Search for third-generation scalar leptoquarks decaying into a top quark and a tau lepton with the CMS detector

(Suche nach skalaren Leptoquarks der dritten Generation im Zerfallskanal Top-Quark und Tau-Lepton mit dem CMS-Detektor)

von

Mareike Meyer

geboren am

29. Juni 1988

Master-Arbeit im Studiengang Physik

Universität Hamburg

2013

1.Gutachter: Prof. Dr. Johannes Haller

2.Gutachter: Prof. Dr. Peter Schleper

Abstract

In this thesis a search for pair production of third-generation scalar leptoquarks is presented. Leptoquarks are new bosons predicted by many theories beyond the Standard Model. They couple to a quark and a lepton. Due to constraints from flavor changing neutral currents and other rare processes it is predicted that leptoquarks decay only into particles of the same Standard Model generation. In this thesis the decay channel into a top quark and a tau lepton is studied. For the analysis the complete data sample from the year 2012 recorded by the CMS experiment at the LHC at a centre-of-mass energy of 8 TeV is used, corresponding to a total integrated luminosity of 19.6 fb^{-1}.

Since a search in this channel has never been done before different selections based on at least one muon candidate and one hadronically decaying tau lepton are studied in order to define a signal region with high sensitivity. After maximizing the sensitivity by calculating expected limits on the production cross section times branching ratio, it is found that the distribution of the transverse momentum of the leading tau lepton shows high sensitivity for different leptoquark masses and this distribution is used for the final interpretation of the result.

No excess over the Standard Model expectation is observed and exclusion limits on the cross section times branching ratio are set. Third-generation leptoquarks decaying to a top quark and a tau lepton are excluded for masses up to 582 $\mathrm{GeV}/\mathrm{c}^2$ at 95% C.L.

Zusammenfassung

In dieser Masterarbeit wird eine Suche nach Paarproduktion von skalaren Leptoquarks der dritten Generation vorgestellt. Leptoquarks sind neue Bosonen, welche von vielen Theorien jenseits des Standardmodells vorhergesagt werden. Sie koppeln an ein Quark und ein Lepton. Durch Grenzen aus Flavor-ändernden neutralen Strömen und anderen seltenen Zerfällen ist vorhergesagt, dass Leptoquarks nur in Teilchen der gleichen Standardmodell-Generation zerfallen. In dieser Arbeit wird der Zerfallskanal in ein Top-Quark und ein Tau-Lepton untersucht. Dazu wird der vollständige im Jahr 2012 vom CMS-Experiment am LHC bei einer Schwerpunktsenergie von 8 TeV aufgenommene Datensatz verwendet. Dies entspricht einer integrierten Luminosität von 19.6 fb^{-1}.

Da bisher in diesem Kanal noch keine Suche durchgeführt wurde, werden zunächst verschiedene Selektionen basierend auf mindestens einem Myonkandidaten und einem hadronisch zerfallenden Tau-Leptonkandidaten studiert, um eine Signalregion mit hoher Sensitivität zu definieren. Nach dem Maximieren der Sensitivität durch die Berechnung von erwarteten Ausschlussgrenzen auf den Wirkungsquerschnitt multipliziert mit dem Verzweigungsverhältnis, zeigt sich, dass die Verteilung des Transversalimpulses der Tau-Leptonen hohe Sensitivität für verschiedene Leptoquarkmassen aufweist. Diese Verteilung wird daher für die finale Interpretation der Ereignisse verwendet.

Es wird kein Überschuss der Daten gegenüber der Standardmodell-Erwartung beobachtet und Ausschlussgrenzen auf den Wirkungsquerschnitt multipliziert mit dem Verzweigungsverhältnis werden berechnet. Leptoquarks mit Massen unter 582 GeV/c^2 sind ausgeschlossen mit 95% C.L.

Contents

Chapter 1

Introduction

The Large Hadron Collider (LHC) is a proton-proton-collider located at CERN [1], where collisions at a centre-of-mass energy of maximal $\sqrt{s} = 14$ TeV and instantaneous luminosities of up to 10^{-34} cm^{-2}s^{-1} can be studied.
The main goals of the LHC are tests and precision studies of the Standard Model, the search for the Higgs-boson and the search for physics beyond the Standard Model. The latter are motivated by questions which are unanswered by the Standard Model, despite the fact that it is in general very successful in explaining all particles which occur in nature and the interactions between them. Examples for these unanswered questions are the asymmetry between matter and antimatter we observe today, the fact that gravity is not included in the Standard Model and that the forces included in the Standard Model cannot be unified so far. Thus many theories beyond the Standard Model are discussed, which address these questions. Many of them predict the existence of leptoquarks, which are new bosons coupling to Standard Model quarks and leptons of the same generation. Such a relationship between quarks and leptons is further motivated by the symmetry of the quark and lepton generations in the Standard Model and by the relationship between the electromagnetic charges of the quarks and the charged leptons.
In this thesis a search for third-generation leptoquarks decaying into a top quark and a tau lepton is performed. Proton-proton-collision data are analyzed which were recorded in the year 2012 by the CMS experiment, one of the multi-purpose experiments at the LHC. The data were taken at a centre-of-mass energy of $\sqrt{s} = 8$ TeV and instantaneous luminosities of up to $7.67 \cdot 10^{33}$ cm^{-2}s^{-1}. In total 19.6 fb^{-1} of data were collected in 2012.
This thesis starts with an introduction of the theoretical aspects (chapter 2). The Standard Model is introduced, the unanswered questions of the Standard Model are addressed, possible theories beyond the Standard Model are mentioned and an introduction into leptoquarks is given. Additionally, the process studied is introduced and the background processes for this analysis are presented. In chapter 3 the LHC and the CMS experiment are introduced. In the following chapter (chapter 4) the particle-flow event reconstruction algorithm used in CMS and the reconstruction and identification of particles in the final state are explained. In chapter 5 the performed search is presented. The last chapter contains a short summary and conclusion.

[1] "Centre européenne pour la Recherche nucléaire", now "European Organization for Nuclear Research"

Chapter 2

Theoretical aspects

The Standard Model of particle physics aims to describe all elementary particles in nature and the interactions between them. It is very successful in explaining all measurements of particle physics. Moreover, it made many predictions, e.g. the existence of the top quark or the tau neutrino, which were experimentally verified later. However, there are still some unanswered questions. This chapter gives a short introduction of the Standard Model and the unanswered questions (section 2.1 and section 2.2). In chapter 2.3 leptoquarks, which are predicted by many extensions of the Standard Model, are introduced. Afterwards the decay channel of third-generation leptoquarks studied in this thesis is presented and the Standard Model backgrounds for this search are described (section 2.4).

2.1 The Standard Model

The Standard Model [1] is formulated as a quantum field theory based on the local gauge group

$$\mathrm{U}(1)_\mathrm{Y} \otimes \mathrm{SU}(2)_\mathrm{L} \otimes \mathrm{SU}(3)_\mathrm{C}. \tag{2.1}$$

With this group three of the four forces which occur in nature can be explained. The three forces included in the Standard Model are the strong force (QCD), the weak force (Weinberg-Salam-Glashow model) and the electromagnetic force (QED), with the corresponding symmetry groups $\mathrm{SU}(3)_\mathrm{C}$, $\mathrm{SU}(2)_\mathrm{L}$ and $\mathrm{U}(1)_\mathrm{Y}$. The electromagnetic force and the weak force can be unified in one theory: the electroweak theory. Only the gravitational force is not included in the Standard Model.

In the Standard Model interactions between particles are explained via the exchange of spin-1 gauge bosons. For the electromagnetic force the exchange particle is the photon. It mediates interactions between particles which carry an electromagnetic charge Q. Since photons are massless, the electromagnetic force has an infinite range.

The weak force is mediated by two massive bosons: the neutral Z^0- and the charged $\mathrm{W}^\pm$-boson. With masses of $91.2\ \mathrm{GeV/c^2}$ and $80.4\ \mathrm{GeV/c^2}$ respectively, these bosons are quite heavy so that the weak force has a very short range. Each particle carrying weak isospin T_3 is interacting via the weak force. In the electroweak theory the charges of the weak and electromagnetic force

force	exchange bosons	charge	gauge group	Quantum field theory
electromagnetic force	photons	electric charge	$U(1)_Y$	QED
weak force	Z^0,W^+,W^-	weak isospin	$SU(2)_L$	electroweak theory
strong force	8 gluons	color	$SU(3)_C$	QCD

Table 2.1: Summary of the forces included in the Standard Model.

are combined to the hypercharge Y, which is defined as $Y = 2(Q - T_3)$.
The strong force is mediated by eight gluons. These are massless and carry color, which is the charge of QCD. Only colored particles can interact via the strong force. Since the gluons themselves carry color, they interact with each other, which has important consequences, as explained later. A summary of the forces included in the Standard Model is given in table 2.1.
The Standard Model contains a second group of particles, which are the constituents of matter. All these particles are fermions with spin $\frac{1}{2}$. They are divided into two different groups: leptons and quarks. The leptons and quarks can be further arranged into three generations. Each lepton generation consists of a particle with charge $-e$ and the corresponding neutral neutrino. A quark generation consists of a quark with charge $+\frac{2}{3}e$ and one quark with charge $-\frac{1}{3}e$. In general the generations share the same properties, only the mass increases from the first to the third generation. The fermion content of the Standard Model is summarized in table 2.2.
Only the quarks carry color and are thus interacting via the strong force. Except for neutrinos, all particles carry electromagnetic charge and interact electromagnetically. All Standard Model particles have a weak isospin quantum number and therefore participate in the weak force. For each Standard Model particle exists an anti-particle, which carries the same quantum numbers but has the opposite electromagnetic charge.
Quarks are bound via the strong force to hadrons, which can be further divided into baryons and mesons. Mesons are bound states of a quark and an antiquark and baryons consist of three quarks, such that color neutral objects are formed.
The strong force has only one free parameter, which is the coupling constant α_s. α_s gets larger with increasing distance between quarks due to the self interaction of gluons. Thus at high energies (~small distance scales) quarks behave as quasi-free particles, which is called "asymptotic freedom". However, the larger the distance between two quarks gets, the stronger the coupling becomes. In high energy interactions, where quarks and gluons get separated through the parton scattering process, the potential of the color field can get very large. It becomes energetically favorable to produce new quark-antiquark-pairs, which results in the production of hadrons. This process is called hadronization. Since the hadrons have a low transverse momentum with respect to the initial parton, they are very collimated and a jet is produced. Thus quarks cannot be separated and can never be found as free-states, which is known as "confinement". An exception is the top quark, which has a very short lifetime ($\approx 10^{-24}$ s) and decays before it hadronizes.
The weak force acts differently on left-handed than on right-handed particles. In the limit of very high velocities left-handed means that the spin of the particle and the vector of its momentum have opposite directions, right-handed means these two vectors point in the same direction. Thus the weak force violates parity(P)-symmetry and CP-symmetry. CP-symmetry means that

	1st generation	2nd generation	3rd generation	charge Q	weak isospin T_3	colored?
leptons	e	μ	τ	$-e$	$+\frac{1}{2}$	✗
	ν_e	ν_μ	ν_τ	0	$-\frac{1}{2}$	✗
quarks	u	c	t	$+\frac{2}{3}e$	$+\frac{1}{2}$	✓
	d	s	b	$-\frac{1}{3}e$	$-\frac{1}{2}$	✓

Table 2.2: Particle content of the Standard Model.

an interaction should be invariant if the particle is replaced with its anti-particle (charge(C)-symmetry) if at the same time the parity transformation ($\vec{r} \to -\vec{r}$) is performed.
Another important feature of the weak force is that the quark eigenstates of the weak force are not the quark mass eigenstates. The convention is to rotate the down-type quarks by a 3×3-matrix, the CKM [1]-matrix. The squared entries of the CKM-matrix give the probability of the transition of one quark flavor to another. The CKM-matrix is given by [2]

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.974 & 0.225 & 0.004 \\ 0.225 & 0.973 & 0.041 \\ 0.009 & 0.040 & 0.999 \end{pmatrix}. \tag{2.2}$$

In local gauge theories all particles have to be massless since mass terms are not invariant under gauge transformations. This would lead to a contradiction with the experimental fact that the fermions and the Z- and W-bosons have masses. Thus a mechanism is needed to explain the masses of the particles. In the Standard Model this mechanism is the Higgs-mechanism, which generates particle masses through electroweak symmetry breaking. The particle which is associated to this mechanism is the Higgs boson. It is a spin 0 particle and was for a long time the only missing particle of the Standard Model. On the 4th of July in 2012 the two general-purpose experiments of the LHC, ATLAS and CMS, claimed the discovery of a new boson with a mass of approximately 125 $\mathrm{GeV/c^2}$, see [3, 4]. This new boson seems to have all of the predicted properties of the Standard Model Higgs boson. It has still to be verified with future data if it is exactly the Standard Model Higgs boson.

Although the Standard Model is very successful in describing all observed phenomena, it leaves some fundamental questions unanswered. These are among others:

- **Gravity:** The gravitational force is not included in the Standard Model.
- **Hierarchy problem and fine-tuning:** When extrapolating the Standard Model to high energies e.g. to the Planck scale the hierarchy problem arises. It can be seen e.g. in the loop corrections of the Higgs mass, which are (in one-loop order)

$$\Delta m_H^2 = -\frac{|\lambda_f|}{8\pi^2}\left[\Lambda^2 - 3m_f \ln\left(\frac{\Lambda}{m_f}\right)\right]. \tag{2.3}$$

  Here λ_f is the coupling of the Higgs to a fermion and m_f is the mass of the fermion. The variable Λ is the scale up to which the Standard Model should be valid. If this scale is

[1] Cabibbo-Kobayashi-Maskawa matrix

large, the loop corrections to the Higgs mass get large too. For high scales the corrections become larger than the Higgs mass itself, which is physically unfavorable.
Moreover, the corrections from the Standard Model bosons and fermions have to cancel to a high degree in order to extrapolate to high energies. This is known as the fine-tuning problem.

- **Asymmetry between matter and antimatter:** In the big bang matter and antimatter were produced in equal amounts, but in the known universe there is more matter than antimatter. This asymmetry can only be explained with the existence of CP violation. Moreover, a phase in the development of the universe is needed, which is called inflation, where the expansion of the universe was very fast. While it is generally accepted that there was a phase like this in the development of our universe, the observed CP violation is much smaller than needed to explain the asymmetry.

- **Dark matter and dark energy:** Measurements of rotation curves of galaxies, X-ray observations of gas bound in galaxies and studies of gravitational lensing and of structure formation in the development of the universe prove the existence of dark matter. From these observations it can be concluded that only approximately 5% of the total matter in the universe is made out of Standard Model particles. 23% is Dark Matter and the rest is Dark Energy. The Standard Model has no particle candidate for Dark Matter and no explanation for Dark Energy.

2.2 Possible extensions of the Standard Model

To solve the problems described above, many extensions of the Standard Model are discussed. These are for example:

- **Supersymmetry:** Supersymmetry (SUSY) postulates for each Standard Model fermion a new partner boson and for each Standard Model boson a partner fermion. The partners have equal quantum numbers, only the spin differs by $\frac{1}{2}$. Since none of these new particles have been observed yet, SUSY has to be a broken symmetry. This means that the masses of the supersymmetric particles have to be larger than the masses of their partners in the Standard Model. SUSY solves the hierarchy problem since supersymmetric particles lead to new contributions in the loop corrections of the Higgs mass. These new loops have opposite signs than the corresponding contributions from loops the Standard Model particles, and the different contributions cancel. Furthermore, SUSY predicts a unification of the coupling constants and the lightest supersymmetric particle is a possible dark matter candidate in most theories. Details about SUSY can be found in [5].

- **Grand Unified Theories:** In Grand Unified Theories (GUTs) [1] the three local gauge groups of the Standard Model are derived from one gauge group G_{GUT},

$$G_{GUT} \supset \mathrm{U}(1)_{\mathrm{Y}} \otimes \mathrm{SU}(2)_{\mathrm{L}} \otimes \mathrm{SU}(3)_{\mathrm{C}}. \tag{2.4}$$

This means that not only the weak and the electromagnetic interaction are unified, but all three forces can be described through one interaction. Since GUTs are described through larger gauge groups, the particle content of the Standard Model is arranged in larger multiplets than in the Standard Model. Thus quarks and leptons appear in common multiplets. This gives an explanation for the relationship between their charges. Moreover, GUTs give possible solutions for the baryon asymmetry in the universe through the prediction of C, CP and baryon number violation.

- **Compositeness:** In compositeness theories [6] quarks and leptons, and sometimes the massive gauge bosons are bound states of new particles, so called preons. These carry a new charge, which is called hypercolor. This charge leads, like the color in QCD, to bound states through confinement. In compositeness models less elementary particles and thus less free parameters than in the Standard Model are predicted. In addition, compositeness establishes a relationship between quarks and leptons through the common preons. It also could give an explanation for the three generations through the assumption that the second and third-generations are excited states of the first one.

- **Technicolor:** In technicolor models [7, 8, 9] no fundamental scalar field is allowed. Instead it is predicted that all scalar fields are bound states of technifermions. These new fermions carry a new charge, so called technicolor, which leads to bound states through confinement. The masses of the gauge bosons are created by global symmetry breaking of the technicolor theory, which is an alternative for the electroweak symmetry breaking in the Standard Model. The fermions get their masses through new four-fermion gauge interactions between two Standard Model fermions and two technifermions.
  Technicolor models provide possible solutions of the baryon asymmetry and they could reduce the number of unknown parameters of the Standard Model by removing the Yukawa couplings, the vacuum expectation value of the scalar and provide a solution for the hierarchy problem [10]. After the discovery of the new boson, extensions of the Technicolor models were introduced which can still explain the existence of the new boson.

Many of the mentioned models beyond the Standard Model predict the existence of leptoquarks. These new particles are introduced in the next sections.

2.3 Leptoquarks

The symmetry between the quark and lepton generations in the Standard Model and the connection of their charges suggest a fundamental relationship between quarks and leptons. In unified theories new bosons which decay into a pair of a lepton and a quark are postulated. These bosons are called leptoquarks. They can occur as scalar particles, which means that their spin is zero, or as vector particles, with spin one. They have lepton and baryon numbers and carry a fractional electromagnetic charge. Leptoquarks are color-triplets. Similar to the three generations of the Standard Model, three generations of leptoquarks are predicted. A summary of all possible leptoquark states in a technicolor model is given in table 2.3. More leptoquark

spin	$SU(3)_C$	T_3	Y	$Q[e]$	decay mode
scalar	$\overline{3}$	0	$\frac{1}{3}$	$\frac{1}{3}$	$\overline{\tau}_R\bar{\mathrm{t}}_R$, $\overline{\tau}_L\bar{\mathrm{t}}_L$, $\overline{\nu}_{\tau,L}\overline{\mathrm{b}}_L$
	$\overline{3}$	0	$\frac{4}{3}$	$\frac{4}{3}$	$\overline{\tau}_R\overline{\mathrm{b}}_R$
	$\overline{3}$	$+1$	$\frac{1}{3}$	$\frac{4}{3}$	$\overline{\tau}_L\overline{\mathrm{b}}_L$
	$\overline{3}$	0	$\frac{1}{3}$	$\frac{1}{3}$	$\overline{\tau}_L\bar{\mathrm{t}}_L$, $\overline{\nu}_{\tau,L}\overline{\mathrm{b}}_L$
	$\overline{3}$	-1	$\frac{1}{3}$	$-\frac{2}{3}$	$\overline{\nu}_{\tau,L}\bar{\mathrm{t}}_L$
	3	$+\frac{1}{2}$	$\frac{7}{6}$	$\frac{5}{3}$	$\mathrm{t_R}\overline{\tau}_\mathrm{L}$, $\mathrm{t_L}\overline{\tau}_\mathrm{R}$
	3	$-\frac{1}{2}$	$\frac{7}{6}$	$\frac{2}{3}$	$\mathrm{b_L}\overline{\tau}_\mathrm{R}$, $\mathrm{t_R}\overline{\nu}_{\tau,\mathrm{L}}$
	3	$+\frac{1}{2}$	$\frac{1}{6}$	$\frac{2}{3}$	$\mathrm{b_R}\overline{\tau}_\mathrm{L}$
	3	$-\frac{1}{2}$	$\frac{1}{6}$	$-\frac{1}{3}$	$\mathrm{b_R}\overline{\nu}_{\tau,\mathrm{L}}$
vector	3	0	$\frac{2}{3}$	$\frac{2}{3}$	$\mathrm{t_R}\overline{\nu}_{\tau,\mathrm{L}}$, $\mathrm{b_R}\overline{\tau}_\mathrm{L}$, $\mathrm{b_L}\overline{\tau}_\mathrm{R}$
	3	0	$\frac{5}{3}$	$\frac{5}{3}$	$\mathrm{t_R}\overline{\tau}_\mathrm{L}$, $\mathrm{t_L}\overline{\tau}_\mathrm{R}$
	3	$+1$	$\frac{2}{3}$	$\frac{5}{3}$	$\mathrm{t_R}\overline{\tau}_\mathrm{L}$, $\mathrm{t_L}\overline{\tau}_\mathrm{R}$
	3	0	$\frac{2}{3}$	$\frac{2}{3}$	$\mathrm{t_R}\overline{\nu}_{\tau,\mathrm{L}}$, $\mathrm{b_L}\overline{\tau}_\mathrm{R}$, $\mathrm{b_R}\overline{\tau}_\mathrm{L}$
	3	-1	$\frac{2}{3}$	$-\frac{1}{3}$	$\mathrm{b_R}\overline{\nu}_{\tau,\mathrm{L}}$
	$\overline{3}$	$+\frac{1}{2}$	$\frac{5}{6}$	$\frac{4}{3}$	$\bar{b}_L\overline{\tau}_L$, $\bar{\mathrm{b}}_R\overline{\tau}_R$
	$\overline{3}$	$-\frac{1}{2}$	$\frac{5}{6}$	$\frac{1}{3}$	$\bar{\mathrm{b}}_L\overline{\nu}_{\tau,L}$, $\bar{\mathrm{t}}_R\overline{\tau}_R$, $\bar{\mathrm{t}}_L\overline{\tau}_L$
	$\overline{3}$	$+\frac{1}{2}$	$-\frac{1}{6}$	$\frac{1}{3}$	$\bar{\mathrm{t}}_L\overline{\tau}_L$, $\bar{\mathrm{t}}_R\overline{\tau}_R$
	$\overline{3}$	$-\frac{1}{2}$	$-\frac{1}{6}$	$-\frac{2}{3}$	$\bar{\mathrm{t}}_L\overline{\nu}_{\tau,L}$

Table 2.3: Summary of possible leptoquark states in a technicolor model and their quantum numbers, taken from [11]. Shown are the spin, the representation in QCD, the weak isospin T_3 of the leptoquark, the weak hypercharge Y, which is defined by $Y = 2(Q - T_3)$, the electromagnetic charge Q, and the allowed decay modes. Shown are here the decay modes for third-generation leptoquarks, but the decay modes for first- and second generation leptoquarks are the same with the corresponding Standard Model particles of the first and second generation, respectively.

states are possible in theories with a larger particle content [12, 11]. In this thesis the scalar leptoquark state mentioned in the first row of table 2.3 is studied.

Many of the possible extensions of the Standard Model mentioned in section 2.2 predict leptoquarks. Since GUTs are described by gauge groups of the form SU(n) with $n > 3$, they predict the existence of $n^2 - 1$ gauge bosons. Among these are leptoquarks, which mediate transitions between quarks and leptons. These are possible since quarks and leptons appear in the same multiplets. Examples are GUTs based on SU(4) models [13], where the lepton number is interpreted as a fourth color so that leptons and fermions can be unified. Other examples are SU(5) models, which were first proposed by H.Georgi and S.L. Glashow [14]. While the minimal SU(5) is excluded due to the prediction of too rapid proton decay, extensions are not excluded [15]. In this model only the third generation of Standard Model fermions can violate lepton and baryon number conservation so that proton decay is not allowed. The corresponding gauge group is $\mathrm{SU}(5) \otimes \mathrm{SM}'$. The first and second-generation fermions are singlets under SU(5), while the third-generation is a singlet under the Standard Model gauge interactions.

Also compositeness and technicolor models are based on larger gauge groups than the Standard Model and leptoquarks are among the new resulting bosons. In compositeness models composite colored bosons are expected since quarks and therefore preons carry color. Furthermore, transitions between quarks and leptons are predicted due to the common set of preons they consist of [6, 11]. In technicolor models new particles which consist of a pair of a techniquark and a technilepton can be found under the new bosons. For these a decay into a Standard Model quark and a Standard Model lepton is predicted [7].

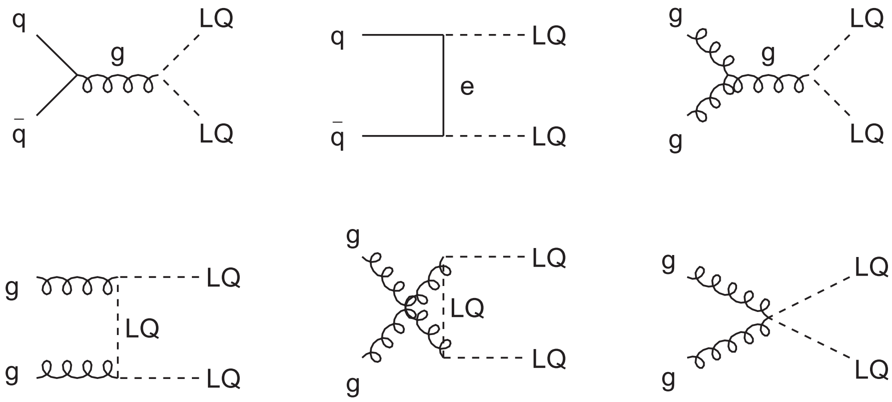


Figure 2.1: Feynman diagrams of leptoquark pair production (LO) through quark-antiquark annihilation and gluon-gluon-fusion [16].

In R-parity violating SUSY [2] a coupling like the one of leptoquarks to Standard Model particles may be realized. Thus bounds on leptoquark masses can also be interpreted as bounds on R-parity violating SUSY models [12].

2.3.1 Leptoquark production

At the LHC leptoquarks would be predominately produced in pairs through quark-antiquark annihilation and gluon-gluon fusion,

$$q+\overline{q} \rightarrow \mathrm{LQ}+\overline{\mathrm{LQ}}, \tag{2.5}$$

$$g+g \rightarrow \mathrm{LQ}+\overline{\mathrm{LQ}}. \tag{2.6}$$

The leading-order Feynman diagrams of these processes are shown in figure 2.1. In general single production of leptoquarks is also possible,

$$g+q \rightarrow \mathrm{LQ}+\ell. \tag{2.7}$$

The Feynman diagrams of this process are shown in figure 2.2. Single-production of leptoquarks is not studied in this thesis since its cross section is much smaller than the one for pair production.

In leading-order the cross sections of the parton processes for the pair production of scalar

[2] R-parity [5] is introduced in SUSY as a new multiplicative quantum number. It is defined as $R_P = (-1)^{3(B-L)+2S}$, where B is the baryon number, L the lepton number and S the spin of the considered particle. For Standard Model particles $R_P = 1$ and $R_P = -1$ for SUSY particles.

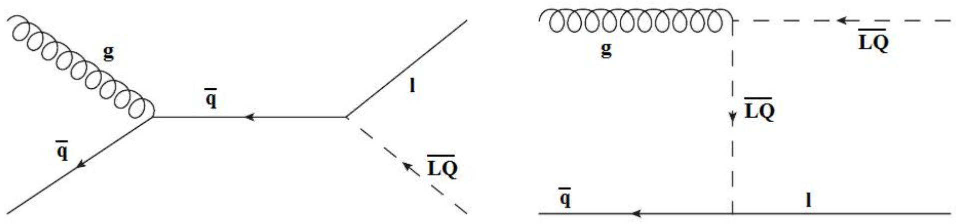


Figure 2.2: Feynman diagrams of single leptoquark production (LO) [11].

leptoquarks for quark-antiquark annihilation and gluon-gluon fusion are given by [16]

$$\widehat{\sigma}_{\mathrm{LO}}^{q\bar{q}} = \frac{2\alpha_s^2\pi}{27\widehat{s}}\beta^3, \tag{2.8}$$

$$\widehat{\sigma}_{\mathrm{LO}}^{gg} = \frac{\alpha_s^2\pi}{96\widehat{s}}[\beta(41-31\beta^2)+(18\beta^2-\beta^4-17)\log\frac{1+\beta}{1-\beta}], \tag{2.9}$$

where β is given by $\beta = \sqrt{1-4\frac{M_{\mathrm{LQ}}^2}{\widehat{s}}}$, $\widehat{s}$ is the energy of the parton process and λ is the Yukawa couplings of the leptoquark to the quarks and leptons.
It can be seen that the cross sections of leptoquark pair production depend only on the mass of the leptoquarks. In contrast, the single production cross section depends on the mass of the leptoquark and its Yukawa coupling. [12, 16]

2.3.2 Decay of leptoquarks

Leptoquarks can decay into a charged lepton and a quark, or into a neutrino and a quark. If leptoquarks decay in particles of different generations, they will induce flavor changing neutral currents (FCNC) like in special kaon decays (e.g. $K^+ \to \pi^+\nu\bar{\nu}$, $K_L \to e^+e^-$), certain pion decays (e.g. $\pi^0 \to \mu^\pm e^\mp$) or in muon decays (e.g. $\mu \to e\gamma$) [17, 18]. Since none of these processes have been observed, strong indirect limits on the existence of leptoquarks can be set. To avoid these strong bounds it is assumed that decays of leptoquarks into a single Standard Model generation are favored [19].

2.3.3 Limits on scalar leptoquark pair production

Searches for leptoquarks have been performed at many colliders e.g. LEP, HERA and Tevatron. So far no evidence for the existence of leptoquarks has been found and limits on their masses and couplings have been set. The best limits on scalar leptoquark pair production come from the CMS and the ATLAS experiments at the LHC [20]-[21].
Searches for first- and second-generation leptoquarks have been performed in the channel with two charged leptons and at least two jets, and in the channel with one charged lepton, at least two jets and missing transverse momentum. Here lepton stands for an electron or a muon. First-generation leptoquarks with masses up to 830 (640) GeV/c^2 [20] are excluded [3] by

[3]Throughout this thesis all exclusion limits are given at 95% confidence level (C.L.).

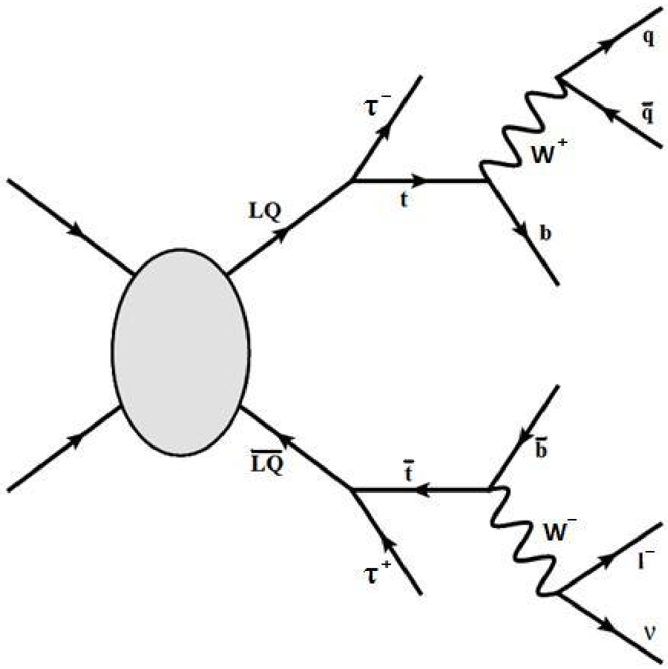


Figure 2.3: Sketch of leptoquark pair production at the LHC with the event signature $pp \rightarrow \mathrm{LQ}_3 + \overline{\mathrm{LQ}}_3 \rightarrow \mu + \tau_{\mathrm{had}} + \mathrm{X}$.

the CMS experiment for a branching ratio of 1 (0.5) for $\mathrm{LQ} \rightarrow \ell + \mathrm{q}$. ATLAS excludes first-generation leptoquarks with masses up to 660 (607) $\mathrm{GeV/c^2}$ [22] for the same branching ratios. 840 (650) $\mathrm{GeV/c^2}$ [20] and 594 (685) $\mathrm{GeV/c^2}$ [23] second-generation leptoquarks are excluded by CMS and ATLAS respectively for branching ratios of 0.5 (1.0). Searches for third-generation leptoquarks have been performed in the channel $\mathrm{LQ}_3 \rightarrow \mathrm{b} + \tau$ and $\mathrm{LQ}_3 \rightarrow \mathrm{b} + \nu_\tau$ by CMS. Third-generation leptoquarks in these channels are excluded up to 525 $\mathrm{GeV/c^2}$ [24] and 450 $\mathrm{GeV/c^2}$ [25] respectively. ATLAS has only performed a search in the $\tau + \mathrm{b}$ channel and excludes third-generation leptoquarks up to masses of 534 $\mathrm{GeV/c^2}$ [21]. In the channel $\mathrm{LQ}_3 \rightarrow \mathrm{t} + \tau$, which is studied in this thesis, no search has been performed so far.

2.4 Studied process and backgrounds

In this thesis the pair production of third-generation scalar leptoquarks decaying to a top quark and a tau lepton is studied. One possible Feynman diagram of this process is shown in figure 2.3.

A search for scalar leptoquarks is performed despite the fact that the cross section for the production of vector leptoquarks is higher than the one for scalar leptoquarks since the cross sections for scalar leptoquark production are calculated to next-to-leading order, while the cross sections for vector leptoquark production are only known to leading-order. Moreover, the interactions of vector leptoquarks with gluons include anomalous couplings, which make the search for vector leptoquarks more complicated than the search for scalar leptoquarks [26].

The signature studied consists of two top quarks and two tau leptons. It is assumed that both top quarks decay into a W-boson and a b quark, which happens, according to the CKM-matrix, in 99.9% of all decays. The W-bosons decay further leptonically in 32.4% and hadronically in approximately 67.6% of all decays [2]. If the W-boson decays leptonically a neutrino and thus missing transverse momentum is produced, otherwise jets are produced. Moreover, additional jets are expected in the event from initial and final state radiation.

The tau lepton decays hadronically in approximately 64.8% of all cases. In these decays also a tau neutrino is produced. In about 17.4% of the decays, the tau lepton decays into a muon, a muon neutrino and a tau neutrino. The probability of a decay into an electron, an electron neutrino and a tau neutrino is 17.8% [2]. Therefore in all tau lepton decays missing transverse momentum is produced.

In the performed search at least one muon and at least one hadronically decaying tau lepton is required. The muon can be produced if one of the W-bosons or one of the tau leptons decays leptonically. In addition to the case of one hadronically decaying tau lepton from the leptoquark decay, it can be produced from one of the W-decays. The resulting branching ratio for the process $\mathrm{LQ}_3 + \overline{\mathrm{LQ}}_3 \rightarrow \mu + \tau_{\mathrm{had}} + \mathrm{X}$ is 42.7%.

Since leptoquarks are expected to be quite heavy, the decay products of the leptoquarks have a high transverse momentum. Thus high values of H_T, which is the scalar sum of the transverse momentum of all jets and leptons in the event plus the missing transverse momentum, are expected.

In addition to the requirement of a large value of H_T in the event, a large amount of $\not{E}_T$ and many jets can be required in order to select the signal events. $\not{E}_T$ is produced in a leptoquark decay through the neutrinos from the tau lepton decays and from leptonically decaying W-bosons. Furthermore, there are many jets in the event, namely the two b-jets from the top quarks decays, jets from hadronically decaying W-bosons and additional jets produced in initial and final state radiation.

The requirements for large values of H_T, $\not{E}_T$ and many jets also helps to reduce the background. This can be explained with figure 2.4, where the cross section for different processes at the LHC is shown in dependence of the mass of the produced particles. It can be seen that the cross section for Standard Model processes like $t\bar{t}$-production, W-boson-production and Z-boson-production is much higher than the cross section for the production of scalar leptoquarks. But by requiring additional jets and high values of H_T and $\not{E}_T$ the cross sections for the Standard Model processes become much smaller, while the cross section for the leptoquark production stays approximately the same. Another feature of the signal events is, that there are many decay combinations in which two leptons with same-sign-charges occur, e.g. the lepton from the anti-top quark decay and the tau lepton from the leptoquark decay have both a negative charge. These same sign events occur in 68.9% of all leptoquark decays with at least one muon and at least one hadronically decaying tau lepton in the final state. In Standard Model processes these same-sign events are relatively rare so that a cut on the sign of the charges of the leptons in the event reduces the Standard Model background significantly.

The studied signature of one muon, one hadronically decaying tau lepton, additional jets and missing transverse momentum can be produced in Standard Model processes, too. These pro-

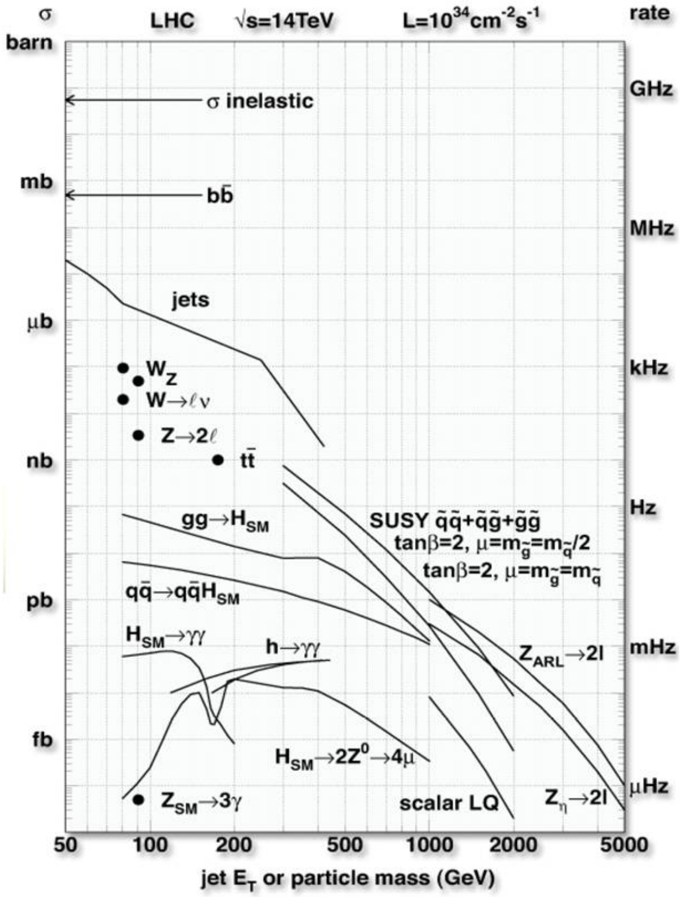


Figure 2.4: Production cross sections of different particles at the LHC at $\sqrt{s} = 14$ TeV in dependence of the mass of the particles [27].

cesses can be divided into irreducible background processes and reducible background processes for the search performed.

Irreducible backgrounds produce exactly the same final state as the studied process. The major irreducible background process in this search is $t\bar{t} \rightarrow WbWb$ with additional jets. If one of the W-bosons decays into a muon and the other one into a hadronically decaying tau lepton, the studied signature is produced. Similar to the leptoquark decay the missing transverse momentum comes from the neutrinos in the leptonically decaying W-bosons and hadronically decaying tau leptons. The same applies for the production of a $t\bar{t}$-pair and a Z-boson, which is an additional irreducible background. The cross section for this process is much lower than the one for $t\bar{t}$-production, but more combinations to produce the studied signature become possible because of decays of the Z-boson into muons or tau leptons.

An additional irreducible background is the production of Z-bosons with additional jets, if the Z-boson decays into a pair of tau leptons and one of them decays further into a muon and the other one hadronically. Missing transverse momentum is produced in the decays of the tau-leptons. In the same way ZZ-production is a irreducible background. Jets are produced in this process through initial and final state radiation and if one of the Z-bosons decays hadronically. Further diboson-production which produces the signature studied is WW-production. The hadronically decaying tau lepton and the muon are produced through the decays of the W-bosons, where

also missing transverse momentum is produced. Jets are again produced in initial or final state radiation. The same applies for WZ-production, which is also an irreducible background.
Reducible background processes are Standard Model processes, where no hadronically decaying tau lepton is produced, but where a jet is misidentified as a hadronically decaying tau lepton. In most of the cases missing transverse momentum is then produced through mismeasurements of the energies of the jets or through leptonically decaying W-bosons.
The most important reducible background process is the production of W-bosons and additional jets, if the W-boson decays into a muon. Similarly, $\mathrm{t\bar{t}}$-production plus additional jets is a reducible background, if one of the W-bosons decays into a muon and the other one hadronically. The same applies for the production of a $\mathrm{t\bar{t}}$-pair and a Z-boson, if the Z-boson decays into a pair of muons or into a pair of tau leptons, where both tau leptons decay further leptonically. Similarly, Z-boson-production with additional jets and ZZ-production are reducible backgrounds. WW-production is a reducible background, if one of the W-boson decays into a muon and a muon neutrino and the second one hadronically. Also WZ-production is a reducible background, if one of the bosons decays into muons and the other one hadronically.
In single top-production the studied signature can be produced if the W-boson decays leptonically into a muon, and one jet, which is the b-jet from the top quark decay or a jet from initial or final state radiation, is faked as a hadronically decaying tau lepton.
Light quark and gluon production is another reducible background. In the process the muon originates from the decay of a c- or b-hadron in the jets.
In conclusion, the search for third-generation leptoquark decaying to a top quark and a tau lepton has a very rich final state. However, many Standard Model processes can fake the typical event signature of $pp \to \mu + \tau_{\mathrm{had}} + \mathrm{jets} + \not{E}_{\mathrm{T}}$ through the misidentification of tau leptons, with an important irreducible background from $\mathrm{t\bar{t}} + \mathrm{jets}$ production. The cross sections for leptoquark pair production is quite small (e.g. 23.5 fb at 8 TeV for a leptoquark with a mass of 600 GeV/c^2 [16]), while the cross section for most of the mentioned background processes is much higher. Thus we expect only a few signal events and a lot of background events, which makes the search for leptoquarks very difficult.

Chapter 3

The LHC and the CMS experiment

This thesis is based on data which was recorded by the CMS experiment at the Large Hadron Collider (LHC) in the year 2012. In the following sections the LHC (section 3.1) and the CMS experiment (section 3.2) are introduced [28].

3.1 The Large Hadron Collider

The LHC is a circular collider located in the former LEP ("Large Electron Positron Collider") tunnel at CERN. It is a proton-proton collider with a circumference of 27 km. Collisions at a centre-of-mass energy of maximal 14 TeV and an instantaneous luminosity of maximal 10^{34} cm^{-2}s^{-1} can be studied.

In order to accelerate the protons to the desired centre-of-mass energy [1] the protons pass a few pre-accelerators before they are injected with an energy of 450 GeV into the LHC. Here the protons are accelerated with the help of superconducting cavities to their final energy. Superconducting dipoles are used to force the protons on a circular trajectory.

In the LHC the protons are grouped in bunches. The instantaneous luminosity can be calculated as

$$L = f \cdot \frac{n_b N^2}{4\pi\sigma_x\sigma_y} \cdot F. \tag{3.1}$$

Here f is the revolution frequency of one of the bunches, n_b stands for the number of bunches and N is the number of particles within one bunch. With the assumption that the distribution of the protons perpendicular to the beam axis can be approximately described by a Gaussian distribution, the factor $4\pi\sigma_x\sigma_y$ in the denominator, with the standard deviations σ_x and σ_y, describes the expansion of the bunches perpendicular to the beam. The variable F is a geometrical correction factor, which takes the inclination of the two beams into account.

At design parameters, there are $n_b = 2808$ bunches with approximately $N = 1.15 \cdot 10^{11}$ protons per bunch in the LHC. The time between two bunches is 25 ns and the revolution frequency of one bunch is $f = 11.25$ kHz. With these parameters the described high instantaneous luminosities can be reached. Focusing superconducting quadrupole magnets are used to keep the circumference of the bunches as small as possible.

[1] The centre-of-mass energy is calculated by $E_{\mathrm{CMS}} = \sqrt{s} = 2 \cdot E_{\mathrm{Proton}}$, where E_{proton}, stands for the energy of a proton.

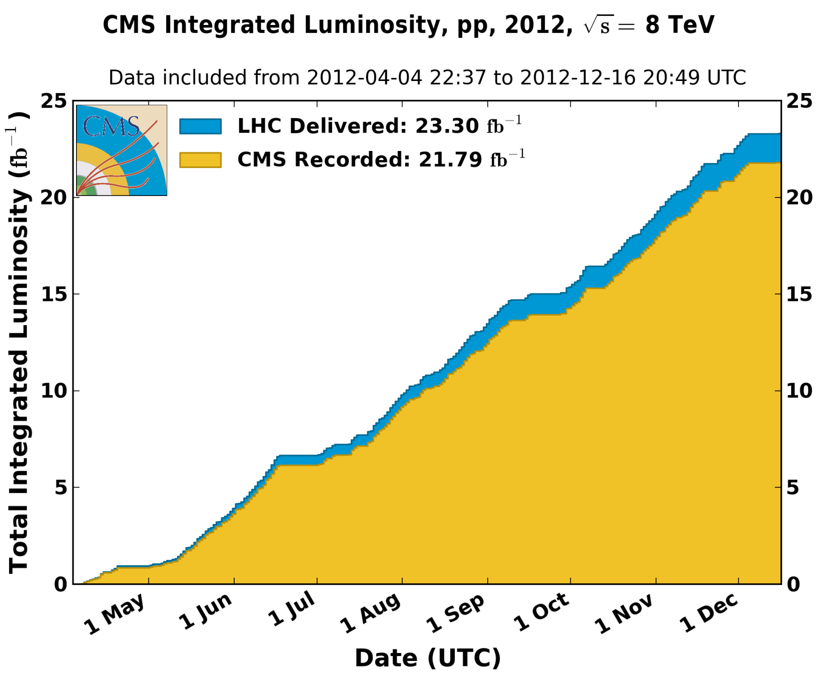


Figure 3.1: Total integrated luminosity delivered by the LHC and recorded by the CMS experiment in the year 2012 as a function of time [29].

The LHC has four main experiments which are CMS ("**C**ompact **M**uon **S**olenoid"), ATLAS ("**A** **T**oroidal **LHC** **A**pparatus"), ALICE ("**A** **L**arge **I**on **C**ollider **E**xperiment") and LHCb ("**L**arge **H**adron **C**ollider **b**eauty"). In their center the two in two different beam lines running counter-rotating beams are brought together, so that 40 million times per second collisions take place in the experiments.
CMS and ATLAS are so called multi-purpose detectors. Their goal is to exploit the full discovery potential of the LHC. In contrast to that, ALICE and LHCb have very special tasks. ALICE is specialized on studies of heavy ion collisions, which also can be performed at the LHC. The goal of LHCb is to study B hadrons and CP violation.
In the year 2012 the LHC ran with a centre-of-mass energy of 8 TeV. A maximum of 1380 bunches with a spacing of 50 ns were used. The highest instantaneous luminosity reached was $7.67 \cdot 10^{33}$ cm^{-2}s^{-1}. In figure 3.1 the total integrated luminosity delivered by the LHC in the year 2012 and the fraction of recorded data by the CMS experiment is shown as a function of time. In total the LHC delivered an integrated luminosity of 23.3 fb^{-1}, from which the CMS experiment recorded 21.79 fb^{-1} [29].

3.2 The CMS experiment

The CMS detector is a multi-purpose detector, which was built to fully exploit the LHC physics goals. It has a length of 21.6 m, a diameter of 14.6 m and weighs 12500 t. A sketch of the detector is shown in figure 3.2. The detector can be divided into a barrel region and two endcaps. It is forward-backward symmetric with respect to the collision point and constructed rotationally symmetrical around the beam axis.

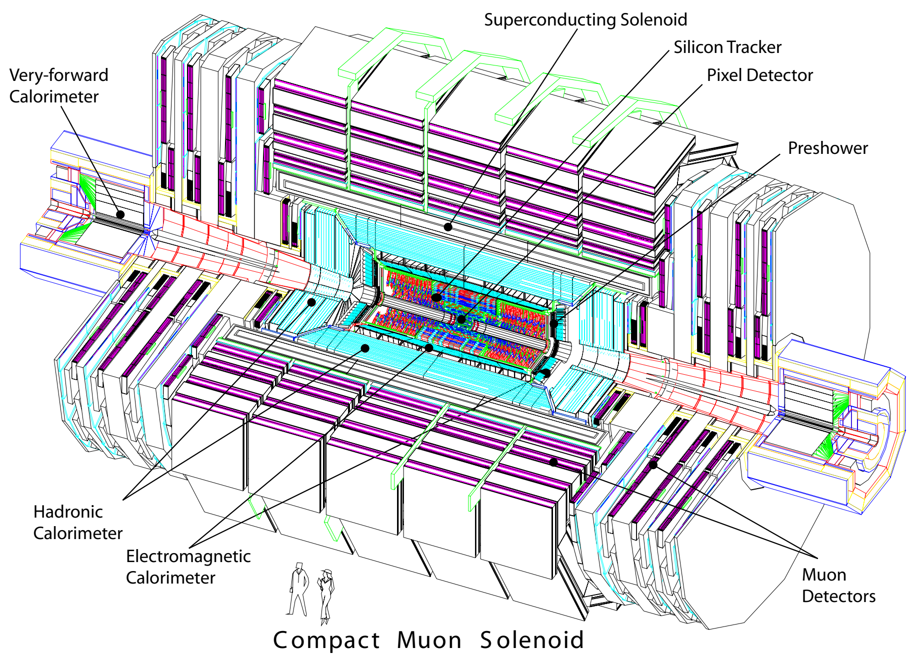


Figure 3.2: Sketch of the CMS detector [28].

The CMS detector consists of many subsystems, which are built around the beam-axis in an onion-like structure. In the center of the detector around the beam axis the pixel detector and the silicon track detector can be found. Together they build the inner track detector. Around the track detector the electromagnetic and hadronic calorimeters are located. A superconducting solenoid surrounds these two components. In the outermost part of the detector the muon system, in which the return yoke for the magnet is embedded, can be found.
The physic goals of the detector are

- the search for the Higgs boson,
- the search for SUSY,
- the search for other new physics (e.g. new massive vector bosons, extra dimensions),
- tests and precision studies of the Standard Model,
- the study of heavy-ion physics.

In order to achieve these goals the detector has to fulfill certain requirements. In the following, the design considerations of the CMS experiment are mentioned. The technical details of each subsystem are explained later.
The track detector has to be able to reconstruct all charged particles very efficiently with a good momentum resolution. Additionally, the track detector should be located close to the beam pipe to guarantee that tau leptons and b-jets are triggered and tagged efficiently to be able to reconstruct for example decays of the Higgs boson to a pair of b quarks or tau leptons.
The electromagnetic calorimeter should have a good energy resolution. Moreover, it should be able to measure the directions photons come from correctly. A good dielectron and diphoton

mass resolution and an efficient lepton and photon isolation are needed. In addition, the rejection of π^0's is important. This should all be possible over a wide geometric range ($|\eta| < 2.5$) and is important for example for the channel $\mathrm{H} \to \gamma\gamma$ and for SUSY searches involving photons. For the hadronic calorimeter a fine lateral segmentation is needed. Furthermore, the hadronic calorimeter should cover a wide geometric range ($|\eta| < 5$) and has to be very hermetic, so that a good dijet mass resolution and $\not{E}_T$ measurement is possible. This is important since many models beyond the Standard Model predict a large amount of $\not{E}_T$ in the event.
The muon system has to be able to identify muons and measure their momenta and their charge such that the dimuon mass spectrum can be reconstructed precisely. This has to be possible over a wide range of muon momenta and for $|\eta| < 2.5$. Muons play an important role in many searches, for example in $Z' \to \mu^+\mu^-$ searches.
The main challenges at the LHC arise from its high energy and luminosity. The high rate of interactions and the high number of particles in a bunch lead to a high flux of particles with large energies in the detectors. Thus all components of the CMS detector have to be radiation-hard.
In total approximately 10^9 interactions per second are expected at the design luminosity, but only 100 events per second can be stored. Thus an efficient readout and trigger system is needed so that the interesting events can be identified and stored.
Moreover, approximately 20 additional vertices per event are expected. To deal with this huge amount of pile-up a detector with high granularity and good time resolution is needed.
In the following sections the subsystems of the CMS detector, which fulfill all the mentioned criteria, are introduced and the coordinate system which is used to describe events in the CMS detector is presented.

3.2.1 Coordinate conventions

The right-handed coordinate system used at CMS to describe the detector and the interactions has its origin at the nominal interaction point inside the detector. The x-axis is directed radially towards the center of the LHC ring, the y-axis is showing upwards and the z-axis is pointing in the direction of the beam.
For the geometric description of the detector and the interactions the angles ϕ and θ are used. The angle ϕ is measured from the x-axis in the x-y-plane, perpendicular to the beam axis. The angle θ is defined with respect to the beam axis. With these parameters the pseudorapidity is defined as

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]. \tag{3.2}$$

Distances between two objects i and j are measured in the η-ϕ-plane as

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}. \tag{3.3}$$

The variables $\Delta\phi$ and $\Delta\eta$ are given by $\phi_i - \phi_j$ and $\eta_i - \eta_j$.
With these coordinate conventions the transverse momentum, which is the momentum of a

particle in the plane perpendicular to the beam axis, is defined by

$$p_T = \sqrt{p_x^2 + p_y^2}, \tag{3.4}$$

where p_x and p_y are the x- and y-components of the total momentum of the particle. This variable is particularly important since at a proton collider the total momentum of the partonic initial state is unknown, but the particles in the initial state have no transverse momentum.

3.2.2 The inner track detector

The track detector is the innermost sub-detector of the CMS experiment. It surrounds the beam pipe and lies inside the magnetic field. It has a length of approximately 540 cm and covers the region up to a radius of approximately 110 cm and $|\eta| < 2.5$. The bending of the trajectory of charged particles in the magnetic field due to the Lorentz force is used to determine their charge and momentum. The CMS track detector is fully based on silicon detector technology.

In the barrel region, where the highest particle flux can be found, hybrid pixel detectors are arranged in three layers. They can be found at radii of 4.4 cm, 7.3 cm and 10.2 cm. The endcaps of the inner track detector consists of two disks of pixel detectors. In the z-direction their positions are at 34.5 cm and 46.5 cm. In total the inner track detector consists of 66 million pixels. The pixels in both the endcaps and the barrel have a size of 100×150 μm^2 in $(r, \phi) \times z$, which leads to a spatial resolution of approximately 10 μm in the r-ϕ-direction and 20 μm in the z-direction.

At bigger radii, where the particle flux is lower, silicon microstrip detectors are placed. In the barrel the microstrip detectors can be divided into two regions: Tracker Inner Barrel (TIB), which lies at a radial distance between 20 cm and 55 cm from the beam pipe, and the Tracker Outer Barrel (TOB), which surrounds the TIB. The TIB consists of four layers of silicon microstrip detectors, the TOB of 6 layers. In the TIB the cells have a minimum size of 10 cm $\times$ 80μm in $(r, \phi) \times z$. In the TOB the cell size is larger, the maximum cell size is 25 cm $\times$ 180μm in $(r, \phi) \times z$. In the TIB the strip pitch (80 - 120 μm) is smaller than in the TOB (120 - 180 μm). Thus in the TOB a resolution of 23-34 μm in the r-ϕ-direction and 23 μm in the z-coordinate is reached, whereas in the TOB the resolution lies between 35-52 μm in r-ϕ and 52 μm in the z-direction.

The endcap silicon microstrip detector can also be divided into two parts: the Tracker End Cap (TEC) and the Tracker Inner Disks (TID). Both consist of rings which are built around the beam pipe perpendicular to it and have radially arranged strips. The TECs have nine disks on each side of the detector, which have z-coordinates between 120 cm and 280 cm. The TID consists of three disks, which are located between the TIB and the TEC, so that no gap exists between the different track detector parts. In total the inner track detector consists of 9.6 million silicon strips.

In general the track detector has a very high granularity and shows an excellent performance [30].

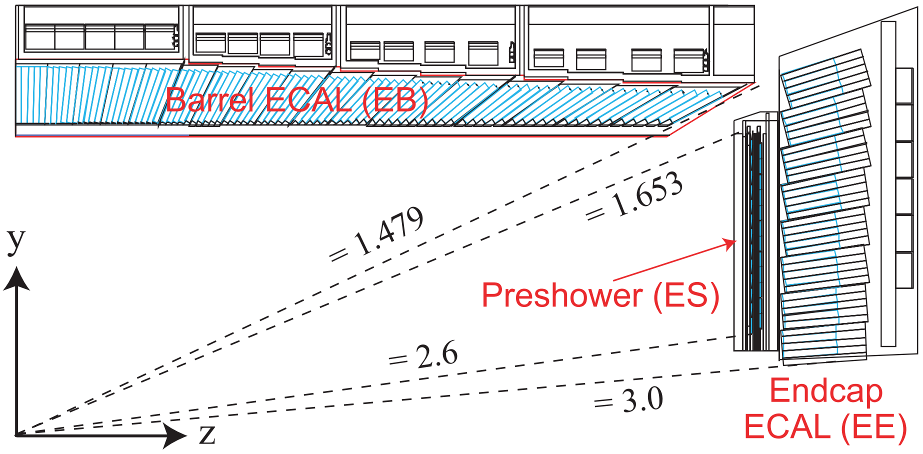


Figure 3.3: Sketch of the electromagnetic calorimeter [28].

3.2.3 The calorimeter

Like the inner track detector the calorimeter is located inside the magnetic field. It consists of an electromagnetic calorimeter, whose task is to measure the energy of photons and electrons, and of a hadronic calorimeter, which measures the energy of hadronically interacting particles.

The electromagnetic calorimeter

The electromagnetic calorimeter (ECAL), which is shown in figure 3.3, covers the region up to $|\eta| < 3.0$. It can also be divided into a barrel part, which goes up to $|\eta| < 1.479$, and 2 endcaps, which cover the region $1.479 < |\eta| < 3.0$. It is a homogeneous, hermetic calorimeter, which uses lead tungstate crystals ($PbWO_4$) as scintillator material. 61200 crystals can be found in the barrel, 7324 in each of the endcaps. These crystals are radiation hard and have a fast response time. Moreover, with $X_0 = 0.89$ cm their radiation length is quite short. Photodiodes which can be used in a magnetic field and which can handle a low light yield are used to detect the signals from the crystals.
The crystals in the barrel are tilted in order to avoid gaps between them. One crystal covers an area of 0.0174×0.0174 in $\Delta\eta \times \Delta\phi$, with a length of 230 mm.
The endcaps are located 314 cm away from the nominal interaction point. The crystals here are arranged in 5×5 clusters. These clusters are also tilted with respect to each other in order to avoid gaps in the acceptance. One crystal has a size of 28.6×28.6 mm^2 in the x-y-plane and a length of 220 mm.
In front of the endcaps additional sampling calorimeters, which serve as preshower devices, can be found. The active material consists of two planes of silicon strip detectors. The pitch size is

1.9 mm. In-between lead absorbers can be found. The task of this sampling calorimeter is the identification of π^0's.
In general the CMS detector has a very granular, fast and hermetic electromagnetic calorimeter which shows an excellent performance [31].

The hadronic calorimeter

The hadronic calorimeter (HCAL) consists of four parts: the Hadron Barrel (HB), the Hadron Outer (HO), the Hadron Endcap (HE) and the Hadron Forward (HF). All of them are sampling calorimeters and, except for the HF, use brass as absorber and plastic scintillator tiles as active material. Embedded in the scintillators tiles are wavelength-shifting fibers. With these fibers the emitted light from the scintillators is carried to the readout system, which is made of multi-channel hybrid photodiodes. The advantage of brass as absorber material is that it has a short interaction length and is non-magnetic. The chosen design of the HCAL results in a hermetic calorimeter, whose only uninstrumented area can be found at 53°. Here a radially running crack can be found through which the cables of the inner track detector and the ECAL run.
The HB covers the range $|\eta| < 1.4$. It is divided into 2304 towers which are segmented in η and ϕ ($\Delta\eta \times \Delta\phi = 0.087 \times 0.087$). The HE consists of the same amount of towers and covers the region $1.3 < |\eta| < 3.0$. The towers at smaller η have segmentations in ϕ of 5° and of 0.087 in η. The inner towers have a segmentation of 10° in ϕ. Here the segmentation in η increases with increasing η from 0.09 to 0.35.
The HO lies inside the barrel part of the muon system, outside the magnet. It covers the pseudorapidity range $-1.26 < \eta < 1.26$. Its task is to measure the energy of showers which escape from the other calorimeters. So the total thickness of the hadronic calorimeters is increased to over 10 interaction lengths. The scintillator tiles in the HO are arranged in sections of 30° in ϕ and in five sections in η, matching the geometry of the muon system, in which it is embedded. The segmentation in η and ϕ is like the one in the HB.
The Hadron Forward calorimeter is a sampling calorimeter consisting of steel and quartz fibers. The calorimeter covers the range $3.0 < |\eta| < 5.0$. The signal is Cerenkov light emitted from the quartz fibers, which is then transferred to photomultipliers. Through this choice the showers are narrower and shorter than in the rest of the calorimeter. This is important for the forward region where a lot of showers within small distances occur. The Hadron Forward calorimeters are located in a distance of 11.2 m in the z-direction measured from the interaction point. The quartz fibers run parallel to the beam pipe and are arranged in a square grid with a distance of 5 mm between two fibers. The HF consists of a total of 900 towers, which results in a segmentation of 10° in ϕ and ≈ 0.175 in η.

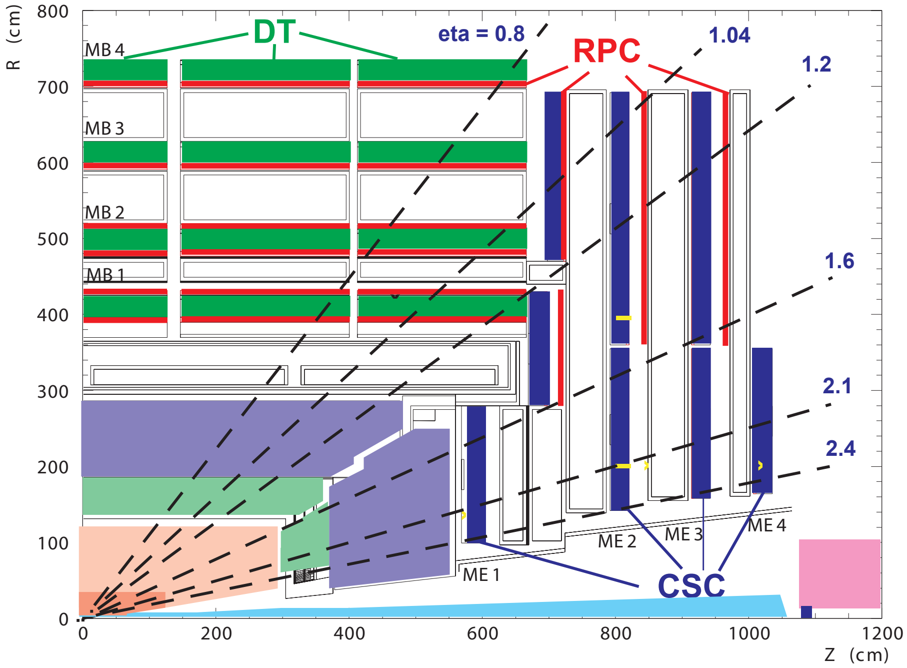


Figure 3.4: Sketch of one quarter of the muon system [28].

3.2.4 The magnet

One of the main features of the CMS experiment is a superconducting solenoid, which generates a magnetic field of 3.8 T. It surrounds the track detector and the calorimeters and has a length of 12.9 m. The inner bore has a diameter of 5.9 m. Due to the resulting magnetic field parallel to the beam charged particles in the track detector and muons in the muon system are bend in the transverse plane so that their transverse momentum can be determined. The strong magnetic field is needed in order to reach the desired performance of the measurements in the muon system and the required resolution of the momentum measurement of charged particles in the track detector.

3.2.5 The muon system

The muon system surrounds the magnet and forms the outermost part of the CMS detector. Three different types of detectors, which all use gas as the active material, are used: aluminum drift tube (DT) chambers, cathode strip chambers (CSCs) and resistive plate chambers (RPCs). In total the active detector material has an area of 25000 m^2.

The layout of one quarter of the CMS muon system is shown in figure 3.4. The Muon Barrel (MB) region consists of four cylindrically arranged layers of detectors at radii between 4 m and 7 m. The layers are segmented in five sections along the beam axis. In between them the return yoke for the magnet is embedded.

In each Muon Endcap (ME) four disks, which are orientated perpendicular to the beam axis, can be found.

The barrel consists mainly out of DT chambers. In total there are 250 DT chambers, which cover the region up to $|\eta| < 1.2$. DT chambers can be used in this region because a small muon

rate occurs, the neutron induced background is relatively low and there is only a small residual magnetic field in the chambers. With the DT chambers the positions of the muons can be measured with a resolution of 100μm in ϕ, their direction can be determined with a precision of 1 mrad. The DT chambers are also used for triggering.
In the first two layers of the barrel each DT chamber is surrounded by two RPCs. In the other two layers they are arranged alternating. Thus a muon with a high transverse momentum passing the barrel is measured at 44 points by traversing four DT chambers and six RPCs.
All used RPCs have double gaps. The operation is done in avalanche mode. The RPCs are very fast, give a good time measurement, but a worse position measurement than the other used muon detectors. The RPCs are used mainly to identify the bunch crossing and for triggering.
In the Muon Endcaps 468 CSCs are used. These cover the pseudorapidity area up to $|\eta| < 2.4$. CSCs were chosen because they are able to operate in a region with a high neutron and muon rate and a high magnetic field. The disks of the endcaps consist of 270 chambers. Each chamber has six gas gaps. In each gap a plane of radial cathode strips can be found. Almost perpendicular to this plane a plane of anode wires is placed. A coarse position measurement is possible with the help of the image charge which is produced at the cathodes when the gas is ionized through a charged particle traversing the chamber. The advantage of the CSCs is that they are very fast so that they can be used for triggering. A more precise measurement is obtained by studying the charge centre-of-gravity of the distribution found on the cathode strips. In each of the CSCs up to six measurements of the space coordinates can be made. This gives a resolution of approximately 200 μm in the r- and z-coordinates. In ϕ the resolution is approximately 10 mrad. In order to avoid gaps in the muon system the chambers overlap in ϕ.
In each of the disks of the endcaps two rings of additional RPCs can be found, with a total of 36 chambers. They cover the pseudorapidity region up to $|\eta| < 1.6$.

3.2.6 The trigger system

At design luminosity an event rate of approximately 10^9 Hz is expected in the CMS detector. Due to limited disk space and computing power this rate has to be reduced to a rate of approximately 100 Hz by the trigger system.
The trigger system of CMS consists of two different parts: the Level-1 trigger and the High-Level trigger (HLT).
The Level-1 trigger is based on custom hardware processors. These use information which can be accessed fast from the calorimeters and the muon system and some combined information from these two sub-systems. However, it is not possible to use the full granularity and the full resolution of the CMS detector, because of timing constraints. With this reduced information so called trigger objects are formed, for example candidates for muons, electrons, photons and jets. The Level-1 trigger searches for such objects above certain p_T or E_T thresholds. Also global sums of $\not{E}_T$ or E_T can be criteria to decide whether an event passes the Level-1 trigger or not. In this way the total event rate is reduced to a rate of 100 kHz. The Level-1 decision is

made in 3.2 μs. In this time the signals are passed from the front-end electronics to the service cavern, where the trigger logic is placed. There the trigger decision is made in less than 1 μs. Afterwards the result is returned to the front-end electronics. The complete data of the studied event is stored in buffers until the final decision of the Level-1 trigger is made.

If the decision is positive the event is passed to the HLT. In contrast to the Level-1 trigger the HLT is software-based. Moreover, the full available granularity and resolution is used in the decision process, but only interesting regions and not the full event are reconstructed.

To rebuild an event the data is transferred to a processor, where the HLT software processes the event. Again it is tested if the event fulfills certain criteria, which can be different than the Level-1 criteria. Each tested set of criteria builds a so called "trigger chain". If the event passes one of these trigger chains it is accepted by the trigger and stored in the corresponding dataset. In the end different datasets are obtained from the available trigger chains and stored for analyzing. In this way the event rate is further reduced to the desired rate of 100 Hz. The decision in the HLT is made in 50 ms.

Chapter 4

Reconstruction and identification of particles candidates and jets

From the signals measured by the different detector components particle candidates are reconstructed and identified. In the following sections the reconstruction and identification algorithms are explained for particle candidates, jets and other observables used in this analysis.
The chapter begins with a short description of the particle-flow event reconstruction algorithm (4.1). Afterwards an introduction of the reconstruction and identification of muon candidates (4.2), jet candidates (4.3) and tau-lepton candidates (4.4) is given. In section 4.5 the measurement of $\not{E}_T$ and H_T is described.

4.1 The particle-flow event reconstruction algorithm

The aim of the particle-flow event reconstruction algorithm [32] is to reconstruct and identify all stable particles produced in the collisions in the CMS detector. To achieve this goal, all available information from all detector components is used. The result of the algorithm is a list of identified candidates for muons, electrons, photons and neutral and charged hadrons. The particle-flow approach is possible with the CMS detector due to the excellent performance of its track detector and its hermetic ECAL with high granularity.
In the first step of the algorithm tracks of charged particles and calorimeter clusters are reconstructed from the measured signals in the detector.
Since the track detector is able to measure the momentum of charged particles and their initial direction very precisely up to high momenta, a reconstruction of the tracks with a high efficiency and a low fake rate is very important for an efficient event reconstruction. To achieve this, an iterative tracking algorithm has been developed. In the first step of this algorithm the reconstruction criteria the tracks have to fulfill are very tight. In this way a very small fake rate with a moderate efficiency is achieved. The hits allocated to the reconstructed tracks are removed and a next reconstruction step with looser criteria is performed. This procedure is repeated several times. In order to reconstruct also charged particles produced in the decays of long living particles, photon conversions and interactions in the track detector material the criteria on the vertex are loosened in the later iteration steps.

In general the reconstruction efficiency gets higher the looser the reconstruction criteria become and by removing the hits assigned to already reconstructed tracks the fake rate stays small because the combinatorics are reduced. In the end a fake rate of a few per cent is achieved. For muons in the track detector acceptance the reconstruction efficiency is 99.5%, charged hadrons in the jets can be reconstructed with an efficiency higher than 90%.
For calorimeter entries a special clustering algorithm has been developed, which is able to separate close energy deposits and to detect even low-energy particles with a high efficiency. In general the algorithm is able to measure the energy and direction of all stable neutral particles, to separate the energy depositions of these neutral particles from those of the charged particles and to reconstruct and identify electrons and Bremsstrahlung photons. In addition, the energy of charged hadrons with low-quality tracks or with high transverse momenta are determined from calorimetrie measurements to improve the resolution.
In each sub-component of the calorimeter, except the HF, the algorithm is applied separately. In its first step a calorimeter cell with an energy deposition which corresponds to a local energy maximum above a certain energy threshold is searched. This cell is the seed for the cluster. Neighboring cells which have an energy two standard deviations higher than the average noise in the calorimeter are added to the cluster. The so produced cluster is called a "topological cluster". This is repeated until no further neighboring cells fulfill the mentioned criteria. Overlapping clusters share the energy of a cell in dependence of the distance of that cell to the mean energy-weighted position of the cluster, where the cluster energies and positions are determined iteratively.
In general, more than one reconstructed particle-flow element corresponds to one particle-flow object. For example an electron produces hits in the track detector and deposits energy in several clusters in the electromagnetic calorimeter due to Bremsstrahlung photons. Therefore an algorithm is needed, which tries to link all particle-flow elements to reconstruct the particles. This algorithm tries to link all particle-flow elements to each other. A track is combined with a calorimeter cluster by extrapolating the track from the last hit to the different parts of the calorimeter. The distance to which the track is extrapolated is determined by the typical longitudinal shower profile of an electron or by one interaction length, which is the typical length of a hadronic shower. If the so extrapolated track ends in a calorimeter cluster the two objects are considered to belong to one particle. Moreover, the algorithm tries to link two clusters in different calorimeters by extrapolating the position in the more granular calorimeter to the less granular calorimeter.
In order to find photons from Bremsstrahlung a tangent to each track is build at each intersection of a track and a track detector layer. If this tangent meets a cluster in the ECAL, the cluster is considered to be a Bremsstrahlung photon.
In the last step of the linking algorithm a χ^2-fit between the tracks in the inner track detector and matching tracks in the muon system is performed.
The combined elements from the linking algorithm build one block. These blocks are then used to identify the different particles. If the identification was successful, the corresponding elements are removed from the block.

In the first step of the identification process it is checked if there is a link between an inner track and a track in the muon system. If this is the case, the momentum determined in the inner track detector measurement and the combined momentum of both systems are compared and if they agree within three standard deviations, the particle is called a particle-flow muon.
In the next step electrons are identified by checking whether the tracks in the blocks are short and if Bremsstrahlung was assigned to the track. Moreover, it is checked if different tracking and calorimeter variables fulfill certain criteria. If this is successful, a particle-flow electron is found.
In the next step tracks are rejected where the expected relative energy resolution of the calorimeter is better than the uncertainty on the transverse momentum measured by the track detector.
Afterwards the momenta of the tracks linked to ECAL or HCAL clusters are compared with the energy deposited in the corresponding clusters. If the track momentum is more than three standard deviations larger than the energy measured in the cluster, a search for fake tracks and for muons with looser criteria is performed. After some additional quality cuts on the tracks all remaining tracks are considered to be particle-flow charged hadrons, whose momentum and energy is determined with the track measurements. If the measurements in the track detector and in the calorimeter are compatible, a fit is done using both measurements and the obtained combined value is taken. After this step the momenta of the linked tracks are again compared to the calorimeter cluster. If the energy in the clusters is higher than the energy of the linked tracks and if this excess is larger than the energy resolution, particle-flow photons and particle-flow neutral hadrons are created from theses excesses. The calorimeter clusters which were not linked to any track become particle-flow photons and particle-flow hadrons.
The resulting list of particles is then used to further reconstruct the event, for example as an input for jet algorithms, for reconstruction of tau leptons, for b-tagging and for calculating the missing transverse momentum.

4.2 Muons

Muons are the first particles reconstructed in the linking step of the particle-flow algorithm described in 4.1. Two different reconstruction approaches are possible: the "global muon reconstruction" and the "tracker muon reconstruction" [33].
The **global muon reconstruction** starts with reconstructed tracks in the muon system and checks whether a matching track in the inner track detector can be found. If this is the case, a fit is done to combine the hits of both tracks. For muons with a transverse momentum higher than 200 GeV/c the resolution of the transverse momentum can be improved significantly with this combination, compared to the resolution achieved with a measurement from the inner track detector alone due to the much longer lever arm.
The **tracker muon reconstruction** starts with tracks measured in the track detector which fulfill $p_T > 0.5$ GeV/c and p $>$ 2.5 GeV/c and extrapolates them to the muon system. In the extrapolation the uncertainty due to multiple scattering and the expected energy loss on the particle's

way from the track detector to the muon system are taken into account. If the extrapolated track can be assigned to a muon segment, the particle is a muon candidate. Especially muons with low momenta ($p \lesssim 5$ GeV/c) are reconstructed with this algorithm, since only one hit in the muon system is needed for the reconstruction.
Most muons can be reconstructed with one of the described algorithms and often they are reconstructed by both. If both algorithms fail, particles with a reconstructed track in the muon system can also be considered as muon candidates.
After the reconstruction, identification criteria with different working points are applied [34]. The tighter the criteria the lower gets the fake rate, but at the same time the efficiency is reduced. In this thesis the loose and the tight working points are used.
The **loose identification criteria** are:

- The muon candidate was reconstructed by the particle-flow algorithm.
- The muon candidate was reconstructed by the global muon reconstruction or by the tracker muon reconstruction.

The additional **tight identification criteria** are:

- The muon candidate was reconstructed by the global muon reconstruction.
- For the fit of the track in the global muon reconstruction the following criteria have to be fulfilled:
  - $\chi^2/n < 10$, where n is the number of degrees of freedoms in the fit.
  - At least one muon chamber hit is included in the fit.
- The track of the muon candidate could be matched to signals in at least two muon stations.
- The muon candidate produced at least one pixel hit.
- There were measurements of the muon candidate in more than five layers of the inner track detector.
- For the distances of the reconstructed tracks with respect to the primary vertex the requirements are:
  - The transverse impact parameter of the trajectory of the muon candidate is smaller than 2 mm.
  - The longitudinal distance to the central track of the muon candidate is smaller than 5 mm.

After the reconstruction of the muon candidates the linking part of the particle-flow algorithm is continued. Afterwards the **isolation of the muon candidate** can be calculated with the reconstructed particle-flow objects. The isolation is defined as the sum of the transverse energy of all particle-flow charged hadrons which were assigned to the primary vertex, all particle-flow neutral hadrons and all particle-flow photons in a cone with radius $\Delta R = 0.4$ around the muon

candidate, divided by the transverse momentum of the muon candidate. This value has to be smaller than 0.12 to fulfill the tight identification criteria. The loose working point does not contain any isolation requirement.
The **momentum of the muon candidates** is in general determined by a fit of the assigned tracks. The determined transverse momentum of the global fit is chosen if it fulfills certain criteria. These criteria are that the fit of the tracks in the track detector and the global fit yield transverse momenta above 200 GeV/c and that the result of the global fit for q/p, where q is the charge of the particle and p the total momentum, is in agreement with the fit in the track detector within two standard deviations.

4.3 Jets

In CMS different types of jets can be used. They differ by the jet algorithms used and the input collections to the jet finders. In this thesis jets based on particle-flow objects are used. The algorithm which is used is the anti-k_t algorithm with a distance parameter of $\Delta R = 0.5$.
The anti-k_t-algorithm [35] is a sequential recombination algorithm. It is collinear and infrared safe, which means that a collinear splitting of a particle or an emission of a particle with low energy does not change the outcome of the algorithm.
The algorithm uses for each pair of particles i and j two different distance measurements,

$$d_{\mathrm{ij}} = min(p_{\mathrm{Ti}}^{-1}, p_{\mathrm{Tj}}^{-1}) \frac{\Delta R_{\mathrm{ij}}^2}{R^2}, \tag{4.1}$$

$$d_{\mathrm{iB}} = p_{\mathrm{Ti}}^{-1}. \tag{4.2}$$

In these equations ΔR_{ij}^2 is given by $\Delta R_{\mathrm{ij}}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ and y_i is the rapidity defined as $y = \frac{1}{2} \ln \frac{E+p_z}{E-p_z}$ of particle i. The parameter R is the chosen cone size of the jet, where in this analysis R is set to 0.5.
The algorithm first calculates the two distances d_{ij} and d_{iB} for each particle pair i and j in the event. From all these values the minimum is chosen. If it corresponds to d_{ij}, the four momenta of the particle i and j are combined to a new particle and the algorithm starts from the beginning. If the minimum is a value from d_{iB}, the particle i is called a jet and it is removed from the list of particles. All distances are calculated again and the procedure starts from the beginning. The algorithm stops when there is no particle left. In this way one gets a list of jets.
In order to suppress pile-up, charged particle-flow hadrons which were not assigned to the primary vertex are subtracted from the reconstructed jets. To suppress also the pile-up from neutral hadrons the average energy of these in a jet is subtracted from the jet.
On the jets obtained, **jet energy corrections** have to be applied because of the non-linear response of the calorimeter to hadronic showers and further instrumental effects. In CMS the jet energy corrections are applied in a staged approach. The corrections are scale factors for the four momenta of jets, which depend on various different jet quantities. The applied corrections are:

- **Offset (L1):** The L1 corrections aim to subtract the average energy from pile-up and from electronic noise from a jet's measured energy.
- **Relative Jet Corrections (L2):** The L2 corrections remove the dependence of the measured jet energy on the pseudorapidity by correcting the energy of all jets relative to jets in the center of the detector ($|\eta| < 1.3$).
- **Absolute Jet Correction (L3):** The L3 corrections are determined from simulated events and they aim to remove the p_T-dependence of the jet response. After the application of the L3 correction, the jet's energy is consistent with the energy of the simulated jet on particle level.
- **Residual corrections:** These corrections are applied to correct for small residual differences in data and MC.

The L1, L2 and L3 corrections are applied to data and MC, the residual corrections are only applied to data. A detailed description of the jet energy corrections can be found elsewhere [36]. On the corrected jets identification criteria are applied. In this analysis jets have to fulfill **the loose identification criteria for particle-flow jets**, which are:

- The neutral hadron fraction of the jet has to be smaller than 99%.
- The neutral electromagnetic fraction of the jet must be smaller than 99%.
- The jet has to have more than one constituent.
- The charged hadron fraction of the jet has to be higher than 0% if $|\eta| < 2.4$, where η is the pseudorapidity of the jet.
- The charged multiplicity of the jet must be larger than zero if $|\eta| < 2.4$
- The charged electromagnetic fraction of the jet has to be smaller than 99% if $|\eta| < 2.4$.

B-jets

Jets originating from b quarks can be distinguished from other jets by requiring a b-tag.
B-tagging algorithms [37] are based on the large lifetime of b-hadrons (≈ 1.5 ps), which are produced in the hadronization process of a b quark. After the resulting long flight distance, which is visible in the detector, the b-hadrons decay and a secondary vertex is produced whose tracks are not compatible with the primary vertex. Another important property of a b quark is its high mass compared to other quarks, which result in high masses of b-hadrons. Moreover, in a b-jet more charged particles are produced in the final state than in other jets and the largest fraction of the b-jet's energy is carried by the b-hadron. These properties of b-jets are used in b-tagging algorithms to find a single variable which can discriminate b-jets from other quark- and gluon-jets.
The **Combined Secondary Vertex (CSV) algorithm** is one of the b-tagging algorithms used in CMS. The algorithm starts with the reconstruction of secondary vertices inside jets. This is

done by using a vertex finder algorithm, which looks at all tracks in the jets, finds tracks which do not belong to the primary vertex and reconstructs new vertices with these outliers. If the so reconstructed vertices fulfill certain quality criteria, the jet is a b-jet candidate.
In the next step these candidates are divided into jets where at least one secondary vertex could be reconstructed, jets where no vertex could be reconstructed but displaced tracks were found such that a pseudo vertex can be build, and jets which could not be assigned to one of the other two categories. Depending on the category different variables are used, which are later combined into one discriminating variable.
A variable for the suppression of charm quark jets is used for all categories except the last one. For the calculation of this variable the tracks are sorted by their impact parameter significance and the invariant mass of the tracks is calculated. If the mass gets higher than the mass of the charm quark, the impact parameter significance of the last added track is taken as an input variable.
Additional variables used for jets of the first category are

- the invariant mass of all charged particles assigned to the vertex and their multiplicity,
- the quotient of the distance between the primary and secondary vertex in the transverse plane and the error of this distance
- and the energy fraction of the charged particles assigned to the secondary vertex, compared to all charged particles in the jet and their pseudorapidity compared to the direction of the jet.

The set of variables for the second category of jets is similar, with the exception that the variable which contains the distance between the two vertices is not considered. For the third category none of these extra variables is used.
All these variables are then combined by a likelihood ratio technique to find one discriminating variable for all b-jet candidates. This variable tends to be one for b-jets and zero for other jets. The value which the variable has to have so that the jet is considered to be a b-jet defines the working point.
In this thesis the loose and the tight working points are used, which means that the discriminating variable has to be higher than 0.244 repectively 0.898.

4.4 Tau leptons

In this thesis hadronically decaying tau leptons play an important role. Tau leptons decay hadronically in 64.8% of all decays. One distinguishes between so called "one-prong decays" and "three-prong decays", where one-prong decays contain one charged hadron in the final state and three-prong decays have three charged hadrons in the final state. The most probable one-prong decays are: $\tau^- \to h^-\nu_\tau$, $\tau^- \to h^-\pi^0\nu_\tau$ and $\tau^- \to h^-\pi^0\pi^0\nu_\tau$. The most important three-prong decays are $\tau^- \to h^-h^+h^-\nu_\tau$ and $\tau^- \to h^-h^+h^-\pi^0\nu_\tau$ [2], where h stands for a hadron. Typically these are charged kaons or pions.

In all these decays there are at least one charged hadron, at least on neutral pion and one neutrino, which, except for the neutrino, can be reconstructed with the particle-flow algorithm and are then clustered into a jet. The tau lepton reconstruction algorithms used in CMS start with these particle-flow jets and check whether the particles in the jet can be assigned to a hadronically decaying tau lepton. The reconstruction of the neutral pions is thereby particularly important. In most of the cases neutral pions decay into two photons, which decay very often further into a pair of electrons while traversing the detector material. This leads to a broadening of the calorimeter depositions due to the magnetic field which bends the tracks of the electrons and positrons.

To take these broadenings into account the algorithm used in this analysis, the **Hadron Plus Strips algorithm (HPS)** [38, 39], reconstructs electromagnetic particle-flow objects in strips in order to find neutral pion candidates. All used strips are growing in ϕ but are narrow in η. The first strip is centered around the electromagnetic particle-flow candidate with the highest energy and it is searched for other electromagnetic particles within the strip. If the search is successful, the highest of the found electromagnetic particles is combined with the first one and the center of the strip is rearranged around the constructed new particle. This procedure is repeated until no further electromagnetic particles are found within the strip. The algorithm searches then for other electromagnetic particles which have not been assigned to a strip yet and repeats the whole process. The constructed strips which fulfill $p_T^{strip} > 1$ GeV/c are combined with the charged particle-flow hadrons in order to reconstruct the different decay modes of the tau leptons. All mentioned decay modes except $\tau^- \to h^- \pi^0 \pi^0 \nu_\tau$ and $\tau^- \to h^- h^+ h^- \pi^0 \nu_\tau$ are considered. Since most of the tau leptons decay through intermediate resonances certain mass constraints can be applied during the combination and energy thresholds have to be fulfilled to increase the background rejection. The reconstructed decay modes have to fulfill a narrowness criterion, which means that all decay products have to lie in a cone of radius $\Delta R = 2.8\ \mathrm{GeV}/(\mathrm{c} \cdot \mathrm{p_T})$, where p_T is the transverse momentum of the reconstructed tau lepton candidate. Additionally, the reconstructed tau lepton candidate has to lie within a distance of $\Delta R = 0.1$ around the axis of the corresponding particle-flow jet.

Afterwards additional identification criteria are applied. For this different working points are available [40]. In this analysis tau lepton candidates who fulfill the so called "decay mode finding" criterion and medium tau lepton candidates are used.

The criteria **"decay mode finding"** means that the only requirement on a tau lepton candidate is that it was reconstructed by the HPS algorithm. No further criteria are required.

Medium tau lepton candidates are defined in the following way:

- The **"decay mode finding"** criterion is applied, thus a hadronically decaying tau lepton candidate found by the HPS algorithm is required.
- A **medium isolation criterion** which contains a correction for pile-up is applied. This means that in a cone of radius 0.5 the pile-up corrected sum of all particle-flow charged particles and photons with $p_T > 0.5$ GeV/c has to be smaller than 1 GeV/c.

- In order to **avoid misidentification with electrons** the electron pion MVA discriminator, which uses information from the track detector and the calorimeter, has to be smaller than -0.1, the pseudorapidity is not allowed to be in $1.4442 < |\eta| < 1.566$ and a special rejection against Bremmstrahlung is applied.
- To **avoid misidentification with muons** a certain amount of energy has to be deposited in the ECAL and HCAL and the leading track of the tau lepton candidate is not allowed to match to a muon.

4.5 Measurement of $\not{E}_T$ and H_T

Important variables to describe an event are $\not{E}_T$ and H_T. $\not{E}_T$ is the missing transverse momentum in the event, which is produced through particles leaving the detector and through mismeasurements of the energies of the particles in the event. The missing transverse momentum is given by

$$\not{E}_T = \left| - \sum_{\text{jets, particle}-\text{flow objects}} \vec{p}_T \right|. \tag{4.3}$$

In the sum all particle-flow objects over the whole pseudorapidity range are considered. If a particle-flow object has been assigned to a jet which fulfills $p_T > 10$ GeV/c, the energy corrected jet is used for the calculation of the missing transverse momentum instead of the particle-flow object.

H_T is defined by

$$H_T = \left(\sum_{\text{jets,leptons}} |\vec{p}_T| \right) + \not{E}_T. \tag{4.4}$$

For this variable all jets and leptons in the event are considered. The value of H_T is therefore, in contrast to $\not{E}_T$, dependent on the applied selection.

Chapter 5

Analysis

In this chapter the search for third-generation leptoquarks decaying to a top quark and a tau lepton with the CMS experiment is presented. In the first section the data and MC samples used are summarized (section 5.1). Afterwards the trigger employed is introduced (section 5.2). In section 5.3 the event cleaning and the pre-selection are presented. Section 5.4 describes why the p_T-spectrum of the leading tau lepton candidate is used for the final statistical interpretation of the results. In section 5.5 different event selections are introduced and compared to each other. In the following section the determination of the tau lepton fake rate in MC and data is described (section 5.6). In section 5.7 the analysis selections are presented. Section 5.8 describes the systematic uncertainties considered. The final results are presented in section 5.9. The chapter ends with a discussion of possible improvements of the performed analysis in section 5.10.

5.1 Data and MC samples

In this thesis data recorded by the CMS experiment at the LHC are analyzed. The data set collected by the trigger "HLT_IsoMu24_eta2p1", which is described in section 5.2, in the data taking periods A, B, C, and D of the year 2012 is used. The periods correspond to the full dataset of the year 2012. The centre-of-mass energy was $E_{\text{CMS}} = 8$ TeV and the total integrated luminosity of the used sample is $L = 19.6\ \text{fb}^{-1}$ [42, 43].
For the signal, MC samples for third-generation leptoquarks decaying to a top quark and a tau lepton with masses from 200 GeV/c^2 to 1000 GeV/c^2 in steps of 50 GeV/c^2 are used. As background, simulated events for all in section 2.4 mentioned processes are considered. The MC samples used are summarized in table 5.1. Shown are the cross sections σ of all considered processes, the order of the cross section, the MC generator which was used to generate the events and the number of events N. With this information the integrated luminosity $\int L dt$ corresponding to each MC sample can be calculated by

$$\int L dt = \frac{N}{\sigma}. \tag{5.1}$$

Process	cross section [pb]	order	MC generator	number of events
LQ ($M = 200$ GeV/c^2)	17.4	NLO	PHYTHIA 6	50160
LQ ($M = 250$ GeV/c^2)	5.26	NLO	PHYTHIA 6	50160
LQ ($M = 300$ GeV/c^2)	1.89	NLO	PHYTHIA 6	50160
LQ ($M = 350$ GeV/c^2)	0.769	NLO	PHYTHIA 6	50160
LQ ($M = 400$ GeV/c^2)	0.342	NLO	PHYTHIA 6	50160
LQ ($M = 450$ GeV/c^2)	0.163	NLO	PHYTHIA 6	50160
LQ ($M = 500$ GeV/c^2)	0.082	NLO	PHYTHIA 6	50160
LQ ($M = 550$ GeV/c^2)	0.0431	NLO	PHYTHIA 6	50160
LQ ($M = 600$ GeV/c^2)	0.0235	NLO	PHYTHIA 6	50160
LQ ($M = 650$ GeV/c^2)	0.0132	NLO	PHYTHIA 6	50160
LQ ($M = 700$ GeV/c^2)	0.00761	NLO	PHYTHIA 6	50160
LQ ($M = 750$ GeV/c^2)	0.00448	NLO	PHYTHIA 6	50160
LQ ($M = 800$ GeV/c^2)	0.00269	NLO	PHYTHIA 6	50160
LQ ($M = 850$ GeV/c^2)	0.00164	NLO	PHYTHIA 6	50160
LQ ($M = 900$ GeV/c^2)	0.00101	NLO	PHYTHIA 6	50160
LQ ($M = 950$ GeV/c^2)	0.000634	NLO	PHYTHIA 6	50160
LQ ($M = 1$ TeV/c^2)	0.000402	NLO	PHYTHIA 6	50160
$\mathrm{t\bar{t}}$	234	approx. NNLO	POWHEG	21591169
W + jets	37509	NNLO	MADGRAPH + PHYTHIA 6	86108825
Z + jets (Z $\to \ell\ell$, 10 GeV/c^2 < m($\ell\ell$) < 50 GeV/c^2)	850	NNLO	MADGRAPH + PHYTHIA 6	7132223
Z + jets (Z $\to \ell\ell$, $m(\ell\ell) > 50$ GeV/c^2)	3504	NNLO	MADGRAPH + PHYTHIA 6	43815824
Single t, s-channel	3.79	approx. NNLO	POWHEG	259961
Single $\bar{\mathrm{t}}$, s-channel	1.76	approx. NNLO	POWHEG	139974
Single t, t-channel	56.4	approx. NNLO	POWHEG	3758229
Single $\bar{\mathrm{t}}$, t-channel	30.7	approx. NNLO	POWHEG	1935072
Single t, tW production	11.1	approx. NNLO	POWHEG	497658
Single $\bar{\mathrm{t}}$, tW production	11.1	approx. NNLO	POWHEG	493460
WW	54.8	NLO	PHYTHIA 6	10000431
ZZ	8.059	NLO	PHYTHIA 6	9799908
WZ	33.2	NLO	PHYTHIA 6	10000283
$\mathrm{t\bar{t}Z}$	0.2057	NLO	MADGRAPH + PHYTHIA 6	210160
QCD (muon enriched)	362389722.8	LO	PHYTHIA 6	73665106

Table 5.1: Summary of the MC samples used. Shown are the generated process, the total cross section of the process, the order in perturbation theory in which the cross section has been computed, the MC generator used and the number of generated events. The cross sections for the signal samples are taken from [16]. For the other MC samples the cross sections are taken from [41].

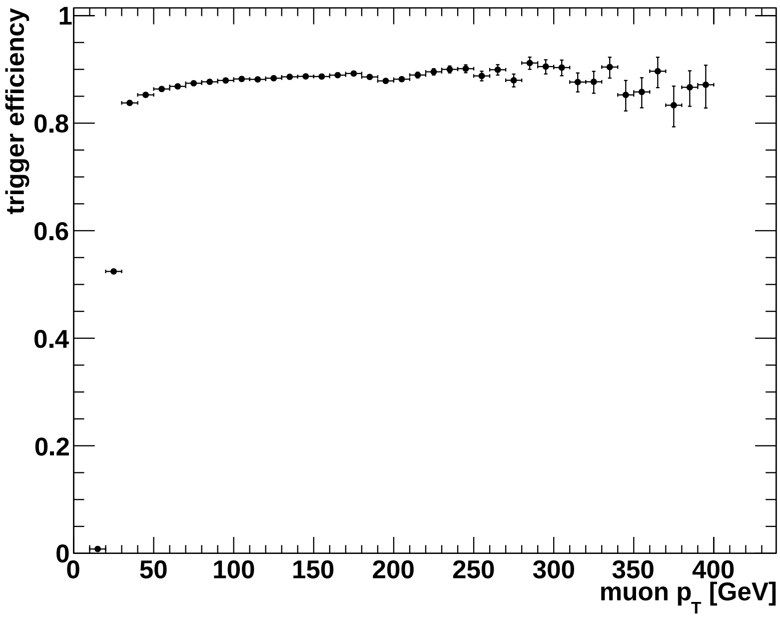


Figure 5.1: Trigger efficiency of the trigger "HLT_IsoMu24_eta2p1" determined in the $t\bar{t}$-MC sample.

All MC events are weighted according to this number and normalized to the integrated luminosity of the data in order to get a sample of background events which can be compared to data.

5.2 Trigger

The data analyzed in this thesis were recorded by a trigger which requires an isolated muon candidate with $p_T > 24$ GeV/c and $|\eta| < 2.1$ ("HLT_IsoMu24_eta2p1"). The isolation in the trigger requirement is defined as the sum of the transverse energy of all particle-flow charged hadrons which were assigned to the primary vertex, all particle-flow neutral hadrons and all particle-flow photons in a cone with radius $\Delta R = 0.5$ around the muon candidate, divided by the transverse momentum of the muon candidate. This value has to be smaller than 0.15 to fulfill the trigger criteria. Isolated muon candidates can be used in this analysis since the top quarks and tau leptons do not have a very high momentum, which means that they are not boosted, and the decay products are thus well separated. The trigger used is very well understood. It has been studied extensively by the muon working group of CMS [44, 45]. They ensured that it worked properly over the full data taking period.

The trigger efficiency in MC is consistent with the efficiency in data determined by CMS [44, 45]. In order to treat all samples in the same way the trigger requirement is applied in data and MC. The trigger efficiency in MC is shown in figure 5.1. It has been determined in the $t\bar{t}$ sample by dividing all triggered events by all events in the MC sample. The trigger efficiency is shown as function of the p_T of the muon candidate with the highest p_T in the event. It can be seen that the turn-on curve starts at $p_T > 10$ GeV/c, the trigger plateau begins at $p_T > 50$ GeV/c. The total efficiency is nearly 90%.

To account for small differences between the two efficiencies scale factors dependent on p_T and η provided by CMS are applied to simulated events [46].

5.3 Event cleaning and pre-selection

As the first step of the analysis an event cleaning is performed. In this step certain quality criteria are applied on the reconstructed particles and the primary vertices.

- For all primary vertices considered in the analysis the number of degrees of freedom in the vertex fit has to be four or higher. The distance from the nominal interaction point to the primary vertex in the z-direction is allowed to be maximal 24 cm. The distance of the primary vertex to the beam axis in the x-y-plane has to be smaller than 2 cm.
- All muon candidates have to fulfill the tight identification criteria described in 4.2. In addition, only muon candidates with $p_T > 30$ GeV/c and $|\eta| < 2.1$ are considered. These cuts are chosen such that the scale factors provided by the muon working group of CMS for the trigger efficiency are applicable [46].
- Jets have to fulfill the loose identification criteria for particle-flow jets described in section 4.3. The jet energy corrections, also described in section 4.3, are applied and the resolution of the jets in MC is deteriorated by roughly 10% by η-dependent factors in order to get the same resolutions of the jets in data and MC [47, 48]. Additionally, all jets have to fulfill $p_T > 30$ GeV/c and $|\eta| < 2.5$.
- For the tau lepton candidates the "decay mode finding" criterion, described in 4.4, $p_T > 20$ GeV/c and $|\eta| < 2.1$ are required.

Additionally, pile-up reweighting [49] is performed to get the same distributions for the number of primary vertices per event in data and MC. In order to do the pile-up reweighting, the luminosity in one lumi-block [1] is calculated. Together with the minimum bias cross section, which is the total inelastic pp cross section (≈ 70 mb), this luminosity is used to calculate the number of interactions in the lumi-block. In MC this number is known, so a scale factor can be calculated for each lumi-block and applied to MC. The scale factors are normalized, so that only the shapes of the distributions not the integrals are changed by applying these factors.

On the cleaned events a pre-selection is applied, which is designed to select as many signal events as possible and reduce the Standard Model background to a reasonable amount. The chosen pre-selection has rather loose cuts, so that later different cuts can be studied to get an optimized event selection.

Each event has to fulfill the following criteria in order to pass the pre-selection:

- The event has to have at least one good primary vertex.

[1] Different runs of data taking are divided into lumi-blocks such that the conditions in the detector were approximately constant in the corresponding time period.

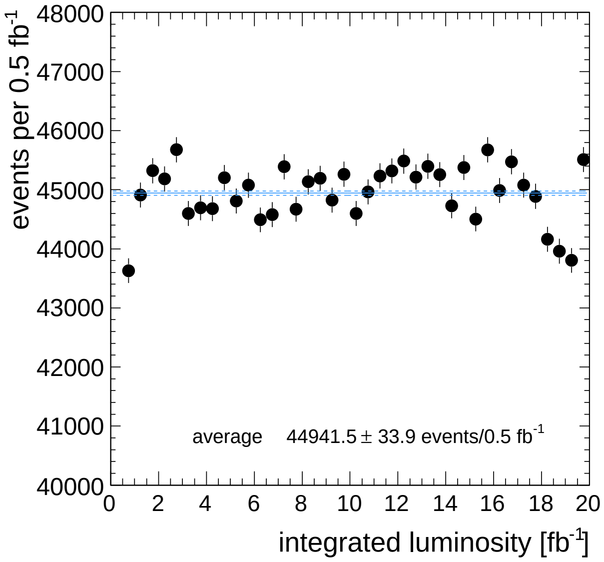


Figure 5.2: Number of events per 0.5 fb^{-1} which passed the pre-selection as function of integrated luminosity.

- There has to be at least one muon candidate.
- The event must have at least two anti-k_t jet candidates with a cone size of $R = 0.5$ which fulfill $p_T > 50$ GeV/c.
- There has to be at least one tau lepton candidate.
- The event has to fulfill $H_T > 350$ GeV/c.

In figure 5.2 the event yield after this pre-selection per integrated luminosity is shown. With this distribution the stability of the selection over time can be checked. The x-axis shows the integrated luminosity where each bin corresponds to 0.5 fb^{-1}. On the y-axis the number of events which fulfill the pre-selection per 0.5 fb^{-1} are plotted. If the conditions were stable during the data taking, the event yield will be constant over time. If there were problems, like a failing detector component or a change in the trigger efficiency, the corresponding data points will deviate from the average value. A fit of the data points with a constant has been done. The result is shown by the blue line. In general, the data points do not show a large deviation from the average. No trend can be observed. This proofs the stability of the performed selection, the conditions of data-taking were stable during the full 2012 running.
Figure 5.3 and 5.4 show control distributions after the pre-selection. In all control distributions the background processes are stacked so that the amount of Standard Model background processes can be compared to data. Moreover, the signal samples for leptoquarks with masses of 300 GeV/c^2 and 600 GeV/c^2 are shown. The grey bands in the ratio plots correspond to the statistical uncertainties of the MC samples.
Shown are the distributions for the number of primary vertices, $\not{E}_T$ and H_T. In addition, the distributions for the tau lepton candidates and the muon candidates with the highest p_T, the number of jets and the number of b-tags in the events are presented. Additionally, the p_T-spectra for the leading jets and the jets with the second and third highest transverse momenta are shown.

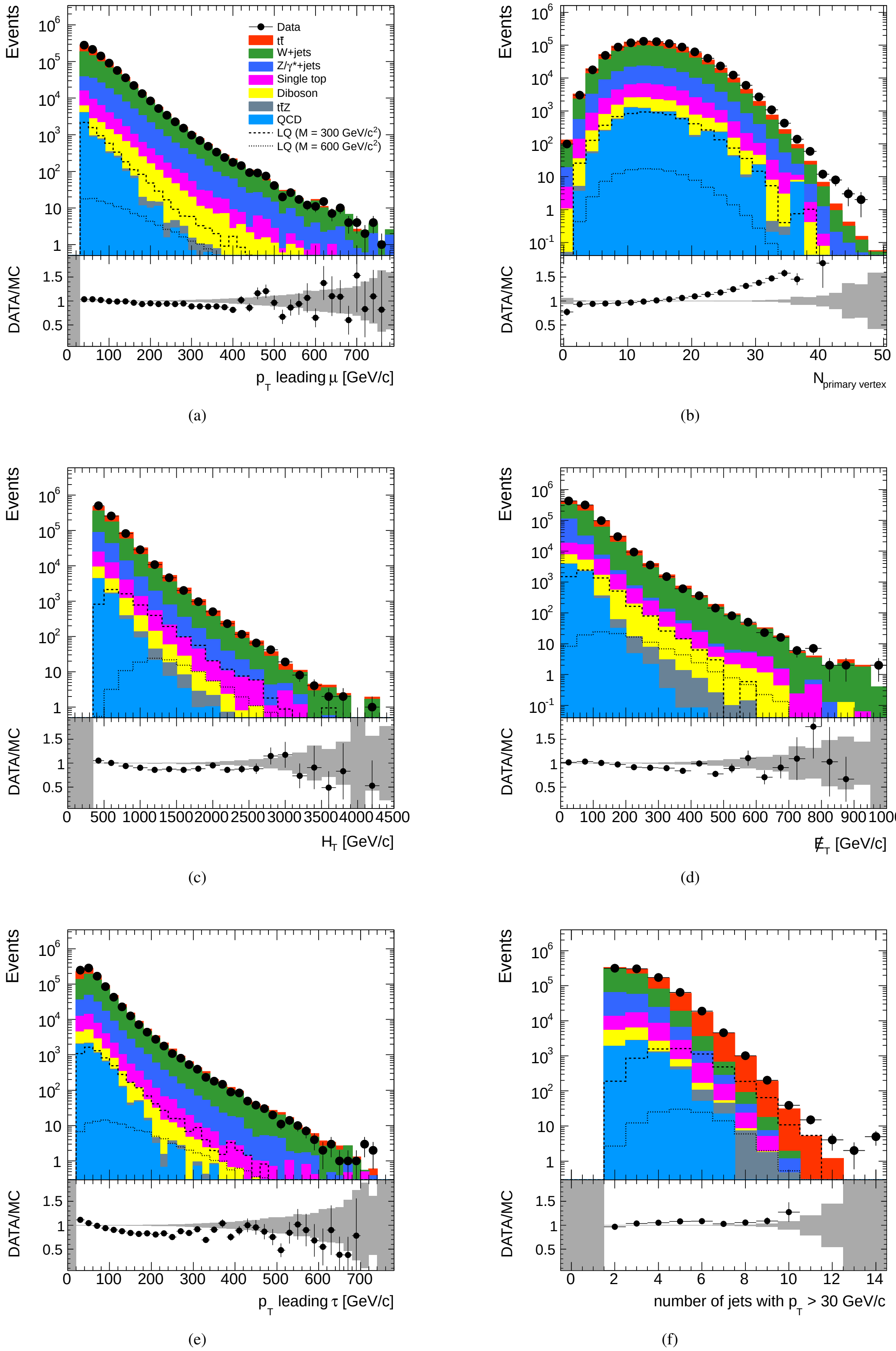
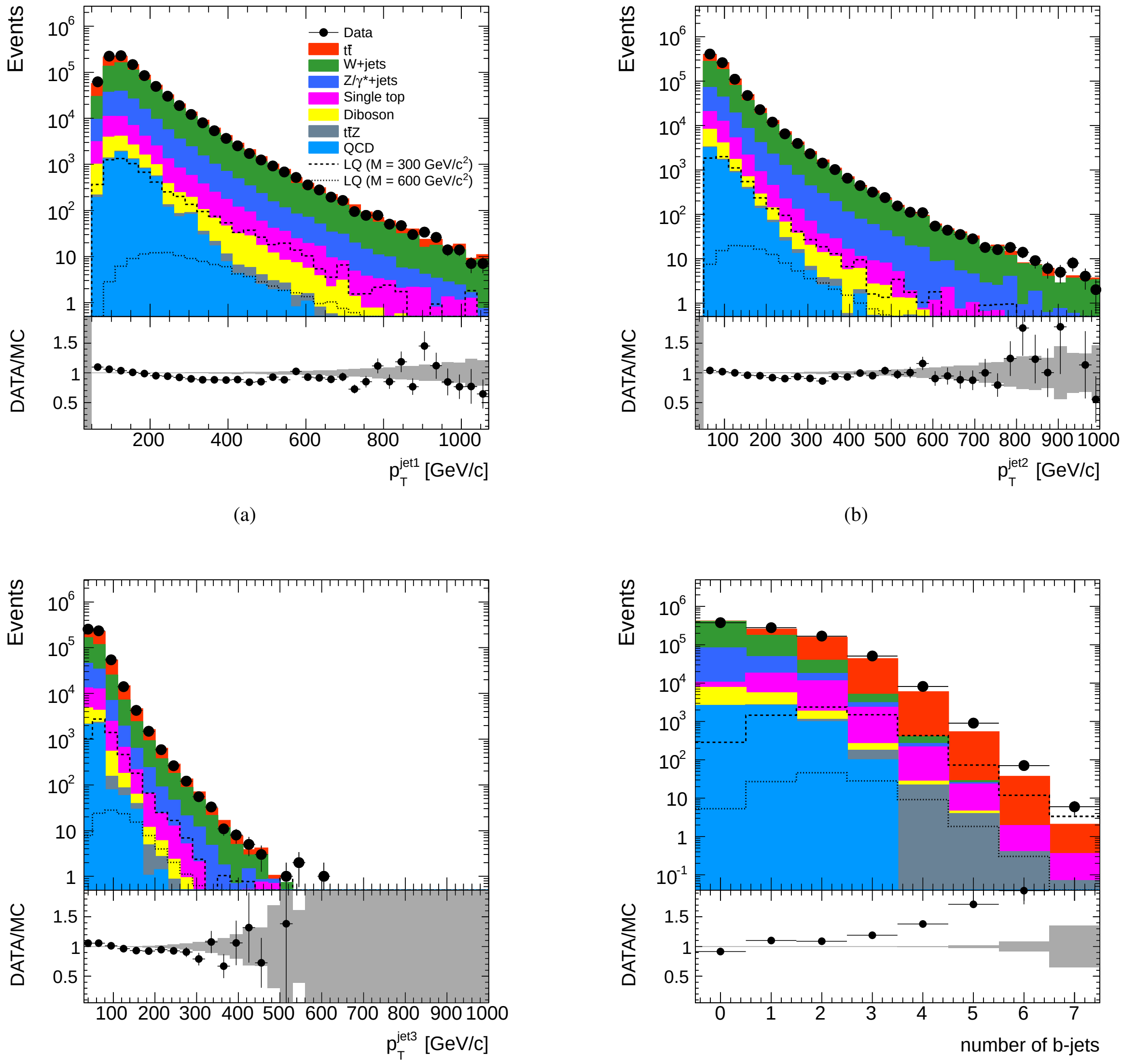


Figure 5.3: Control distributions after the pre-selection. Shown are the distributions for (a) the p_T of the leading muon candidate, (b) the number of primary vertices, (c) H_T, (d) $\not{E}_T$, (e) the p_T of the leading tau lepton candidate and (f) the number of jets.



(a) (b)

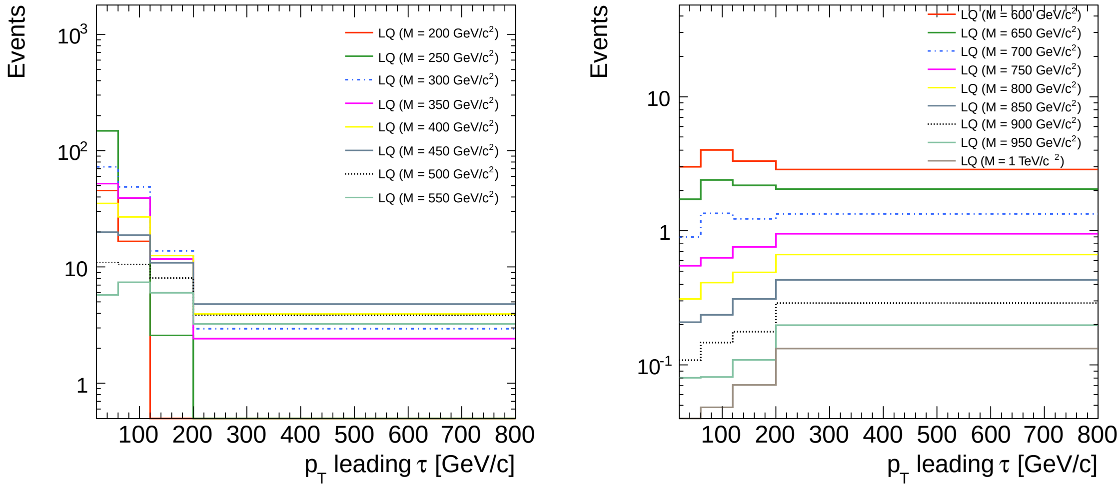
(c) (d)

Figure 5.4: Control distributions after the pre-selection. Shown are the distributions for (a) the leading jet, (b) the jet with the second highest p_T, (c) the jet with the third highest p_T and (d) the number of b-jets, where the loose working point of the CSV-algorithm was used.

In the distribution for the number of primary vertices one can see that the pile-up reweighting works reasonably well. At high values of the number of primary vertices a disagreement between data and MC can be seen. The disagreement disappears if a different minimum bias cross section is used. This will be covered by a systematic uncertainty assigned later.

In general, data and MC agree well in all distributions. One can see that the most important background processes are $\mathrm{t\bar{t}}$-production and W+jets-production. $\mathrm{t\bar{t}}$-production becomes particularly important with increasing number of jets in the event. Another important background is the production of Z+jets. Despite the huge cross section, the background from light quark and gluon production is negligible due to the requirement of one muon candidate in the event.

The selections studied in the following are optimized for leptoquarks with a mass of 600 $\mathrm{GeV/c^2}$. In the control distributions one can see that signal events with these leptoquark masses tend to have higher values of H_T and $\not{E}_T$, more jets and b-jets, and tau lepton candidates and jets with



(a) (b)

Figure 5.5: Distributions for the p_T of the leading tau lepton candidate in the event for signal samples with leptoquarks of masses (a) from 200 GeV/c^2 to 550 GeV/c^2 and (b) from 600 GeV/c^2 to 1 TeV/c^2 after applying the hard selection (see section 5.5).

higher transverse momenta than Standard Model background processes. Thus, with cuts on these variables the Standard Model background can be reduced, while most of the signal events can be kept.

5.4 The p_T distribution of the leading tau lepton

In the performed search the reconstruction of the leptoquark mass is not possible. The main reason is that there are at least two and up to six neutrinos in the event. From these neutrinos only the sum of their transverse momenta can be reconstructed through the measurement of $\not{E}_T$. The directions and the values of the transverse momenta of the individual neutrinos cannot be determined. This information is needed for the reconstruction of the mass of the leptoquark. Moreover, due to the rich final state in the performed search the combinatorics are very large. This makes the reconstruction of M_{LQ} difficult.
However, it has been found that the p_T spectrum of the leading tau lepton candidate in the events shows high sensitivity to the different leptoquark masses and Standard Model backgrounds. This distribution is used for the statistical interpretation of the results. All Standard Model background processes produce tau leptons with low transverse momenta compared to leptoquark decays, while the tau leptons produced in leptoquarks decays tend to have higher transverse momenta. This can be seen in figure 5.24(a), where the transverse momenta of the leading tau leptons for leptoquarks with masses of 300 GeV/c^2 and 600 GeV/c^2 and for the Standard Model background processes are shown after applying the hard selection described in 5.5. Due to the different shapes of the distributions for the different Standard Model processes and the signal samples a good separation between background and signal events can be reached, which is essential for the statistical interpretation of the search.

A comparison of the p_T-distributions of the leading tau lepton candidates for the different signal samples in figure 5.5 shows that the shapes and rates differ considerably. For example, the signal for a leptoquark with a mass of 250 GeV/c^2 peaks at a transverse momentum of 60 GeV/c, while there are no tau lepton candidates with p_T above 200 GeV/c. The signal for a leptoquark with a mass of 550 GeV/c^2, peaks at a transverse momentum of around 100 GeV/c with a considerable tail at higher p_T. Thus, the p_T spectrum of the leading tau lepton candidate and therefore the sensitivity is highly dependent on the leptoquark's mass, which is good for the final interpretation of the result.
Different variables have been studied, but the p_T-distribution of the leading tau lepton candidate in the event showed the best separation between signal and background together with the best sensitivity for different leptoquark masses. The following studies on the optimization of the selection are therefore based on this distribution and it will also be used for the final interpretation of the results of this search.

5.5 Optimization of the event selection

Since the performed search has not been done before, first a selection has to be found which suppresses the Standard Model backgrounds but has a high signal efficiency. Due to the rich final state, there are many possibilities to select events based on cuts on different final state particles. Therefore, five different selections are studied in this analysis. The optimization of the selections studied is based on the expected limit calculated with the theta package [50]. As input to theta the distributions of the p_T of the leading tau lepton candidate in the event is used. Two of the five selections studied are rather loose while in the other three selections in general the same but harder cuts are used.
The studied selections are called soft selection, soft selection with same-sign requirement, medium selection, medium selection with same-sign requirement and hard selection in the following. They are based on the pre-selection described in 5.3. The applied cuts are listed in the following.

- For the **soft selection** the requirements are:
  - At least one medium tau lepton candidate, described in 4.4, is identified in the event. Additionally, all tau lepton candidates which are considered fulfill the medium working point. All tau lepton candidates which have a distance $\Delta R < 0.5$ to a muon candidate are removed in order to reject fake tau lepton candidates which were also reconstructed as muon candidates.
  - A third anti-k_t jet with a cone size of $R = 0.5$, which fulfills $p_T > 30$ GeV/c and $\eta < 2.5$ is present. In order to remove jets which have also been reconstructed as tau lepton candidates or muon candidates all jets which have a distance $\Delta R < 0.5$ to a tau lepton candidate or a muon candidate are removed.

  - The missing transverse momentum in the event has to fulfill $\not{E}_T > 50$ GeV/c.
  - At least one jet has to have a b-tag, where the tight working point of the CSV-algorithm, described in 4.3, is used.

- For the **medium selection** the events have to pass the soft selection without the b-tag criterion. Additionally, the events have to fulfill the following cuts:
  - the leading jet must have $p_T > 100$ GeV/c and
  - $H_T > 400$ GeV/c.

- In order to pass the **hard selection** the events have to to fulfill the medium selection criteria and
  - the leading jet must have $p_T > 150$ GeV/c,
  - the jet with third highest transverse momentum has to fulfill $p_T > 50$ GeV/c,
  - $H_T > 700$ GeV/c and
  - $\not{E}_T > 100$ GeV/c.

Two thirds of the background events contain events with one tau lepton candidate and one muon candidate with opposite-sign charges. Thus, by requiring two lepton candidates [2] with same-sign charges reduces the background by two thirds. However, also the signal would be reduced by one half. An additional advantage of this cut would be that afterwards the background is dominated by fake tau leptons, which could make an estimation of the number of background events easier. In order to find out if this selection would help to improve the sensitivity, two additional selections are studied: the **soft selection with same-sign requirement** and the **medium selection with same-sign requirement**. In these selections the same-sign requirement is applied to the soft and the medium selections and the resulting sensitivity is checked.

5.5.1 Comparison of the expected limits

In order to find out which selection gives the best result, the expected limits on the production cross section of third-generation leptoquarks times branching ratio after applying the selections are compared. The expected limits are calculated using the theta package [50]. A bayesian method, which is explained in [51], is used. Only statistical uncertainties are considered in the limit calculation, the background is determined completely from simulated events. As input the distributions for the p_T of the leading tau lepton candidates in the events after the five selections studied have been applied are used. These distributions are shown in figure 5.6. In figures 5.6(a) and 5.6(c) the input histograms for the limit calculation for the soft and the medium selection are shown. The difference between these selections are harder cuts on the p_T of the leading jet, on H_T and the requirement of a jet with a b-tag. In the soft selection the background is reduced significantly through the requirement of a b-tag while in the medium selection the background

[2] Here lepton means muon candidate and or tau lepton candidate

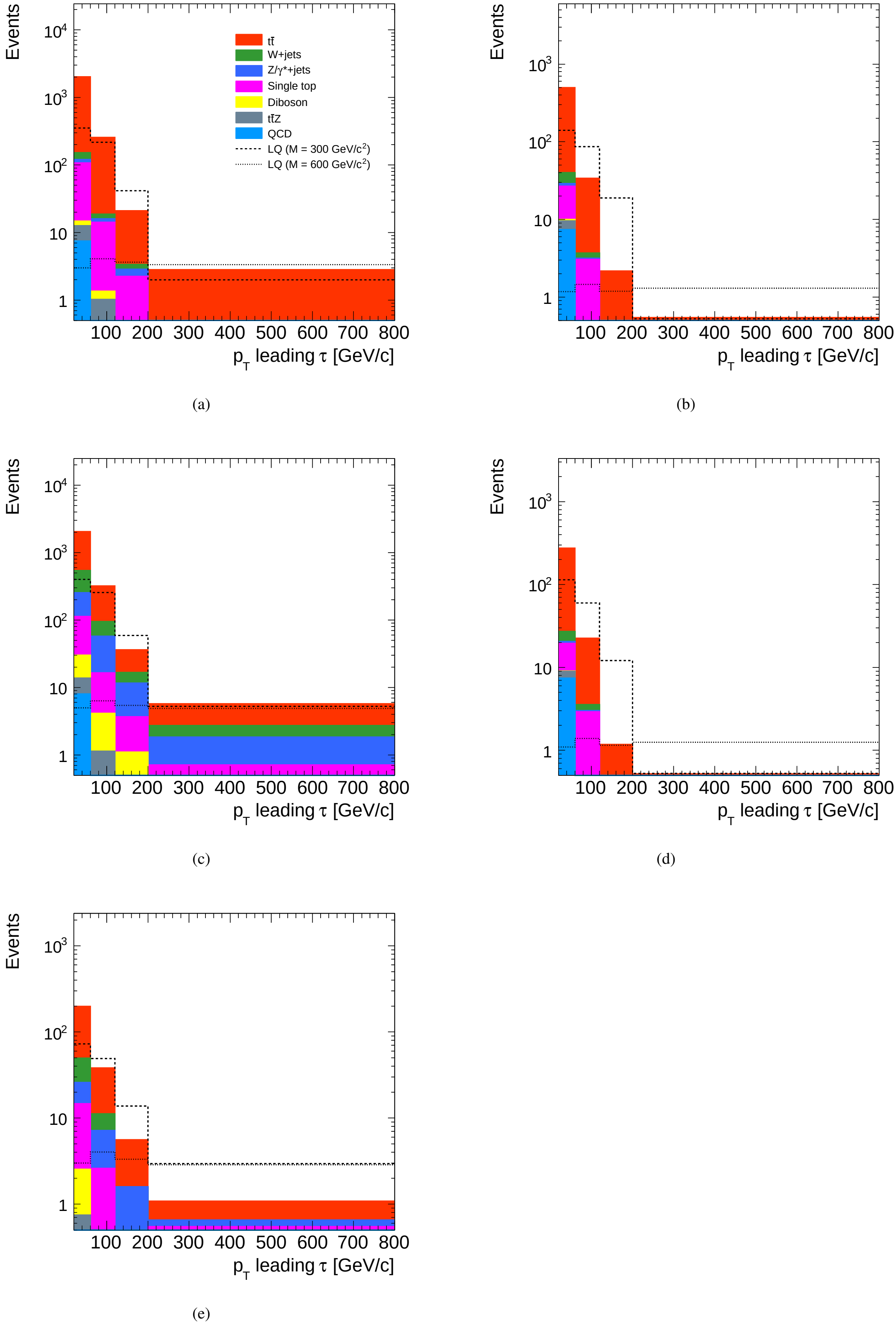


Figure 5.6: Distributions of the p_T spectra of the leading tau lepton in the event after applying (a) the soft selection, (b) the soft selection with the same-sign requirement, (c) the medium selection, (d) the medium selection with the same-sign requirement and (e) the hard selection.

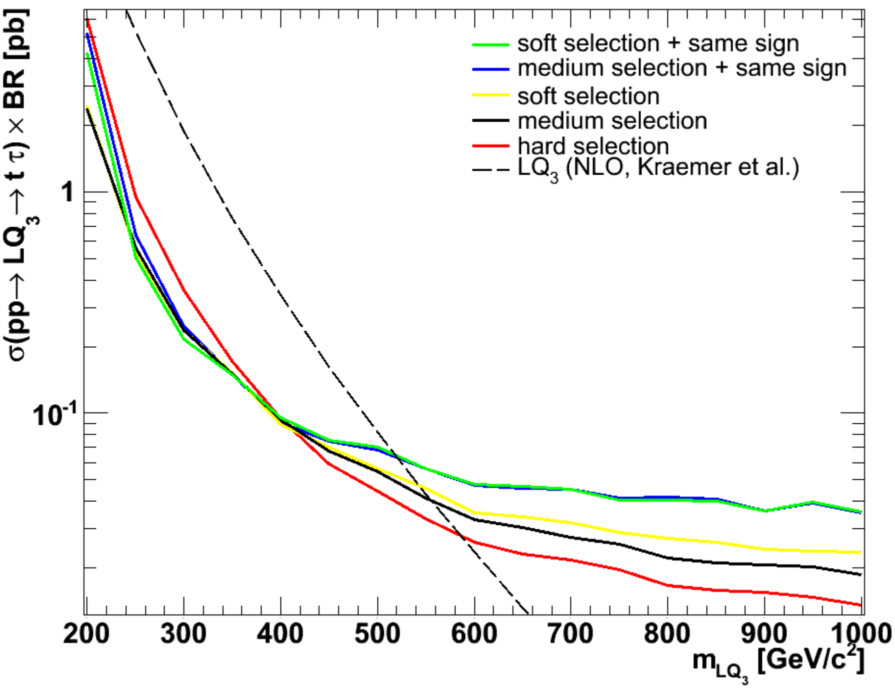


Figure 5.7: Comparison of the expected limits on the cross section times branching ratio of the different studied selections as described in 5.5. The theory curve displays the next-to-leading order cross section of the studied process [16].

is reduced by a similar amount through the harder cuts on H_T and the p_T of the leading jet. Due to the requirement of a b-tag, the main background after applying the soft selection is $t\bar{t}$-production, while the main backgrounds after applying the medium selection are $t\bar{t}$-production, W+jet- and DY-events. The distributions after applying the same-sign cuts in addition to the soft selection and the medium selection are shown in figures 5.6(b) and 5.6(d). In both cases the background is reduced significantly through this additional cut. For example, after the selections with the same-sign requirement only one event with $p_T > 200$ GeV/c is expected for leptoquarks with masses of 600 GeV/c^2, whereas 3-5 events are expected for the selections without this requirement. In figure 5.6(e) the p_T distribution of the leading tau lepton candidate is shown after the hard selection has been applied. The harder cuts on jet p_T and H_T reduce the Standard Model background over the full p_T spectrum. The signal is reduced at lower values of p_T but retained at high values of p_T. This results in an increase of the sensitivity on higher leptoquark masses and a decrease of the sensitivity at lower masses.

In figure 5.7 the expected limits on the cross section times branching ratio are shown as a function of the leptoquark's mass. The cross sections above the expected limits can be excluded at 95% C.L.. The theory curve displays the next-to-leading order cross section of the studied process [16]. By comparing the limit with the theory cross sections the limits on the production cross section can be converted into an upper limit on the mass of the leptoquarks.

It can be observed that the selections with the same sign requirement give the worst limits almost over the whole leptoquark mass range. In general the soft and the medium selections give similar limits. The hard selection gives the worst limit for leptoquark masses smaller than

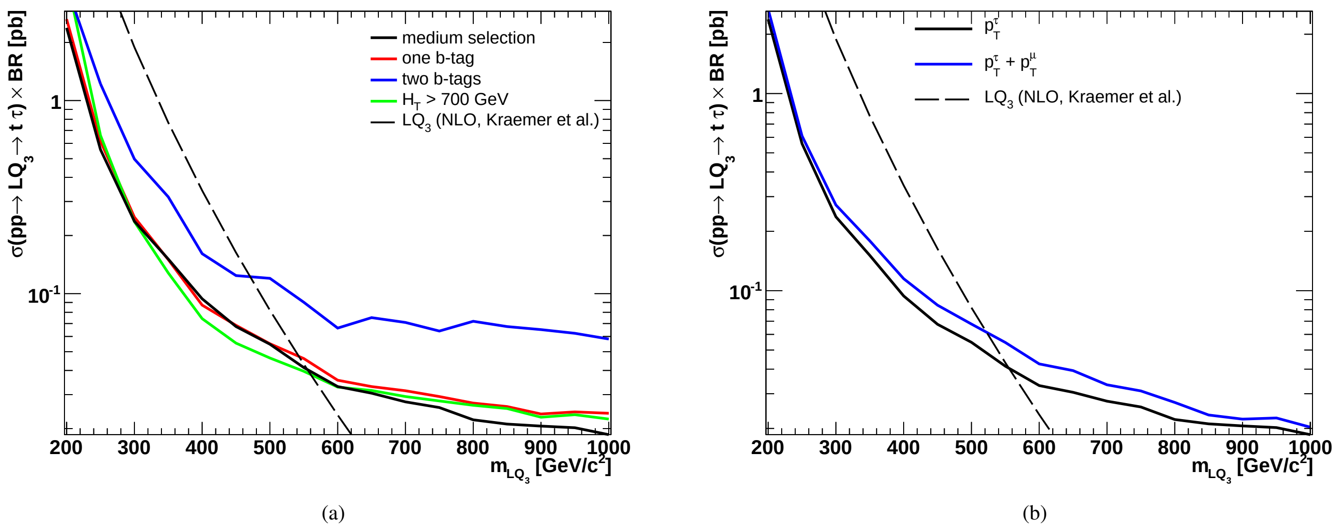


Figure 5.8: (a) Comparison of the expected limits of the medium selection with different selection criteria as described in 5.5.2. (b) Comparison of the expected limits based on the distributions of the p_T of the leading tau lepton candidate and of the sum of the p_T of the leading tau lepton candidate and the muon candidate after applying the medium selection. The theory curve displays the next-to-leading order cross section of the studied process [16].

400 GeV/c^2, but for masses higher than 400 GeV/c^2 the expected limit gets best.
The result of this study is that the best expected limits can be reached by taking the medium selection for leptoquark masses lower than 400 GeV/c^2 and the hard selection for masses higher than 400 GeV/c^2. In general the medium and the soft selection give almost the same results for lower leptoquark masses. The medium selection is chosen because the cuts on the jets are higher so that the effect from the jet energy scale uncertainty is expected to be smaller.

5.5.2 Further optimization of the selections

In order to see if the selections can be optimized further, some of the cuts of the medium selection are varied. The result is shown in figure 5.8(a). The expected limits on the cross section times branching ratio are shown for the medium selection with a cut on $H_T > 700$ GeV/c instead of $H_T > 400$ GeV/c. The limit is improved by this cut for leptoquark masses higher than approximately 300 GeV/c^2. This shows again that by a higher cut on H_T the sensitivity is improved for higher masses and deteriorated for lower masses, which can be also explained by looking at figure 5.3(c). Here one can see that the signal for higher leptoquark masses is concentrated at higher H_T-values and that most of the signal at lower leptoquark masses has lower H_T-values.
The second cut which is varied is the b-tag requirement. The expected limit is calculated with the requirement of one and of two b-tags in the event. For the b-tag the tight working point of the CSV-algorithm is used. If one b-tag is required, it can be seen that the change in the limit for leptoquark masses lower than 500 GeV/c^2 is negligible compared to the expected uncertainties. For higher leptoquark masses the medium selection without the b-tag requirement

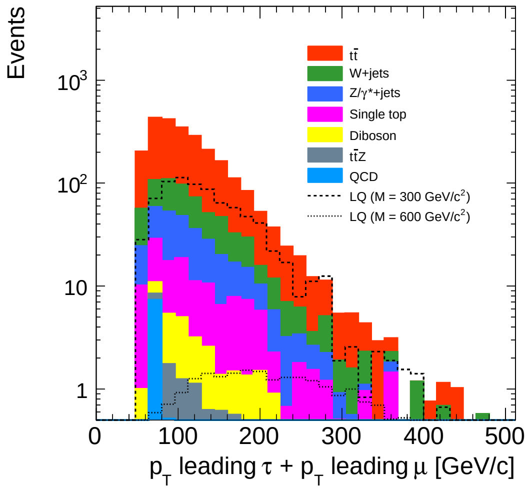


Figure 5.9: Distribution of $p_T^\tau + p_T^\mu$, where the leading tau lepton candidate and the leading muon candidate are considered, after applying the medium selection.

gives the better limit. With two b-tags the limit gets worse over the whole mass range, since too much signal efficiency is lost. From this study it can be concluded that the medium selection without any b-tag requirement has the highest sensitivity. This is an advantage since in this case no b-tagging uncertainties have to be considered.
An alternative approach to optimize the expected limit would be to change the variable on which the limit calculation is based. The variable $p_T^\tau + p_T^\mu$ could be an interesting variable for this. The distributions after the pre-selection is shown in figure 5.9. The expected limits for the medium selection based on p_T^τ and $p_T^\tau + p_T^\mu$, where always the leading tau lepton candidate and the leading muon in the event are considered, are compared in figure 5.8(b). It can be seen that the limit is better if it is calculated based on the p_T^τ-spectrum, especially at masses between 300 GeV/c^2 and 700 GeV/c^2, where this analysis has the highest sensitivity.
As a conclusion the medium selection is used for low leptoquark masses. For high masses the hard selection is used. The transition is chosen based on the best expected limit. No combination of the outcome of the two selections is performed since the two samples are highly statistically correlated. The limit will be calculated based on the p_T^τ-spectra of the leading tau lepton candidate.

5.6 Determination of the tau lepton fake rate in data and simulation

For the determination of the fake rate a background sample is needed with a negligible contribution of prompt tau leptons. The goal is to get a clean W+jets sample which is kinematically similar to the final analysis selection. The same data and MC samples and the same trigger are used, as described in the previous sections (5.1 and 5.2). The event cleaning is done in the same way as described in section 5.3. To pass the selection the events are required to have:

- at least one good primary vertex,

- exactly one tight muon candidate,
- $\not{E}_T > 40\,\mathrm{GeV/c}$,
- at least two anti-k_t jets with a cone size of $R = 0.5$,
- exactly one tau lepton candidate, where only the decay mode finding criterion is applied,
- $M(\mu\tau) < 80\,\mathrm{GeV/c^2}$ and $M(\mu\tau) > 100\,\mathrm{GeV/c^2}$.

In addition, a veto against b-tags is applied, where the loose working point of the CSV-algorithm is used in order to reject the b-jets as efficiently as possible. By requiring exactly one muon candidate it is ensured that the W-boson decays into a muon (not into a tau). Additionally, the requirement for the missing transverse momentum helps to select only events were the muon candidate is produced in the decay of the W-boson and not in a jet from the decay of b- or c-hadrons. All these cuts are chosen in a way that the background and the signal events are reduced as much as possible and that the selection is kinematically similar to the analysis selection (described in 5.3). The veto against b-tags helps to reduce the background from $t\bar{t}$. This has to be done because in a leptonically decaying $t\bar{t}$-pair real tau leptons can be produced in addition to the muon. The cut on the invariant mass of the tau lepton candidate and the muon candidate reduces the background from DY-processes since events where the Z-boson decays into two tau leptons and one tau lepton decays subsequently into a muon are rejected.

In this way a quite clean W+jets sample can be obtained. The control distributions for data and simulated events are shown in figure 5.10. The remaining background processes to the W+jets sample are mainly Z+jets events, but also $t\bar{t}$, single top, diboson events and a small contribution from QCD events are present. Furthermore, it can be seen that only a negligible amount of signal events is left after the selection.

The tau leptons studied are identified through their hadronic decay. Thus, the tau lepton fake rate of jets is particularly important. Since the statistical interpretation of the result will be based on the p_T spectrum of the leading tau lepton candidate, it is important that the tau lepton fake rate is well simulated. Therefore, the tau lepton fake rate has to be determined in data and MC.

Since most of the selected events are W+jets events, one can be sure that in the resulting sample most of the selected tau lepton candidates are fakes. To estimate the relative contribution of fake tau leptons, a matching of the selected tau lepton candidates to generated tau leptons is done on the MC samples [3]. The events with a fake tau lepton candidate are shown as a function of p_T and η of the leading tau lepton candidate in figure 5.11. The events which contain a real tau lepton are shown in figure 5.12. It can be seen that events with a real tau lepton make up only approximately 1% of all selected events. The events which contain a real tau lepton candidate are mostly DY events.

In the next step a medium tau lepton candidate is required in addition to the described selection cuts. In this way the fake rate of the medium working point can be calculated with respect to

[3] If a generated tau lepton is found in a distance of $\Delta R < 0.5$ to a selected tau lepton candidate, the tau lepton candidate is considered to come from a real tau lepton.

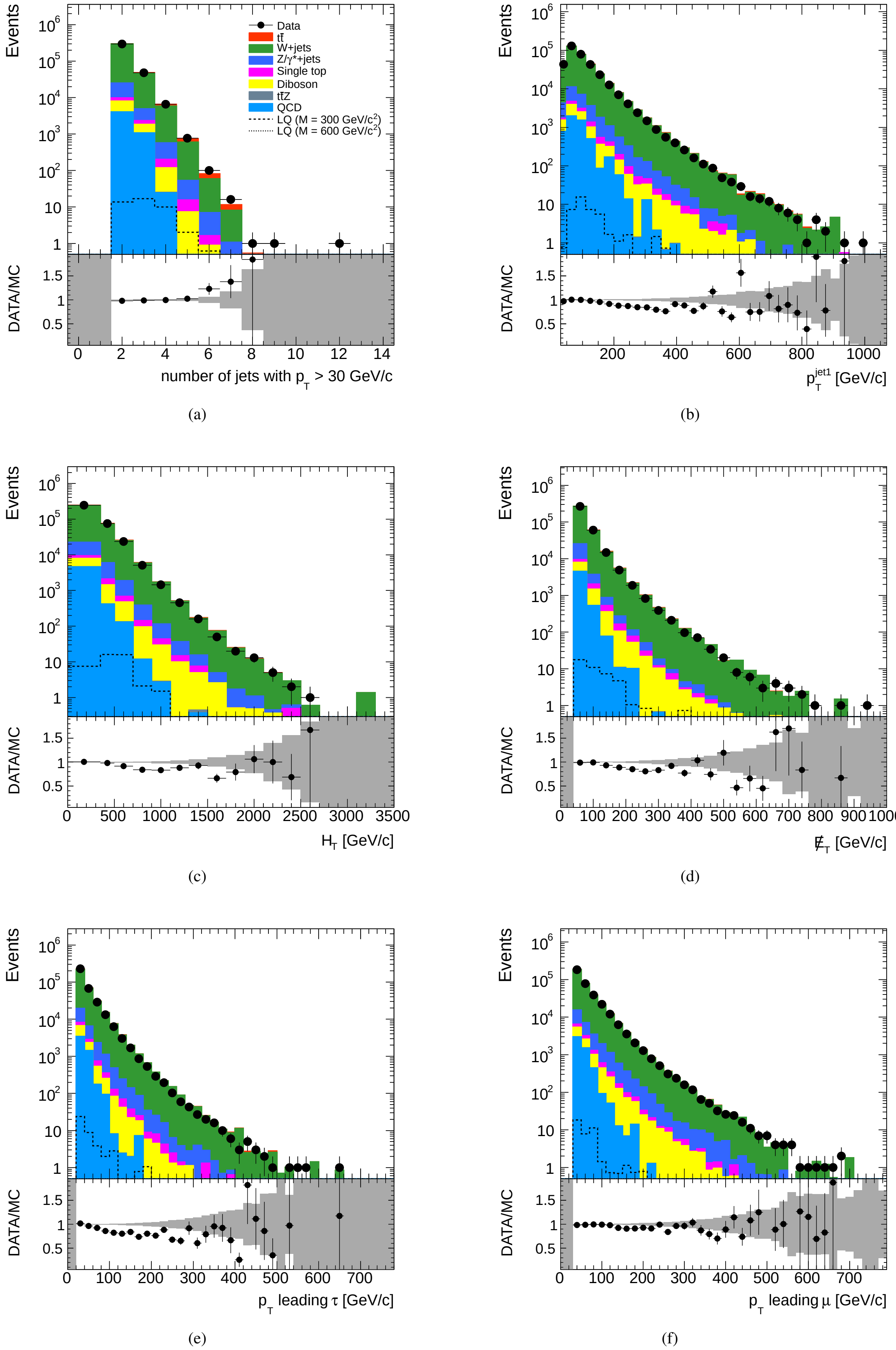
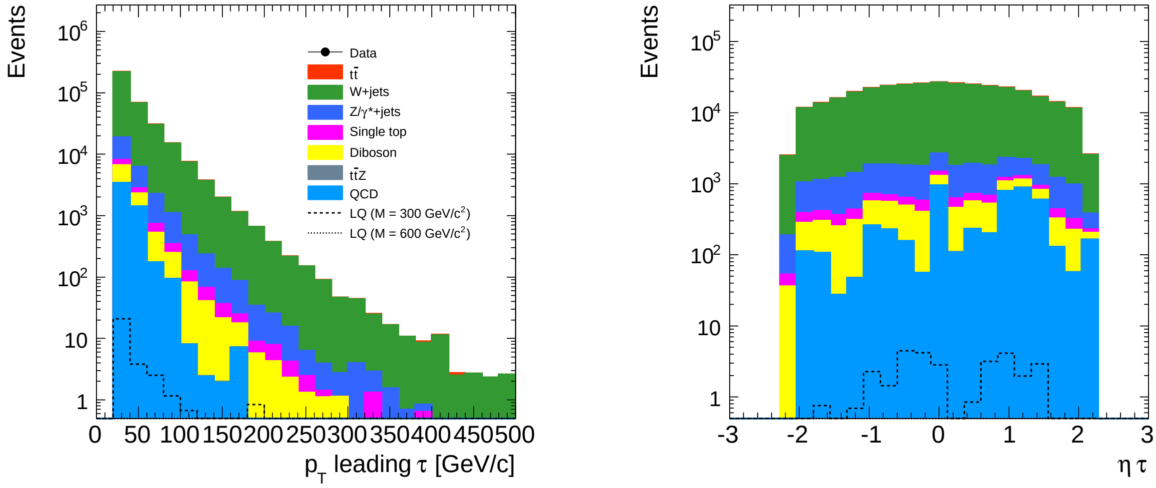
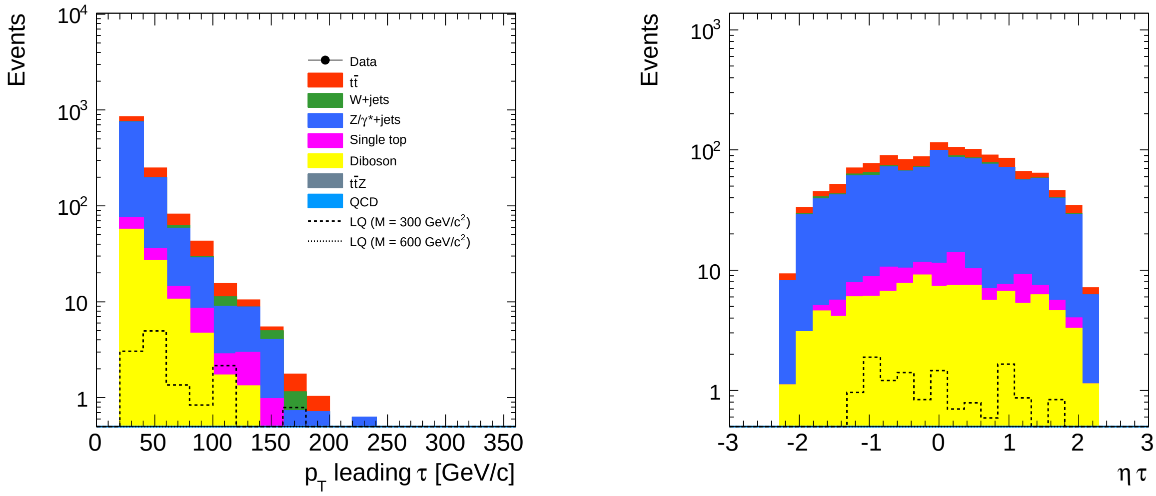


Figure 5.10: Control distributions for the background enriched sample: (a) the number of jets, (b) the p_T of the leading jets, (c) H_T, (d) $\not\!\!E_T$, (e) the p_T of the leading tau lepton candidate and (f) the p_T of the leading muon candidate after applying the selection described in 5.6.



(a) (b)

Figure 5.11: Distributions after applying the selection described in 5.6 of (a) the p_T of the leading tau lepton candidate and (b) η of the leading tau lepton candidate, where only events which contain a fake tau lepton are shown.



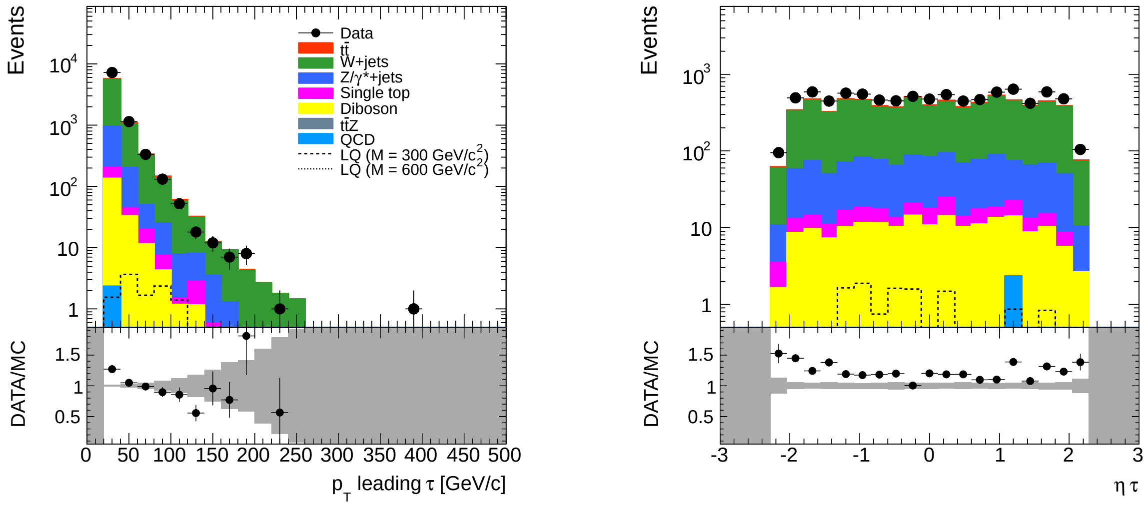
(a) (b)

Figure 5.12: Distributions after applying the selection described in 5.6 of (a) the p_T of the leading tau lepton candidate and (b) η of the leading tau lepton candidate, where only events which contain a real tau lepton are shown.

the "decay mode finding"-criterion by applying

$$\varepsilon = \frac{\text{\#events(medium tau lepton)}}{\text{\#events(tau, only decay mode finding applied)}}, \tag{5.2}$$

where #events(medium tau lepton) is the number of events which contain a medium tau lepton candidate and #events(tau, only decay mode finding applied) corresponds to the number of events which contain a tau lepton candidate which fulfills the "decay mode finding criterion". Here it is assumed that the fake rate of the "decay mode finding" criterion is the same in data and MC.



(a) (b)

Figure 5.13: Control distributions after requiring a medium tau lepton candidate of (a) the p_T of the leading tau lepton candidate and (b) η of the leading tau lepton candidate.

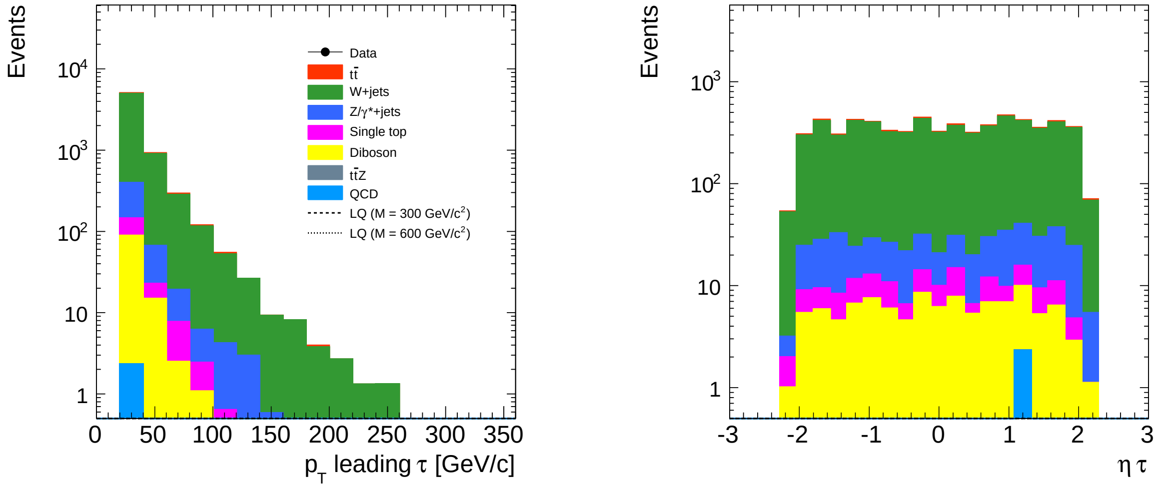
The distributions after the requirement of a medium tau lepton candidate are shown in figure 5.13. Data and MC do not agree very well in these distributions. This is due to the different fake rates in data and MC. The amount of fake tau leptons in the events is shown in figure 5.14. After selecting medium tau lepton candidates, the contribution of real tau leptons is with roughly 17% quite high, as can be seen in figure 5.15, where the events which contain a real tau lepton are shown. In order to be able to apply equation 5.2 the real tau leptons have to be subtracted from the total number of events, where the number of real tau leptons has to be taken from MC since it cannot be determined in data. The fake rate can then be calculated by

$$\varepsilon = \frac{\#\text{events}(\text{medium tau lepton}) - \#\text{events}(\text{real tau lepton})}{\#\text{events}(\text{tau lepton, only decay mode finding applied}) - \#\text{events}(\text{real tau lepton})}, \quad (5.3)$$

where #events(real tau lepton) is the number of MC events where a real tau lepton candidate has been found.

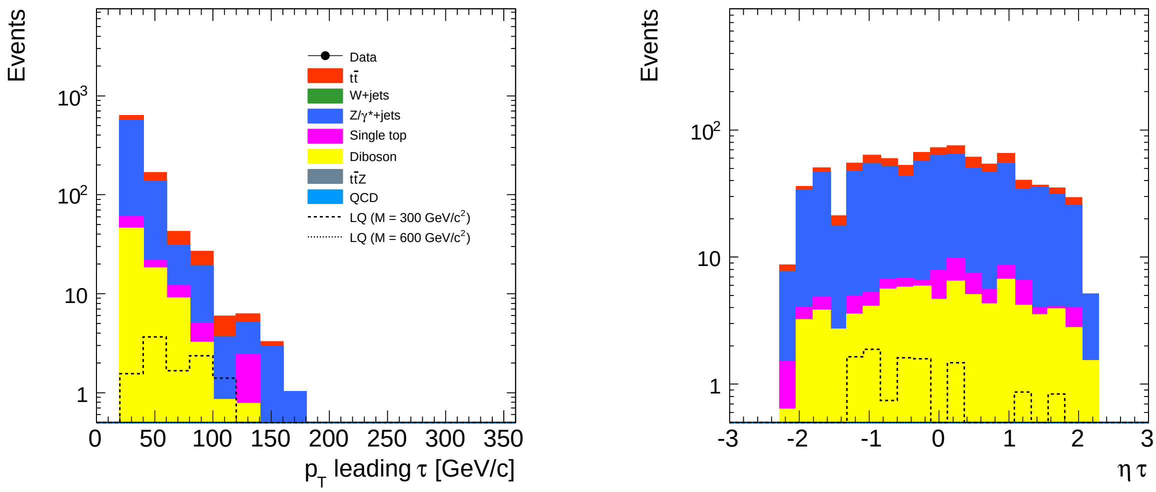
The fake rate can be calculated in data and MC as a function of different variables by dividing the corresponding histograms. In figure 5.16 the fake rates are shown as a function of the p_T and η of the leading tau lepton candidate. In both distributions it can be seen that the mistag rates are different in data and MC and of the order of a few per cent.

Since the number of real tau lepton candidates in data could deviate from the one in MC, an uncertainty has to be assigned on the subtraction of the real tau lepton candidates. Since most of the real tau lepton candidates originate from the remaining Z +jet events, the normalization of this sample has been varied by $\pm 50\%$ and the fake rate has been calculated again with the formula from equation 5.3. The result is shown in figure 5.17. The difference between the fake rates before and after the variation are calculated and taken as systematic uncertainties.



(a) (b)

Figure 5.14: Distributions after requiring a medium tau lepton candidate of (a) the p_T of the leading tau lepton candidate and (b) η of the leading tau lepton candidate where only events which contain a fake tau lepton candidate are shown.



(a) (b)

Figure 5.15: Distributions after requiring a medium tau lepton candidate of (a) the p_T of the leading tau lepton candidate and (b) η of the leading tau lepton candidate where only events which contain a real tau lepton candidate are shown.

The obtained values for the fake rates in MC are

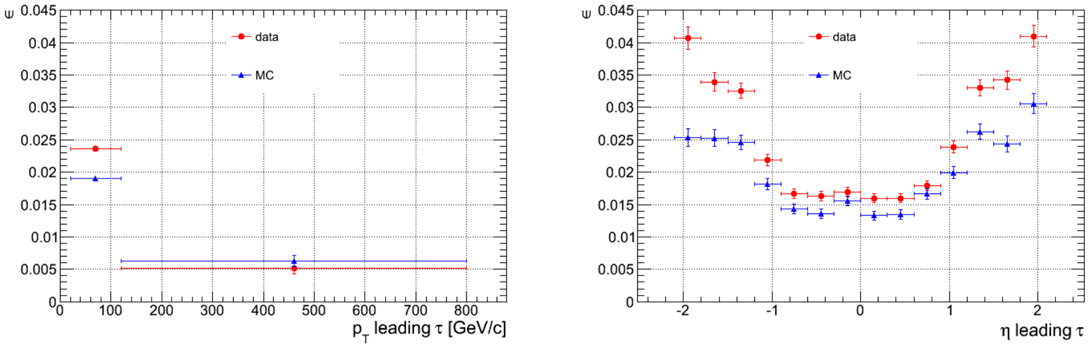
$$\varepsilon \approx 0.0190 \pm 0.0003(\mathrm{stat})\ ^{+0.0012}_{-0.0013}(\mathrm{sys}) \text{ for } \mathrm{p}_\mathrm{T}^\tau < 120\ \mathrm{GeV/c} \tag{5.4}$$

$$\varepsilon \approx 0.0062\ ^{+0.0009}_{-0.0008}(\mathrm{stat})\ ^{+0.0007}_{-0.0006}(\mathrm{sys}) \text{ for } \mathrm{p}_\mathrm{T}^\tau > 120\ \mathrm{GeV/c}. \tag{5.5}$$

For data the result is

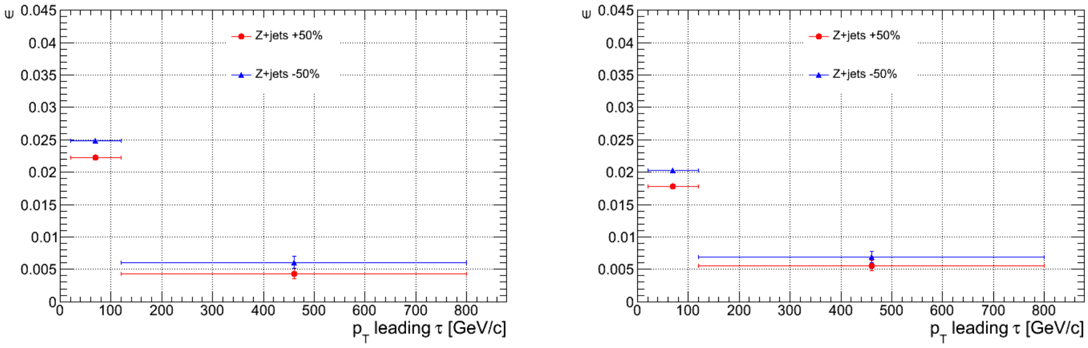
$$\varepsilon \approx 0.0236 \pm 0.0003(\mathrm{stat})\ ^{+0.0012}_{-0.0013}(\mathrm{sys}) \text{ for } \mathrm{p}_\mathrm{T}^\tau < 120\ \mathrm{GeV/c} \tag{5.6}$$

$$\varepsilon \approx 0.0052\ ^{+0.0009}_{-0.0008}(\mathrm{stat})\ ^{+0.0008}_{-0.0009}(\mathrm{sys}) \text{ for } \mathrm{p}_\mathrm{T}^\tau > 120\ \mathrm{GeV/c}. \tag{5.7}$$



(a) (b)

Figure 5.16: Tau lepton fake rates as a function of (a) the p_T of the leading tau lepton candidate and (b) η of the leading tau lepton candidate.



(a) (b)

Figure 5.17: Tau lepton fake rates as a function of the p_T of the leading tau lepton candidate with a variation of the normalization of the DY-background by 50% up and down in (a) data and (b) MC.

From these obtained fake rates scale factors for the simulation can be derived by dividing the fake rates in data by the ones in MC. Since the limit will be derived from the p_T distribution of the leading tau lepton candidate and thus this is the most important variable the scale factors are calculated for this variable. The result is shown in figure 5.18. The scale factors are

$$s \approx 1.2405 \pm 0.0215(\text{stat}) \; ^{+0.1007}_{-0.1091}(\text{sys}) \text{ for } \mathrm{p}_{\mathrm{T}}^{\tau} < 120\,\mathrm{GeV/c}, \tag{5.8}$$

$$s \approx 0.8298 \; ^{+0.1823}_{-0.1763}(\text{stat}) \; ^{+0.1601}_{-0.1663}(\text{sys}) \text{ for } \mathrm{p}_{\mathrm{T}}^{\tau} > 120\,\mathrm{GeVc}. \tag{5.9}$$

These factors can be applied to the fake tau leptons in MC in order to correct for the different fake rates in data and MC.
As a cross check the obtained scale factors can be applied on the W+jets sample which was used to determine the factors. The fake rates can then be calculated again with the formula in equation 5.3. If data and MC agree afterwards, the method works fine and no further scale factors in dependence of other variables have to be determined. The result of this cross check is shown in figure 5.19, where it can be observed that the fake rates agree in data and MC within the errors. Thus, the method is fine and the scale factors can be applied to the sample described

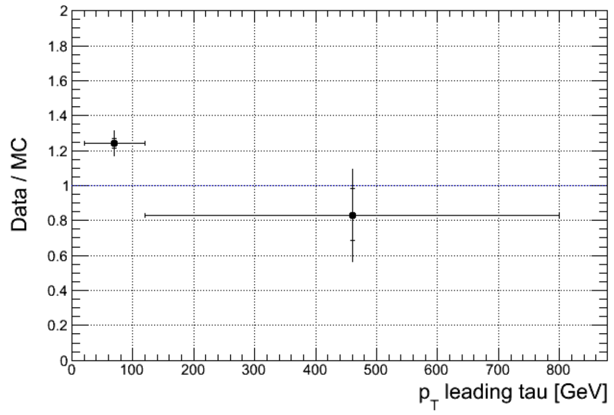
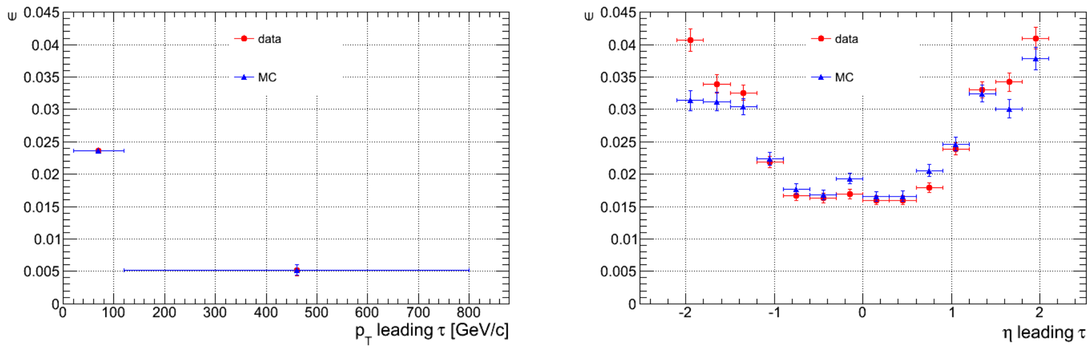


Figure 5.18: Scale factors for the MC samples to account for the different fake rates in data and MC as a function of the p_T distribution of the leading tau lepton candidate. The inner error bars represent the statistical uncertainties, the outer error bars the total uncertainty.



(a) (b)

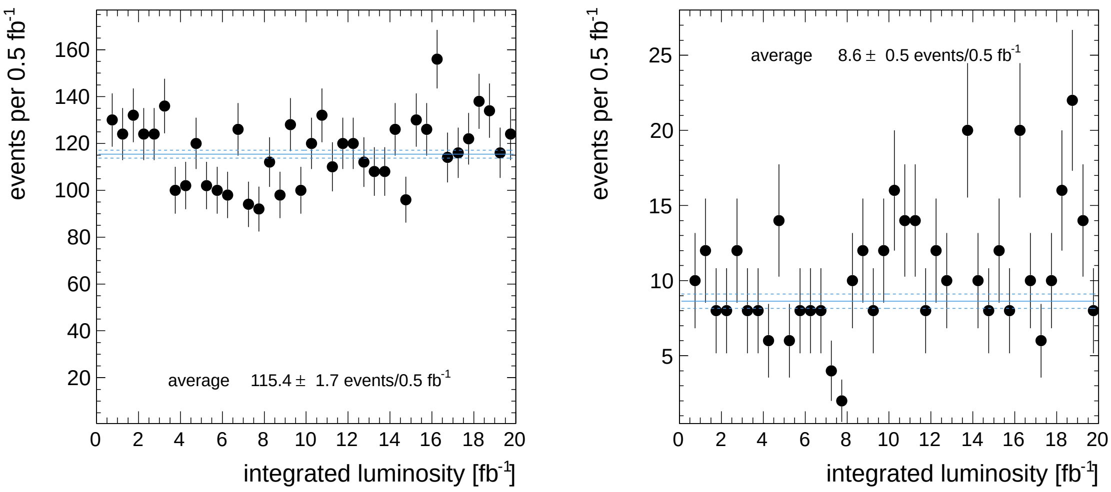
Figure 5.19: Tau lepton fake rates as a function of (a) the p_T of the leading tau lepton candidate and (b) η of the leading tau lepton candidate after applying the obtained scale factors on the fake tau leptons in MC.

in 5.3.

The tau lepton identification efficiency has been studied in simulated events in a different analysis [52]. The only difference in the efficiency of the leptoquark reconstruction in the signal sample and the Standard Model processes is the tau lepton isolation, which is well described by the MC.

5.7 Analysis selections

The obtained scale factors are applied to the MC samples after a medium tau lepton candidate is required. The scale factors are applied event-wise, depending on the number of fake tau lepton candidates in MC. To identify fake tau lepton candidates a matching between the reconstructed tau lepton candidates and the generated tau leptons is performed, as already described in section 5.6. Afterwards the medium selection cuts and the hard selection cuts are applied. The event yields per 0.5 fb^{-1} are shown in figure 5.20 for both selections. No significant deviations of single data points from the average yield are observed. The selections therefore are stable over



(a) (b)

Figure 5.20: Number of events per 0.5 fb^{-1} which passed the (a) medium selection and (b) the hard selection as function of integrated luminosity.

the whole data-taking period.
The impact of each selection step on the signal samples for a leptoquark mass of 300 GeV/c^2 and 600 GeV/c^2 are shown in the tables 5.2 and 5.3. The efficiency of each selection cut with respect to the previous step and the total selection efficiency after each cut are given. Furthermore, the value $S/\sqrt{B}$, which is a measure of the sensitivity of the selection, is given after each selection step. In this formula S is the number of signal events and B the number of background events. For the leptoquarks with a mass of 600 GeV/c^2 the value for $S/\sqrt{B}$ increases with each cut of the medium selection and is rather stable for each cut of the hard selection. For the lower leptoquark mass the value increases with each cut of the medium selection, whereas the value of $S/\sqrt{B}$ decreases for each cut of the hard selection. This is expected because the hard selection is optimized for higher leptoquark masses, while the medium selection is optimized for the lower leptoquark mass regions.
The efficiencies of the cuts are in general quite high (over 80%). Only the cuts of the pre-selection and the requirement of a medium tau lepton candidate have lower efficiencies. Taking into account the branching ratio of $\mathrm{LQ}_3\overline{\mathrm{LQ}}_3 \rightarrow \mu + \tau_{\mathrm{had+X}}$ the efficiency of these cuts is 50-60%. However, this is acceptable, because the cuts of the pre-selection reduce the Standard Model background significantly. Especially, the background from light quark and gluon production is reduced by a huge amount due to the requirement of a tight muon candidate. The requirement of a medium tau lepton candidate has an efficiency of 50%, but it is needed since the distribution of the p_T of the leading tau lepton candidate has a very high sensitivity for different leptoquark masses and shows a good separation between background and signal events as described in section 5.4. After the medium selection 2% of the events with leptoquarks with a mass of 300 GeV/c^2 are left, with an excellent $S/\sqrt{B}$ of 15.21. After the hard selection 3% of the events with leptoquarks of a mass of 600 GeV/c^2 are left, while the value for $S/\sqrt{B}$ is 0.876.

Selection	Cut	$S/\sqrt{B}$	Efficiency of the cut	Selection efficiency
pre-selection		6.952	0.167	0.167
medium selection	one medium tau lepton candidate	11.559	0.239	0.040
	third jet with $p_T > 30\text{GeV/c}^2$	15.141	0.864	0.035
	$\not{E}_T > 50$ GeV/c^2	14.603	0.737	0.025
	leading jet with $p_T > 100$ GeV/c^2	14.592	0.777	0.020
	$H_T > 400$ GeV/c^2	15.208	0.987	0.020
hard selection	$H_T > 700$ GeV/c^2	13.706	0.410	0.008
	leading jet with $p_T > 150$ GeV/c^2	12.050	0.86	0.007
	third jet with $p_T > 50\text{GeV/c}^2$	12.164	0.913	0.006
	$\not{E}_T > 100$ GeV/c^2	9.202	0.596	0.004

Table 5.2: Cut-flow table for the signal sample for a leptoquark with a mass of 300 GeV/c^2, given are the value $S/\sqrt{B}$, where S is the number of signal events and B the number of background events, the efficiency of each selection cut with respect to the previous step and the total selection efficiency after each cut.

Selection	Cut	$S/\sqrt{B}$	Efficiency of the cut	Selection efficiency
pre-selection		0.136	0.259	0.259
medium selection	one medium tau candidate lepton	0.213	0.231	0.060
	third jet with $p_T > 30\text{GeV/c}^2$	0.289	0.894	0.053
	$\not{E}_T > 50$ GeV/c^2	0.351	0.923	0.050
	leading jet with $p_T > 100$ GeV/c^2	0.431	0.964	0.047
	$H_T > 400$ GeV/c^2	0.456	1.0	0.047
hard selection	$H_T > 700$ GeV/c^2	0.942	0.944	0.045
	leading jet with $p_T > 150$ GeV/c^2	0.866	0.894	0.040
	third jet with $p_T > 50\text{GeV/c}^2$	0.842	0.879	0.035
	$\not{E}_T > 100$ GeV/c^2	0.876	0.813	0.029

Table 5.3: Cut-flow table for the signal sample for a leptoquark with a mass of 600 GeV/c^2, given are the value $S/\sqrt{B}$, where S is the number of signal events and B the number of background events, the efficiency of each selection cut with respect to the previous step and the total selection efficiency after each cut.

In figures 5.21 to 5.24 control distributions after the medium selection and the hard selection are presented. Shown are the distributions for the number of jets, the number of b-jets, H_T, $\not{E}_T$, the p_T of the leading muon candidate in the event, the product of the charges of all tau lepton candidates and muon candidates in the event and the p_T of the leading tau lepton candidate in the event $\Pi_\tau(q_\tau) \cdot \Pi_\mu(q_\mu)$. In addition, p_T spectra for the jets with the highest, second and third highest p_T in the event are presented. In all distributions data and MC agree well. After applying the selections $t\bar{t}$ is the main background.
The event with the high values of H_T and $\not{E}_T$ and the high jet p_T has been studied in the event display. Nothing conspicuous could be seen.

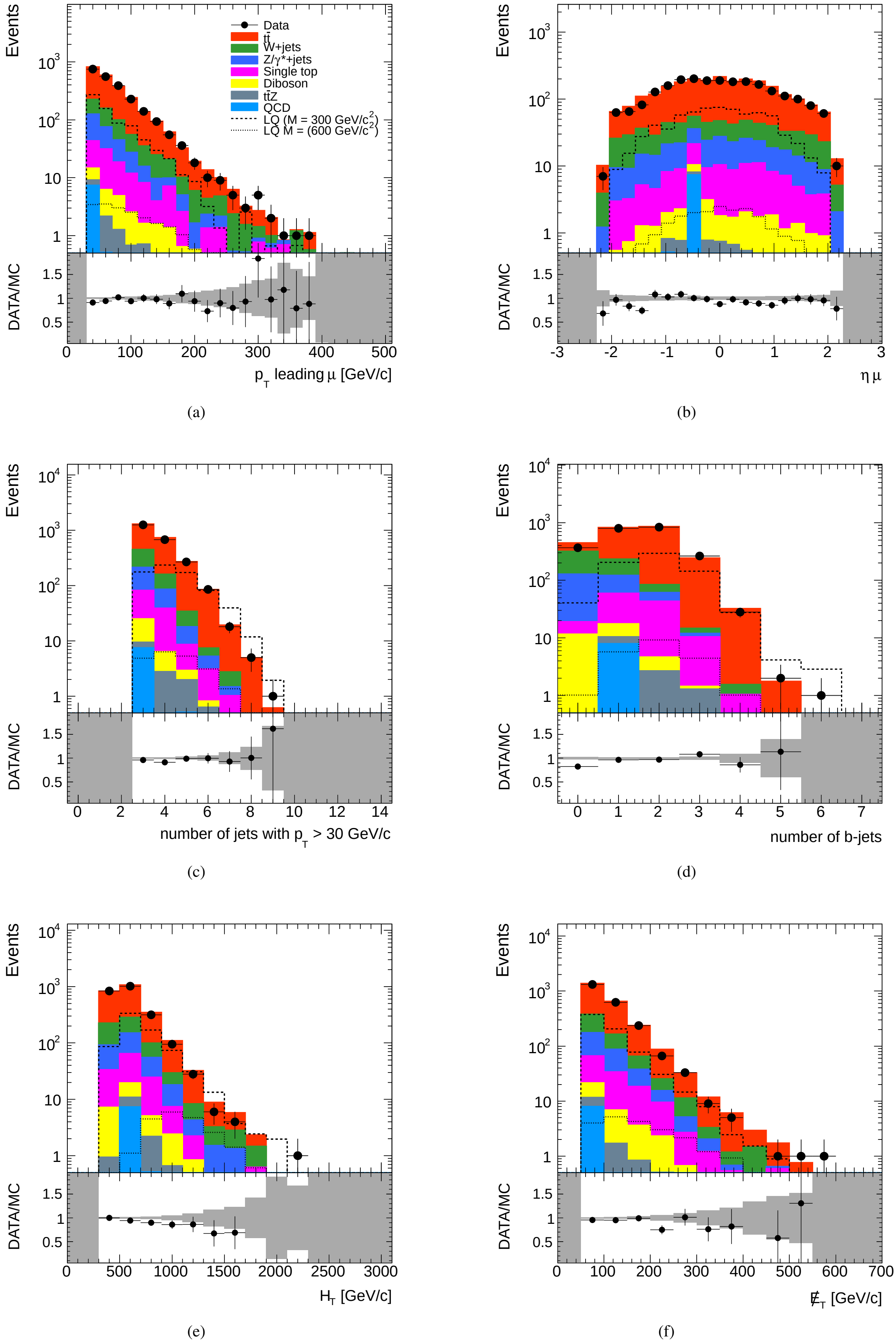
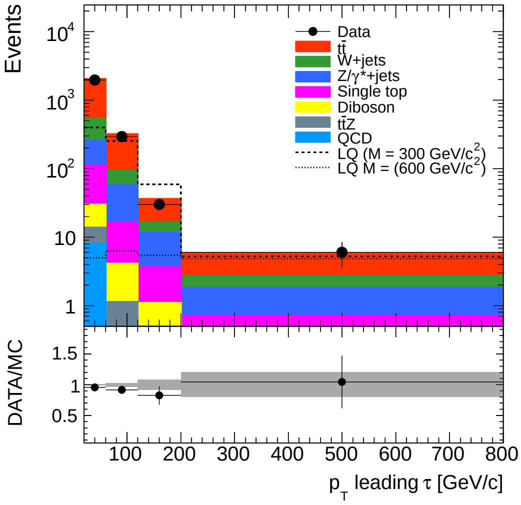
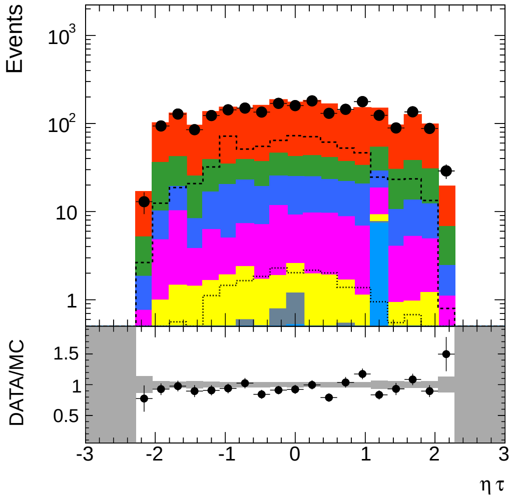


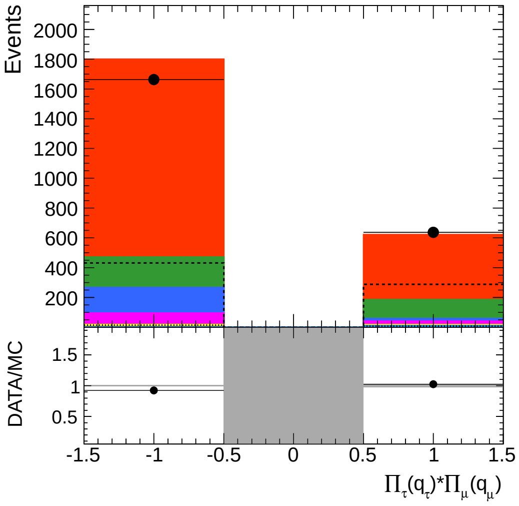
Figure 5.21: Control distributions after the medium selection. Shown are the distributions for (a) the p_T of the leading muon candidate in the event (b) η of the leading muon candidate in the event, (c) the number of jets, (d) the number of b-jets, where the loose working point of the CSV-algorithm was used, (e) H_T and (f) $\not{E}_T$.



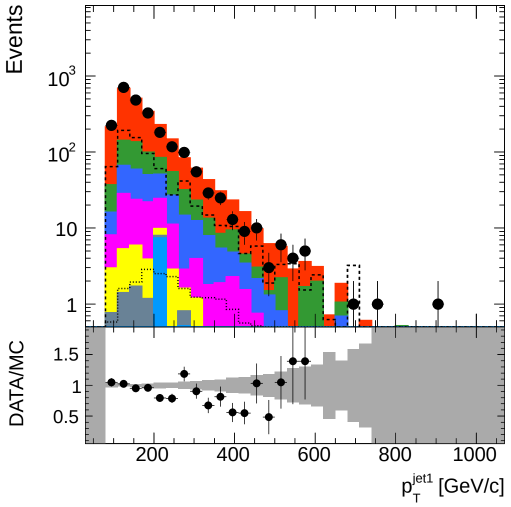
(a)



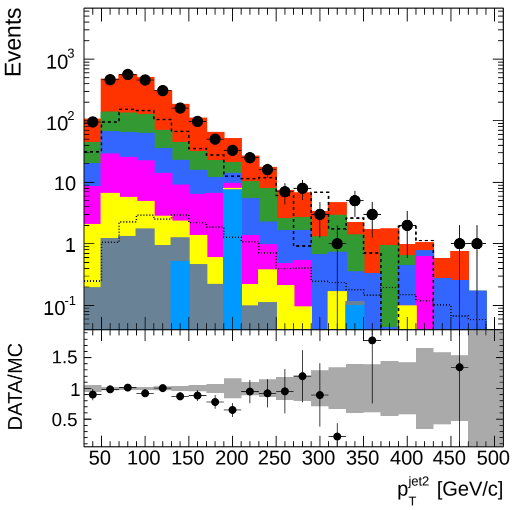
(b)



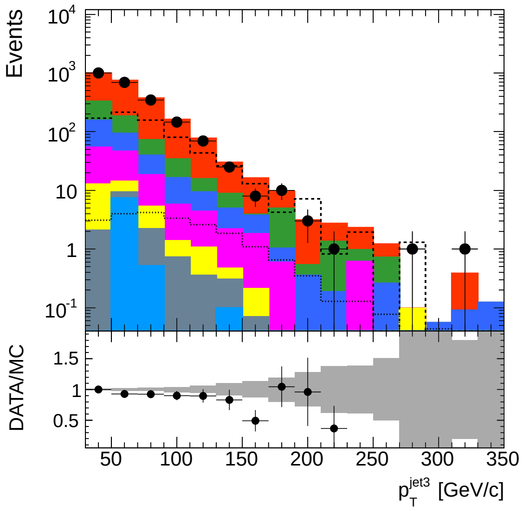
(c)



(d)



(e)



(f)

Figure 5.22: Control distributions after the medium selection. Shown are the distributions for (a) p_T of the leading tau lepton candidate, (b) η of the leading tau lepton candidate, (c) the product of the charges of all tau lepton candidates and muon candidates, (d) the p_T of the leading jet, (e) the p_T of the jet with the second highest transverse momentum and (f) the p_T of the jet with the third highest transverse momentum.

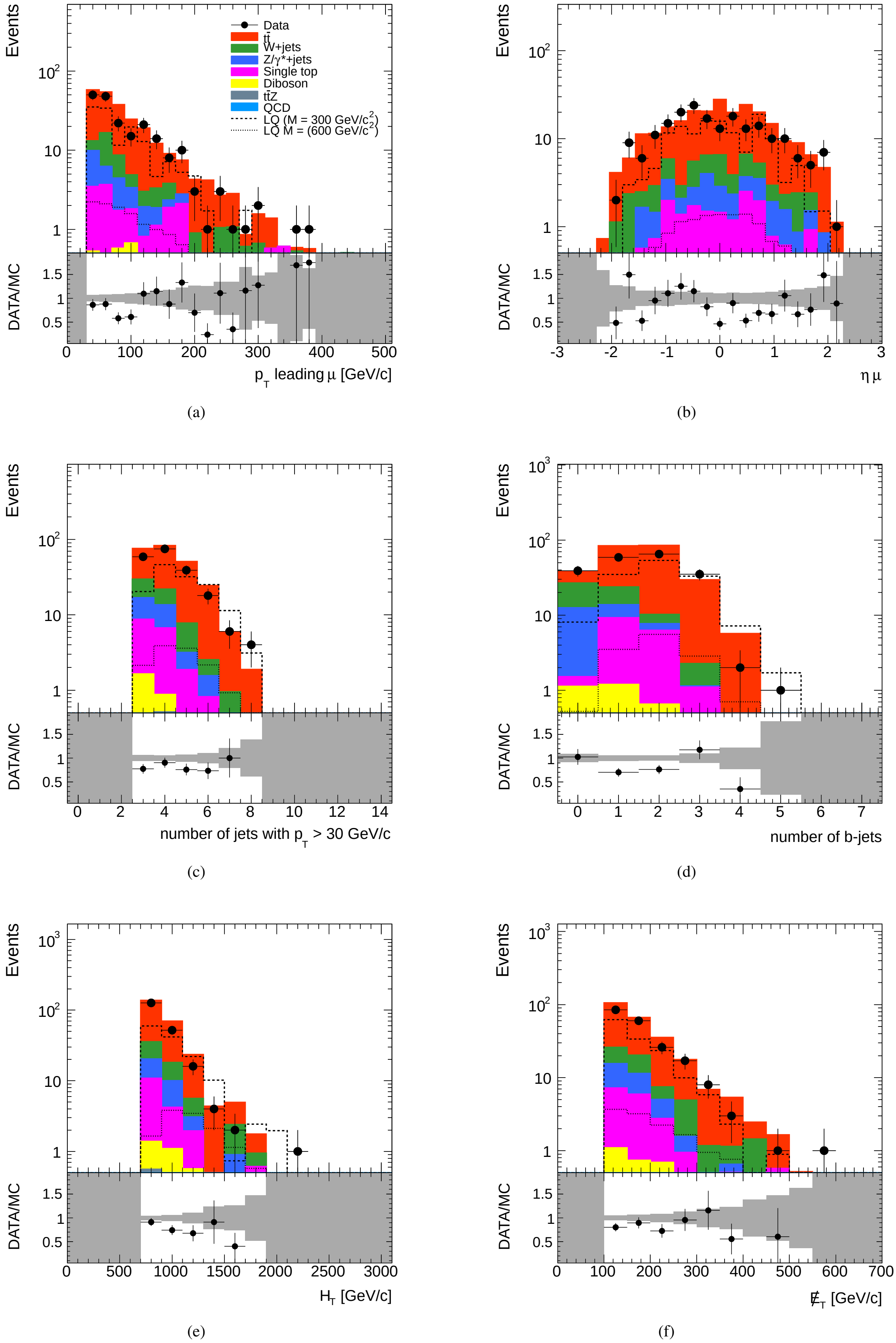


Figure 5.23: Control distributions after the hard selection. Shown are the distributions for (a) the p_T of the leading muon candidate in the event (b) η of the leading muon candidate in the event, (c) the number of jets, (d) the number of b-jets, where the loose working point of the CSV-algorithm was used, (e) H_T and (f) $\not{E}_T$.

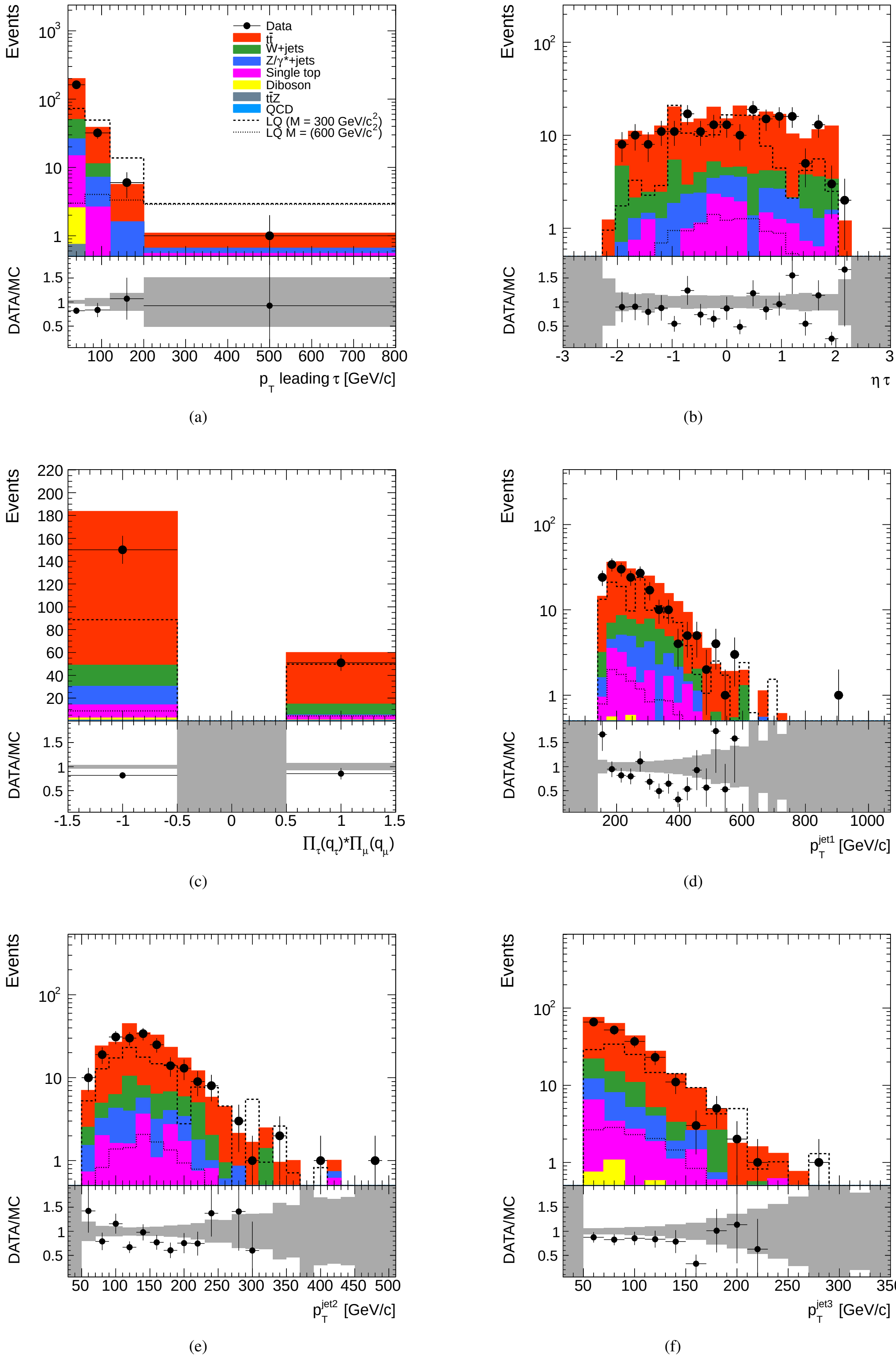


Figure 5.24: Control distributions after the hard selection. Shown are the distributions for (a) p_T of the leading tau lepton candidate, (b) η of the leading tau lepton candidate, (c) the product of the charges of all tau lepton candidates and muon candidates, (d) the p_T of the leading jet, (e) the p_T of the jet with the second highest transverse momentum and (f) the p_T of the jet with the third highest transverse momentum.

5.8 Systematic uncertainties

Systematic uncertainties on the p_T distribution of the leading tau lepton candidate have to be estimated and considered in the final statistical interpretation of the result. The systematic uncertainties considered are listed in the following.
Uncertainties which change only the rate of the p_T^τ distribution are:

- uncertainty on the **lumiosity** of 4.4% [53],
- conservative estimates are used for the uncertainties on the **cross section** for Standard Model processes (estimates based on [54, 55] and [56])
  - $\mathrm{t\bar{t}}$-production: 15%,
  - single top-production: 50%,
  - W+jets-production: 50%,
  - Z+jets-production: 50%,
  - diboson-production: 100%,
  - $\mathrm{t\bar{t}Z}$-production: 100%,
  - QCD-production: 100%.

Uncertainties which vary the rate and the shape of the p_T^τ distribution are taken into account by varying the corresponding quantity by one standard deviation up and down. The amount by which the number of events changes in a certain bin gives the result in the corresponding uncertainty.
The following uncertainties vary the rate and the shape of the p_T distribution of the leading tau lepton candidate:

- **Tau lepton fake rate:** On the determined tau lepton fake rate an uncertainty of $\pm\sigma$ is assigned. See section 5.6 for a detailed description of the determination of this uncertainty.
- **Tau identification:** An uncertainty of 6% is applied for each real tau lepton in the event in order to account for uncertainties in the efficiency of the tau lepton identification.
- **Tau lepton energy**: The uncertainty on the tau lepton energy is taken into account by varying the energy of all tau leptons by 3%.
- **Muon identification:** The uncertainties on the muon identification, isolation and on the trigger are determined by CMS as a function of p_T and η of the muon candidate [46]. These values are added quadratically to the statistical uncertainty of the muon scale factors. Additionally a normalization uncertainty of 0.5% and an uncertainty of 0.2% on the tag and probe method which was used to determine the efficiencies is taken into account.
- **PU-reweighting:** In order to take uncertainties due to the pile-up reweighting into account the minimum bias cross section of 69.3 ± 4.1mb, on which the PU-reweighting is based, is varied by one standard deviation.

- **Jet energy resolution (JER):** The η-dependent factors which are used to adjust the jet resolution in MC to the one in data [47, 48] are varied by one standard deviation to obtain the uncertainty on the p_T distribution of the leading tau lepton candidate. This variation is also transferred to the calculation of the missing transverse momentum.
- **Jet energy scale (JEC):** The jet energy scale is varied by one standard deviation as a function of η and p_T [47, 48], which corresponds to a variation of 1-3%. This variation is also transferred to the calculation of the missing transverse momentum.
- **Scale uncertainty on the $\mathrm{t\bar{t}}$-production:** The factorization and renormalization scale are varied by a factor 2 and by a factor 0.5 in order to determine the scale uncertainties on the $\mathrm{t\bar{t}}$-production.

In figure 5.25 and 5.27 the relative uncertainties on the $\mathrm{t\bar{t}}$ sample of the p_T of the leading tau lepton candidate are shown after applying the medium and the hard selection, respectively. In figure 5.26 and figure 5.28 the same is shown for the signal sample of a leptoquark with a mass of 500 $\mathrm{GeV/c^2}$. High uncertainties for the $\mathrm{t\bar{t}}$ sample are the uncertainties on the tau lepton fake rate. At high transverse momenta the uncertainty is almost 20%. In this p_T-range almost all tau leptons are fakes. Thus, the uncertainty on the fake rate determination is transferred directly to the uncertainty in this bin. For the leptoquark sample the uncertainty on the tau lepton fake rate is very small, because there are only a few fake tau leptons in this sample. Since in these samples most of the tau lepton candidates are real tau leptons, the uncertainties on the tau lepton identification and on the tau lepton energy are the highest uncertainties.

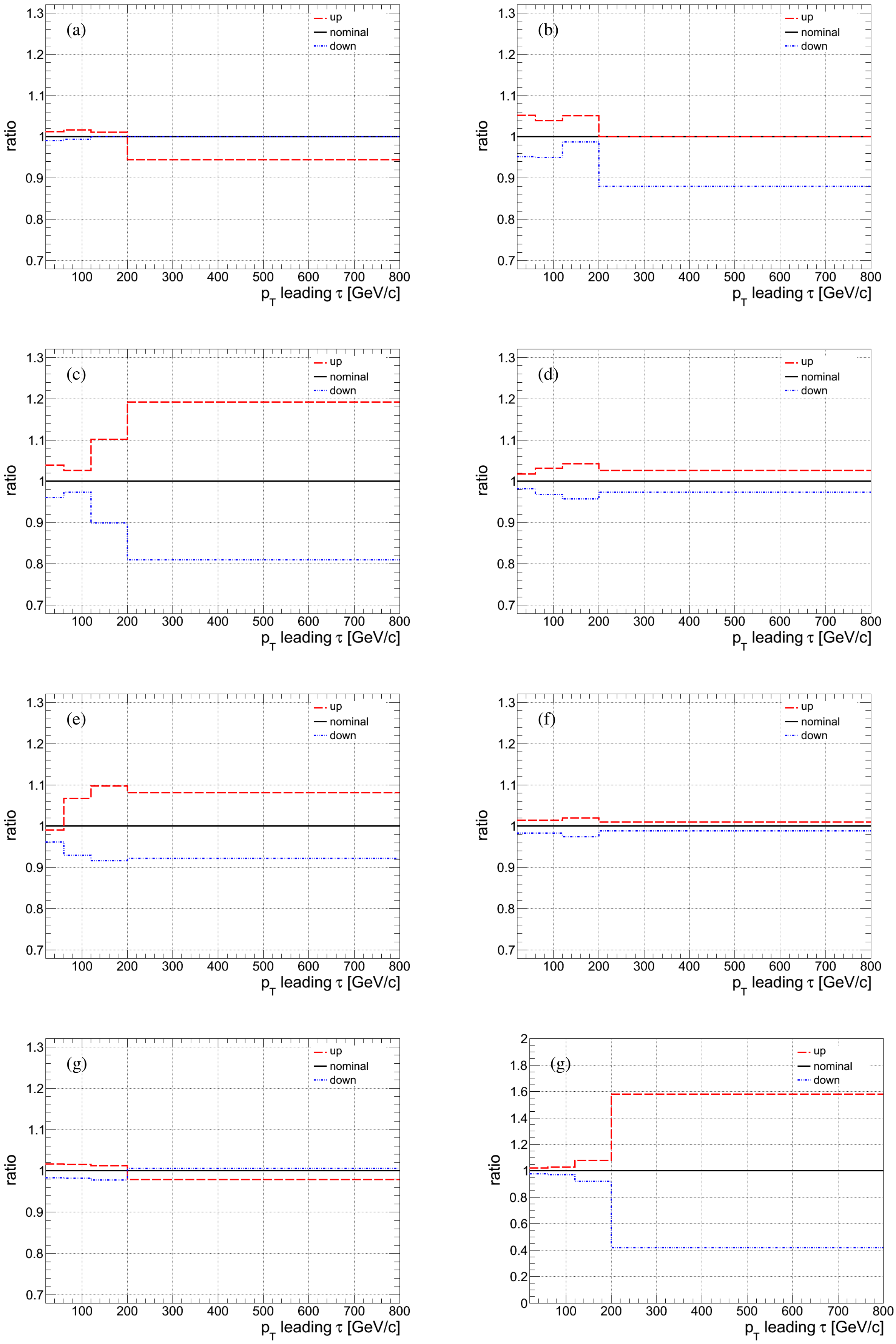


Figure 5.25: Relative uncertainties as a function of the p_T of the leading tau lepton candidate for the $t\bar{t}$ sample for (a) JER, (b) JEC, (c) tau lepton fake rate, (d) tau lepton identification, (e) tau lepton energy, (f) muon identification, (g) pile-up and (h) the scale uncertainty after applying the medium selection cuts.

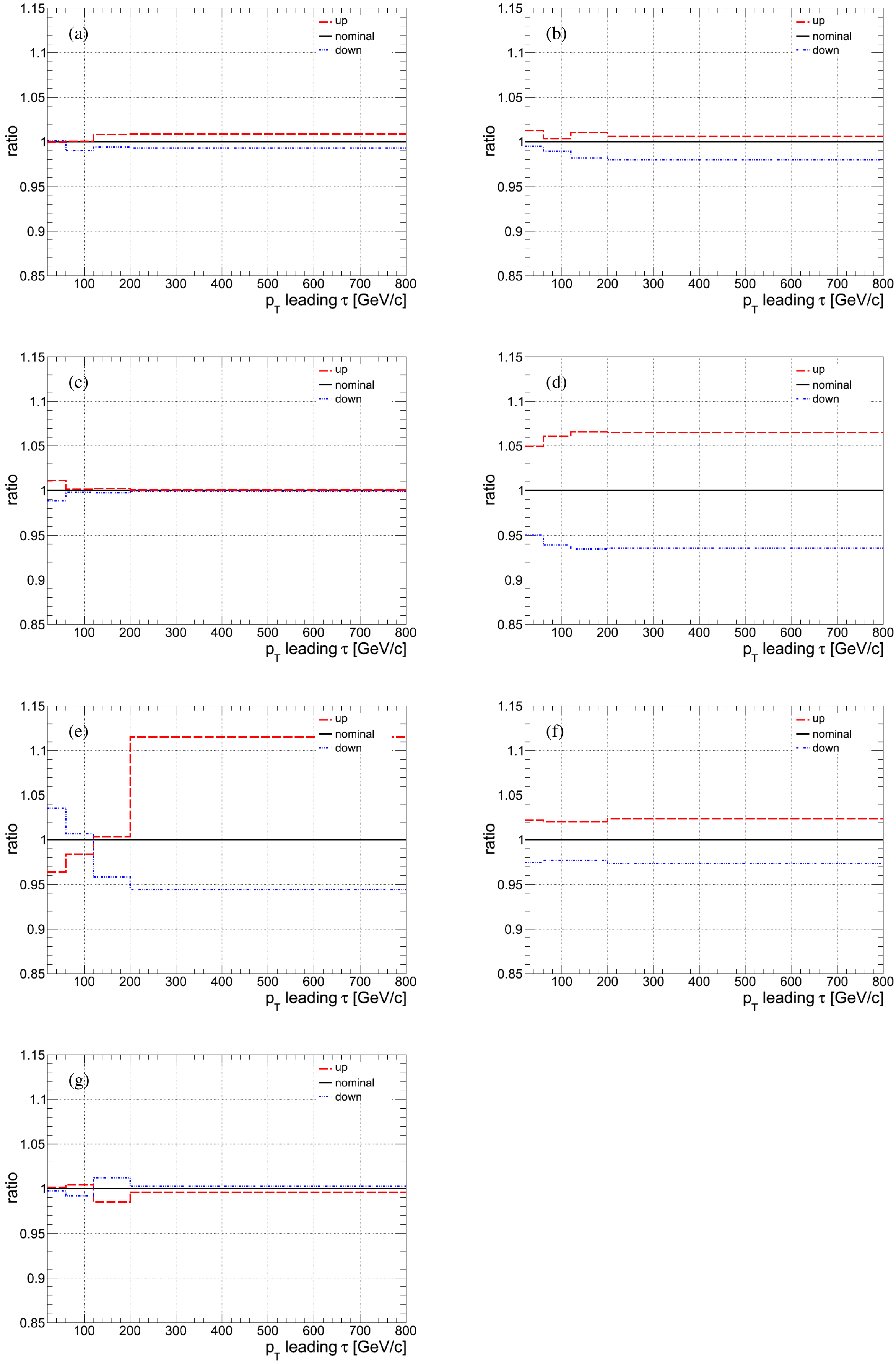


Figure 5.26: Relative uncertainties as a function of the p_T of the leading tau lepton candidate for the signal sample for a leptoquark with a mass of 500 GeV for (a) JER, (b) JEC, (c) tau lepton fake rate, (d) tau lepton identification, (e) tau lepton energy, (f) muon identification and (g) pile-up after applying the medium selection cuts.

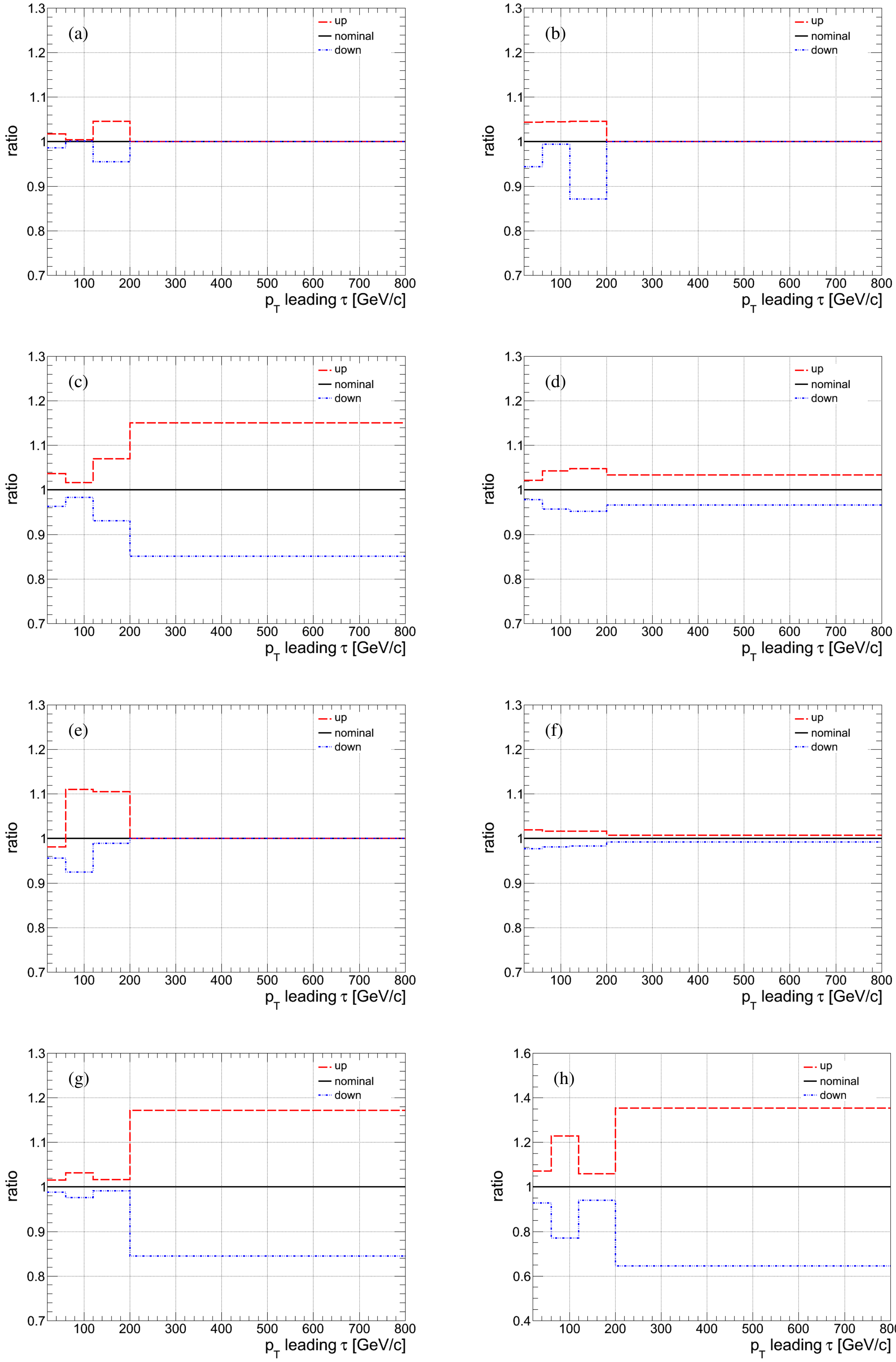


Figure 5.27: Relative uncertainties as a function of the p_T of the leading tau lepton candidate for the $t\bar{t}$ sample for (a) JER, (b) JEC, (c) tau lepton fake rate (d) tau lepton identification, (e) tau lepton energy, (f) muon identification (g) pile-up and (h) the scale uncertainty after applying the hard selection cuts.

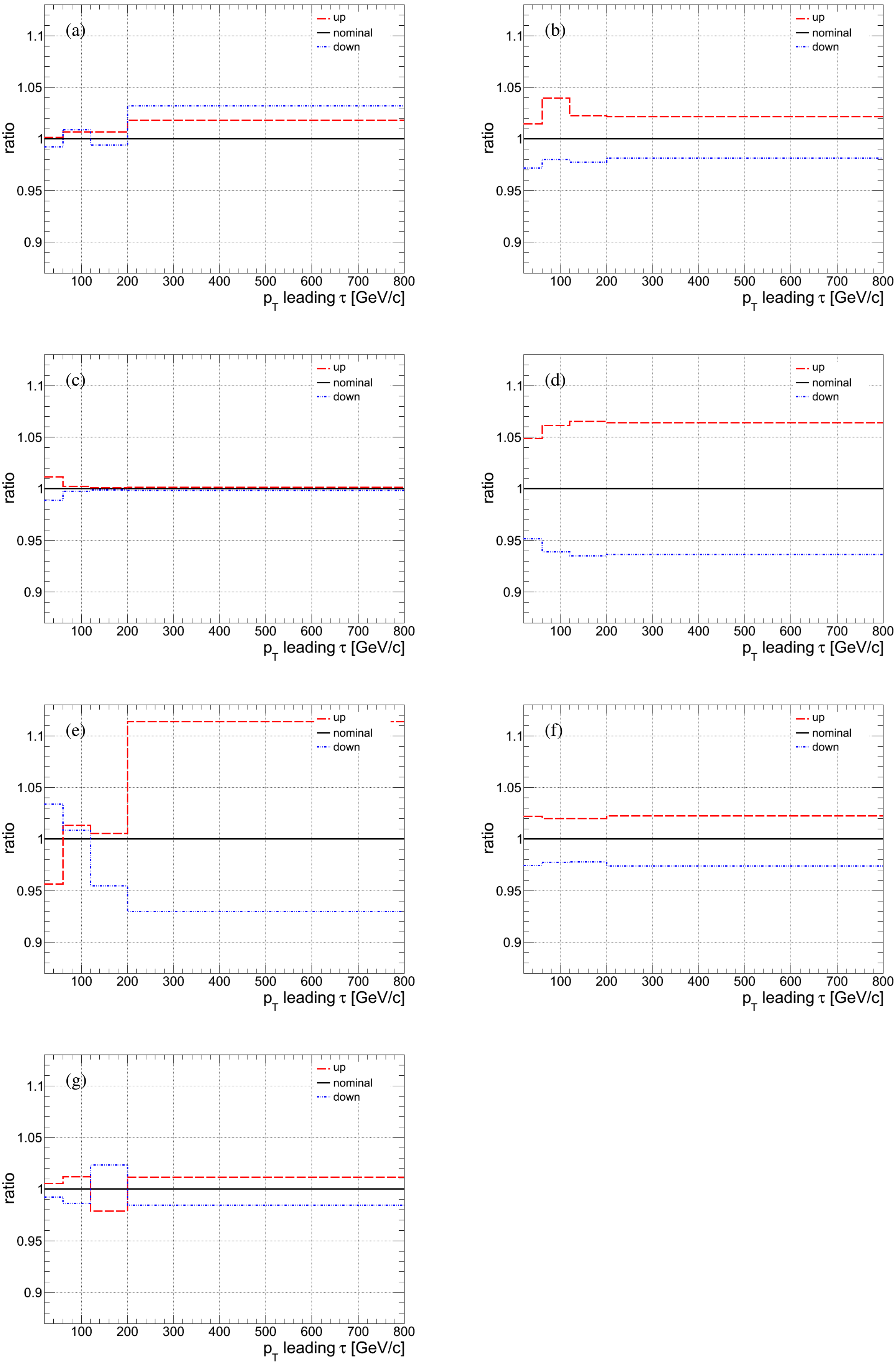


Figure 5.28: Relative uncertainties as a function of the p_T of the leading tau lepton candidate for the signal sample for a leptoquark with a mass of 500 GeV for (a) JER, (b) JEC, (c) tau lepton fake rate, (d) tau lepton identification, (e) tau lepton energy, (f) muon identification and (g) pile-up after applying the hard selection cuts.

Sample	Number of events after the medium selection	Number of events after the hard selection
LQ ($M = 200$ GeV/c^2)	659.8	62.0
LQ ($M = 250$ GeV/c^2)	1014.0	190.3
LQ ($M = 300$ GeV/c^2)	719.2	138.6
LQ ($M = 350$ GeV/c^2)	410.7	105.7
LQ ($M = 400$ GeV/c^2)	232.5	78.7
LQ ($M = 450$ GeV/c^2)	125.1	54.5
LQ ($M = 500$ GeV/c^2)	67.6	33.4
LQ ($M = 550$ GeV/c^2)	39.3	22.4
LQ ($M = 600$ GeV/c^2)	21.6	13.2
LQ ($M = 650$ GeV/c^2)	12.4	8.4
LQ ($M = 700$ GeV/c^2)	7.1	4.8
LQ ($M = 750$ GeV/c^2)	4.1	2.9
LQ ($M = 800$ GeV/c^2)	2.6	1.9
LQ ($M = 850$ GeV/c^2)	1.6	1.2
LQ ($M = 900$ GeV/c^2)	0.9	0.7
LQ ($M = 950$ GeV/c^2)	0.6	0.5
LQ ($M = 1$ TeV/c^2)	0.4	0.3
$t\bar{t}$	1765.0	180.0
W + jets	331.2	27.9
Z + jets	190.9	17.1
Single top	97.9	15.0
Diboson	20.4	2.2
$t\bar{t}$Z	7.1	1.0
QCD	8.0	0
Total background	$2421\ +464\ -411$	$243\ +47\ -39$
Data	2300	201

Table 5.4: Summary of the number of events after the medium selection and the hard selection have been applied.

5.9 Results

After the determination of the systematic uncertainties the final statistical interpretation of the results can be performed. This is done based on the distributions of the p_T of the leading tau lepton candidate after the medium selection and the hard selections have been applied. In table 5.4 the total number of events is given after the selection cuts have been performed. After the medium selection 2300 data events are observed. The Standard Model expectation is with 2421^{+464}_{-411} events a little bit higher but the observed number of events is well within the errors. The same applies for the hard selection, where 201 data events are observed compared to the Standard Model expectation of 243^{+47}_{-39} events.
The distributions of the p_T of the leading tau lepton candidate are shown in figure 5.29 and 5.30. The grey shaded areas in the distributions represent the total uncertainties, where the statistical uncertainties of the MC samples and the determined systematic uncertainties have been added in quadrature. In the ratio plots the dark grey area corresponds to the statistical error and the light grey area illustrates the total error. In both distributions the data is described by the Standard Model prediction within the errors over the full range of p_T^{τ}. In conclusion, no excess above the Standard Model expectation is observed.

In the following, a statistical interpretation of the results is performed where limits on the cross section times branching ratio are calculated. For the limit calculation the theta package [50] is used. As input to theta the histograms in figure 5.29 and 5.30 are used and a shape analysis is

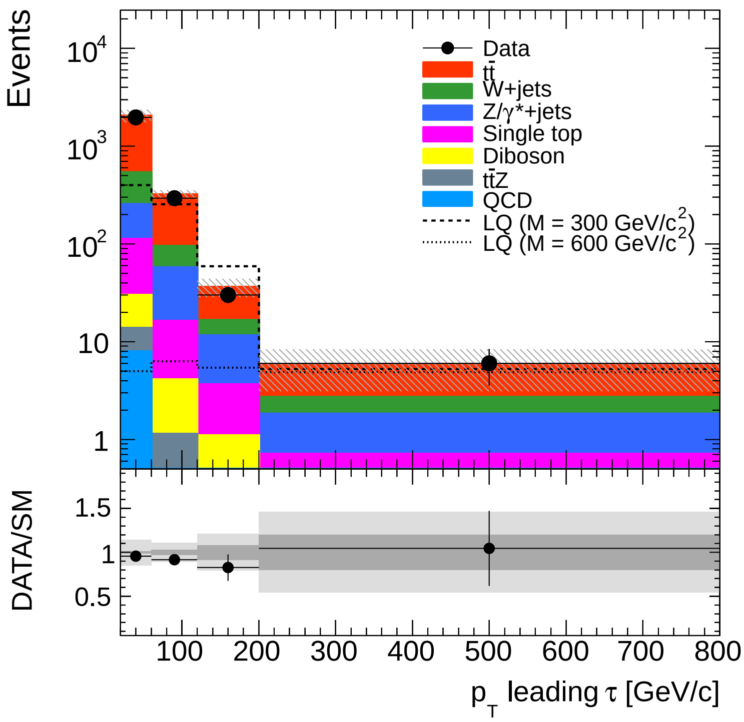


Figure 5.29: Distributions for the p_T of the leading tau lepton candidate after applying the medium selection. The grey shaded areas in the distributions represent the total errors. In the ratio plot the dark grey area corresponds to the statistical error while the light grey area illustrates the total error.

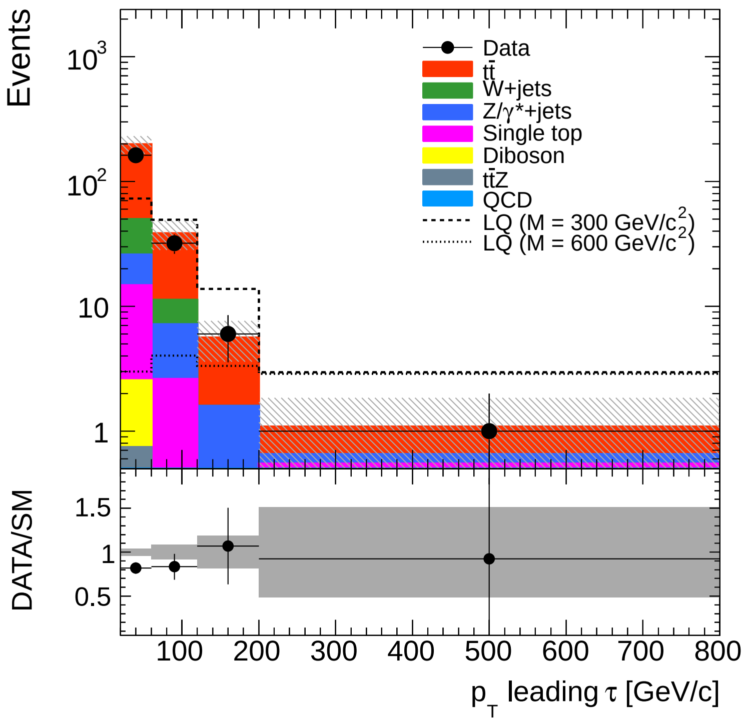
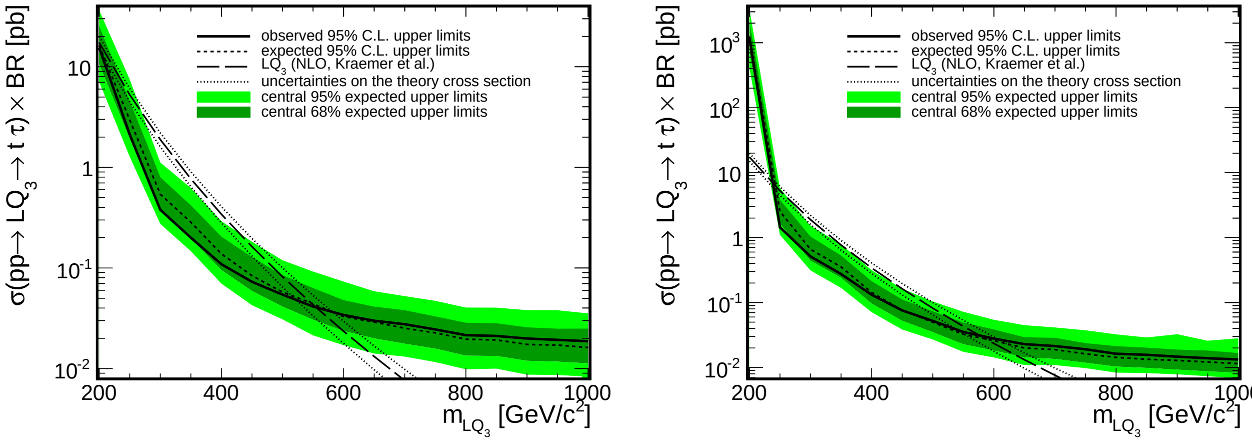


Figure 5.30: Distributions for the p_T of the leading tau lepton candidate after applying the hard selection. The grey shaded areas in the distributions represent the total errors. In the ratio plot the dark grey area corresponds to the statistical error while the light grey area illustrates the total error.



(a) (b)

Figure 5.31: Limit on the cross section times branching ratio based on the distribution of the p_T of the leading tau lepton candidate after (a) the medium selection and (b) the hard selection. The theory curve displays the next-to-leading order cross section of the studied process [16]. The uncertainties on the theory production cross section are determined by adding the scale uncertainties and the pdf uncertainties in quadrature [57].

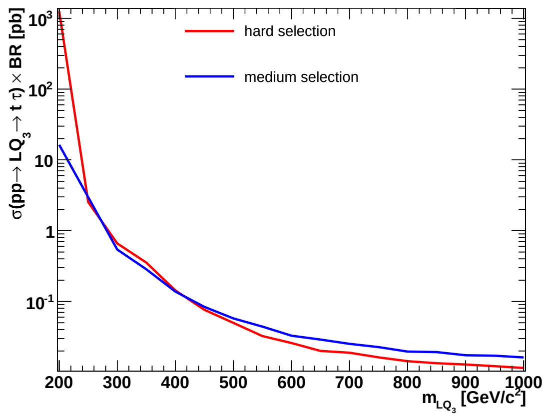


Figure 5.32: Comparison of the expected limits after applying the medium selection and the hard selection.

performed. All systematic uncertainties described in 5.8 and the statistical uncertainties of the MC samples are taken into account as nuisance parameters in the limit calculation. The limit is set on the production cross section times branching ratio for third-generation leptoquarks decaying to a top quark and a tau lepton.

The results of the limit calculation are shown in figure 5.31 for the medium and the hard selections. Shown are the observed and expected limits on the cross section times branching ratio for the process $LQ_3 \rightarrow t + \tau$. The green bands represent the uncertainties corresponding to one and two standard deviations. The theory curve displays the next-to-leading order cross section of the studied process [4] [16]. The observed limit lies well inside the uncertainty bands over the whole mass range. No indication for new physics can be found. It can be observed that the medium selection is more sensitive for lower leptoquark masses while the hard selection is

[4] For the uncertainties on the theory prediction the scale uncertainties and the pdf uncertainties are considered. The scale uncertainties are determined by varying the factorization and renormalization scale by a factor 2 and by a factor 0.5. The pdf uncertainties are obtained by calculating the theory cross sections for each uncertainty eigenvector from the pdf and by adding the differences to the theory cross section in quadrature [57]. In order to get the total uncertainty on the theory cross section the scale uncertainties and the pdf uncertainties are added in quadrature.

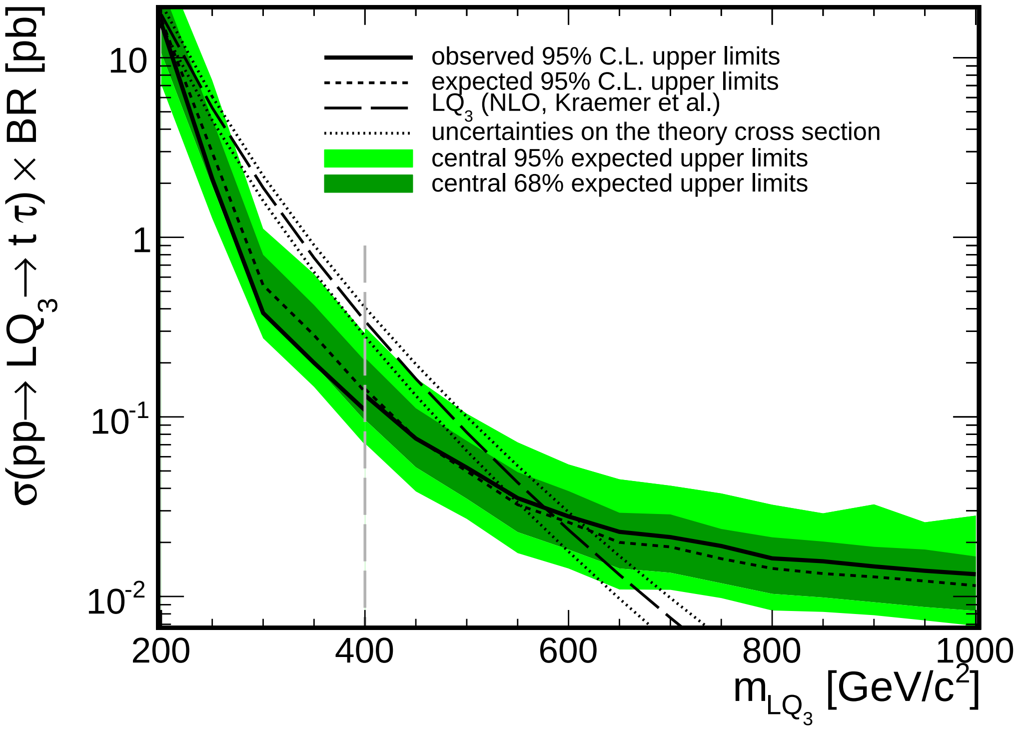


Figure 5.33: Combination of the limits. The medium selection is used for leptoquarks masses lower than 400 GeV, the hard selection is used for leptoquarks with masses higher than 400 GeV. The transition region between the two selections is indicated with a gray vertical line. The theory curve displays the next-to-leading order cross section of the studied process [16]. The uncertainties on the theory production cross section are determined by adding the scale uncertainties and the pdf uncertainties in quadrature [57].

more sensitive to higher leptoquark masses.

In order to find out which selection is best for which mass points, the expected limits are directly compared. The result is shown in figure 5.32. The two expected limits meet at a leptoquark mass of 400 GeV/c^2. In conclusion, the medium selection is used to quote limits for leptoquarks with masses up to 400 GeV/c^2, the hard selection is used for leptoquarks with higher masses.

The combination of both limits is shown in figure 5.33. The transition region between the two selections is indicated with a grey vertical line.

Production cross sections for third-generation leptoquarks decaying to a top quark and a tau lepton above approximately 10 pb can be excluded for leptoquark masses of about 200 GeV/c^2. For masses up to 1 TeV cross sections above up to 0.014 pb can be excluded. By comparing the observed limit with the theory cross sections the limits on the production cross section can be converted into an upper limit on the mass of the leptoquarks. Third-generation leptoquarks decaying to a top quark and a tau lepton can be excluded up to masses of 582 GeV/c^2 at 95% C.L. All these results are given under the assumption of $\mathrm{BR}(\mathrm{LQ}_3 \rightarrow \mathrm{t} + \tau) = 1$.

5.10 Possible improvements of the analysis

The following aspects of this analysis can still be improved:

- In order to not rely on simulated events and the associated uncertainties as well as to get higher statistics, the background could be estimated from data.
- One could try to optimize the event selection further. For example, multivariate-techniques could be used to get a better separation power between the signal and the MC processes.
- The tau lepton identification at high p_T could be improved and the efficiency of the tau lepton reconstruction could be calculated in data and MC.
- The samples after the selections could be split in two orthogonal samples. This would lead to more degrees of freedom in the limit calculation and could lead to a better limit if a good partition of the samples can be found. An idea is to separate the events into events where the tau lepton candidate and the muon candidate have charges with opposite signs and in events where at least one muon candidate and/or at least one tau lepton candidate can be found with same-sign-charges. This would have the advantage that the background estimation is easier in the same-sign sample since here almost all tau lepton candidates are fakes and the irreducible background is negligible [58]. By estimating this background from data the limit could be improved further.
- More leptoquark models and more decay channels could be included into the analysis. The statistics could be doubled by requiring at least one electron candidate instead of one muon candidate and then combining the results.
- In the appendix A, a study is shown based on an event selection with top tagging. The application of top-tagging in the search might be an interesting option for the design centre-of mass energy of the LHC of 14 TeV. In the performed analysis at a centre-of-mass energy of 8 TeV the top tagger used is unfortunately not efficient enough in the p_T-range of the top quarks from the leptoquark decays, such that not enough events are left after the selection to calculate a limit. However, with more data, a higher centre-of-mass energy and a top tagger which is more efficient at lower transverse momenta, like for example the HEP-Top tagger [59, 60], the application of top-tagging might be an interesting option for a search for leptoquarks. This could also help to extend this search into an all-hadronic + τ channel. In combination with the presented search this could improve the result further.

Chapter 6

Conclusion

In this thesis a search was presented for pair production of third-generation scalar leptoquarks decaying into a top quark and a tau lepton using proton-proton collision data recorded by the CMS experiment at the LHC in the year 2012. The search for leptoquarks is motivated by the symmetry between quarks and leptons in the Standard Model. Many theories beyond the Standard Model like GUTs, Compositeness and Technicolor predict the existence of leptoquarks. Searches for leptoquarks have been performed in many channels and at many colliders. A search for third-generation leptoquarks decaying into a top quark and a tau lepton has never been performed so far. Therefore, and because of the very rich final state, different selections were studied in order to find an optimized selection for the search. All studied selections are based on one hadronically decaying tau lepton, one muon, three jets and high H_T and $\not{E}_T$. A looser selection which is more sensitive to lower leptoquark masses and a selection with harder cuts more sensitive for higher leptoquark masses were chosen based on the best expected exclusion limits. The final interpretation of the result is based on the distribution of the transverse momentum of the leading tau lepton, which shows a high sensitivity for different leptoquark masses. Since tau leptons play an important role in this study, the fake rates for tau leptons were determined in data and MC. They were found to be at the order of a few per cent, but different in Data and MC. Therefore scale factors were derived depending on p_T of the leading tau lepton. These were applied on the fake tau leptons in the simulated samples in order to eliminate the impact of the different fake rates in data and MC simulations.

After applying the chosen selection no excess over the Standard Model expectation was observed. In a statistical interpretation of the results limits on the cross section times branching ratio were set. Third-generation leptoquarks decaying to a top quark and a tau lepton could be excluded up to masses of 582 $\mathrm{GeV/c^2}$ at 95% C.L. Since this was the first search in this channel, this is the highest available limit up to date.

Appendix A

Search for leptoquarks with top tagging

An alternative approach to the presented analysis strategy is to search for third-generation leptoquarks decaying to a top quark and a tau lepton by requiring a hadronically decaying top quark, with the usage of top tagging in the event selection. A study on this is shortly introduced in this chapter. The idea of top tagging is explained in section A.1. The clustering of jets with the Cambridge/Aachen (CA) algorithm, which are the input for the top tagging algorithm, is explained in section A.2. The Johns Hopkins Top Tagger, which is the top tagging algorithm used in this study, is introduced in section A.3. Afterwards the studied selection with top tagging is presented in section A.4.

A.1 Top tagging

The current limits on leptoquarks (see section 2.3.3) indicate that leptoquarks will be very heavy if they exist. This means that the top quarks into which they decay have a very high transverse momentum. The decay products of the top quarks are therefore highly boosted and very often clustered in one jet. These jets have a very characteristic substructure, which can be used to distinguish top-jets from the light quark and gluon background. The top-jet contains three subjets where one is a b-jet and the other two have the invariant mass of a W-boson. This information can be used to distinguish top-jets from jets from light quark and gluon production which in general do not have this particular substructure.
There are different top tagging algorithms available, a detailed overview can be found for example in [61]. In this study the so called CMS Top Tagger, which is based on the Johns Hopkins Top Tagger, is used [62]. It uses C/A jets as input and reconstructs subjets by reversing the clustering sequence.

A.2 The Cambridge/Aachen jet clustering algorithm

The Cambridge/Aachen algorithm [35] is a sequential recombination algorithm. It is collinear and infrared safe. Like the anti-k_t-algorithm, described in 4.3, particle-flow objects are used as input for the jet clustering. For each pair of particles i and j the distance

$$d_{\mathrm{ij}} = \frac{\Delta R_{\mathrm{ij}}^2}{R^2} \tag{A.1}$$

is calculated, where ΔR_{ij}^2 is given by $\Delta R_{\mathrm{ij}}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$. The parameter R is the chosen cone size of the jet. If the minimum of the distance between particle i and j is smaller than one, the corresponding particles are combined to a new particle and the algorithm starts from the beginning. This procedure is repeated until all particles have a distance $\Delta R_{\mathrm{ij}} > R$. Each of the final particles is called a jet.

There is no weight according to the momentum in the distance measure, as for example in the anti-k_t-algorithm and the order of the clustering steps depends only on the angular distance between the particles. By reversing the clustering history hard subjets can be found with less ambiguity than in other jet algorithms. Therefore, C/A jets are well suited as an input for top tagging algorithms.

A.3 The Johns Hopkins Top Tagger

The Johns Hopkins Top Tagger [62] is based on C/A jets with a cone size of $R = 0.8$. The CMS Top Tagger reverses the clustering of the C/A jets. Two decomposition steps are performed. In the first step the jet clustering is reversed until the algorithm finds two clusters which carry a significant amount of the momentum of the initial jet and which are well separated. If such clusters cannot be found, the initial jet is considered to have only one subjet, namely the initial jet, and the algorithm stops. If the algorithm finds two clusters, it tries to decluster both of them further. In this way a maximum of four subjets can be found.

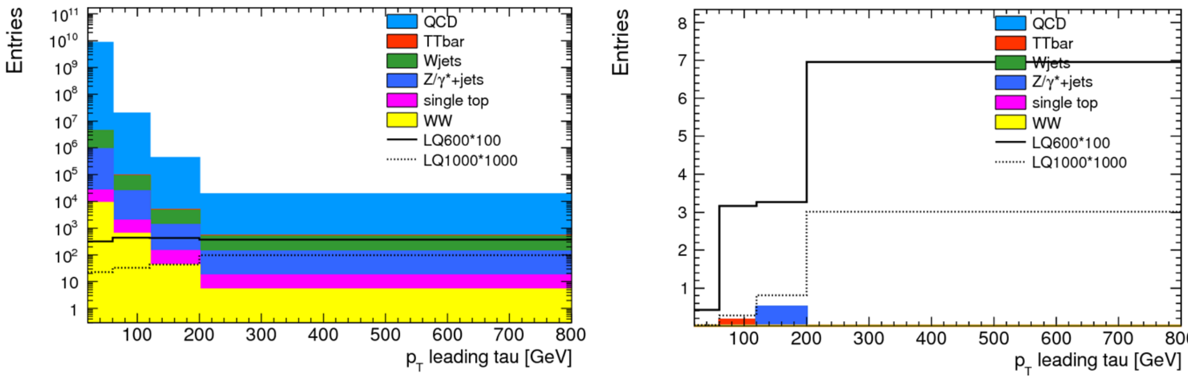
In order to test the separation of the clusters often a p_T-dependent criterion is used, which is in the case of the CMS tagger

$$\sqrt{(\eta_1 - \eta_2)^2 + (\phi_1 - \phi_2)^2} > 0.4 - 0.0004 \times p_T^C. \tag{A.2}$$

In this equation p_T^C is the transverse momentum of the jet from which the corresponding clustering step started. This ensures that the distance is smaller for initial jets with higher momentum, which is motivated by the fact that jets with higher transverse momentum are more collimated. If only one of the found clusters satisfies the momentum criterion, only this one is considered in the following step of the algorithm. The other cluster is ignored.

With the so found subjets the mass of the initial jet (m_{jet}) is calculated by adding the four-vectors of the individual subjets. Moreover from the three subjets with the highest transverse momentum the invariant mass of each pair is calculated if at least three subjets were found. From these three masses the minimum m_{min} is considered in the following.

In order to be top tagged the initial jet has to fulfill the following criteria.



(a)

(b)

Figure A.1: Distribution for the p_T of the leading tau lepton candidate before and after the selection described in A.4. The signal samples are scaled with a factor 100 for leptoquarks with a mass of 300 GeV/c^2 and with a factor 1000 for a leptoquark with a mass of 1 TeV.

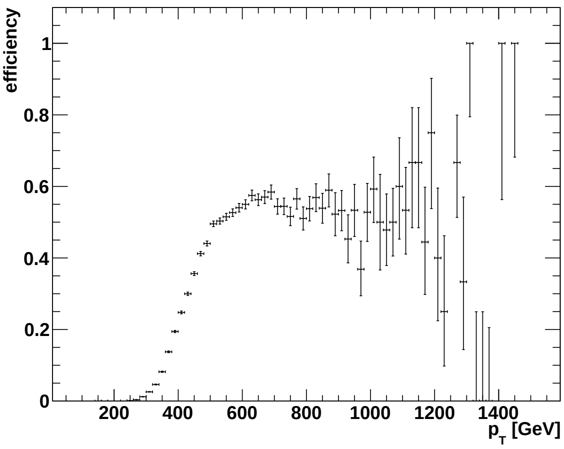
- The total jet mass has to fulfill 140 GeV/c^2 $<$ m$_{\text{jet}}$ $<$ 250 GeV/c^2, corresponding to the top quark mass.
- The variable m_{min} has to be higher than 50 GeV/c^2, so that two subjets have a mass near the mass of the W-boson.
- At least three subjets have been found.

A.4 Selection with top tagging

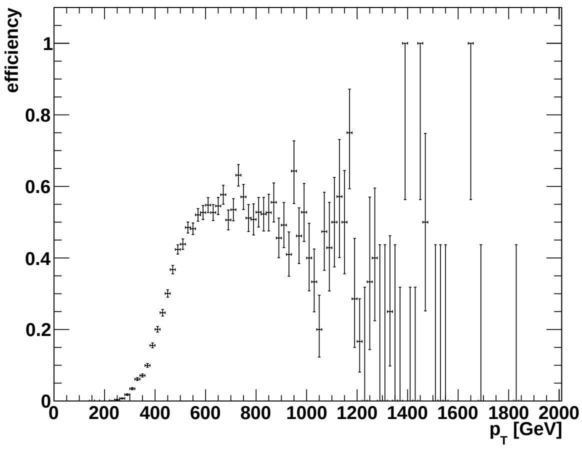
Instead of requiring a lepton in the event selection, a top tag and therefore a hadronically decaying top quark, can be required in order to select the leptoquark candidate events. Different selections have been tried. The one which gave the best result for the signal over background ratio is introduced here. The studies have been performed on the first 4.4 fb^{-1} of the data set recorded in the year 2012 by the CMS experiment.
An event cleaning on the primary vertices, jets and tau leptons is performed as described in section 5.3. The only difference is that tau leptons with $|\eta| < 2.3$ instead of $|\eta| < 2.1$ are used. In order to pass the selection the events have to have

- at least one good primary vertex,
- at least one C/A jet with $p_T > 400$ GeV/c^2, which has been tagged by the Johns Hopkins Top Tagger,
- at least two hadronically decaying tau lepton candidates with $p_T > 40$ GeV/c^2.

The distributions for the p_T of the leading tau lepton candidate in the event before and after the selection are shown in figure A.1. Before the selection the main background is the production of light quark and gluon jets. The value for $S/\sqrt{B}$, where $S \approx 15.5$ is the number of signal events and $B \approx 8.6 \cdot 10^9$ is the number of background events, is with $1.7 \cdot 10^{-4}$ very small. After the selection only a few background events are left. These are mainly events from $t\bar{t}$-production and



(a)



(b)

Figure A.2: Efficiencies of the Johns Hopkins Top Tagger for (a) the $t\bar{t}$ sample and (b) the signal sample for a lepoquark with a mass of 600 GeV/c^2

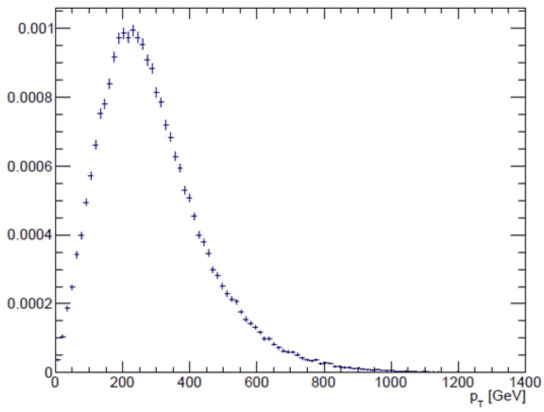
DY events. The value for $S/\sqrt{B}$ is with 0.17 much higher than before the selection. However, only a few signal events ($S \approx 0.1$, $B \approx 0.7$) are left after the selection. Thus too less signal events are left to be able to find leptoquarks or to exclude them.

The reason for this is the low top tagging efficiency of the Johns Hopkins Top Tagger for the top quarks originating from the leptoquark decays. The top tagging efficiencies have been determined in MC. In order to determine the top tagging efficiency a matching between generated top quarks and the reconstructed C/A jets is done. If a generated top quark decaying to hadrons is found in a distance of $\Delta R < 0.8$ to a C/A jet, it is checked if this jet was tagged by the top tagger. The number of tagged jets divided by the number of jets which could be matched to a generated top quark gives the efficiency of the top tagger. The efficiencies are shown as a function of the p_T of the jets in figure A.2. They have been determined in the $t\bar{t}$ sample and in the signal sample for leptoquarks with a mass of 600 GeV/c^2. The top tagger reaches the highest efficiency (nearly 60%) for top quarks with a p_T higher than 600 GeV/c. For very high transverse momenta the efficiencies decrease again, because at such high momenta the decay products are so close to each other that not all three subjets can be reconstructed anymore.

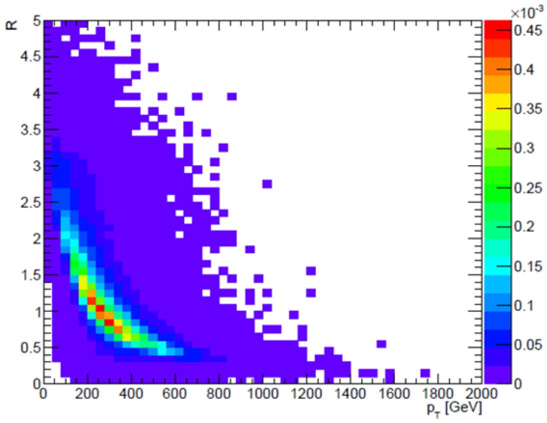
The p_T distribution for top quarks originating from the decay of letpoquarks with a mass of 600 GeV/c^2 is shown in figure A.3(a). Most of the top quarks have a transverse momentum around 300 GeV/c^2 and only in a fraction of events top quarks have transverse momenta of 600 GeV/c^2 and more.

In figure A.3(b) the distance between the decay products of the top quarks are shown as a function of the p_T of the top quark. In order to be clustered into a C/A jet the b-quark and the W-boson must have a distance smaller than $\Delta R = 0.8$. Only very few top quarks with a transverse momentum of less than 300 GeV/c fulfill this requirement. So most of the top quarks in the signal sample for a leptoquark with a mass of 600 GeV/c^2 are not tagged by the John Hopkins Top Tagger.

In conclusion, the studied top tagger is too inefficient in the p_T region of top quarks from lep-



(a)



(b)

Figure A.3: (a) P_T distribution for the top quarks from the decay of a leptoquark with a mass of 600 GeV$/c^2$, (b) Distance $R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ between the decay products of the top quarks from a leptoquark decay with a mass of 600 GeV$/c^2$ as a function of the p_T of the top quark.

toquark decays with masses between 300 and 600 GeV$/c^2$ since the decay products of the top quark are not collimated enough to be clustered in a C/A jet with a cone size of $\Delta R = 0.8$. Thus too few events are selected such so that a statistical interpretation of the results after applying a selection with top tagging is not possible. However, with data at higher centre-of-mass energies or with a top tagger which is more efficient at low transverse momenta, the application of top tagging might be an interesting option for the search for third-generation leptoquarks.

List of Figures

List of Tables

Bibliography

[1] B. R. Martin and G. Shaw, "Particle physics". Wiley. John & Sons, second edition, 2005.

[2] Particle Data Group Collaboration, "Review of Particle Physics (RPP)", *Phys.Rev.* **D86** (2012) 010001. doi:10.1103/PhysRevD.86.010001.

[3] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys.Lett.* **B716** (2012) 1–29, arXiv:1207.7214. doi:10.1016/j.physletb.2012.08.020.

[4] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys.Lett.* **B716** (2012) 30–61, arXiv:1207.7235. doi:10.1016/j.physletb.2012.08.021.

[5] S. P. Martin, "A Supersymmetry primer", arXiv:hep-ph/9709356.

[6] B. Schrempp and F. Schrempp, "LIGHT LEPTOQUARKS", *Phys.Lett.* **B153** (1985) 101. doi:10.1016/0370-2693(85)91450-9.

[7] E. Farhi and L. Susskind, "Technicolor", *Phys.Rept.* **74** (1981) 277. doi:10.1016/0370-1573(81)90173-3.

[8] E. Eichten and K. D. Lane, "Dynamical Breaking of Weak Interaction Symmetries", *Phys.Lett.* **B90** (1980) 125–130. doi:10.1016/0370-2693(80)90065-9.

[9] S. Dimopoulos and L. Susskind, "Mass Without Scalars", *Nucl.Phys.* **B155** (1979) 237–252. doi:10.1016/0550-3213(79)90364-X.

[10] S. Dimopoulos, "Technicolored Signatures", *Nucl.Phys.* **B168** (1980) 69–92. doi:10.1016/0550-3213(80)90277-1.

[11] B. Gripaios, A. Papaefstathiou, K. Sakurai et al., "Searching for third-generation composite leptoquarks at the LHC", *JHEP* **1101** (2011) 156, arXiv:1010.3962. doi:10.1007/JHEP01(2011)156.

[12] S. Rolli and M. Tanabashi, "Leptoquarks". http://pdg.lbl.gov/2013/reviews/rpp2012-rev-leptoquark-quantum-numbers.pdf.

[13] J. C. Pati and A. Salam, "Lepton Number as the Fourth Color", *Phys.Rev.* **D10** (1974) 275–289. doi:10.1103/PhysRevD.10.275, 10.1103/PhysRevD.11.703.2.

[14] H. Georgi and S. Glashow, “Unity of All Elementary Particle Forces”, *Phys.Rev.Lett.* **32** (1974) 438–441. doi:10.1103/PhysRevLett.32.438.

[15] S. Chakdar, T. Li, S. Nandi et al., “Unity of elementary particles and forces for the third family”, *Phys.Lett.* **B718** (2012) 121–124, arXiv:1206.0409. doi:10.1016/j.physletb.2012.10.021.

[16] M. Kramer, T. Plehn, M. Spira et al., “Pair production of scalar leptoquarks at the CERN LHC”, *Phys.Rev.* **D71** (2005) 057503, arXiv:hep-ph/0411038. doi:10.1103/PhysRevD.71.057503.

[17] S. Davidson, D. C. Bailey, and B. A. Campbell, “Model independent constraints on leptoquarks from rare processes”, *Z.Phys.* **C61** (1994) 613–644, arXiv:hep-ph/9309310. doi:10.1007/BF01552629.

[18] M. Leurer, “A Comprehensive study of leptoquark bounds”, *Phys.Rev.* **D49** (1994) 333–342, arXiv:hep-ph/9309266. doi:10.1103/PhysRevD.49.333.

[19] W. Buchmüller and D. Wyler, “Constraints on SU(5) Type Leptoquarks”, *Phys.Lett.* **B177** (1986) 377. doi:10.1016/0370-2693(86)90771-9.

[20] CMS Collaboration, “Search for pair production of first- and second-generation scalar leptoquarks in pp collisions at $\sqrt{s} = 7$ TeV”, *Phys.Rev.* **D86** (2012) 052013, arXiv:1207.5406. doi:10.1103/PhysRevD.86.052013.

[21] ATLAS Collaboration, “Search for third generation scalar leptoquarks in pp collisions at sqrt(s) = 7 TeV with the ATLAS detector”, arXiv:1303.0526.

[22] ATLAS Collaboration, “Search for first generation scalar leptoquarks in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector”, *Phys.Lett.* **B709** (2012) 158–176, arXiv:1112.4828. doi:10.1016/j.physletb.2012.03.023, 10.1016/j.physletb.2012.02.004.

[23] ATLAS Collaboration, “Search for second generation scalar leptoquarks in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector”, *Eur.Phys.J.* **C72** (2012) 2151, arXiv:1203.3172. doi:10.1140/epjc/s10052-012-2151-6.

[24] CMS Collaboration, “Search for pair production of third-generation leptoquarks and top squarks in pp collisions at $\sqrt{s} = 7$ TeV”, *Phys. Rev. Lett.* **110** (2013) 081801, arXiv:1210.5629.

[25] CMS Collaboration, “Search for third-generation leptoquarks and scalar bottom quarks in pp collisions at $\sqrt{s} = 7$ TeV”, *JHEP* **1212** (2012) 055, arXiv:1210.5627. doi:10.1007/JHEP12(2012)055.

[26] CDF Collaboration, “Search for second and third generation leptoquarks including production via technicolor interactions in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV”, *Phys.Rev.Lett.* **85** (2000) 2056–2061, arXiv:hep-ex/0004003. doi:10.1103/PhysRevLett.85.2056.

[27] https://confluence.slac.stanford.edu/pages/viewpage.action?pageId=38897376.

[28] "CMS Physics: Technical Design Report Volume 1: Detector Performance and Software". Technical Design Report CMS. CERN, Geneva, 2006.

[29] https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults#2012_Proton_Proton_Collisions.

[30] CMS Collaboration, "CMS Tracking Performance Results from early LHC Operation", *Eur.Phys.J.* **C70** (2010) 1165–1192, arXiv:1007.1988. doi:10.1140/epjc/s10052-010-1491-3.

[31] CMS Collaboration, "Electron and photon energy calibration and resolution with the CMS ECAL at $\sqrt{s} = 7$ TeV", *CMS-PAS-EGM-11-001* (2011).

[32] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET", *CMS-PAS-PFT-09-001* (2009).

[33] CMS Collaboration, "Performance of muon reconstruction and identification in pp collisions at sqrt(s) = 7 TeV", *CMS-PAS-MUO-10-004* (2011).

[34] https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMuonId.

[35] G. P. Salam, "Towards Jetography", *Eur.Phys.J.* **C67** (2010) 637–686, arXiv:0906.1833. doi:10.1140/epjc/s10052-010-1314-6.

[36] CMS Collaboration, "Plans for Jet Energy Corrections at CMS", *CMS-PAS-JME-07-002* (2008).

[37] C. Weiser, "A combined secondary vertex based B-tagging algorithm in CMS", *CERN-CMS-NOTE-2006-014* (2006).

[38] CMS Collaboration, "Study of tau reconstruction algorithms using pp collisions data collected at sqrt(s) = 7 TeV", *CMS-AN-PFT-10-004* (2010).

[39] CMS Collaboration, "Tau identification in CMS", *CMS-PAS-TAU-11-001* (2011).

[40] https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuidePFTauID.

[41] https://twiki.cern.ch/twiki/bin/view/CMS/StandardModelCrossSectionsat8TeV.

[42] https://twiki.cern.ch/twiki/bin/viewauth/CMS/LumiCalc.

[43] https://twiki.cern.ch/twiki/bin/viewauth/CMS/PdmV2012Analysis.

[44] E. N. D. Martino, "Single Trigger Muon Efficiencies of Run2012D". https://indico.cern.ch/getFile.py/access?contribId=2&resId=0&materialId=slides&confId=228197, talk at internal CMS meeting, January 2013.

[45] I. J. Eduardo Navarro De Martino, "Muon Efficiencies 2012 in 53X". `https://indico.cern.ch/getFile.py/access?contribId=3&resId=0&materialId=slides&confId=214870`, talk at internal CMS meeting, 2012.

[46] CMS Muon POG, "Reference muon id and isolation efficiencies". `https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonReferenceEffs`, 2012.

[47] CMS Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS", *JINST* **6** (2011) P11002, `arXiv:1107.4277`. `doi:10.1088/1748-0221/6/11/P11002`.

[48] `https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetResolution`.

[49] `https://twiki.cern.ch/twiki/bin/viewauth/CMS/PileupMCReweightingUtilities`.

[50] `www.theta-framework.org`.

[51] J. Ott and T. Muller, "Search for Resonant Top Quark Pair Production in the Muon+Jets Channel with the CMS Detector", Dec, 2012. Presented 14 Dec 2012.

[52] E. Bender, "Effizienz des Tau-Lepton-Nachweises beim CMS-Detektor für Leptoquarkproduktion", Universität Hamburg, 2012.

[53] "CMS Luminosity Based on Pixel Cluster Counting - Summer 2012 Update", technical report, CERN, Geneva, (2012).

[54] CMS Collaboration, "Measurement of the $\mathrm{t\bar{t}}$ Production Cross Section in pp Collisions at 7 TeV in Lepton + Jets Events Using b-quark Jet Identification", *Phys.Rev.* **D84** (2011) 092004, `arXiv:1108.3773`. `doi:10.1103/PhysRevD.84.092004`.

[55] CMS Collaboration, "Measurement of the Top-antitop Production Cross Section in *pp* Collisions at $\sqrt{s} = 7$ TeV using the Kinematic Properties of Events with Leptons and Jets", *Eur.Phys.J.* **C71** (2011) 1721, `arXiv:1106.0902`. `doi:10.1140/epjc/s10052-011-1721-3`.

[56] CMS Collaboration, "Search for $\mathrm{t\bar{t}}$ resonances in semileptonic final states in *pp* collisions at $\sqrt{s} = 8$ TeV", *CMS-PAS-B2G-12-006* (2013).

[57] S. Alekhin, S. Alioli, R. D. Ball et al., "The PDF4LHC Working Group Interim Report", `arXiv:1101.0536`.

[58] H. Saka, "Status Update on $\mathrm{LQ_3\overline{LQ}_3} \rightarrow \mathrm{t\bar{t}}\tau^+\tau^-$". `https://indico.cern.ch/conferenceDisplay.py?confId=222392`, talk at internal CMS meeting, December 2012.

[59] T. Plehn, M. Spannowsky, M. Takeuchi et al., "Stop Reconstruction with Tagged Tops", *JHEP* **1010** (2010) 078, `arXiv:1006.2833`. `doi:10.1007/JHEP10(2010)078`.

[60] T. Plehn, M. Spannowsky, and M. Takeuchi, “How to Improve Top Tagging”, *Phys.Rev.* **D85** (2012) 034029, arXiv:1111.5034. doi:10.1103/PhysRevD.85.034029.

[61] T. Plehn and M. Spannowsky, “Top Tagging”, *J.Phys.* **G39** (2012) 083001, arXiv:1112.4441. doi:10.1088/0954-3899/39/8/083001.

[62] CMS Collaboration, “Study of tau reconstruction algorithms using pp collisions data collected at sqrt(s) = 7 TeV”, *CMS-AN-PFT-10-004* (2010).

Hiermit bestätige ich, dass die vorliegende Arbeit von mir selbstständig verfasst wurde und ich keine anderen als die angegebenen Hilfsmittel - insbesondere keine im Quellenverzeichnis nicht benannten Internet-Quellen - benutzt habe und die Arbeit von mir vorher nicht einem anderen Prüfungsverfahren eingereicht wurde. Die eingereichte schriftliche Fassung entspricht der auf dem elektronischen Speichermedium. Ich bin damit einverstanden, dass die Masterarbeit veröffentlich wird.

Ort, Datum Unterschrift

Danksagung

Ich möchte mich an dieser Stelle bei allen bedanken, die diese Arbeit durch ihre Unterstützung ermöglicht haben.

Bei **Prof. Dr. Johannes Haller** möchte ich mich bedanken für die Möglichkeit die Arbeit in seiner Gruppe durchführen zu können. Außerdem möchte ich mich für die Betreuung der Arbeit und die vielen hilfreichen Hinweise bedanken.

Ein besonderer Dank gilt **Dr. Roman Kogler** für die Betreuung meiner Arbeit, die vielen sehr hilfreichen und interessanten Erklärungen und Diskussionen über die Ergebnisse und die Geduld beim Suchen von Programmierfehlern.

Bei **Prof. Dr. Peter Schleper** möchte ich mich für die Übernahme des Zweitgutachtens bedanken.

Des Weiteren möchte ich mich bei **Dr. Thomas Peiffer** bedanken für die Hilfe bei Problemen mit "SFrame" und für die Beantwortung vieler Fragen.

Bei **Kristin Heine, Teresa Lenz und Rebekka Höing** möchte ich mich für die nette Atmosphäre im Büro, die vielen Diskussionen und für die lustigen Abende außerhalb des Büros bedanken.

Außerdem möchte ich mich bei **meiner Familie und Nils** dafür bedanken, dass sie mich in Allem unterstützen und immer für mich da sind.