MUSiC –
Model Unspecific Search in CMS with τ Leptons

von
Michael Brodski

Masterarbeit im Fach Physik

vorgelegt der

Fakultät für Mathematik, Informatik und Naturwissenschaften
der Rheinisch-Westfälischen Technischen Hochschule

im
September 2012

angefertigt am

III. Physikalischen Institut (IIIA)

bei

Prof. Dr. Thomas Hebbeker

Zweitgutachter
Prof. Dr. Christopher Wiebusch

Zusammenfassung

Der **L**arge **H**adron **C**ollider ist die größte je gebaute Kollisionsmaschine und ist seit dem Jahr 2007 in Betrieb. Bei Protonenkollisionen am LHC überprüfen Physiker das Standardmodell der Physik und suchen nach neuen Phänomenen – der Physik jenseits des Standardmodells.

Der hierzu in dieser Arbeit verwendete Ansatz MUSiC – eine modellunabhängige Suche – untersucht eine Vielzahl von Endzuständen, gemessen mit dem CMS-Detektor des LHC, systematisch nach Abweichungen von der Standardmodellvorhersage. Im Rahmen dieser Arbeit wird ein neues Objekt – das τ-Lepton – in die MUSiC-Umgebung eingeführt. Anschließend wird eine Untersuchung des kompletten Datensatzes des CMS-Detektors aus dem Jahr 2011 durchgeführt. Zur Überprüfung der Sensitivität der MUSiC-Umgebung wird ferner eine Studie mit einem simulierten W$'$-Signal vorgenommen. Dabei ist das W$'$ ein schweres Eichboson und ist eine hypothetische Erweiterung des Standardmodells. Die Ergebnisse der Untersuchung des kompletten Datensatzes sowie der Sensitivitätsstudie werden in dieser Masterarbeit vorgestellt und diskutiert.

Abstract

The **L**arge **H**adron **C**ollider is the biggest collider machine to be built so far and is in service since 2007. Creating proton-proton collisions physicists check the Standard Model of physics and search for new phenomena – physics beyond the Standard Model.

The approach used in the present thesis is called MUSiC – **M**odel **I**ndependent **S**earch **i**n **C**MS – and performs a systematic search for deviations from the Standard Model prediction of multiple final states recorded by the CMS detector at the LHC. In the context of this work, a new object – the τ Lepton – is introduced into the MUSiC environment. Subsequently, a scan of the full CMS dataset of 2011 is performed. In order to inspect the sensitivity of the MUSiC environment a sensitivity study with a W' signal is conducted. Hereby, the W' is a heave gauge boson – a hypothetical extension of the Standard Model. The results of the data analysis and of the sensitivity study are presented and discussed in the present master thesis.

Contents

CHAPTER 1

Introduction

In 2007, the Large Hadron Collider (LHC), the biggest collider in history of high energy physics has been successfully launched. Since then, a huge amount of data from proton-proton and lead ion collisions has been recorded by the four experiments conducted at the LHC: **C**ompact **M**uon **S**olenoid (CMS), **A T**oroidal **L**HC **A**pparatu**S** (ATLAS), **L**arge **H**adron **C**ollider **b**eauty (LHCb) and **A L**arge **I**on **C**ollider **E**xeriment (ALICE). The experimental setup of the CMS detector is discussed in section 3 of the present thesis.

This work concentrates on the data collected by the CMS detector in 2011 using a Model Unspecific approach for the data analysis. Over $5\,\text{fb}^{-1}$ of data have been collected and very many different dedicated analyses have been performed by scientists participating in the CMS experiment. In contrast to dedicated analyses, Model Unspecific Search in CMS uses a different approach – the framework is explained in section 4. The goal of the present work is to introduce a new physical object – the τ lepton to the MUSiC Framework. Subsequently, the data selection and analysis will be presented in section 5. The conclusion and the outlook will be given in section 6.

1.1 Notations and Conventions

In the present analysis, some conventions which are common in high energy physics will be used. In particular,

$$\hbar = \mathrm{c} = 1 \qquad (1.1)$$

where $\hbar$ is the Planck constant and c is the speed of light. As a consequence, the unit of mass $\frac{\mathrm{eV}}{\mathrm{c}^2}$ and the unit of momentum $\frac{\mathrm{eV}}{\mathrm{c}}$ both obtain the unit of energy eV. Furthermore, a four-vector momentum p of a particle is defined by

$$p \equiv \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} E \\ \vec{p} \end{pmatrix} \tag{1.2}$$

with the scalar product

$$p^2 \equiv p_0^2 - (p_1^2 + p_2^2 + p_3^2) = E^2 - \vec{p}^2. \tag{1.3}$$

The energy-momentum relation reads

$$E^2 = m^2 + \vec{p}^2 \tag{1.4}$$

and therefore Eq. 1.3 becomes

$$p^2 = m^2 \tag{1.5}$$

In addition to that, a coordinate system will be used in order to describe the directions within the particle detector CMS. A well established quantity for this purpose is the *pseudorapidity* η defined by

$$\eta \equiv -\ln\left[\tan\left(\frac{\theta}{2}\right)\right] \tag{1.6}$$

where θ is the angle to the particle beam in the $z - r$ plane, z being the direction of the proton-proton beam. The different values of η are, for example, visible in Figure 3.6.The angle in the $x - y$ plane (also called the transverse plane) is given the symbol ϕ. Thus, the coordinates x, y and z will be translated into η, ϕ and z. In addition to this, the variable $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ is used.

CHAPTER 2

Theory

2.1 The Standard Model

Today's understanding of modern high energy physics rests upon the Standard Model - the most successful theory to describe particle interactions. The latest discoveries at the Large Hadron Collider (LHC) at CERN[1] showing strong evidence for the existence of a Higgs boson once more highlighted the Standard Model as (for now) a theory closest to explaining the nature of particle physics.

In this model, three fundamental interactions can be described - the weak interaction, the strong interaction and the electromagnetic interaction. Before describing the experimental setup, a short introduction (based on the lectures [1],[2] and [3]) will be given in the following.

2.1.1 The Particle Zoo

The so called Particle Zoo is the set of all elementary particles described within the Standard Model (cf. Fig. 2.1). These particles are assumed to be pointlike and interact with each other via bosons – the force carriers. Each Standard Model particle has a mass m, and electrical charge q and carries a further set of quantum numbers allowing it to interact in a certain manner. Such quantum numbers are the spin S, the weak isospin I, the color and a number of further quantities. The particles are classified into three generations. The electron and the electron neutrino as well as the up- and the down-quark belong to the first generation. Correspondingly, the muon and its neutrino and the charm- and the strange-quark belong to the second, and the tau with its neutrino and the top- and bottom-quark belong to the third generation. To every particle with exceptions for the γ, the Z boson, the hypothetical Higgs

---

[1] French for European Organization for Nuclear Research – **C**onseil **E**uropéen pour la **R**echerche **N**ucléaire

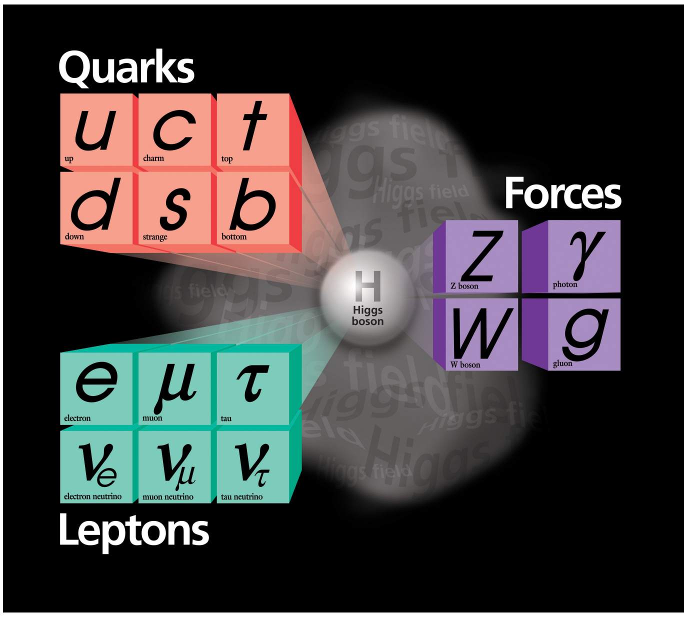

Figure 2.1: A sketch of the particles in the Standard Model. The interactions are transmitted via bosons (force carriers) - the gluon, the photon (γ) and the Z and $W^{\pm}$ bosons [4]

bosons and a few more hypothetical particles[2], there is one antiparticle.

2.1.2 Quantum Electrodynamics

The first and the most successful part of the Standard Model is **Q**uantum **E**lectro**d**ynamics (QED) - the theory describing the electromagnetic interaction with the symmetry $U(1)$. The history of QED goes back all the way to the 19th century when the first experimental hints of existence of an elementary charged particle, the electron, were obtained. However, the first theoretical approach successfully describing the electromagnetic interactions within the scope of perturbation theory was done in the first decades of the 20th century by such physicists as Dirac, Bethe and later on Feynman, Dyson and many more.

One introduces the Lagrangian $\mathscr{L}$ of a free electron and a free photon

$$\mathscr{L} = \bar{\Psi}\left(i\gamma^{\mu}\partial_{\mu} - m\right)\Psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.1}$$

where Ψ is the wave function of the electron, and γ^{μ} are the Dirac matrices. The indices μ and ν run through the values $\mu, \nu = 0, 1, 2, 3, 4$. Furthermore, tensor $F^{\mu\nu}$ is given by

$$F^{\mu\nu} = \partial^{\mu}A^{\nu} - \partial^{\nu}A^{\mu} \tag{2.2}$$

$$\tilde{F}^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\alpha\beta}F_{\alpha\beta} \tag{2.3}$$

---

[2] It is not known for sure whether neutrinos have antiparticles or ν and $\bar{\nu}$ are actually the same particle

where $A = A^\mu = \left(\phi, \vec{A}\right)$ is the four-dimensional potential. Using this notation, one obtains the lorentz invariant Maxwell equations in a compact notation

$$\partial_\mu F^{\mu\nu} = 0 \tag{2.4}$$

$$\partial_\mu \tilde{F}^{\mu\nu} = J \tag{2.5}$$

with the four-dimensional current $J = (\rho, \vec{j})$. In the introduced notation, Eq. 2.4 corresponds to the magnetic and Eq. 2.5 to the electric Maxwell equations.

The free (non-interacting) electron obeys the Dirac equation, the equation of motion, stating

$$\left(i\gamma^\mu \partial_\mu - m\right)\Psi = 0. \tag{2.6}$$

One very important statement which is made by quantum field theory is the postulation of *local* gauge invariance. Let us consider the Lagrangian from Eq. 2.1 and introduce a gauge transformation of the following form

$$\Psi \rightarrow \Psi' = U\Psi \tag{2.7}$$

$$\text{where } U = \mathrm{e}^{i\alpha} \tag{2.8}$$

Now, two cases are possible:

1. The parameter α does not depend on spacetime. In this case, the Lagrangian keeps its invariance.
2. The parameter α depends on the spacetime $\alpha = \alpha(x)$. In this case, the Lagrangian $\mathscr{L}$ does *not* stay invariant since there appears a term proportional to $\partial_\mu \alpha(x) \neq 0$.

The postulate now requires the local invariance of $\mathscr{L}$ for the second case[3]. One extends the differential operator ∂^μ by:

$$\partial^\mu \rightarrow \mathrm{D}^\mu = \partial^\mu - iQA^\mu. \tag{2.9}$$

From this one follows that the potential A_μ has to be transformed by

$$A_\mu \rightarrow A'_\mu = A_\mu + \frac{1}{Q}\partial_\mu \alpha(x) \tag{2.10}$$

in order to keep the Lagrangian $\mathscr{L}$ invariant. This results in an interaction between the electron an the photon fields. The Lagrangian now reads:

$$\mathscr{L} = \bar{\Psi}\left(i\gamma^\mu \partial_\mu - m\right)\Psi + Q\bar{\Psi}\gamma^\mu \Psi A_\mu - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \tag{2.11}$$

---

[3] At last, there is no fundamental argument to do so. The only convenience of that proceeding is the result which is in great agreement with the experiment

where the middle contribution corresponds to the coupling of the photon field A_μ to the electron Ψ arising from the demand for local gauge invariance.

This approach is very important and fundamental for the construction of a quantum field theory.

2.1.3 Quantum Flavourdynamics

The theory of weak interaction is called Quantum Flavourdynamics (QFD). In 1896, the α-decay of Uranium was discovered by Becquerel. This fundamental discovery corresponds to a new interaction which was later found to be the cause of the β-decay

$$\mathrm{n} \rightarrow \mathrm{pe}^- \bar{\nu}. \tag{2.12}$$

The discovery of the β-decay had a great influence on physics in general. Since the neutrino was not detectable back then (and is barely detectable now due to a very small cross section), there was a threat to the law of conservation of energy – there was no way to explain the continuous energy spectrum of the measured particles. As an "emergency" solution, Pauli postulated the neutrino in 1930 which was found in 1956 by Clyde Cowan and Frederick Reines and earned Reines[4] the Nobel Prize in the year 1995. As a further step, Glashow, Salam and Weinberg introduced the electroweak theory and deduced the states W^0, $W^\pm$ and B. The charged W bosons of the weak interaction can be described as a linear combination of two states W^1 and W^2:

$$W^+ = \frac{1}{\sqrt{2}}\left(W^1 + iW^2\right) \tag{2.13}$$

$$W^- = \frac{1}{\sqrt{2}}\left(W^1 - iW^2\right) \tag{2.14}$$

Furthermore, one can show that for the photon field A and the Z boson the following relation is valid

$$Z = -\sin\theta_\mathrm{W}\, B + \cos\theta_\mathrm{W}\, W_0 \tag{2.15}$$

$$A = +\cos\theta_\mathrm{W}\, B + \sin\theta_\mathrm{W}\, W_0 \tag{2.16}$$

where θ_W is the Weinberg angle with $\sin^2\theta_\mathrm{W} \approx 0.23$. The $W^\pm$ and Z bosons are particularly remarkable because the theory predicts them to be massless. However, the experiment states that all the force carriers of the weak interaction possess a (rather large) mass. This will lead to a further extension of the Standard Model discussed in 2.1.5.

---

[4] Clyde Cowan died in 1974

2.1.4 Quantum Chromodynamics

The first theoretical approach to the strong interaction (QCD), the interaction responsible for the forces that keep the nuclei of atoms together, was performed by H.Yukawa in 1935 and gained him the Nobel Prize in the year 1949. Yukawa described the scatter process between a proton and a neutron

$$n + p \rightarrow n + p \tag{2.17}$$

by the exchange of a spinless particle. In 1963, Gell-Mann and Zweig introduced the quarks, the constituents of hadrons. The quarks carry the charge $\pm 1/3$ or $\pm 2/3$ and, as found out later, a color charge and couple to gluons, the force carriers of the strong interaction. Hereby, the color charge consists of two quantum properties – a color and an anti-color. The hadrons which, as stated earlier, consist of quarks are always colorless. This means that only combinations of quarks with colors and the corresponding anti-colors are realized in nature. An important difference to QED arises from the fact that gluons, unlike photons, couple to each other. The structure of QCD is nevertheless similar to QED with the difference that its corresponding symmetry group is the $SU(3)$. Due to a strong coupling (relative to QED), the predictions of QCD possess the larger uncertainties since a perturbative approach is not always justified. Furthermore, the coupling of the QCD decreases with high energies leading to so called "asymptotic freedom" – quarks can not exist as free particles. The energy stored in their binding is used to create new quarks – this phenomenon is called confinement. Modelling the QCD processes is therefore a big challenge, especially at the Large Hadron Collider.

2.1.5 The Higgs Mechanism

In order to lend mass to the Z and $W^{\pm}$ bosons, one extends the Lagrangian by:

$$\mathscr{L} = \left(\mathrm{D}_{\mu}\phi\right)\left(\mathrm{D}_{\mu}\phi\right)^{*} - V(\phi) \tag{2.18}$$

$$\text{where } \phi = \frac{1}{\sqrt{2}}\left(\phi_1 + i\phi_2\right) \tag{2.19}$$

with the potential $V(\phi)$:

$$V(\phi) = -\mu^2\phi\phi^{*} + \lambda^2\left(\phi\phi^{*}\right)^2. \tag{2.20}$$

Hereby, ϕ is a complex scalar field describing a spinless particle and its anti-particle. The potential $V(\phi)$ from Eq. 2.20 is given the colloquial name "The Mexican Hat Potential" (cf. Fig. 2.2).

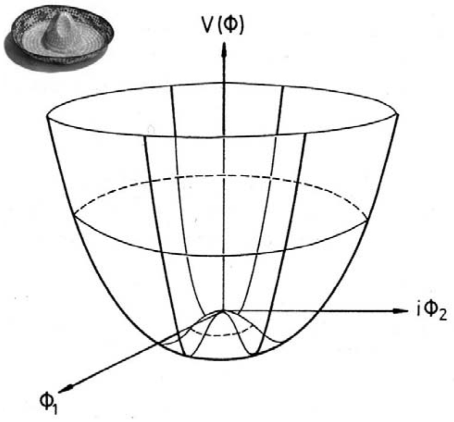

Figure 2.2: A sketch of the Higgs "Mexican Hat" potential as function of ϕ_1 and ϕ_2 [1]

This extension of the Standard Model Lagrangian introduces masses to the mass-carrying bosons Z and $W^{\pm}$ also predicting the following relation:

$$\frac{m_{\mathrm{W}}}{m_{\mathrm{Z}}} = \cos\theta_W \tag{2.21}$$

where θ_W is the Weinberg angle. The ratio of the measured masses $m_{\mathrm{W}} = 80.399 \pm 0.023$ GeV and $m_Z = 91.1876 \pm 0.0021$ GeV [5] is in good agreement with the value of the Weinberg angle θ_{W} discussed earlier. Furthermore, the new Lagrangian introduces a new particle with spin 0, the Higgs boson. While making various predictions about the consequences of the new contributions to the Lagrangian, the extended theory does not predict the mass of the Higgs boson – it has to be determined experimentally.

2.1.5.1 The discovery of a new boson at the Large Hadron Collider

On July 4th 2012, the CMS and ATLAS collaborations at the LHC announced the observation of a new heavy boson with a mass of $\approx 125\,\mathrm{GeV}$ (cf. Fig. 2.3) [6]. While it is still to investigate closely whether the new particle is the predicted Higgs boson, there are many evidences, such as coupling strengths to other particles, decay modes etc. that the observed boson is in fact the Higgs boson. A final confirmation of this hypothesis would be one of the greatest triumphs of the Standard Model in history of physics.

2.2 Physics Beyond the Standard Model

As stated earlier, the Standard Model is an excellent theory being able to predict many phenomena found in the experiment. However, many theoretical issues remain unresolved and also not fully understood by the scientists. Even more, certain predictions of the Standard Model have been disproved by the experiment. For instance, the Standard Model is not

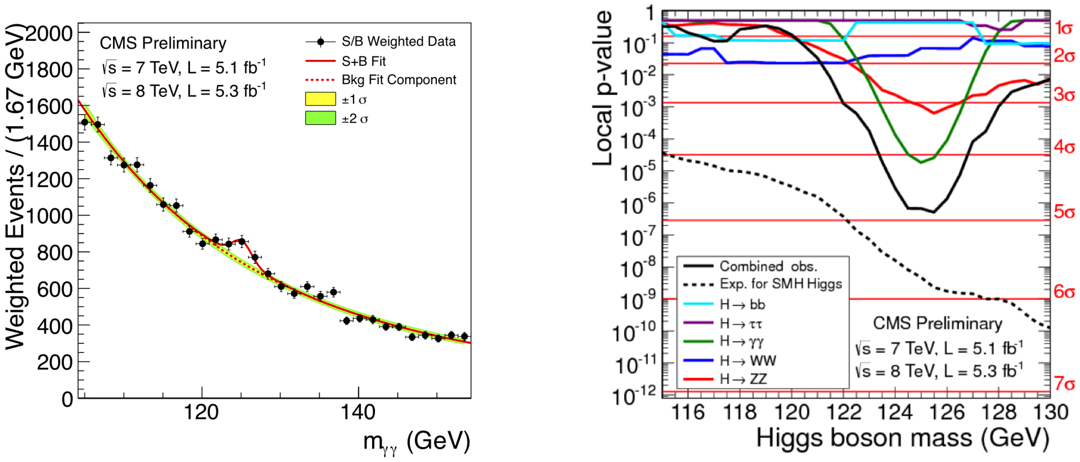

Figure 2.3: On the left: weighted number of observed events as a function of the invariant mass of two photons in the decay channel $H \to \gamma\gamma$
On the right: the local significance of the combined hypothetic Higgs signal for different decay channels [6]

in agreement with the general theory of relativity and its certain predictions which have been confirmed by the experiment, such as gravitation itself, dark matter and the fact that neutrinos are not massless like the Standard Model suggested at first. These facts imply that there might be phenomena beyond the Standard model. In the following, some unresolved issues of the Standard Model will be discussed and some possible scenarios for New Physics will be specified.

2.2.1 Unresolved issues of the Standard Model

Astronomic observations state that a large fraction of the universe is filled with massive "dark" matter which only presents itself through its mass – it is not visible in any other way. Furthermore, an expansion of the universe is observed. As one possible reason for that, so called "dark" energy is postulated – energy responsible for the expansion. The Standard Model does not have an explanation for neither one of these observations.

In addition to that, all efforts to unify all four forces into one unifying theory (**G**rand **U**nifying **T**heory) did not succeed. It is still unknown how to include the gravitation into the modern theory of particle interactions. Also in the experiments performed so far, violation of charge parity symmetry is observed. While its nature is studied and understood, this effect is not sufficient in order to explain the asymmetry between matter and anti-matter in the universe.

A further problem of the Standard Model is the so called fine-tuning. It appears that many parameters of the Standard Model have to be chosen very precisely in order to agree with

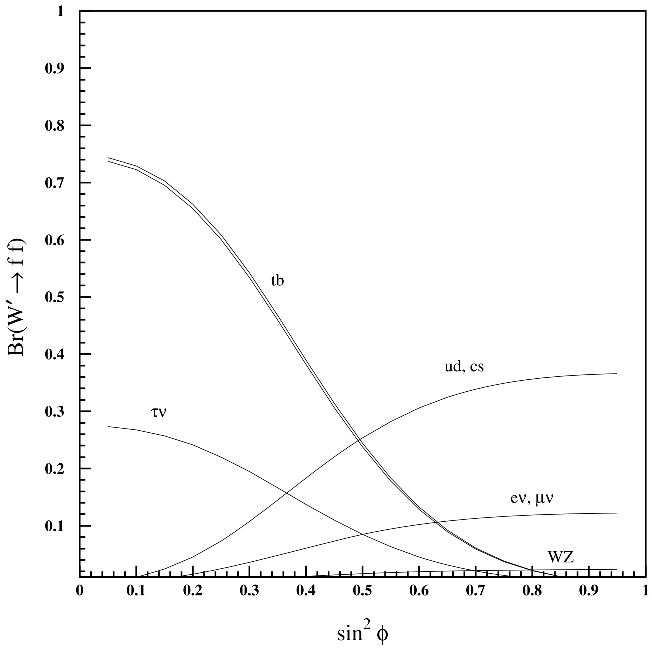

Figure 2.4: The branching ratio of the W′ into fermions as a function of the sine of the mixing angle ϕ. The branching ratio does not depend on the mass of the W′. For small values of ϕ, the W′ only couples to the third generation of fermions [7].

experimental results. It is considered that a theory would be much more satisfying if all its parameters could be deduced theoretically from only one or a few fundamental constants. Since there is no elementary reason for the fine-tuning of so many parameters, this approach of the Standard Model is not yet justified.

Also, the Standard Model does not give an explanation for the fact that the strengths of the different interactions including gravitation lie orders of magnitude apart – the so called hierarchy problem.

2.2.2 New Heavy Gauge Bosons

There are several approaches for solving the theoretical problems of the Standard Model such as Supersymmetry, introduction of extra space dimension or String theories. One possible solution of problems listed above is an extension of the Standard Model. In many theories, the gauge sector is broadened corresponding to new gauge bosons. One potential approach is the W′ model proposed by Altarelli [8]. In this model, a new heavy "carbon" copy of the W′ is introduced. All properties of the W′ are as far as possible identical to those of the Standard Model W. Because of the high mass of the W′ the hadronic decay into top quark is possible. The decay into Standard Model bosons like W and Z is allowed in general and can be the dominant process for heavy W′ bosons. However, in models considered in the present thesis, this decay channel is assumed to be suppressed. One particular interesting extension of the Standard Model featuring heavy gauge bosons introduces a mixing angle ϕ for the branching ratios into the Standard Model particles [7] as a new parameter. In this case, the branching ratio into fermions of the first and second generation is given by

$$\Gamma(\mathrm{W}' \to f\bar{f}') = \Gamma_0 \frac{m_{\mathrm{W}'}}{m_{\mathrm{W}}} \cdot \tan^2\phi \tag{2.22}$$

and the branching ratios into fermions of the third generation

$$\Gamma(\mathrm{W}' \to f\bar{f}') = \Gamma_0 \frac{m_{\mathrm{W}'}}{m_{\mathrm{W}}} \cdot \tan^2\phi \cdot \left(1 - \frac{1}{\sin^2\phi}\right)^2 . \tag{2.23}$$

Hereby, the following relation applies for the mass of the W′:

$$m_{\mathrm{W}'}^2 = \frac{m_0^2}{\lambda \sin^2\phi \cos^2\phi} \tag{2.24}$$

where m_0 is the mass of the ordinary W boson in leading order. $\lambda = \frac{v^2}{u^2}$ is a ratio of the electroweak scale v and u is a higher scale than v. Note that in order to keep the approach perturbative the following relation has to be satisfied:

$$0.03 < \sin^2\phi < 0.96 \tag{2.25}$$

since in cases where the latter relation is not satisfied the coupling

$$G_L' = -\frac{g}{\sqrt{2}} \tan\phi, \qquad \text{for 1st and 2nd generations} \tag{2.26}$$

$$= -\frac{g}{\sqrt{2}} \tan\phi \left(1 - \frac{1}{\sin^2\phi}\right) \qquad \text{for 3rd generation} \tag{2.27}$$

is too strong for the perturbative approach.

For small values of ϕ, the coupling to the first and second generations disappears and the heavy boson couples almost exclusively to the third generation (cf. Fig. 2.4). This will be particularly interesting for this work since the τ lepton will be investigated more closely.

Another scenario proposes a new $U(1)$ group corresponding to a new heavy Z boson [9]. In this theory, the Z′ is a heavier copy of the Standard Model boson. Its most sensitive decay channels at the LHC would be Z′ → $l^+ l^-$.

CHAPTER 3

Experimental Setup

The data used in the present work has been recorded by the Compact Muon Solenoid (CMS) detector at the LHC in Cern, Geneva. The Large Hadron Collider is an outstanding experimental machine, a result of hard work of very many physicists all over the world. In the following, a brief overview of the collider and the detector will be given [10][11].

3.1 General Background

In the first years of particle physics, the physicists used cosmic particles for the experiments since there were no machines for producing particle beams. Later on, fixed target experiments were built in different research centers, such as CERN in Geneva. However, it is energetically disadvantageous to use one high energy particle beam aiming at a fixed target because a very large fraction of the beam energy will be converted into the momentum of the new particles due to the momentum conservation and only a small fraction of the beam energy will be available for particle production. From this point of view, the collider experiments using two beams accelerated towards each other have a big advantage – here, the sum of momenta of the colliding particles is 0 and all energy of the beams is available for particle production. This energy is called the center of mass energy $\sqrt{s}$:

$$\sqrt{s} = 2E_{\text{Beam}}. \tag{3.1}$$

If the beam consists of pointlike particles, e.g. electrons, $\sqrt{s}$ is the energy available for the particle production. If the beam constists of composed particles, e.g. protons at the LHC, the energy E_{Beam} is distributed among the constituents. One then defines an effective center of mass energy $\sqrt{\hat{s}}$ which is, unfortunately, unknown since one does not know which energy fraction the interacting constituents are carrying in a particular collision event. While reconstructing the interaction inside the beam pipe one therefore can only consider the law of conservation of energy and momentum in the plane perpendicular to the beam.

One important property of a particle colliding machine is its instantaneous luminosity $\mathscr{L}$. This quantity is a measure for the quality of the optics and the general setup of the collider. In other words, it rates the ability of the collider to collimate the beams in order to achieve the best experimenting setup and is approximately given by

$$\mathscr{L} = \frac{nN^2 f}{4\pi\sigma_x\sigma_y} \tag{3.2}$$

where σ_x and σ_y are the cross sections of the beam in the corresponding directions, f is their rotational frequency, n the number of bunches and N the number of particles per bunch. The corresponding rate of interactions (or so called "events") $\dot{N}$ then reads

$$\dot{N} = \mathscr{L} \cdot \sigma \tag{3.3}$$

where σ is the interaction cross section of the respective process. The total number of events N then corresponds to the luminosity integrated over the time. Luminosity has the unit barn^{-1} with

$$1\mathrm{b} = 10^{-28}\mathrm{m}^2. \tag{3.4}$$

Consequently, a unit usually used for the ranking of the amount of data taken is the barn^{-1}. As one challenging goal of the experimentalist is the measurement of the cross section, the determination of the luminosity is a crucial task.

3.2 The Large Hadron Collider

The Large Hadron Collider (LHC) is located at CERN and is the biggest colliding machine to be built and operated so far. The LHC beam tunnel has a circumference of 27 km and is designed to accelerate two proton beams achieving a center of mass energy of up to $\sqrt{s} = 14\,\mathrm{TeV}$. In the year 2011, collisions at $\sqrt{s} = 7\,\mathrm{TeV}$ were performed and over 5fb^{-1} of data have been recorded by the CMS detector [13]. In 2012, collisions at $\sqrt{s} = 8\,\mathrm{TeV}$ are being performed and the amount of recorded data exceeded the size of the 2011 dataset. In the present work, only the data recorded in the year 2011 will be discussed.

The proton beam is injected via an injection chain since the apparatus of the LHC is only suited for high-energetic particles. Several particle accelerators which used to be in service earlier are now used as pre-accelerators (cf. Fig. 3.1). The LHC is supplied with protons which come from the injector Linac2 to the Proton Synchrotron Booster and Proton Synchrotron and eventually are accelerated to 450 GeV by the Super Proton Synchrotron before they enter the beam pipe of the LHC. The particles are kept on their trajectory using 1232 superconducting

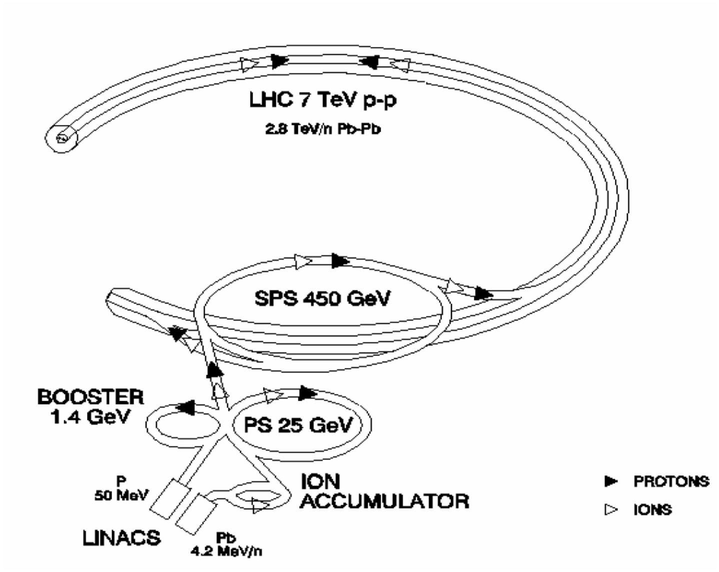

Figure 3.1: The injection chain of the LHC collider [12]

magnets being able to deliver magnetic fields of up to 4T. Each beam consists of 2808 bunches each containing $N = 1.15 \cdot 10^{11}$ particles. The total energy stored in the beam at $\sqrt{s} = 7$ TeV corresponds to 362 MJ.

3.3 Compact Muon Solenoid

The data analyzed in this work was recorded by the Compact Muon Solenoid (CMS) detector in 2011. The CMS experiment is a general-purpose detector and is installed in an underground cavern (around 100 m depth) near Cessy, France. Its main specifications are

- muon detection with a good dimuon mass resolution of roughly 1% at 100 GeV and capability of determining the muon charge at momenta $p < 1$ TeV
- good resolution in the momentum reconstruction of charged-particle. This corresponds to an efficient triggering and especially offline tagging of τ- and b-jets
- good electromagnetic energy resolution[1]
- efficient measurement of missing-transverse-energy and dijet-mass resolution

In the following, a brief overview of the detector parts is given (cf. Fig. 3.2).

3.3.1 Inner Tracker System

The inner tracker system of the CMS detector consists of a pixel detector with three barrel layers and a silicon strip tracker and is responsible for the measurement of the trajectories of

---

[1] especially in dielectron and diphoton channels since the Higgs boson should present itself in the $\gamma\gamma$ channel

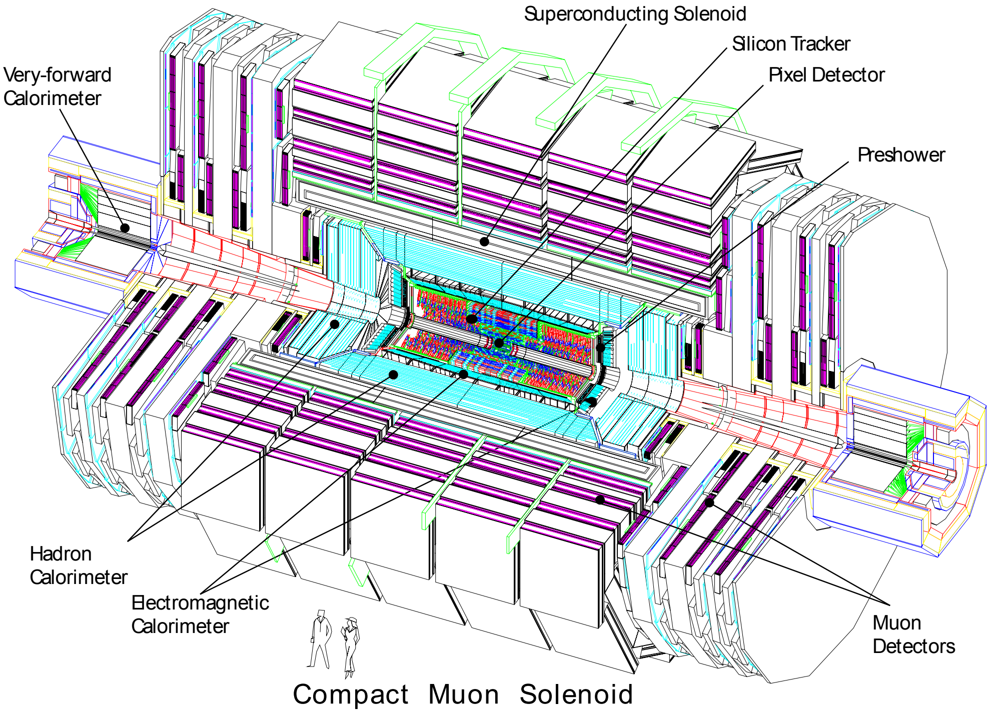

Figure 3.2: The CMS detector at CERN [11]

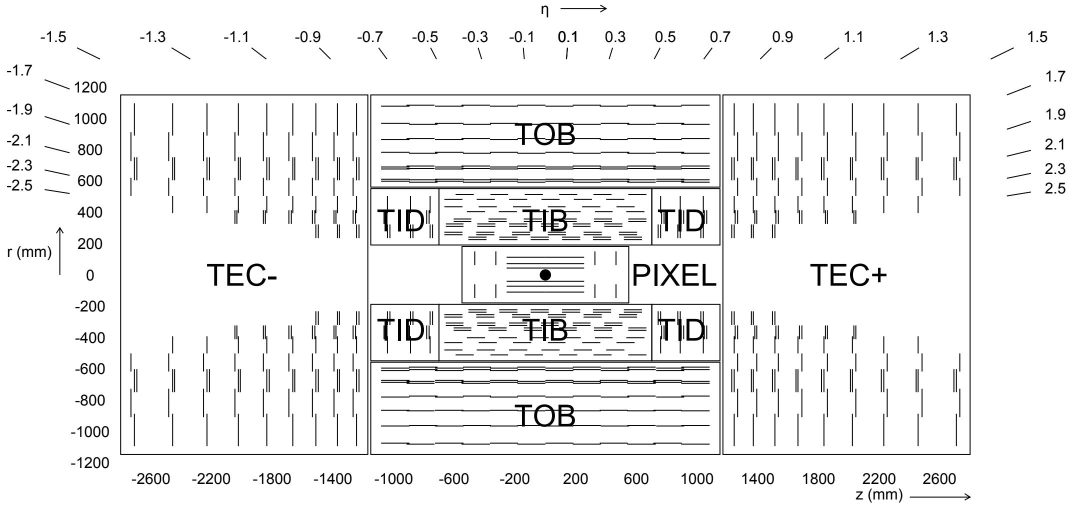

Figure 3.3: A schematic cross section of the CMS inner tracker system. Hereby, TIB and TID correspond to Tracker Inner Barrel and Discs, TOB corresponds to Tracker Outer Barrel and TEC stands for Tracker EndCaps [11]

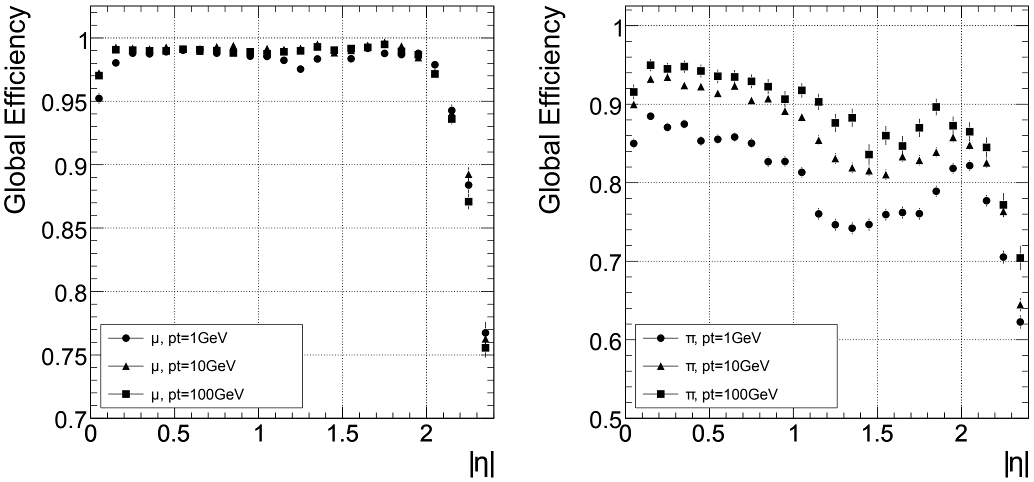

Figure 3.4: Global track reconstruction efficiency for muons (on the left) and charged π (on the right) with different transverse momenta as function of η [11]

charged particles originating from the interaction point (cf. Fig. 3.3).

Silicon Pixel Tracker

The silicon pixel tracker is located next to the interaction point and consists of 66 million pixels, each having the dimensions of $100\,\mu\text{m} \times 150\,\mu\text{m}$. This system is very important for the measurement of the impact parameter of the interaction allowing for good secondary vertex reconstruction. The pixel tracker has a pseudorapidity range of $-2.5 < \eta < 2.5$. Since the silicon pixel tracker is the part of detector closest to the interaction point, it is optimized for the high event rate. The strong magnetic field of 4 T causes a Lorentz drift of charged particles passing through the silicon pixel tracker strong enough to hit more than one pixel. Together with other optimizations, this allows for a three-dimensional resolution of about $20\,\mu$m.

Silicon Strip Tracker

The silicon strip tracker consists of three different subsystems and is built of single sided p-doped on n-doped sensors. The Tracker Inner Barrel and Discs (TIB and TID respectively) have each 4 layers and 4 discs at the end. In TIB, TID (cf. Fig.3.3) and inside the TEC, the sensors have the thickness of $\approx 320\,\mu$m. The overall single point resolution of the TIP and TID varies between 23μm and 35μm. The Tracker Outer Barrel (TOB) surrounding the TIB and TID consists of 6 barrel layers of sensors with the thickness of $\approx 500\,\mu$m. The single point resolution of TOB varies within $35\,\mu$m and $53\,\mu$m.

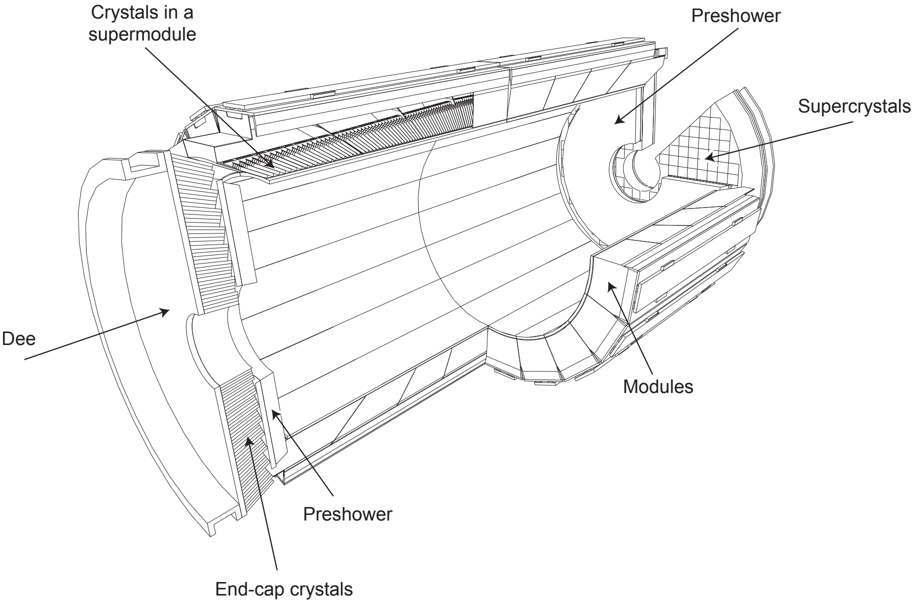

Figure 3.5: The ECAL of the CMS experiment [11]

3.3.2 The electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) (cf. Fig. 3.5) is the next layer after the tracker and has the aim of measuring the energy of particles passing through it interacting electromagnetically, e.g. electrons or photons. It is made of 61200 lead tungstate crystals in the central barrel of the CMS detector and has 7324 crystals in each of the two endcaps (cf. Fig.3.2). While passing through the ECAL, the electrons interact via bremsstrahlung and ionisation and photons via pair production, Compton effect and photo effect. Lead tungstate ($PbWO_4$) has a high density of $8.28\,\mathrm{g/cm^3}$ and a short radiation length of 8.9mm. Larger improvements of $PbWO_4$ scintillation characteristics in the recent years make it a suitable material for the LHC – about 80% of photons are emitted within the bunch crossing time of 25 ns. The barrel and the end caps of the ECAL have pseudorapidity ranges of $-1.479 < \eta < 1.479$ and $1.479 < |\eta| < 3.0$ respectively while covering the full 360° angle in ϕ. The resolution of the ECAL has been determined to [11]:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right)^2 + \left(\frac{0.12}{E}\right)^2 + (0.3\%)^2 \tag{3.5}$$

where E is in GeV.

3.3.3 The hadronic calorimeter

The hadronic calorimeter (HCAL) is built around the ECAL of the CMS detector. Its purpose is the measurement of the energy of particles interacting via the strong force, for example

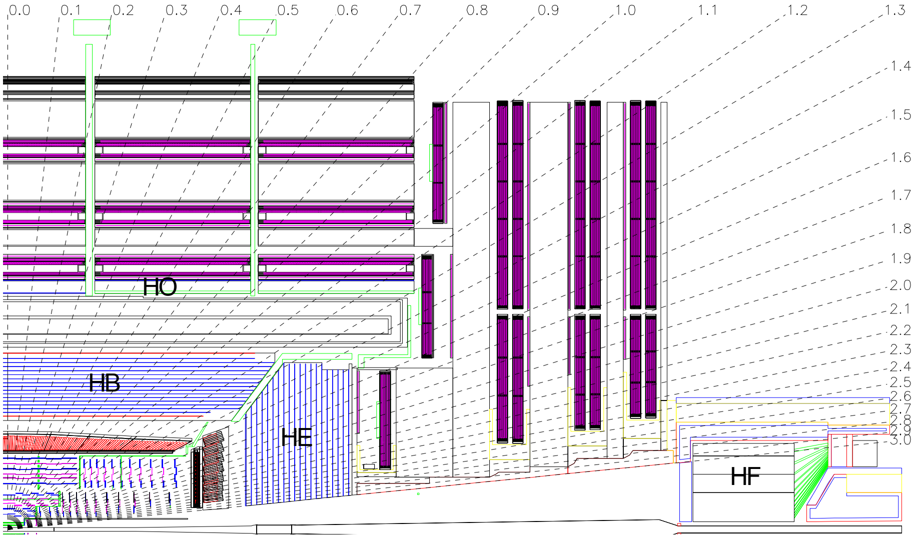

Figure 3.6: The HCAL of the CMS experiment [11]

pions, which will appear as jets in the final state. The HCAL consists of 70% copper and 30% zinc resulting into a radiation length of roughly 1.49 cm and an interaction length of 16.42 cm. The barrel of the HCAL has the pseudorapidity coverage of $|\eta| < 1.3$, the endcap covers $1.3| < \eta| < 3.0$ with a partial overlap, as can be seen in Fig.3.6. Furthermore, forward hadron calorimeters are installed in the back and front of the detector having 11.2 m distance to the interaction point. These calorimeters cover pseudorapidities up to $|\eta| = 5.2$. They are not used in the present analysis.

3.3.4 CMS Magnet

The superconducting CMS magnet is installed between the HCAL and the muon system, having a length of 12.5m and storing the energy of up to 2.6 GJ while being operated at very low temperature of 1.8 K. The magnet produces a 3.8T magnetic field which is essential for a high resolution measurement of charged particles which are produced at the interaction point. In Figure 3.4, track reconstruction efficiencies for muons and π-mesons are demonstrated. As one will see later on, reconstructing charged pions is crucial for τ reconstruction and is therefore of a special interest for the present work.

3.3.5 The Muon System

As the name of the detector suggests, the detection of muons is of special importance. Originally designed to be able to find the "gold plated" decay of the Standard Model Higgs boson into four muons, the CMS detector has a distinguished muon detection system allowing for

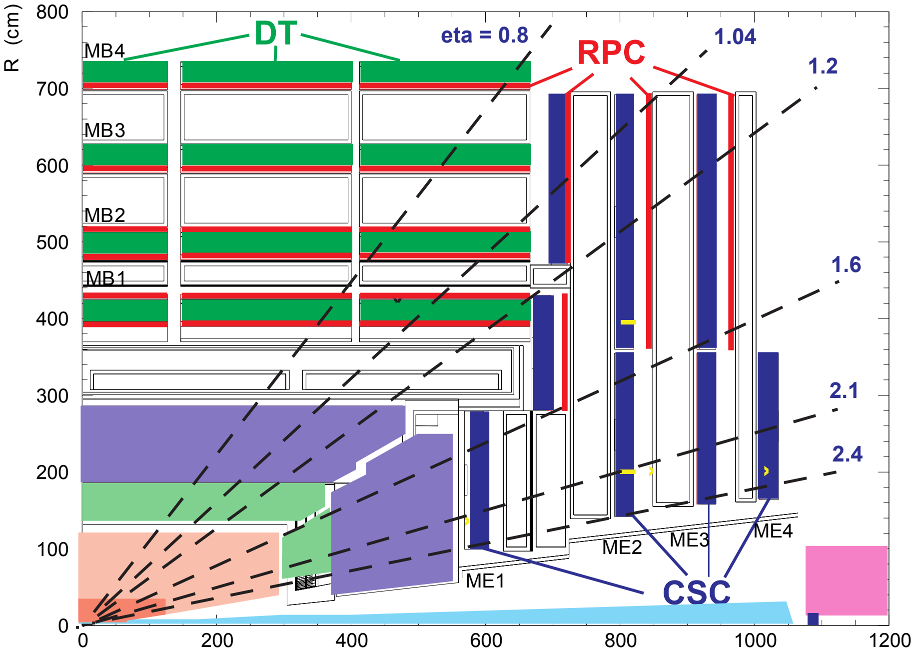

Figure 3.7: A representative part of the muon system of CMS [14]

a precise measurement of muon's charge and momentum[2]. The muon detection system is located as the last layer of the detector. Since muons have a relatively long mean lifetime and barely interact with the ECAL and the HCAL material, they will travel through the whole detector and can be registered in the muon system. Furthermore, one expects the muons to have enough hits in the tracker.

The muon system consists of three kinds of gaseous detectors - a drift tube system (DT), a cathode strip chamber system (CSC) and a resistive plate chamber system (RPC) (cf. Fig.3.7).

Drift tube system (DT)

The drift tube system is used in the barrel region of CMS since in this region the muon rate is expected to be low and the magnetic field to be uniform. The DTs in the barrel cover an η interval of $|\eta| < 1.2$. They are formed of four stations which are arranged as concentric cylinders around the beam pipe each. Each of the inner three cylinders have 60 drift chambers (cf. Fig. 3.8) whereas the outer cylinder has 70. The first inner stations contain eight chambers grouped in clusters of four and provide the measurement of the coordinates of the muon in the $r-\phi$ plane and additional four chambers which are responsible for the measurement of the z-coordinate of the muon. In order to achieve the best coverage, the drift cells of each chamber are displaced by a half length relative to each other (cf. Fig. 3.9).

---

[2] RWTH Aachen provided an eminent contribution to the muon detection system of CMS

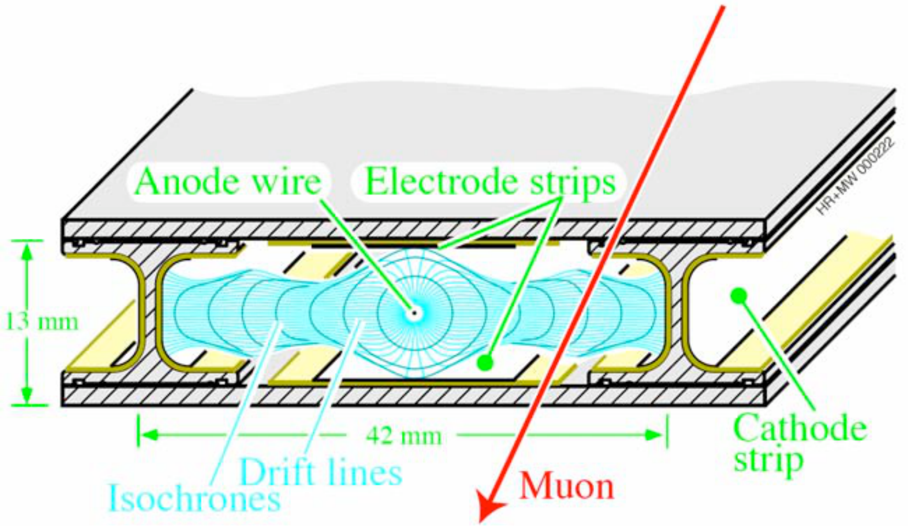

Figure 3.8: A muon detection cell. The voltage of the electrodes is +3600 V for wires, −1200 V for cathodes and +1800 V for strips [11]

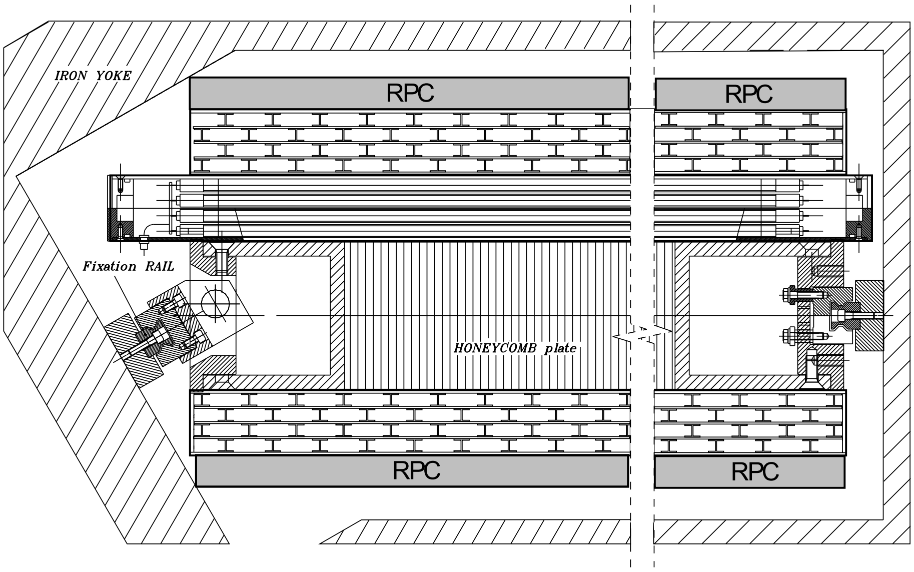

Figure 3.9: A DT chamber installed at CMS. RPCs attached to the DT chambers with plates are not shown [11]

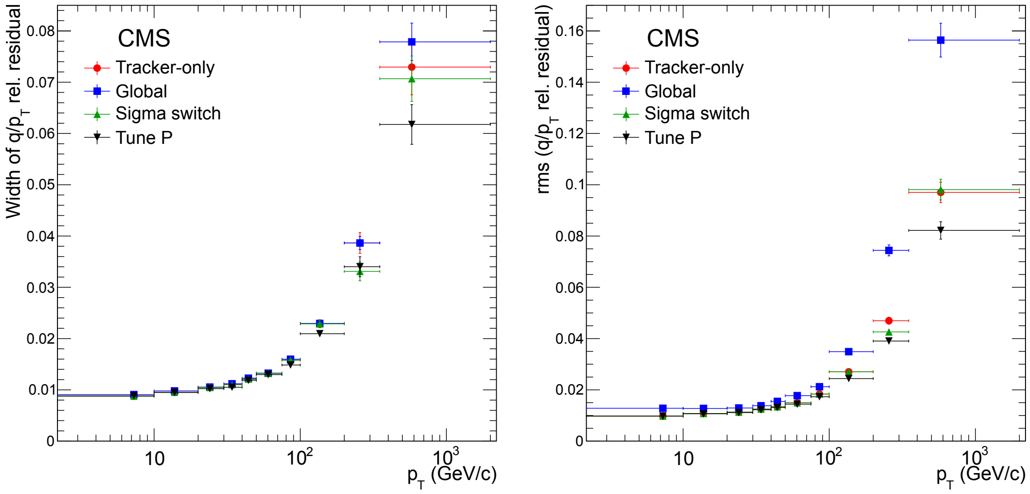

Figure 3.10: Width of the Gaussian and root mean square fit for different muon reconstruction algorithms used by CMS as a function of transverse momentum [15]

Cathode strip chambers (CSC)

In the endcap regions of CMS, CSCs are used. Since the muon rate in these regions is higher, one profits from their fast response time, fine segmentation and radiation resistance. The CSCs cover the η range of $0.9 < |\eta| < 2.4$. Four stations of CSC placed in each endcap are arranged perpendicularly to the beam direction and alternate with flux return plates. The wires of the cathode strips are installed radially outward and allow for a precise measurement in the $r - \phi$ plane. The anode wires, however, are installed approximately perpendicular to the strips and are used for the measurement of η and beam-crossing time of the muon.

The η region covered by the muon system is $|\eta| < 2.4$ and guarantee a robust muon reconstruction achieving an efficiency of up to 99% [15]. However, the momentum resolution can vary depending on the information used during the reconstruction (cf. Fig. 3.10). Furthermore, both DT and CSC are capable of triggering on the transverse momentum p_T of a muon independently from the rest of CMS.

Resistive Plate Chamber system (RPC)

In order to reduce the uncertainty of the background rates and the measurement of the beam-crossing time, an additional, dedicated trigger system of resistive plate chambers is used in the barrel and in the endcaps. This system allows for fast triggering covering pseudorapidity ranges up to $|\eta| < 1.6$. The RPC consists of double-gap chambers which are embedded in 6 layers in the barrel region of the detector. Hereby, two layers are installed in the first and the second station making the first two stations redundant and allowing for low-p_T triggering since a low-energetic muon might not reach the outer station.

3.3.6 Particle flow reconstruction algorithms

The general strategy of CMS for reconstruction of stable particles within an event, e.g. electrons, charged hadrons etc. is to combine the information from all subdetectors of CMS into one interrelated object – the so called "Particle Flow" (PF) technique [16]. This approach allows for a much better reconstruction efficiency, as one will see later. The physical objects produced by this algorithm are colloquially called PFjets, PFmuons and so on.

3.3.7 The Trigger System of CMS

In order to keep the size of measured data in acceptable ranges, a sophisticated trigger-system is used in CMS. The design beam crossing interval of 25 ns corresponds to an interaction frequency of 40 MHz making efficient triggering an essential equipment for CMS. In 2011, the beam crossing interval was 50 ns. Furthermore, several interactions can occur during one bunch crossing leading to so called "pile-up". The data rate reduction is performed in two steps: Level-1 (L1) Trigger and High-Level-Trigger (HLT). While the L1 trigger mostly consists of special designed programmable electronics, the HLT is a large software system using over one thousand processors. The factor of the rate reduction achieved by both L1 and HLT triggers has the order of magnitude of 10^6.

3.4 Computing Environment

Handling the huge amount of data produced by the CMS detector in 2011 is a very challenging and exciting task. In this section, a brief overview of the software used for data analysis will be given.

3.4.1 The Worldwide LHC Computing Grid

The first step in order to analyze the recorded data is the assembling of a big computational network since it is by far not possible to satisfy the computational needs of over two thousand scientists with one or two even very large computing elements [17]. 15 Petabytes of data are produced by the LHC annually and are processed in several steps. Starting with the biggest computational center located at CERN, the so called `Tier0`, the data is transfered to several `Tier1` centers and then made available to the analyzer at `Tier2` and `Tier3` centers. RWTH Aachen maintains a `Tier2` computing center with over 2000 processors.

3.4.2 CMSSW and PXL

The CMS collaboration provides an own software framework, CMSSW[3], in order to ensure a common computational basis for all collaborators [18]. In the present work, CMSSW is used in order to prepare the reconstructed data for analysis (the so called "skimming") [19]. The "skimmed" data is then stored in PXL containers [20][21] and can be viewed with the VISPA software [22]. After this step, a specially designed software, the MUSiC framework (see chapter 4), is used for further data analysis.

3.4.2.1 Production of Monte Carlo samples

Monte Carlo (MC) samples are of a big importance for the analysis. In order to predict the data later measured in the experiment, large samples of physical interactions are simulated. First, the physical process including initial and final state radiation is simulated. In the next step, the interaction of the particles produced by the interaction with the detector material is simulated [23]. The simulated events are then reconstructed in the same way as real data. Comparison of the simulated prediction with the data plays a central role in the present thesis.

3.4.3 ROOT

ROOT is a standard high energy physics analysis tool used by the CMS collaboration at CERN [24]. ROOT offers a broad variety of tools for data analysis, plotting and further calculations.

---

[3] Compact Muon Solenoid SoftWare

CHAPTER 4

The MUSiC Framework

4.1 Introduction into MUSiC

The Model Unspecific Search in CMS (MUSiC) framework [25] is the main software tool used in the present work . In past years, many dedicated analyses have been performed aiming to analyze one special final state at the CMS detector and searching for one or a few new physics processes. Being based on a particular theory predicting new physics in one exclusive way, dedicated analyses, however, mostly have the potential to find evidence for new physics in the expected region. In contrast to this, MUSiC offers an alternative *unspecific* approach and is capable of searching for new physics without having to presume a certain exotic theory. Starting with the Standard Model, MUSiC systematically compares the prediction of Monte Carlo simulations with the data measured by CMS. This way, also new phenomena not predicted by a certain theory can be detected by MUSiC. Similar searches have been performed at experiments conducted in the past at Tevatron [26], LEP [27], and at HERA (for example [28][29] [30]). However, MUSiC is the first global model independent approach performed at the CMS detector at LHC. Various studies have already been performed using the MUSiC framework [31][32][33].

4.2 Workflow

The main idea of MUSiC is a systematic search for deviations in a large number of final states. Based on the physical objects contained in an event, the final states are sorted into "Event Classes" (cf. Fig. 4.1). At the moment, MUSiC uses the following physical objects:

- electrons - e
- muons - μ

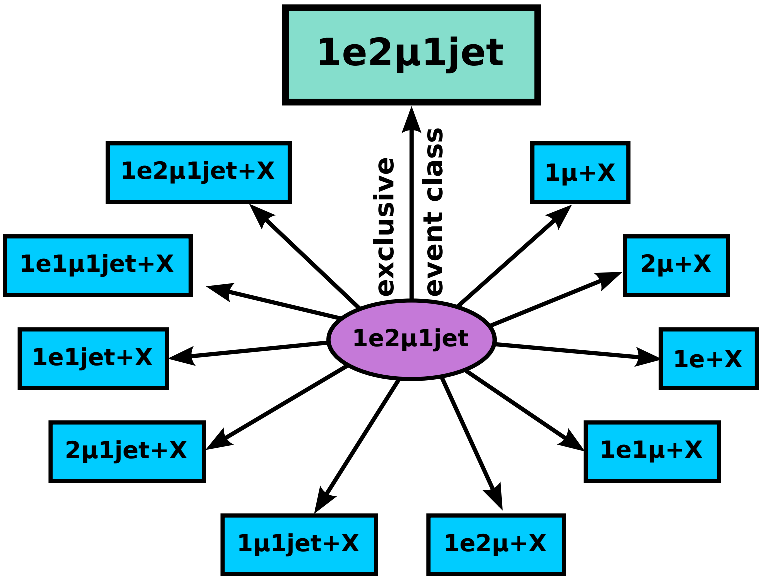

Figure 4.1: An example of how one final state is sorted into event classes. Here, a final state of 1e2μ1jet is stored in exactly one exclusive and several inclusive classes.

- photons - γ
- jets
- Missing Transverse Energy - MET or $\not{E}_T$[1]

In addition to those objects, b-jets can be optionally used [34]. MUSiC also accounts for the charge of leptons. Both these features are not used in the present work.

In this thesis, a new object, the τ lepton in its hadronic decay is introduced to the MUSiC framework (the formal definition of τ in MUSiC is discussed in section 4.3).

In the following, the workflow of MUSiC will be explained [31][32][33].

First of all, the CMS data has to be prepared for the analysis. This step is colloquially called "skimming" and is performed by a special software piece – the skimmer. The skimmer saves all important information about the event and the particles it contains. In particular, the kinematic properties of the objects listed above are saved. Furthermore, all relevant signals from the CMS triggering system are stored. Those trigger signals are used in the next step – during the classification.

Based on physical objects, one final state is stored (one also says "classified") in one exclusive class which only contains the objects specified in the name of the class, and in several inclusive classes containing only the minimal set of the objects in the final state.

---

[1] Neutrinos can not be detected by CMS and will appear as missing transverse energy

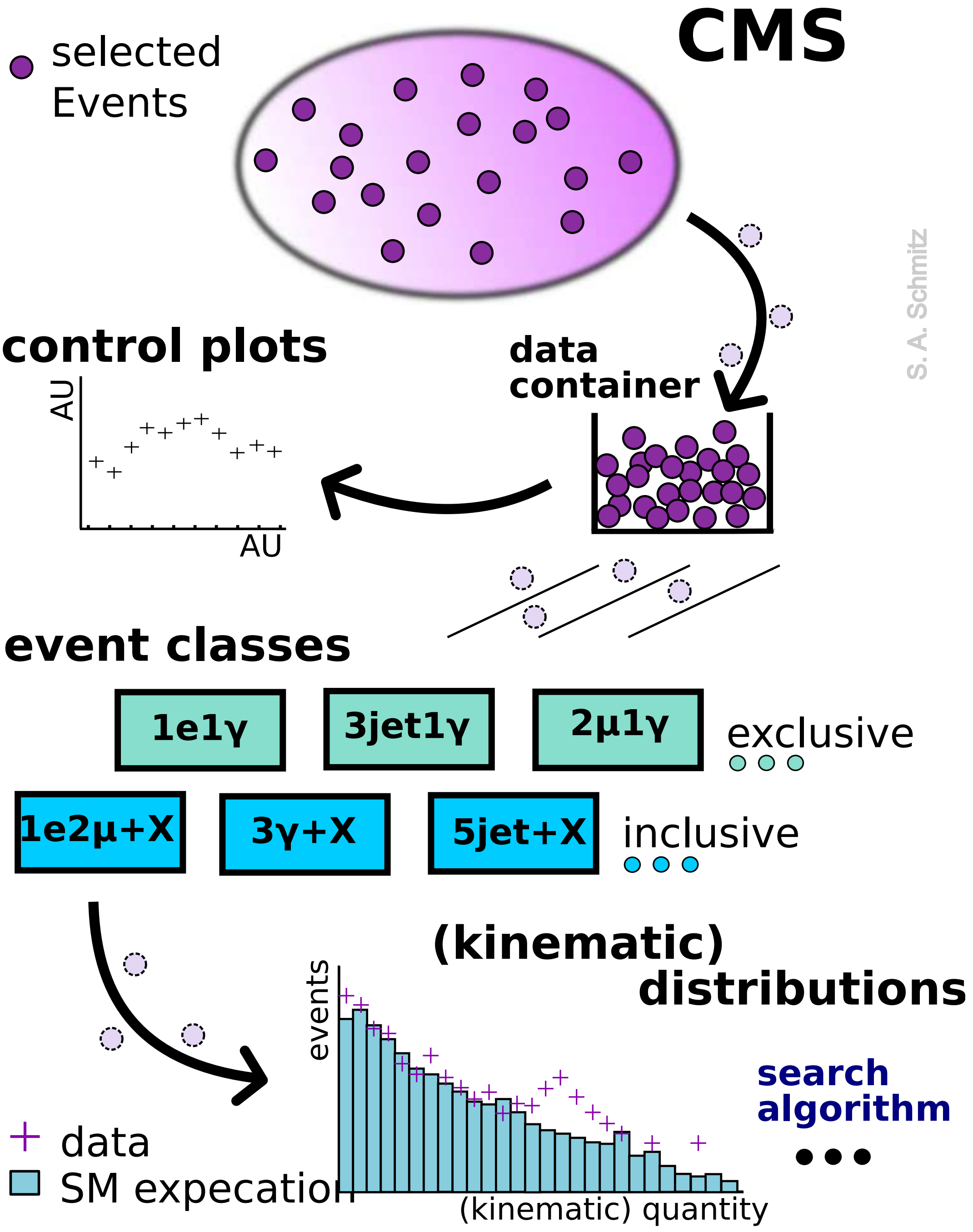

Figure 4.2: Workflow of the MUSiC Algorithm. MC simulation and data from the CMS detector are skimmed and stored in event classes depending on the physical objects contained in the final state. For each class, three distributions, $\sum p_T$, MET and M_{inv}^T are systematically scanned for deviations. Optionally, certain control plots can be viewed by the analyzer in order to inspect the analysis chain. [33]

Inclusive classes are expected to be helpful since one does not expect new physics to appear in one predetermined final state but in a number of final states. In fact, it is probable that new phenomena will show up in more than one event class. Therefore, inclusive classes can act as an indicator for new physics and are normally more significant. However, they are more difficult to interpret due to the large overlaps between classes. Since the present work concentrates on introducing a new object to MUSiC, inclusive classes will not be discussed here. Nevertheless, the inclusive classes remain an important part of the MUSiC framework.

The classified events are triggered by HLT Triggers of CMS. Since MUSiC normally uses unprescaled single muon and unprescaled single electron triggers, a special trigger interpretation routine is used. Hereby, several triggers are declared in the configuration file of MUSiC. An event is accepted for further analysis if one of the required triggers has fired. Hereby, an additional procedure makes sure that no event is counted twice. A further explanation is given in section 4.4.

Furthermore during the classification, the Monte Carlo samples are rescaled to the integrated luminosity of the data using the integrated form of equation 3.3

$$N_{\text{total}} = \sigma \int \mathscr{L} \mathrm{d}t. \tag{4.1}$$

with the respective theoretical cross section σ. Correspondingly, the total number of Monte Carlo events in a sample $N_{\text{total}}^{\text{MC}}$ has to be rescaled by a scale factor f where

$$f = \frac{N_{\text{total}}^{\text{MC}}}{N_{\text{total}}}. \tag{4.2}$$

The larger the factor $1/f$ (especially for $1/f \gg 1$), the poorer is the statistical quality of the Monte Carlo sample since all statistical fluctuations will be scaled up and appear larger in the final distribution. Especially for QCD samples, as one will see later, the scale factor $1/f$ can reach values up to 100 making the final distribution "spiky" - very discontinuous due to small sample size.

During the classification, control plots are produced. They serve as an additional check and aid for the end user. This way, the whole analysis chain remains manageable.

In the next step, all the contributions from the Monte Carlo simulations and the CMS data for all event classes are each merged into one root file consisting of all event classes. A specially designed scanner then performs a general statistical analysis of the

1. $\sum p_T$
2. MET
3. M_{inv}

distributions. In the first distribution, the sum of the transverse momenta of all physical objects which passed the selection cuts and appear in the event class are filled into a histogram.

In the second case, the missing transverse energy (in case there is any) of the corresponding event class is filled into the distribution. In the third distribution, the invariant mass of the objects of the event class is calculated and filled into a histogram. If there is missing transverse energy is the event class, only the transverse invariant mass is calculated.

Those three distributions are the most general distributions yet accounting for the most suggested new physics scenarios. New particles like W′ should appear as some kind of excess in the $\sum p_T$ and M_{inv} distributions while not interacting particles, as such, for example, predicted by supersymmetrical theories could manifest themselves in the MET distribution.

4.2.1 Scanning algorithm in MUSiC – the *p*-Value

The scanning algorithm of MUSiC systematically runs over the event classes calculating the significance of the deviations between the data and the Monte Carlo – the p-value:

The p-value is a number between 0 and 1 indicating a probability, assuming a certain hypothesis H is true, of obtaining a test statistic value with a more significant deviation from H as the actually measured value.

Here, the hypothesis H is the agreement of the data with the Standard Model Monte Carlo simulation. Data measurement and Monte Carlo generation both obey Poisson statistics. Therefore, the probability P to obtain n number of events in a bin region is

$$P = \frac{\mathrm{e}^{-\lambda} \lambda^n}{n!}. \tag{4.3}$$

While calculating the p-value, one has to presume a certain uncertainty distribution. In MUSiC, a Gaussian shape is assumed and the convoluted probability then is

$$P = A \int_0^\infty \mathrm{d}\lambda \, \mathrm{e}^{\frac{(\lambda-b)^2}{2\sigma^2}} \frac{\mathrm{e}^{-\lambda} \lambda^n}{n!} \tag{4.4}$$

where

- A is a normalization factor
- b is the expectation value of the event count. Using the rescaling factor f from Eq. 4.2 yields

$$b = \sum_{\mathrm{MC}(i)} \frac{N_{\mathrm{SM},i}}{f_i} \tag{4.5}$$

- σ is the combined uncertainty of the Monte Carlo samples

$$\sigma = \sqrt{\sum \sigma^2_{\text{syst.}} + \sum \sigma^2_{\text{stat,MC}}}. \tag{4.6}$$

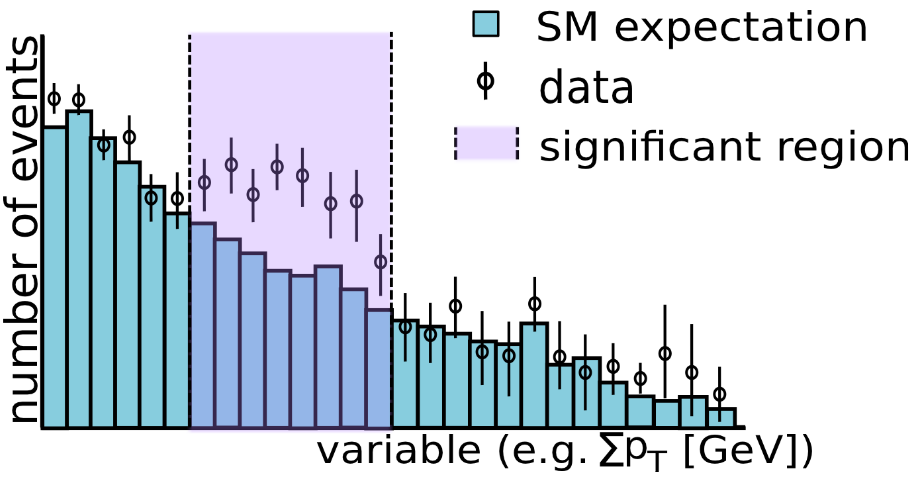

Figure 4.3: The Region of Interest in MUSiC (here) marked in light purple. The p-value in this region is the smallest for all possible connected bin combinations. Furthermore, MUSiC is also sensitive to deficits [35].

The p-value is then the sum of all probabilities P over all event numbers N in the region

$$p = \sum_{i=N_{\text{data}}}^{\infty} \int_0^{\infty} \mathrm{d}\lambda\, A \cdot \mathrm{e}^{\frac{(\lambda-b)^2}{2\sigma^2}} \frac{\mathrm{e}^{-\lambda}\lambda^i}{i!} \tag{4.7}$$

in case $b \leq N_{\text{data}}$ and

$$p = \sum_{i=0}^{N_{\text{data}}} \int_0^{\infty} \mathrm{d}\lambda\, A \cdot \mathrm{e}^{\frac{(\lambda-b)^2}{2\sigma^2}} \frac{\mathrm{e}^{-\lambda}\lambda^i}{i!} \tag{4.8}$$

in case $b > N_{\text{data}}$. The p-value is calculated for all connected regions of the distribution. The region with the smallest p-value is then named the Region of Interest (RoI) and is marked accordingly (cf. Fig. 4.3).

A p-value calculated this way is a good estimator for evaluating the significance of a distribution. However, this p-value has to be corrected for the so called "Look Elsewhere Effect", which will be presented in the next paragraph.

4.2.2 Look Elsewhere Effect and $\tilde{p}$

The p-value as it is described so far is a good estimate for a local significance of a distribution. Since MUSiC is evaluating many regions in many distributions, a significance of a single distribution has to be corrected for the "Look Elsewhere Effect" (LEE). This effect arises from the fact that, if searching for deviations in a large number of distributions, a particular fraction

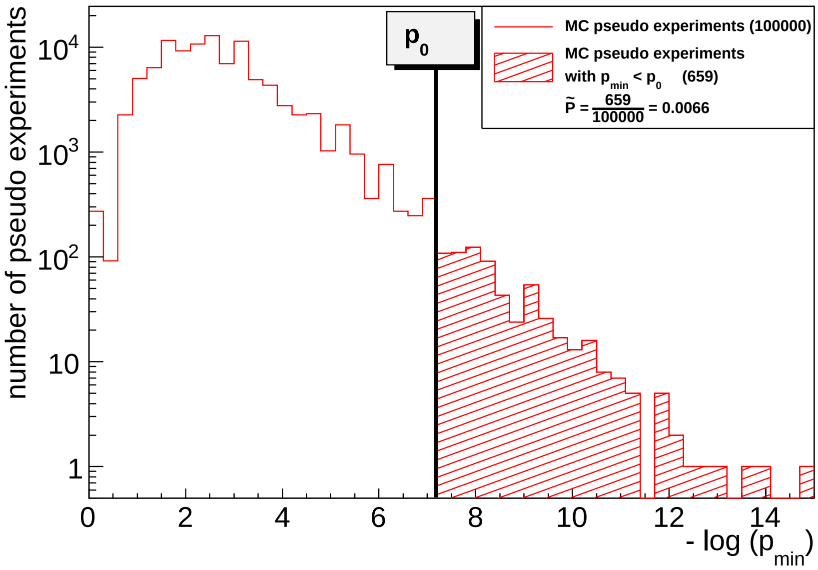

Figure 4.4: Visualization of the $\tilde{p}$ calculation. On the x-axis, the negative logarithm of p-values is displayed. On the y-axis, one finds the number of the dicing experiments corresponding to the respective p-value.

of the studied distributions is expected to deviate due to the statistical fluctuations *only*[2].

Thus, when dealing with a large number of statistical distributions, one has to take the Look Elsewhere Effect in account in order to provide a correct understanding of the deviations found. In order to take this into account, the $\tilde{p}$ is computed as follows:

1. The p-value as it is introduced in the section before is calculated and saved as a reference p_0
2. (a) Every bin in the distribution is diced within its uncertainties
   (b) The p-value is computed and compared to the reference value p_0. A p-value from a dicing experiment smaller than p_0 is considered a "hit".

---

[2] **An example for the Look Elsewhere Effect**

When calculating the chances to win the German lottery, one finds the probability to guess 6 numbers out of possible 49 and miss the "Superzahl" to be

$$\frac{1}{\binom{49}{6}} \cdot \frac{9}{10} = \frac{1}{15{,}537{,}573} \tag{4.9}$$

which is a rather poor chance for a participant. However, assuming that a large number (≈ 50 Mio.) of people participate and that their guesses are not correlated, one finds a much greater probability for *at least* one person to win – the expectation number of winners is larger than three. In fact, it seems yet rather unlikely that nobody wins.

The steps 2(a) and 2(b) are repeated for a substantial number of times N, which can be chosen by the user, in order to achieve a good statistical quality. The $\tilde{p}$ value is then defined as the ratio of the dicing experiments with a lower p-value than the reference p_0 (the hits) to the total number of dicing experiments. In case no hits were observed, only the upper threshold of $\tilde{p}$ defined by

$$\tilde{p} < \frac{1}{N}, \tag{4.10}$$

where N is the number of conducted dicing experiments is known. A visualization of this procedure is represented in Figure 4.4. In this figure, p_0 is the reference p-value. The red labeled area represents the number of dicing experiments with p-values smaller than p_0 (in this case 659). The resulting $\tilde{p}$ is then the ratio of the number of hits to the total number of dicing experiments $\frac{659}{100000} = 0.0066$. From experience, it is found that values of p_0 and $\tilde{p}$ can differ by up to approximately five or six orders of magnitude. However, since the determination of $\tilde{p}$ is a very computationally intensive task, $\tilde{p}$ is not calculated if p_0 has a very low value (the threshold can be chosen by the user)[3]. It is furthermore important to mention that MUSiC is *not* an "automated discovery tool". Every deviation found by MUSiC has to be examined closely and serves as a hint both for correct modelling of the Standard Model and the search for new physics. In no case, the $\tilde{p}$ value can be interpreted as a "discovery" display. It is very important to understand the nature of the deviation and examine the shape of the occurring discrepancies between the measurement and the Standard Model prediction.

Further studies have been conducted in order to find an alternative statistical method for evaluating deviations in MUSiC [32]. These methods will, however, not be used in the present work.

4.2.3 Bin width

MUSiC uses variable bin widths. Earlier studies have shown that in order to achieve the best possible sensitivity a bin size should be chosen to be as large as possible while not exceeding the resolution. The bin width is chosen to be an integer multiple of 10 GeV. The number of events is normalized to the bin width in order to obtain a smooth distribution. The resolution for the $\sum p_T$ distribution is estimated by

$$\sigma_{\sum p_T} \approx \sqrt{\sum_i N_i \cdot \sigma_i^2(p'_T)} \tag{4.11}$$

where N_i is the number ob objects of one type and σ_i is the corresponding resolution. For all three distributions, $\sum p_T$, invariant mass and $\not{E}_T$, the resolution is assumed to be equal.

---

[3] Thus, if $p_0 = 10^{-40}$, which can happen during Monte Carlo studies, and the number of pseudo-experiments to be conducted is, for instance, $N = 10^6$, it is very improbable that a "hit" will be diced. The computational effort, on the other hand, would be eminent. This way, needless dicing is prevented.

Contribution	Value	Remarks
MC statistics	various	sample dependent
Luminosity	2.2%	
PDF	various	PDF reweighting method
jet energy scale	3-5%	p_T, η dependent
reconstruction efficiencies	1-4%	object dependent
misreconstruction probabilities	30-100%	object dependent
W-boson cross sec.	5%	NNLO
DY cross sec.	5%	NNLO
$t\bar{t}$ cross sec.	10%	NNLL
Multi-boson cross sec.	10%	LO/NLO
Υ cross sec.	30%	measured
QCD-multijet cross sec.	50%	LO
photon+jets cross sec.	50%	LO

Table 4.1: Various relative systematic uncertainties taken into account by the MUSiC Framework.

4.2.4 Uncertainties

A number of uncertainties are taken into account by the MUSiC Framework:

- Uncertainty of the integrated luminosity measurement of 2.2% [36]. This uncertainty is the same for all bins in a distribution and therefore all bins are scaled correspondingly.
- The uncertainties of the total cross section of Monte Carlo samples which can vary depending on how good the theory prediction for the corresponding process is. Being only calculated to leading order, QCD Monte Carlo samples are assigned a large uncertainty of up to 50%. The list of the uncertainties can be found in Tab. 4.1.
- Jet energy scale (JES) describing the difference between the measured energy and the actual energy of a jet. In order to evaluate the Jet Energy scale uncertainty, the jet energies are rescaled up and down and the classification is performed again since bin content or even a content of event classes can change due to the energy variation and some particles may drop from the selection.
- The uncertainty of misidentification probability. Due to many effects, a particle type might be reconstructed incorrectly (e.g. a photon as a jet). The probability for a wrong reconstruction is colloquially called the "fake rate". The fake rate is calculated using the Monte Carlo samples during the classification by comparing the generated particle content with the reconstructed. This probability, however, has a further systematic uncertainty which is taken into account by MUSiC. Values of the fake rate uncertainty

can vary up to 100% and are provided by the working groups of CMS responsible for the reconstruction algorithms (**P**article **O**bject **G**roups).

- Monte Carlo statistics uncertainty. Obeying the Poisson statistics, as stated earlier, the statistical uncertainty of the Monte Carlo is taken into account with the standard deviation $\frac{\sqrt{N_{\text{SM}}}}{f}$ and the scaling factor f (cf. Eq. 4.5).
- Parton distribution functions (PDF) which describe the longitudinal momentum distribution of partons also resulting in an influence on the cross section. There are procedures for determining the uncertainty of the PDFs [37][38]. In MUSiC, the reweighting method is used calculating a weight w basing on a collision of two protons:

  $$w^j = \frac{\text{PDF}^j(x_1, f_1, Q)}{\text{PDF}^0(x_1, f_1, Q)} \cdot \frac{\text{PDF}^j(x_2, f_2, Q)}{\text{PDF}^0(x_2, f_2, Q)} \tag{4.12}$$

  where x_i is the fraction of longitudinal momentum carried by the parton, f_i its flavor and Q the factorization scale. The bin content X_j is weighted by w^j for each j. The uncertainty then reads

  $$\Delta X^+ = \sqrt{\sum_{i=1}^{40} [\max(X_i - X_0, 0)]^2} \tag{4.13}$$

  $$\Delta X^- = \sqrt{\sum_{i=1}^{40} [\max(X_0 - X_i, 0)]^2} \tag{4.14}$$

  for the higher and lower bin content respectively. MUSiC uses the "CTEQ6.1" parton distribution function sets [39]. Here, 40 is the number of eigenvector sets used in the parton distribution functions. The greater value of $\Delta X^{\pm}$ is eventually taken as the PDF uncertainty for the corresponding bin.
- MUSiC uses a fill up procedure. Hereby, an uncertainty on a certain MC sample with its last contribution in a bin x is filled up to a bin number $2 \cdot x$. Hereby, the assumed uncertainty for N bins is

  $$\sigma = \frac{1}{\sqrt{N f}} \tag{4.15}$$

  where f is the scaling factor (cf. Eq.4.2). This way, a Monte Carlo sample with its last contribution in bin x will produce an additional uncertainty σ from the latter equation for all following bins up to the bin $2 \cdot x$. So, insufficient statistics of Monte Carlo samples, especially for high energies, are taken into account. Also, this approach is general and identic for all event classes.

The effect of the luminosity uncertainty is correlated over all bins in all event classes since all bin are rescaled with the corresponding scale factor. The uncertainties of the different

Mass m_τ	1.778 GeV
Life time τ_τ	291 fs
Distance	$c \cdot \tau_\tau = 87\,\mu$m

Decay channel	Branching Ratio	
$\tau \rightarrow \mu\nu_\mu\nu_\tau$	$(17.36 \pm 0.05)\%$	} Not fully reconstructable
$\tau \rightarrow e\nu_e\nu_\tau$	$(17.85 \pm 0.05)\%$	
$\tau \rightarrow$ Hadrons	$\approx 65\%$	HPS Algorithm

Table 4.2: The properties (top) and the branching ratios (bottom) of the τ [5]. The HPS algorithm for reconstructing hadronic decays of the τ is described in section 4.3.2

Monte Carlo samples are assumed to be uncorrelated. However, the uncertainty of each Monte Carlo process is taken as correlated in all bins and all event classes. The uncertainty of the bin is obtained by adding in quadrature the individual uncertainties of the contributing samples in the corresponding bin. It is furthermore assumed that the PDF uncertainty is fully correlated for all bins in all event classes. In analogy, the Jet Energy Scale corrections are assumed to deviate in one direction distorting the whole distribution at once.

4.3 Introduction of τ Leptons into the MUSiC framework

To this moment, only electrons e, muons μ, photons γ, missing transverse energy MET and jets were used within the MUSiC Framework. Introduction of the τ lepton is a central topic of this thesis. While searching for new physics beyond the Standard Model, τ is an important ingredient for a physics analysis since a number of theories make prognoses about a number of exotic decay modes including the τ (cf. section 2.2).

4.3.1 τ Lepton

The τ lepton is the third and the heaviest member of the lepton family [5]. The Feynman graphs for τ production are displayed in Fig. 4.5. The physical characteristics of the τ are summarized in table 4.2. The relatively long mean life time of τ allows for a travel distance away from the interacting point which is big enough to be recognized as a secondary vertex within the resolution of the tracker. However, in most cases, no τ-tagging is performed by the reconstructing algorithm (cf. paragraph 4.3.2). In about 1/3 of all cases, the τ decays leptonically to an electron or a muon. In case the τ originates from a $W^\pm$, the decay into leptons will produce an additional neutrino (cf. Fig. 4.6). Two neutrinos in one event do not allow for a kinematically unique reconstruction of both neutrinos. Also in case of a Z decay,

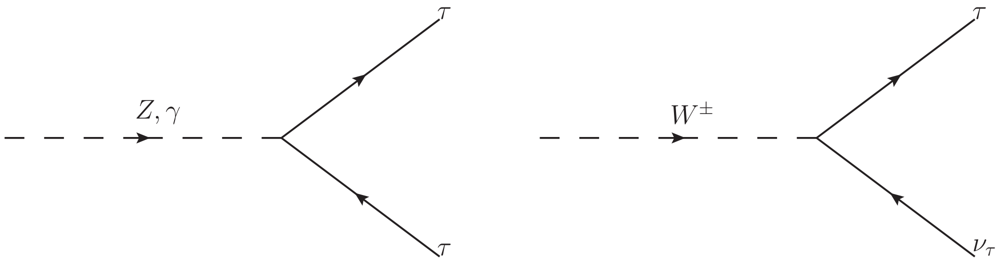

Figure 4.5: Feynman graphs for τ production. On the left, from the Z or γ and on the right from the $W^\pm$.

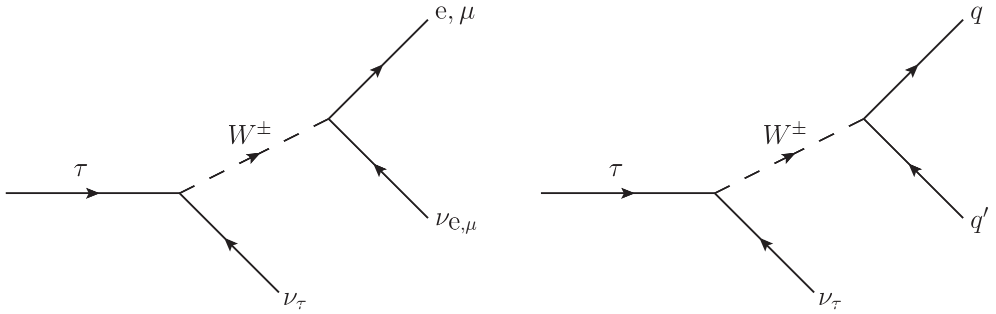

Figure 4.6: Feynman graphs for τ decay. On the left, for the leptonic decay and on the right for the hadronic decay. In case of hadronic decay, several decay scenarios are possible.

each tau will produce a neutrino resulting in the same problem. The hadronic decay of the τ can produce either one (so called One-Prong decay) or three (Three-Prong decay) charged hadrons. Hereby, one has also to keep in mind that the cross section for $W^\pm$ production at the LHC is by factor ten larger than the production cross section of Z [40] and therefore most τ leptons will come from $W^\pm$.

4.3.2 Reconstruction of the τ – HPS Algorithm

A well-established algorithm for reconstructing hadronic τ decays τ_{had} in CMS is the **H**adron **P**lus **S**trips algorithm [41]. The algorithm starts with a Particle Flow jet (PFJet) [16][42] and proceeds as follows (cf. Fig. 4.7)

1. Within a PFJet, the strip is associated with the most energetic charged hadron.
2. The algorithm searches for other electromagnetic particles within the window of $\Delta\eta = 0.05$ and $\Delta\phi = 0.2$ around the strip center. In case such are found, they are associated with the strip and its momentum is updated.
3. The procedure is repeated until no charged hadrons or their decay products are found.

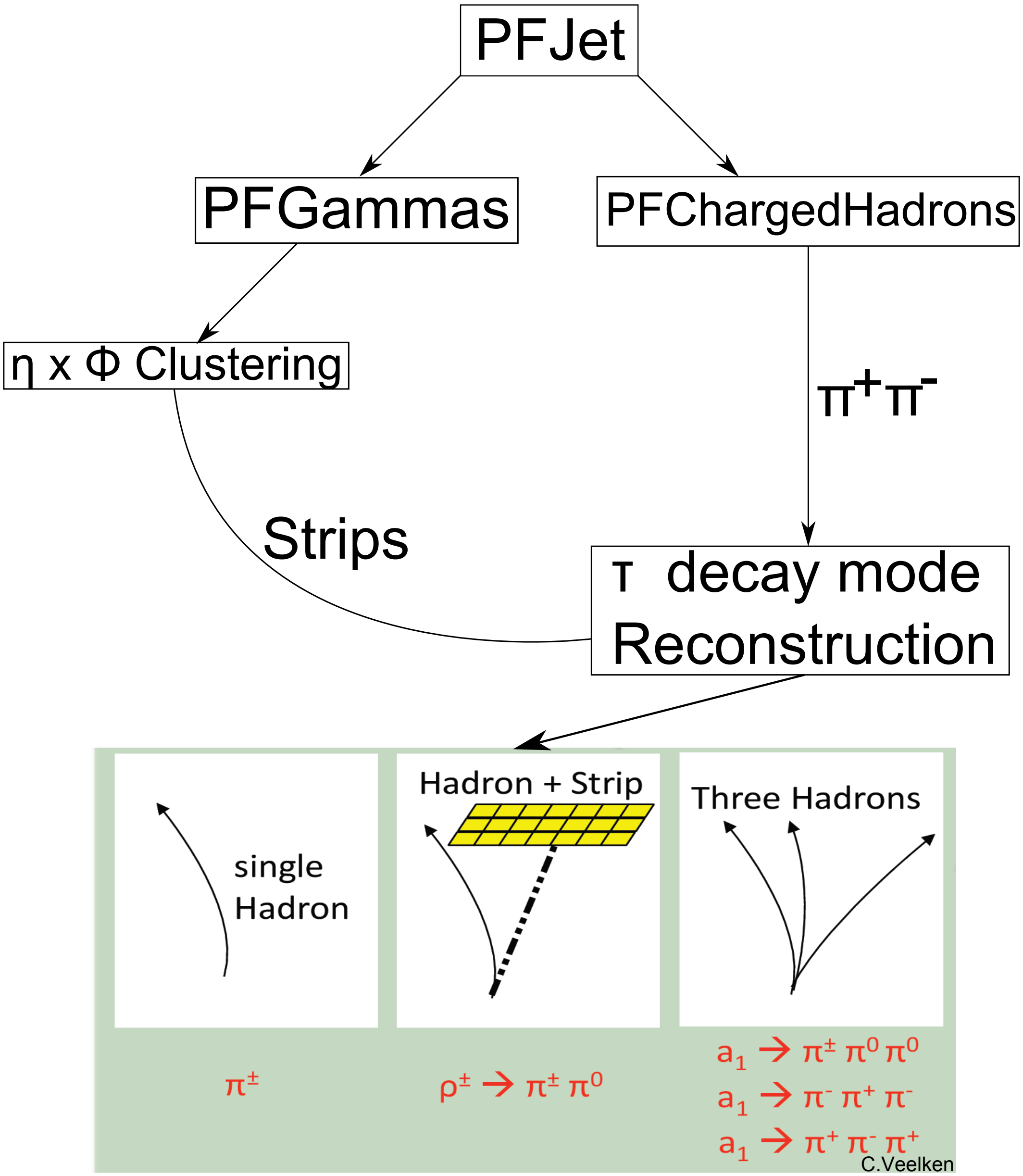

Figure 4.7: A sketch of the Hadron Plus Strips algorithm working procedure . At the bottom, the possible decay modes of hadronic τ decays are visualized. The corresponding decay modes can be found in Table 4.3 [43]

Decay mode	Resonance	Mass(MeV/c^2)	Branching fraction(%)
$\tau^- \to h^- \nu_\tau$			11.6%
$\tau^- \to h^- \pi^0 \nu_\tau$	ρ^-	770	26.0%
$\tau^- \to h^- \pi^0 \pi^0 \nu_\tau$	a_1^-	1200	9.5%
$\tau^- \to h^- h^+ h^- \nu_\tau$	a_1^-	1200	9.8%
$\tau^- \to h^- h^+ h^- \pi^0 \nu_\tau$			4.8%

Table 4.3: Branching fractions of the most common hadronic decays of the τ lepton and the resonances (without loss of generality for τ^-). Hereby, h can be either π or K while the π mass is assumed for all charged particles [41].

4. Strips with $p_T > 1\,\text{GeV}$ are eventually combined with the charged hadrons.

The following decay scenarios are hereby considered by the HPS algorithm:

- A single hadron decay corresponding to a charged hadron h^-, a neutrino ν_τ and a possible neutral hadron (π^0) with not enough energy to pass the strip reconstruction criteria. In this case, the π^0 is not visible.
- One hadron and one strip corresponding to a $h^- \nu_\tau \pi^0$ decay where the photons from the decay of the π [5] are close together.
- One hadron and two strips corresponding to the same decay with two well separated photons.
- Three hadron decay corresponding to the decay mode $h^- h^+ h^- \nu_\tau$ where all three hadrons have to be reconstructed from the same secondary vertex.

The HPS algorithm requires all charged hadrons and strips to be within a cone of size $\Delta R = 2.8\,\text{GeV}/p_T^{\tau_\text{had}}$ where $p_T^{\tau_\text{had}}$ is the transverse momentum of the hadronic τ decay. Furthermore, the τ momentum has to be within a $\Delta R = 0.1$ distance from the reconstructed PFJet. During the reconstruction of momentum, all charged hadrons are assumed to be pions and required to be compatible with the τ decay resonances listed in Table 4.3. Here, π^0 are allowed to be within a mass range of 50–200 MeV, ρ within 0.3–1.3 GeV and a_1 within 0.8–1.5 GeV.

In addition to the mentioned criteria, there is an isolation requirement for the τ_had candidates. Within an isolation cone of $\Delta R = 0.5$, the HPS algorithm vetoes on charged hadrons or photons not being part of the τ_had decay. Based on three thresholds for their p_T, working points "loose", "medium " and "tight" are defined. The expected efficiency for those working points is presented in Fig. 4.8. For determining the efficiency, the HPS algorithm is applied to events coming from inclusive production of $Z \to \tau\tau$. The efficiency ϵ is then defined by

$$\epsilon = \frac{N_\text{pass}}{N_\text{pass} + N_\text{fail}} \tag{4.16}$$

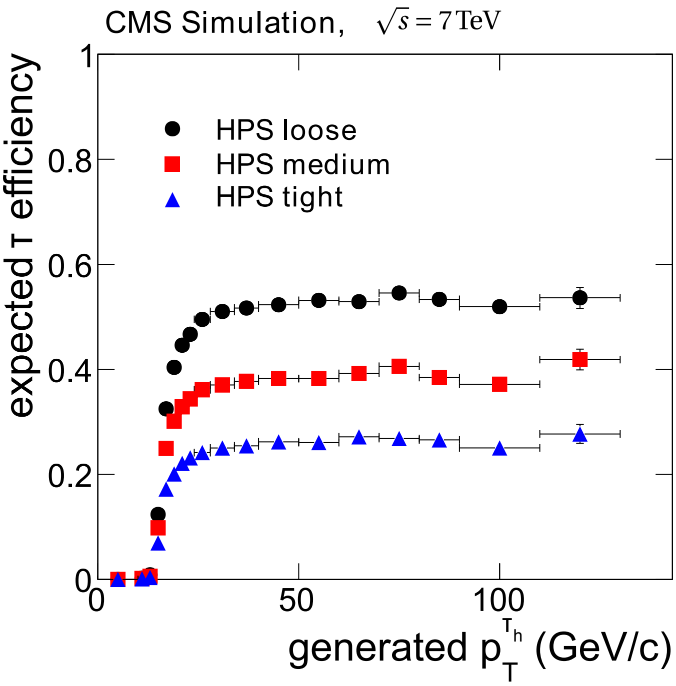

Figure 4.8: Expected efficiency of different working points of the HPS algorithm as function of p_T of τ_{had}. Low efficiency in very low p_T ranges can be explained by a small $\gamma = \frac{E}{m}$ parameter of low p_T τ – the τ-jet is then not boosted enough and too broad to be reconstructed by the HPS algorithm [41].

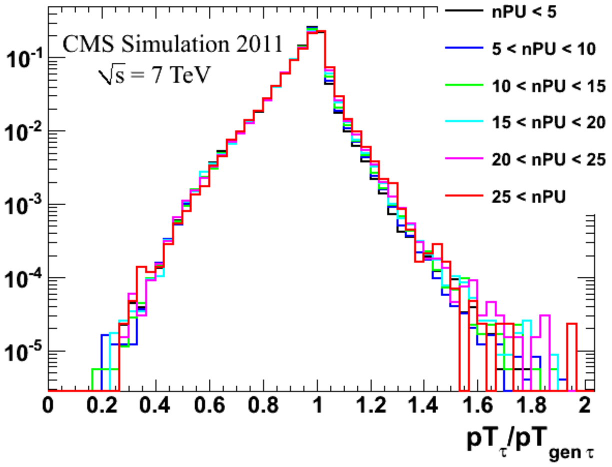

Figure 4.9: Resolution of 1-prong $+1\pi^0$ τ leptons coming from $Z \to \tau^+\tau^-$ for different pile-up scenarios – various number of soft interactions. The τ energy scale is almost not affected by pile-up [44].

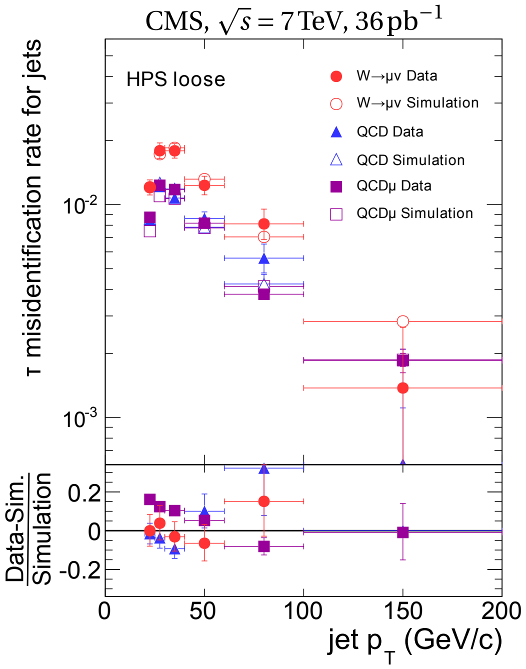

Figure 4.10: Misidentification rate for jets to be identified as τ_{had} by the HPS algorithm as function of jet p_T. Starting around 2% at $p_T \approx 30\,\text{GeV}$, the fake rate falls down to few times $10^{-3}$for higher transverse momenta [41].

where $N_{\text{pass,fail}}$ are the numbers of $Z \to \tau\tau$ events. This way, only hadronic decays of τ leptons are considered. The HPS algorithm generates boolean discriminators indicating how well the isolation (cf. Table . 6.1) of the τ_{had} is and whether the decay mode requirement is met [45]. The reconstructed energy of τ_{had} is found to be within 2% systematic uncertainty to the true visible energy of τ_{had} for the HPS algorithm [41]. However, the η dependence of the reconstructed τ_{had} energy is underestimated by up to 5% with increasing η. The fake rate of the HPS algorithm is fairly low and varies around 1% for HPS "tight" and "medium" depending on the p_T of τ_{had} 4.10. All in all, the τ physics object group specifies the resolution of the τ reconstruction with $10\% \cdot \tau_{p_T}$ (cf. Fig. 4.9). The misidentification rate of the "tight" working point varies around 1‰. The difference between the three working points "loose", "medium" and "tight" is presented in Fig. 4.11. Here, it is visible that the misidentification rate can vary within almost an order of magnitude, depending on the working point. Since the jets are very frequent at LHC, reducing the QCD background is a crucial task for any τ-analysis [46]. Furthermore, the choice of the working point is an important task for a physics analysis in general and especially relevant for model independent searches since they use the most generalized approach.

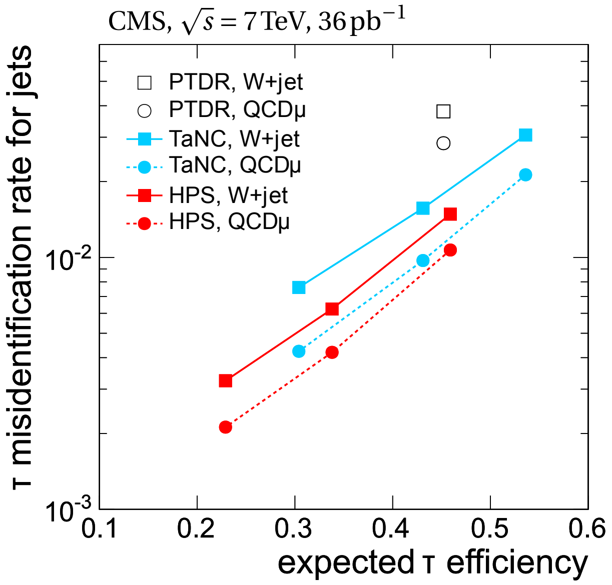

Figure 4.11: Misidentification rate for jets to be identified as τ_{had} for different working points of the HPS algorithm. The HPS algorithm delivers the lowest fake rate compared to older algorithms used by CMS. The left bottom working point of the HPS algorithm is "tight", the middle is "medium" and the top right is "loose" (cf. Fig. 4.8) [41].

4.3.3 Implementation of τ in MUSiC

Introducing τ to MUSiC starts in the skimming procedure. Originating from a jet τ_{had} has to be reconstructed and saved in the corresponding collection within the data container of CMSSW. The energy, momentum and the discriminators (cf. Tab.6.1) produced by the HPS algorithm have to be read out by the skimmer and saved in PXL containers in order to be available during the classification. Correspondingly, the collection of jets is corrected for the extracted τ_{had} jets.

In the next step, τ is introduced as an object to the classification procedure resulting in classes containing taus, e.g. $1\text{mu} + 1\tau + 1\text{MET}$. The rest of the procedure is performed by the MUSiC algorithm automatically. One special feature which has to be implemented into MUSiC is the cross-triggering. It is presented in the following section.

4.4 Introduction of a new triggering system into the MUSiC framework

During the classification procedure of MUSiC, the data selection is triggered by HLT triggers of CMS. So far, a "true" entry by the corresponding trigger was the signal MUSiC used in order to analyze the event. MUSiC then performed a further "trigger" cut since the trigger efficiency can vary at the trigger threshold. A possibility to use more than one "trigger" cut per one HLT trigger did not exist. In fact, MUSiC did not link the trigger signal to a certain physics object which caused the trigger to fire – the trigger signal was only the signal to perform the selection on the event. Since during the 2011 data taking CMS only offered a cross-trigger triggering on τ **and** missing transverse energy MET at the same time, an update on the triggering procedure of MUSiC was performed as a part of the present thesis.

In the course of this update, a linking procedure between the trigger particles and the HLT trigger was implemented into MUSiC. In contrast to the previous situation, now the trigger signal of a certain trigger is linked to a flexible number of cuts on each of six physical objects used by MUSiC. This way, the signal from the HLT trigger is directly connected to the cuts which the user wants to apply to the triggering object. The most important advantage of this approach is the new feature of using more than one physical object for triggering events. Any conceivable combination of triggering particles can now be used in MUSiC, e.g. double-μ or double-$\gamma\gamma$ triggers as prominent examples for Higgs analyses or a τ-μ trigger for the search for new physics. One further advantage of this approach is, among others, a better usability – the triggering cut is set in the same configuration file as the trigger and the chance of a user mistake is reduced.

CHAPTER 5

Data Selection and Analysis

As already stated before, a sophisticated data selection is crucial for a physics analysis. Depending on the goal of a particular analysis, the cuts applied to data and MC can vary a lot. In MUSiC, it is tried to implement a general but yet efficient data selection. The data selection and the results obtained after introducing the τ lepton to the MUSiC framework are presented in this chapter.

5.1 Monte Carlo Samples

MUSiC uses a many different of Monte Carlo samples. The most important contributions are made by the following Monte Carlo samples:

- Drell-Yan $q\bar{q} \rightarrow \mathrm{e}^+\mathrm{e}^-, \mu^+\mu^-, \tau^+\tau^-$
- W-boson - $W + \mathrm{Jets} \rightarrow$Leptons
- Di-boson: WW, WZ, ZZ, Wγ, Zγ, $\gamma\gamma$
- $t\bar{t}$ and single top
- QCD-multijet: inclusive & EM-, b-, Mu-enriched samples
- Upsilon :$\Upsilon(1\mathrm{S}), \Upsilon(2\mathrm{S}), \Upsilon(3\mathrm{S}) \rightarrow \mu\mu, \mathrm{ee}$

Note that the $W + \mathrm{Jets}$ contains the W boson production in presence of up to four jets. The Drell-Yan sample contains the Drell-Yan process in presence of up to four jets. The QCD EM-, b- and Mu-enriched samples are produced by demanding a corresponding object (electrons, b-quarks or muons) directly on generator level. This is done because MUSiC does not trigger on jets at the moment. The cross section is corrected correspondingly. The samples are generated by MADGRAPH [47], Powheg [48][49][50] and Pythia [51]. Before the implementation of the τ lepton, all events containing at least an electron or a muon were

Name of the trigger	Trigger cut in GeV
HLT_Ele27_CaloIdVT_CaloIsoT_TrkIdT_TrkIsoT_vX	82
HLT_Ele42_CaloIdVT_CaloIsoT_TrkIdT_TrkIsoT_vX	
HLT_Ele52_CaloIdVT_TrkIsoT_vX	
HLT_Ele65_CaloIdVT_TrkIsoT_vX	
HLT_Ele80_CaloIdVT_TrkIsoT_vX	
HLT_IsoMu12_vX	32
HLT_IsoMu17_vX	
HLT_IsoMu20_eta2p1_vX	
HLT_IsoMu24_vX	
HLT_IsoMu30_eta2p1_vX	
HLT_IsoPFTau35_Trk20_MET45_vX	$\tau_{p_T} > 40$
HLT_MediumIsoPFTau35_Trk20_MET60_vX	$\mathrm{MET}_{p_T} > 90$

Table 5.1: Triggers used for the analysis of 2011 CMS data. Hereby, the p_T cut of the trigger is normally part of the name – HLT_IsoMu12_v1 sets a cut of 12 GeV. "vX" stands for different version of the same trigger setup. Two muon triggers also have η cuts of $|\eta| < 2.1$.

analyzed by MUSiC. In this work, events triggered by electrons, muons or $\tau + \mathrm{MET}$ will be presented. In order to gather the best possible statistics, MUSiC always uses the lowest unprescaled trigger. This strategy sometimes results in rather high p_T trigger cuts or even trigger cuts on pseudorapidity. Nevertheless, this approach grants the most general access to the data. The criteria set for the event selection of the present work are summarized in Table 5.2. The trigger selection is presented in Table 5.1. Furthermore, an event with more than 10 tracks is rejected if more than 25% of the tracks are badly reconstructed. Also a good primary

Object	p_T / GeV	$\lvert\eta\rvert$	other	trigger cut p_T/GeV
e	> 25	< 2.5	shape/track/isolation	82
μ	> 20	< 2.1	isolation/track	32
τ	> 20	< 2.3	shape/isolation	40[1]
γ	> 25	< 1.442	shape/isolation	–
jet (anti-k_T, $R = 0.5$)[2]	> 50	< 2.5	energy fraction	–
MET	> 50	–		90[1]

Table 5.2: Object cuts performed by MUSiC. For triggering objects, a special cut is applied.

---

[1] Cross-trigger for τ and MET

[2] Cf. section 5.2

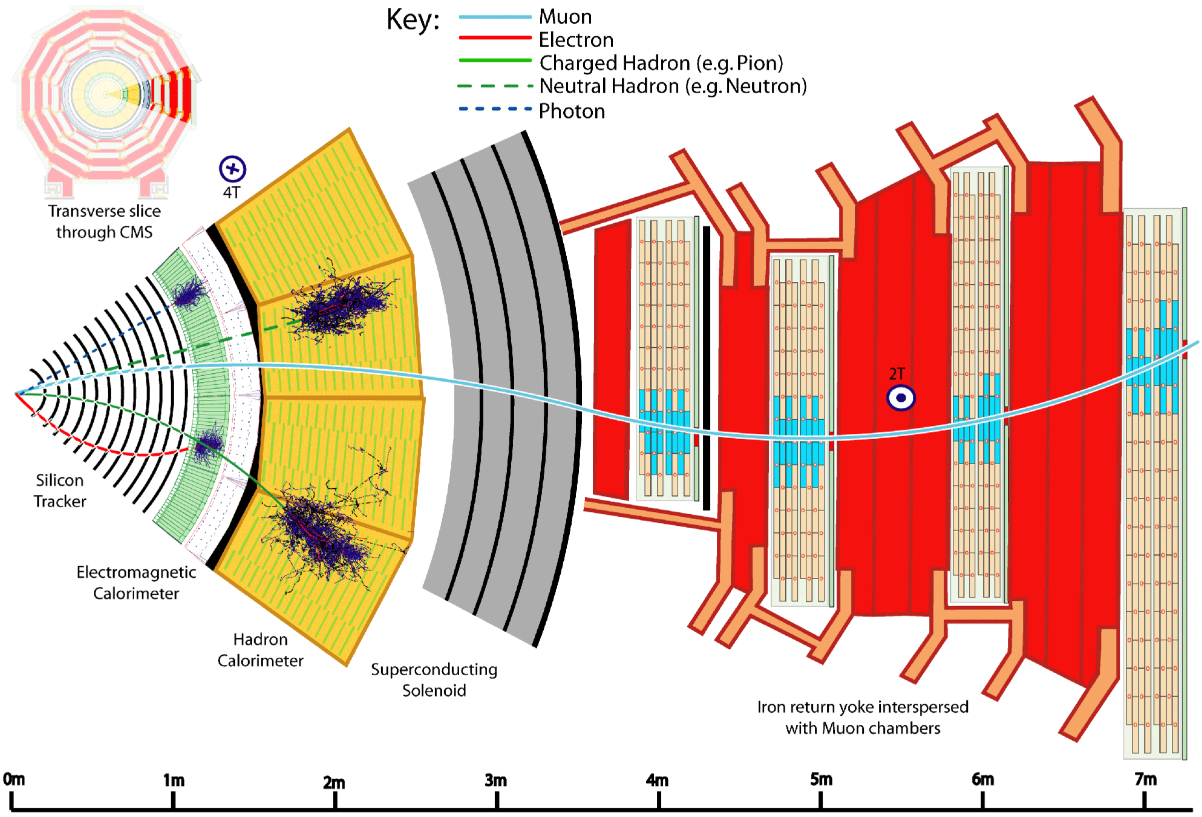

Figure 5.1: A slice of the CMS detector. Charged hadrons are identified via their interaction with the tracker and the calorimeters. The photons are only visible in the ECAL, neutral hadrons in the HCAL. Electrons are visible in the tracker and the ECAL and muons are visible in the tracker and the muon system. Neutrinos leave CMS undetected [52].

vertex has to be within 2 cm away from the nominal interaction point in the $r\phi$ plane and not more than 24 cm in z-direction. In addition, at least four tracks with a fit quality of $\frac{\chi^2}{\text{ndf}} < 10$ for each are demanded. A primary vertex is required by MUSiC. It is demanded to have at least five tracks and to be within 2 cm in the vertical direction from the beam and 24 cm in both directions along the beam from the nominal interaction point.

5.2 Object Selection

In the following, a short overview of the reconstruction of other objects used by MUSiC (electrons, muons, photons, τ, jets and missing transverse energy) is given. A short overview of the CMS detector and how the particles interact with the detector material is given in Figure 5.1.

Muons

MUSiC uses global muons meaning that the muon has to be reconstructed in the silicon tracker and in the muon system. First, the signals in the muon system (DT and CSC) are combined to segments. From those segments, a trajectory of the muon inside the muon system is reconstructed and combined into a track. If this tracks satisfies the criteria for the track reconstruction, a so called "standalone" muon object is created. The trajectories from the silicon tracker are then matched with the standalone muon in order to yield a global muon. If there is more than only one track matching the standalone muon, the track with the smallest χ^2 of the track fit is chosen for the global muon [53]. Furthermore, the following cuts are applied on the global muon:

- A transverse momentum cut $p_T > 20\,\text{GeV}$. Note that this cut only applies on all muons in the event except the leading trigger muon. The latter has to satisfy the trigger cut discussed in Table 5.2.
- Pseudorapidity cut of $|\eta| < 2.1$ corresponding to the trigger coverage of CMS.
- The distance to the primary vertex d has to satisfy $d < 0.2\,\text{cm}$ in order to reject cosmic muons and muons coming from pile-up and having a different vertex.
- Number of hits in the tracker of at least 11 in order to ensure a good reconstruction.
- At least two hits within the muon system.
- Isolation requirement: within a cone of $\Delta R = 0.3$, no particle with $p_T > 3\,\text{GeV}$ should exist. This cut reduces the number of muons being decay products of particles in the detector not belonging to the hard interaction.
- At least one hit in the pixel tracker. This cut rejects muons which are produced by other particles while travelling through the detector.

Electrons

Electrons are detected in the tracker and the ECAL where they shower and deposit their energy. While interacting with the ECAL material, the electrons will loose energy via bremsstrahlung. This effect has to be taken into account by the reconstructing algorithm since it has a great influence on the measured energy of the electron [54]. An electron of around 100 GeV will lose over 95% of its energy within 5 × 5 crystals of the ECAL. Those crystals are combined to a cluster by the reconstruction algorithm. The clusters containing the radiated photons are eventually combined with these to a supercluster. The most energetic cluster is selected as a so called "seed" cluster and the clusters around are added in case they pass certain criteria [55]. In the next step, the path of the electron inside the tracker is reconstructed. Using the information from the ECAL, the track of a possible electron or a positron is reconstructed starting in the first layers of the pixel tracker. At least 5 hits are required and a χ^2 fit is performed. An electron reconstructed this way has to satisfy a number of further conditions:

- Transverse momentum cut of $p_T > 25\,\text{GeV}$.
- An η range of $|\eta| < 1.442$ for the barrel and $1.56 < |\eta| < 2.5$ for the endcap
- $\Delta\phi < 0.06$ between the trajectory of the electron measured by the tracker and the. supercluster in the ECAL.
- $\Delta\eta < 0.005$ for for the barrel and $\Delta\eta < 0.007$ for the endcap.
- The ratio of the energy deposit in the HCAL to the energy in the ECAL $H/E < 0.05$ since a bigger energy deposit in the HCAL would indicate a jet.

Also, a number of isolation cuts is used.

Photons

Photons are reconstructed in the ECAL. The same reconstruction algorithm is used for the cluster building in the ECAL while no hits in the tracker are expected. Furthermore, in a number of cases a photon can convert to an electron-positron pair. Originally having the same direction of flight as the mother photon, the electron and the positron will be deflected in different ϕ directions due to the magnetic field. Those pairs are reconstructed by the algorithm described in section 5.2 [56]. The energy of the measured photons is corrected for the converted photons. A photon has to satisfy the following criteria:

- Transverse momentum cut of $p_T > 25\,\text{GeV}$.
- An η range of $|\eta| < 1.442$. Endcap photons are not a part of the MUSiC framework at the moment due to their poor reconstruction quality.
- The ratio of the energy deposit in the HCAL to the energy in the ECAL $H/E < 0.05$.
- No pixel seed in the tracker as this would indicate a charged particle, e.g. an electron

and a number of further isolation cuts.

Jets

Jets are reconstructed using the Particle Flow anti-k$_T$ algorithm [42]. This algorithm is collinear and infrared safe meaning that it is stable in case the energy of the jet is split into two parts and also in case a low energy particle is radiated. The anti-k$_T$ algorithm looks for two entities i and j defining the distance d_{ij} by

$$d_{ij} = \min(k_{T,i}^{-2}, k_{T,j}^{-2}) \cdot \frac{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}{R^2} \tag{5.1}$$

where $k_{T,i}$ is the transverse momentum of the entity i, η is the pseudorapidity and ϕ is the azimuth angle as defined in section 1.1. The cone parameter R used by MUSiC is $R = 0.5$. Furthermore, the distance between an entity and the beam is defined by

$$d_{i,B} = k_{T,i}^{-2}. \tag{5.2}$$

Two entities i and j with the smallest distance are combined to a single pseudojet. Their four-momenta are added. In case $d_{i,B}$ is the minimal distance, the corresponding pseudojet originating from the entity i is considered a jet and removed from the sample. The algorithm repeats the described routine until no particles are left.

Particle Flow jets are reconstructed using not only the calorimeter entries but all parts of the CMS detector. A jet energy factor is applied by MUSiC in order to correct the energy as a function of η and p_T. The correction has the order of magnitude of 10% which is an improvement compared to the HCAL jet energy measurment only which has to be corrected by up to 50%. The following jet criteria are applied by MUSiC:

- Transverse momentum cut of $p_T > 50\,\mathrm{GeV}$
- Pseudorapidity cut of $|\eta| < 2.5$
- Neutral electromagnetic and hadron fractions of < 0.99 in order to remove jets caused by ECAL and HCAL noise.
- A jet has to consist of more than 1 particle in order to suppress the misidentification of single particles as jets.

The electromagnetic and hadron fractions are determined using matched tracks to ECAL or HCAL energy depositions, respectively. The fraction of the energy contribution matched to the tracker system divided by the total jet energy yields the corresponding fraction.

Missing Transverse Energy

Building the negative vectorial sum of all momenta of particles originating from the hard interaction one obtains the Missing Transverse Energy (MET or $\not{E}_T$). This missing energy is caused

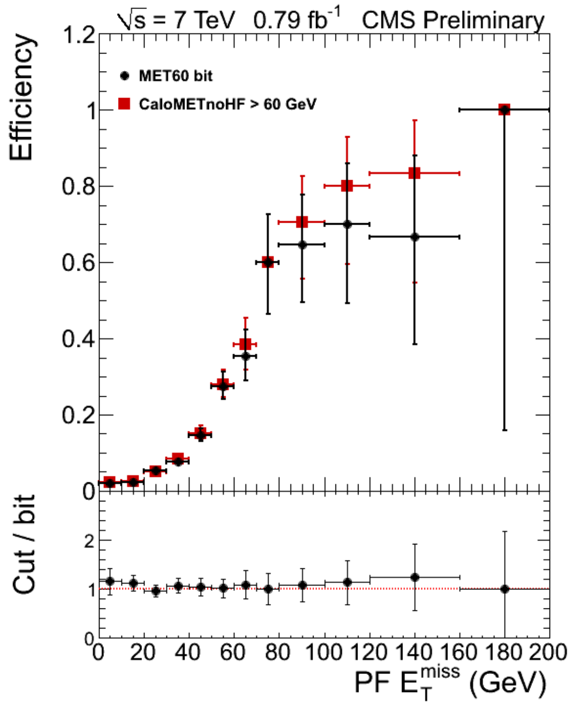

Figure 5.2: The HLT efficiency as a function of $\not{E}_T$. It is noticeable that the trigger has a long turn-on and the uncertainty of the efficiency grows with increasing values of $\not{E}_T$. The variable "Cut/bit" is the ratio between the CaloMETnoHF and MET60, two different $\not{E}_T$reconstruction algorithms [57].

by particles not accessible to detection like, for example, neutrinos. Many particles postulated by exotic physics models are also not visible in the detector. The $\not{E}_T$is reconstructed using the global energy information [16]. The following two requirements are set:

- Transverse momentum cut of $p_T > 50\,\text{GeV}$ since lower MET can be caused by the detector resolution.
- The angle difference $\Delta\phi > 0.1$ to any electron within the event in order to suppress $\not{E}_T$ caused by electrons interacting with parts of calorimeter which are not capable of measuring the energy contribution.

The offline trigger cut is set to 90 GeV as a consequence of poor MET trigger efficiency below this point (cf. 5.2).

Tau

The τ reconstruction has been described before in section 4.3.2. The selection criteria are

- Transverse momentum cut of $p_T > 20\,\text{GeV}$ due to the insufficient efficiency of the HPS algorithm at lower momenta (cf. Fig. 4.8)

- A cut of $|\eta| < 2.3$
- Decay Mode matching discriminator (cf. [58] and Table 4.3)
- byMediumCombinedIsolationDeltaBetaCorr discriminator:
  isolation cone of 0.5 , Delta Beta corrected $\sum p_T$ of PF charged and PF γ isolation candidates ($p_T > 0.5\,\text{GeV}$) less than 1 GeV
- againstElectronTight discriminator:
  electron pion MVA discriminator < -0.1 and **not** $1.4442 < |\eta| < 1.566$ and Brem pattern cuts [59]
- againstMuonTight discriminator
  τ lead track not matched to global/tracker muon and large enough energy deposit in ECAL + HCAL

Hereby, "Delta Beta" stands for the correction for multiple interaction producing neutral hadrons which are also measured by the HCAL. The Decay Mode matching discriminator makes sure that the energy of the τ decay matches a decay resonance (cf. Table 4.3) and is responsible for suppressing the QCD background. The byMediumCombinedIsolation-DeltaBetaCorr discriminator ensures the isolation of τ-jet taking the energy contribution of pile-up into account. Both againstElectronTight and againstMuonTight discriminators check whether any electron or muon tracks are associated with the τ-jet. A full description of the discriminator working points can be found in Table 6.1 in the Appendix. The fake rate uncertainty is assumed to be 30% as recommended by the Tau Physics Object Group of CMS [46].

5.3 A full scan of 2011 CMS data

In this section, the results of a full scan of 2011 CMS data corresponding to an integrated luminosity of $5\,\text{fb}^{-1}$ are presented. The total number of events observed by MUSiC in the 2011 CMS dataset is $\approx 210 \cdot 10^6$ producing all in all 681 event classes. Of these, 321 event classes contain at least one τ lepton, the focus of this thesis. As it is not possible to present all classes within the present thesis, only the most interesting classes will be discussed. Three scans have been performed, each for the $\sum p_T$, invariant mass and missing transverse energy classifications. The scanning settings are summarized in Table 5.3. The "Hits threshold" setting is used in order to avoid unnecessary dicing experiments after reaching sufficient statistical quality. Therefore, if the named number of hits (cf. section 4.2.2) is reached after a certain number of conducted pseudo-experiments, the upper limit of remaining pseudo-experiments is set to twice the number of experiments needed to reach the threshold. For instance, in case 3000 dicing experiments were needed to reach the hits threshold, the total number of dicing experiments to be conducted is set to 6000. However, this procedure does

	# of dicing experiments	# of hits threshold	Minimal bin width
$\sum p_T$	10^6	150	3
Invariant Mass	10^5	150	1
MET	10^6	150	3

Table 5.3: Scan settings for the different distributions. The scan of invariant mass distributions is computionally more intensive and therefore the number of pseudo experiments is smaller compared to $\sum p_T$ and $\not{E}_T$scans.

not set upper limits higher than the initial maximal number of dicing experiments entered by the user. This way, unneeded computations are avoided. The minimal bin width of the region of interest is set to three for $\sum p_T$ and MET distributions due to the fact that deviations are expected to be broader in those distributions. In contrast, the deviations in the invariant mass distributions can have the widths comparable to the width of one bin. Therefore, the minimal bin width is set to one for those distributions resulting in a higher computational complexity since more potential regions of interest have to be investigated by the scanning algorithm. First, a W dominated class is presented.

5.3.1 τ+MET Class

Before introducing the first results, a short overview of the standard MUSiC scanning image is given (cf. Figure 5.3). On the top left, one finds the name of the considered event class, in this case $1\tau + 1\text{met}$[3]. Next to the name of the event class, the center of mass energy $\sqrt{s}$ is shown. On the left bottom, the integrated luminosity of the shown data is presented. On the top right, one finds the p-value (cf. Eq. 4.7 and 4.8) and the $\tilde{p}$ value (cf. section 4.2.2). The plots are generally logarithmic on the y-axis and the Monte Carlo contributions are stacked up to yield the global MC prediction in the corresponding bins. The "blind stitched" areas represent the uncertainties of the Monte Carlo prediction. Beneath the p and $\tilde{p}$-values, the legend of contributing Monte Carlo samples is given. On the x-axis, the quantity of interest, $\sum p_T$, invariant mass or missing transverse energy are plotted. The Region of Interest (RoI) lies between the two vertical dashed blue lines.

The τ-MET class is an important representative for electroweak calculations and also for testing and understanding the HLT τ-MET trigger of CMS. The distribution of the transverse momenta of this class is shown in Figure 5.3. The bins $(170-180)$ GeV, $(220-230)$ GeV and $(310-310)$ GeV contain relatively large QCD contribution. As stated before, these discontinuous contributions ("spikes") are due to a big rescaling factor $1/f$ (cf. Eq. 4.2) of the

---

[3] It should be noted that an event class can not contain more than one MET since all transverse energies of undetectable particles appear as one missing transverse energy.

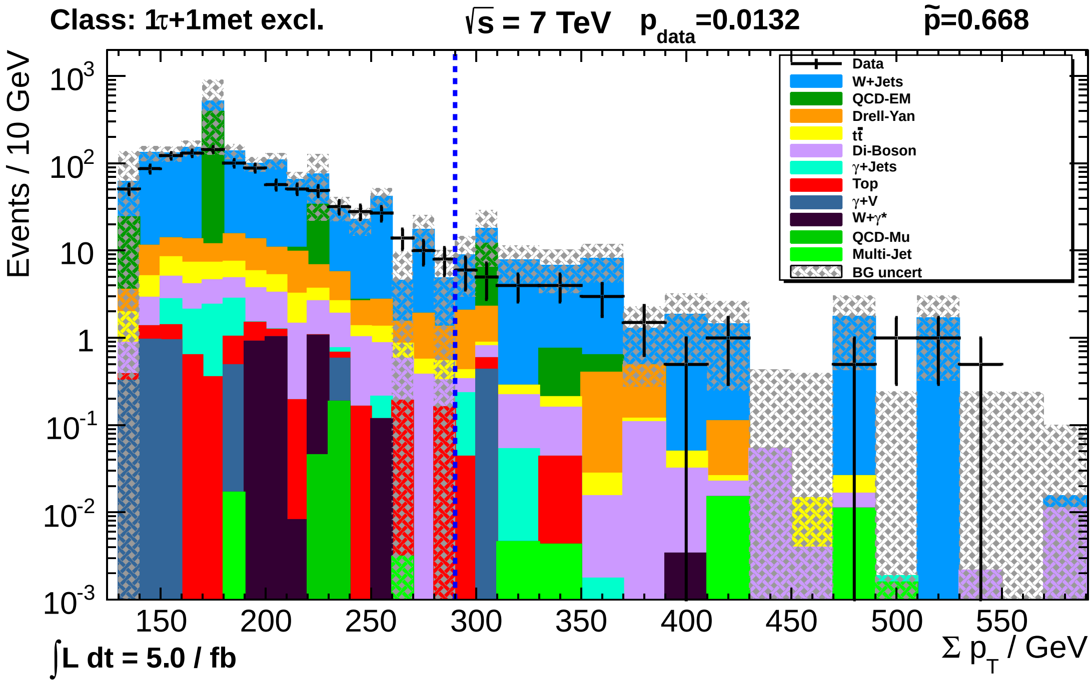

Figure 5.3: The $\sum p_T$ distribution of the $1\tau+\text{MET}$ event class. Dominated by the $W \rightarrow \tau\nu_\tau$ production this event class is the first to profit from the new triggering routine of MUSiC

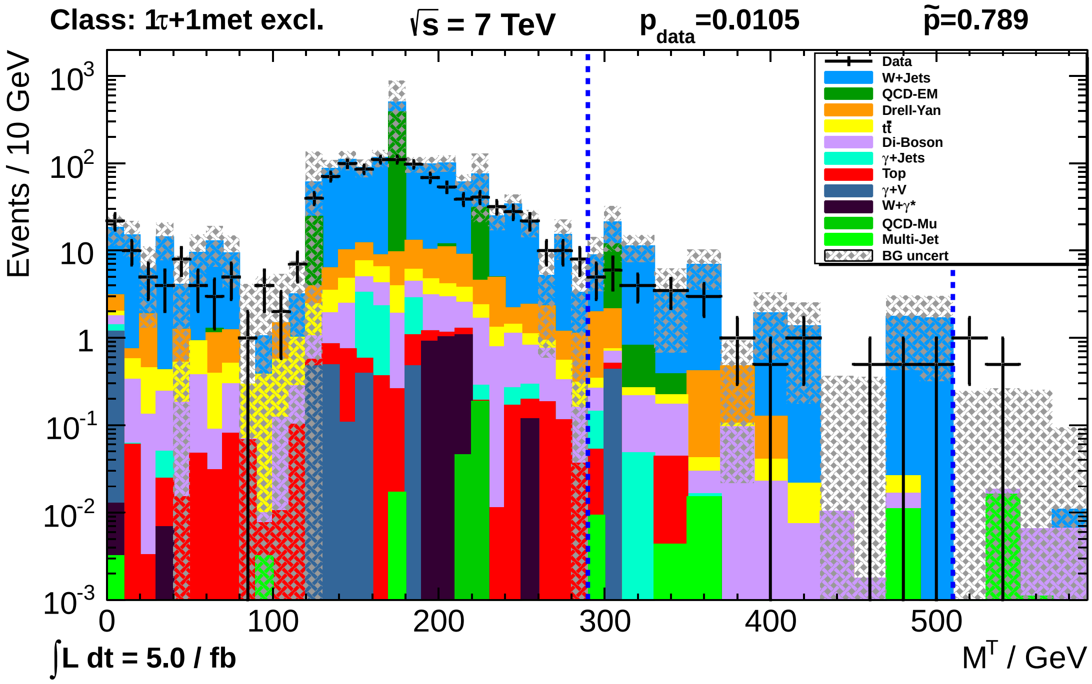

Figure 5.4: The transverse invariant mass distribution of the $1\tau+\text{MET}$ event class. Due to high trigger thresholds, the Jacobian peak of the W can not be seen

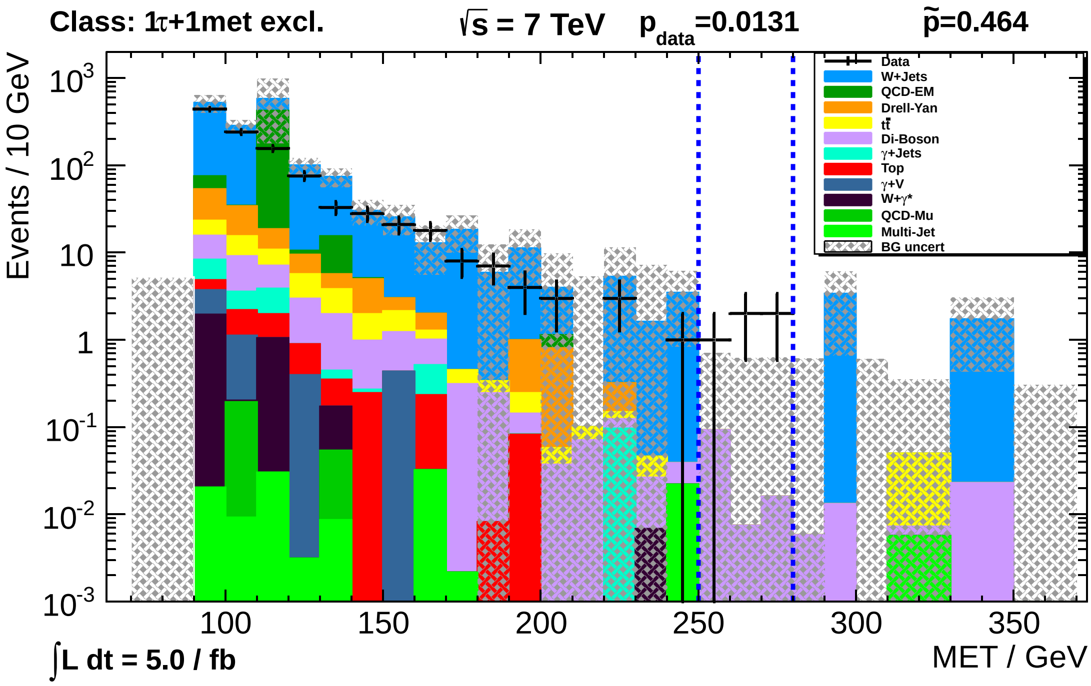

Figure 5.5: The MET distribution of the $1\tau + \text{MET}$ event class. Good agreement between data and the Monte Carlo prediction is observed.

corresponding QCD Monte Carlo samples. Nevertheless, since the QCD samples are assigned a large uncertainty of 50%, the data measurement in those bins still agrees well with the Monte Carlo prediction. The origin of QCD contributions can be jets misidentified as τ leptons or τ leptons produced during the showering of jets. However, those contributions do not dominate the overall distribution. The QCD contributions are mostly suppressed by the use of the decayModeFinding discriminator. Earlier studies with MUSiC have shown that not requiring this discriminator increases the QCD contribution in this class by two or even more orders of magnitude.

As expected, this class is dominated by the W-boson production (cf. Fig.4.5). In addition, a Drell-Yan contribution is observed. This contribution can be due to $Z \to \tau^+\tau^-$ decay with one τ decaying into low energetic $\mu\nu_\mu$ which does not pass the selection criteria. The region of interest starts at 300 GeV and ends at over 3500 GeV. For display purposes, the upper bound is not shown in the figure. The reasons for such a big range are the small expected MC contributions, e.g. around 800 GeV and further. In these terms, a minor deficit of data is observed by the algorithm resulting in the minimal p in this interval. However, the $\tilde{p}$-value of $\tilde{p} = 0.668$ indicates a good agreement between data and Monte Carlo prediction stating that the recorded deviations are insignificant. It is further noticeable that the data points are systematically beneath the Monte Carlo expectation. This effect is also visible in other classes dominated by $W+$ Jets, especially in those containing jets or τ jets in the final state. In event classes without jets in the final state, e.g. $\mu+$ MET, this effect is observed for momenta up to 200 GeV. However, the measured data is still within the range of the Monte Carlo uncertainty. The Monte Carlo prediction and data are altogether compatible to each other.

The distribution of the transverse invariant mass is presented in Figure 5.4. As in the $\sum p_T$ distribution, the transverse mass distribution is dominated by the $W+$ Jets sample. Unfortunately, it is not possible to resolve the W mass peak since the threshold of the $\tau+$ MET HLT trigger already suppresses the W resonance[4] and the offline trigger cut on the missing energy ($\text{MET}_{\text{trigger}} > 90$ GeV) enlarges this effect. Over all, the data is well described by the Standard Model Monte Carlo.

The $\not{E}_T$distribution is shown in Figure 5.5. In analogy to the $\sum p_T$ and transverse invariant mass distributions, the $W+$ Jets is dominant here. All in all, a good agreement between the data and the Standard Model Monte Carlo prediction is observed in this distribution.

5.3.2 μ+τ-MET Class

A further interesting event class is the $1\mu+1\tau+1$MET class. The $\sum p_T$ distribution is depicted in Figure 5.6 and shows an overall good agreement between data and the Monte Carlo prediction. A contribution to these classes is double W production (Di-Boson in the Figure) with one $W\to\mu\nu_\mu$ and one $W\to\tau\nu_\tau$ decay. A strong contribution is also expected by the Drell-Yan process ($qq\to Z/\gamma\to\tau\tau$) with one τ decaying hadronically and one leptonically into $\mu\nu_\mu$.

In contrast to the expectation, the leading background in this class is the $W+$ Jets background and not the Drell-Yan process. This is caused by a relatively high misidentification rate of the HPS algorithm (cf. Fig. 4.11). $W+$ Jets samples contain the decay of the W into either e, μ or τ in presence of up to four jets. From Figure 4.11, one can expect the probability for a jet to be identified as a τ by mistake to have an order of magnitude around $1-2\%$. The assumption that the $W+$ Jets contribution to this event class mainly consists of fake τ leptons is supported by the $\sum p_T$ distribution of $1\mu+1$jet$+1$MET shown in Figure 5.8. Due to a higher cut on the p_T of a jet, a direct comparison is not possible. Nevertheless, one can clearly see that the order of magnitude of the $W+$ Jets contribution reaches up to over $\mathcal{O}(10^4)$ whereas its contribution in the $1\mu+1\tau+1$MET has the order of magnitude of $\mathcal{O}(10^2)$. This supports

---

[4] The $\tau+$ MET HLT trigger is not intensively used by the CMS community due to a worse reconstruction efficiency and resolution of τ leptons in comparison to muons and electrons in general. Most analyses using this HLT trigger are searches for new exotic physics in higher energy ranges. Hence, these searches consider the production of the W an unneeded background. A W mass peak is clearly visible in the $\mu+$ MET transverse mass distribution.

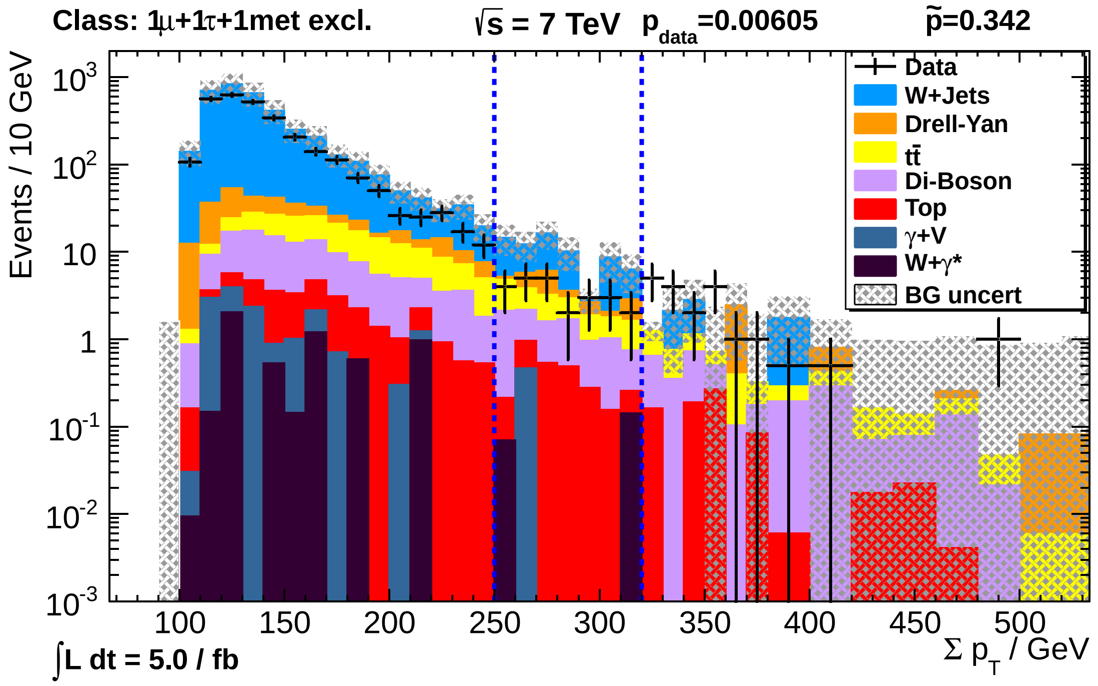

Figure 5.6: The $\sum p_T$ distribution of $1\mu + 1\tau + 1$MET. The Drell-Yan process is visible here – leptonic decays of the τ can be observed. However, it is not possible to identify such a decay unambiguously due to two neutrinos in the final state which are not visible.

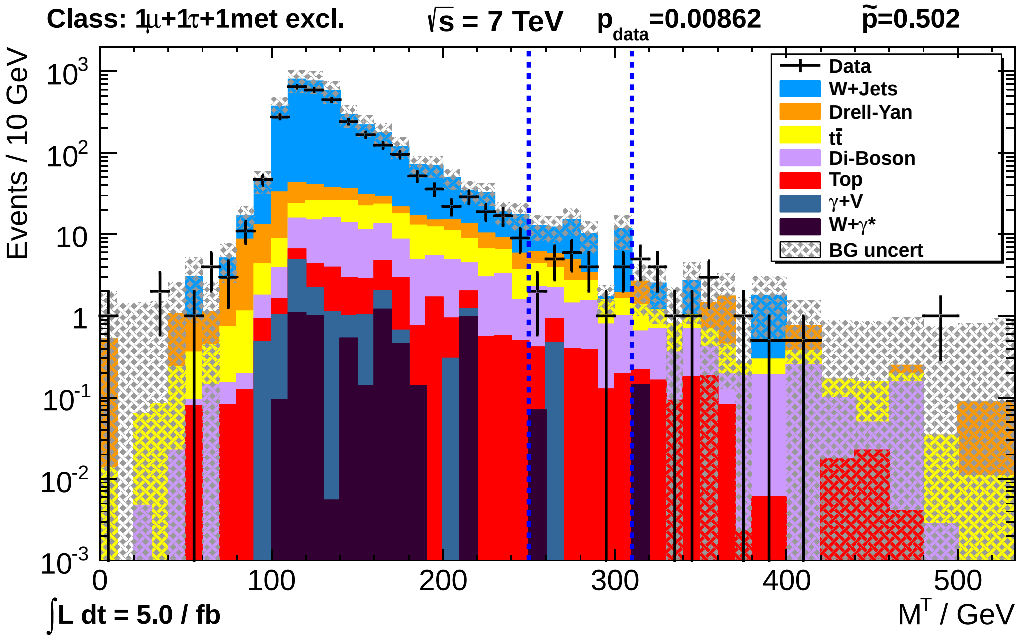

Figure 5.7: The transverse invariant mass distribution of $1\mu + 1\tau + 1$MET. Unfortunately, the Z peak can not be resolved because of the large W background and the high cuts on the transverse missing energy, the p_T of the τ and the trigger muon p_T.

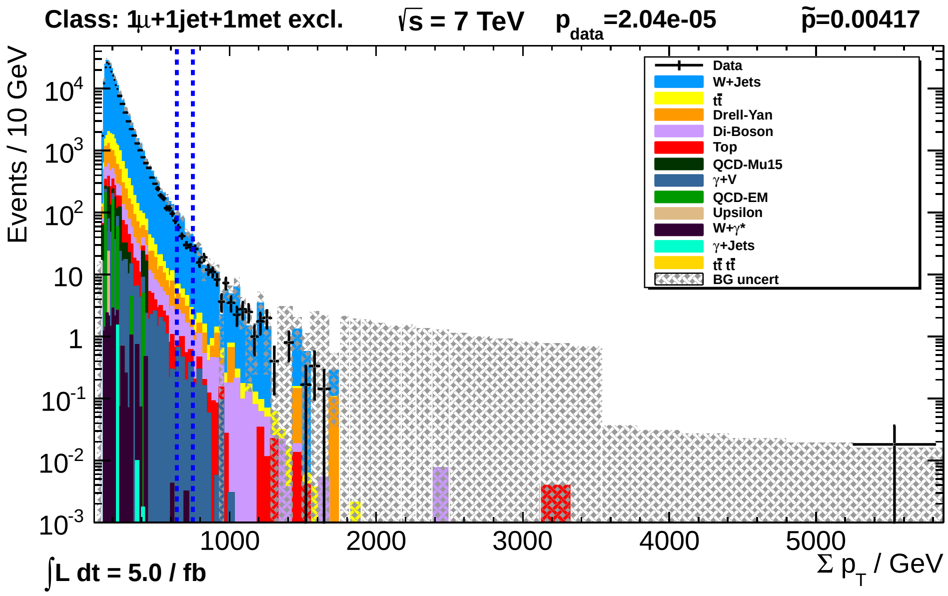

Figure 5.8: The $\sum p_T$ distribution of $1\mu + 1\text{jet} + 1\text{MET}$. The difference to the Figure 5.6 allows to estimate the order of magnitude of the misidentification rate of the HPS algorithm and could explain why W is the leading background in the latter distribution. The very high energetic event around 5000 GeV is likely a result of misreconstruction.

the latter hypothesis.

However, the contribution of the Drell-Yan process is visible, as expected, and one can also see the expected drop of its contribution with growing transverse momenta. A further interesting contribution is made by the $t\bar{t}$ sample. Knowing that the top quark decays almost exclusively into a W boson and a b quark [5], one would expect at least two additional jets in the final state assuming both W bosons to decay leptonically into $\mu\nu_\mu$ and $\tau\nu_\tau$. A plausible explanation would be that both jets do not pass the jet p_T cut of 50 GeV and therefore do not appear in the distribution. The region of interest is rather small in this distribution and is built around four bins around 300 GeV – in this case a deficit of data. However, the $\tilde{p}$ value of 0.342 indicates that this deviation is insignificant and this class is over all well described by the Monte Carlo prediction.

The distribution of the transverse mass is shown in Figure 5.7. Due to relatively high cuts on the μ, τ and MET it is unfortunately not possible to see the Z boson peak. The class is, as stated earlier, dominated by the W production. An overall $\tilde{p}$ of 0.502 demonstrates a good agreement between data and Monte Carlo.

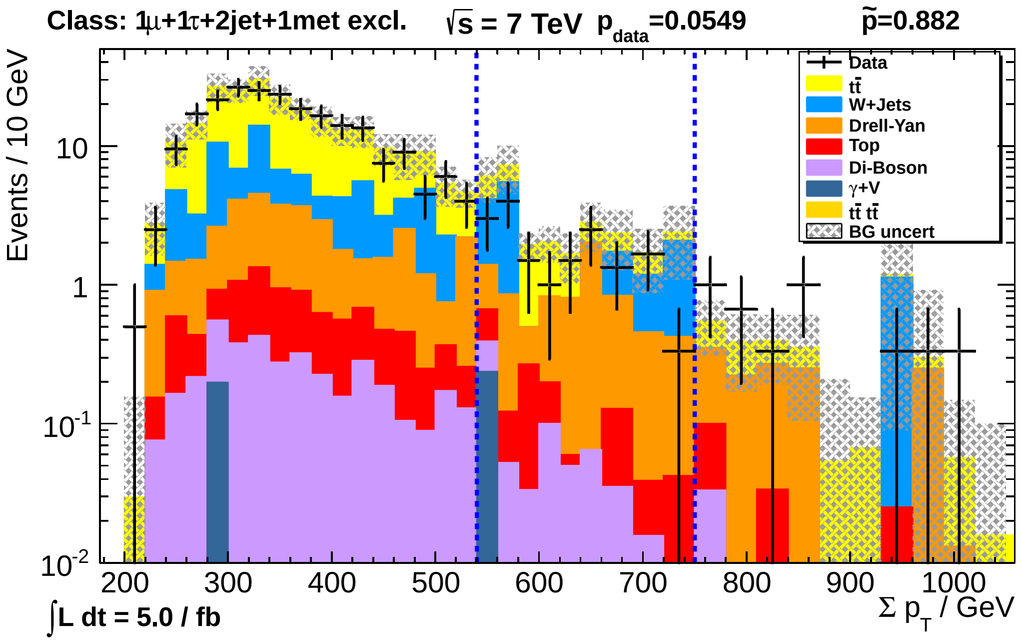

Figure 5.9: The transverse invariant mass distribution of $1\mu + 1\tau + 2$jet. This event class is clearly dominated by the $t\bar{t}$ production and presents a good agreement between the data and Monte Carlo. Altogether, approximately 100 events are observed by an integrated luminosity of 5.0/fb

5.3.3 μ+τ-2Jet-MET Class

The $1\mu + 1\tau + 2$jet + MET event class is interesting since it is expected to be dominated by $t\bar{t}$ production. Since the t quark is the heaviest particle in the Standard Model with $m_t \approx 172$ GeV [5] its understanding is crucial for searches for new physics. The $\sum p_T$ distribution is shown in Figure 5.9. As expected, the distribution is dominated by $t\bar{t}$ production. Over the energies of up to 1000 GeV, one observes a good agreement between data and Monte Carlo prediction. As expected, the main contribution is due to $t\bar{t}$. The second dominant background consisting of W + Jets is assumed to have its origin in jets misidentified as τ while the W boson decays into $\mu\nu_\mu$. The region of interest contains four bins around 800 GeV where a small insignificant excess of data is observed. The conclusions of the M^T and MET analyses (not shown) are similar.

5.3.4 μ+τ Class

The main purpose of the MUSiC framework is to search for new physics – physics beyond the Standard Model. Therefore, an "exotic" event class is presented here. There is no interaction within the Standard Model which would lead to this final state at the LHC – a $\mu + \tau$ final state would imply a violation of the lepton family number conservation. Only if some objects of the

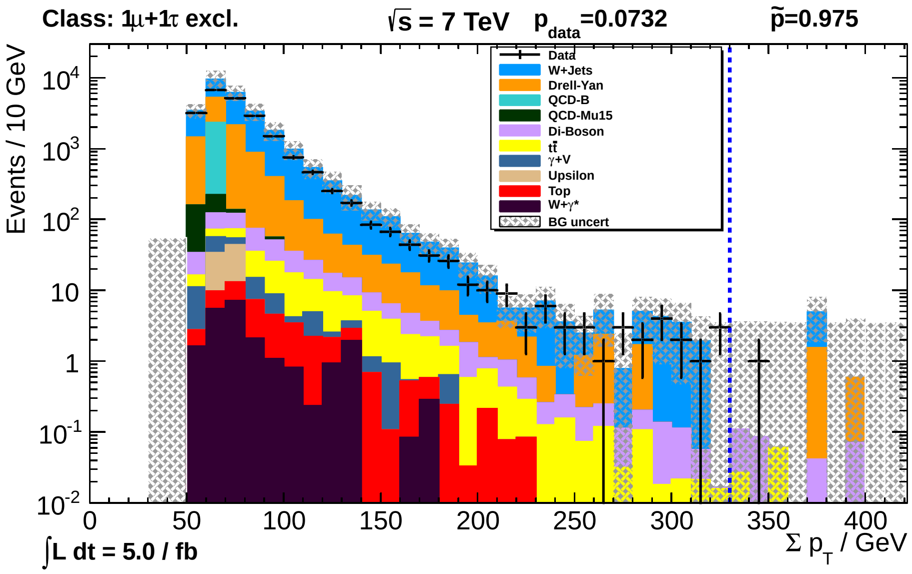

Figure 5.10: The $\sum p_T$ distribution of the $1\mu-1\tau$ event class. Being a promising candidate for exotic models this final state has to be understood well when searching for new physics. Overall, good agreement between data and Monte Carlo is observed. Dominating contributions are made by the W boson production. No evidence for new physics is observed.

initial Standard Model interaction did not pass the selection criteria, an exclusive $\mu+\tau$ class can appear. One dedicated analysis of this final state has been performed at III. Physikalisches Institut A [60], a search for τ sneutrinos performed within the SUSY research group.

The $\sum p_T$ distribution is shown in Figure 5.10. The main contribution is due to the $W+$Jets background. Since only a $\tau+$MET HLT crosstrigger was used during the analysis and no MET is present in the class, this event class is entirely triggered by the μ and not by the τ. Therefore, it is most probable that the triggering muon originated from the W decay while the τ is a misidentified jet. A comparable contribution is made by the Drell-Yan process. In this case, one could expect a $Z\to\tau\tau$ decay where one τ decays hadronically and the MET does not pass the selection criteria, or a $Z\to\mu\mu$ in presence of a radiated jet or a misidentified τ-jet with one μ not passing the selection. Some small contributions are also expected by QCD processes.

No significant deviations from the Monte Carlo prediction are detected. Most notably, a small deficit of data in high p_T region is observed. As a consequence, the region of interest starts around 350 GeV and spreads up to 1400 GeV (the whole region is not shown for demonstration purposes). The $\tilde{p}=0.975$ indicates that the data is well described by the Monte Carlo prediction. As it is not a primary aim of MUSiC framework to compute exclusion limits[5], one can not speak of exclusion of a certain theory. Nevertheless, the fact that MUSiC does not observe an excess in a given final state is a strong evidence for validity of the Standard Model in this range.

5.3.5 Observed deviations

The event classes described so far show a good agreement between the Monte Carlo prediction and observed data encouraging the evidence that the τ lepton is well understood both within the CMS experiment and within the MUSiC framework. However, there are event classes showing deviations which are very significant and therefore improbable to be due to statistical fluctuations only. Those event classes show that there is room for improvement of the understanding of τ physics at the LHC and at CMS. In the following, the most significant deviations are discussed. All deviations are listed in the Table 5.4.

5.3.5.1 $1e+1\mu+1\tau+1\gamma+3\text{jet}+\text{MET}$-Event Class

The only event class with τ leptons containing data but no Monte Carlo prediction is the $1e+1\mu+1\tau+1\gamma+3\text{jet}+\text{MET}$ event class. One data event is observed. Speaking statistically this event class has an infinite significance since no Monte Carlo sample predicts such a final state. However, it is natural that not every physical interaction can be simulated properly – a simulation can not reproduce nature completely. This event class contains seven physical objects along with missing transverse energy and thus has a very complex topology. The presence of three jets and a τ jet indicates a large QCD contribution and, as stated earlier, QCD interactions are the most difficult to simulate. In fact, it is almost remarkable that only one event class with such a complex topology has no Monte Carlo prediction in the MUSiC framework. In event classes with two or only one jet instead of three, small Monte Carlo prediction ($\mathcal{O}(10^{-2}\text{events})$) exists. However, there are no data events observed in neither one of those event classes. There is neither a Monte Carlo prediction nor data observed in event classes with four jets or more.

The event display is shown is Figure 5.11. As one can clearly see, the event is very busy. The transverse momentum of "Jet 4" is very close to the selection cut threshold. Hence, it is possible that this event could belong to the $1e+1\mu+1\tau+1\gamma+2\text{jet}+\text{MET}$ event class or an event class with less jets. Many low energetic jets which do not pass the selection cuts are observed. In the last section of the present work, suggestions are made for dealing with high jet multiplicities.

---

[5] First of all, the computational effort of calculating limits for that many final states would exceed the available resources. Furthermore, MUSiC uses a general approach not allowing for many selection criteria which are essential for particular final states.

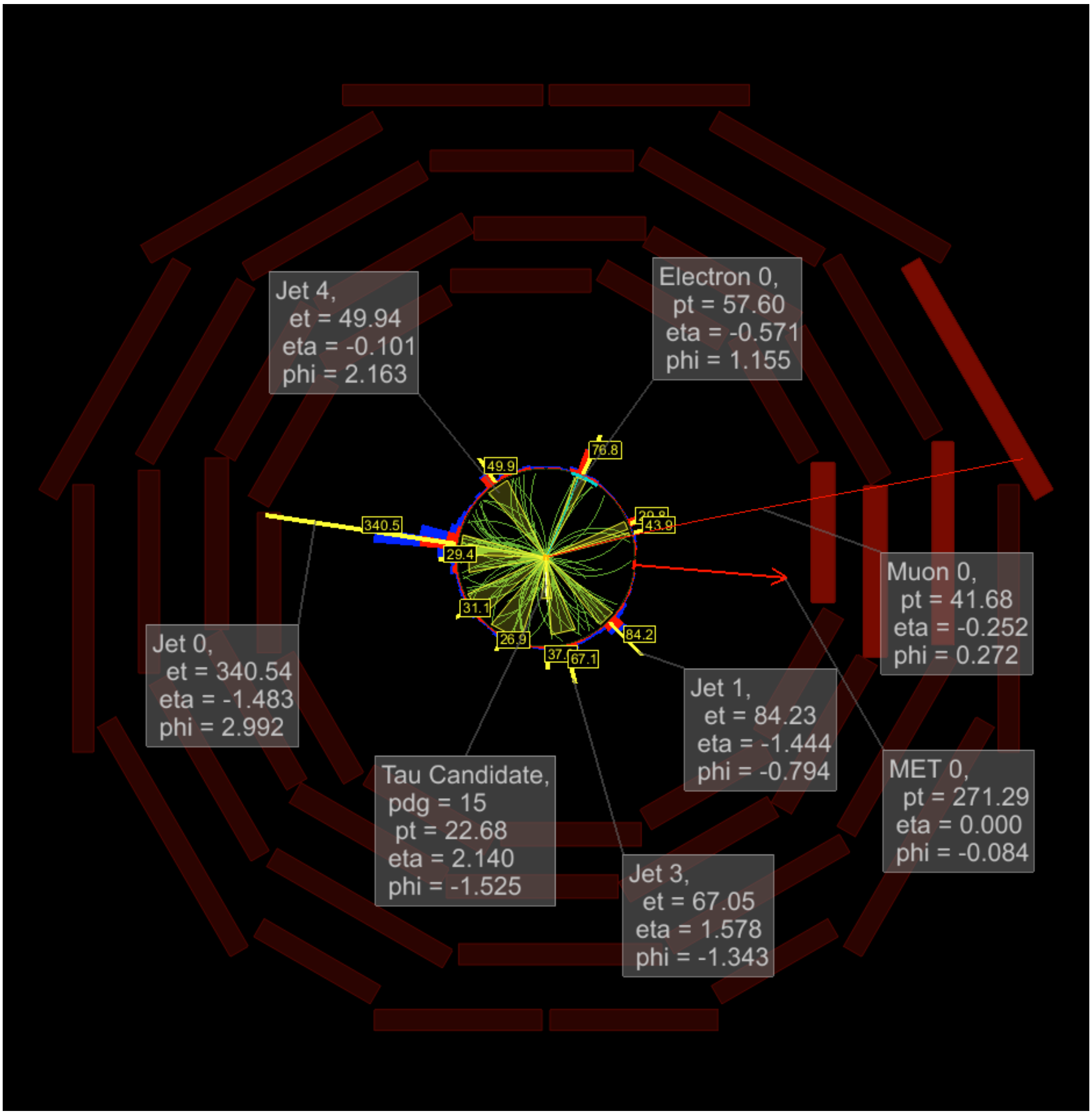

Figure 5.11: The event display of the final state with no Monte Carlo prediction. The event is very "busy". The p_T of "Jet 4" is close to the selection cut. It is possible that multiple interactions (pile-up) are responsible for the many hits in the tracker.

Event Class	Distribution	MC prediction	Evts observed	RoI/ GeV	$\tilde{p}$
$2\mu + 1\tau + 3$jet	$\sum p_T$	0.41 ± 0.20	7	$410 - 700$	$6.4 \cdot 10^{-5}$
$1e + 2\tau + 1$jet		0.057 ± 0.034	3	$330 - 600$	$8.2 \cdot 10^{-4}$
$2e + 1\tau + 1\gamma + 1$jet		$(2.8 \pm 2.4) \cdot 10^{-4}$	1	$210 - 240$	0.0012
$2e + 1\tau + 1\gamma + 3$jet		0.020 ± 0.020	2	$350 - 410$	0.0015
$2\tau + 1\gamma + 1$met		$< 1.4 \cdot 10^{-4}$	1	$450 - 480$	0.0015
$1\mu + 1\tau + 1\gamma + 3$jet		0.014 ± 0.014	2	$270 - 330$	0.0020
$1\tau + 1\gamma + 3$jet $+ 1$met		0.24 ± 0.10	4	$600 - 1050$	0.0033
$1\mu + 2\tau + 3$jet		$< 9.6 \cdot 10^{-4}$	1	$230 - 290$	0.0049
$1\tau + 1\gamma + 2$jet $+ 1$met		0.22 ± 0.09	4	$400 - 460$	0.0060
$1\mu + 1\tau + 5$jet $+ 1$met		1.42 ± 0.41	8	$700 - 860$	0.0061
$1\tau + 1\gamma + 3$jet $+ 1$met	MET	0.15 ± 0.08	3	$120 - 150$	0.0077
$1e + 1\tau + 1$jet	M^{T}	17.6 ± 13.0	124	$570 - 630$	$< 10^{-5}$
$1e + 1\tau + 3$jet		3.00 ± 0.92	20	$660 - 750$	$< 10^{-5}$
$1e + 1\tau + 2$jet		17.7 ± 13.8	106	$690 - 960$	$4 \cdot 10^{-5}$
$2\mu + 1\tau + 3$jet		0.32 ± 0.13	6	$490 - 760$	0.00019
$1e + 1\tau + 1\gamma + 2$jet		0.012 ± 0.024	3	$520 - 640$	0.0003
$2e + 1\tau + 1\gamma + 1$jet		$< 2.1 \cdot 10^{-5}$	1	$270 - 280$	0.00067
$1e + 1\tau$		0.63 ± 1.43	16	$510 - 710$	0.001
$1\tau + 1\gamma + 2$jet $+ 1$met		0.096 ± 0.052	4	$420 - 460$	0.001
$2\tau + 1\gamma + 1$met		$< 7.3 \cdot 10^{-5}$	1	$450 - 460$	0.0017
$1e + 2\tau$		0.16 ± 0.060	4	$140 - 180$	0.0024
$1\mu + 2\tau + 1$jet		$1.8 \cdot 10^{-4} \pm 0.011$	2	$160 - 170$	0.00284
$1e + 2\tau + 1$jet		0.0072 ± 0.012	2	$430 - 470$	0.0030
$1\mu + 2\tau$		0.037 ± 0.063	3	$60 - 80$	0.0055
$2e + 1\tau + 1$jet		0.12 ± 0.09	4	$570 - 610$	0.0067
$1e + 1\tau + 1\gamma + 3$jet		$< 3.6 \cdot 10^{-4}$	1	$780 - 810$	0.0073
$1e + 1\tau + 5$jet		0.105 ± 0.059	3	$630 - 750$	0.008

Table 5.4: A summary of deviations observed in the full scan of the CMS data from 2011. All event classes with at least one τ lepton and $\tilde{p} < 0.01$ are presented. In cases where an upper limit is set for the MC prediction, it has 68%C.L.

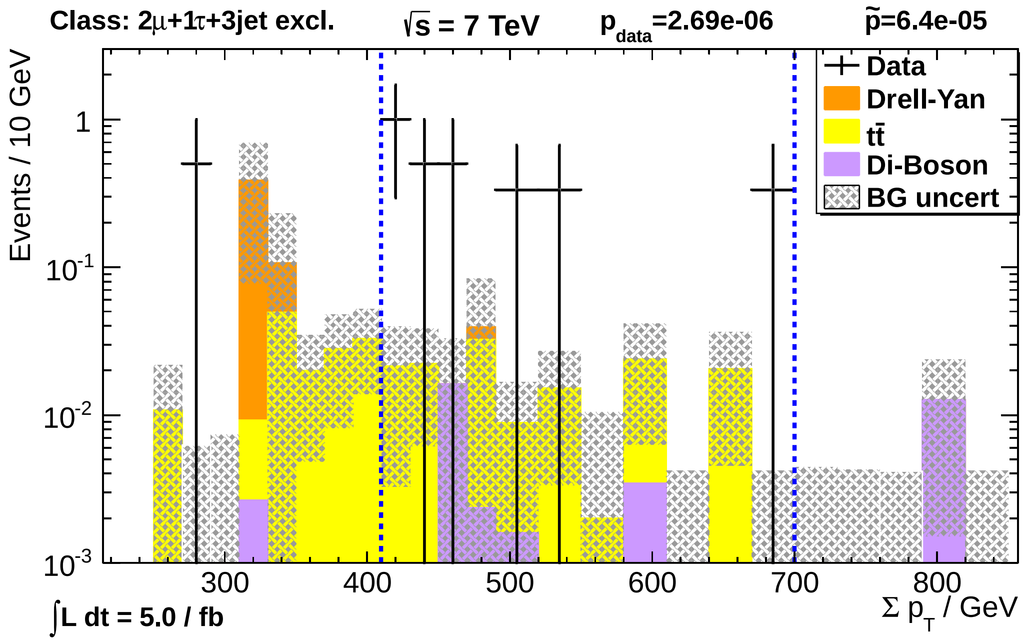

Figure 5.12: The $\sum p_T$ distribution of the $2\mu+1\tau+3$jet event class. Only a few data events between 400 GeV and 700 GeV provide a very significant $\tilde{p}$ since the Monte Carlo prediction in this range has the order of magnitude of $\mathcal{O}(10^{-1})$

5.3.5.2 $2\mu+1\tau+3$jet-Event Class

The lowest $\tilde{p}$ value of all $\sum p_T$ distributions is observed in the $2\mu+1\tau+3$jet event class (cf Fig. 5.12). In the region of interest between 400 GeV and 700 GeV, a significant excess of data is observed. Hereby, 7 data events appear versus 0.41 ± 0.20 events expected from the Monte Carlo prediction. The $\tilde{p} = 6.4 \cdot 10^{-5}$ value indicates that it is very unlikely that this excess is of statistical nature. One possible explanation for such behavior could be the misidentification probability of the HPS algorithm. Indeed, the misidentification probability of the τ lepton is taken into account by MUSiC. However, it is not easy to take into account the fact that a wrongly identified hadronic τ erases the corresponding jet from the jet collection or that a true τ is still considered a jet. This behavior leads to different final states and thus different event classes. The $2\mu+4$jet shows a good agreement between Monte Carlo and data ($\tilde{p} = 0.394$). In case one jet out of four has been wrongfully identified as a τ in data but not in Monte Carlo, such deviation could arise. In event classes with only one or two jets, however, a good agreement between data and Monte Carlo is observed ($\tilde{p} > 0.8$). In the $2\mu+1\tau+1$jet event class, ≈ 100 events are observed. The $2\mu+1\tau+2$jet contains ≈ 10 events. This fact could be a hint that the misidentification rate discrepancies between simulation and the actual measurement of the HPS algorithm are not well understood yet. A further investigation of the τ reconstruction algorithm is needed in order to understand the nature of this excess.

5.3.5.3 $1e+1\tau+\{1/2/3\}$jet-Event Classes

A further interesting deviation is observed in the $1e+1\tau$ along with one, two or three jets. Those event classes are the most significant event classes in the invariant mass distributions. Hereby, the classes with one and three jets have a $\tilde{p} < 10^{-5}$ value which means that no hit has been achieved during the dicing of pseudo-experiments (cf. Figure 5.13). The invariant mass distribution of the $1e+1\tau+2$jet event class is depicted in Figure 6.1 in the appendix. The corresponding $\tilde{p}$ value is $4 \cdot 10^{-5}$. While the event classes with one or two jets are dominated by the QCD samples (therefore "smeared") and suffer from insufficient statistics of QCD Monte Carlo samples, the event class with three jets is dominated by the $t\bar{t}$ production and still has a very high significance. The region of interest lies approximately around the same bins in all three event classes. The fact that all those three very nearby final states are dominated by different Standard Model backgrounds and still show the highest significance in narrow regions of interest is very startling. A definite statement can not be made here since further careful investigations of the contributing Monte Carlo samples and a comparison with the LHC 2012 data are necessary in order to understand the nature of those deviations. Yet, it is for sure that this is how new physics phenomena might present themselves.

5.3.5.4 $1e+2\tau+1$jet-Event Class

A similar deviation is observed in the $1e+2\tau+1$jet event class. Here, the Monte Carlo prediction is very low (consistent with 0) and a few data events between 320 GeV and 600 GeV result in a significant deviation. As in the event class in section 5.3.5.2, a relatively high jet multiplicity could be responsible for such deviation.

5.3.5.5 $1e+1\tau$-Event Class

In section 5.3.4, an exotic event class is discussed. $1e+1\tau$ is a further exotic final state. In contrast to to the $1\mu+1\tau$ event class, a significant deviation is observed in the transverse invariant mass distribution (cf. Fig. 5.15). While in the mass region up to 500 GeV an overall good agreement between data and Monte Carlo prediction is visible, a significant deviation is found by the MUSiC algorithm in the region of interest which starts at 510 GeV and ends at 710 GeV. Also, the uncertainties are larger compared to the $1\mu+1\tau$ final state due to the fact that it is almost impossible for a jet to fake a muon since the muon system is located outside the CMS magnet. In contrast to that, electrons originating in jets have a big uncertainty and are clearly dominant in many bins of this event class. A possible reason for the deviation observed in the region of interest could be the low statistics of the QCD samples. As one can see, all contributions made by the QCD samples are "spikes" – their shape is not smooth. From this point of view, it would be reasonable to suggest that better statistics could resolve

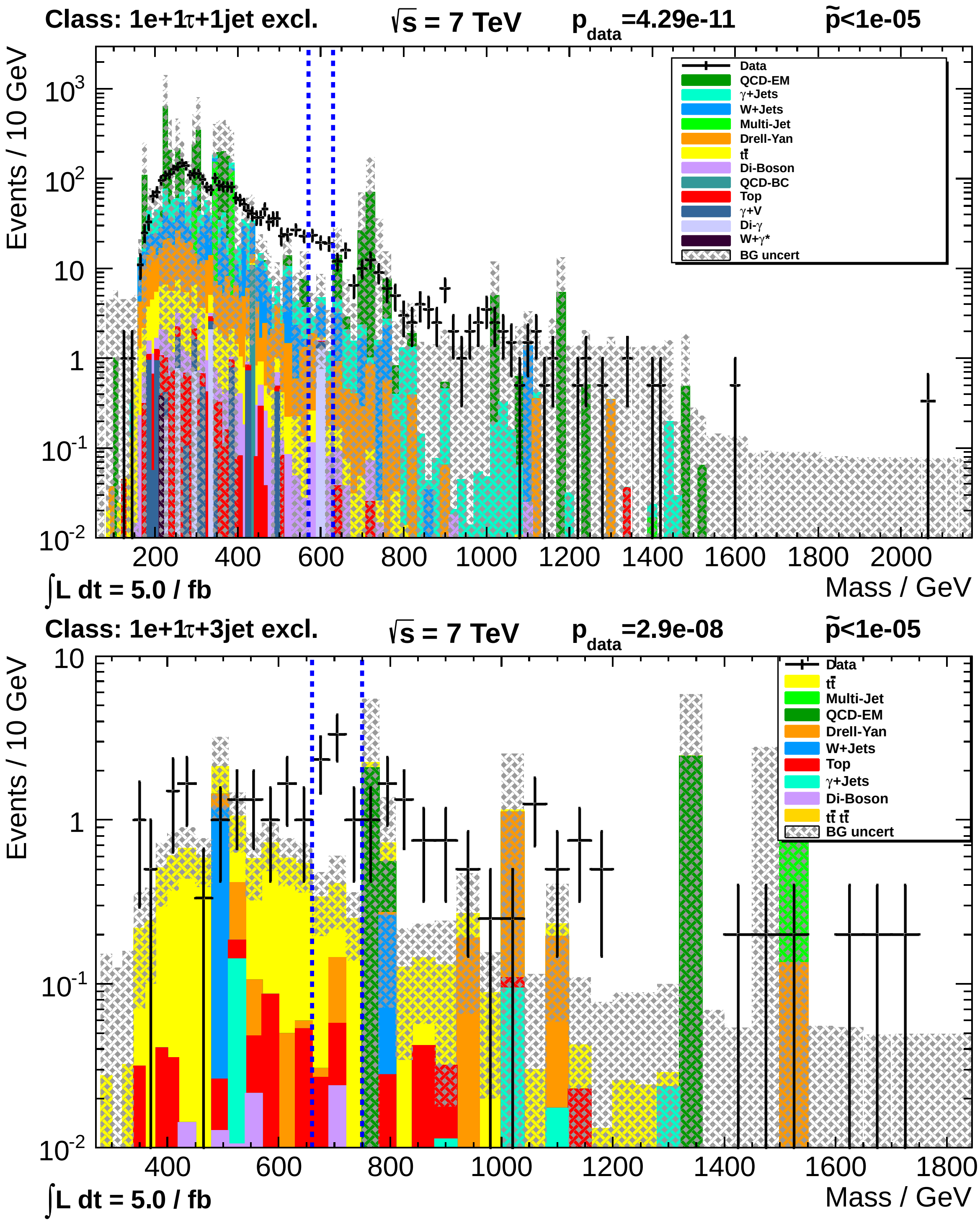

Figure 5.13: The invariant mass distributions of the $1e+1\tau+1$jet (top) and $1e+1\tau+3$jet (bottom) event classes. These event classes possess the highest significance out of all invariant mass distributions.

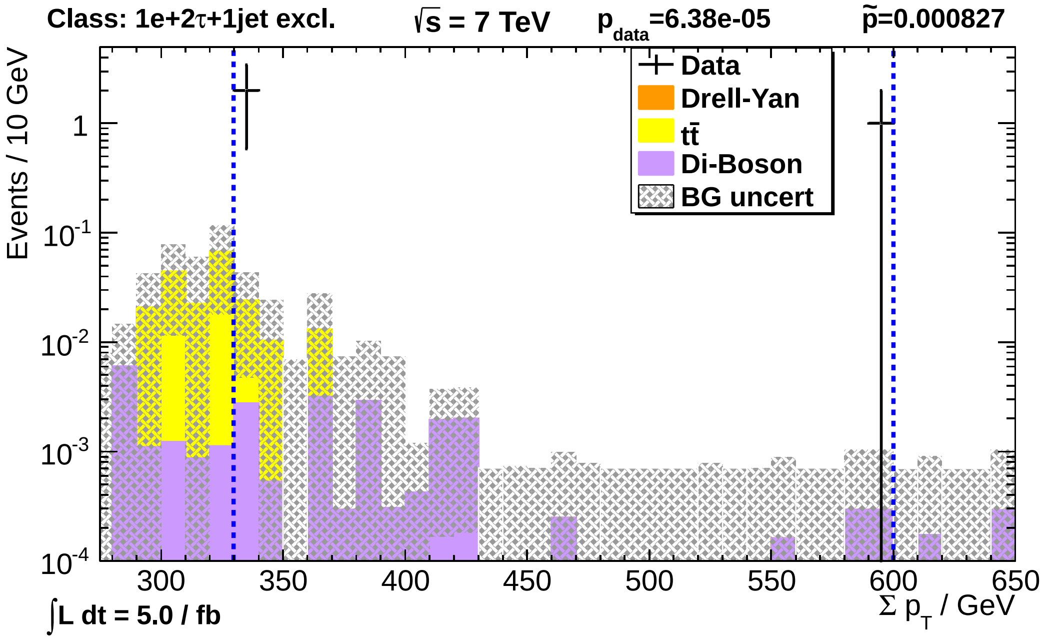

Figure 5.14: The $\sum p_T$ distribution of the $1e + 2\tau + 1$jet event class. The situation observed here is similar to the one observed in section 5.3.5.2. While the expectations is 0.057 ± 0.035 events, 3 data events are observed.

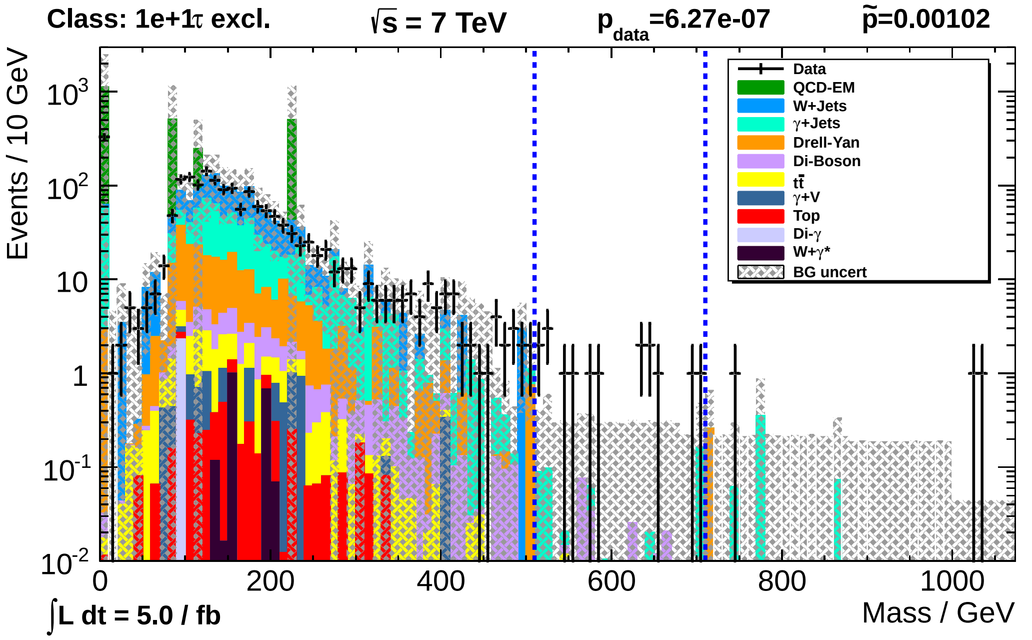

Figure 5.15: The invariant mass distribution of the $1e + 1\tau$ event class. Large contributions are made by the QCD samples.

this issue. However, this deviation is very interesting and should be investigated in the CMS 2012 data. A hint of new physics could indeed be the reason for this observation.

In the majority of cases, significant discrepancies between the Monte Carlo prediction and data occur in event classes with at least one jet and one τ jet. This supports the hypothesis that distinguishing between hadronic τ decays and jets in one final state and especially the simulation of jet radiation at the LHC and its interpretation within the MUSiC framework have to be improved. Also, a bigger dataset (over $20\,\mathrm{fb}^{-1}$ are planed for 2012 at $\sqrt{s} = 8$ TeV) might resolve some issues or at the same time give a better prospect on the nature of the observed deviations. A summary of deviations is given in Table 5.4. Also, some further deviating event classes are depicted in Figures 6.1 and 6.2 in the Appendix.

5.3.6 $\tilde{p}$ distribution

Not every (deviating) distribution can be discussed within the scope of the present thesis. All $\tilde{p}$ values are summarized in a distribution comparing the expectation and the observed values. The $\tilde{p}$ distribution for the $\sum p_T$ scan is shown in top of Figure 5.16. Overall, deviations from the expectation are observed. Starting with the forth bin the measured $\tilde{p}$ values are higher than one would expect from a pure statistical point of view. Due to the reasons discussed in earlier sections, such as QCD modelling, the possibly underestimated misidentification rate of the HPS algorithm, one can not expect a perfect agreement. The $\tilde{p}$ value for the missing energy distribution is shown at the bottom of Figure 5.16. Also here, deviations are observed. Unlike the $\sum p_T$ distribution, more event classes show 2σ or 3σ deviations than expected but there are no larger deviations from the expectation. The $\tilde{p}$ of the (transverse) invariant mass is shown in the middle of Figure 5.16. Most of the deviations are observed in this category. It is noticeable that this is the only distribution where the bin with the highest significance goes up compared to the one before. Invariant mass is a derived quantity and not directly measured by CMS. The observed deviation could be interpreted as a "sum up" of issues described so far. Nevertheless, the shape of the measured $\tilde{p}$ for $\sum p_T$ and the missing transverse energy still shows that a global understanding of τ physics is provided within the MUSiC framework. Even more, many processes with τ leptons show a good agreement and give strong evidence that searches for new physics can be performed with MUSiC.

5.4 Study of sensitivity of MUSiC

As it is presented in section 5.3, no strong evidence for physics beyond the Standard Model could be found with the MUSiC framework. However, it is advisable to conduct a study testing the ability of MUSiC to find a potential signal in final states containing τ leptons. In the present work, one model is tested:

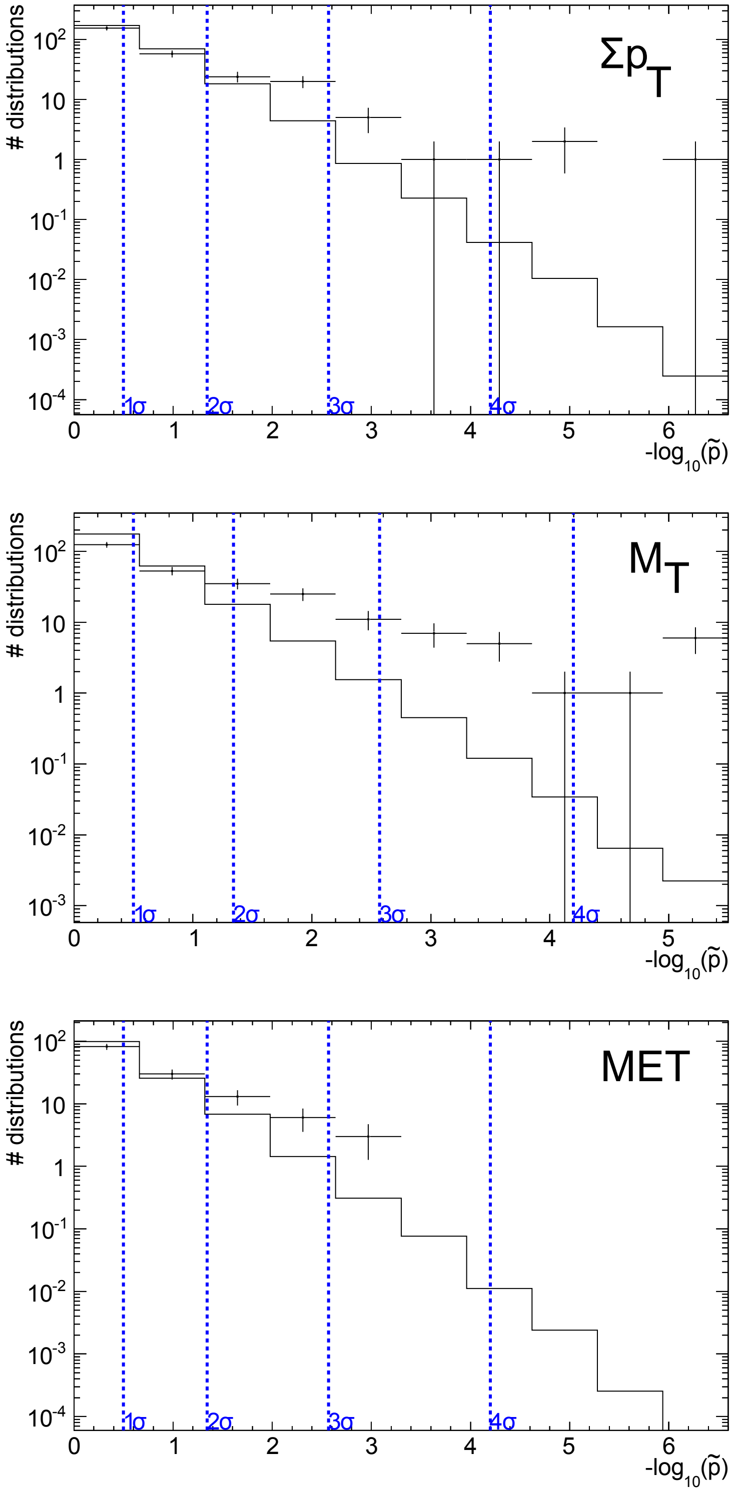

Figure 5.16: The $\tilde{p}$ distributions for the full scan of the 2011 CMS data. On the top, the $\tilde{p}$ of all $\sum p_T$ distributions, in the middle the $\tilde{p}$ of the (transverse) invariant mass and the $\tilde{p}$ of the missing transverse energy at the bottom. On the x-axis one finds the significance and on the y-axis the number of classes with the corresponding $\tilde{p}$ values. The boxed bins represent the expected distribution of the $\tilde{p}$ values. The crosses are the observed $\tilde{p}$ values. The last bin of the distribution is the overflow bin.

Sample	Mass in GeV	cross section in pb	$\sin^2(\phi)$	Events in the sample
W′ enhanced	1800	0.13	0.031	50000
	1900	0.11		
	2000	0.090		
W′	1300	0.22	–	
	1400	0.14		
	1500	0.095		

Table 5.5: The summary of examined W′ samples with their properties. The cross sections of the W′ with enhanced third generation coupling have been calculated by the W′ group of the IIIA institute. The cross sections of the W′ are taken from [61].

- W′ with enhanced coupling to the third generation of leptons.

For this coupling (cf. section 2.2.2), a mixing angle of $\sin^2(\phi) = 0.031$ is chosen corresponding to an almost exclusive coupling of the W′ to the τ. The characteristics are summarized in Table 5.5. For sensitivity studies, pseudo data is diced by the MUSiC framework. Hereby, for each bin a value is diced within the assumed uncertainties (both statistical and systematic) of the Monte Carlo prediction and a p-value is calculated. The distribution with the median p-value is then chosen as the final distribution. Subsequently, the usual MUSiC algorithm reviews the distribution and calculates the $\tilde{p}$ value.

The pseudo data was diced 100 times. The number of dicing experiments for the computation of $\tilde{p}$ was set to 10^6 with a "hit threshold" of 100 hits.

5.4.1 Results of the W′ sensitivity study

Several studies have been performed in order to detect the W′ signal. The one with the highest W′ mass of $m_{W'} = 2000$ GeV is presented in the following.

Significant deviations are observed in the $1\tau + 1$MET event class only. The corresponding distributions are shown in Figure 5.17, 5.18 and 5.19. In all three distributions, the dominating background is the Standard Model W boson production. The $\sum p_T$ shows the most convincing prospect. Here, a broad deviation is observed and a very broad region of interest is selected by the MUSiC algorithm. In both the missing transverse energy and the transverse invariant mass distributions, the MUSiC scanning algorithm was also able to locate the signal even if it did not mark a broad range on the x-axis. Interestingly, the missing transverse energy shows the lowest $\tilde{p} = 1.1 \cdot 10^{-5}$ value even though the region of interest is rather tight in this particular pseudo-experiment. The $\sum p_T$ and the M^T distributions show a somewhat larger $\tilde{p} \approx 10^{-4}$. Combining this information one obtains strong evidence for the fact that these deviations are *not* of statistical nature and that a new physics phenomenon might have been

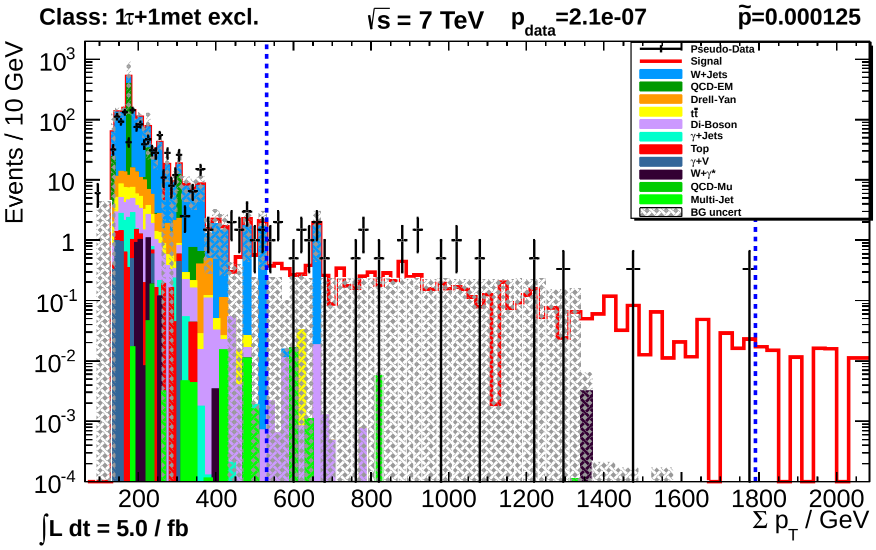

Figure 5.17: The $\sum p_T$ distribution of the $1\tau + 1$MET event class of the W′ sensitivity study for $m_{W'} = 2000$ GeV. The last event in the region of interest rests upon a fill up uncertainty of $\approx 10^{-4}$ which is not shown for demonstration purposes. The $\tilde{p}$ indicates a significant deviation from the Monte Carlo prediction resulting in a very broad (over 1000 GeV) region of interest.

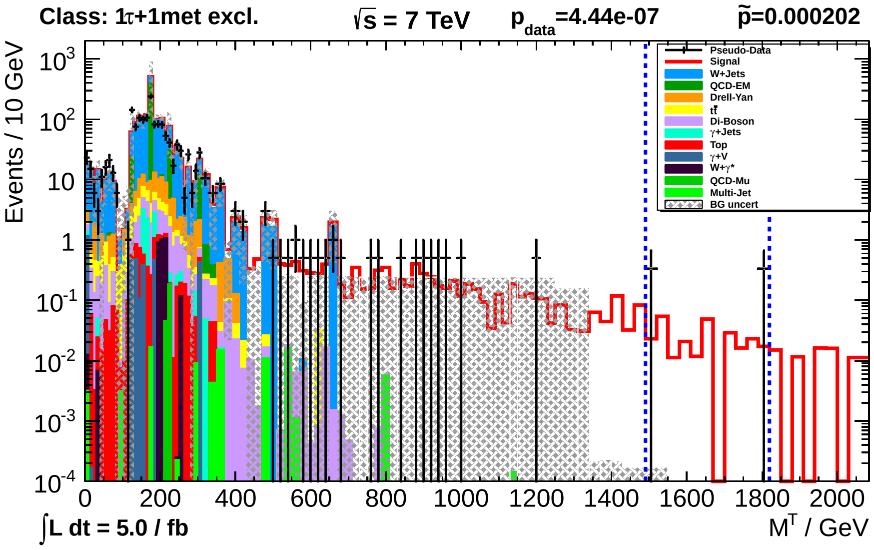

Figure 5.18: The transverse invariant mass distribution of the $1\tau + 1$MET event class of the W′ sensitivity study for $m_{W'} = 2000$ GeV. The second event in the region of interest rests upon a fill up uncertainty of 10^{-6} which is not shown for demonstration purposes. A significant deviation is observed in the region interest.

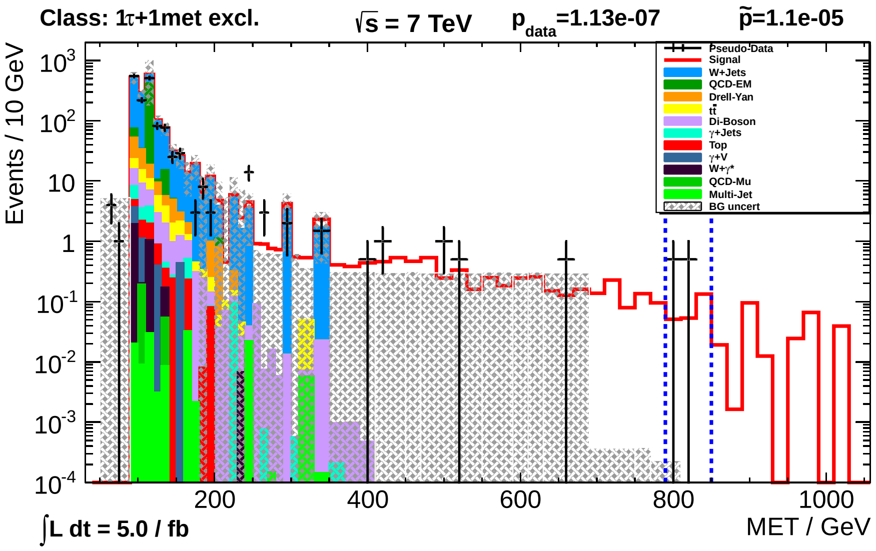

Figure 5.19: The missing transverse energy distribution of the $1\tau + 1\mathrm{MET}$ event class of the W' sensitivity study for $m_{\mathrm{W}'} = 2000$ GeV. The second event in the region of interest rests upon a fill up uncertainty of $\approx 10^{-5}$ which is not shown for demonstration purposes. A significant deviation is observed in the region interest.

found by MUSiC. The $\tilde{p}$ distribution is shown in Figure 5.20.

All in all, MUSiC with τ leptons is sensitive to W' with a mass around $m_{\mathrm{W}'} = 2000$ GeV and the corresponding cross section of $\sigma_{\mathrm{W}'} = 0.090$ pb. In cases where a carbon copy of the Standard Model W' was tested, a similar sensitivity is measured. Here, a W' with a mass $m_{\mathrm{W}'} = 1500$ GeV and a corresponding cross section of $\sigma_{\mathrm{W}'} = 0.095$ pb (cf. Table 5.5) is visible. In dedicated W' studies performed by the members of the III. Physikalisches Institut A, a preliminary exclusion limit for the W' mass has been set to $m_{\mathrm{W}'} > 1.9$ TeV for the $\mathrm{W}' \rightarrow \tau\nu_\tau$ channel [62]. This threshold is indeed higher than the one observed by MUSiC. However, the potential of MUSiC for discovery of new physics in τ channels is evident. For the $\mathrm{W}' \rightarrow \mu\nu_\mu$ decays, the exclusion limit is higher since the measurement of μ leptons has smaller uncertainties at CMS. The exclusion limit based on 2011 CMS data is $m_{\mathrm{W}'} > 2.2$ TeV at 95% confidence level [63].

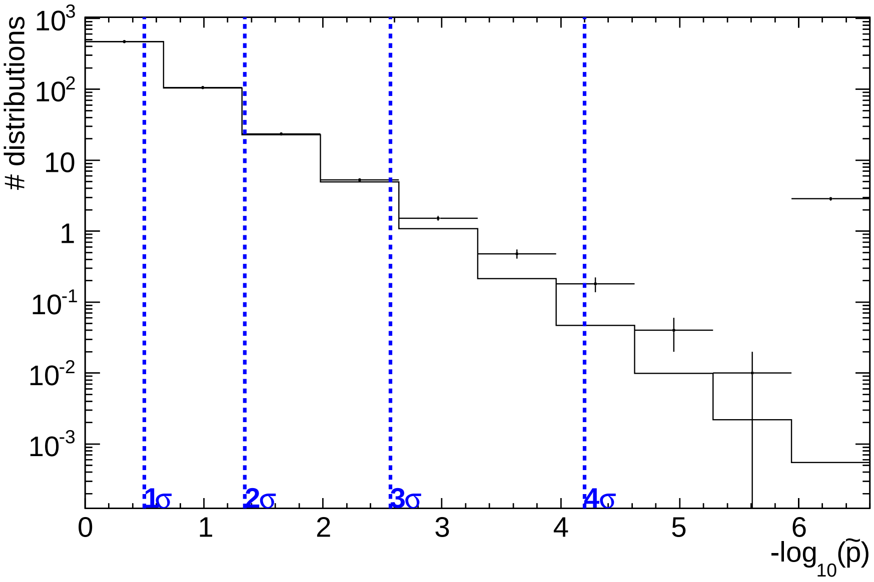

Figure 5.20: The $\tilde{p}$ distribution for the sensitivity study summed up for all distributions $\sum p_T$, M^T and MET. A clear deviation is visible.

CHAPTER 6

Conclusion and Outlook

A new object – the hadronic decay of the τ lepton has been introduced to the MUSiC framework. As a result, the sensitivity of the MUSiC framework has increased and the variety of the final states which can be analyzed by MUSiC has been expanded.

In the present thesis, a full scan of the CMS 2011 data corresponding to 5.0 /fb has been conducted. In general, good agreement between data and the Monte Carlo prediction is observed. Some deviations between the prediction and the measurement were found. Those can be traced back to difficulties during the QCD modelling and the reconstruction of the hadronic decays of the τ lepton but might also be indications for new physics. A sensitivity study has been performed showing that the MUSiC framework is also capable of detecting new physics in τ channels – a significant improvement since some exotic models predict enhanced or even exclusive third generation couplings.

Further studies with the MUSiC framework could concentrate on the following issues since these could not be resolved during the present analysis:

- **Jet multiplicities**
  A possible approach to the modelling of QCD processes at the LHC could be the introduction of event classes with a lower limit for the number of jets, e.g. $1\mu + 3\text{jets}\oplus$ meaning that final states with at least three jets are all stored in this event class. Especially for high jet multiplicities, this approach could save a lot of computing time and ameliorate the problems in generating high jet multiplicities.
- **Data driven QCD background determination**
  The present work has shown, among others, that simulating QCD processes at the LHC is a very challenging task not always manageable at the moment. A very promising approach would be to determine the QCD background with data-driven methods as they are used by many analyses in the CMS community. However, this approach would

have to be adapted to the MUSiC framework since MUSiC wants to keep its generality while searching for new physics.

Finally, in the course of the present thesis τ- and multi-object triggering has been introduced into MUSiC, further enhancing its purpose as a tool in searches for physics beyond the Standard Model.

APPENDIX

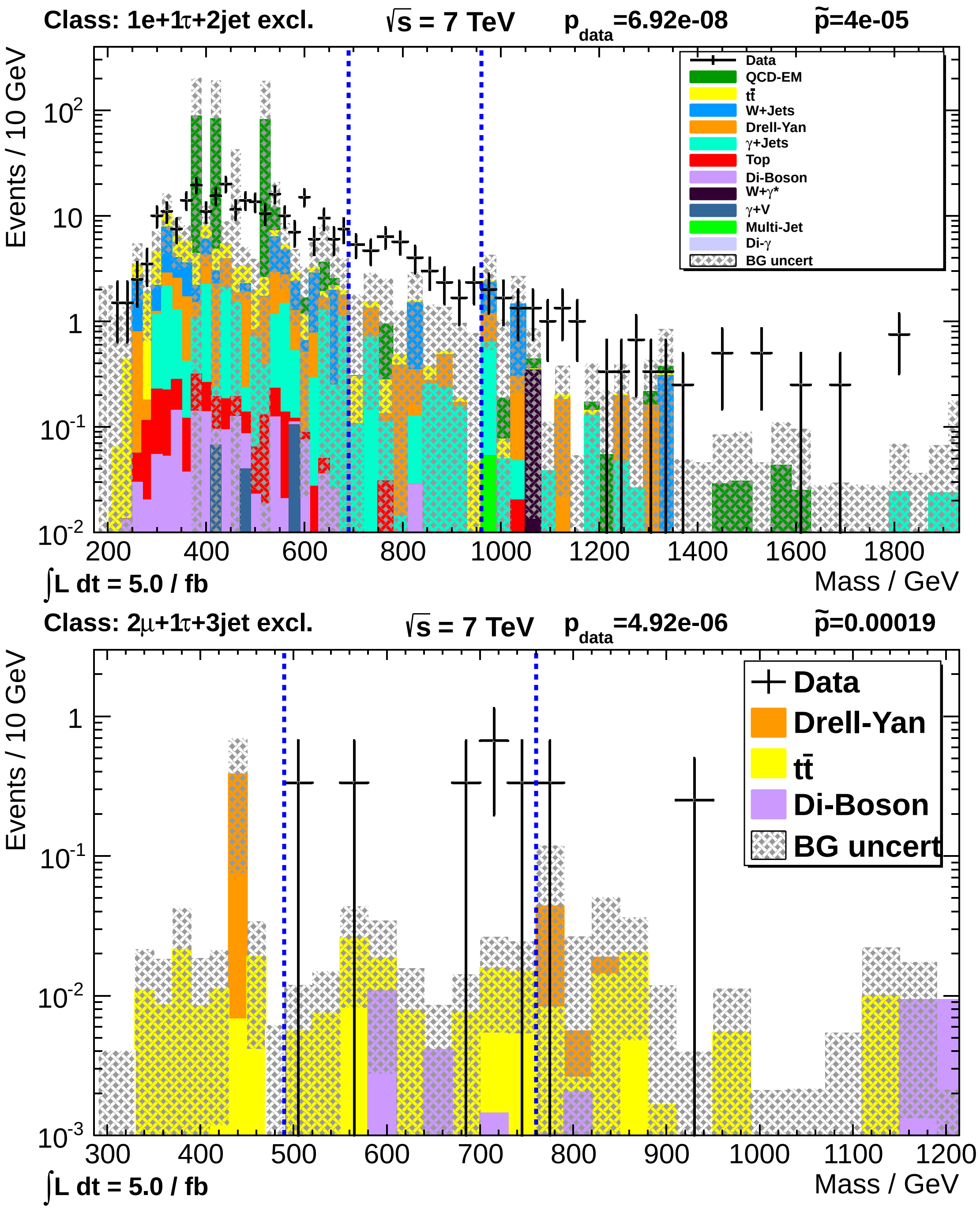

Figure 6.1: Further deviating event classes (cf. Table 5.4).

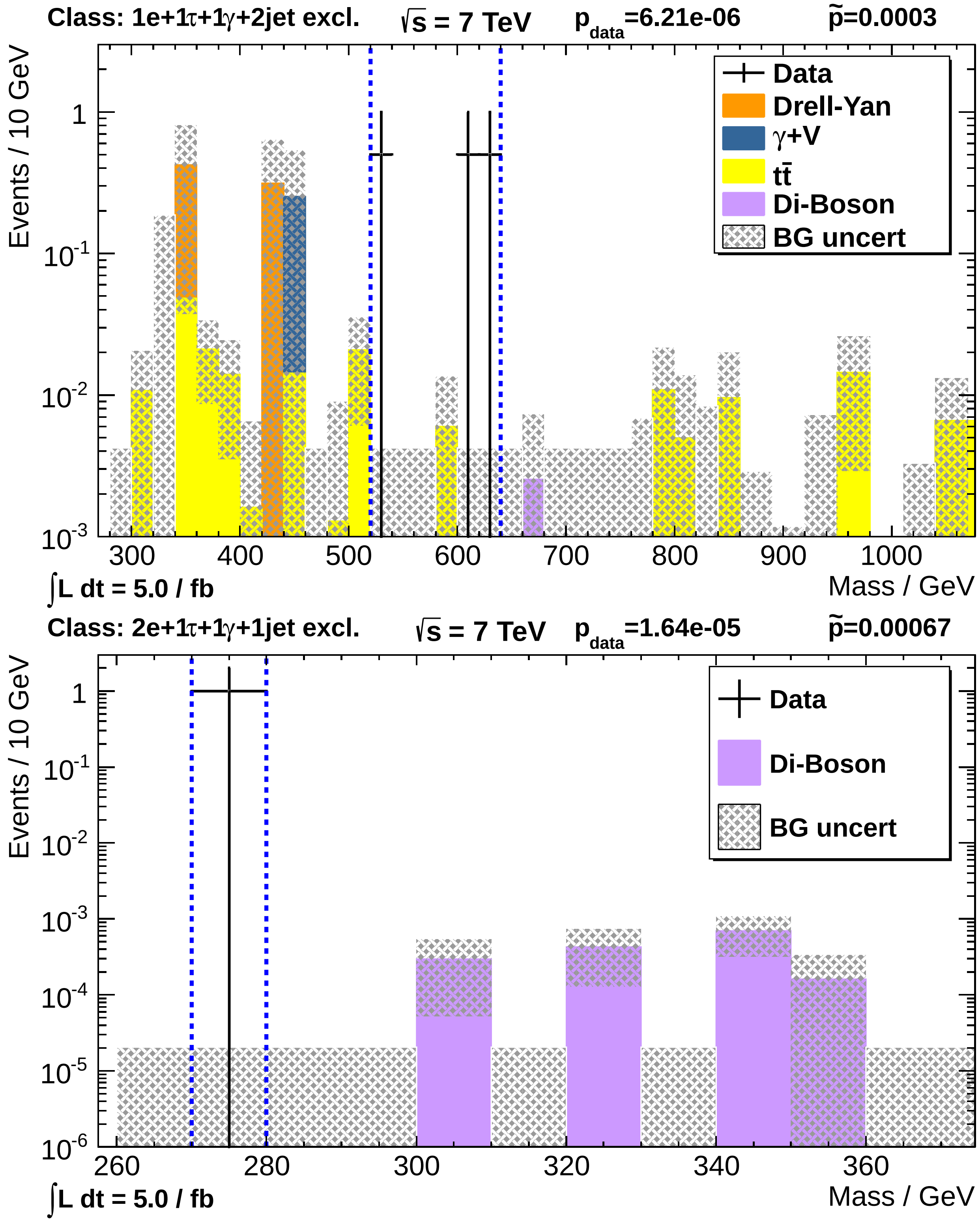

Figure 6.2: Further deviating event classes (cf. Table 5.4).

Name of the discriminator	Short description
AgainstElectronLoose	electron pion multivariate discriminator < 0.6
AgainstElectronMedium	electron pion MVA discriminator < -0.1 and **not** $1.4442 < \|\eta\| < 1.566$
AgainstElectronTight	electron pion MVA discriminator < -0.1 and **not** $1.4442 < \|\eta\| < 1.566$ and Brem pattern cuts [59]
AgainstElectronMVA	anti-electron MVA discriminator [64]
AgainstMuonLoose	τ lead track not matched to chamber hits
AgainstMuonMedium	τ lead track not matched to global/tracker muon
AgainstMuonTight	τ lead track not matched to global/tracker muon and large enough energy deposit in ECAL + HCAL
ByDecayModeFinding	Decay mode matching [58]
ByVLooseIsolation	isolation cone of 0.3 , no PF Charged Candidates with $pT > 1.5\,\mathrm{GeV}$ and no PF γ candidates with $E_T > 2.0\,\mathrm{GeV}$
ByLooseIsolation	no PF Charged Candidates with $pT > 1.0\,\mathrm{GeV}$ and no PF γ candidates with $E_T > 1.5\,\mathrm{GeV}$
ByMediumIsolation	no PF Charged Candidates with $pT > 0.8\,\mathrm{GeV}$ and no PF γ candidates with $E_T > 0.8\,\mathrm{GeV}$
ByTightIsolation	no PF Charged Candidates with $pT > 0.5\,\mathrm{GeV}$ and no PF γ candidates with $E_T > 0.5\,\mathrm{GeV}$
ByVLooseCombinedIsolationDBSumPtCorr	isolation cone of 0.3 , Delta Beta corrected $\sum p_T$ of PF charged and PF γ isolation candidates ($p_T > 0.5\,\mathrm{GeV}$) less than $3\,\mathrm{GeV}$
ByLooseCombinedIsolationDBSumPtCorr	isolation cone of 0.5 , Delta Beta corrected $\sum p_T$ of PF charged and PF γ isolation candidates ($p_T > 0.5\,\mathrm{GeV}$) less than $2\,\mathrm{GeV}$
ByMediumCombinedIsolationDBSumPtCorr	isolation cone of 0.5 , Delta Beta corrected $\sum p_T$ of PF charged and PF γ isolation candidates ($p_T > 0.5\,\mathrm{GeV}$) less than $1\,\mathrm{GeV}$
ByTightCombinedIsolationDBSumPtCorr	isolation cone of 0.5 , Delta Beta corrected $\sum p_T$ of PF charged and PF γ isolation candidates ($p_T > 0.5\,\mathrm{GeV}$) less than $0.8\,\mathrm{GeV}$

Table 6.1: Boolean τ discriminators produced by the HPS algorithm. These discriminators are stored in PXL containers during the skimming process and can be used as cut parameters for the τ selection during the analysis. Hereby, the "Delta Beta" correction is a correction for the contribution of neutral hadrons produced by pile-up (multiple interactions during a bunch crossing) [45][65].

Acknowledgements

An dieser Stelle möchte ich gerne Herrn Professor Hebbeker für die Möglichkeit danken, diese Arbeit am III Physikalischen Institut A schreiben zu dürfen. Die Zusammenarbeit mit Professor Hebbeker empfand ich stets als sehr angenehm und konstruktiv.

Ein besonderer Dank gilt Arnd Meyer, der immer ein offenes Ohr für meine Probleme und Vorschläge hatte. Großer Dank geht an das gesamte MUSiC-Team – Deborah Duchardt, Paul Papacz, Mark Olschewski, Holger Pieta und Johannes Kirchmair. Sie alle haben mich während der Arbeit sehr unterstützt. Ohne Euch wäre diese Arbeit undenkbar – vielen Dank für Eure Betreuung!

Ein Dankeschön geht an Simon Knutzen, der sich mit einem ähnlichen Thema beschäftigt und mein Weiterkommen ebenfalls kräftig unterstützte. Auch möchte ich Sebastian Thüer, Klaas Padeken, Matthias "Klausowitsch" Endres und Johannes Hellmund für ihren Input und die interessanten Diskussionen danken. Markus Merschmeyer und Martin Weber danke ich für die unterhaltsame und unkomplizierte Zusammenarbeit. Außerdem möchte ich Kerstin Hoepfner für viele wertvolle Tipps während der gesamten Masterarbeitszeit danken.

Ein besonderer Dank geht an Adriana del Piero, die Sekretärin unseres Institutes. Sie hat von Anfang an dafür gesorgt, dass ich mich in diesem Institut wohlfühlte – das weiß ich sehr zu schätzen. Vielen Dank, Adriana!

Für fleißiges und unermüdliches Korrekturlesen danke ich Nadja Saborova. Du hast viel mehr Fehler gefunden als mir lieb wäre – danke! Deine ironischen Kommentare haben dafür gesorgt, dass ich selbst beim Korrigieren der scheinbar korrigierten Fehler sehr viel Spaß hatte.

Für die Unterstützung während meines gesamten Studiums möchte ich außerdem vor allem meinen Eltern, meinen beiden Brüdern Alexander und Boris und meiner Schwester Renata danken. Ohne die Musik mit Renata, Alexander und kürzlich auch Boris wäre ich jemand anders. Boris danke ich außerdem für seine Programmiertipps. Last but not least: vielen herzlichen Dank an Michael Novoselsky – solche Freunde sind im Leben sehr selten.

Bibliography

[1] T. Hebbeker, "Skript zur Vorlesung Elementarteilchenphysik". RWTH Aachen, 2010.

[2] A. Stahl, "Skript zur Vorlesung Elementarteilchenphysik I". 2010.

[3] A. Stahl, "Skript zur Vorlesung Elementarteilchenphysik II". 2011.

[4] Wikipedia Collaboration, "Standard Model From Fermi Lab", July, 2012. `http://en.wikipedia.org/wiki/File:Standard_Model_From_Fermi_Lab.jpg`.

[5] Particle data group, "Particles Physics Booklet", July, 2010.

[6] CMS Collaboration, "Observation of a New Particle with a Mass of 125 GeV", July, 2012. `http://cms.web.cern.ch/news/observation-new-particle-mass-125-gev`.

[7] Y. G. Kim and K. Y. Lee, "Early LHC bound on the boson mass in the nonuniversal gauge interaction model", *Physics Letters B* **706** (2012), no. 4–5, 367 – 370. `doi:10.1016/j.physletb.2011.11.032`.

[8] G. Altarelli, B. Mele, and M. Ruiz-Altaba, "Searching for new heavy vector bosons in $p\bar{p}$ colliders", *Z.Phys. C - Particles and Fields* **45** (1989) 109–121. `doi:doi:10.1007/BF01556677`.

[9] P. Langacker, "The physics of heavy Z' gauge bosons", *Rev. Mod. Phys.* **81** (Aug, 2009) 1199–1228. `doi:10.1103/RevModPhys.81.1199`.

[10] L. Evans and P. Bryant, "LHC Machine", *Journal of Instrumentation* **3** (2008), no. 08, S08001. `http://stacks.iop.org/1748-0221/3/i=08/a=S08001`.

[11] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (August, 2008) S08004. `doi:10.1088/1748-0221/3/08/S08004`.

[12] LHC Collaboration, "LHC Design Report". `http://lhc.web.cern.ch/lhc/LHC-DesignReport.html`.

[13] CMS Collaboration, "CMS Luminosity Collision Data 2011", June, 2012. `https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults2011`.

[14] CMS Collaboration, "CMS Physics Technical Design Report Volume I : Detector Performance and Software", volume 1. CMS, 2006.

[15] CMS Collaboration, “Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7 TeV$”, June, 2012. `http://cms.cern.ch/iCMS/jsp/analysis/admin/analysismanagement.jsp?ancode=MUO-10-004`.

[16] CMS Collaboration, “Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and Missing Transverse Energy”, *CMS PAS* **PFT-009-001** (April, 2009).

[17] WLCG Collaboration, “Worldwide LHC Grid Website”. `http://wlcg.web.cern.ch/`.

[18] CMS Collaboration, “CMSSW Application Framework”, September, 2011. `https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookCMSSWFramework`.

[19] CMS Collaboration, “Analysis Overview: an Introduction”, February, 2011. `https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookAnalysisOverviewIntroduction`.

[20] O. Actis, M. Erdmann, A. Hinzmann et al., “Physics eXtension Library (PXL): User’s Guide”, 2008.

[21] O. Actis, M. Erdmann, A. Hinzmann et al., “PXL:Physics eXtension Library”. `http://pxl.sourceforge.net/doxygen3a/` and `https://forge.physik.rwth-aachen.de/projects/pxl/wiki/`.

[22] H. P. Bretz, M. Brodski, M. Erdmann et al., “A development environment for visual physics analysis”, *Journal of Instrumentation* **7** (2012), no. 08, T08005.

[23] CMS Collaboration, “Thirty Seconds Introduction to Generation and Simulation”, August, 2011. `https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookGenIntro`.

[24] “ROOT Website”. `http://root.cern.ch/drupal/`.

[25] CMS Collaboration, “Model Unspecific Search for New Physics in $p\bar{p}$ Collisions at $\sqrt{s} = 7\,\mathrm{TeV}$”, *CMS PAS* **EXO-10-021** (2011). `http://cms-physics.web.cern.ch/cms-physics/public/EXO-10-021-pas.pdf`.

[26] P. Biallass, “Model Independent Search for Deviations from the Standard Model at the Tevatron: Final States with Missing Energy”, December, 2004. `http://web.physik.rwth-aachen.de/~hebbeker/theses/biallass_diploma.pdf`.

[27] J. H. Field, “Model-independent analysis of CERN LEP and SLAC SLD data on Z decays: Is the standard model confirmed?”, *Phys. Rev. D* **58** (Oct, 1998) 093010. `doi:10.1103/PhysRevD.58.093010`.

[28] D0 Collaboration, “Quasi-Model-Independent Search for New High p_T Physics at D0”, *Phys. Rev. Lett.* **86** (Apr, 2001) 3712–3717. `doi:10.1103/PhysRevLett.86.3712`.

[29] G. Choudalakis, “Model Independent Search For New Physics At The Tevatron”. PhD thesis, MIT, May, 2008. `http://arxiv.org/abs/0805.3954v1`.

[30] A. Aktas, V. Andreev, T. Anthonis et al., "A general search for new phenomena in ep scattering at HERA", *Physics Letters B* **602** (2004), no. 1ñ2, 14 – 30. `doi:10.1016/j.physletb.2004.09.057`.

[31] H. Pieta, "MUSiC - a Model Unspecific Search in CMS based on 2010 LHC data.". PhD thesis, RWTH Aachen University, 2012.

[32] M. Olschewski, "Study of Alternative Statistical Methods for a Model Unspecific Search in CMS, Diplomarbeit", July, 2011. `http://web.physik.rwth-aachen.de/~hebbeker/theses/olschewski_diploma.pdf`.

[33] S. A. Schmitz, "Model Unspecific Search for New Physics with High p_T Photons in CMS, Diplomarbeit", October, 2009. `http://web.physik.rwth-aachen.de/~hebbeker/theses/schmitz_diploma.pdf`.

[34] E. Dietz-Laursonn, "Model Unspecific Search for New Physics with b-Hadrons in CMS, Diplomarbeit", October, 2010. `http://web.physik.rwth-aachen.de/~hebbeker/theses/dietz-laursonn_diploma.pdf`.

[35] S. A. Schmitz, "Model Unspecific Search for New Physics with High p_T Photons in CMS, Diplomarbeit", October, 2009. Modified Figure; `http://web.physik.rwth-aachen.de/~hebbeker/theses/schmitz_diploma.pdf`.

[36] CMS Collaboration, "Absolute Calibration of the Luminosity Measurement at CMS: Winter 2012 Update",. `http://cdsweb.cern.ch/record/1434360/`.

[37] D. Stump, J. Huston, J. Pumplin et al., "Inclusive jet production, parton distributions, and the search for new physics", *Journal of High Energy Physics* **2003** (2003), no. 10, 046. `http://stacks.iop.org/1126-6708/2003/i=10/a=046`.

[38] J. Pumplin, D. R. Stump, J. Huston et al., "New Generation of Parton Distributions with Uncertainties from Global QCD Analysis", *Journal of High Energy Physics* **2002** (2002), no. 07, 012. `http://stacks.iop.org/1126-6708/2002/i=07/a=012`.

[39] D. Stump, J. Huston, J. Pumplin et al., "Inclusive jet production, parton distributions, and the search for new physics", *JHEP* **0310** (2003) 046, `arXiv:hep-ph/0303013`.

[40] CMS Collaboration, "Measurements of Inclusive W and Z Cross Sections in $p\bar{p}$ Collisions at $\sqrt{s} = 7\,\text{TeV}$", *arXiv:1012.2466* (January, 2011). `http://arxiv.org/pdf/1012.2466.pdf`.

[41] CMS Collaboration, "Performance of τ-lepton reconstruction and identification in CMS", *CMS Paper Draft* **TAU-11-001v29** (December, 2011). `http://cms.cern.ch/iCMS/analysisadmin/analysismanagement?awg=TAU&awgyear=2011&dataset=All`.

[42] M. Cacciari, G. P. Salam, and G. Soyez, "The anti- k t jet clustering algorithm", *Journal of*

High Energy Physics **2008** (2008), no. 04, 063.
`http://stacks.iop.org/1126-6708/2008/i=04/a=063`.

[43] C. Veelken, "CMS Tau Physics Object Group approval Talk and private communication. Modified Figure.", 2012. `https://indico.cern.ch/getFile.py/access?contribId=3&resId=0&materialId=slides&confId=172430`.

[44] CMS Collaboration, C. Calabria, "Tau trigger, reconstruction and identification at CMS", September, 2012.
`https://hypernews.cern.ch/HyperNews/CMS/get/AUX/2012/09/09/11:07:15-23577-slides_Tau2012_v1.pdf`.

[45] CMS Collaboration, "Software Guide for Tau Reconstruction", May, 2012.
`https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuidePFTauID`.

[46] CMS Collaboration, "TauID: recommendation from the Tau POG", June, 2012.
`https://twiki.cern.ch/twiki/bin/viewauth/CMS/TauIDRecommendation`.

[47] J. Alwall, M. Herquet, F. Maltoni et al., "MadGraph 5: going beyond", *Journal of High Energy Physics* **2011** (2011) 1–40. 10.1007/JHEP06(2011)128.

[48] P. Nason, "A New method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* **0411** (2004) 040, `arXiv:hep-ph/0409146`. `doi:10.1088/1126-6708/2004/11/040`.

[49] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method", *JHEP* **0711** (2007) 070, `arXiv:0709.2092`. `doi:10.1088/1126-6708/2007/11/070`.

[50] S. Alioli, P. Nason, C. Oleari et al., "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **1006** (2010) 043, `arXiv:1002.2581`. `doi:10.1007/JHEP06(2010)043`.

[51] S. M. Torbjörn Sjöstrand and P. Skands, "PYTHIA 6.4 physics and manual", *JHEP05* **2006** (2006) 026. `doi:10.1088/1126-6708/2006/05/026`.

[52] CMS Collaboration, "Interactive Slice of the CMS detector". `https://cms-docdb.cern.ch/cgi-bin/PublicDocDB/ShowDocument?docid=4172`.

[53] G. Abbiendi, N. Adam, J. Alcaraz et al., "Muon Reconstruction in the CMS Detector", Technical Report CMS AN-2008/097, 2008. `http://cms.cern.ch/iCMS/jsp/openfile.jsp?tp=draft&files=AN2008_097_v4.pdf`.

[54] S. Baffioni, C. Charlot, F.Ferri et al., "Electron reconstruction in CMS", Technical Report CMS-NOTE-2006-040, 2007. `http://hal.in2p3.fr/in2p3-00140484/`.

[55] E. Meschi, T. Monteiro, C. Seez et al., "Electron Reconstruction in the CMS Electromagnetic Calorimeter", Technical Report CMS-NOTE-2001-034, CERN, Geneva, June, 2001.

[56] J. Nysten, “Photon Reconstruction in CMS”, *CMS CR* **004** (2004).

[57] A. Attikis, R. Kinnunen, and M. Kortelainen, “Search for light charged Higgs bosons with the $H^{\pm} \rightarrow \tau^{\pm}\nu_{\tau}$ decay in the fully hadronic final state”, *CMS AN* **AN-11-470** (2011).

[58] M.Bachtis, S.Dasu, and A.Savin, “Prospects for measurement of $\sigma(pp \rightarrow Z) \cdot \beta(Z \rightarrow \tau^{+}\tau^{-})$ with CMS in $p\bar{p}$ Collisions at $\sqrt{s} = 7\,\text{TeV}$”, *CMS AN* **AN-2010/082** (April, 2010). `http://cms.cern.ch/iCMS/jsp/openfile.jsp?tp=draft&files=AN2010_082_v1.pdf`.

[59] M.Bachtis, J.Swanson, A.Savin et al., “Study of the di-tau spectrum using muon-tau and electron-tau final states with CMS detector at $\sqrt{s} = 7\,\text{TeV}$”, *CMS AN* **AN-10-387** (February, 2011). `http://cms.cern.ch/iCMS/jsp/openfile.jsp?tp=draft&files=AN2010_387_v4.pdf`.

[60] J. Schulz, “Suche nach resonanter τ -Sneutrino Produktion im $\mu\tau$ Endzustand mit CMS bei $\sqrt{s} = 8$ TeV”, August, 2012. Bachelor Thesis.

[61] K. Hoepfner, C. Leonidopoulos, S. Thüer et al., “Search for a heavy gauge boson W′ in the leptonic final states with W-W′ interference in pp collisions at $\sqrt{s} = 7 TeV$”, *CMS AN* **AN-10-227** (2011).

[62] S. Knutzen, “Study of the $W' \rightarrow \tau\nu_{\tau}$, Thesis in progress”, Master’s thesis, RWTH, 2012.

[63] CMS Collaboration, “Search for W′ in the leptonic channels in $p\bar{p}$ Collisions at $\sqrt{s} = 7$ TeV”, *CMS AN* **EXO-11-024** (2011). `http://cms-physics.web.cern.ch/cms-physics/public/EXO-11-024-pas.pdf`.

[64] CMS Higgs to Tau Tau Group, “Search for Higgs Bosons Decaying to Tau Pairs in $p\bar{p}$ Collisions at $\sqrt{s} = 7$ TeV”, *CMS AN* **AN-2011/390** (December, 2011). `http://cms.cern.ch/iCMS/jsp/openfile.jsp?tp=draft&files=AN2011_390_v8.pdf`.

[65] CMS Collaboration, “SWGuideMuonId”. `https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMuonId#Muon_Isolation`.

Hiermit versichere ich, Michael Brodski, dass ich die vorliegende Arbeit selbstständig verfasst und nur die angegebenen Quellen und Hilfsmittel benutzt habe.

Aachen, den 25.09.2012