Vrije Universiteit Brussel

Faculteit Wetenschappen
Departement Natuurkunde

# Top Quark physics with B-Jets containing muons with CMS

**Stijn Blyweert**

Promotor: Prof. Dr. Jorgen D'Hondt
Academiejaar: 2008 – 2009

Proefschrift ingediend met het oog op het behalen
van de academische graad Master in de Fysica

# Contents

# Introduction

The Standard Model of particle physics is a relativistic quantum field theory which gives a description of the behaviour of all known subatomic particles. In the past decades it has been extensively tested with very high precision by different experiments.

Further tests of this theory will be performed at the Large Hadron Collider (LHC) which will become operational in the fall of 2009. The LHC is a proton-proton collider which is constructed in the former LEP tunnel of the CERN laboratory near Geneva. In its first year of operation, it will provide collisions at a centre-of-mass energy of about 10 TeV and afterwards, the energy will be increased to 14 TeV. One of the two general purpose experiments is the Compact Muon Solenoid (CMS). This large particle detector will be used to improve our knowledge about particle physics at the TeV-scale and try to solve some fundamental problems of the Standard Model. One of the main goals of the CMS experiment is the detailed observation of the properties of the heaviest quark, the top quark.

In this thesis, events containing two top quarks will be used to study b-jets where a muon is produced in the decay. In the first chapter, a general introduction to the Standard Model, the Large Hadron Collider and the CMS experiment will be given. The second chapter contains some more detailed information about top quark physics and how this is done with the Compact Muon Solenoid detector. In the third and the fourth chapter, these b-jets with a muon will be used to provide a better measurement of the energy of the b-quark and the mass of the top quark. The fifth chapter contains an event selection which uses these b-jets containing a muon. Finally, the conclusions from this thesis can be found in the sixth and last chapter.

# Dankwoord

Eerst en vooral dank aan mijn promotor, Prof. Dr. Jorgen D'Hondt, voor de massa's de tijd en moeite die hij heeft gestoken in hulp en feedback voor mijn thesis. Zonder hem had ik misschien nooit kennis gemaakt met het onderzoek binnen de experimentele deeltjesfysica. Verder wil ik ook de doctoraatsstudenten en de post-docs van het "Top Quark Team" bedanken, vooral Joris en Petra, voor de hulp die steeds ter beschikking stond (en voor de productie van de data samples). Natuurlijk ook bedankt aan mijn thesisgenoot Michael voor de `C++` en CMSSW hulp en ook bedankt voor de vruchtbare discussie's.

Ook bedankt aan alle mensen die voor de nodige schouderklopjes zorgden, op alle momenten van de dag. Bedankt ook aan Gregory voor het nalezen en verbeteren van mijn thesis. Ten slotte wil ik ook mijn ouders bedanken, eerst en vooral om mij de kans te geven om te studeren wat ik wou en natuurlijk ook omdat ik zonder hun steun dit nooit tot een goed einde had kunnen brengen.

# Chapter 1

# Particle physics with CMS

Elementary particle physics is the branch of physics where people aim to understand the smallest constituents of matter. Almost forty years ago, theorists came up with a model which nowadays still describes all the known experimental facts about elementary particle physics. This model is known as the Standard Model of particle physics. Many experiments are making precision measurements of the Standard Model or are searching for physics beyond the Standard Model. The Compact Muon Solenoid detector (CMS), located at the Large Hadron Collider (LHC) at CERN near Geneva, is combining both.

One of the particles of the Standard Model was only discovered 14 years ago: the top quark. One of the main goals of the CMS experiment is to measure the top quark properties in detail. With a precision beyond the current one.

## 1.1 The Standard Model

The description of the smallest constituents of matter and their interactions is nowadays done by the Standard Model [1, 2]. This is a quantum field gauge theory describing all the known particles and their interactions.

### 1.1.1 The particles and forces of the Standard Model

All the visible matter in the universe is, according to the Standard Model, built from only 12 elementary particles, which are fermions[1] with spin 1/2. Each of these 12 particles also has its anti-particle, which has the same mass but opposite electrical charge. As can be seen from Table 1.1, the 12 mat-

---

[1] Fermions are particles with half-integer spin.

| Generation | Quarks | |
|---|---|---|
| | Name | Mass |
| 1st | up $u$<br>down $d$ | $2.55^{+0.75}_{-1.05}$ MeV/c$^2$<br>$5.04^{+0.96}_{-1.54}$ MeV/c$^2$ |
| 2nd | charm $c$<br>strange $s$ | $1.27^{+0.07}_{-0.11}$ GeV/c$^2$<br>$105^{+25}_{-35}$ MeV/c$^2$ |
| 3rd | top $t$<br>bottom $b$ | $173.1 \pm 2.2$ GeV/c$^2$<br>$4.20^{+0.17}_{-0.07}$ GeV/c$^2$ |

| Generation | Leptons | |
|---|---|---|
| | Name | Mass |
| 1st | electron neutrino $\nu_e$<br>electron $e$ | $< 2$ eV/c$^2$ (95 % C.L.)<br>$0.510998910 \pm 0.000000013$ MeV/c$^2$ |
| 2nd | muon neutrino $\nu_\mu$<br>muon $\mu$ | $< 0.19$ MeV/c$^2$ (90 % C.L.)<br>$105.6583668 \pm 0.0000038$ MeV/c$^2$ |
| 3rd | tau neutrino $\nu_\tau$<br>tau $\tau$ | $< 18.2$ MeV/c$^2$ (95 % C.L.)<br>$1776.84 \pm 0.17$ MeV/c$^2$ |

Table 1.1: The 3 generations of fermions of the Standard Model and their measured masses [3, 4].

ter particles can be classified into three generations. The particles of the first generation are the building blocks of all the matter around us. The up and down quarks combine to form the protons and the neutrons, which also combine to make all the nuclei of the chemical elements. These nuclei together with the electrons, make all the atoms. The electron neutrino can be observed in the $\beta$-decay of some nuclei. The particles of the other two generations have the same properties as those from the first generation, except for their mass, which is higher for the second and the third generation. Apart from the matter particles, the Standard Model also describes the interactions between these matter particles. There are four known interactions: electromagnetism-magnetism, the strong interaction, the weak interaction and gravity. The Standard Model describes all these interactions, except gravity, because gravity is extremely weak compared to the other 3 interactions at small distances, so its effects are immeasurable small. According to the Standard Model, these interactions happen through the exchange of gauge bosons[2] with spin 1. An overview of the forces of the Standard Model together with the corresponding bosons and their mass can be found in Table 1.2.

| Force | Force carrier | Mass (GeV/c$^2$) |
|---|---|---|
| electromagnetism-magnetism | Photon $\gamma$ | 0 |
| Weak | $W^{\pm}$ | $80.399 \pm 0.025$ |
| | $Z^0$ | $91.1876 \pm 0.0021$ |
| Strong | eight gluons $g$ | 0 |

Table 1.2: The bosons of the Standard Model, responsible for carrying the fundamental forces, and their masses [3, 4]. The masses of the gluons and the photon are the masses which are predicted by the Standard Model.

The 12 matter particles of the Standard Model can also be divided into two other groups, based on their interactions: the leptons and the quarks. The leptons are insensitive to the strong force. The uncharged leptons (the neutrino's) are also insensitive to the electromagnetic-magnetic interaction. The quarks, on the other hand, are sensitive to all the interactions.

There also exists another way of dividing the fermions of the Standard Model into two groups: by their chirality. The chirality of a particle (left-handed or right-handed) depends on the way it transforms under the Poincaré group[3]. Chirality is an important property because experiments pointed out that the weak interaction only acts on left-handed fermions and not on right-handed fermions.

The Standard Model predicts the existence of another particle: the Higgs boson. It is a massive, electrically neutral spin 0 particle and its existence was predicted by the mechanism of spontaneous symmetry breaking. The idea of spontaneous symmetry breaking was first worked out by R. Brout, F. Englert and P. W. Higgs [5, 6] and it is a mechanism which gives an explanation for the masses of the weak gauge bosons and of the fermions. Despite many searches, there is still no experimental evidence of the existence of the Higgs boson. The search for the Higgs boson is one of the biggest motivations for the construction of the LHC.

---

[2] Bosons are particles with integer spin.

[3] The Poincaré group is the group of all the transformations in Minkowski space: translations, rotations and boosts.

### 1.1.2 Mathematical description of the Standard Model

The Standard Model is a quantum field theory based on the principle of local gauge invariance. This means it is a combination of classical field theory and quantum mechanics. Classical field theory is based on Hamilton's principle of least action. The action is defined as $S = \int \mathcal{L}\left(\psi(x), \partial^\mu \psi(x)\right) d^4x$. Here, $\mathcal{L}$ is the Lagrangian density, which depends on the particle's field wave functions $\psi(x)$ and its first derivatives $\partial^\mu \psi(x)$. Hamilton's principle states that the action $S$ is stationary for variations of the fields ($\delta S = 0$) and this requirement gives rise to the equations of motion, the so-called Euler-Lagrange equations.

The fermions of the Standard Model are described by Dirac spinors $\psi(x)$ and the Dirac Lagrangian:

$$\mathcal{L}^{dirac} = i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi \tag{1.1}$$

Here, $\gamma^\mu$ are the Dirac matrices[4] and $m$ is the fermion mass. The Lagrangian in Eq. (1.1) only describes freely moving fermions. In order to include interactions, the Lagrangian must be made invariant under local gauge transformations. A local gauge transformation with rotation parameters $\vec{\epsilon}(x)$ and characterized by the generators $\vec{\tau}$ of a Lie-group is given by

$$\psi'(x) = U\psi(x) = e^{i\vec{\epsilon}(x)\cdot\frac{\vec{\tau}}{2}}\ \psi(x) \tag{1.2}$$

It is clear that quantum mechanical observables, who only depend on $|\psi|^2$ are invariant under these transformations. Lagrangians, however, are in general not invariant and to restore the invariance, the normal partial derivative $\partial_\mu$ needs to be replaced by a co-variant derivative:

$$\mathcal{D}_\mu = \partial_\mu - ig\frac{\vec{\tau}}{2}\cdot\vec{A}_\mu \tag{1.3}$$

In this equation, $\vec{A}_\mu$ is a new interacting vector field and $g$ is the coupling constant, representing the interaction strength. When this is applied to the Dirac Lagrangian of Eq. (1.1), it gives

$$\begin{aligned}\mathcal{L}^{dirac} &= i\bar{\psi}\gamma^\mu \mathcal{D}_\mu \psi - m\bar{\psi}\psi \\ &= i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi - ig\bar{\psi}\gamma^\mu\frac{\vec{\tau}}{2}\cdot\vec{A}_\mu\psi \end{aligned} \tag{1.4}$$

The last term in this equation represents the coupling between the fermion field and the new interacting vector field. For the Dirac Lagrangian to be

---

[4] Dirac matrices are $4 \times 4$ matrices which obey: $\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}$ and $(\gamma^\mu)^\dagger = \gamma^0\gamma^\mu\gamma^0$.

invariant under the transformation in Eq. (1.1), the covariant derivative needs to transform according to $\mathcal{D}'_\mu\psi' = U(\mathcal{D}_\mu\psi)$. From this one can derive the transformation rules of the interacting vector fields under the local gauge transformation.

When all this is applied to the group $SU(2)_L \times U(1)_Y$, the electroweak Lagrangian of the Standard Model follows. The subscript $L$ means that the $SU(2)$ transformations only act on left-handed spinors[5]. Since the left-handed and the right-handed spinors both transform in a different way under $SU(2)_L \times U(1)_Y$, mass terms like in Eq. (1.1) are not invariant under $SU(2)_L \times U(1)_Y$ and to preserve the invariance, they need to be dropped by setting all the masses equal to zero. The fact that they are not invariant can be seen when the mass term is rewritten in terms of left- and right-handed spinors:

$$m\bar{\psi}\psi = m\left(\bar{\psi}_L\psi_L + \bar{\psi}_R\psi_R + \bar{\psi}_R\psi_L + \bar{\psi}_L\psi_R\right) \tag{1.5}$$

The last two terms are clearly not invariant. The mass of the gauge bosons and the fermions will be introduced by means of the Brout-Englert-Higgs mechanism.

For the strong interaction, the Lagrangian again needs to be invariant under the local gauge transformations of a group. This time the group is $SU(3)_C$ and the resulting theory is called QCD or Quantum Chromodynamics. A remarkable property of QCD is that the quarks (the only particles which interact strongly) can only be observed in bound states called hadrons.

The next thing that needs to be done, is the inclusion of mass for the weak gauge bosons. As already mentioned, this is done by the Brout-Englert-Higgs mechanism, which is based on the idea of spontaneous symmetry breaking, according to which the vacuum state of a system does not possess the same symmetry as the Lagrangian density. The Brout-Englert-Higgs mechanism predicts the spontaneous symmetry breaking of the $SU(2)_L \times U(1)_Y$ symmetry and it predicts the existence of an extra scalar doublet on top of the already existing spinors and vector fields. The weak gauge bosons and the fermions acquire their mass through their interaction with the Higgs boson field. The theory also has another very important prediction: the existence of a new particle, the Higgs boson, but the theory does not predict the mass of the Higgs boson. An estimate of the mass of the Higgs boson can be found by performing a global fit on al electroweak data [4]. The resulting $\Delta\chi^2$ of this fit can be found in Figure 1.1. The yellow bands are the regions which are already excluded by direct searches. This fit predicts a value for the mass of the Higgs boson of $90^{+36}_{-27}$ GeV/c$^2$.

---

[5] Left-handed spinors are defined as $\psi_L = \frac{1}{2}(1-\gamma^5)\ \psi$ and right-handed as $\psi_R = \frac{1}{2}(1+\gamma^5)\ \psi$. Here $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$.

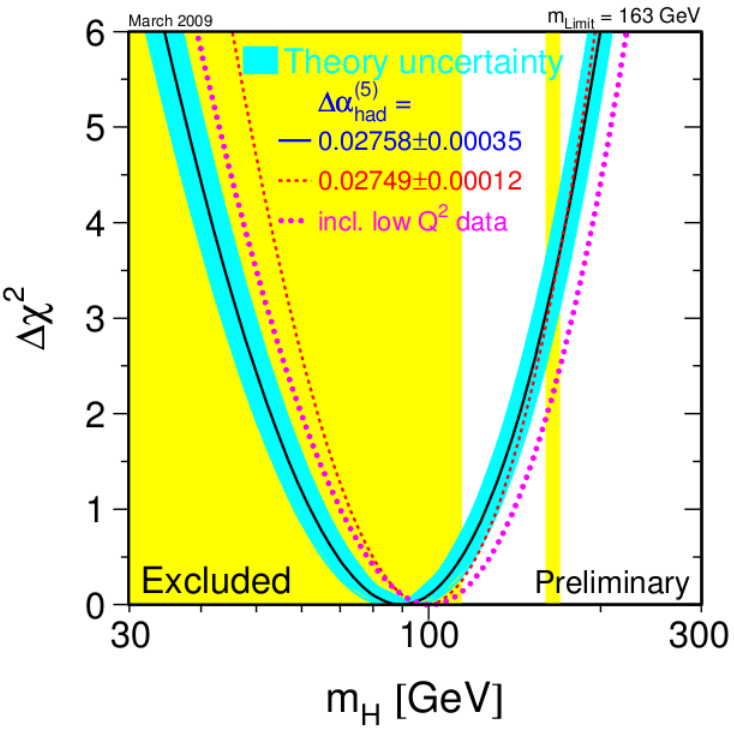

Figure 1.1: The $\Delta\chi^2 = \chi^2 - \chi^2_{min}$ curve in function of the mass of the Higgs boson after a global fit to all the electroweak data.

Some final words need to be told about deriving predictions and calculations from this beautiful theory. Calculations are done by means of Feynman diagrams and Feynman rules. This is a diagrammatic approach to calculate probabilities from certain processes. When performing these calculations, quantum-mechanical corrections need to be incorporated and this done by the inclusion of extra loops and vertices in the Feynman diagrams. By ordering those diagrams as a function of the number of vertices, a series is formed where each higher order diagram gives a smaller correction than the previous one. The problem is that except the lowest order diagrams, all diagrams give rise to divergences which can always be absorbed into unobservable, bare parameters, but they make the calculations extremely difficult.

### 1.1.3 Shortcomings and possible extensions of the Standard Model

Although it has proven to be a very successful and precise theory about the interactions of fundamental particles, the Standard Model is not a complete theory of nature. There are several reasons for this, for example:

- Gravity is not included in the Standard Model.
- Why are there three generations of fermions?
- Astrophysical observations have proved the existence of dark matter [7]. Other observations also excluded the Standard Model Particles as good dark matter candidates.
- Why is there a disequilibrium of matter and anti-matter in the universe?

There are many good candidates for solving some of the problems of the Standard Model. The most popular ones are theories with Supersymmetry (SUSY) [8]. This is an extra symmetry between bosons and fermions and has as a consequence that for each fermion in nature, there exists a bosonic partner and vice versa. One of the interesting aspects of SUSY is that some theories with SUSY predict the existence of the so-called Lightest Supersymmetric Particle (LSP), which is a good candidate for dark matter.

Other extensions predict the existence of extra dimensions [9], other gauge bosons [10], ...

## 1.2 The Large Hadron Collider

The Large Hadron Collider (LHC) [11, 12] is a particle accelerator located at CERN (the European Organization for Nuclear Research), near Geneva, which will produce proton-proton and heavy ion (Pb-Pb) collisions. It is build in the same tunnel as the former LEP electron-positron accelerator. This tunnel has a circumference of about 26.7 km and lies between 45 m and 170 m underground between the Lake of Geneva and the Jura Mountains. The construction is finished and on the $10^{th}$ of September 2008, the first proton beams were successfully circulated through the accelerator. However, on the $19^{th}$ of September 2008 a serious fault developed and damaged a number of superconducting magnets. Those magnets need to be replaced and the LHC will not see any beam before September 2009.

When the LHC is operational, it will collide protons at a total centre-of-mass energy of 14 TeV. During the first year of operation, however, the centre-of-mass energy will be about 10 TeV. Even with this lower energy, the LHC will still be the world's most energetic particle collider, replacing Fermilab's Tevatron $p$-$\bar{p}$ collider [13], which has a centre-of-mass energy of 1.96 TeV.

### 1.2.1 Physics motivation

As already mentioned, the main motivation to build the LHC is the search for the Standard Model Higgs boson. With a luminosity of $2 \times 10^{33} cm^{-2} s^{-1}$, around $10^5$ Higgs bosons are expected to be produced every year. This is two orders of magnitude smaller than the expected $t\bar{t}$ production rate. Even when the Higgs boson is not found at the LHC, this is an important discovery. We know that the weak gauge bosons are massive, hence when the Higgs boson is not found, there has to be another mechanism responsible for their masses.

Also the search to Supersymmetry and other extensions of the Standard Model, as already mentioned in Section 1.1.3, is one of the important motivations for the LHC's construction. There are many theories who predict a lot of interesting things to happen in the energy range of the LHC. Even without the discovery of new particles, the outcoming of the LHC will learn us a lot about particle physics.

Another very important task is the confirmation and better understanding of the already known physics. For example, the LHC will be the first collider with such a high top quark production rate and this will allow a very precise measurement of the top quark's properties.

The LHC will also provide high energy Pb-Pb collisions. The aim of these collisions is the observation of a new state of matter: the Quark-Gluon plasma. This is a deconfinement of quarks and gluons and is predicted by QCD at very high temperature. An overview of the current knowledge in the research of the Quark-Gluon plasma can be found in [14]. The current experimental observations might be due to the formation of a Quark-Gluon plasma, but further research is necessary to confirm the discovery of the Quark-Gluon plasma.

### 1.2.2 Design of the accelerator

Like in every high energy particle collider, the protons used in the LHC are first accelerated in lower energy accelerators. The CERN accelerator complex is shown in Figure 1.2. The first acceleration is provided by the LINAC

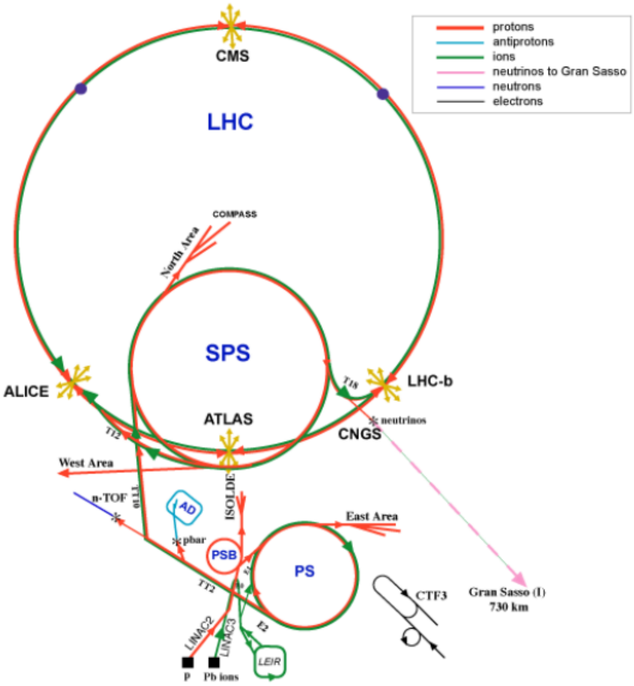

Figure 1.2: The accelerator complex of CERN.

2, which will deliver 50 MeV protons. The next step is the Proton Synchrotron Booster and this will deliver beams with an energy of 1.4 GeV. The Proton Synchrotron (PS) will further increase the energy until the particles reach an energy of 26 GeV and are ready for injection into the Super Proton Synchrotron (SPS). This last pre-accelerator will accelerate the protons until they are ready for injection into the LHC. The 450 GeV beams of the SPS are then further accelerated in the LHC until they reach their final energy of 7 TeV.

The LHC, unlike a lot of other colliders, has two separate beam pipes. One for the circulating beam in each direction. This is because the LHC needs to circulate the same particles in each direction, hence the dipoles need to have a magnetic field in opposite directions for each beam. The LHC consists of 1232 superconducting dipoles and over 2500 other magnets. The design of these dipoles is illustrated in Figure 1.3. They will operate at a temperature of 1.9 K and will produce a magnetic field of 8.33 T when the LHC beams circulate with an energy of 7 TeV. When the LHC will run at its design luminosity of $10^{34}$ cm$^{-2}$ s$^{-1}$, 2 808 bunches of about $10^{11}$ protons each will circulate in each direction in the beam pipes. The spacing between the different bunches will be 25 ns. An overview of the main design parameters for the use as a proton collider can be found in Table 1.3.

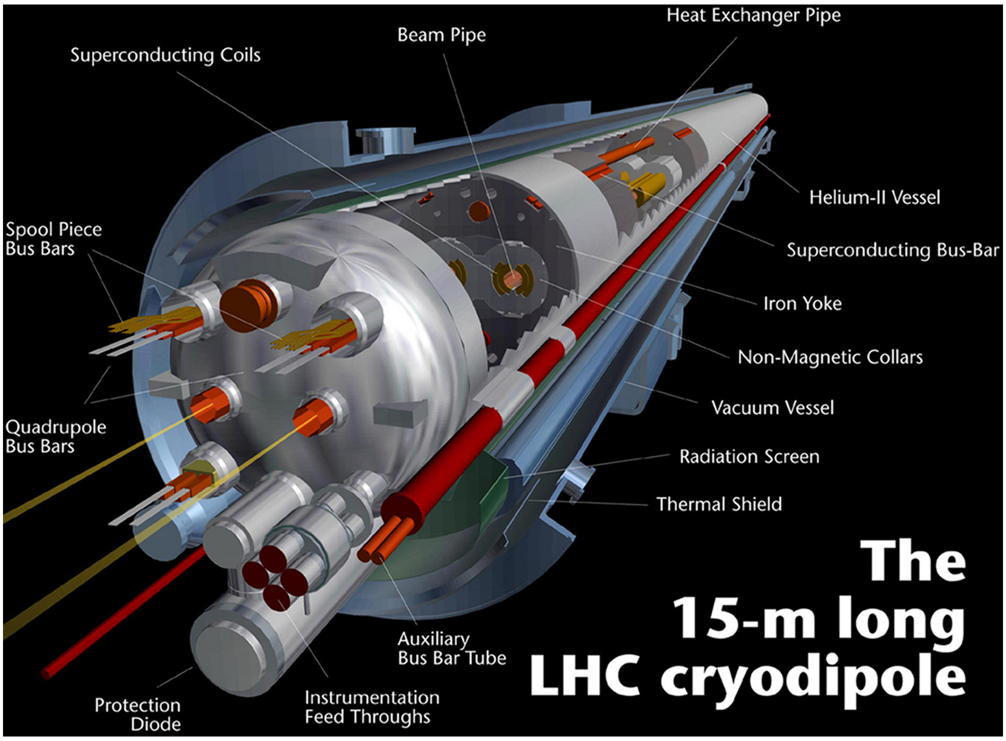

Figure 1.3: Overview of a LHC dipole.

### 1.2.3 Experiments at the LHC

The LHC will provide proton-proton and heavy ion collisions at four locations. Around these four locations, six experiments are constructed. The location of the different experiments along the LHC-ring is shown in Figure 1.2. The two biggest experiments, CMS [15, 16] and ATLAS [17], are so-called general purpose detectors. They are designed to be able to detect almost all of the predicted signals of many theories beyond the Standard Model. Of course, they will also be used to improve our understanding of the Standard Model by, for example, a precise measurement of the mass of the top quark. Both detectors will also be used during the heavy ion collisions.

The ALICE detector [18] is a specialized detector dedicated to the measurement of the heavy ion collisions. The other big specialized detector is the LHC-b detector [19]. As the name already suggests, this experiment will look at b-physics. CP violation and rare decays of particles containing b-quarks, for example.

| Parameter | Design value |
|---|---|
| Circumference | 26.7 km |
| Design Luminosity | $10^{34}$ cm$^{-2}$ s$^{-1}$ |
| Average number of inelastic collisions per bunch crossing at design luminosity | $\approx$ 22 |
| Beam lifetime at design luminosity | $\approx$ 14.9 h |
| Dipole magnetic field at 7 TeV | 8.33 T |
| Magnet temperature | 1.9 K |
| Injection beam energy | 450 GeV |
| Acceleration time | $\approx$ 20 min |
| Beam energy at collision | 7 TeV |
| Number of bunches in each beam | 2808 |
| Number of protons per bunch | $10^{11}$ |
| Bunch spacing | 25 ns |

Table 1.3: Some of the important parameters of the LHC in the proton-mode.

Although they are not shown on Figure 1.2, there are also two much smaller experiments conducted at the LHC. The first one is TOTEM [20]. It is located at the same interaction point as the much bigger CMS detector. The second one, the LHC-f experiment [21], is located at the interaction point of ATLAS. Both experiments are looking at physics in the forward region of the proton-proton collisions. TOTEM will focus on the measurement of the total cross section of the proton-proton collisions and on elastic scattering, while LHC-f will high energy photons and neutrons which are produced in the very forward region of the proton-proton collisions to study the validity of nuclear interaction models used in Monte Carlo simulations of air showers induced by ultra-high energy cosmic-rays.

## 1.3 The Compact Muon Solenoid

The CMS (Compact Muon Solenoid) experiment is one of the two general purpose detectors which are constructed near the LHC. As the name already suggests, it is more or less compact (ATLAS is more than two times bigger in length and in diameter), it has a very good muon system and it has a solenoidal magnetic field with a strength of about 3.8 T. Furthermore, it has an outstanding ECAL (Electromagnetic calorimeter) energy resolution, for the possible detection of the Higgs boson decaying to two photons, and a hermetic calorimetry system, for the detection of non-interacting high energetic particles. It also has a full-silicon-based inner tracking system. An overview of the CMS detector and the different subsystems can be found in Figure 1.4. The detector has a cylindrical shape around the beam pipe. It has a length of 21.6 m, a diameter of 14.6 m and a total weight of 12 500 tons and it consists of 3 main parts: the barrel and the two endcaps. The different sub-detectors also have a barrel-part and two endcap-parts.

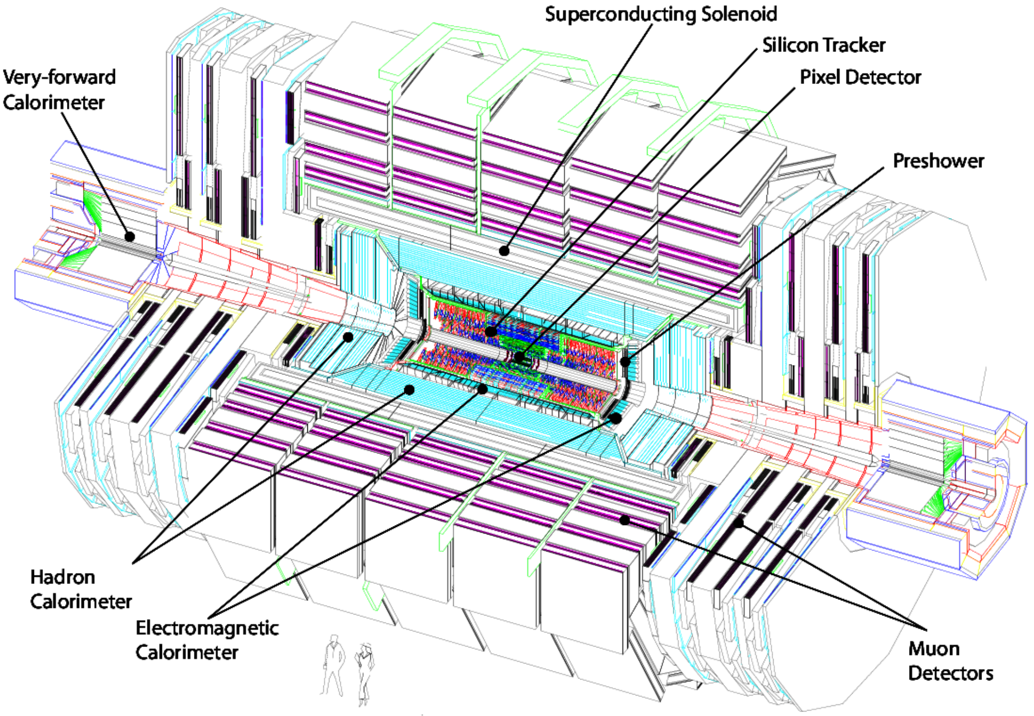

Figure 1.4: Overview of the CMS detector.

This thesis will use the standard CMS coordinate system. This coordinate system has the $z$-axis pointing along the beam in the direction of the Jura mountains. The $y$-axis is pointing upwards and the $x$-axis is pointing towards the center of the LHC ring. The azimuthal angle $\phi$ is measured from the $x$-axis in the $x$-$y$ plane. The polar angle $\theta$ is measured from the $z$-axis. The pseudorapidity, defined as $\eta = -\ln\tan(\theta/2)$, is another useful quantity is because its numerical value for a particle is a good approximation of the particle's rapidity when the particle is relativistic.

### 1.3.1 The Magnet System

One of the most striking parts of the CMS detector is the big solenoidal magnet which has an inner diameter of 5.9 m and a length of 12.9 m. It is made of reinforced NbTi superconductor which is cooled down to a operational temperature of 4.5 K. These conductors will house a current of 19.5 kA to reach the design value of 4 T of the magnetic field. With this high field, the magnet will store an energy of 2.7 GJ. The whole magnet system has a cold mass of 220 tons. When the CMS detector is operational, the solenoid will produce a field of 3.8 T instead of 4 T to improve the lifetime of the solenoid.

The flux of the magnetic field is returned through a 10 000 ton return yoke surrounding the solenoid. This return yoke is interleaved with the muon stations.

### 1.3.2 The Silicon Tracker

This part of CMS is the one which is closest to the interaction point. The main goal is the reconstruction of the tracks of all charged particles emerging from the collisions. Later on, these tracks can be used to reconstruct the momentum of the particles and all the vertices. In order to ensure an efficient pattern recognition, it has been designed to keep the occupancy of the silicon sensor channels small. Based on this occupancy (or on the particle flux at various radii), three regions can be distinguished, each with a different type of silicon tracking detector:

- Closest to the interaction vertex where the particle flux is the highest ($\approx 10^7$/s at $r \approx 10$ cm), pixel detectors are placed. The size of each pixel is $\approx 100 \times 150\ \mu\text{m}^2$, so this gives an occupancy of about $10^{-4}$ per pixel per LHC crossing.

- In the intermediate region ($20 < r < 55$ cm), the particle flux is low enough to use silicon microstrip detectors with a minimum cell size of 10 cm × 80 $\mu$m. This leads to an occupancy of 2–3 % per LHC crossing.

- In the outermost region ($r > 55$ cm), the particle flux has dropped sufficiently to allow the use of larger silicon microstrips with a maximum cell size of 25 cm × 180 $\mu$m, while keeping the occupancy to $\approx 1$ %.

In order to achieve such a small occupancy, the CMS tracker consists of 20 000 silicon sensors which have altogether a surface of 210 m$^2$. It is the largest silicon tracker ever build, having a diameter of 2.4 m and a length of 5.4 m. It has a pseudorapidity coverage of $|\eta| < 2.5$. A graphical overview of the tracker can be found in Figure 1.5. It shows the cross section of one fourth of the silicon tracker. The location of the different silicon sensors can clearly be seen.

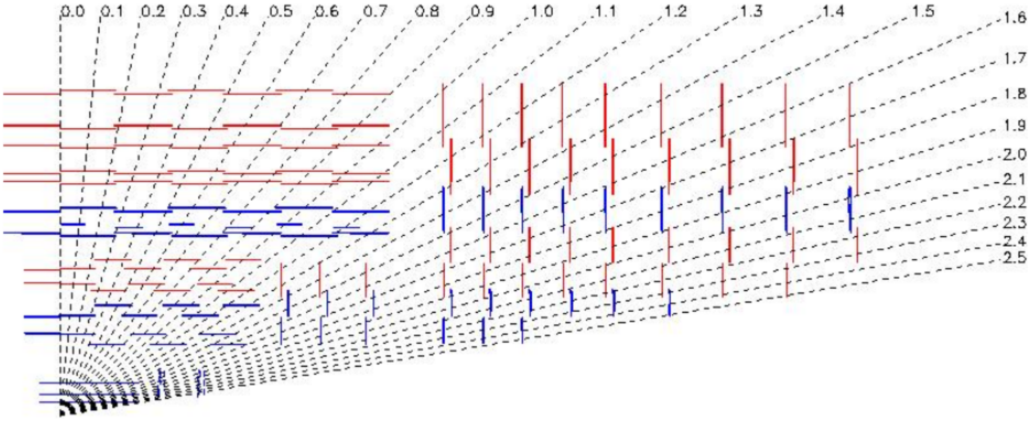

Figure 1.5: Part of the CMS Tracking system.

Together with the high magnetic field, the silicon tracker will provide a precise measurement of the tracks and the momentum of all the charged particle. The tracker information will also be used to allow the identification of jets originating from a b-quark (b-tagging).

### 1.3.3 The Electromagnetic Calorimeter

The CMS ECAL was designed in order to achieve the best sensitivity for the decay of the Higgs boson into two photons. The result is a hermetic homogeneous calorimeter made of lead tungstate ($PbWO_4$) crystals. A pre-shower detector is placed in front of the endcap crystals. An overview of the CMS Electromagnetic Calorimeter is shown in Figure 1.6.

The ECAL barrel consists of 61 200 crystals, each with a surface of 22 mm × 22 mm and a depth of 23 cm. Each ECAL endcap, on the other hand, consists of 7324 crystals with a surface of 29 mm × 29 mm and a depth of 22 cm. The ECAL provides a full coverage in the region $|\eta| < 3.0$. The pre-shower detector has a smaller pseudorapidity coverage than the ECAL endcaps ($1.653 < |\eta| < 2.6$ for the pre-shower compared to $1.479 < |\eta| < 3.0$ for the ECAL endcap). Its role is to provide a high accuracy position measurement of the electromagnetic shower and this is used to discriminate e.g. photons produced in a Higgs boson decay from photons produced in $\pi^0 \rightarrow \gamma\gamma$. The photons in the last process are produced very close to one another and a precise position measurement of the electromagnetic shower can distin-

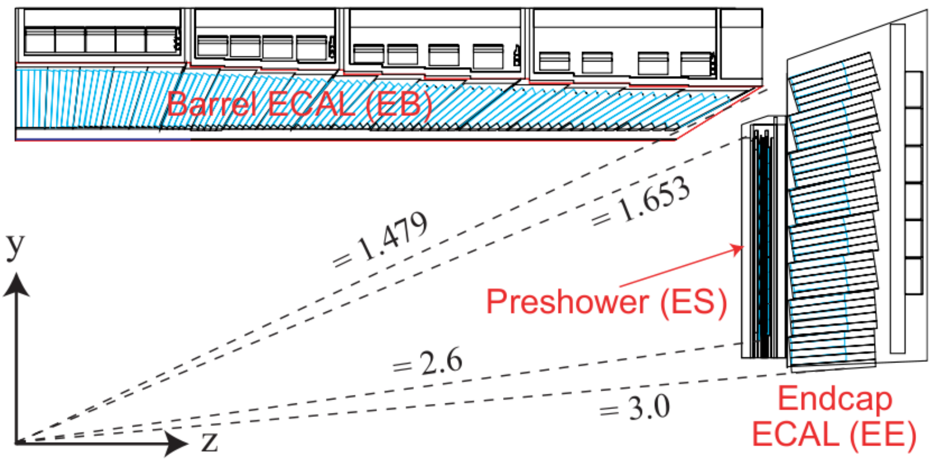

Figure 1.6: The ECAL calorimeter layout.

guish both signals. The pre-shower detector consists of thin lead radiators to initiate the shower interleaved with silicon strip sensors to measure the hit position of the shower.

The Electromagnetic Calorimeter will be used to identify e.g. electrons and photons and measure their energy and the direction in which they are produced.

### 1.3.4 The Hadronic Calorimeter

The Hadronic Calorimeter (HCAL) is placed around the ECAL. It is responsible for measuring the energy of hadrons and their products. The design was strongly influenced by the decision to place the main part of the calorimeter inside the solenoid. This leads to little space for the detector. Figure 1.7 shows the different hadronic calorimeters located in the CMS detector. The HCAL consists of four main parts: the Hadron Barrel (HB), the Hadron Endcap (HE), Hadron Outer (HO) and Hadron Forward (HF). The HB and HE calorimeters provide a coverage up to $|\eta| = 3$ and the HF extends the coverage op to $|\eta| = 5.2$.

Both the HB and the HE are sampling calorimeters with brass as absorber material and plastic scintillator as active layer. The calorimeter is divided into towers with a dimension of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$. Behind the coil of the magnet are also scintillators to detect penetrating showers leaking through the the rear of the calorimeters. This part of the calorimeter is called the HO and it provides a coverage up to $|\eta| = 1.26$. The HF calorimeter has

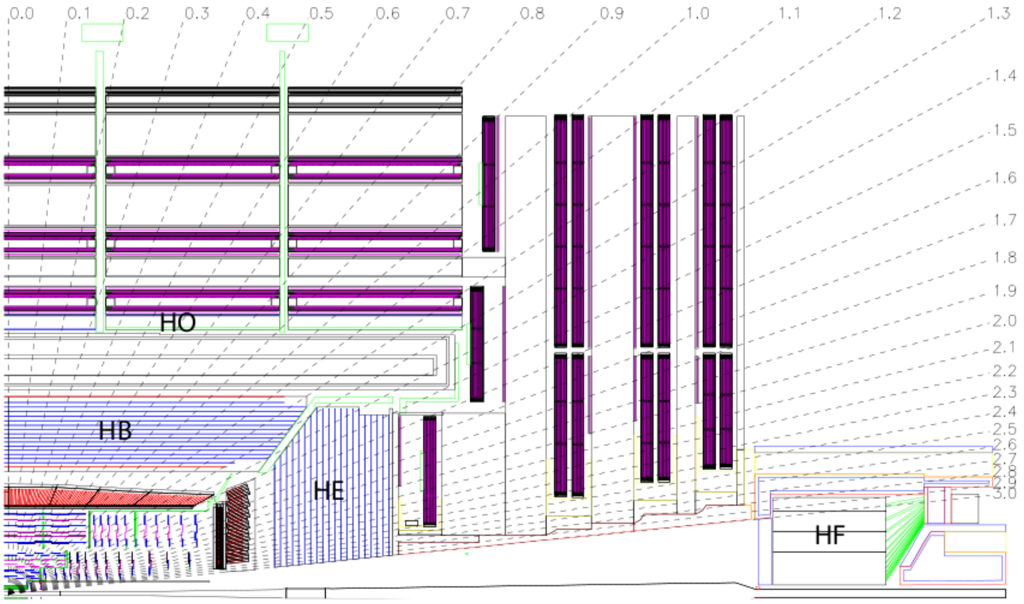

Figure 1.7: The different HCAL calorimeters.

a totally different design. It consists of steel absorber and quartz fibers emitting Cerenkov light as the active medium. The main reason for this different design is the better radiation hardness of the quartz fibers with respect to the plastic scintillators of the other parts of the HCAL detector.

The main purpose of the Hadronic Calorimeter will be to measure the energy and direction of jets which will be produced in a huge amount of events and of course also in $t\bar{t}$ events. The hermeticity of the calorimetric system will also allow a measurement of the missing transverse energy.

### 1.3.5 The Muon System

Muon detection is a very powerful tool for recognizing signatures of interesting processes over the background. Therefore, as the name of the experiment already suggests, CMS possesses a very good muon system. A graphical overview of the muon system can be found in Figure 1.8. The muon system uses three different gaseous detector technologies. The muon detectors are interleaved with the iron yoke to return the magnetic flux. In the barrel region, Drift Tubes (DT) are used with a coverage in the region $|\eta| < 1.2$, while in the endcaps, the choice was made to use Cathode Strip Chambers (CSC) with a coverage of $0.9 < |\eta| < 2.4$. Both type of detectors are assisted by Resistive Plate Chambers (RPC) extending to $|\eta| < 1.6$. The RPC plates

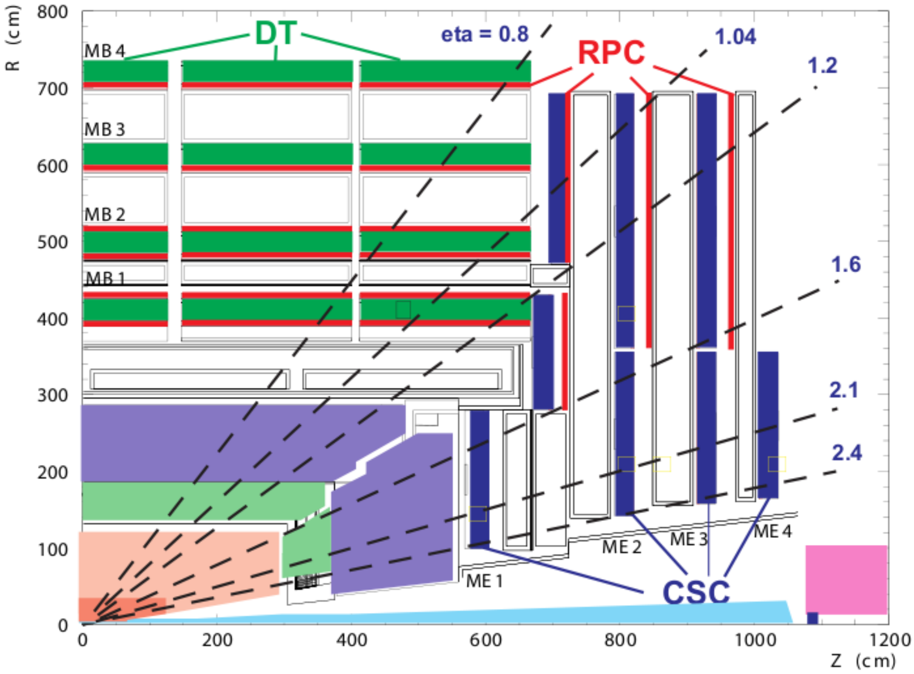

Figure 1.8: Overview of the CMS Muon System.

are, due to their very fast response and time resolution of the order of 1 ns, dedicated for trigger purpose.

In the muon identification and reconstruction, the information of the dedicated muon system will be combined with the information of the silicon tracking system to form complete muon tracks through the whole CMS detector. In a lot of analyses, like in top quark physics, muons will be very intensively used.

### 1.3.6 The Trigger System

The purpose of the Trigger System is to reduce the data of the 40 MHz event rate down to about 100 Hz, which is the maximum amount that can be stored for offline analysis. The rate is reduced in two steps called the Level-1 Trigger (L1) and the High-Level Trigger (HLT).

The Level-1 Trigger consists of custom-designed electronics which use segmented data from the calorimeters and from the muon system to decide to keep an event for further analysis or not. No tracker information is used during the L1 trigger. The L1 trigger has 3.2 $\mu$s to decide and in the meantime, the event information is stored in so-called pipelines. It has a designed output rate limit of 100 kHz. The L1 trigger electronics are located partly on the detector and partly in the underground control room located at a distance of approximately 90 m of the detector.

The High Level Trigger, on the other hand, consists of a software system implemented in a filter farm of about one thousand commercial computers, which are located at the surface next to the CMS control room. It has access to complete read-out data and can therefore perform complicated calculations, similar to those used in the offline reconstruction and analysis, to decide to keep an event or not. The main advantage of this software system is the adaptability compared to the custom-designed electronics of other trigger systems. The HLT will have a total output rate of about 100 Hz and those events will be stored for offline analysis.

# Chapter 2

# Top Quark Physics with CMS

As already mentioned in the first chapter, measuring the top quark properties in detail is one of the important tasks of the CMS experiment. There is only one particle collider in the world where the top quark can currently be produced, the Tevatron $p\bar{p}$-collider. This is the place where it was also discovered in 1995 by its two experiments, CDF [22] and D∅ [23]. The combination of all the Tevatron data until now gives a top quark mass of $173.1 \pm 1.3$ GeV/c$^2$ [4].

## 2.1 Top Quark production and decay

Since the centre-of-mass energy of the LHC is much higher than that of the Tevatron (1.96 TeV compared to 10 or 14 TeV), the cross section for top quark pair and single top quark production will be much higher at the LHC. In a hadron collision, $t\bar{t}$ or top quark pairs are produced via the strong interaction. Figure 2.1 shows the leading order Feynman diagrams for $t\bar{t}$ production. The total NLO (Next to Leading Order) and NNLO (Next to Next to Leading Order) $t\bar{t}$ production cross section at the Tevatron and the LHC can be found in Table 2.1.

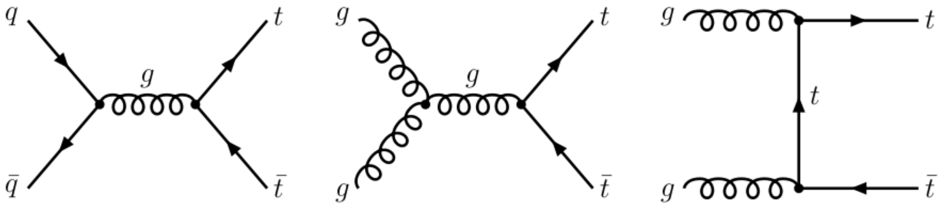

Figure 2.1: Leading order Feynman diagrams for top quark pair production.

| | NLO (pb) | NNLO (pb) |
|---|---|---|
| Tevatron ($\sqrt{s} = 1.96$ TeV) | $7.36^{+0.51}_{-0.89}$ | $7.80^{+39}_{-45}$ |
| LHC ($\sqrt{s} = 14$ TeV) | $900^{+111}_{-108}$ | $968^{+80}_{-52}$ |
| LHC ($\sqrt{s} = 10$ TeV) | $414 \pm 53$ | $446^{+33}_{-25}$ |

Table 2.1: The $t\bar{t}$ production cross section at Tevatron and LHC [24] with $M_{top} = 172$ GeV/c$^2$.

The decay of a top quark is almost exclusively (99.8 %) into a $W$-boson and a $b$-quark. The other decays ($t \rightarrow Ws$ and $t \rightarrow Wd$) will not be taken into account in what follows. The $W$-boson, on the other hand, has two decay channels: the hadronical channel ($\mathcal{B}(W \rightarrow q\bar{q}) \approx 2/3$) and the leptonical one ($\mathcal{B}(W \rightarrow l\nu_l) \approx 1/3$). The leptonic channel can be divided further into $W \rightarrow e\nu_e$, $W \rightarrow \mu\nu_\mu$ and $W \rightarrow \tau\nu_\tau$. All this means of course that $t\bar{t}$ pairs have different decay channels. An overview of the different decay channels and the corresponding branching ratios can be found in Table 2.2.

| Decay Channel | Branching Ratio |
|---|---|
| $ee$ | 1.2 % |
| $\mu\mu$ | 1.2 % |
| $\tau\tau$ | 1.2 % |
| $e\mu$ | 2.5 % |
| $e\tau$ | 2.5 % |
| $\mu\tau$ | 2.5 % |
| $e$ + jets | 14.8 % |
| $\mu$ + jets | 14.8 % |
| $\tau$ + jets | 14.8 % |
| All jets | 44.4 % |

Table 2.2: The $t\bar{t}$ decay channels and their branching ratios.

The Standard Model predicts that the top quark has a lifetime of about $10^{-25}$ s [25]. Since this is even an order of magnitude smaller than the characteristic hadronisation time of QCD, the top quark will decay before it can hadronize and form jets. It is the only quark with such a small decay time and it is therefore the only quark that can be studied as a free quark.

## 2.2 Muon Reconstruction

In top quark physics, like in a lot of other physics channels in CMS, the reconstruction of muons is very important. An overview of the muon reconstruction in CMS can be found in [26]. The reconstruction can be divided into three categories: Stand-alone reconstruction, Global reconstruction and Tracker Muon reconstruction.

In the Standalone reconstruction, a Kalman filter technique [27] is used to reconstruct the tracks, using only information from the muon system (drift tubes, resistive plate chambers and cathode strip chambers). When the track is reconstructed, it is extrapolated to the interaction point and a vertex constraint is also applied.

The Global reconstruction uses the tracks from the Stand-alone reconstruction and combines them with information from the inner tracker. The first step is a matching of the tracks from the Stand-alone muon reconstruction with silicon tracker tracks. The next step is a global refit of the track using the hits from both the muon system and the silicon tracker.

In the reconstruction of Tracker Muons, the starting points are tracks in the silicon tracker. For those tracks, the algorithm looks in the calorimeters and the muon system for compatible signatures. No combined track fit is performed. This algorithm is for example useful when a muon leaves the detector through one of the gaps between the barrel wheels.

A comparison of the efficiency in function of pseudorapidity of the Standalone and the Global reconstruction can be found in Figure 2.2. The different places of degradation of the reconstruction efficiency are located at the transition between the different barrel wheels (where small gaps in the muon system exists) and in the transition from the barrel to the endcap.

The momentum resolution versus the muon momentum for barrel and endcap and for the different reconstructions can be found in Figure 2.3. It is clear that the Global reconstruction gives the best resolution, since here the information of both are reconstructions is combined.

In most of the analyses presented in this thesis, only the GlobalMuon reconstruction will be used.

## 2.3 Jet Reconstruction

In each $t\bar{t}$ decay, between two and six quarks will be produced from the top decays. Since quarks are confined by QCD, they will hadronize and form so-called jets. Other jets will also be produced, by initial state radiation (ISR), final state radiation (FSR) or from the remnants of the colliding protons.

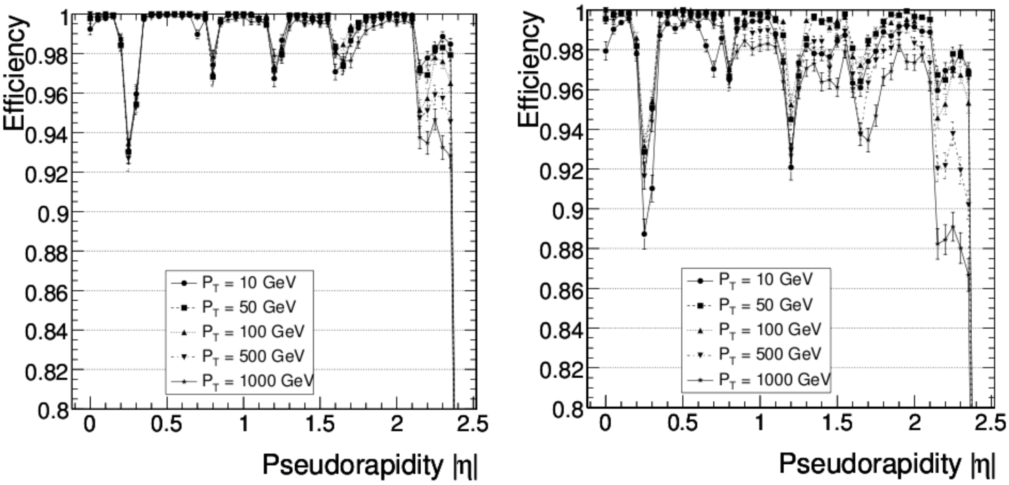

Figure 2.2: Muon reconstruction efficiency as a function of pseudorapidity for various values of the muon transverse momentum and for the Standalone reconstruction (left) and the Global reconstruction (right).

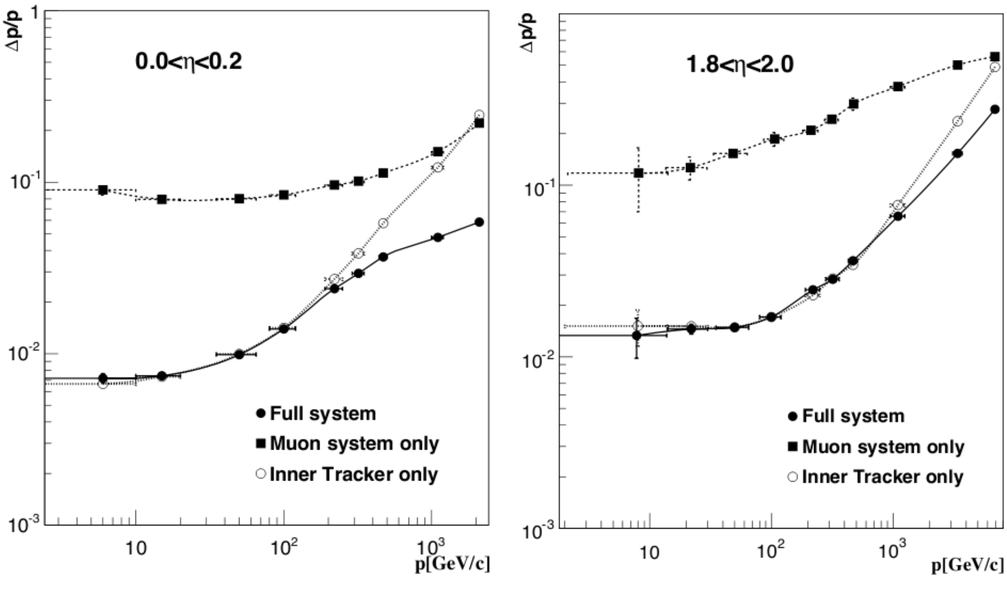

Figure 2.3: The muon momentum resolution versus the muon momentum using the muon system only, the inner tracker only, or both (full system) for the barrel, $|\eta| < 0.2$ (left) and the endcap, $1.8 < |\eta| < 2.0$ (right).

The reconstruction of those jets is a very important and non-trivial step. An overview of the jet reconstruction in CMS can be found in [29]. Those reconstruction algorithms use only calorimeter information from both the ECAL and the HCAL. An overview of the granularity of the calorimeter towers which are used as the input for the jet reconstruction can be seen in Figure 2.4. The overlap between the endcap and the forward HCAL region around $|\eta| = 3$ is clearly visible.

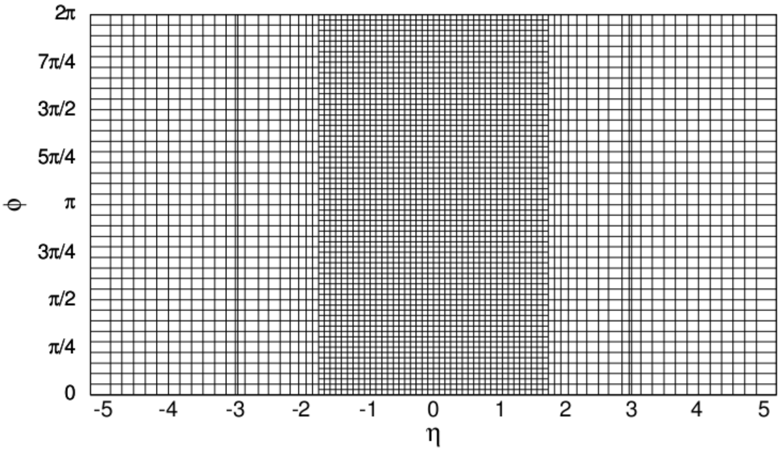

Figure 2.4: The granularity of the CMS calorimeter.

Other options which use also information from other detectors exist, like for example the Particle Flow reconstruction, which will also be used in CMS [28].

### 2.3.1 Jet Reconstruction Algorithm

In the biggest of this thesis, the jet reconstruction was done with the Iterative Cone algorithm. The input of the algorithm is an $E_T$-ordered list of input objects (calorimeter towers or particles). A cone of size $R$ in $(\eta, \phi)$-space[1] is cast around the input object with the highest $E_T$ and above a certain threshold ($E_{T,seed}^{min}$). The objects inside this cone are combined into a so-called proto-jet. They are used to calculate the proto-jet direction (the sums run over all the proto-jet constituents):

$$\eta = \frac{\sum_i E_T^i \eta^i}{\sum_i E_T^i}, \qquad \phi = \frac{\sum_i E_T^i \phi^i}{\sum_i E_T^i} \tag{2.1}$$

---

[1] The distance between two points in $(\eta, \phi)$-space is given by $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

Around this new proto-jet direction, a new cone of radius $R$ is cast and the objects inside this new cone are combined into a proto-jet, from which the new direction is calculated again. This is repeated until a stable proto-jet is found. This means the direction of the proto-jet changes by $\Delta\eta < 0.001$ and $\Delta\phi < 0.001$ between two successive iterations. The stable proto-jet is added to the list of jets and its constituents are removed from the list of input objects. This whole procedure is repeated until the list of input objects is empty or contains no more objects with an $E_T$ above $E_{T,seed}^{min}$. The parameters used in this thesis are $R = 0.5$ and $E_{T,seed}^{min} = 1$ GeV.

The next step in the reconstruction of a jet is the recombination scheme. Here, all the constituents of the jet are added together to calculate the properties of the jet. In the reconstruction used for this thesis, the energy recombination scheme is used. This means that the constituents are simply added as four-vectors. From the resulting four-vector, the desired kinematical properties of the jet can be calculated. Other recombination schemes are also used in CMS.

### 2.3.2 Jet Energy Corrections

CMS is developing a factorized multi-level jet correction. The correction factors will initially be derived from simulation tuned on testbeam data, determined directly from collision data when available, and ultimately from a simulation tuned on collision data. These corrections must be applied in the following sequence:

1. **Offset:** Required correction for the offset due to pile-up and electronic noise.

2. **Relative $\eta$:** Required correction to make the jet response flat in function of $\eta_{jet}$ for a fixed jet $p_T$.

3. **Absolute $p_T$:** Required correction to make the jet response in function of jet $p_T$ flat and equal to unity for jets in the control region ($|\eta| < 1.3$).

4. **EMF:** Optional correction for variations in jet response with electromagnetic energy fraction.

5. **Flavor:** Optional correction for different types of jets (light quark, c, b, gluon).

6. **Underlying Event:** Optional correction for underlying event energy.

7. **Parton:** Optional correction to parton level to make the GenJet[2] response to a parton ($p_T^{GenJet}/p_T^{parton}$) equal to unity on average.

Since there is no collision data yet, the correction factors are derived from simulations. More information on the jet corrections can be found in [30]. In this thesis, only the Level-2, Level-3 and Level-7 (relative $\eta$, absolute $p_T$ and parton level corrections respectively) are used. In the forthcoming, these corrections will be abbreviated to L2L3L7 corrections.

### 2.3.3 Differences between light jets and b-jets

Since each $t\bar{t}$ event contains two b-jets (jets originating from a b-quark), a good identification algorithm to distinguish them from other jets (the light jets) is necessary. Those algorithms are called b-tagging and rely on some of the special properties of b-jets.

A b-jet contains a b-hadron which will decay. In a first approximation, the other quarks of the b-hadron are just spectators of the decay and the b-quark decays via the weak interaction: $b \rightarrow W^{*}X$. The average lifetime of b-hadrons lies around 1.5 ps, corresponding to $c\tau \approx 450\mu$m. This rather long lifetime results experimentally in a displaced secondary vertex and tracks not originating from the primary vertex, as shown in Figure 2.5.

Another important property of b-jets is that part of them contain non-isolated leptons, originating from the leptonic decay of the $W$-boson: $W^{-*} \rightarrow l^{-}\bar{\nu}_l$. Those leptons inside b-jets can originate from three processes. The processes and the corresponding branching ratio's (BR) for each type of lepton are, according to [31]:

- The direct $b \rightarrow l$ decay: $b \rightarrow W^{-*}X,\ W^{-*} \rightarrow l^{-}\bar{\nu}_l$, $BR \approx 10.7$ %.
- The cascade $b \rightarrow c \rightarrow l$ decay: $b \rightarrow W^{-*}c,\ c \rightarrow l^{+}\nu_l X$, $BR \approx 8.0$ %.
- The "wrong sign" cascade $b \rightarrow \bar{c} \rightarrow l$ decay: $b \rightarrow W^{-*}X$, $W^{-*} \rightarrow q\bar{c},\ \bar{c} \rightarrow l^{-}\bar{\nu}_l X$, $BR \approx 1.6$ %.

All these processes together give a total branching ratio for the decay in at least one charged lepton of about 19.3 % for each type of lepton.

B-jets also have other properties to distinguish them from light jets, like for example a high mass and high charged track multiplicity.

---

[2] A GenJet is a jet where the input of the jet algorithm are stable generated particles from an event generator, which are located at the vertex, instead of the more usual calorimeter towers, which are of course located at the calorimeter surface. These particles are then clustered into jets by the jet algorithm. This means that GenJets are not affected by deflection in the magnetic field.

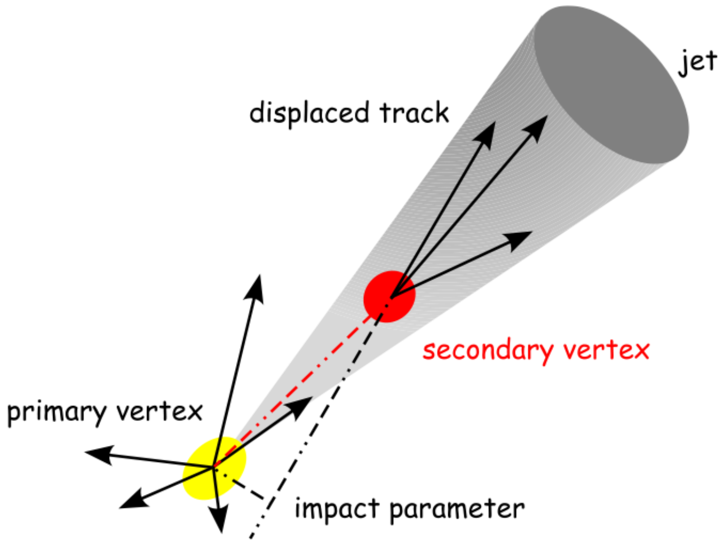

Figure 2.5: Schematic representation of a b-hadron decaying inside a b-jet.

## 2.4 Event Generation and Detector Simulation

For the studies made in this thesis, simulated events were used. These events were produced with the so-called CMS Software Framework (CMSSW). This performs all the necessary steps in the production of simulated data: the event generation, the simulation of the full CMS detector and of course the reconstruction as described in the previous sections.

The first step in the simulation of data is the generation of the event. In this step, all the outgoing particles as a result of the proton-proton collisions are generated. This step also includes the hadronization of the quarks and gluons. In this work, the event generation is done by `PYTHIA` [32] and `MadGraph` [33]. In fact, `MadGraph` only produces the parton-level events. The hadronization is also done by `PYTHIA`.

The next step uses the generated particles from the first step as input. The interactions of those particles as they cross the detector are simulated with a program based on `GEANT-4` [34]. The final step before the reconstruction is the digitization step. In this step, the response of the electronics to the hits of the previous step is simulated. This results in a data-stream as is expected from the real detector with real collision data.

One of the main disadvantages of this full simulation is that it is very CPU intensive, so it is quite time-consuming. For this reason, the CMSSW framework also contains a faster version of the simulation in which a number of simplifications are made. This FastSim was not used for the work in this thesis.

# Chapter 3

# B-Jets with muons

As already mentioned in the previous chapter, each $t\bar{t}$-production will give rise to two b-jets. The hadronization of the b-quark will produce a b-hadron inside a b-jet and in the decay of this b-hadron, a muon can be formed with a total branching ratio of about 19.3 %. The first goal of this chapter is to distinguish these b-jets from other b-jets, which do not contain a muon. When a muon is produced in the weak decay of a b-hadron, a neutrino will be produced together with the muon. The next goal of this chapter is to work out a method to estimate the energy of that neutrino. In a further step this can be used to try to improve the estimation of the b-jet energy.

## 3.1 Data Sample

The simulated data sample used in this chapter contains 78 174 fully-hadronically decaying $t\bar{t}$-events. It was generated with `PYTHIA` and it contains no pile-up (other proton collisions which happen during the same bunch crossing). The mass of the top quark was set to 172.4 GeV/c$^2$. The simulation of the CMS detector was performed with the standard CMSSW full simulation, as explained in the previous chapter. The reconstruction was also done as described in the previous chapter and the L2L3L7 corrections were applied on the jets, which were reconstructed using the Iterative Cone algorithm.

For all the reconstructed objects, in particular the reconstructed jets and muons, an association with the generated particles[1] was performed. For this, the standard association procedure within the CMSSW framework was used. A reconstructed muon was matched to a generated muon if $\Delta R(\mu_{reco}, \mu_{gen}) < 0.3$ and if $|(p_T^{\mu_{reco}} - p_T^{\mu_{gen}})/p_T^{\mu_{gen}}| < 3$. These cut-values are higher than the angular and transverse momentum resolution of the muon

---

[1] In what follows, generated particles will also be called genParticles.

reconstruction, which means that only muons originating from a genParticle are matched with the corresponding genParticle. A reconstructed jet was matched to a genParticle (only $u$, $d$, $c$, $s$ and $b$ quarks and also gluons) if $\Delta R(jet, genParticle) < 0.3$ and if $|(p_T^{jet} - p_T^{genParticle})/p_T^{genParticle}| < 3$.

## 3.2 Selection of b-jets with muons

The logical first step is a selection to find out which b-jets contain muons from the weak decay of a b-hadron and which b-jets do not contain such muons. First some initial selection cuts on the b-jets and on the muons will be worked out and afterwards a matching between the b-jets and the muons will be presented.

### 3.2.1 Initial cuts on b-jets and on muons

Originally, each of the 78 174 events contains two b-jets originating from top quark decay. On these b-jets, some cuts will be applied. First of all, only the jets who are matched to a b-quark which originates from a top decay will of course be used. On the remaining b-jets, three cuts are applied:

1. $p_T^{b-jet} > 20$ GeV/c
   This first cut is done to select only the b-jets with high enough transverse momentum. B-jets with low transverse momentum are buried inside the background of low-$p_T$ jets originating from the remnants of the two colliding protons.

2. $|\eta^{b-jet}| < 2.4$
   The goal of this cut is to remove all b-jets which do not leave tracks in the Silicon tracker. These tracks are used in a lot of analyses for b-tagging. Another reason for this cut is that the b-jets need to overlap at least partially with the CMS muon system for the matching of those b-jets with muons.

3. $\Delta R(b-jet, b-quark) < 0.3$ This is done to remove b-jets which change too much in direction by the hadronization and the reconstruction. Also very bad reconstructed jets will be removed by this cut.

After these cuts, 126 244 b-jets still remain, which means that these initial cuts have an efficiency of 80.80 %. The energy and pseudorapidity distributions of those b-jets can be found in Figure 3.1.

The muons will also be subjected to some initial cuts:

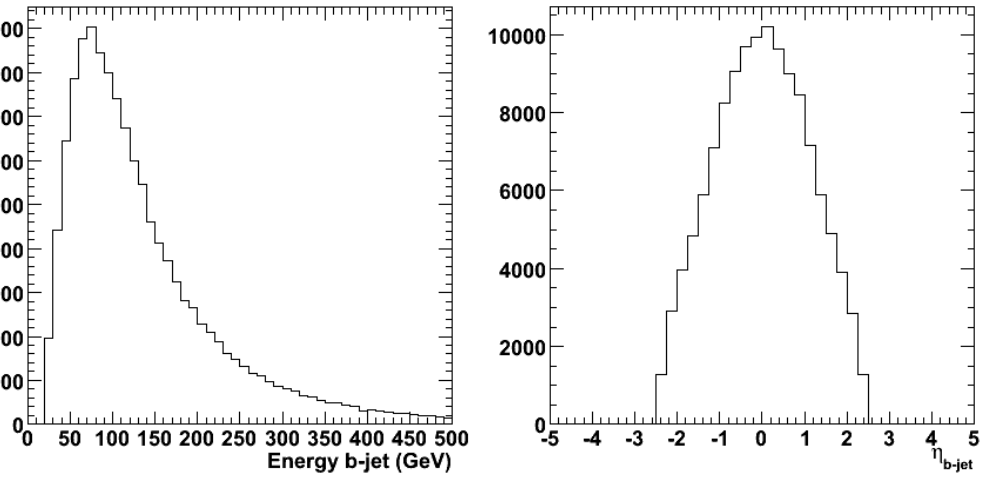

Figure 3.1: The energy (left) and pseudorapidity (right) distributions of the b-jets after the initial cuts.

1. Only GlobalMuons will be used in the further analysis.
   GlobalMuons are reconstructed muons for which a track in the muon system is matched to a track in the silicon tracker. Those muons are better reconstructed since they have more hits in the detectors.
2. $(\chi^2/ndf)_{globalTrack} < 5$
   This is to remove the muons with a bad reconstruction. The globalTrack is the combination of the track in the Silicon tracker with the track in the muon system.
3. Number of valid hits in the Silicon tracker $> 10$
   This cut is also to reject muons with a bad reconstruction.

The transverse momentum and and pseudorapidity of the 27 656 muons which survive these initial cuts are distributed as shown in Figure 3.2.

### 3.2.2 B-jet – muon matching

The muons surviving the initial cuts of the previous section will be matched with the b-jets which also survived the initial cuts. A muon is matched to a b-jet if $\Delta R(\mu, b-jet) < 0.5$. The $\Delta R(\mu, b-jet)$ between all the b-jets and all the muons is shown in Figure 3.3, together with the number of muons matched to each b-jet. In this last plot, the first bin, containing the number of b-jets without muons, is left empty to make the other bins more visible. The number of b-jets not matched with one or more muons is 109 485. From

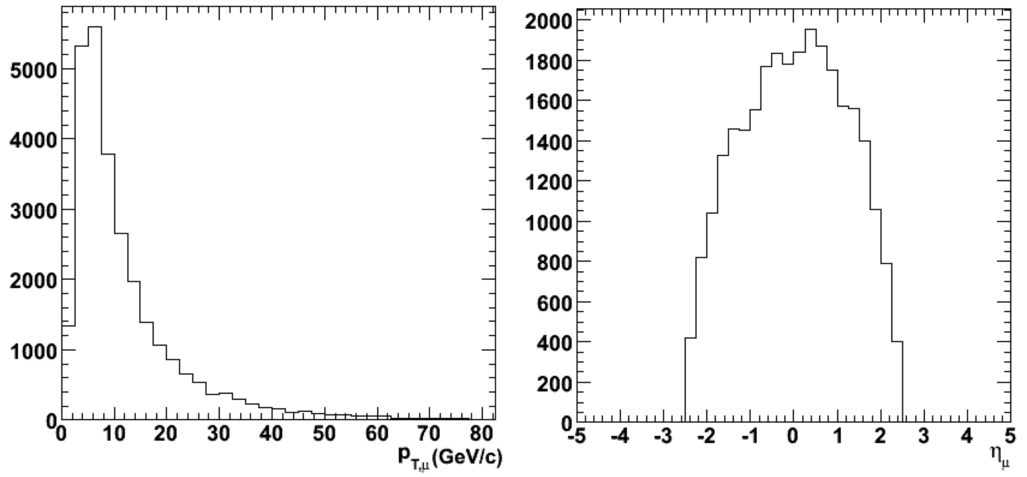

Figure 3.2: The $p_T$ (left) and pseudorapidity (right) distributions of the muons after the initial cuts.

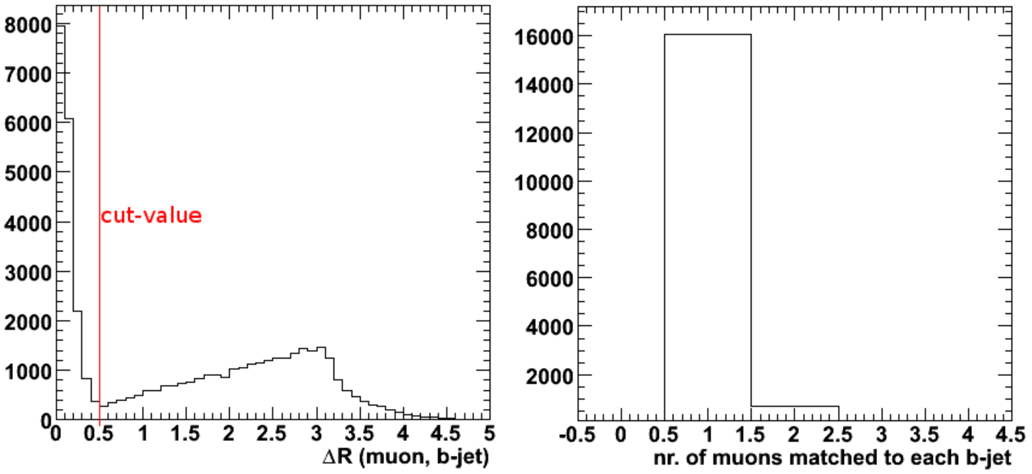

Figure 3.3: $\Delta R(\mu, b-jet)$ (left) and the number of muons matched to each b-jet (right). The bin containing the number of b-jets without muons is empty as explained in the text.

the 126 244 b-jets which survived the initial cuts, 16 759 b-jets are matched with one or more muons. This means that 13.28 % of all the b-jets are matched with one or more muons. When this value is compared to the total fraction of b-jets that should contain at least one muon from Section 2.3.3 (19.3 %), it is clear that the selection and matching is far from optimal. This can partially be explained by the fact that part of the b-jets are exceeding the maximum pseudorapidity coverage of the muon system ($|\eta| = 2.4$).

There are also 72 events where the two b-jets are matched to the same muon. These events will not be used in the further analysis. In a later step they can be used, for example by matching the muon to the b-jets which is closest to the muon in $(\eta, \phi)$-space. It was not checked if this really matches the muon with the good b-jet.

### 3.2.3 B-jets with and without muons

For the b-jets matched with a muon and for those not matched with a muon, the transverse energy[2] (of course after the L2L3L7 calibration) and the pseudorapidity is drawn in Figure 3.4. The pseudorapidity of the b-jets with muons shows that there is no significant difference in pseudorapidity between both types of b-jets. The transverse energy plot shows that b-jets without muons have, on average, a higher transverse energy when compared to b-jets containing muons. This can be explained by the fact that for b-jets containing muons, part of the energy is not measured in the calorimeter due to the muons and the neutrino's inside the b-jet. This explanation is also supported by what is shown in Figure 3.5. Here, the relative energy difference is shown for b-jets with and without muons. The relative energy difference is defined as:

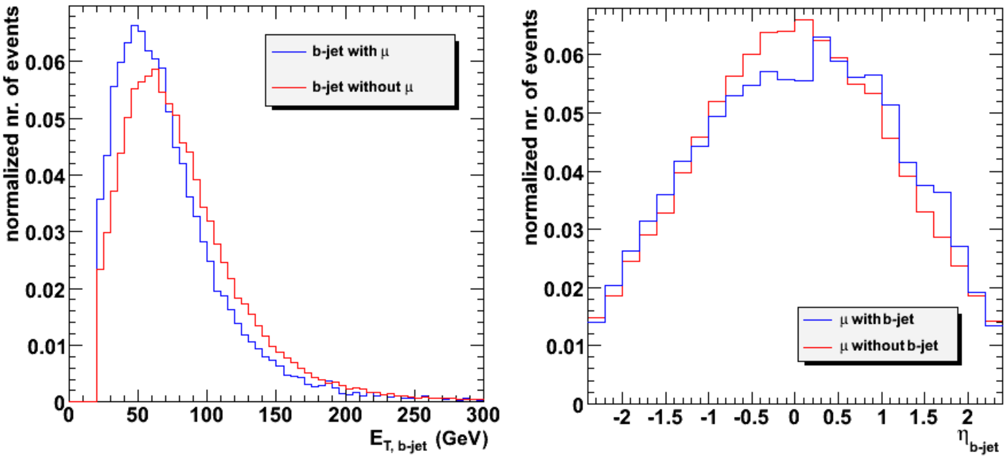

Figure 3.4: Transverse energy (left) and pseudorapidity (right) of the b-jets with and without muons.

$$\text{Relative energy difference} = \frac{E_{b-jet} - E_{b-quark}}{E_{b-quark}} \tag{3.1}$$

---

[2] Transverse energy is defined as: $E_T = E\ \sin\theta$

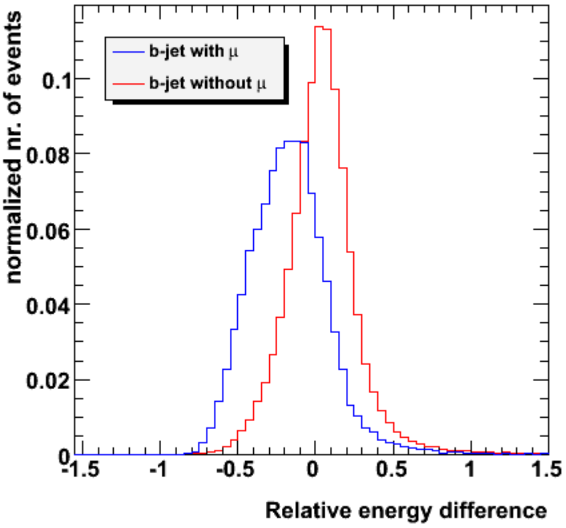

Figure 3.5: Relative energy difference for b-jets with and without muons.

This relative energy difference is significantly narrower and closer to zero for b-jets with muons than for b-jets without muons. The reason for this difference is that those muons (and the neutrino's) do not deposit their energy in the calorimeters, so the b-jet's energy does not take into account the energy of the muon (and the neutrino).

Figure 3.6 shows the relative energy difference as function of the b-quark energy and pseudorapidity. These plots again show that the b-jets without

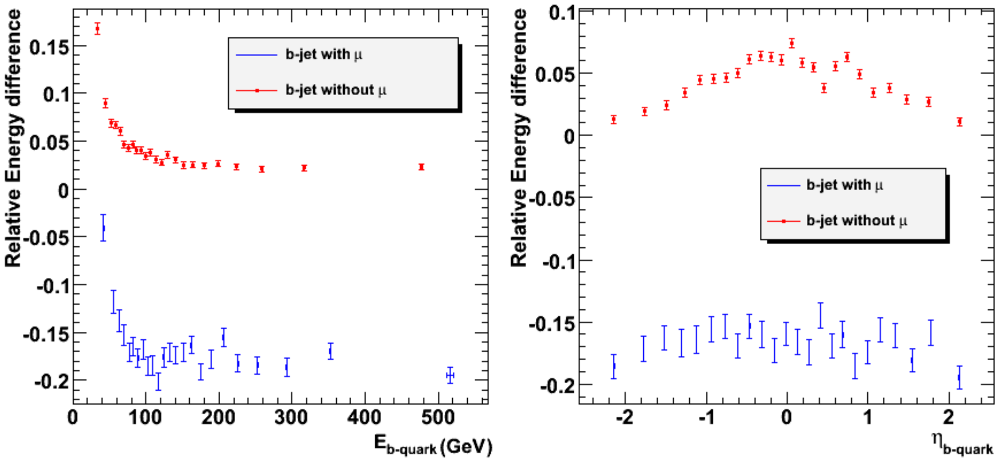

Figure 3.6: Relative energy difference in function of the b-quark energy (left) and pseudorapidity (right) for b-jets with and without muons.

muons have a more narrow relative energy difference distribution than the b-jets with muons. The first plot shows that the relative energy difference is significantly different from zero for b-jets with lower energy. As the energy of the b-jet rises, this difference becomes much smaller. Part of this effect is caused by the high magnetic field of CMS. This bends the positively charged particles and the negatively charged particles each in another direction and this broadens the jet. This broadening effect is higher for low energy jets than for high energy jets. This means that for lower energy jets, a higher fraction of the jet's energy will lie outside the cone of the jet algorithm when compared to high energy jets. The second plot shows that the relative energy difference in function of the pseudorapidity is closer to zero for b-jets without muons than for b-jets with muons. The plot is also symmetric around $\eta = 0$, as expected.

These last two plots have a special horizontal binning to have the same amount of data points inside each bin. When all the data points are binned, the mean x and y value is calculated for each bin, together with the statistical uncertainties on both means. The final step is plotting those x and y values together with their uncertainties. These kind of plots will be used quite a lot in the remainder of this thesis, except in Section 3.2.5, where normal profile plots (with a fixed bin size) will be used.

### 3.2.4 Comparison between muons inside a b-jet and muons not inside a b-jet

The transverse momentum and pseudorapidity of muons matched to a b-jet and of muons not matched to a b-jet are plotted in Figure 3.7. The pseudorapidity plot shows that the muons matched to a b-jet have a more narrow pseudorapidity distribution which is closer to $\eta = 0$ than the other muons. The histogram of the transverse momentum shows that, on average, muons with a slightly higher transverse momentum are matched to a b-jet.

### 3.2.5 Correlation between some variables of the b-jets matched with muons

Before looking at these correlations, another variable needs to be defined. The $p_{Trel}$ variable gives the transverse momentum of the muon relative to the direction of the total muon-jet momentum vector, as shown in Figure 3.8. $p_{Trel}$ can be calculated by

$$p_{Trel} = \frac{|p^{\mu} \times p^{\mu+jet}|}{|p^{\mu+jet}|} \tag{3.2}$$

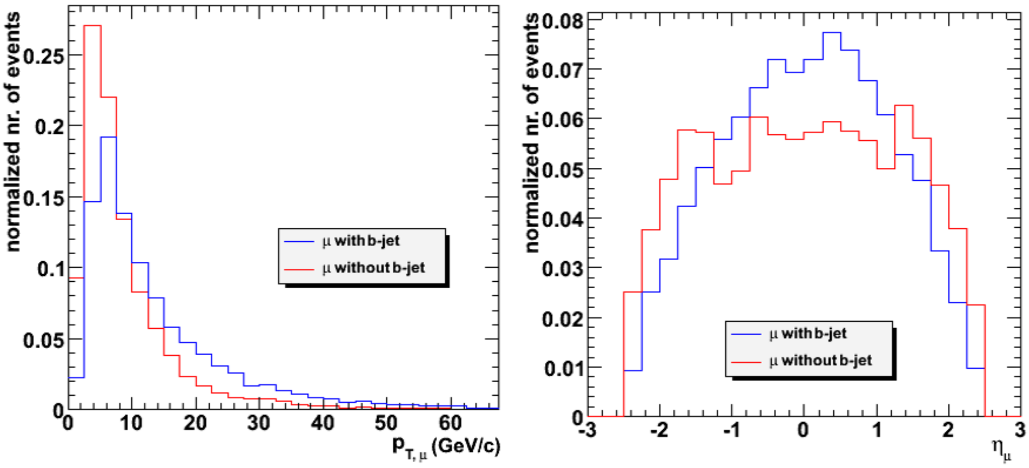

Figure 3.7: Transverse momentum (left) and pseudorapidity (right) of the muons matched to a b-jet and of the muons not matched to a b-jet.

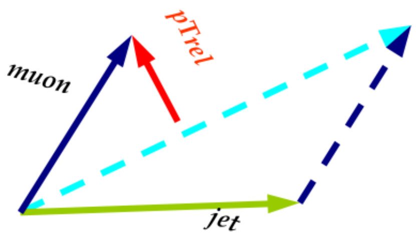

Figure 3.8: Definition of the $p_{Trel}$ variable.

In the following plots, profile plots will be used, so some explanation on how they are made is necessary. They are constructed starting from a two-dimensional histogram with only binning on the x-axis. For every bin, the average and the statistical uncertainty on the average is calculated and plotted. As already explained at the end of Section 3.2.3, the difference between these plots and the plots from other sections, is that these plots use a fixed bin size and those from other sections have a variable bin size to have the same amount of data points inside each bin.

The relation between the $p_{Trel}$ variable and the distance in $(\eta, \phi)$-space between the muon and the b-jet ($\Delta R(\mu, b - jet)$) can be observed in Figure 3.9. The correlation coefficient is 51.50 %. Both plots and the correlation coefficient show that there is some correlation between both variables. When thinking about the definition of both variables, this is indeed expected. A b-jet and a muon which are closer to each other in $(\eta, \phi)$-space will have a smaller $p_{Trel}$, because this variable also depends on the angle between the

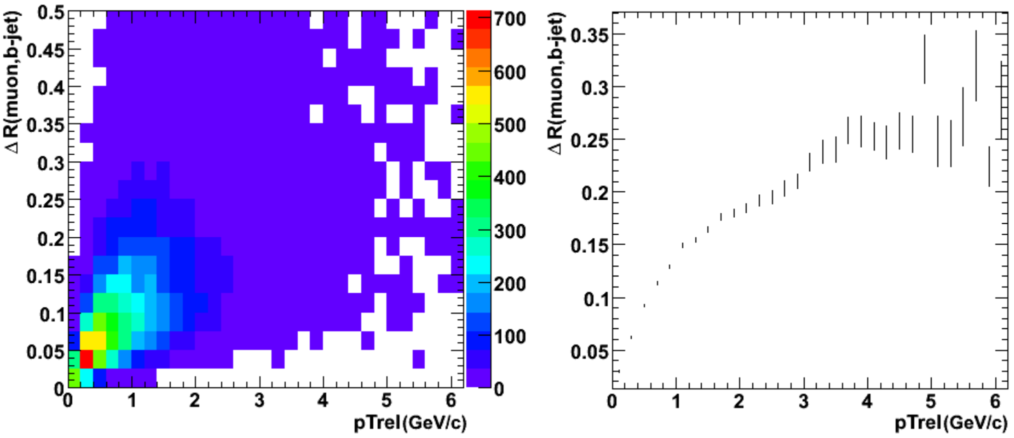

Figure 3.9: Two-dimensional histogram (left) and profile plot (right) of $p_{Trel}$ versus $\Delta R(\mu, b-jet)$.

muon and the b-jet. The same holds for a b-jet and a muon which has a large separation in $(\eta, \phi)$-space. Here the $p_{Trel}$ will also be higher, as observed in the plots.

Another relation that deserves some attention is the relation between the transverse momentum of the b-jet and the distance in $(\eta, \phi)$-space between the muon and the b-jet. This can be seen in Figure 3.10. The correlation coefficient equals -19.03 %. Both variables are almost uncorrelated, but the plots suggest a small connection between both variables. When a b-jet has lower transverse momentum, the distance in $(\eta, \phi)$-space between the muon

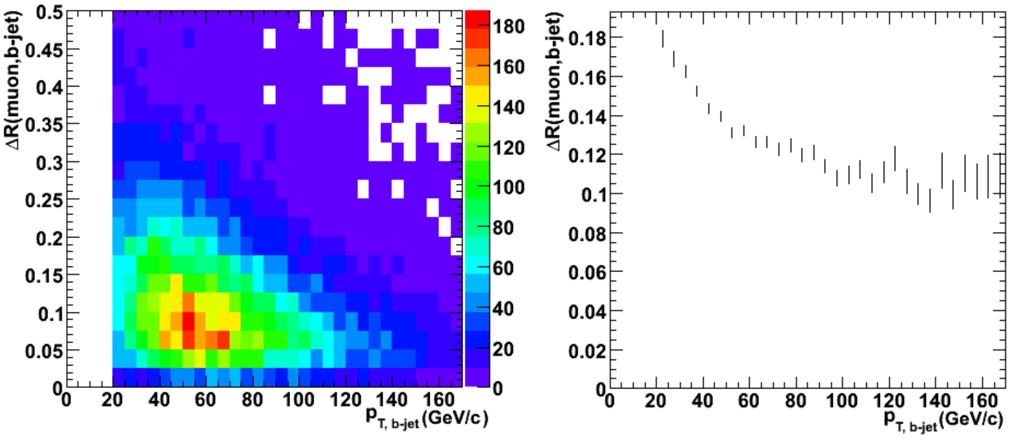

Figure 3.10: Two-dimensional histogram (left) and profile plot (right) of $p_{T,b-jet}$ versus $\Delta R(\mu, b-jet)$.

and the b-jet is somewhat higher and vice versa. This can be expected, because when a b-jet has a higher transverse momentum, it is more boosted and this results in a b-jet inside a cone with a smaller opening angle. For b-jets with lower transverse momentum, the opposite happens. They are less boosted and they result in a b-jet inside a cone with a larger opening angle.

## 3.3 Adding the muon to the b-jet

The logical first step when trying to make a better estimate of the b-jet energy and momentum, is of course adding the energy and momenta of the muon, which is matched to the b-jet, to the energy and momentum of the b-jet. For this, a distinction will be made between b-jets with one muon and b-jets with more than one muon. This is done because in a later stage, the corrections for the neutrino energy will be worked out first for b-jets with one muon to keep the corrections as simple as possible in a first stage. Afterwards, the results for all b-jets containing one or more muons will also be given.

### 3.3.1 B-jets containing one muon

As previously said, the energy and momenta of the muon is added to those of the b-jet. In this section, only the b-jets containing exactly one muon are used.

The relative energy difference of the b-jet, as defined in Eq. 3.1, is shown in Figure 3.11 before and after the addition of the muon. As expected, the addition of the muon gives a much better distribution of the relative energy difference. It has a lower standard deviation (0.25 before and 0.23 after adding the muon) and a mean which is much closer to zero ($-0.163$ before and 0.003 after the inclusion of the muon).

In Figure 3.12, the relative energy difference in function of the b-quark energy and pseudorapidity for b-jets with one muon is shown. These plots are made with a variable bin size to have about the same amount of data points inside each bin, as explained in Section 3.2.3. The plots show that adding the muon to the b-jet gives a relative energy difference which is much closer to zero. The first plot shows again that the relative energy difference is significantly higher for lower energy b-jets when compared to higher energy b-jets, as already noted in Section 3.2.3. The second plot shows that adding the muon to the b-jet gives a relative energy difference which is much closer to zero in each pseudorapidity bin and the plot shows also the expected symmetry around $\eta = 0$. It also shows that the muons make up about 15 %

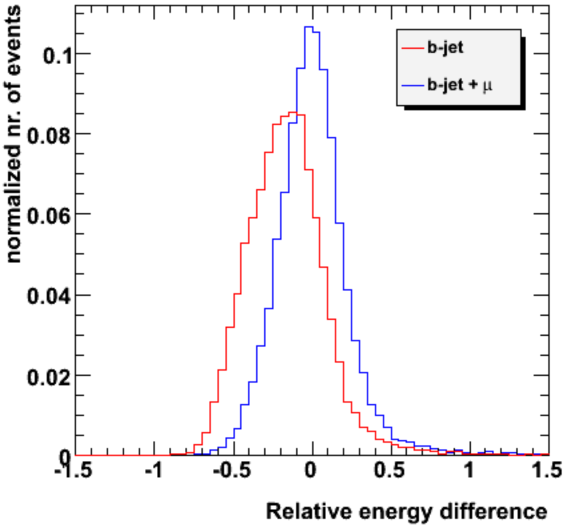

Figure 3.11: Relative energy difference of the b-jets before and after the addition of the muon, only for b-jets with one muon.

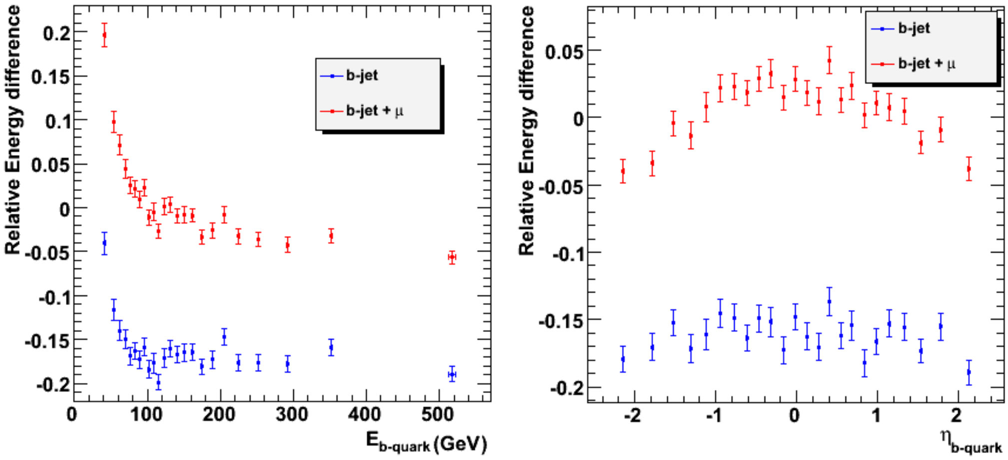

Figure 3.12: Profile plot of the relative energy difference versus the b-quark energy (left) and versus the b-quark pseudorapidity (right), only for b-jets with one muon.

of of the total energy of the jet.

### 3.3.2 B-jets containing one or more muons

For completeness, the same plots as in the previous section are also shown with b-jets with one and more than one muon. Figure 3.3 shows that there is only a small amount of b-jets with more than one muon, so the effect of including these b-jets will be small. The relative energy difference before and after the addition of the muons is shown in Figure 3.13 and the relative energy difference in function of the b-quark's energy and pseudorapidity is shown in Figure 3.14. The conclusion of these three plots is the same as the

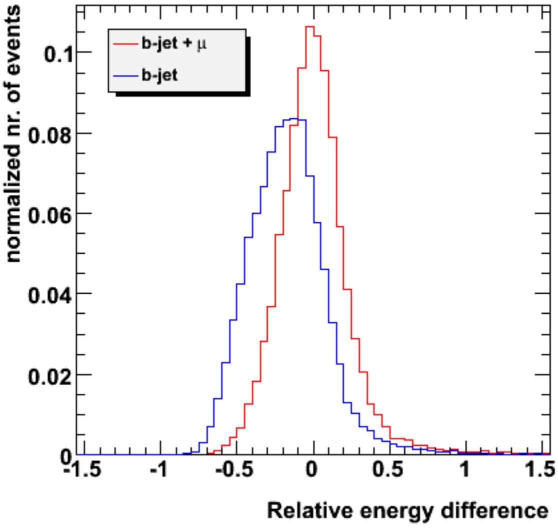

Figure 3.13: Relative energy difference of b-jets with one or more muons before and after the addition of the muons.

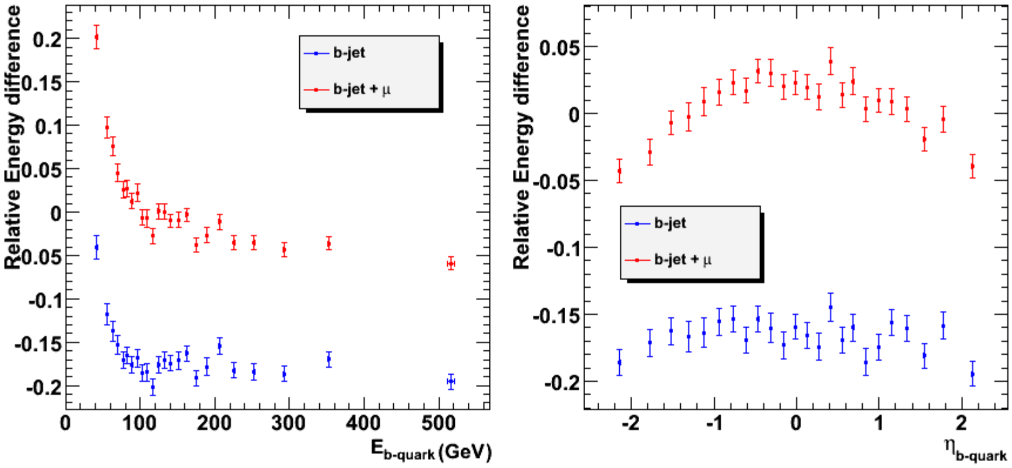

Figure 3.14: Profile plot of the relative energy difference versus the b-quark energy (left) and versus the b-quark pseudorapidity (right), for b-jets with one or more muons.

conclusion of the corresponding plots in the previous section. The addition of the muon to the b-jets gives a narrower relative energy difference

Another thing to study is how these b-jets with the addition of the muons compare with b-jets without muons. The comparison between b-jets with muons and b-jets without muons (without any addition of muons to the b-jets) can be found in Section 3.2.3. The relative energy difference for b-jets without muons and for b-jets with muons and with the addition of the muons can be found in Figure 3.15. The mean of the histograms is 0.042 for b-jets without muons and 0.005 for b-jets with muons. a possible explanation for this is that the energy of the neutrino still needs to be added to the b-jet with muons. The standard deviation is slightly larger for b-jets with muons (0.239) than for b-jets without muons (0.237), which gives a distribution which is a bit broader for b-jets with muons than for b-jets without muons. This shows that both classes of b-jets have, on average, almost the same relative energy difference. The relative energy difference for those b-jets in function of the b-quark energy and pseudorapidity can be found in Figure 3.16. The statistical uncertainties are larger for b-jets with muons because there are less b-jets with muons. The first plot shows that the relative energy difference is flatter for b-jets without muons than for b-jets with muons and it shows also that for high energy b-quarks, the relative energy difference is significantly lower for b-jets with muons than for b-jets without muons. With the inclusion of the energy of the neutrino that escapes detection, this

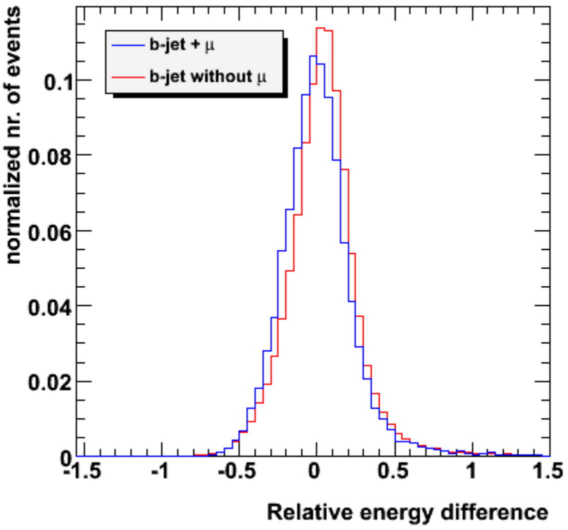

Figure 3.15: Relative energy difference of b-jets with and without muons. For the b-jets with muons, the muons are added to the b-jet.

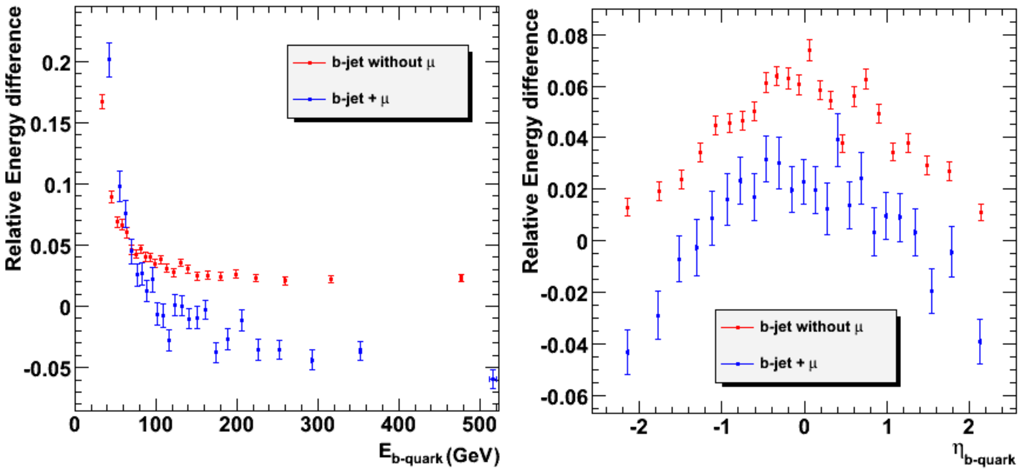

Figure 3.16: Profile plot of the relative energy difference versus the b-quark energy (left) and versus the b-quark pseudorapidity (right), for b-jets with and without muon. For the b-jets with muons, the muons are added to the b-jet.

can possibly be improved. The second plots, showing the relative energy difference in function of the pseudorapidity of the b-quark, shows that the relative energy difference is higher for b-jets without muons in the whole pseudorapidity range. As already mentioned, this can be explained by the fact that the energy of the neutrino still needs to be added to the b-jets with muons.

### 3.3.3 Adding the muons not passing the initial cuts

For completeness, the result of simply adding all the muons (without any initial cut) with $\Delta R(\mu, b-jet) < 0.5$ is shown, but only for b-jets which are matched with one muon (which survives the initial muon cuts). This muon will be called the good muon and the muons not surviving these initial cuts will be called the other muons. This nomenclature will only be used in this small section. Figure 3.17 shows the relative energy difference of the b-jets without any addition of muons, the b-jets with the addition of the good muon and the b-jets with the addition of the good muon and the other muons with $\Delta R(\mu, b-jet) < 0.5$. The mean (standard deviation) of the distributions are $-0.163$ (0.254) for the b-jets, 0.005 (0.237) for b-jets + good $\mu$ and 0.042 (0.252) for b-jets + good $\mu$ + other $\mu$. The addition of these other muons broadens the distribution and it also shifts the distribution to a value which is more different from zero than before the addition of these other muons.

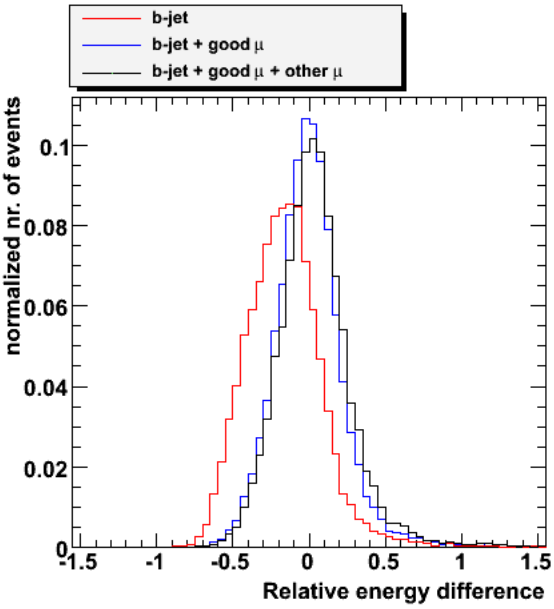

Figure 3.17: Relative energy difference of b-jets with one muon. The muons surviving the initial cuts and the muons not surviving these cuts are added to the b-jet.

The addition of the good muons gives a much larger effect (about 16 %) in the shift of the distribution of the relative energy difference than the addition of the other muons (about 4 %).

## 3.4 The neutrino's

The weak decay of a b-hadron inside a b-jet which gives rise to a muon, will also give rise to a neutrino, which remains of course undetected. The aim of this section is to estimate the energy of this neutrino and to use this estimate to improve the relative energy difference. In this section, only b-jets matched with exactly one muon will be used to keep the estimation as simple as possible.

### 3.4.1 Finding the neutrino's in the genParticles

The first step is to see whether the muon which is matched to the b-jet is really originating from the decay of a b-hadron inside this b-jet. This is checked by looking if the muon's genParticle originates from the decay of a b-hadron. This b-hadron also needs to originate from the hadronization of the b-quark, which on his turn originates from the top quark decay.

From the 16 020 muons of the b-jets which are matched to one muon, 2 428 muons are not matched to a genParticle, so no information on the origin of the muon can be given. The fact that they are not matched to a genParticle can have several causes, for instance, the muon can originate from the decay of a long-lived particle, like a pion. Another possibility is that the reconstructed object that is identified as a muon, is not really a muon but some other particle that is wrongly identified as a muon. It is also possible that the reconstructed muon differs so much from the generated muon that the matching between generated and reconstructed muons does not match the reconstructed to the generated muon anymore.

The 13 592 muons which are matched to a genParticle originate from different sources. An overview of these sources can be found in Table 3.1. In what follows, only the muons 8 011 which are produced directly from b-hadron decay will be used. In a later stage, the other muons can be included and certainly those from c-hadron decay, since these c-hadrons can originate from the decay of a b-hadron inside the b-jet.

| Source | Number of muons | Fraction of total |
|---|---|---|
| b-hadron decay | 8 011 | 58.94 % |
| c-hadron decay | 5 140 | 37.82 % |
| $\tau$ decay | 434 | 3.19 % |
| $\rho$, $\eta$ and $\phi$ decay | 7 | 0.05 % |
| Total | 13 592 | 100.00 % |

Table 3.1: The origin of the muons inside b-jets.

The b-hadrons from which the muons originate are also asked to be produced in the hadronization of a b-quark which is a decay product of a top quark. From the initial 8 011 muons which are produced in b-hadron decays, there are 3 232 muons which do not pass this extra constraint. The b-hadrons from which these muons originate, are mainly produced in the hadronization of the remnants of the initial protons. There still remain 4 779 muons which are produced in the decay of a b-hadron which originates from the b-jet from a top quark. In the decay of all these b-hadrons, a neutrino is also produced, as expected. This neutrino will be used in the remaining sections of this chapter.

### 3.4.2 Differences between b-jets with a muon from b-hadron decay and b-jets with a muon from another origin

As described in the previous section, the use of the genParticles gives access to the origin of the muons inside the b-jets. In this section, some differences between the b-jets with one muon from b-hadron decay (where the b-hadron originates from the hadronization a b-quark, which on his turn originates from the decay of a top quark) and the other b-jets with one muon will be presented. These differences are shown in Figure 3.18. In these histograms, the b-jets with one muon from b-hadron decay and originating from the top quark, are called "good $\mu$" and all the other b-jets with one muon are called "other $\mu$".

The first plot, shows the distance in $(\eta, \phi)$-space between the b-jet and the muon for both types of muons. The b-jets with "good $\mu$" have a slightly larger separation between b-jets and muons than the b-jets with "other $\mu$". This can be explained by the fact that these "good $\mu$" originate from the decay of a heavier particle and that the "other $\mu$" originate from the decay of a lighter particle. These lighter particles will, in general, be more boosted than these heavier particles and their decay products will be produced inside a smaller cone than the decay products of the heavier particles. The same explanation applies for $p_{Trel}$, which is defined in Eq. 3.2 and which is shown for both types of b-jets in the second plot. The "good $\mu$" will be more separated from the b-jet axis and this will result in a higher momentum transverse to the b-jet + muon axis. The "other $\mu$" will be closer to the b-jet axis and this will give a lower value of $p_{Trel}$.

The third figure shows that the transverse energy of the b-jets with "good $\mu$" is, on average, lower than the transverse energy of b-jets with "other $\mu$", while the fourth figure shows that the transverse momentum of the "good $\mu$" is, on average, higher than the transverse momentum of the "other $\mu$". From these two plots, it can be concluded that the muon which originates from the b-hadron decay will have, on average, a higher transverse momentum than the "other $\mu$". For b-jets with one "good $\mu$", this means that a higher fraction of the b-jet's energy will escape detection in the calorimeters and this gives rise to a lower transverse energy for b-jets with one "good $\mu$". For the b-jets with one "other $\mu$", the opposite can be concluded: the muon has a lower transverse momentum on average and this results in a higher energy deposition of the b-jet in the calorimeters and of course a higher transverse energy. This also explains what is shown in the fifth plot. The b-jets with one "good $\mu$" will have a lower b-jet energy and a higher muon momentum,

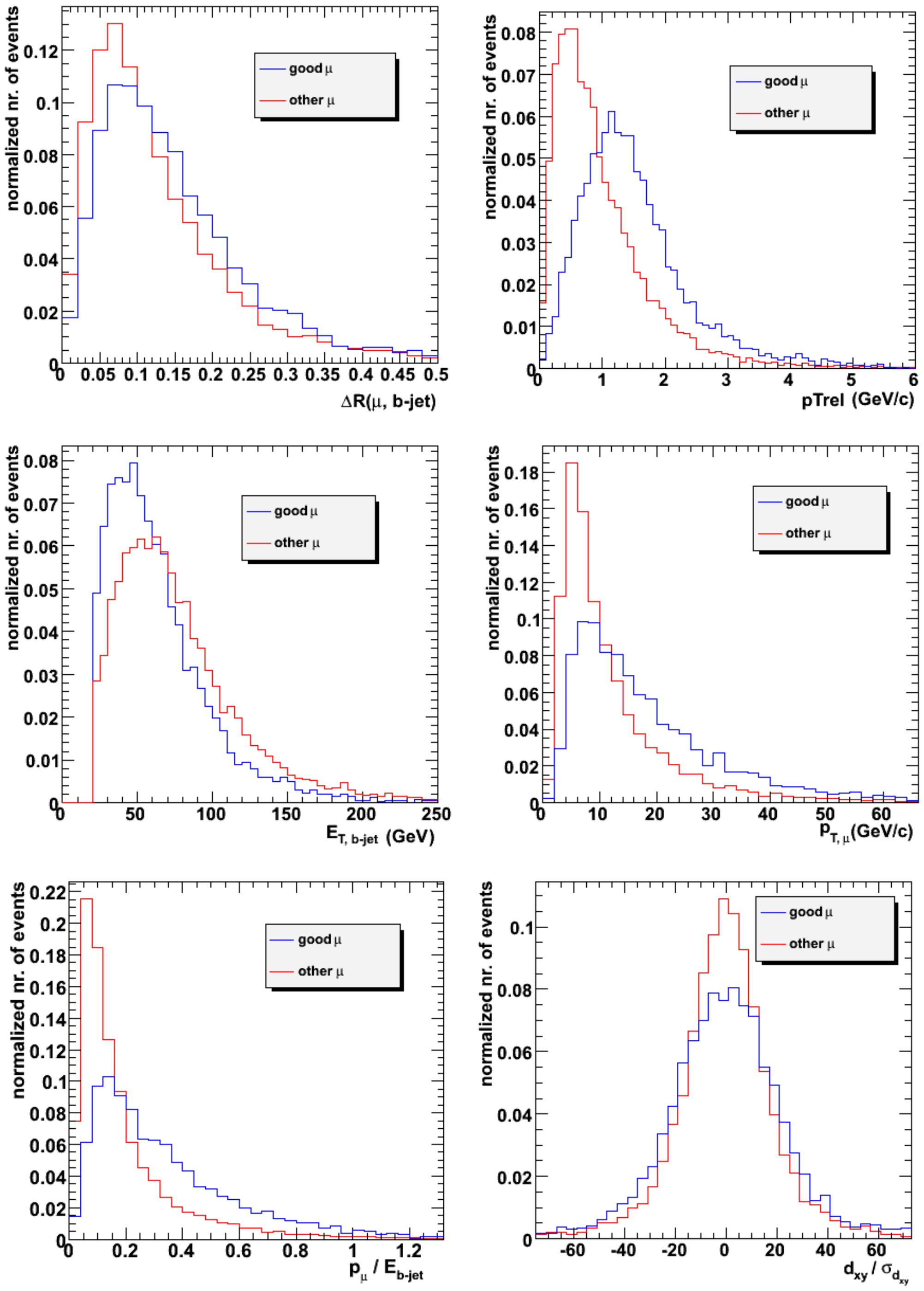

Figure 3.18: Differences between b-jets with one “good $\mu$” and b-jets with one “other $\mu$”.

as explained in the previous, and this results in a higher value of $p_\mu/E_{b-jet}$. For the b-jets with one “other $\mu$”, the reverse is true.

The sixth and final plot shows the significance of the transverse impact parameter ($\sigma_{d_{xy}}/d_{xy}$) for both types of muons. The “good $\mu$” have a slightly larger value of this significance, which means they have a larger transverse impact parameter and/or a smaller uncertainty on this transverse impact parameter, while it is the opposite for the “other $\mu$”. The “good $\mu$” are produced in the decay of a long-lived b-hadron, which gives a larger significance, while a significant fraction of the “other $\mu$” are muons which are not produced in the decay of long lived particles, which gives a smaller significance. The difference is not very larger, since the “other $\mu$” also contain a significant amount of muons which are produced in the decay of a c-hadron, which on his turn was produced in the decay of a b-hadron. Since they also originate from the decay of long-lived particles, these muons will also have a larger significance.

### 3.4.3 Adding the generated neutrino to the b-jet

For the b-jets with one “good $\mu$”, a method will be worked out to estimate the energy of the neutrino. This will be done in the next section, but first the energy of the generated neutrino is added to the b-jet. As in Section 3.3, the energy of the matched muon is also added to the b-jet.

The relative energy difference, as defined in Eq. 3.1, is shown in Figure 3.19 for b-jets with one “good $\mu$”, before and after the addition of the muon an the generated neutrino neutrino energy. The average (standard deviation) of the different distributions is $-0.240$ (0.247) for the b-jet, $-0.013$ (0.240) for the b-jet + $\mu$ and 0.144 (0.207) for the b-jet + $\mu$ + $\nu$. As already showed in the previous sections, the addition of the muon energy gives a relative energy distribution which is narrower and closer to zero. The addition of the neutrino energy makes the distribution even more narrow, but it gives a distribution which is significantly shifted from zero. This means that a good estimation of the neutrino energy can give a more narrow distribution of the relative energy difference, which means a better estimation of the energy of the b-quark.

The relative energy difference in function of the b-quark energy and the b-quark pseudorapidity for b-jets with one “good $\mu$”, before and after the addition of this “good $\mu$” and the generated neutrino, is shown in Figure 3.20. Like in the previous plots, both these plots show that the relative energy difference is shifted to higher values by the addition of the generated neutrino. This shift happens in every b-quark energy and pseudorapidity bin. The shift due to the addition of the muon to the b-jet is about the same

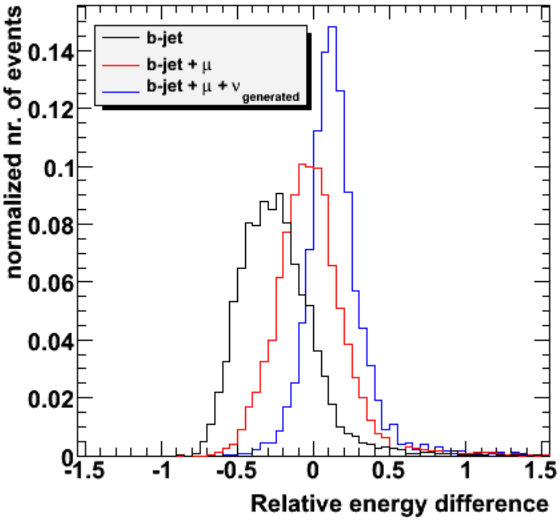

Figure 3.19: Relative energy difference of b-jets with one "good $\mu$".

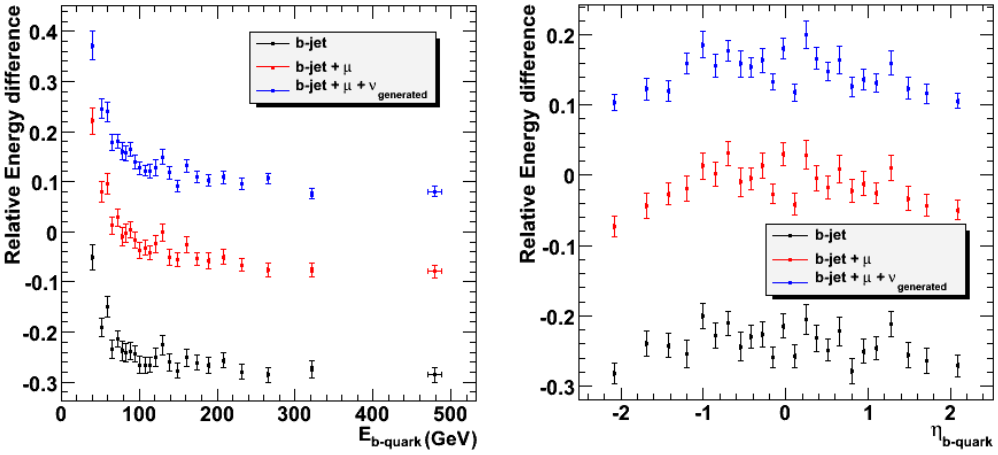

Figure 3.20: Relative energy difference in function of the b-quark energy (left) and pseudorapidity (right) for b-jets with one "good $\mu$".

(slightly smaller) as the shift due to the addition of the generated neutrino to the b-jet and the muon.

### 3.4.4 Estimating the energy of the neutrino

In this last section of this chapter, a method will be presented to estimate the energy of the neutrino and this estimate will be used to improve the estimation of the energy of the b-quark. Only the b-jets which have one "good $\mu$" (and of course also a neutrino) will be used. It is expected that the absolute energy difference between the b-jet + muon and the b-quark is correlated with the energy of the generated neutrino. This dependency is shown in Figure 3.21. There is a clear correlation between the absolute energy difference and the energy of the generated neutrino, with a correlation coefficient of -62.28 %. This again shows that a proper estimation of the neutrino energy will improve the estimation of the energy of the b-quark.

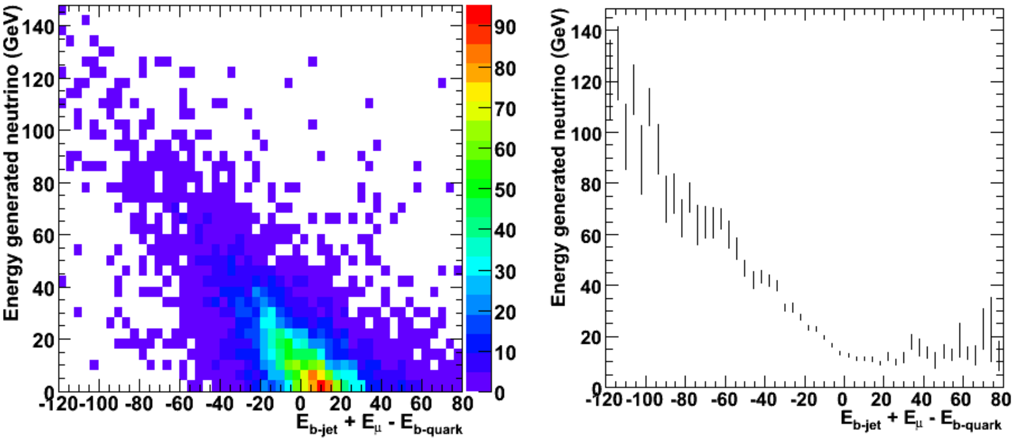

Figure 3.21: Two-dimensional histogram (left) and profile plot with constant bin size (right) of the absolute energy difference of b-jets with one "good $\mu$" versus the energy of the generated neutrino.

The method to estimate the neutrino energy consists out of several steps. First, a variable which is correlated with the energy of the generated neutrino needs to be found. As a consequence of the previous plot, this variable should also be correlated to the relative energy difference between the b-jet + muon and the b-quark. When such a variable is found, a profile plot with the same amount of data points in each bin is made (as explained in Section 3.2.3), with the energy of the generated neutrino on the y-axis and the variable on the x-axis. This plot is fitted with a straight line or with a parabola and the result of this fit is used as relation between the variable and the neutrino energy to estimate the neutrino energy using this variable. The estimated neutrino energy will be added to the b-jet (also the muon will be added of course) and the relative energy difference will be used as a benchmark to check if this gives a better estimation of the energy of the b-quark.

Figure 3.22 shows the plots used for the estimation of the neutrino energy with the energy of the b-jet as a variable. The results of this estimation is also shown. First it is checked if $E_{b-jet}$ is correlated with the energy of the generated neutrino and with the absolute energy difference between the b-jet

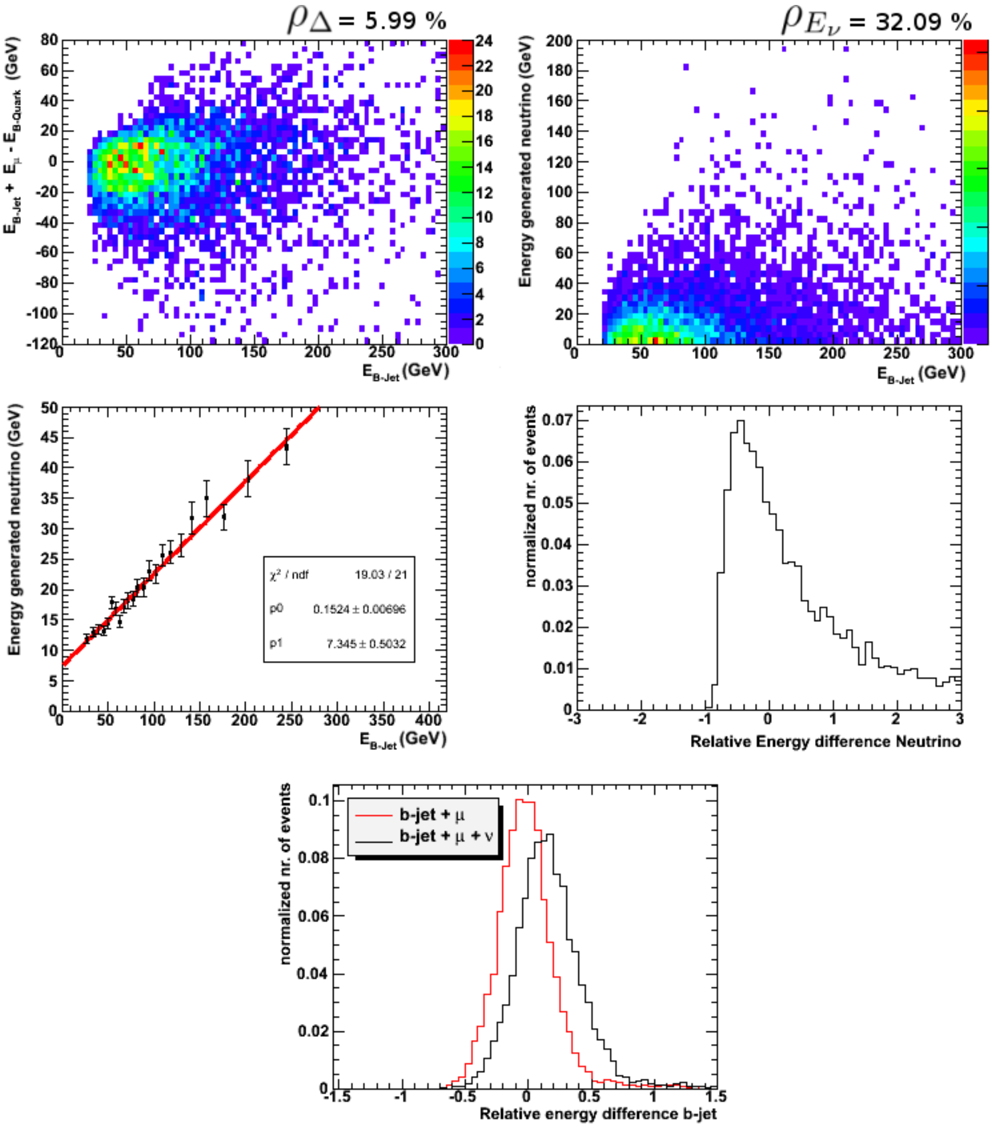

Figure 3.22: Estimation of the neutrino energy with $E_{b-jet}$ as variable. The function which is used in the fit is: $E_\nu = p_1 + p_0 \cdot E_{b-jet}$.

+ muon and the b-quark ($E_{b-jet} + E_\mu - E_{b-quark}$). The result is that $E_{b-jet}$ has a correlation coefficient of 32.09 % with the generated neutrino and 5.99 % with the absolute energy difference. These correlation coefficients will be called $\rho_{E_\nu}$ and $\rho_\Delta$ respectively in what follows. The next step is a profile plot of the energy of the generated neutrino in function of $E_{b-jet}$. This plot is made with the same amount of data point in each bin, as explained in Section 3.2.3. Afterwards, this plot is fitted with, in this case, a straight line ($E_\nu = p_1 + p_0 \cdot E_{b-jet}$) and the result of this fit is shown on the plot. This gives a relation between $E_{b-jet}$ and the generated neutrino energy, so for each b-jet with one "good $\mu$", this relation is used to estimate the neutrino energy from the b-jet energy. The relative energy difference of the neutrino ($(E_{\nu,estimated} - E_{\nu,generated})/E_{\nu,generated}$) is also shown. This shows that in the case of $E_{b-jet}$ as variable, the estimation is not so good. The relative energy of the b-jet, with addition of the muon energy and before and after the addition of the estimated neutrino energy is also plotted. Here it is also clear that the addition of the estimated neutrino energy with $E_{b-jet}$ as a variable gives no good results. The relative energy difference is wider and more shifted to higher values after the addition of the estimated neutrino energy.

This method was also applied with other variables. In Figure 3.23, the energy of the muons is used, while in Figure 3.24, the result of using the sum of the b-jet and muon energy as a variable is shown. Figure 3.25 contains the result of using the $p_{Trel}$ and, finally, in Figure 3.26, the pseudorapidity of the muon is used. These results are also summarised in Table 3.2. For each variable, the correlation coefficient between the variable and $E_{b-jet} + E_\mu - E_{b-quark}$ ($\rho_\Delta$), the correlation between the variable and the energy of the generated neutrino ($\rho_{E_\nu}$) and the mean and the standard deviation of the relative energy distribution after the addition of the estimated neutrino energy. The mean and the standard deviation of the relative energy distribution before any addition of the estimated neutrino energy is also shown for comparison.

| Variable | Correlation | | Relative energy difference | |
|---|---|---|---|---|
| | $\rho_\Delta$ | $\rho_{E_\nu}$ | Average | Standard deviation |
| *none* | \ | \ | $-0.013$ | 0.240 |
| $E_{b-jet}$ | 5.99 % | 32.09 % | 0.172 | 0.282 |
| $E_\mu$ | $-3.01$ % | 17.65 % | 0.190 | 0.283 |
| $E_{b-jet} + E_\mu$ | 4.24 % | 31.16 % | 0.169 | 0.278 |
| $pTrel$ | 18.11 % | $-9.80$ % | 0.308 | 0.295 |
| $\eta_\mu$ | 0.63 % | $-0.47$ % | 0.177 | 0.274 |

Table 3.2: Overview of the results of the method to estimate the energy of the neutrino.

All these corrections give similar results and they do not provide a good estimation of the real neutrino energy and they do not give an improved

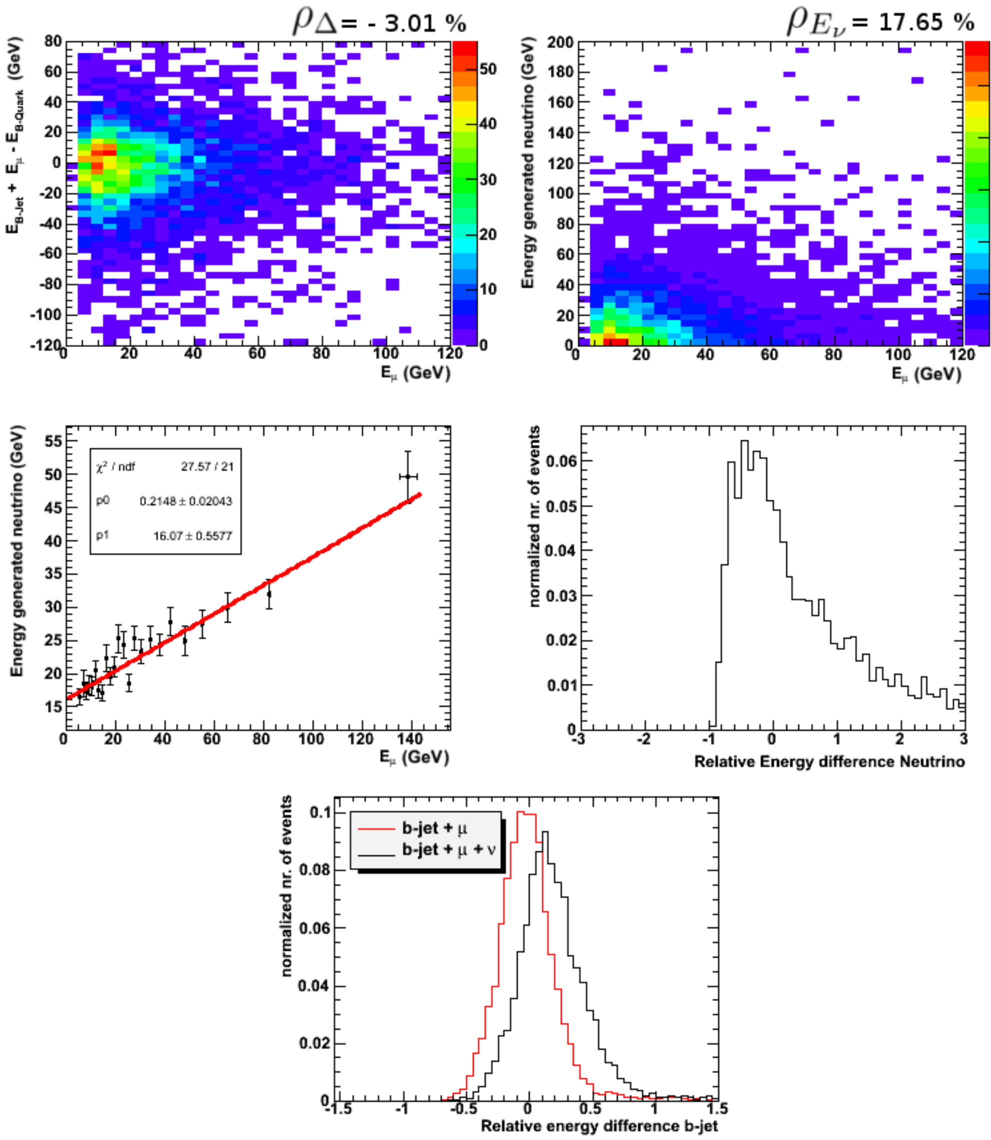

Figure 3.23: Estimation of the neutrino energy with $E_\mu$ as variable. The function which is used in the fit is: $E_\nu = p_1 + p_0 \cdot E_\mu$.

estimation of the energy of the b-quark. This method was also applied to a number of other variables, all with similar results. No single variable was found where this method gives an improved estimation of the b-quark energy.

It was illustrated that it would be good to add the neutrino energy to the b-jet, but the accurate estimation of this neutrino energy is far from trivial. This means that another method should be developed and tested to estimate the energy of the generated neutrino.

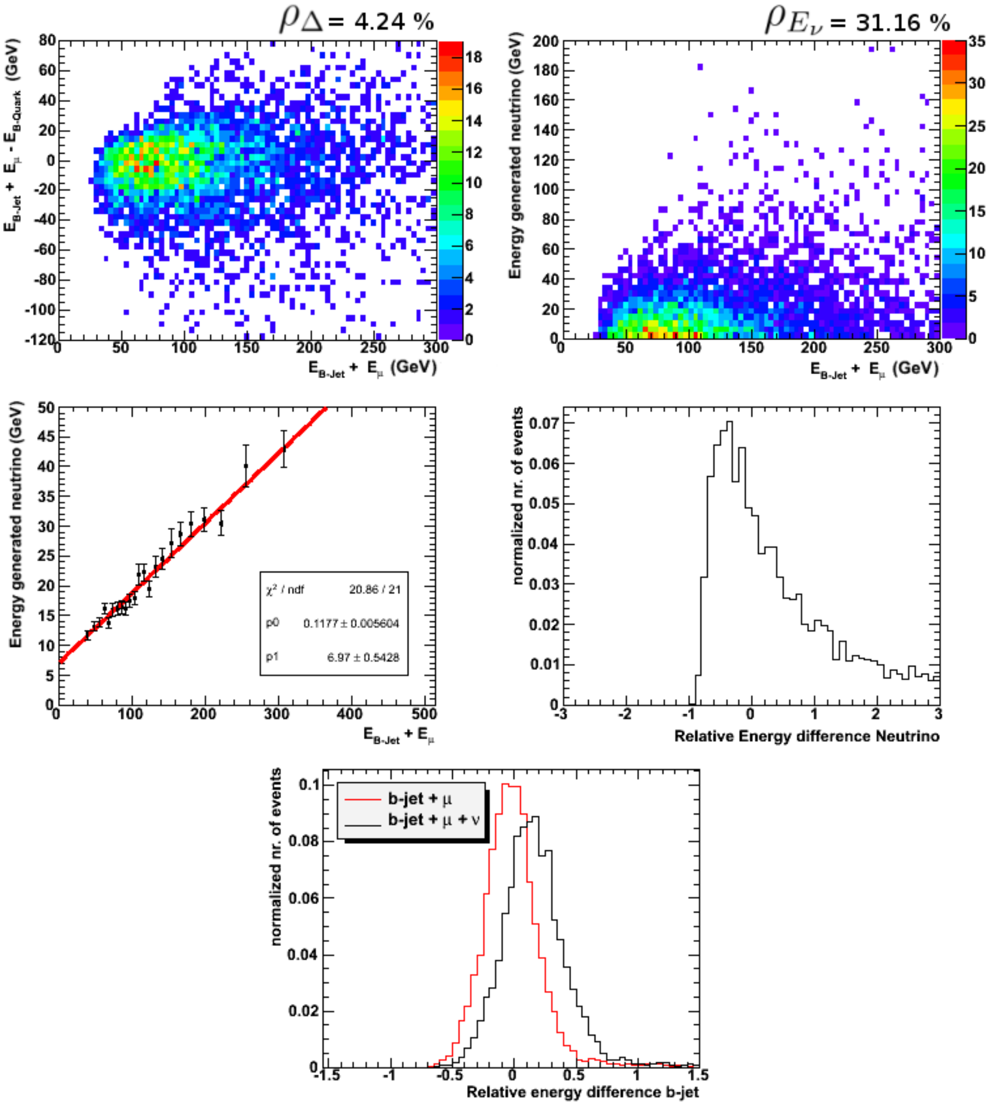

Figure 3.24: Estimation of the neutrino energy with $E_{b-jet} + E_{\mu}$ as variable. The function which is used in the fit is: $E_{\nu} = p_1 + p_0 \cdot (E_{b-jet} + E_{\mu})$.

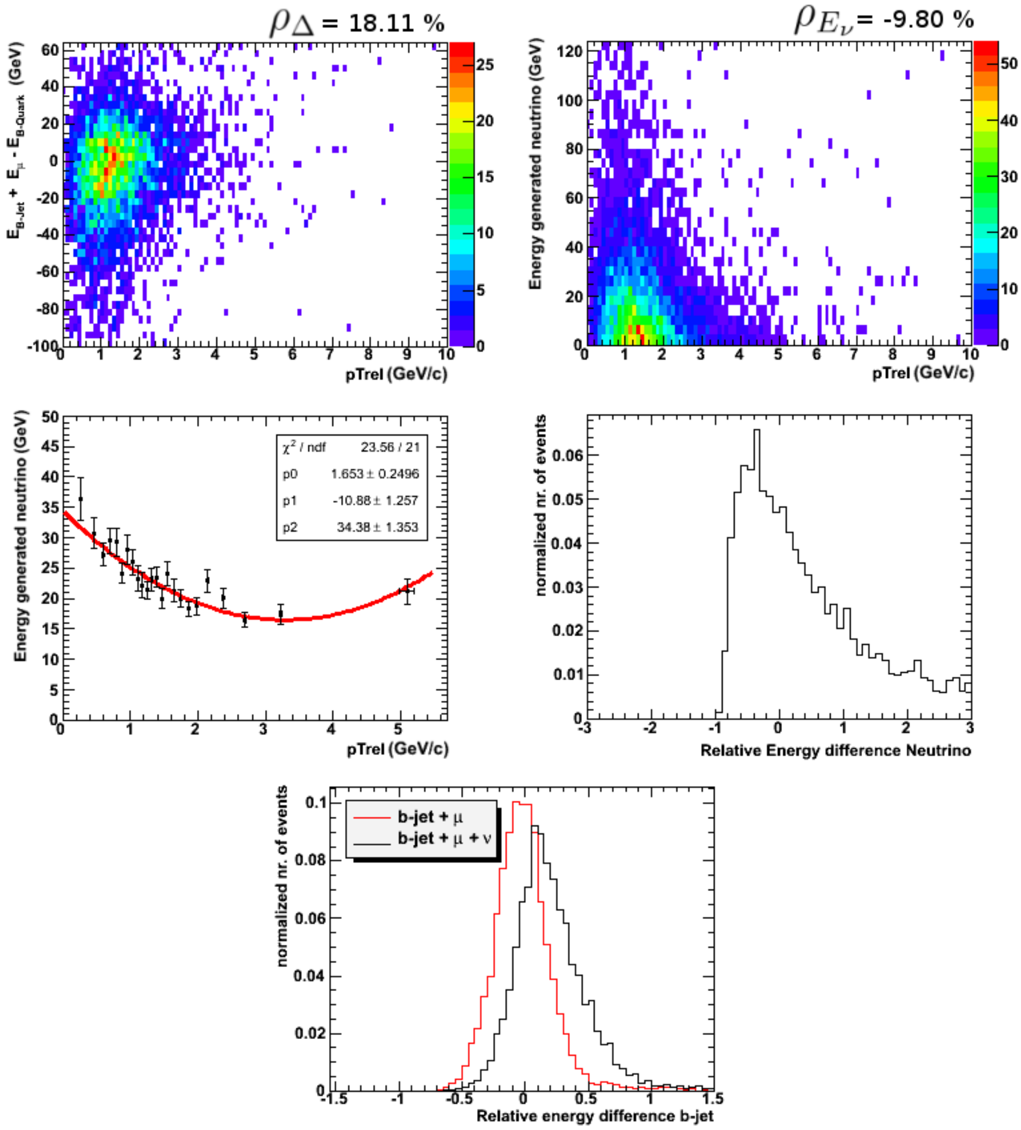

Figure 3.25: Estimation of the neutrino energy with $p_{Trel}$ as variable. The function which is used in the fit is: $E_\nu = p_2 + p_1 \cdot p_{Trel} + p_0 \cdot p_{Trel}^2$.

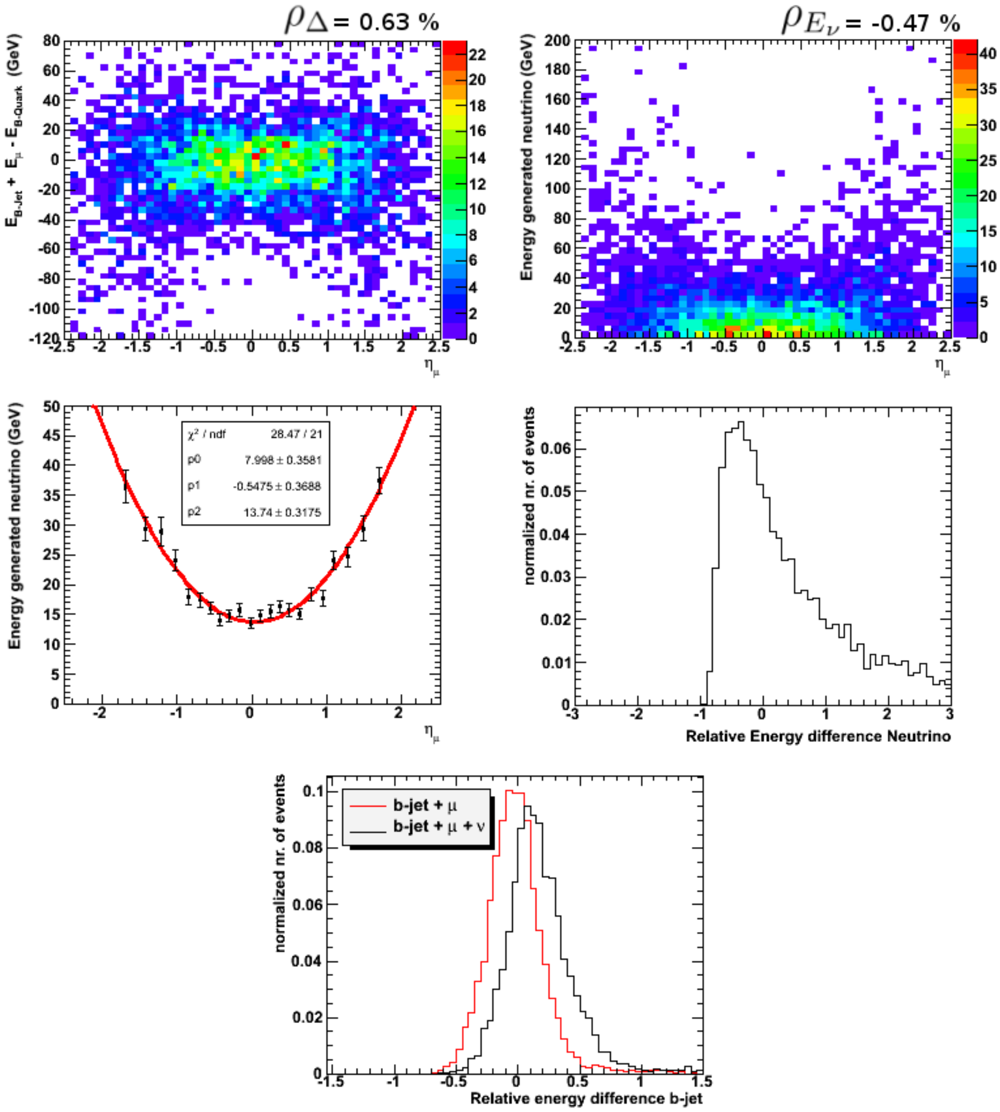

Figure 3.26: Estimation of the neutrino energy with $\eta_\mu$ as variable. The function which is used in the fit is: $E_\nu = p_2 + p_1 \cdot \eta_\mu + p_0 \cdot \eta_\mu^2$.

# Chapter 4

# B-Jets with muons in top quark physics

In the previous chapter, a matching between b-jets and muons was presented and the results of this matching were discussed. The effect of adding the matched muons to the b-jets was also given and finally, a method to estimate the energy of the neutrino which is produced together with this muon was worked out. Unfortunately, this method did not give a good estimate of the energy of the neutrino.

In this chapter, the tools from the previous one will be used to explore the effect on the top quark mass distribution. This was done with the same data sample as the previous section. Only the b-jets and the muons which survive the initial cuts from Section 3.2.1 will be used. For each b-jet, two other jets are looked for which, according to the genParticles, both originate from the hadronic decay of the W-boson, which on his turn originates from the same top quark as the b-jet. On the other two jets, the same initial cuts as on the b-jets are applied. The top quark mass will only be calculated if both other jets are found.

## 4.1 B-jets with and without muons

Like in the previous section, the first step is a comparison between b-jets which are matched with muons and b-jets which are not matched with muons. The matching procedure is exactly the same as in the previous chapter. The mass of the top quark which is calculated for both types of b-jets, is shown in Figure 4.1. Both histograms are fitted with a Gaussian. The range of this fit goes from the maximum of the histogram minus 50 GeV/c$^2$ to the maximum of the histogram plus 50 GeV/c$^2$. All the histograms of the top quark mass

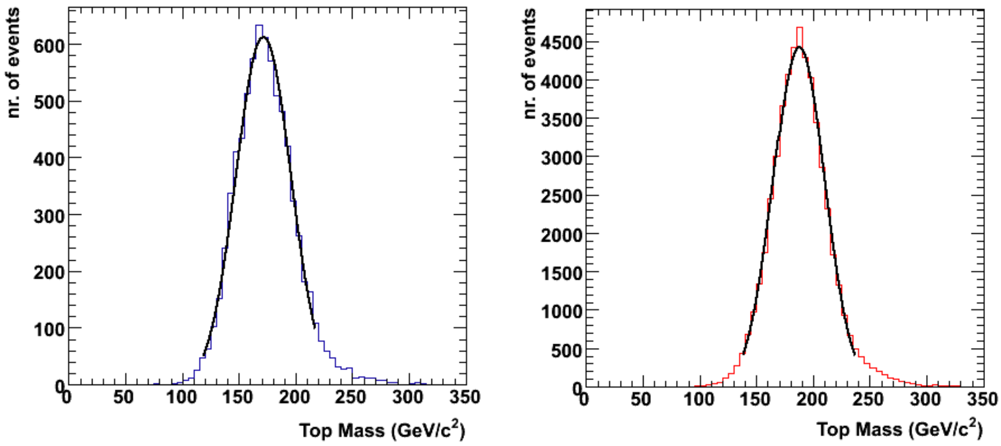

Figure 4.1: Mass of the top quark for b-jets with one or more muons (left) and for b-jets without muons (right).

in this section will be fitted within this range.

For b-jets without muons, the fit gives a top quark mass equal 187.2 ± 0.1 GeV/$c^2$, while for b-jets with muons, it equals 171.3 ± 0.3 GeV/$c^2$. When this is compared to the generated value of the top quark mass, being 172.4 GeV/$c^2$, it is clear that the bias is much larger for b-jets without muons than for b-jets with muons. This shift can be reduced by a better estimation of the jet energy scale. A lower value of the top quark mass for the b-jets without muons is expected, because here part of the b-quark energy escapes detection in the calorimeter by the production of one or more muons and one or more neutrino's. When the width of the fit of both histograms is compared to each other, it is also clear that the histogram of b-jets without muons is more narrow (a width of 22.8 ± 0.1 GeV/$c^2$) than the histogram of the b-jets with muons (a standard deviation of 23.9 ± 0.3 GeV/$c^2$). The goal of the better estimation of the b-quark energy is to result in a narrower the top quark mass distribution for b-jets with muons.

## 4.2 Adding the muons to the b-jets

The next step is adding the energy and momenta of the muons which are matched to a b-jet, to this b-jet and recalculate the mass of the top quark after this addition. The result of this addition is shown in Figure 4.2 for b-jets with one or more muons. As more or less expected, the addition of the muons to the b-jets gives a more narrow distribution, with a fitted width of 23.9 ± 0.3 GeV/$c^2$ before and 22.4 ± 0.3 GeV/$c^2$ after this addition. It

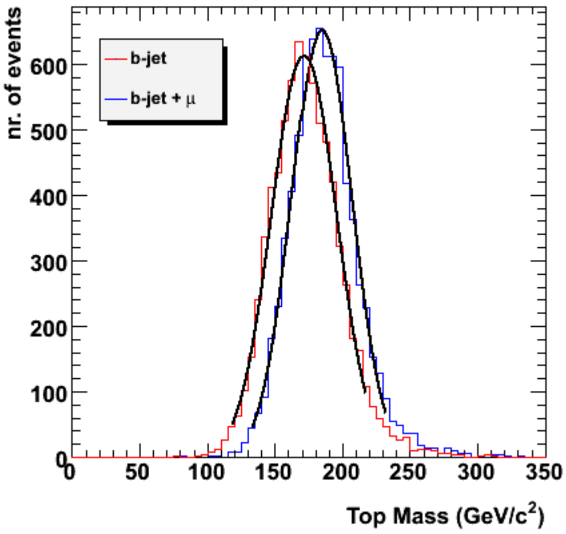

Figure 4.2: Mass of the top quark for b-jets with one or more muons before and after the addition of the muon momenta.

also introduces a shift in the fitted top quark mass, which equals 171.3 ± 0.3 GeV/c$^2$ before and 184.4 ± 0.3 GeV/c$^2$ after the addition of the muons. This shifted value is much closer to the value of the b-jets without muons, but what is more important, the width after the addition of the muons is smaller than the width of distribution with b-jets without muons. This means that the addition will give a clear improvement in a proper estimation of the top quark mass.

Like in the previous chapter, the same results are also shown for b-jets which are matched with exactly one muon. This is done in Figure 4.3. Here, the fitted top mass shifts from 171.7 ± 0.3 GeV/c$^2$ to 184.5 ± 0.3 GeV/c$^2$ by the addition of the muon and the width of the Gaussian reduces from 23.6 ± 0.3 GeV/c$^2$ to 22.3 ± 0.3 GeV/c$^2$. Both changes are about the same as for the previous histograms.

## 4.3 The neutrino's

When a good estimation of the energy and momentum of the neutrino which is produced together with the muon is available, this estimated energy and momentum should also be added to the b-jet. The method from the previous chapter did not give a good estimation of the neutrino energy. A method to estimate the momenta of the neutrino was not yet worked out, because the addition of the estimated neutrino energy to the b-jet gave a worse (wider

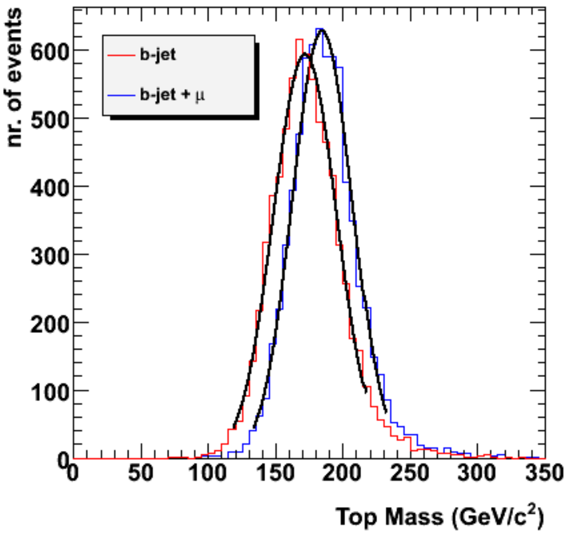

Figure 4.3: Mass of the top quark for b-jets with one muon before and after the addition of the energy and momentum of this muon.

and shifted to higher values) distribution of the relative energy difference.

As a motivation that the addition of a good estimation of energy and momenta of the neutrino can give good results, the generated neutrino is added to the b-jet + muon system and from this the value of the top quark mass is calculated. This is only done for the b-jets with one "good $\mu$", as explained in Section 3.4. This means that the muon needs to be produced together with a neutrino by the decay of a b-hadron. This b-hadron needs to be formed by the hadronization of a b-quark, which is a result of a top quark decay.

The mass of the top quark of these b-jets before and after the addition of the muon and the generated neutrino is shown in Figure 4.4. The fitted top quark masses for these three categories are $165.1 \pm 0.5$ GeV/c$^2$ for the b-jets, $182.6 \pm 0.5$ GeV/c$^2$ after the addition of the muon and $193.9 \pm 0.5$ GeV/c$^2$ after the addition of the muon and the generated neutrino. The widths of the fitted Gaussian change from $22.9 \pm 0.5$ GeV/c$^2$ for the b-jets, $22.1 \pm 0.5$ GeV/c$^2$ with the addition of the muons and $20.8 \pm 0.4$ GeV/c$^2$ when the generated neutrino and the muon is added to the b-jet. The same behaviour as in the previous plots of the top quark mass is found. The addition of the muon introduces a shift to a higher value of the top quark mass, but it results in a more narrow distribution and when the generated neutrino is added, the shift becomes larger, but the distribution also becomes more narrow. This shows that a good estimation of the energy and momentum of the neutrino,

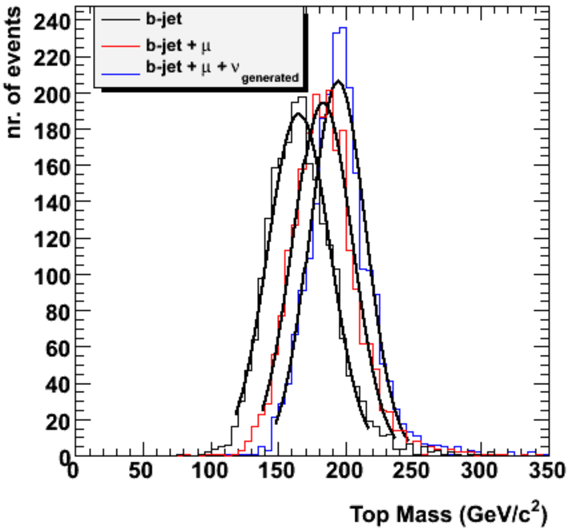

Figure 4.4: Mass of the top quark for b-jets with one “good $\mu$”, before and after the addition of the muon and the generated neutrino.

which is produced together with the muon, can give good results when it is, like the muon, added to the b-jet. The shift can be reduced with a better estimation of the total jet energy scale, but after the removal of this shift, the width of the distribution is expected to remain smaller after the addition of the muon and the neutrino.

# Chapter 5

# Top quark selection

When CMS will start taking data, the identification of $t\bar{t}$ events will be far from trivial, because they will be buried inside the background, which has a much higher cross section. In this chapter, a selection will be presented which aims to reduce this background using jets which contain muons. The idea is to use the fact that in each $t\bar{t}$ event, two b-jets are produced and part of these b-jets will have a muon inside them. The background events are not expected to have such a large amount of jets containing a muon. The method will be applied for the selection of semi-muonic $t\bar{t}$ events.

## 5.1 Data Samples

All the data samples for this study are produced with `MadGraph` [33]. In these samples, the mass of the top quark was set to 172.4 GeV/c$^2$. The generated events were simulated with the full CMS detector simulation and the jets were reconstructed with the SISCone algorithm [35] and corrected with the L2L3 corrections. There were also some loose initial cuts applied on these data samples to remove events which will certainly not be used. An event was asked to have at least one muon with a transverse momentum higher than 15 GeV/c and at least four jets with a transverse momentum also higher than 15 GeV/c. An overview of all the samples is shown in Table 5.1, together with the cross-section, number of events after these initial cuts, the selection efficiency of these initial cuts and the corresponding integrated luminosity of each sample. The names in the table need some extra explanation of which events are inside each sample. The TTJets sample contains $t\bar{t}$ pairs which are produced together with up to three extra jets from radiation effects. The cross section for this sample is the NLO value from Section 2.1. The three SingleTop samples contain, as their name suggests, events where only one top

| Data Sample | $\sigma$ (pb) | # of events | $\epsilon_{selection}$ | $L$ (pb$^{-1}$) |
|---|---|---|---|---|
| TTJets | 414 | 218 404 | 0.2307 | 2 287 |
| SingleTop s channel | 5 | 1 352 | 0.1127 | 2 400 |
| SingleTop t channel | 130 | 21 868 | 0.0780 | 2 167 |
| SingleTop tW channel | 29 | 27 932 | 0.1652 | 5 829 |
| VQQ | 290 | 11 947 | 0.0125 | 3 300 |
| WJets | 45 600 | 57 876 | 0.0059 | 214 |
| Wc | 1 487 | 33 683 | 0.0116 | 1 960 |
| ZJets | 3 700 | 13 857 | 0.0110 | 299 |
| QCD100-250 | 15 000 000 | 22 005 | 0.0018 | 0.82 |
| QCD250-500 | 400 000 | 114 583 | 0.0226 | 12.7 |
| QCD500-1000 | 14 000 | 447 904 | 0.0976 | 327 |
| QCD1000-Inf | 370 | 259 769 | 0.2507 | 2 856 |

Table 5.1: The data samples used in this section and their cross section, number of events after the initial cuts, selection efficiency of these initial cuts and the corresponding integrated luminosity.

quark is produced via the corresponding channel. The VQQ sample contains events where a leptonically decaying vector boson ($Z$ or $W$) is produced together with two heavy quarks ($b$ or $c$). The WJets contains a leptonically decaying $W$ boson which is produced together with up to four jets, where at least one of these jets originates from an up quark, a down quark or a gluon. B-jets are not included here. The Wc sample also contains leptonically decaying $W$ bosons, but here they are produced together with one c-jet. The ZJets sample contains a leptonically decaying $Z$ boson, which is produced together with up to four jets. Again, b-jets are not included. Finally, the QCD samples each contain multi-jet events which have a scalar sum of the transverse energy of the jets ($H_T$) within the interval (in GeV) of the sample's name. From these explanations, it is clear that some samples contain parts which are overlapping with other samples. The VQQ sample has a small overlap with the WJets and the ZJets samples.

As already mentioned, an event was asked to have at least one muon with a transverse momentum higher than 15 GeV/c and at least four jets with a transverse momentum also higher than 15 GeV/c. Not only the events, but also the individual muons and jets not passing these criteria were removed from the data files. This is not optimal, since a significant fraction of the muons which are matched with a jet will have a lower transverse momentum, but these data files were originally not produced for this purpose. The effect of these cuts will be discussed in the final section of this chapter.

## 5.2 Selection of the $t\bar{t}$ signal

The selection of semi-muonic $t\bar{t}$ events consists out of several steps. First, the standard cuts, which are used in a lot of analyses in CMS, will be applied. These consist out of cuts on the jets and the muon isolation cuts. Afterwards, some additional cuts will be presented, which will use a matching between jets and muons. As already mentioned in the description of the data samples, the events were already subjected to some initial cuts. To pass these cuts, an event was asked to have at least one muon with a transverse momentum higher than 15 GeV/c and at least four jets with a transverse momentum also higher than 15 GeV/c.

The number of events passing the cuts of each data sample will all be rescaled to an integrated luminosity of 200 $\text{pb}^{-1}$. This is the foreseen integrated luminosity that will be collected with the CMS detector during the first (2009 - 2010) LHC run.

### 5.2.1 Standard muon isolation cuts

CMS has developed standard muon isolation cuts [36], which can be used in analyses where an isolated muon needs to be identified, like for example in the selection of semi-muonic $t\bar{t}$ events. Together with these isolation cuts, some cuts on the jets are also applied.

An event is passing the cuts if it contains:

- At least 4 jets with $p_T > 30$ GeV/c and $|\eta| < 2.4$.
- At least 1 muon with:
  - $p_T > 30$ GeV/c and $|\eta| < 2.1$.
  - The muon is a GlobalMuon.
  - A globalized $\chi^2/ndf < 10$.
  - Transverse impact parameter $|d_0| < 0.2$ mm.
  - Number of hits in the silicon tracker $\geq 11$.
  - In HCAL veto cone ($\Delta R = 0.1$) around muon: $E_T < 6$ GeV.
  - In ECAL veto cone ($\Delta R = 0.07$) around muon: $E_T < 4$ GeV.

- Combined Relative Isolation ($CombRelIso$) $< 0.1$.
  $CombRelIso = \frac{EcalIso+HcalIso+TrkIso}{p_T^{\mu}}$, where $EcalIso$ and $HcalIso$ are the transverse energies deposited inside a cone of $\Delta R = 0.3$ around the muon track in the ECAL and HCAL respectively and $TrkIso$ is the sum of the transverse momenta of all the tracks inside a cone of $\Delta R = 0.3$ around the muon track.

The number of events before and after applying these cuts can be found in Table 5.2. The signal-to-background ratio, the significance of the signal ($\mathrm{S}/\sqrt{\mathrm{B}}$) and the significance of the signal when taking into account a systematic uncertainty equal to 20 % of the background rate ($\mathrm{S}/\sqrt{\mathrm{B} + (0.2\ \mathrm{B})^2}$) is shown when only semi-muonically decaying $t\bar{t}$ events are considered as signal, when all $t\bar{t}$ events are considered as signal and when all events containing a top quark ($t\bar{t}$ and single top) are considered as the signal. The denominator in the formula for the significance of the signal when taking into account a systematic uncertainty of 20 % is the square root of the quadratic sum of the statistical uncertainty on the background ($\sqrt{\mathrm{B}}$) and the systematical uncertainty on the background (0.2 B).

These cuts clearly improve the signal-to-background ratio and the two significances of the signal, since a large fraction of the other events are not expected to contain a significant amount of isolated muons. The main background is WJets, as this is also expected to contain a significant amount of jets and also isolated muons produced in the decay of a W boson to a muon and a neutrino. The second largest background are all the other events containing one or more top quarks (other $t\bar{t}$ decays and single top), since they will also give rise to a significant amount of isolated muons by the same decay mechanism as the semi-muonic $t\bar{t}$ events.

### 5.2.2 Additional cuts

After the application of the standard cuts from the previous section, some additional cuts will be worked out to improve the selection of semi-muonic $t\bar{t}$ events. The first step is, like in the previous chapters, a $\Delta R$ matching between the jets and the muons. This matching will only be performed between the jets surviving the standard cuts from the previous section ($p_T >$ 30 GeV/c and $|\eta| < 2.4$) and the muons which are not isolated, according to the cuts from the previous section. Only GlobalMuons will be used in the matching. A muon is matched to a jet if

$$\Delta R(jet, \mu) < 0.5 \tag{5.1}$$

| Data Sample | initial cuts | | standard cuts | |
|---|---|---|---|---|
| | sample | 200 pb$^{-1}$ | sample | 200 pb$^{-1}$ |
| TTJets semi-$\mu$ | 100 048 | 8 749 | 32 116 | 2 809 |
| TTJets other | 118 356 | 10 350 | 7 903 | 691 |
| SingleTop s | 1 352 | 113 | 79 | 7 |
| SingleTop t | 21 868 | 2 018 | 1 611 | 149 |
| SingleTop tW | 27 932 | 958 | 3 127 | 107 |
| VQQ | 11 947 | 724 | 141 | 9 |
| WJets | 57 876 | 54 090 | 1 429 | 1 336 |
| Wc | 33 638 | 3 089 | 429 | 39 |
| ZJets | 13 857 | 9 269 | 421 | 282 |
| QCD100-250 | 22 005 | 5 367 073 | 1 | 244 |
| QCD250-500 | 114 583 | 1 804 457 | 4 | 63 |
| QCD500-1000 | 447 904 | 273 947 | 10 | 6 |
| QCD1000-Inf | 259 769 | 18 191 | 3 | 0 |
| S/B (S = semi-$\mu$ $t\bar{t}$) | | 0.001 | | 0.958 |
| S/$\sqrt{\mathrm{B}}$ (S = semi-$\mu$ $t\bar{t}$) | | 3.19 | | 51.9 |
| S/$\sqrt{\mathrm{B} + (0.2\ \mathrm{B})^2}$ (S = semi-$\mu$ $t\bar{t}$) | | 0.006 | | 4.77 |
| S/B (S = all $t\bar{t}$) | | 0.003 | | 1.56 |
| S/$\sqrt{\mathrm{B}}$ (S = all $t\bar{t}$) | | 6.96 | | 73.9 |
| S/$\sqrt{\mathrm{B} + (0.2\ \mathrm{B})^2}$ (S = all $t\bar{t}$) | | 0.013 | | 7.77 |
| S/B (S = all top) | | 0.003 | | 1.90 |
| S/$\sqrt{\mathrm{B}}$ (S = all top) | | 8.09 | | 84.6 |
| S/$\sqrt{\mathrm{B} + (0.2\ \mathrm{B})^2}$ (S = all top) | | 0.014 | | 14.8 |

Table 5.2: The number of events before and after the standard cuts. Both the number of events remaining in the data sample and the number of events for 200 pb$^{-1}$ are shown.

For each event, the number of jets which are matched to at least one muon will be counted and this number will be used to select events. In a first step, only events containing at least one jet matched to one or more muons are selected. The results of this cut can be seen in Table 5.3. It clearly raises the signal-to-background ratio, it reduces the significance of the signal, but it improves the significance of the signal when a systematic uncertainty is

| Data Sample | standard cuts | | $\geq 1$ jet with $\mu$ | | $\mathrm{M}_{\mathrm{inv}}^{\mu\mu}$ cut | |
|---|---|---|---|---|---|---|
| | sample | 200 pb$^{-1}$ | sample | 200 pb$^{-1}$ | sample | 200 pb$^{-1}$ |
| TTJets semi-$\mu$ | 32 116 | 2 809 | 2 787 | 244 | 2 600 | 227 |
| TTJets other | 7 903 | 691 | 867 | 76 | 814 | 71 |
| SingleTop s | 79 | 7 | 5 | 0 | 5 | 0 |
| SingleTop t | 1 611 | 149 | 154 | 14 | 145 | 13 |
| SingleTop tW | 3 127 | 107 | 233 | 8 | 212 | 7 |
| VQQ | 141 | 9 | 25 | 2 | 16 | 1 |
| WJets | 1 429 | 1 336 | 18 | 17 | 17 | 16 |
| Wc | 429 | 39 | 12 | 1 | 10 | 1 |
| ZJets | 421 | 282 | 54 | 36 | 30 | 20 |
| QCD100-250 | 1 | 244 | 0 | 0 | 0 | 0 |
| QCD250-500 | 4 | 63 | 0 | 0 | 0 | 0 |
| QCD500-1000 | 10 | 6 | 1 | 1 | 1 | 1 |
| QCD1000-Inf | 3 | 0 | 1 | 0 | 1 | 0 |
| S/B (semi-$\mu$ $t\bar{t}$) | | 0.958 | | 1.58 | | 1.79 |
| S/$\sqrt{\mathrm{B}}$ (semi-$\mu$ $t\bar{t}$) | | 51.9 | | 19.6 | | 19.9 |
| S/$\sqrt{\mathrm{B} + (0.2\ \mathrm{B})^2}$ (semi-$\mu$ $t\bar{t}$) | | 4.77 | | 7.31 | | 7.97 |
| S/B (all $t\bar{t}$) | | 1.56 | | 4.05 | | 5.01 |
| S/$\sqrt{\mathrm{B}}$ (all $t\bar{t}$) | | 73.9 | | 36.0 | | 38.7 |
| S/$\sqrt{\mathrm{B} + (0.2\ \mathrm{B})^2}$ (all $t\bar{t}$) | | 7.77 | | 17.7 | | 21.0 |
| S/B (all top) | | 1.90 | | 6.08 | | 8.30 |
| S/$\sqrt{\mathrm{B}}$ (all top) | | 84.6 | | 45.6 | | 51.5 |
| S/$\sqrt{\mathrm{B} + (0.2\ \mathrm{B})^2}$ (all top) | | 14.8 | | 37.7 | | 48.1 |

Table 5.3: The number of events before and after asking at least one jet which is matched to one ore more muons and after the cut on $\mathrm{M}_{\mathrm{inv}}^{\mu\mu}$. Both the number of events remaining in the data sample and the number of events for 200 pb$^{-1}$ are shown.

taken into account. As mentioned in the description of all the data samples, the VQQ sample overlaps slightly with the WJets and the ZJets samples, but since after this cut, only two VQQ events remain, the effect of this overlap on the signal-to-background ratio and on the significances can be neglected.

Except from other events containing top quarks, the main background are ZJets events, so an additional cut will be applied to remove a significant fraction of these events containing a $Z$ boson. In each event passing the standard cuts and containing at least one jet matched with one or more muons, the two muons with the highest transverse momentum will be combined to calculate the invariant mass of this 2 muon system, $\mathrm{M}_{\mathrm{inv}}^{\mu\mu}$. This invariant mass is shown in Figure 5.1. A clear peak can be observed in the ZJets events around an invariant mass of 90 GeV/c$^2$, originating from the decay of the $Z$ boson into two muons. This can be used as an additional cut. The events which have a value of $\mathrm{M}_{\mathrm{inv}}^{\mu\mu}$ which is close to the mass of the $Z$ boson will be removed, so to pass the cut, an event must have

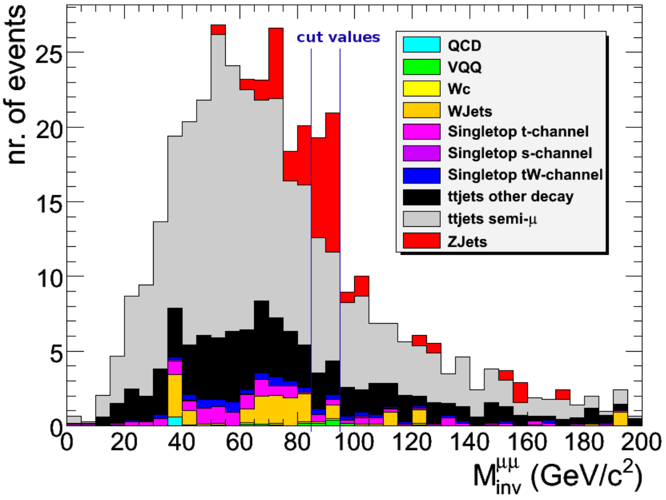

Figure 5.1: The invariant mass of the two muons with the highest transverse momentum. The events between the two blue lines of the cut values will be removed in the selection.

$$\mathrm{M}_{\mathrm{inv}}^{\mu\mu} < 85\ \mathrm{GeV/c}^2 \text{ or } \mathrm{M}_{\mathrm{inv}}^{\mu\mu} > 95\ \mathrm{GeV/c}^2 \tag{5.2}$$

The results of this cut can also be found in Table 5.3. This cut clearly improves the signal-to-background ratio and the significances. The significance of the signal without the systematic uncertainty is still below the values immediately after the standard cuts, but after the inclusion of the systematic uncertainty, the significance improves even further. Despite this cut, ZJets still remains the biggest background which does not contain top quarks.

For completeness, it was also checked if the requirement to have at least 2 jets containing one or more muons gives good results. This cut was applied after the cut on $\mathrm{M}_{\mathrm{inv}}^{\mu\mu}$ and the results can be found in Table 5.4, together with the results of the previous cuts. After this cut, only 16 events remain for 200

| Data Sample | $\geq 1$ jet with $\mu$ | | $M_{\text{inv}}^{\mu\mu}$ cut | | $\geq 2$ jets with $\mu$ | |
|---|---|---|---|---|---|---|
| | sample | 200 pb$^{-1}$ | sample | 200 pb$^{-1}$ | sample | 200 pb$^{-1}$ |
| TTJets semi-$\mu$ | 2 787 | 244 | 2 600 | 227 | 114 | 10 |
| TTJets other | 867 | 76 | 814 | 71 | 50 | 4 |
| SingleTop s | 5 | 0 | 5 | 0 | 0 | 0 |
| SingleTop t | 154 | 14 | 145 | 13 | 6 | 1 |
| SingleTop tW | 233 | 8 | 212 | 7 | 6 | 0 |
| VQQ | 25 | 2 | 16 | 1 | 1 | 0 |
| WJets | 18 | 17 | 17 | 16 | 0 | 0 |
| Wc | 12 | 1 | 10 | 1 | 0 | 0 |
| ZJets | 54 | 36 | 30 | 20 | 1 | 1 |
| QCD100-250 | 0 | 0 | 0 | 0 | 0 | 0 |
| QCD250-500 | 0 | 0 | 0 | 0 | 0 | 0 |
| QCD500-1000 | 1 | 1 | 1 | 1 | 0 | 0 |
| QCD1000-Inf | 1 | 0 | 1 | 0 | 0 | 0 |
| S/B (semi-$\mu$ $t\bar{t}$) | | 1.58 | | 1.79 | | 1.70 |
| S/$\sqrt{\text{B}}$ (semi-$\mu$ $t\bar{t}$) | | 19.6 | | 19.9 | | 4.12 |
| S/$\sqrt{\text{B} + (0.2\ \text{B})^2}$ (semi-$\mu$ $t\bar{t}$) | | 7.31 | | 7.97 | | 3.71 |
| S/B (S = all $t\bar{t}$) | | 4.05 | | 5.01 | | 9.63 |
| S/$\sqrt{\text{B}}$ (S = all $t\bar{t}$) | | 36.0 | | 38.7 | | 11.8 |
| S/$\sqrt{\text{B} + (0.2\ \text{B})^2}$ (all $t\bar{t}$) | | 17.7 | | 21.0 | | 11.4 |
| S/B (S = all top) | | 6.08 | | 8.30 | | 20.7 |
| S/$\sqrt{\text{B}}$ (S = all top) | | 45.6 | | 51.5 | | 17.7 |
| S/$\sqrt{\text{B} + (0.2\ \text{B})^2}$ (all top) | | 37.7 | | 48.1 | | 23.9 |

Table 5.4: The number of events after asking at least two jets which are matched to one or more muons, together with the results of the previous cuts. Both the number of events remaining in the data sample and the number of events for 200 pb$^{-1}$ are shown.

pb$^{-1}$ of integrated luminosity. The significances are strongly reduced for all the signals used. For the semi-muonic $t\bar{t}$, the signal-to-background ratio is even reduced. This clearly shows that this cut is too stringent, so it will not be used anymore in what follows.

## 5.3 Reconstructing the top quark mass

The mass of the top quark will be reconstructed with the M3 method, as explained in, for example, [37]. This method aims at reconstructing the mass of the hadronically decaying top quark in semi-muon $t\bar{t}$ events. The three jets which give the highest vectorial summed $E_T$ are combined and used to calculate an invariant mass. This variable is called M3 and it is plotted in Figure 5.2 for the events after the standard cuts, the events after the cut asking one jet containing one or more muons and the events after the cut on the invariant mass of the two muons with the highest transverse momentum ($\mathrm{M}_{\mathrm{inv}}^{\mu\mu}$). These histograms show that the three M3 distributions clearly show a peak, but at a value which is higher than the generated mass of the top quark (172.4 GeV/c$^2$), like in the previous section. The peak is not very narrow, since there will be a significant background caused by a wrong combination of three jets.

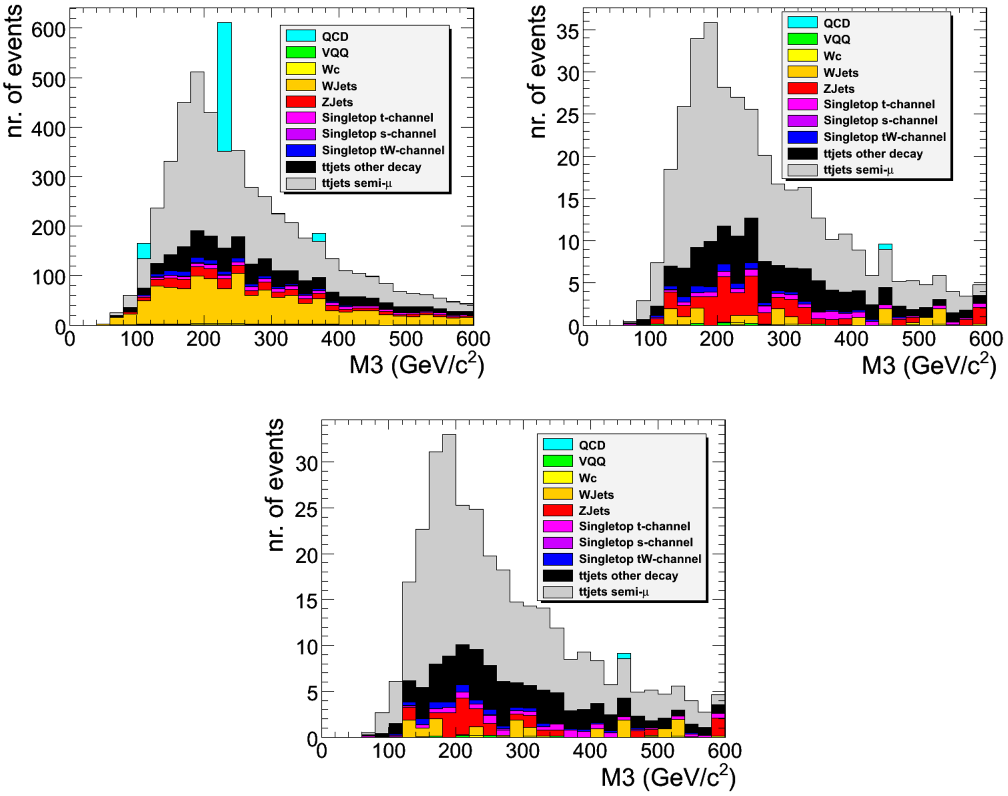

Figure 5.2: The M3 variable for the events after the standard cuts (top left), after the cut asking one jet containing at least one muon (top right) and after the cut on $\mathrm{M}_{\mathrm{inv}}^{\mu\mu}$ (bottom).

The histogram with the M3 variable for the events immediately after the standard cuts (the top left histogram of Figure 5.2) shows a sharp QCD peak around a value of 230 GeV/c$^2$ and also some smaller QCD peaks around values of 110 GeV/c$^2$ and 370 GeV/c$^2$. These peaks are there because the QCD100-250 and QCD250-500 samples are only generated for a very small integrated luminosity (0.82 pb$^{-1}$ and 12.7 pb$^{-1}$ respectively) and the scaling to an integrated luminosity of 200 pb$^{-1}$ gives a very large weight to these events. The M3 invariant mass of the QCD100-250 and QCD250-500 data samples after asking more than one jet with at least one muon is shown in Figure 5.3. These are the histograms before the rescaling to 200 pb$^{-1}$ and they have events at exactly the same places as the peaks in the top left histogram of Figure 5.2. For the rescaling to 200 pb$^{-1}$, these two histograms are multiplied by a factor 243.90 (200/0.82) and 15.75 (200/12.7), respectively. This explains these peaks and also the height of the peaks. When these QCD data samples would be generated for a much higher integrated luminosity, these peaks would not be there, but since they are already gone in the other M3 histograms, the effect of the small integrated luminosity on these other M3 histograms will be small.

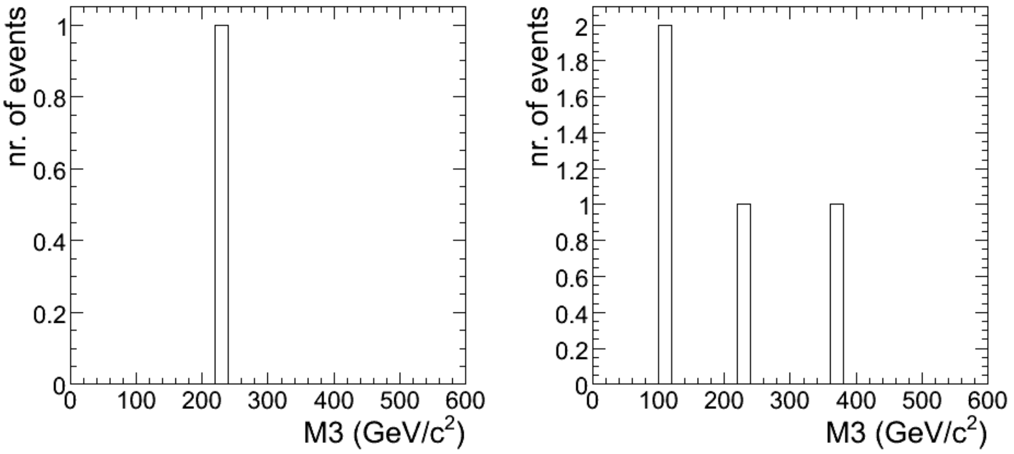

Figure 5.3: The M3 variable for the QCD100-250 sample (left) and for the QCD250-500 sample (right), both before rescaling to 200 pb$^{-1}$.

## 5.4 Effect of the initial muon transverse momentum cut

In the description of the data samples, it was explained that all the muons which did not pass the initial muon cuts ($p_{T,\mu} > 15$ GeV/c) were discarded in the production of the data samples, even for events which did pass all the initial cuts. Since a large amount of muons which are produced inside b-jets will have a transverse momentum which is lower than 15 GeV/c, the effect of this initial cut needs to be studied. For this, a TTJets data sample without the initial cuts, which is produced on exactly the same way, will be used and it will be compared to the TTJets data sample with the initial cuts. The background samples could however not be produced in the scope of this thesis.

On these data samples, the same cuts are applied, except for the matching between jets and muons. In Section 5.2.2, the muons did not have to pass a cut on the transverse momentum, since they already have a transverse momentum above 15 GeV/c. Here, the muons used in the matching need to have a transverse momentum above 3 GeV/c. The transverse momentum of the muons which are matched to a jet is shown in Figure 5.4, for semi-muonic $t\bar{t}$ events which are not subjected to the initial cuts and which are passing the standard cuts. It is clear that a cut of 15 GeV/c on the transverse momentum of all the muons removes a large amount of semi-muonic $t\bar{t}$ events, so it should be checked what the effect is of lowering or removing this cut.

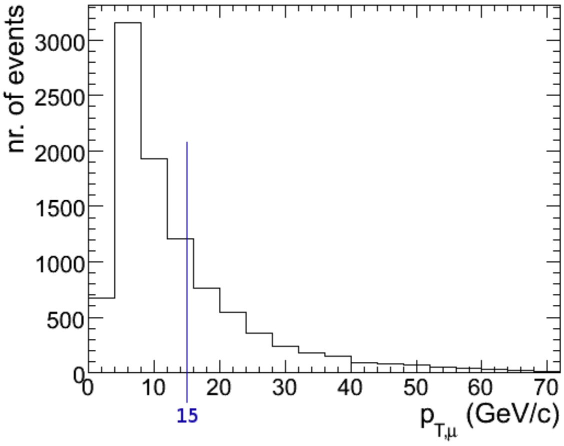

Figure 5.4: The transverse momentum of the muons which are matched to a jet, only for semi-muonic $t\bar{t}$ events without initial cuts.

The number of events and the ratio of semi-muonic to other $t\bar{t}$ events after each cut can be found in Table 5.5. Without the initial cuts, there are

| Data Sample | # of events | standard cut | $\geq 1$ jet with $\mu$ | $M_{\text{inv}}^{\mu\mu}$ cut |
|---|---|---|---|---|
| semi-$\mu$, with cuts | 100 048 | 32 116 | 2 787 | 2 600 |
| other, with cuts | 118 356 | 7 903 | 867 | 814 |
| $\frac{\text{semi-}\mu}{\text{other}}$, with cuts | 0.845 | 4.064 | 3.215 | 3.194 |
| semi-$\mu$, no cuts | 140 001 | 32 116 | 8 431 | 8 134 |
| other, no cuts | 806 643 | 7 903 | 2 142 | 2 061 |
| $\frac{\text{semi-}\mu}{\text{other}}$, no cuts | 0.174 | 4.064 | 3.936 | 3.947 |

Table 5.5: The number of events and the ratio of semi-$\mu$ to other $t\bar{t}$ events after each cut of the TTJets data samples with and without initial cuts.

almost three times more events surviving the cut which asks at least one jet containing a muon. The ratio of the number of semi muonic $t\bar{t}$ events to the number of other $t\bar{t}$ events is also larger for the events without the initial cuts. This shows that lowering the cut on the transverse momentum of the muons gives good results when only other $t\bar{t}$ decays are considered as background. It is hard to predict what the effect of lowering this cut will be on the other backgrounds and since no samples without these cuts are available, it can not be checked in the context of this thesis.

# Chapter 6

# Conclusion

In this thesis, a study of b-jets (jets originating from the hadronization of a b-quark) where a muon is produced in the decay of the b-hadron is presented. In a first step, these b-jets with muons are used to give a better estimation of the energy of the b-quark and the effect of this better estimation on the top quark mass is also studied. In a second step, these b-jets containing muons are used to work out an event selection which aims at selecting $t\bar{t}$ pairs which decay semi-muonically and reducing all the backgrounds of this process.

## 6.1 Using b-jets with muons to improve the estimation of the b-quark energy

In Section 3, the estimation of the b-quark energy was improved by using b-jets where muons are produced in the decay of a b-hadron inside the b-jet. First, some initial cuts were performed on the muons and on the b-jets, mainly to remove badly reconstructed muons and jets and also to remove the b-jets which lay outside the acceptance of the tracker, since the tracker is heavily used in the identification of b-jets (b-tagging). After these initial cuts, a matching between b-jets and muons is worked out, which is based on the distance in $(\eta, \phi)$-space between both objects. After this matching, the effect of adding the energy of the matched muons to the b-jet energy is studied and this addition gives, as expected, a distribution of the relative energy difference between the b-jet and the b-quark which is more narrow and closer to zero. The addition gives a relative improvement of about 22 % in the b-jet energy resolution $(\sigma(E_{b-jet}/E_{b-quark})/(E_{b-jet}/E_{b-quark}))$ for b-jets containing one or more muons. After this addition, the muons make up about 15 % of the total b-jet energy. This addition gives an improvement especially for b-jets with an energy above about 60 GeV. For the lower energy b-jets, a slight overestimation of the b-quark energy is found.

In the decay of a b-hadron to a muon, a neutrino is also produced and this neutrino escapes the CMS detector without leaving any signal, so its energy is not accounted for in the energy of the b-jet. After adding the generated neutrino to the b-jet, the energy resolution is relatively improved by 24 %. The generated neutrino's make up about 14 % of the total energy of the b-jet.

A method to estimate the energy of the neutrino is worked out, as explained in Section 3.4.4. This method is tested for several variables, but none of them gives a proper estimation of the neutrino energy. Consequently, the addition of this badly estimated neutrino energy to the energy of the b-jet (where the muon energy was already added to) does not give good results, because the relative energy distribution broadens due to this addition. The width is raised by about 16 %. Therefore, a good estimation of the energy of the generated neutrino is far from trivial, but once it is found, it can give a better estimation of the total energy of the b-jet.

The effect on the distribution of the top quark mass after the addition of the energy and momenta of the muon to the energy and momenta of the b-jet was also studied and this addition shifts the distribution about 13 GeV/c$^2$ to higher values, but it also becomes more narrow (a reduction of the width of about 6 %). The resolution of the top quark mass ($\sigma(M_{top})/M_{top}$) improves relatively by about 13 % due to the addition of the muons. The energy and momenta of the generated neutrino were also added to those of the b-jet and the muon and this extra addition shifted the top quark mass distribution to higher values, but made the distribution even more narrow. This shows that a good estimation of the energy and momenta of the neutrino can be very useful in top quark physics.

## 6.2 Selection of semi-muonic $t\bar{t}$ events

In the final section of this thesis, a selection of semi-muonic decaying $t\bar{t}$ events is worked out for an integrated luminosity of 200 pb$^{-1}$. This is the expected amount of data that will be taken during the first physics run of the Large Hadron Collider, hence by the end of 2010. The idea behind this selection is to use the fact that $t\bar{t}$ events contain a large amount of b-jets (two b-jets for each event) when compared to the backgrounds like QCD and $W$ or $Z$ production. A significant amount of these b-jets will have muons originating from the decay of a b-hadron, so it was checked if asking for one or more jets which contains at least one muon improved the already existing selection (which requires at least four centrally produced jets and an isolated muon). After an additional cut to remove the largest background (ZJets events), the signal-to-background ratio was improved from 0.958 with only the existing selection to 1.79 after the additional cuts, the significance of the signal ($\mathrm{S}/\sqrt{\mathrm{B}}$) became smaller: 51.9 before and 19.9 after the additional cuts, but after the inclusion of a systematic uncertainty, which was taken equal to 20 % of the background rate, the significance with this systematic uncertainty ($\mathrm{S}/\sqrt{\mathrm{B}+(0.2\ \mathrm{B})^2}$) equals 4.77 before and 7.97 after the additional cuts. With this improved significance and signal-to-background ratio, a potential improvement exists for a better estimation of the cross section of $t\bar{t}$ in proton-proton collisions at the LHC.

# Summary

In this thesis the use of b-jets containing muons in top quark physics with the CMS detector was discussed. The CMS detector will start taking data in the fall of 2009, when the LHC will start producing proton-proton collisions in the CERN laboratory near Geneva.

In the first analysis, a better estimation of the energy of the b-quarks is envisaged for b-jets which contain muons originating from the decay of a b-hadron inside the b-jet. The addition of the energy of the muons which lay within the b-jet, to the energy of the b-jets already gives a relative energy difference which is shifted about 15 % closer to zero and gives a relative improvement of about 22 % in the b-jet energy resolution for b-jets with a muon inside. Since in the decay of a b-hadron to a muon also a neutrino is produced, a method is worked out to estimate the energy of the neutrino. Before this estimation, the generated neutrino which is produced together with the muon was added (together with the muon) to the b-jet. This extra addition induces a shift in the relative energy distribution of about 14 % and reduces the width by about 13 %. The method which is presented to estimate the energy of the neutrino did not give any good results, so another method needs to be worked out. It was also shown that the addition of the muon and of the generated neutrino to the b-jet gives a more narrow distribution of the top quark mass. The addition of the muon improves the resolution of the top quark mass by about 13 %.

In the second analysis, a selection of semi-muonically decaying $t\bar{t}$ events was worked out. This was done by asking, after the standard muon isolation cuts, for at least one jet which has one or more muons lying inside the cone of the jet and by applying an additional cut to remove the ZJets background. This enlarged the signal-to-background ratio from 0.958 after the standard muon isolation cuts to 1.79 after the additional cuts. It also enhanced the significance of the signal when a systematic uncertainty of 20 % of the background rate was taken into account. This significance evolved from a value of 4.77 after the muon isolation cuts to a value of 7.97 after the additional cuts, allowing a potential better $t\bar{t}$ cross section estimation.

# Samenvatting

In deze thesis werd het gebruik van b-jets die muonen bevatten besproken voor het gebruik in top quark fysica met de CMS detector. De CMS detector zal data beginnen nemen in de herfst van 2009, wanneer de LHC proton-proton botsingen zal produceren in het CERN laboratorium in Genéve.

In de eerste analyse werd een betere meting van de energie van de b-quarks beoogd, met behulp van b-jets die muonen bevatten die ontstaan in het verval van een b-hadron in de b-jet. Het optellen van de energie van de muonen die in de b-jet liggen bij de energie van de b-jet geeft reeds een relatief energieverschil dat ongeveer 15 % dichter bij nul ligt en een relatieve verbetering van ongeveer 22 % in de b-jet energieresolutie. Aangezien in het verval van een b-hadron naar een muon ook een neutrino ontstaat, werd er ook een manier om de energie van dit neutrino te schatten uitgewerkt. Voor deze schatting werd het gegenereerd neutrino, dat samen met het muon ontstaat, samen met dit muon bij de b-jet geteld. Deze extra optelling veroorzaakt een verschuiving in de verdeling van het relatieve energieverschil van ongeveer 14 % en verkleint de breedte met ongeveer 13 %. De methode die werd toegepast om de energie van het neutrino te schatten gaf geen goede resultaten en dus moet er een andere methode uitgewerkt worden. Er werd ook aangetoond dat de optelling van het muon en van het gegenereerd neutrino bij de b-jet een smallere verdeling van de top quark massa geeft. De optelling van het muon verbetert de resolutie van de top quark massa met ongeveer 13 %.

In de tweede analyse werd een selectie van semi-muonisch vervallende $t\bar{t}$ gebeurtenissen uitgewerkt. Er werd, na de standaard muon isolatie snede's, gevraagd naar tenminste één jet die één of meer muonen bevat. Ook werd er een snede toegepast om de ZJets achtergrond te reduceren. Deze extra snede's vergroten de signaal-over-achtergrond verhouding van 0.958 na de standaard muon isolatie snede's naar 1.79 na de extra snede's. Ze verhoogden ook de significantie van het signaal wanneer een systematische onzekerheid gelijk aan 20 % van de achtergrond in rekening was genomen. De significantie evolueerde van 4.77 na de muon isolatie snede's naar 7.97 na de extra snede's, waardoor de $t\bar{t}$ werkzame doorsnede mogelijk beter kan gemeten worden.

# Bibliography

[1] F. Mandl and G. Shaw, *Quantum Field Theory*, (1984) 358 p.

[2] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, (1995) 815 p.

[3] *Particle Data Group*, `http://pdg.lbl.gov/`.

[4] *The LEP Electroweak Working Group*, `http://lepewwg.web.cern.ch/LEPEWWG/`.

[5] F. Englert and R. Brout, Phys. Rev. Lett. **13** (1964) 321 – 322.

[6] P. W. Higgs, Phys. Rev. Lett. **13** (1964) 508 – 509.

[7] G. Bertone, D. Hooper and J. Silk, Phys. Rep. **405** (2005) 279390.

[8] S. Weinberg, *The Quantum Theory of Fields, Volume 3: Supersymmetry*, (2000) 419p.

[9] H. Davoudiasl and T. G. Rizzo, *Signatures of Spherical Compactification at the LHC*, `arXiv:hep-ph/0702078v2`.

[10] T. G. Rizzo, *Z' Phenomenology and the LHC*, `arXiv:hep-ph/0610104v1`.

[11] *The Large Hadron Collider Project*, `http://lhc.web.cern.ch/lhc/`.

[12] Lyndon Evans and Philip Bryant (editors), 2008 JINST **3** S08001.

[13] *The Fermilab Tevatron collider*, `http://www-bdnew.fnal.gov/tevatron/`.

[14] R. Granier De Cassagnac, *What's the matter at RHIC?*, `arXiv:0707.0328v2 [nucl-ex]`.

[15] *The CMS Experiment*,
`http://cms.cern.ch/`.

[16] *The CMS Collaboration*, S. Chatrchyan *et al.*, 2008 JINST **3** S08004

[17] *The ATLAS Experiment*,
`http://atlas.web.cern.ch/`.

[18] *The ALICE Experiment*,
`http://aliceinfo.cern.ch/`.

[19] *The LHC-b Experiment*,
`http://lhcb.web.cern.ch/`.

[20] *The TOTEM Experiment*,
`http://totem.web.cern.ch/`.

[21] *The LHC-f Experiment*,
`http://www.stelab.nagoya-u.ac.jp/LHCf/`.

[22] *The CDF Experiment*,
`http://www-cdf.fnal.gov/`.

[23] *The D∅ Experiment*,
`http://www-d0.fnal.gov/`.

[24] N. Kidonakis and R. Vogt, *The theoretical top quark cross section at the Tevatron and the LHC*, `arXiv:0805.3844v1 [hep-ph]`.

[25] A. H. Hoang, *Top Threshold Physics*,
`arXiv:hep-ph/0604185v1`.

[26] G. Abbiendi *et al.*, *Muon Reconstruction in the CMS Detector*, CMS Analysis Note AN-2008/097.

[27] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. **A262** (1987) 444 – 450.

[28] The CMS Particle Flow Physics Object Group, *Particle Flow Reconstruction of Jets, Taus, and MET*, CMS Analysis Note AN-2009/039.

[29] A. Heister *et al.*, *Measurement of Jets with the CMS Detector at the LHC*, CMS-NOTE-2006/036.

[30] The CMS Collaboration, *Plans for Jet Energy Corrections at CMS*, CMS Physics Analysis Summary CMS PAS JME-07-002.

[31] P. Demin, S. De Visscher, A. Bocci, and R. Ranieri, *Tagging b jets with electrons and muons at CMS*, CERN-CMS-NOTE-2006-043.

[32] T. Sjstrand, S. Mrenna and P. Skands, JHEP05 (2006) 026.

[33] *The* `MadGraph` *Matrix Element Generator*, `http://madgraph.phys.ucl.ac.be/`.

[34] GEANT4 Collaboration, S.Agostinelli et al. *GEANT4: A simulation toolkit*, Nucl. Instrum. and Methods **A506** (2003) 250 – 303.

[35] G. P. Salam and G. Soyez, *A practical Seedless Infrared Safe Cone jet algorithm*, `arXiv:0704.0292v2 [hep-ph]`.

[36] *Recommended Muon Selection*, `https://twiki.cern.ch/twiki/bin/view/CMS/VplusJets`.

[37] The CMS Collaboration, *Observability of Top Quark Pair Production in the Semileptonic Muon Channel with the first 10* $pb^{-1}$ *of CMS Data*, CMS Physics Analysis Summary CMS PAS TOP-08-005.