", *CMS Physics Analysis Summary: TAU-16-002* (2016) . https://cds.cern.ch/record/2196972.

[112] CMS Collaboration, "Performance of tau reconstruction and identification in pp collisions at $\sqrt{s} = 8$ tev", Tech. Rep. CMS AN -2014/008, CERN, Geneva, May, 2014.

[113] CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker", *JINST* 9 no. 10, (2014) P10009, arXiv:1405.6569 [physics.ins-det].

[114] GEANT4 Collaboration, *Physics Reference Manual: GEANT4 v10.2*. Geant User Guides. CERN, Geneva, 2015. http://geant4.web.cern.ch/geant4/UserDocumentation/UsersGuides/PhysicsReferenceManual/fo/PhysicsReferenceManual.pdf.

[115] CMS Collaboration, Tau POG, "Performance Validation of the new MVA training using Spring15 MC sample", 2015. https://indico.cern.ch/event/446548/contributions/1953055/.

[116] CMS Collaboration, "Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV", *JINST* 10 no. 06, (2015) P06005, arXiv:1502.02701 [physics.ins-det].

[117] M. Cacciari and G. P. Salam, "Pileup subtraction using jet areas", *Phys. Lett.* B659 (2008) 119–126, arXiv:0707.1378 [hep-ph].

[118] D. Giordano and G. Sguazzoni, "An innovative seeding technique for photon conversion reconstruction at CMS", *J. Phys. Conf. Ser.* 396 (2012) 022015.

[119] CMS Collaboration, "Search for physics beyond the standard model in dilepton mass spectra in proton-proton collisions at $\sqrt{s} = 8$ TeV", *JHEP* 04 (2015) 025, arXiv:1412.6302 [hep-ex].

[120] CMS Collaboration, N. Neumeister, "Presentation on Global Muon reconstruction", 2012. https://indico.cern.ch/event/210563/contributions/418066/.

[121] CMS Collaboration, "CMSSW configuration for the high p_T muon identification", 2016. https://github.com/cms-sw/cmssw/blob/CMSSW_7_6_X/DataFormats/MuonReco/src/MuonCocktails.cc.

[122] M. Cacciari, G. P. Salam, and G. Soyez, "The Anti-k_T jet clustering algorithm", *JHEP* 0804 (2008) 063, arXiv:0802.1189 [hep-ph].

[123] CMS Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS", *JINST* 6 (2011) P11002, arXiv:1107.4277 [physics.ins-det].

[124] CMS Collaboration, "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV", *Submitted to: JINST* (2016) , arXiv:1607.03663 [hep-ex].

[125] Aachen 3A group, "ACSusy Analysis Framework", 2016. https://github.com/aachen3aCMS.

[126] I. Antcheva *et al.*, "ROOT: A C++ framework for petabyte data storage, statistical analysis and visualization", *Comput. Phys. Commun.* 180 (2009) 2499–2512, arXiv:1508.07749 [physics.data-an].

[127] CMS Collaboration, "Model Unspecific Search for New Physics in pp Collisions at sqrt(s) = 7 TeV", *CMS Physics Analysis Summary: EXO-10-021* (2011) . https://cds.cern.ch/record/1360173.

[128] H. P. Bretz *et al.*, "A Development Environment for Visual Physics Analysis", *JINST* 7 (2012) T08005, arXiv:1205.4912 [physics.data-an]. http://vispa.physik.rwth-aachen.de/.

[129] Aachen 3A group, "Three A Physics Analysis Software", 2016. https://gitlab.cern.ch/aachen-3a/tapas.

[130] CMS Collaboration, "CMSSW GIT reposetory", 2016. https://github.com/cms-sw/cmssw.

[131] K. Padeken, "GIT reposetory for the Run I analyis of the W' signal in $\tau + E_T^{miss}$", 2016. https://github.com/padeken/WprimeTau8TeV.

[132] K. Padeken, "GIT reposetory for the Run II analyis of the W' signal in $\tau + E_T^{miss}$", 2016. https://gitlab.cern.ch/aachen-3a/wprime-analyser/tree/qcd_cleanup.

[133] CMS Collaboration, "Measurement of the underlying event in the Drell-Yan process in proton-proton collisions at $\sqrt{s} = 7$ TeV", *Eur. Phys. J.* C72 (2012) 2080, arXiv:1204.1411 [hep-ex].

[134] GEANT4 Collaboration, "GEANT4—a simulation toolkit", *Nucl. Instrum. Meth. A* 506 (2003) 250.

[135] N. Davidson, G. Nanava, T. Przedzinski, E. Richter-Was, and Z. Was, "Universal Interface of TAUOLA Technical and Physics Documentation", *Comput. Phys. Commun.* 183 (2012) 821, arXiv:1002.0543 [hep-ph].

[136] CMS Collaboration, "Utilities for Accessing Pileup Information for Data. TWiki Website, 2012.", *https://twiki.cern.ch/twiki/bin/viewauth/CMS/PileupJSONFileforData2012_Pileup_JSON_Files* (2014) .

[137] CMS Collaboration, "Standard Model Cross Sections for CMS at 8 TeV", *https://twiki.cern.ch/twiki/bin/viewauth/CMS/StandardModelCrossSectionsat8TeV* (2014) .

[138] G. Balossini, G. Montagna, C. Carloni Calame, M. Moretti, O. Nicrosini, F. Piccinini, M. Treccani, and A. Vicini, "Combination of electroweak and qcd corrections to single w production at the fermilab tevatron and the cern lhc", *JHEP* 01 (2010) 013, arXiv:0907.0276.

[139] J. Butterworth, G. Dissertori, S. Dittmaier, D. de Florian, N. Glover, *et al.*, "Les Houches 2013: Physics at TeV Colliders: Standard Model Working Group Report." 2014.

[140] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, "W Physics at the LHC with FEWZ 2.1", *Comput. Phys. Commun.* 184 (2013) 208, arXiv:1201.5896 [hep-ph].

[141] S. Catani, L. Cieri, G. Ferrera, D. de Florian, and M. Grazzini, "Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO", *Phys.Rev.Lett.* 103 (2009) 082001, arXiv:0903.2120 [hep-ph].

[142] C. M. Carloni Calame, G. Montagna, O. Nicrosini, and A. Vicini, "Precision electroweak calculation of the charged current Drell-Yan process", *JHEP* 12 (2006) 016, arXiv:hep-ph/0609170 [hep-ph].

[143] J. Pumplin, A. Belyaev, J. Huston, D. Stump, and W. Tung, "Parton distributions and the strong coupling: CTEQ6AB PDFs", *JHEP* 0602 (2006) 032, arXiv:hep-ph/0512167 [hep-ph].

[144] Kim, DongHee and Oh, Youngdo and Yang, Yuchul and Lee, Jeongeun, "PDF Uncertainties and K-factor for the W Search at 8 TeV Collisions", *CMS AN-12-172* (2012) .

[145] CMS Collaboration, Jet MET POG, "CMS MET Filters TWiki Website, 2016.", 2016. https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETOptionalFilters.

[146] CMS Collaboration, Jet MET POG, "CMS MET Filters Run II TWiki Website, 2016.", 2016. https://twiki.cern.ch/twiki/bin/view/CMS/MissingETOptionalFiltersRun2.

[147] CMS Collaboration, Tau POG, "TauID: recommendation from the Tau POG ", 2014. https://twiki.cern.ch/twiki/bin/viewauth/CMS/TauIDRecommendation.

[148] CMS Collaboration, Jet MET POG, "Jet energy scale uncertainty sources. TWiki Website, 2016.", 2016. https://twiki.cern.ch/twiki/bin/view/CMS/JECUncertaintySources.

[149] CMS Collaboration, "Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in pp Collisions at $\sqrt{s} = 7$ TeV", *JINST* 8 (2013) P09009, arXiv:1306.2016 [hep-ex]. [JINST8,9009(2013)].

[150] CMS Collaboration, "CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update", *CMS Physics Analysis Summary: LUM-13-001* (2013) . http://cdsweb.cern.ch/record/1598864.

[151] NNPDF Collaboration, R. D. Ball *et al.*, "Parton distributions for the LHC Run II", *JHEP* 04 (2015) 040, arXiv:1410.8849 [hep-ph].

[152] A. Accardi, L. T. Brady, W. Melnitchouk, J. F. Owens, and N. Sato, "Constraints on large-x parton distributions from new weak boson production and deep-inelastic scattering data", arXiv:1602.03154 [hep-ph].

[153] ZEUS, H1 Collaboration, "Combination of measurements of inclusive deep inelastic $e^{\pm}p$ scattering cross sections and QCD analysis of HERA data", *Eur. Phys. J.* C75 no. 12, (2015) 580, arXiv:1506.06042 [hep-ex].

[154] J. Butterworth *et al.*, "PDF4LHC recommendations for LHC Run II", *J. Phys.* G43 (2016) 023001, arXiv:1510.03865 [hep-ph].

[155] The PDF4LHC working group, "http://www.hep.ucl.ac.uk/pdf4lhc/", 2016.

[156] M. Botje, J. Butterworth, A. Cooper-Sarkar, A. de Roeck, J. Feltesse, *et al.*, "The PDF4LHC Working Group Interim Recommendations", `arXiv:1101.0538 [hep-ph]`.

[157] CMS Collaboration, "CMS Luminosity Measurement for the 2015 Data Taking Period", *CMS Physics Analysis Summary: LUM-15-001* (2016) . `https://cds.cern.ch/record/2138682`.

[158] ATLAS, CMS Collaboration, "Procedure for the LHC Higgs boson search combination in summer 2011", *ATL-PHYS-PUB-2011-011,CMS-NOTE-2011-005* (2011) .

[159] ATLAS, CMS Collaboration, "Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments", *Phys. Rev. Lett.* 114 (2015) 191803, `arXiv:1503.07589 [hep-ex]`.

[160] RooStats Collaboration, "RooStats", *https://twiki.cern.ch/twiki/bin/view/RooStats/WebHome* (2016) .

[161] Y. G. Kim and K. Y. Lee, "Early LHC bound on W' boson in the nonuniversal gauge interaction model", *Phys. Lett. B* 706 (2012) 367, `arXiv:1105.2653 [hep-ph]`.

[162] CMS Collaboration, S. Chatrchyan *et al.*, "Search for W' $\to$ tb decays in the lepton + jets final state in pp collisions at $\sqrt{s} = 8$ TeV", *JHEP* 5 (2014) 108, `arXiv:1402.2176 [hep-ex]`.

[163] Kim, DongHee and Lee, Jeongeun and Oh, Youngdo and Yang, Yu Chul, "PDF Uncertainties and K-factor for the W Search at 13 TeV Collisions", *CMS AN-14-240* (2014) .

[164] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J.* C71 (2011) 1554, `arXiv:1007.1727 [physics.data-an]`.

[165] CMS Collaboration, "Search for W' $\to$ tb decays in the lepton + jets final state in pp collisions at $\sqrt{s} = 8$ TeV", *JHEP* 05 (2014) 108, `arXiv:1402.2176 [hep-ex]`.

[166] CMS Collaboration, "Search for W$'$ boson resonances decaying into a top quark and a bottom quark in the leptonic final state at $\sqrt{s} = 13$ TeV", *CMS Physics Analysis Summary: CMS-PAS-B2G-15-004* (2015) . `http://cds.cern.ch/record/2114803`.

[167] ATLAS Collaboration, "Search for $W' \to t\bar{b}$ in the lepton plus jets final state in proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 8$ TeV with the ATLAS detector", *Phys. Lett.* B743 (2015) 235–255, `arXiv:1410.4103 [hep-ex]`.

[168] ATLAS Collaboration, "Search for $W' \to t\bar{b}$ in the lepton plus jets final state in proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 8$ TeV with the ATLAS detector", *Phys. Lett.* B743 (2015) 235–255, `arXiv:1410.4103 [hep-ex]`.

[169] ATLAS Collaboration, "Search for new particles in events with one lepton and missing transverse momentum in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector", *JHEP* 9 (2014) 37, `arXiv:1407.7494 [hep-ex]`.

[170] ATLAS Collaboration, "Search for new resonances in events with one lepton and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *Phys. Lett.* B762 (2016) 334–352, arXiv:1606.03977 [hep-ex].

[171] ATLAS Collaboration, "Search for new phenomena in the dijet mass distribution using $p - p$ collision data at $\sqrt{s} = 8$ TeV with the ATLAS detector", *Phys. Rev.* D91 no. 5, (2015) 052007, arXiv:1407.1376 [hep-ex].

[172] ATLAS Collaboration, "Search for new phenomena in dijet mass and angular distributions from pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *Phys. Lett.* B754 (2016) 302–322, arXiv:1512.01530 [hep-ex].

[173] CMS Collaboration, "Search for resonances and quantum black holes using dijet mass spectra in proton-proton collisions at $\sqrt{s} = 8$ TeV", *Phys. Rev.* D91 no. 5, (2015) 052009, arXiv:1501.04198 [hep-ex].

[174] CMS Collaboration, "Search for narrow resonances in dilepton mass spectra in proton-proton collisions at $\sqrt{s} = 13$ TeV and combination with 8 TeV data", arXiv:1609.05391 [hep-ex].

[175] CMS Collaboration, "Search for narrow resonances in dilepton mass spectra in proton-proton collisions at $\sqrt{s} = 13$ TeV and combination with 8 TeV data", arXiv:1609.05391 [hep-ex].

[176] ATLAS Collaboration, "Search for high-mass dilepton resonances in pp collisions at $\sqrt{s} = 8$TeV with the ATLAS detector", *Phys. Rev.* D90 no. 5, (2014) 052005, arXiv:1405.4123 [hep-ex].

[177] ATLAS Collaboration, "Search for high-mass new phenomena in the dilepton final state using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *Phys. Lett.* B761 (2016) 372–392, arXiv:1607.03669 [hep-ex].

[178] ATLAS Collaboration, "A search for high-mass ditau resonances decaying in the fully hadronic final state in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector", *ATLAS CONF note: ATLAS-CONF-2013-066* (Jul, 2013) . http://cds.cern.ch/record/1562841.

[179] CMS Collaboration, "Search for new physics with high-mass tau lepton pairs in pp collisions at sqrt(s) = 13 TeV with the CMS detector", *CMS Physics Analysis Summary: PAS-EXO-16-008* (2016) . http://cds.cern.ch/record/2160363.

Acknowledgements

I want to thank all the people who made this thesis possible. First I want to thank Prof. Hebbeker, who made it possible for me to work on this thesis by providing guidance and knowledge throughout this research. I also want to thank Prof. Lutz Feld, agreeing to be the second referee of this thesis.

A special thanks also goes to Dr. Kerstin Hoepfner. She not only managed the W' group, but also guided every discussion and navigated the stormy waters of publication. I also want to thank all the members of the W' team Fabian Bispink, Sören Erdweg, Simon Knutzen, Viktor Kutzner, Philipp Millet, Swagata Mukherjee and Mark Olschewski. We have worked as a team and made this analysis into a success. To our newcomers Marcel Materok and Sebasian Wiedenbeck good luck and strong nerves.

I also want to thank the men and women from the tau POG, who worked with me on the high p_T taus and made it possible to have this analysis.

I also want to thank the other members of this institute, who not only made the work atmosphere light, but also were always up for the after work activities. Not only was every door always open for discussion, but you also provided the constant incentive for better results. I want to name especially those who I have been working with: Andreas Albert, Michael Brodski, Deborah Duchardt, Matthias Endres, Thomas Esch, Andreas Güth, Markus Merschmeyer, Arnd Meyer, Paul Papacz, Holger Pieta, Tobias Pook, Markus Radziej, Stefan Schmitz, Daniel Theyssier, Sebastian Thüer and Martin Weber.

I also want to thank the theorists who helped me with a lot of background knowledge and scientific eagerness: Leila Ali Cavasonza, Lisa Edelhäuser, Alexander Knochel, Alexander Mück, Lennart Oymanns. The non universal and dark matter models would not have been possible without your help. Also I want to thank all those who contributed to our joint experimental and theoretical meetings, it was a great institution.

I also want to thank all the members of the CMS experiment, who made it possible to record the data and provide the resources and support to perform such a high level analysis. The decades of work building this machine have resulted in a beautiful experiment and a wonderful collaboration.

At this place I also want to apologize to all those who had to endure my English.

At last I want to thank my family for their constant support. My wife Regina endured me over all this time and provided the love and backing needed to make this possible. Finally to my son Lukas: if you can read this you have learned English, congratulations.