COMENIUS UNIVERSITY BRATISLAVA

FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS

Evidence no.: 945b5424-92ff-4521-a044-b121e60ab19e

STUDY OF THE TOP QUARK ELECTRIC CHARGE AT THE CDF EXPERIMENT

2011

Mgr. Pavol Bartoš

Acknowledgements

Abstrakt

V tejto práci prezentujeme meranie elektrického náboja top kvarku použijúc metódu tagovania náboja jetov na eventoch, ktoré obsahujú jeden leptón. Tieto eventy boli nazbierané CDF II detektorom vo Fermilabe v čase od februára 2002 do februára 2010 pri energii hmotného stredu $\sqrt{s} = 1.96$ TeV. Metódu, ktorú požívame, možno rozdeliť do troch zložiek: určenie náboja W bozónu (použijúc náboj leptónu), párovanie W bozónu a b-jetu, ktoré pochádzajú z jednej top kvarkovej rozpadovej vetvy a napokon určenie náboja b-jetu Jet Charge algoritmom. Na dátovej vzorke 5.6 fb^{-1} sme určili p-value pre Štandardný model (13,4%) a p-value pre Exotický model (0,014%). Na základe kritérií, všeobecne akceptovanými CDF kolaboráciou a definovanými pred získaním výsledku z nameraných dát, môžme povedať, že získaný výsledok je konzistentný so Štandardným modelom, zakiaľ čo hypotézu exotického kvarku vylučujeme s 95% konfidenciou. Prezentovaná metóda má najvačšiu citlivosť na elektrický náboj top kvarku v porovnaní s doposiaľ prezentovanými výsledkami analýz týkajúcich sa náboja top kvarku.

Kľúčové slová: náboj top kvarku, experiment CDF, váhovanie náboja trekov, b-jet.

Abstract

We report on the measurement of the top quark electric charge using the jet charge tagging method on events containing a single lepton collected by the CDF II detector at Fermilab between February 2002 and February 2010 at the center-of-mass energy $\sqrt{s} = 1.96\,TeV$. There are three main components to this measurement: determining the charge of the W (using the charge of the lepton), pairing the W with the b-jet to ensure that they are from the same top decay branch and finally determining the charge of the b-jet using the Jet Charge algorithm. We found, on a sample of $5.6\,\mathrm{fb}^{-1}$ of data, that the p-value under the standard model hypothesis is equal to 13.4%, while the p-value under the exotic model hypothesis is equal to 0.014%. Using the *a priori* criteria generally accepted by the CDF collaboration, we can say that the result is consistent with the standard model, while we exclude an exotic quark hypothesis with 95% confidence. Using the Bayesian approach, we obtain for the Bayes factor $(2ln(BF))$ a value of 19.6, that favors very strongly the SM hypothesis over the XM one. The presented method has the highest sensitivity to the top quark electric charge among the presented so far top quark charge analysis.

Key words: top quark charge, CDF experiment, track charge weighing technique, b-jet

Contents

List of Figures

List of Tables

Introduction

In early 1960s, a lot of different particles were known from from cosmic rays and accelerator's experiments. It looked like there are too many fundamental particles, but in 1964 Gell-Mann and Zweig came with the theory that all the known particles are built of the smaller parts - quarks. This theory later became a part of the Standard Model (SM), which also describes all fundamental interactions, except gravity - the processes evoked by electromagnetic, weak and strong forces.

In the SM we have 12 fundamental fermions (spin 1/2) - 6 quarks with a fractional electric charge (-1/3 or +2/3) and 6 leptons with integer charge (0 or 1). We divide them into three generations as it is shown in Figure 1. The first generations is the lightest one and all stable matter is made of the quarks and the leptons from this generation. The heavier quarks from the second and third generations, which can be observed in cosmic rays or are produced in high energy collisions, play a important role in the early stages of the Universe. Each of these fundamental particles has also its "own" antiparticle.

Figure 1: The fundamental particles and intermediate bosons in Standard Model.

To describe the interactions in SM we have two theories based on calibrations symmetries:

- electroweak (EW) theory based on $SU(2)_L \times U(1)_Y$ calibration symmetry

- and quantum chromodynamics (QCD) based on $SU(3)_C$ calibration symmetry.

Each free symmetry group parameter corresponds to the one intermediate boson (spin 1). Finally we have a photon γ, $W^\pm$ and Z bosons in the EW theory and eight gluons in the QCD.

An important ingredient of SM is the idea of the spontaneous symmetry breaking, which is explained by a Higgs field. This field has non-zero vacuum expectation value and interaction with it leads to non-zero masses of the quarks, leptons, W and Z bosons. The Higgs boson (with spin 0) is the only undiscovered particle of the SM.

However the SM does not explain everything - the hierarchy of the particle masses, number of particle generations. It does not include the gravity. These and other reasons leaded to creation of the other theories (beyond SM) like models with 4th generation of quarks, GUT, Supersymmetry, String Theory.

The top quark as the heaviest quark with the mass close to the scale of EW symmetry breaking is an important object of the research. For example, the precise measurement of its mass leads to constrain the limits on the Higgs boson mass, the top quark events are an important part of the background in the Higgs boson search, the measuring of the single top quark production cross-section can lead to confirm the existence of only 3 fermion generations.

In this work we want to distinguish between two hypothesis - in the null (SM) hypothesis the electric charge of the top quark measured in the Fermilab is equal to +2/3, while in other (exotic) hypothesis the charge is -4/3, what would leads to confirmation of a new quarks generation existence. Due to the fact that we reconstruct the top quark charge from the electric charges of its decay products, it is possible to combine W boson with b-quark in the ways which lead to these two hypothesis.

Goals of the thesis

The main goal of the thesis is to analyze the CDF experimental data and on the sample of 5.6 fb^{-1} to setup the top quark charge in the lepton + jet channel. The ultimate goal of the analysis is to distinguish between the Standard Model and exotic model scenarios. In the previous published result done on 1.5 fb $^{-1}$ it was seen a contradiction between the muon data and the electron ones. The electron data gave a clearly favor for the SM scenario while in the case of the muons it was not so clear, though the combined electron and muon 1.5 fb^1 lepton+jets data excluded XM with the confidence level of 87% [68].

In frame of the main goal I predominantly concentrate on the following partial goals:

1. Validation of the jet energy scale - check the jet correction for each new collected data sample and if needed, contribute to updating the relevant scripts. Mainly it was needed to check the absolute and out-of-cone corrections (for detail see section 2). The validation is based on the γ-jets events, because the photon p_T^γ is very precisely measured in electromagnetic calorimeter and due to the momentum conservation law the jet p_T^{jet} should be the same.

2. To increase the muon acceptance by adding a non-triggered muons, which are reconstructed off-line.

3. To improve the jet charge algorithm and it's purity.

4. To calibrate jet charge algorithm using the new scale factors, because the latest calibration has been done only for the 1.9 fb^{-1} sample and also the tracking scripts were changed, so we need to compare the scale factors for both - old and new tracking scripts. The calibration is done by using the dijet samples, where one jet is required to contain a soft muon and the second jet has to be back-to-back to the first one.

5. To update the background using also a high luminosity MC samples (generated with higher instantaneous luminosity).

6. To update the systematic error studies.

Chapter 1

Tevatron accelerator and CDF experiment

1.1 Tevatron

Tevatron is the circular particle accelerator situated at the Fermi National Accelerator Laboratory near Chicago, USA. Before the LHC started the protons-antiprotons ($p\bar{p}$) collisions, Tevatron was the most powerful $p\bar{p}$ accelerator in the world. It collides protons and antiprotons with center-of-mass energy $\sqrt{s} = 1.96$ TeV in the cross-points placed in the centers of the CDF (**C**ollider **D**etector experiment at **F**ermilab) and the D0 (the technical name of the detector position in the Tevatron ring) detectors, which are positioned around the beam pipe at two different locations. [1]

To reach the energy of 980 GeV for both protons or antiprotons, they are accelerated in several steps by the chain of accelerators (Figure 1.1).

1.1.1 Proton source

The accelerator chain begins with the *Cockcroft-Walton pre-accelerator*. It serves as the source of the negatively charged hydrogen ions, which are accelerated here by electric field to the energy of 750 keV. Hydrogen ions continue to the linear accelerator *Linac* constructed from the radio-frequency cavities, where they reach the energy of 400 MeV. Then they are stripped off the electrons in the first circular accelerator in the chain, *Booster*, so only protons remain and are accelerated to the 8 GeV. The next increase of the energy is provided by the second circular accelerator, *Main injector*,

which accelerates protons to the energy of 120 GeV or 150 GeV. In the former case protons continue to the antiproton source while in the latter one they are injected to the *Tevatron.* [2]

1.1.2 Antiproton source

Protons accelerated in *Main Injector* to the energy of 120 GeV are directed to the nickel target, where collisions produce the spray of all sorts of secondary particles. The magnets are used to select 8 GeV antiprotons, which are then stored in the *Recycler.* The *Recycler* is placed in the same tunnel as the *Main Injector* and does not accelerate antiprotons, but rather store them at a constant kinetic energy of 8 GeV. When sufficient amount of antiprotons has been produced, they are accelerated in the *Main Injector* to the energy of 150 GeV and injected to the *Tevatron.* [2]
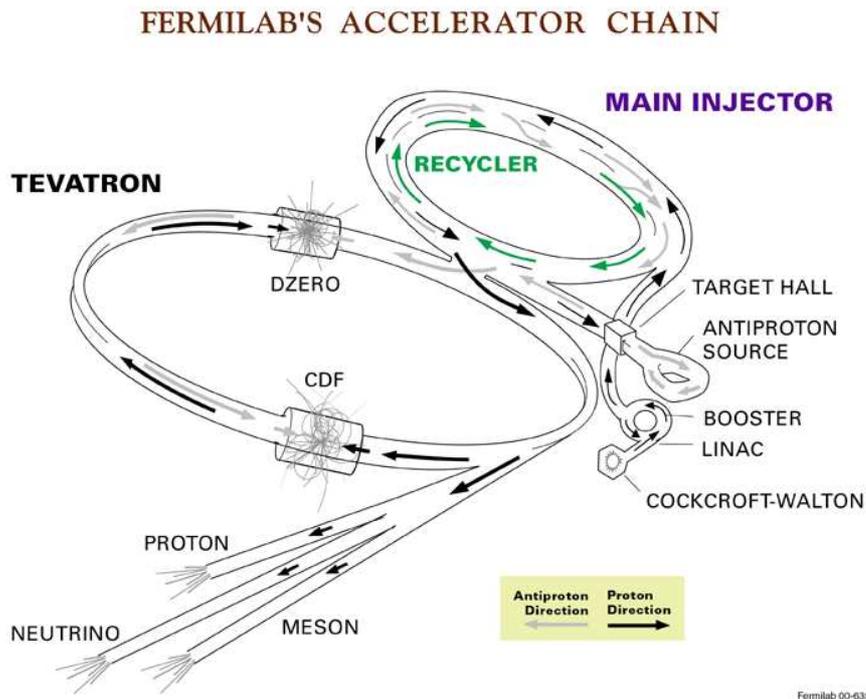
Figure 1.1: Fermilab's accelerator chain

1.1.3 The Tevatron

The Tevatron is the largest of the Fermilab accelerators, with a circumference of approximately 6 kilometers. Energy of protons and antiprotons is increased from 150

GeV to 980 GeV by eight accelerating cavities placed along the tunnel. The primary purpose of the *Tevatron* is to act as a storage ring where protons and antiprotons can collide with each other and produce interesting secondary particles. Once the final energy is reached, the two counter-rotating particle beams pass through each other for hours at a time. This stable situation of 980 GeV proton and antiproton collisions is called a "store". We define the instantaneous luminosity of the store by the following equation:

$$\mathcal{L} = fF\frac{N_B N_p N_{\bar{p}}}{2\pi(\sigma_p^2 + \sigma_{\bar{p}}^2)} \tag{1.1}$$

where f is the revolution frequency, F is a form factor dependent upon the bunch length, N_B the number of bunches, N_p ($N_{\bar{p}}$) the number of protons (antiprotons) per bunch and σ_p ($\sigma_{\bar{p}}$) is the standard deviation of the beam spatially at the interaction point.

The instantaneous luminosity is not constant during the store, but is decreasing with the time. After it drops too low to be useful for the experimenters, the store is ended and the Tevatron is being prepared for a new store. [2]

The Table 1.1 summarizes the Tevatron parameters for so called Run II data taking period (2002 - 2011), while in the Figure 1.2 is shown the integrated luminosity delivered by the Tevatron and also luminosity written to the tapes by CDF.

parameter	Run II value
Bunches	36
Bunch spacing [ns]	396
Bunch length [m]	0.38
Protons/bunch (N_p)	2.7 x 10^{11}
Anti-protons/bunch ($N_{\bar{p}}$)	7 x 10^{10}
Typical instantaneous luminosity [$cm^{-2}sec^{-1}$]	2 x 10^{32}

Table 1.1: The Tevatron parameters for Run II data taking period.

1.2 The CDF experiment

The CDF is an international collaboration of about 500 physicist from ~60 institutes from all around the world. One of big successes of this experiment during the Run I
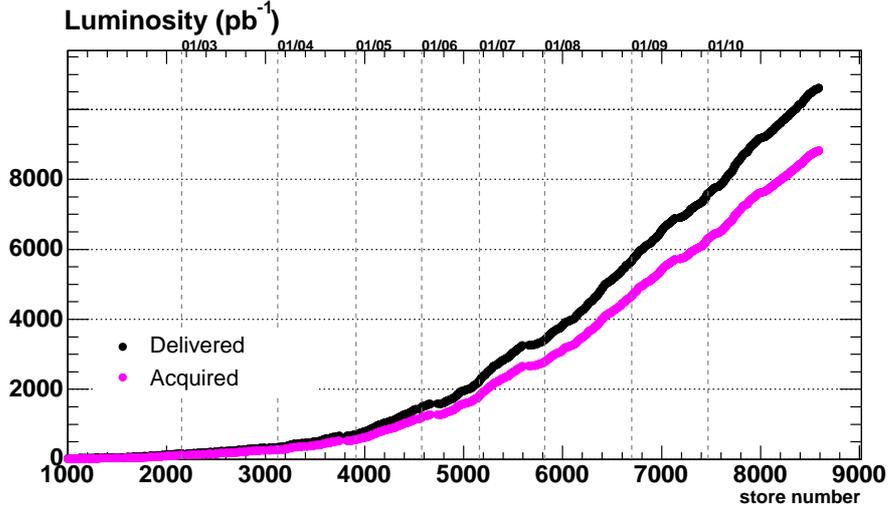
Figure 1.2: Integrated Luminosity for Run II data taking. The black curve is luminosity delivered by Tevatron, the pink curve is the luminosity written to the tapes by CDF.

data taking (1992 - 1996) was discovering of the top quark in 1995, together with D0 experiment [3, 4]. The CDF experiment's measured top quark mass and $t\bar{t}$ production cross-section were $m_t = 176 \pm 8(stat) \pm 10(syst)$ GeV/c^2, $\sigma_{p\bar{p}}^{t\bar{t}} = 6.8^{+3.4}_{-2.4}$ pb, respectively [3].

The end of the Run I data taking was followed by upgrades of the detectors. The real physics data taking (Run II) started six years later, in 2002, but after the 9 years of excellent work and progress in particle physics research Tevatron is going to finish data taking in September 2011. Many goals of this Run II data taking were already reached, so the new one were set. In the following we summarize the original main goals [5] and also some results in the corresponding field of research:

- Study of the properties of the top quark (see Section 3), precision measurements, e.g. the top quark mass. While the original expected precision $\delta m_t \leq 3$ GeV/c^2 was reached with only 1 fb^{-1} of data, the ten times higher statistics by the end of the Run II leads to expect top mass measurement uncertainty be lower then 1 GeV/c^2 [6]. The current CDF value of top mass uncertainty is 1.1 GeV/c^2 [7].

- A global precision electroweak program, e. g. W boson mass $\delta m_W \leq 40$ MeV/c^2. The W boson mass was measured by CDF on 200 pb^{-1} with precision of 48 MeV/c^2 [8], then the combination with D0 measurement and with result from LEP yields the world average of $m_W = 80399 \pm 23$ MeV [9].

7

- Direct search for new phenomena like high mass resonances (Z', W', sneutrino, graviton), super-symmetric particles (stop, chargino), new or excited fermions (t', b'), etc. Up today no new particles were observed, results were consistent with Standard Model expectations or leads to the setting limits to the particles masses (e. g. graviton, stop, W', Z', b', t'), which can be find in [10, 11, 12, 13, 14, 15].

- Tests of perturbative QCD at Next-to-Leading Order and the large Q^2. The results shows very good agreement with NLO perturbative QCD, e.g. the sub-structure of high transverse momentum jets, measurement of $Z/\gamma^* \rightarrow \mu^+\mu^- + jets$ production cross-section [18, 19].

- Constrain of the CKM matrix with high statistics B decays, e.g. observation of $B_S - \bar{B}_S$ oscilation in 2006 [20].

- Search for the Higgs boson. The Tevatron latest results excluded Higgs boson mass in the region of 158 - 173 GeV/c^2 with 95% confidence [21], while in case of a fourth sequential generation of fermions with large masses, the CDF experiment exclude at the 95% confidence a SM-like Higgs boson in the mass range between 124 - 202 GeV/c^2 [22].

Many of the Tevatron results maybe will be overcome by the coming LHC results in the coming years, but there is still a lot of work needed to be done to get final results from data collected by the Tevatron which are of common interest.

1.3 The CDF detector

The CDF II detector, shown in Figure 1.3, is a general purpose solenoidal detector which combines precision charged particle tracking with fast projective calorimetry and fine grained muon detection. Tracking systems are contained in a superconducting solenoid, 1.5 m in radius and 4.8m in length, which generates a 1.4 T magnetic field parallel to the beam axis. Calorimetry and muon systems are all outside the solenoid. [5]

The coordinate system used at the CDF is right-handed cylindrical system. The positive direction of the z axis is parallel with the incoming protons, while the horizontal x axis has positive direction pointing outward from the accelerator ring. Then we
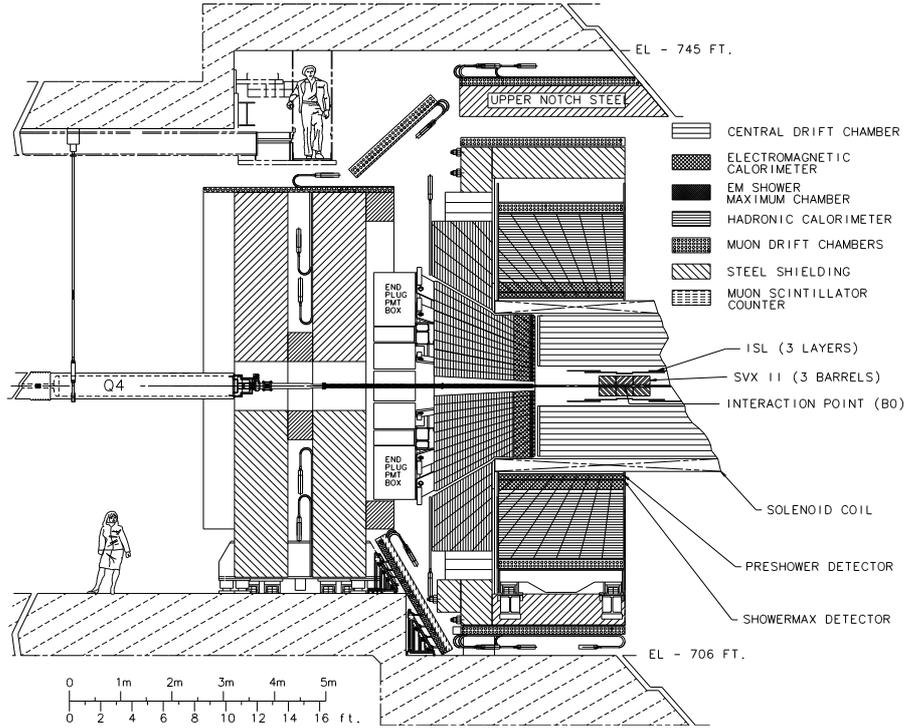
Figure 1.3: Side view of the CDF II detector.

define cylindrical coordinates r, θ, ϕ as follows. The distance in transverse plain $r = \sqrt{x^2 + y^2}$, the polar angle θ is measured from proton direction and the azimuthal angle ϕ is measured from the x axis.

Using these coordinates we can also define pseudo-rapidity η, which is very useful because pseudo-rapidity difference $\Delta\eta$ is Lorentz invariant:

$$\eta = -ln(tan(\theta/2)) \tag{1.2}$$

1.3.1 Tracking system

As we already mentioned the tracking system is placed inside the magnetic field of 1.4 T pointing parallel with beam direction, what cause that particles travel on the circle in the transverse plane. From the curvature we can measure for example the charge and the transverse momenta of the particles.

Figure 1.4 shows the tracking system which consists of two detectors: *Silicon Vetex Detector and Central Outer Tracker (COT).*

9

CDF Tracking Volume
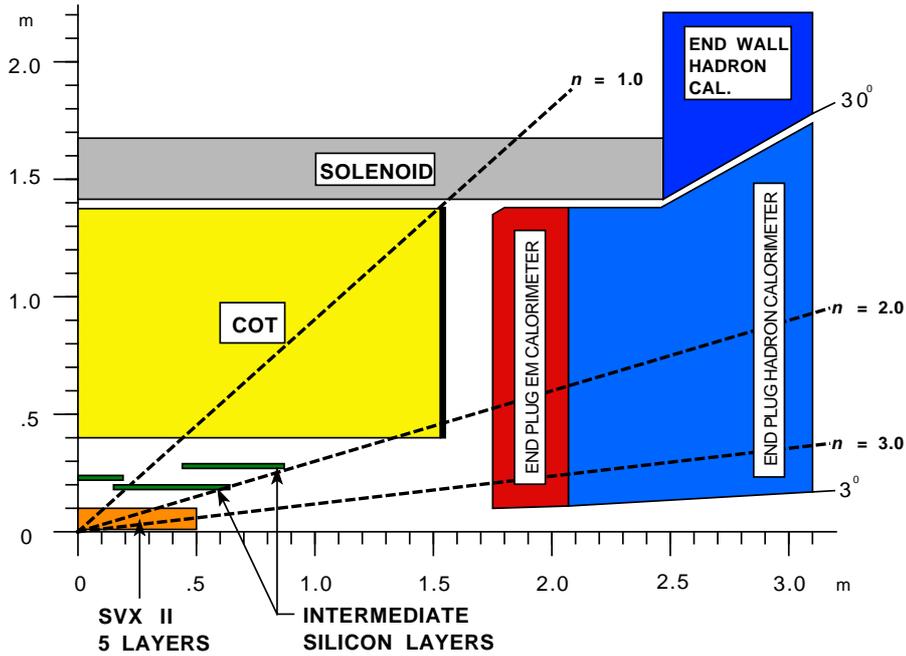
Figure 1.4: Longitudinal view of the CDF II tracking system.

Silicon Vertex detector

The Silicon vertex detector consists of 8 layers of silicon sensors (p-n junction diodes) divided into three sub-detectors (Figure 1.5).

Layer 00 (L00) is mounted closest to the beam (15 mm) and comprises of one-sided silicon micro-strips. [16, 5]

The next five layers compose *SVXII* system. These layers are constructed from double-sided strips. Three of them combine an $r - \phi$ measurements on one side with 90° stereo measurement on the other, while the remaining two layers combine an $r - \phi$ with small angle stereo ±1.2°. [16, 5]

The last sub-detector *Intermediate Silicon Layers* (ISL) consists of a single layer placed at a radius 22 cm in the central region and two layers at radii of 20 cm and 28 cm in the plug region ($1.0 \leq |\eta| \leq 2.0$). Double sided strips, with 0° and small angle ±1.2° with respect to the beam axis, are used. [16, 5]

The excellent identification of secondary vertices is done by the $40\mu m$ resolution of the impact parameter.

10
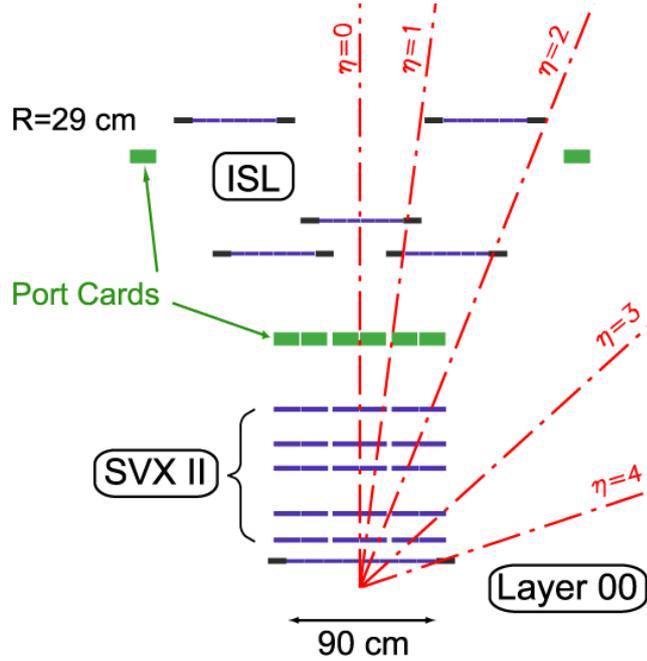
Figure 1.5: The silicon tracking detector projected on the $r-z$ plane.

Central Outer Tracker

The Central Outer Tracker (COT) is an open cell drift chamber covering radii between 44 and 132 cm and the region of $|\eta| \leq 1.0$. It is composed of 96 sense wire layers in radius that are grouped into eight "superlayers" (see Figure 1.6). Four superlayers are axial, wires are parallel with beam axis and four layers are stereo layers tilled by $\pm 2°$ angle with respect to the beam axis. Each superlayer is divided in ϕ into "supercells", and each supercell has 12 sense wires. A maximum drift distance is approximately the same for the all superlayers. The supercells are tilled by $35°$ with respect to the radial direction to compensate for the Lorentz angle of the drifting electrons in the magnetic field. The standard gas mixture argon / ethane (50:50) is used [17].

1.3.2 Calorimetry

Scintillator-based calorimetry covers the region $|\eta| \leq 3.6$. As seen in Figure 1.7, the calorimetry consists of an electromagnetic section, which measure the energy of electrons and photons, and is followed by the hadronic section used for the jets' (hadrons') energy measurement. By the position with respect to the interaction point we separate calorimeters into two groups - central and plug [5].

11
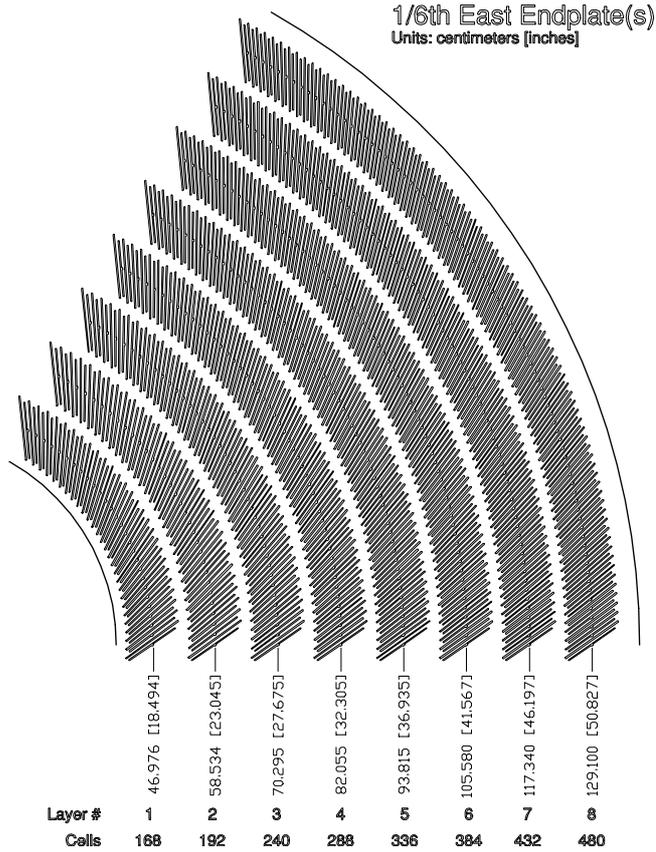
Figure 1.6: The COT end plate.

The CDF calorimetric system consists of 5 calorimeters, which are enumerated together with their parameters - the energy resolutions and covered η region in Table 1.2 [5, 23]. All calorimeters are composed of the smaller segments - towers by η and ϕ. The segmentation is described in Table 1.3.

calorimeter	acronym	covered η region	energy resolution (%)
central electro-magnetic calorimeter	CEM	$|\eta| < 1.1$	$13.5 / \sqrt{E_T} \oplus 2$
plug electro-magnetic calorimeter	PEM	$1.1 < |\eta| < 3.6$	$16 / \sqrt{E_T} \oplus 1$
central hadron calorimeter	CHA	$|\eta| < 0.9$	$50 / \sqrt{E_T} \oplus 3$
end-wall hadron calorimeter	WHA	$0.7 < |\eta| < 1.2$	$75 / \sqrt{E_T} \oplus 4$
plug hadron calorimeter	PHA	$1.2 < |\eta| < 3.6$	$80 / \sqrt{E_T} \oplus 5$

Table 1.2: The CDF II calorimeters, with their acronyms, η region, which they cover and energy resolutions.

The electro-magnetic calorimeters (CEM, PEM) layers are made of the lead and

12
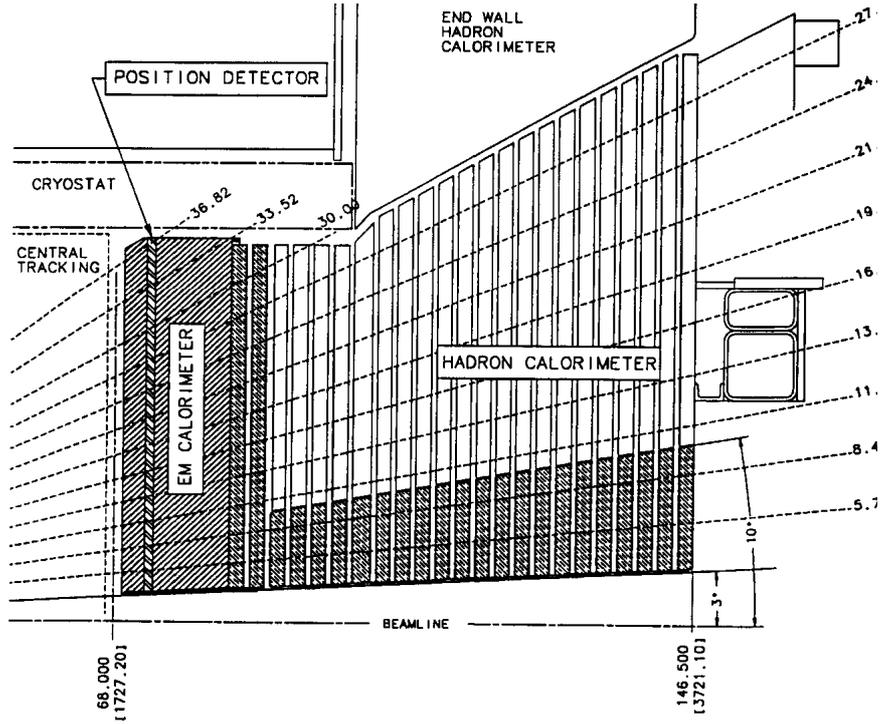
Figure 1.7: Cross-section of upper part of plug calorimeter.

$|\eta|$ range	$\Delta\phi$	$\Delta\eta$
0. - 1.1 (1.2 had.)	15°	~ 0.1
1.1 (1.2 had.) - 1.8 (1.2 had.)	7.5°	~ 0.1
1.8 - 2.1 (1.2 had.)	7.5°	~ 0.16
2.1 - 3.64 (1.2 had.)	15°	0.2 - 0.6

Table 1.3: The CDF II calorimeter segmentation.

polystyrene scintillator. The total thickness of CEM is about $18X_0$ (radiation lengths), while PEM of about $21X_0$. In approximately $6X_0$ depth shower-max position detectors (CES, PES) are placed in both calorimeters. The position measurements and the transverse shower profile is used to separate electrons and photons [5].

The central hadron CHA and WHA calorimeters consists of steel and PMMA polystyrene scintillator, while the plug PHA calorimeter is made of the same material like electro-magnetic calorimeters (lead and polystyrene) [5].

13

1.3.3 Muons detectors

Muon system is placed at the most outer part of the CDF detector and consists of four drift chamber detectors and scintillators counters.

The Central Muon Detector (CMU), situated right behind the Central Hadron Calorimeter, consists of 144 moduled wits 16 rectangular cells per module. It covers the region of $|\eta| \leq 0.6$. There is a requirement on minimum transverse momentum p_T of the muon to be detected, which is 1.4 GeV/c [5].

The Central Muon Upgrade (CMP) consists of a second set of muon chambers behind an additional 60 cm of steel. The chambers are of fixed length in z and form a box around the central detector. The pseudo-rapidity coverage is up to $|\eta| < 0.6$ but it varies with the azimuth, as is shown in Figure 1.8. A layer of scintillator counters (CSP) with 2.5 cm thickness is installed on outside surface of the detector. To be able to detect muon by the CMP, the minimum p_T of the muon has to be > 2.2 GeV/c [5].

The Central Muon Extension (CMX) consists of canonical sections of drift tubes and scintillator counters (CSX) located at each end of the central detector and extending in polar angle from 42° to 55°. At 55° the CMX/CSX system slightly overlaps the coverage provided by the central muon system and extends its pseudo-rapidity coverage from 0.6 to 1.0 [5].

The Intermediate Muon Detector (IMU) is design to trigger on muons with $|\eta| \leq 1.5$ and to identify off-line muons with $|\eta| \leq 2.0$. As in previous case the detector consist of a drift chambers barrel (BMU) and scintillator counters (WSU, BSU, TSU). The non-used toroidal magnet from previous Run I data taking together with additional steel plates are used to shield the IMU detector from hadrons [5].

1.3.4 Triggers

The CDF data acquisition system can store data at a maximum rate of 18 MB/s. With an average event size of 170 kB, this translate into an event rate of 100 Hz. Therefore, in processing the 1.7 MHz of collision data the CDF trigger system must reject more than 99.99% of the events. The CDF trigger system has a three-level architecture with each level providing a rate reduction sufficient to allow for processing at the next level with minimal dead-time. The block diagram of the trigger system is presented in Figure 1.9 [24].
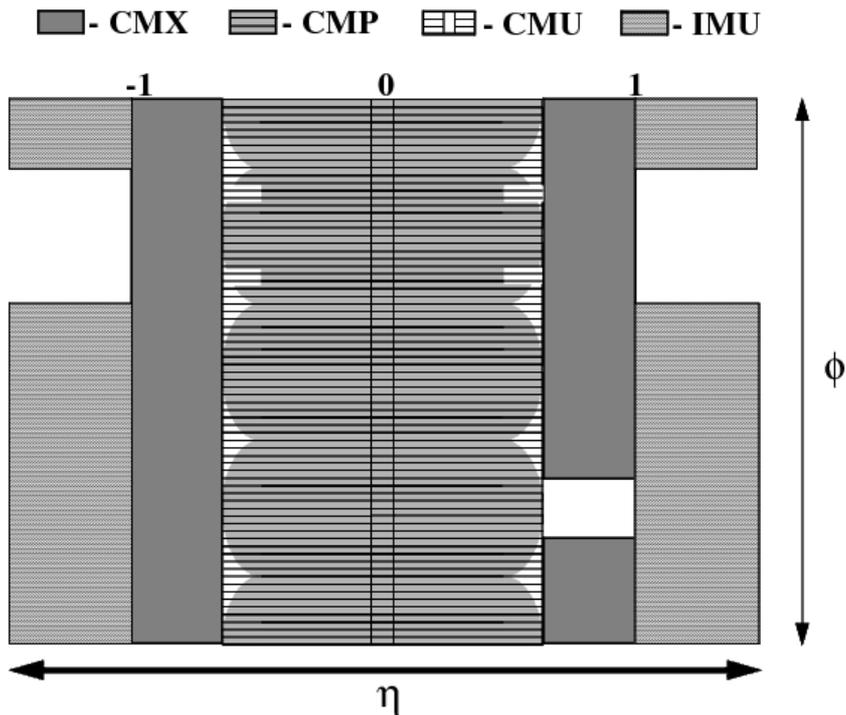
Figure 1.8: CDF Muon detectors.

The *Level-1* operates on every beam crossing and uses a custom designed hardware to find physics objects based on a subset of the detector information. The input comes from the calorimeters, tracking chamber and muon detectors. The decision to retain an event for further processing is based on the number and energies of track, electron, photon, muon, τ lepton and jet candidates, as well as the total energy and missing transverse energy in the event. The lepton identification at this level is done by extrapolating tracks to the calorimeter and muon systems. The accept rate of Level-1 is 25 kHz with $\leq 5\%$ deadtime [24].

The *Level-2* consists of four on-board buffers. The global Level-1 information is enriched by the information from the shower-max position detector (CES) and silicon vertex detector (SVX), which provides the ability to trigger on displaced tracks arising from the decay of long-lived particles. The Level-2 decision node is the first place where software algorithms are utilized to process the event. Its accept rate is 350 Hz [24].

The *Level-3* system is a computing farm consists of about 400 processor nodes. Here the events are reconstructed and filtered using full event reconstruction. After

that they are written to the permanent data storage (tapes) with rate 75 Hz [24].

<u>RUN II TRIGGER SYSTEM</u>

Detector Elements

```
CAL    COT    MUON    SVX    CES

              XFT    MUON           XCES
                     PRIM.

              XTRP

       L1     L1     L1
       CAL    TRACK  MUON

          GLOBAL
          LEVEL 1

       L2                    SVT
       CAL

          GLOBAL             TSI/CLK
          LEVEL 2
```
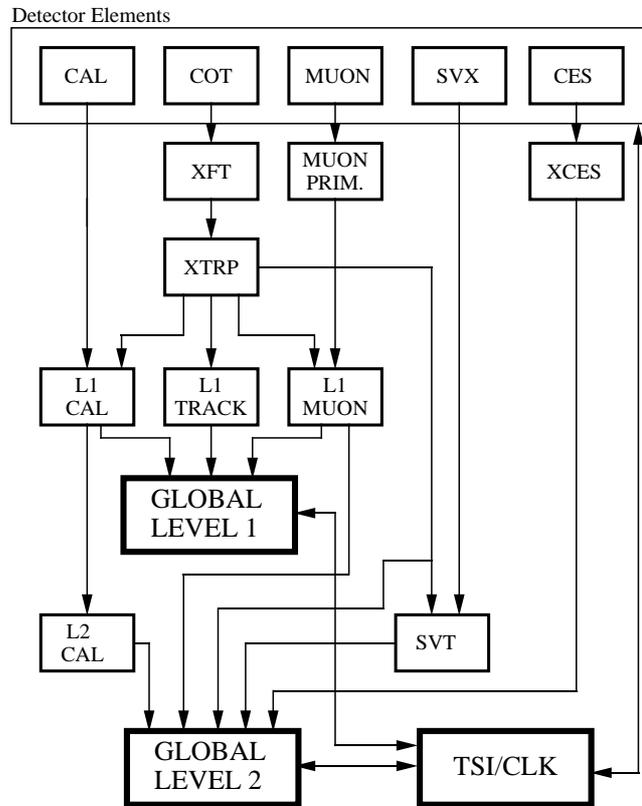
PJW 9/23/96

Figure 1.9: The trigger system block diagram.

16

Chapter 2

Particle identification

Colliding protons with antiprotons at center-of-mass energy 1.96 GeV leads to creation of many sorts of particles. The detectors are constructed in such a way to be able to detect and perform particle identification of photons, leptons, hadrons and muons. In the case of photon and electron almost all energy is released in one of the electromagnetic calorimeters, while hadrons, which are usually detected in a form of a jets, pass the electromagnetic calorimeters loosing only a relatively small part of their energies and the rest of their energies are lost in the hadronic calorimeters. In the case of muons, they can deposit some part of energy in both type of calorimeters (electromagnetic and hadronic) but the decision about the muon's presence is taken from the muon chambers information. The tracking system is also involved. Each charged particle has its track in this system so we can combine the calorimeters' or muon chambers' information with that from the COT or the silicon detector. There can be also a neutrinos produced by the interactions, but due to the fact that they interact only weak, we don't see their signal. However we can predict the amount of energy which was carried by neutrinos.

In this section we describe the identification criteria for different objects found in the detector.

2.1 Electron identification

The electrons can be identified by a COT track which corresponds to an electromagnetic (EM) calorimeter cluster. This cluster is formed from a seed tower (a tower with the highest energy deposition) and adjacent neighbor towers. In the case of the CEM only

2 adjacent towers are taking into account (one on each side in the η direction), while in the PEM case the maximum size of a cluster is 3 x 3 towers.

There are further selection criteria, which help to identify the electrons, but let us first define a variables needed for using these criteria:

- $E_T = E.sin\theta$.

  The transverse electromagnetic energy, E_T, is obtained from the energy of the EM cluster and the polar angle of the track pointing to the seed tower of the cluster.

- $p_T = P.sin\theta$

  The transverse momentum of the track, where P is the track momentum.

- E_{had}/E_{em}

  The ratio of the hadronic and electromagnetic energy deposition. This variable helps to separate electrons from jets, which leave more energy in the HAD calorimeter.

- E/P

  The ratio of the transverse energy of the EM cluster and transverse momentum of the COT track. The value of this fraction can be large if the electron radiate a soft photon, which usually deposits its energy in the same tower as the electron, but the the radiation lower the value of the electron's momentum. For the high energy electrons the value of the fraction is close to 1.

- $\Delta x_{CES} * Q$, Δz_{CES}

  Distances between the extrapolated COT track and the best matching CES cluster in the $r - \phi$ or the $r - z$ planes, respectively. The distance Δx_{CES} in the $r - \phi$ plane is multiplied by the track charge Q.

- χ^2_{strip}

  The χ^2 comparison of the CES shower profile with the shower profile obtained from the test beam measurements.

- L_{shr}

  The comparison of the measured $E_i^{measured}$ and the expected $E_i^{expected}$ energy of

the towers in the electromagnetic shower:

$$L_{shr} = 0.14 \sum \frac{E_i^{measured} - E_i^{expected}}{\sqrt{(0.14\sqrt{E})^2 + \sigma^2_{E_i^{expected}}}} \qquad (2.1)$$

where i runs over all towers, and the error of the energy measurement, represented by the denominator, is the uncertainty in the energy estimate. The reason to do this is evoked by the fact that the expected energy distributions in the towers around the seed tower are different for the EM and HAD showers.

- z_0

  The z-coordinate position of the electron associated track's intersection with the beam axis.

- COT track quality cut

  The electron associated track is required to pass trough a certain number of the COT super-layers (SLs).

- Isolation $Isol$

  The isolation is defined by the expression:

$$Isol = \frac{E_T^{cone} - E_T^{electron}}{E_T^{electron}} \qquad (2.2)$$

where $E_T^{electron}$ is the energy of the electron cluster and E_T^{cone} is the transverse energy in a cone of radius $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \leq 0.4$ around the electron cluster.

- Fiduciality

  This variable checks if the electron cluster position is in the region where the energy is well measured. For example there is the region where the two halves of the central calorimeter meet or the region at the edge of the calorimeter, which we don't use to an electron selection.

The following variables are used only for a plug electron identification (a plug electron is an electron detected by the PEM calorimeter):

- PES 2D η

  The η of the best matching 2D PES cluster

- χ^2_{3x3}

  The χ^2 comparison of the PEM shower profile with the shower profile obtained from the test beam measurements. (As was mentioned the maximum size of the cluster is the PEM is 3x3 towers.)

- U_{5x9} and V_{5x9}

  The isolation variables for the PES cluster defined separately for the U strips and V strips in the PES:

  $$U_{5x9}\ (V_{5x9}) = \frac{\text{energy sum in the 5 central strips}}{\text{energy sum of all 9 strips of the PES cluster}} \qquad (2.3)$$

- $\Delta R(PEM, PES)$

  The difference between the PEM and PES coordinates of the electron cluster.

- PhxMatch

  The variable which tells us if the electron cluster was matched by a Phoenix track. The Phoenix track is a track calculated by so called Phoenix algorithm [25], which is matched by the silicon detector hits. The Phoenix algorithm constructs two tracks (for the positive and negative charges) using the primary vertex position and the center of the PEM energy cluster. If the both constructed tracks are matched by the silicon detector hits, we use the one with the better fit. An electron candidate which energy cluster was matched by the Phoenix track is called Phoenix (PHX) electron.

- N^{Si}_{hits}

  The number of the silicon detector hits

We summarize all the cuts used for the CEM electron identification (an electron with the energy deposition in the CEM) in Table 2.1, while the cuts for the PHX electron are in Table 2.2.

The efficiencies of the CEM and PHX electrons identification cuts are determined from $\gamma/Z^* \rightarrow e^+e^-$ data sample. For the CEM electrons the efficiency is $\epsilon_{CEM} = 0.799 \pm 0.002$, while for the PHX electron it is $\epsilon_{PHX} = 0.658 \pm 0.004$. However due to the reconstruction differences between MC and data we need a scale factor to correct obtained acceptances from MC. The scale factor is defined as ratio of efficiency obtained from data over that one obtained from MC. The values of the scale

variable	value of the CEM cut
$|\eta|$	< 1.1
CES Fiduciality	$= 1$
E_T	≥ 20 GeV
$|z_0|$	≤ 60 cm
track p_T	≥ 10 GeV/c
E/P	≤ 2 if track $p_T < 50$ GeV/c
E_{had}/E_{em}	$\leq 0.055 + (0.00045.E)$
$Isol$	≤ 0.1
Conversion	$\neq 1$
L_{shr}	≤ 0.2
$|\Delta z_{CES}|$	< 3 cm
$\Delta x_{CES} * Q$	$-3.0 \leq \Delta x_{CES} * Q \leq 1.5 cm$
χ^2_{strip}	≤ 10
COT track quality	3 axial SLs and 2 stereo SLs with 5 hits per SL

Table 2.1: The CEM electron identification criteria.

factors are $SF_{\epsilon_{CEM}} = 0.981 + -0.003(stat) + -0.004(syst)$ for CEM electrons and $SF_{\epsilon_{PHX}} = 0.952 + -0.006(stat) + -0.012(syst)$ for PHX electrons.

2.2   Muon identification

Muons are minimum ionizing particles so they leave only a small energy deposits in the EM or HAD calorimeters. They can be identified by the COT track which corresponds to the muon track segment (called also stub) in the muon chambers.

Depending on the muon chamber where the stub is present, we distinguish the CMU, CMP, CMX and BMU muons. The CMUP muon is defined as the muon with stubs in both, the CMU and CMP chambers. It's also possible to use a stubless muons, which have no stub in any muon chamber, but are identified by the high p_T track. This type of muons is also called CMIO (central minimum ionizing object).

Similar to the electron case we define several variables, which are used to the muons identification:

variable	value of the CEM cut
PES 2D η	$1.2 \leq |\eta| \leq 2.8$
E_T	≥ 20 GeV
E_{had}/E_{em}	≤ 0.05
χ^2_{3x3}	≤ 10
U_{5x9}	≥ 0.65
V_{5x9}	≥ 0.65
$Isol$	≤ 0.1
$\Delta R(PEM, PES)$	< 3 cm
PhxMatch	$= 1$
N^{Si}_{hits}	≥ 3
$|z_0|$	≤ 60 cm

Table 2.2: The PHX electron identification criteria.

- p_T, E_{em}, E_{had}

  The transverse momentum, p_T, of the best matched track associated to the muon, the EM (HAD) calorimeter energy, E_{em} (E_{had}), corresponding to the muon.

- COT track quality cut

  The muon associated track is required to pass through a certain number of the COT super-layers (SLs).

- z_0

  The z-coordinate of the muon associated track at the distance of closest approach to the beamline.

- d_0

  The impact parameter d_0 is the distance of the muon associated track to the primary vertex (interaction point) in the $r - \phi$ plane. This variable helps to cut out the cosmic muons.

- Δx_{CMU}, Δx_{CMP}, Δx_{CMX}, Δx_{BMU}

  Distances between the COT track extrapolated to the muon chamber (CMU, CMP, CMX, or BMU) and the stub in the muon chamber.

22

- ρ_{COT}

  The radius of the muon associated track at the point where this track leaves the COT

- $Isol$

  The ratio of the energy deposited in the calorimeter tower within the cone ΔR = 0.4 around the muon associated track and the muon associated track p_T.

We summarize all identification cuts of the CMUP and CMX muons in Table 2.3, the CMIO and BMU muons identification cuts are present in Tables 2.4 and 2.5, respectively. The efficiencies of these identification cuts are determined from $\gamma/Z^* \to \mu^+\mu^-$ data sample.

variable	value of the CMUP or CMX cut
p_T	$\geq$ 20 GeV
E_{em}	$\leq 2.0 + Max(0, (p - 100) * 0.0115)$ GeV
E_{had}	$\leq 6.0 + Max(0, (p - 100) * 0.028)$ GeV
$|z_0|$	$\leq$ 60 cm
$|d_0|$	$\leq$ 0.02 cm (if $N_{hits}^{Si} > 0$)
$|d_0|$	$\leq$ 0.2 cm (if $N_{hits}^{Si} = 0$)
COT track quality	3 axial SLs and 2 stereo SLs with 5 hits per SL
$Isol$	$\leq$ 0.1
$|\Delta x_{CMU}|$	< 3 cm (only for CMUP muons)
$|\Delta x_{CMP}|$	< 5 cm (only for CMUP muons)
$|\Delta x_{CMX}|$	< 6 cm (only for CMX muons)
ρ_{COT}	$\geq$ 140 cm (only for CMX muons)

Table 2.3: The CMUP and CMX muons identification criteria.

2.3 Jet reconstruction and corrections

One can define jet as a shower of hadrons created in a quark or gluon hadronization process. Due to this fact the hadrons in a jet travels along the direction of the original parton.

variable	value of the CMIO cut
p_T	≥ 20 GeV
E_{em}	$\leq 2.0 + Max(0, (p - 100) * 0.0115)$ GeV
E_{had}	$\leq 6.0 + Max(0, (p - 100) * 0.028)$ GeV
$E_{em} + E_{had}$	> 0.1
$|z_0|$	≤ 60 cm
$|d_0|$	≤ 0.02 cm (if $N_{hits}^{Si} > 0$)
$|d_0|$	≤ 0.2 cm (if $N_{hits}^{Si} = 0$)
COT track quality	3 axial SLs and 2 stereo SLs with 5 hits per SL
$Isol$	≤ 0.1

Table 2.4: The CMIO muons identification criteria.

There are many algorithms which can be used for reconstruction of jets. The CDF II collaboration choose the so-called *jet clustering* algorithm as the default one. In this algorithm we look for the calorimeter tower with the maximum energy deposition which is denoted as a seed tower. Then the towers within a radius of $\Delta R = 04.$ with respect to the seed tower position are used to build clusters. We can define the centroid of the cluster by η and ϕ coordinates as follows:

$$E_T^{centroid} = \sum_{i=0}^{N_{tow}} E_{Ti} \qquad \phi^{centroid} = \sum_{i=0}^{N_{tow}} \frac{E_{Ti}\phi_i}{E_T^{centroid}} \qquad \eta^{centroid} \sum_{i=0}^{N_{tow}} \frac{E_{Ti}\eta_i}{E_T^{centroid}} \qquad (2.4)$$

where E_{Ti} is the transverse energy of the tower placed at (η_i, ϕ_i) and N_{tow} is number of towers. If the cluster centroid is different from the seed tower's coordinates, we define a new cluster with $\Delta R = 0.4$ around the $(\eta^{centroid}, \phi^{centroid})$ and calculate a new cluster centroid, which is compared with the previous one. This procedure is repeated iteratively until the geometrical center of the towers correspond to the cluster centroid [26].

After the finding of the final cluster, we can define the jet energy, E_{jet}, and momentum coordinates, $(p_x^{jet}, p_y^{jet}, p_z^{jet})$, as follows:

$$E_{jet} = \sum_{i}^{N_{tow}} E_i \qquad (2.5)$$

$$p_x^{jet} = \sum_{i}^{N_{tow}} E_i sin\theta_i cos\phi_i \qquad p_y^{jet} = \sum_{i}^{N_{tow}} E_i sin\theta_i sin\phi_i \qquad p_z^{jet} = \sum_{i}^{N_{tow}} E_i cos\theta_i \qquad (2.6)$$

24

variable	value of the BMU cut
p_T	≥ 20 GeV
E_{em}	$\leq 2.0 + Max(0, (p-100)*0.0115)$ GeV
E_{had}	$\leq 6.0 + Max(0, (p-100)*0.028)$ GeV
$|z_0|$	≤ 60 cm
$|d_0|$	≤ 0.02 cm (if $N_{hits}^{Si} > 0$)
$|d_0|$	≤ 0.2 cm (if $N_{hits}^{Si} = 0$)
COT hit fraction	> 0.6
$Isol$	≤ 0.1
$|\Delta x_{BMU}|$	< 9 cm

Table 2.5: The BMU muons identification criteria.

where E_i is the energy of the tower i. Using these definitions we can express the jet transverse energy, $E_{T,jet}^{raw}$, by equation:

$$E_{T,jet}^{raw} = E_{jet} \sin\theta_{jet} = E_{jet} \frac{p_T^{jet}}{\sqrt{(p_x^{jet})^2 + (p_y^{jet})^2 + (p_z^{jet})^2}} \tag{2.7}$$

where $p_T^{jet} = \sqrt{(p_x^{jet})^2 + (p_y^{jet})^2}$ is the jet transverse momentum. We call the thansverse energy *raw* because it needs to be corrected due to the detector and physics effects. The corrections scale the raw jet energy to the energy of the jet origin parton, so we often use also an expression of a jet energy scale (JES) [26].

Below we describe different types of corrections which are used by the CDF collaboration:

- The η dependence correction C_η (also called as relative correction) takes care of different calorimeters response. The central and plug calorimeters have different response to the same energy particle.

- The multiple interaction correction C_{MI} - the number of extra interactions which occurs in one bunch crossing, depends on instantaneous luminosity. At the Tevatron luminosity of 2×10^{32} cm^{-2}s^{-1} it means 6 extra interactions in average. This leads to an increase of the jet energy. The C_{MI} correction, dependent on the measured number of primary vertices, is applied to extract the extra energy.

- The absolute correction C_{ABS} converts the calorimeter energy to the particle-level energy. It is a correction of the calorimeter response which covers the calorimeter non-linearity effects and particle leakage effects. After applying this correction the jet energy is detector independent.

- The underlying events correction C_{UE} - The underlying events energy comes from beam remnants - the spectator partons with color connection to the other partons from proton (antiproton). This energy can be included into the jet cluster, so we need to subtract it from the jet energy.

- The out-of-cone correction C_{OOC} takes into account the fact that some particles can leave the jet cone during the fragmentation process because they have low p_T and their trajectories are bended in the magnetic field. This correction adds the lost energy to the energy of the jet.

All these corrections are determined as a function of the jet transverse momentum, p_T^{jet}, but they apply to all components of the four-momentum of the jet. After applying the corrections we can express the transverse momenta of the original parton, p_T^{parton}, as follows:

$$p_T^{parton} = (p_T^{jet} \times C_\eta - C_{MI}) \times C_{ABS} - C_{UE} + C_{OOC} \qquad (2.8)$$

Figure 2.1 shows the JES systematic uncertainties as a function of the corrected jet p_T. These uncertainties play important role in the systematic uncertainties in almost all analysis.

2.4   Primary vertex

The primary vertex is the point from which all prompt tracks originate. In order to find it we reconstruct the tracks' points of origin (vertices). However more than one of such vertices can be found. In this case we first identify which of them is nearest to the identified high momentum electron or muon. If no high momentum lepton is present we use the vertex with highest total scalar sum of transverse momentum of associated tracks [30].

The position of the primary vertex is determined by fitting together the tracks within a ± 1 cm window in z around this vertex. The procedure starts by fitting the
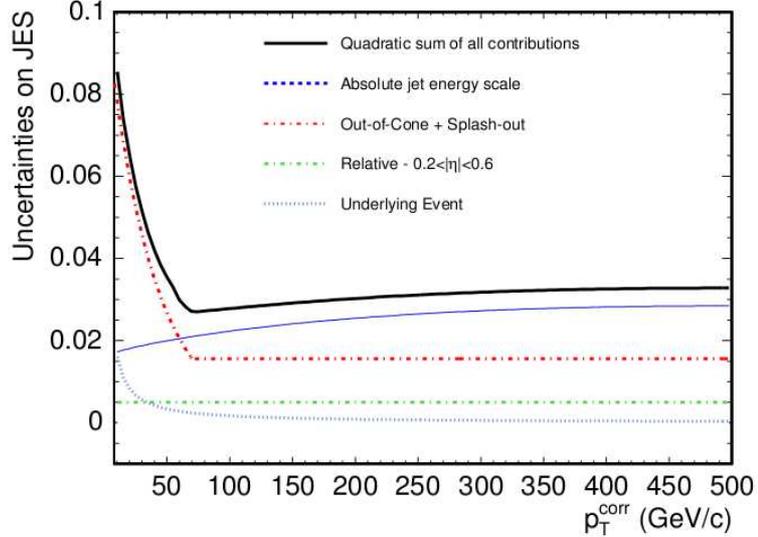
Figure 2.1: Systematic uncertainties on jet energy scale as a function of the p_T in the region of $0.2 < |\eta| < 0.6$ [26].

vertex using all tracks within the z window and with the impact parameter significance (relative to the beamline) $S_{d_0} = |d_0/\sigma_{d_0}| < 3$, where σ_{d_0} includes the uncertainty on both the track and the beamline positions. A pruning stage removes tracks which contribute with $\chi^2 > 10$ to the fit. After the initial pruning, the fit is repeated using only the remaining tracks until the vertex with no tracks over the χ^2 cut is found. If no tracks survive the pruning stage then the beamline profile is used for the primary vertex position estimate [30].

2.5 Missing transverse energy

As we already mentioned neutrino interacts only weakly, so it leaves the detector without any imprint. However using the energy and momenta conservation laws, we can determine the missing energy, which was carried by the neutrinos. It is defined as the transverse energy needed to have the total transverse energy in the event equal to 0. The missing energy can be expressed as:

$$\vec{\slashed{E}_T}^{\,raw} = -\sum_{i=0}^{N_{tow}} E_{Ti}^{raw} \vec{n}_i sin\theta_i \qquad (2.9)$$

where N_{tow} is the number of all towers in the event jet clusters, E_{Ti}^{raw} is the raw energy of the tower i with its center having the polar angle coordinate θ_i. The $\vec{n}_i$ is the unit

27

vector pointing to the tower center.

The raw missing E_T has to be corrected in each event individually. If the muon as a minimum ionizing particle is present in the event, we need to extract the calorimeters' energy deposits E_T^μ corresponding to the muon from the raw transverse energy, but we also need to include the muon transverse momentum p_T^μ:

$$\vec{\not{E}_T}^\mu = -\left(\sum_{i=0}^{N_{tow}} E_{Ti}^{raw} \vec{n}_i sin\theta_i - E_T^\mu + p_T^\mu \right) = \vec{\not{E}_T}^{raw} + E_T^\mu - p_T^\mu \qquad (2.10)$$

The second correction is connected to the jet energy corrections, which was mentioned in the previous section. Due to the fact that the missing transverse energy was calculated by using only the raw energies of the jets, we need to include the difference between the raw E_{Tjets}^{raw} and final E_{Tjets}^{corr} (corrected) jets' energies. That leads to the final expression for the missing transverse energy:

$$\vec{\not{E}_T} = \vec{\not{E}_T}^\mu - E_{Tjets}^{raw} + E_{Tjets}^{corr} \qquad (2.11)$$

2.6 b-jet tagging algorithms

At CDF, there are three methods usually used to identify a jet as a b-jet - the jet probability, the soft lepton tagger and the secondary vertex tagger. In this section we describe only the basic idea of the first two methods, while the last one will be described in detail, due to the fact, that we use it in our analysis.

Jet probability

This algorithm calculates the jet probability that a jet is consistent with a zero lifetime hypothesis. To determine the probability we use the impact parameters (d_0) of the tracks in the jet and their uncertainties. For the jets having only prompt tracks, this probability is uniformly distributed in the interval [0-1], while for the jets which originate in heavy partons, the probability distribution peaks at 0.

In detail this approach is described in [27].

Soft lepton tagger

In the method, the b-jets are identified by a presence of an electron or muon coming from the semileptonic decays of the b-hadrons inside the jet. These leptons have lower

(softer) transverse momenta, so we call this method the *soft lepton* b-tagging. We distinguish the soft electron SLT_e and soft muon SLT_μ tagger depending on the soft lepton type.

In both cases the algorithm begins by selecting the, so-called, taggable tracks which have to pass several criteria (e.g. the cuts for z_0, d_0, COT super layer hits, ...) different from those described in the electron or muon identification section.

Using the SLT_μ algorithm we extrapolate the taggable track into the muon chambers, where we look for stub(s). The differences between the extrapolated and measured positions in the x and z coordinates as well as the slope difference, $\Delta\phi_L$, are used to construct the global χ^2 quantity L. The final decision whether the jet is b-tagged or not, depends on the L value.

Something similar is done in the SLT_e algorithm. The taggable track extrapolation should be matched with an electromagnetic energy cluster. To distinguish an electron from a low p_T hadron we use the χ^2_{strip} and the distance Δ (in cm) between the extrapolated track and the position of the cluster energy centroid. The distributions of these variables help us to decide whenever the jet is b-tagged or not.

More details about the electron SLT_e or muon SLT_μ tagger are described in [28] or [29], respectively.

Secondary vertex tagger

This technique take advantage of the long lifetime of a b-hadrons to identify jets from the bottom quark hadronization through the presence of a decay vertex displaced from the primary interaction. It operates on a per-jet basis, where only tracks within the jet cone are considered for each jet in the event. The displaced vertex can be produced by at least two tracks, which passed a tracks' quality cuts and are selected using the significance S_{d_0} of their impact parameter with respect to the primary vertex [30].

The algorithm uses a two-pass approach to find the secondary vertices. In the first pass, using the tracks with $p_T > 0.5$ GeV/c and $|S_{d_0}| > 2.5$, it attempts to reconstruct a secondary vertex which includes at least three tracks (at least one of the tracks must have $p_T > 1$ GeV/c). If the first pass is unsuccessful, it performs a second pass which makes tighter track requirements ($p_T > 1$ GeV/c and $|S_{d_0}| > 3$) and attempts to reconstruct a two-tracks vertex (one track must have $p_T > 1.5$ GeV/c) [30].
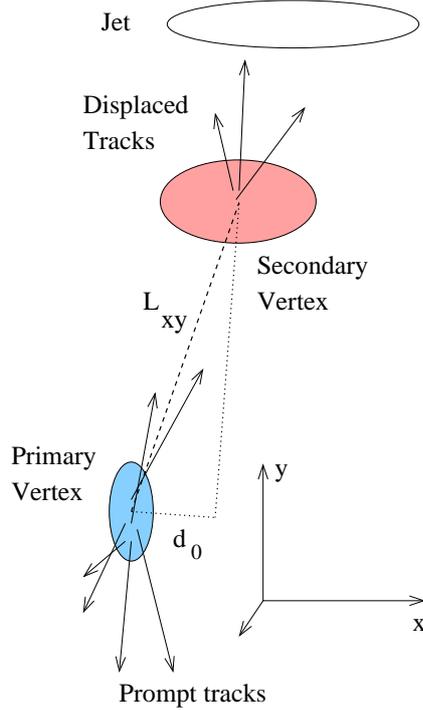
Figure 2.2: The two-dimensional decay length of the secondary vertex L_{xy}.

Once a secondary vertex is found in a jet, the two-dimensional decay length of the secondary vertex L_{xy} is calculated as its distance from the primary vertex in the transverse plane (see Figure 2.2). Secondary vertices corresponding to the decay of b and c hadrons are expected to have large L_{2D} while the secondary vertices from random mis-measured tracks are expected to be less displaced from the primary vertex. The tagged jet is then defined as a jet containing a secondary vertex with significance of the two-dimensional decay length $|S_{L_{xy}}| \equiv |L_{xy}/\sigma_{L_{xy}}| > 3$ [30].

This algorithm has several version, which differs from each other in the tracks selection cuts. In Table 2.6 we summarize the cuts for the tight and loose version of the tagger algorithm.

To measure the efficiency for tagging heavy flavor hadrons, we use a sample of low-p_T inclusive muon data which is enriched in semileptonic decays of the bottom and charm hadrons. This method is described in detail in [73]. In Table 2.7 we show the efficiencies of the tight and loose SecVtx tagger for the central jets ($|\eta| < 1.$) with $E_T = 60$ GeV/c^2 obtained from the $t\bar{t}$ MC sample. Figure 2.3 presents the dependence of these efficiencies on jet E_T. The dependence on jet η can be find elsewhere [31].

There is a possibility that b-tagged jets does not result from the fragmentation of a

cuts	loose SecVtx tagger	tight SecVtx tagger
χ^2 threshold for primary vertex finding	20	10
first pass for secVtx finding	$p_T^{track} > 0.5$ GeV/c $\quad$ $|S_{d_0}^{track}| > 2.5$	$p_T^{track} > 1$ GeV/c $\quad$ $|S_{d_0}^{track}| > 2$
second pass for secVtx finding	$|S_{d_0}^{track}| > 3$	$|S_{d_0}^{track}| > 2.5$
track $|S_{L_{xy}}|$	> 3	> 7.5

Table 2.6: The tracks selection cuts for the tight and loose version of secondary vertex (SecVtx) tagger.

heavy quark. We call such jets *Mistags*. They are caused mostly by a random overlap of the tracks which are displaced from the primary vertex due to the tracking errors, or they comes from K_S and Λ decays and nuclear interactions with the detector material (the beam-pipe or the inner silicon layers) as well [30]. In Table 2.7 we summarize the mistag rates (fractions) obtained from data for the tight and loose version of the SecVtx tagger.
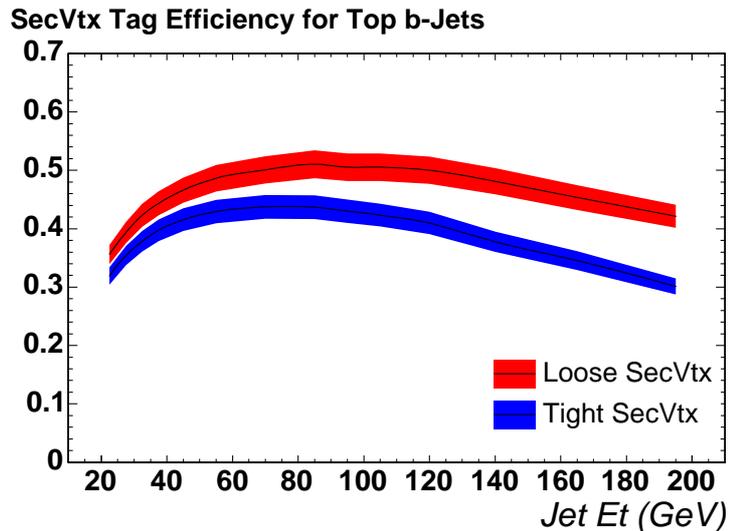
Figure 2.3: The b-tagging efficiency using the tight or loose SecVtx tagger, respectively for b-jets from top quark decays as a function of jet E_T.

SecVtx tagger	b-tag efficiency	Mistag rate
Loose	$\approx 50\%$	$\approx 2.8\%$
Tight	$\approx 45\%$	$\approx 1.4\%$

Table 2.7: The b-tag efficiency and mistag rate for a central jet ($|\eta| < 1$) with $E_T = 60$ GeV/c^2 from top quark events. The numbers for the tight and loose version of the SecVtx tagger are present.

2.7 JES validation

The jet energy corrections are mostly derived from the dijet samples. However we verify their validity and systematic uncertainties using γ-jet, Z-jet and $t\bar{t}$ events. Below we shortly describe the validation done using the γ-jet events, what presents my service work at the CDF [26].

The γ-jet events can be produced by the processes represented by the diagrams shown in Figure 2.4. The first diagram represents is the quark-antiquark annihilation where the jet originates in the produced gluon and the photon (γ) is radiated before the collision. In the second process gluon interaction with quark leads to the production of photon-quark pair (the photon comes from a virtual quark radiation).
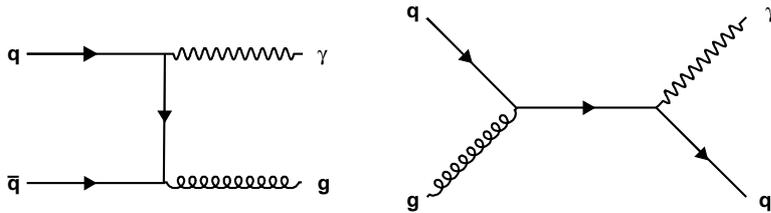
Figure 2.4: The Feynman diagrams of the γ-jet events production.

The photon transverse energy, E_T^γ, is measured accurately in the CEM calorimeter and thus provides a perfect reference for the jet transverse energy E_T^{jet}. If we require no addition jets, the energy conservation law leads to the fact that the jet energy should always balance the photon energy [26]. We can express the balance as follows[1]: .

---

[1]Due to the fact that the jet energy corrections are determined as a function of jet p_T (as was

$$\frac{p_T^{jet}}{p_T^{\gamma}} - 1 = 0 \tag{2.12}$$

where p_T^{jet} (p_T^{γ}) corresponds to the transvese momentum of the jet (photon). Figure 2.5 shows the p_T balance comparison of the data and the MC samples generated with HERWIG or PYTHIA for different jet cone sizes $\Delta R = 0.4$, 0.7 and 1.0. The difference is less than 3% [26].
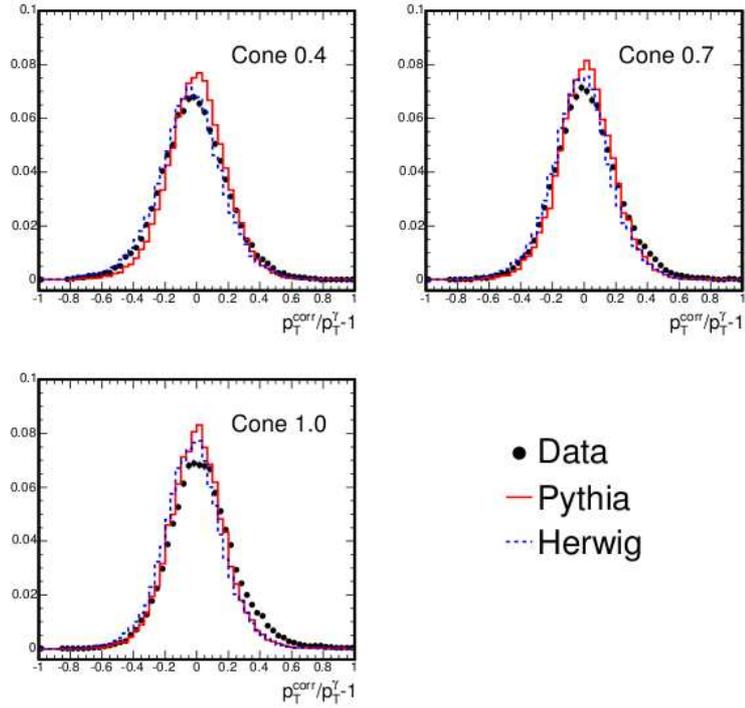
Figure 2.5: γ-jet balance in data, PYTHIA and HERWIG using $\Delta R = 0.4$, 0.7 and 1.0 after the η-dependent, absolute and OOC+UE corrections are applied [26].

.

My duty in the JES field was to validate the JES for a newly collected data using these γ-jet events. In the following we present the results after applying the all jet energy corrections.

As it is shown in Figure 2.6 we compare the latest collected data with the previous ones using the Gauss fit in the p_T balance range [-0.4, 0.4]. The data collected between February 2002 and September 2005 were set as the default. Table 2.8 shows the results of the fits. We can see that there is 2.8σ discrepancy between the latest collected data

mentioned before), we validate them in terms of the transverse momentum

and the default one. However it is not a problem, because the p_T balance in the latest data is closer to 0, what means a better result.
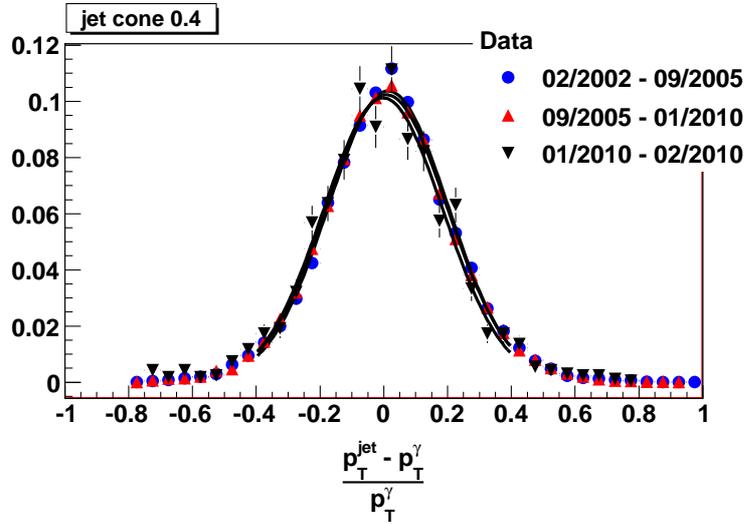
Figure 2.6: The p_T balance obtained from data for the central jets ($0.2 < |\eta| < 0.6$) after applying the all corrections.

Time	Gauss fit
02/2002 - 09/2005	0.0142 ± 0.0018
09/2005 - 01/2010	0.0088 ± 0.0011
01/2010 - 02/2010	-0.0017 ± 0.0054

Table 2.8: The mean values of the p_T balance Gauss fit obtained from data for the central jets ($0.2 < |\eta| < 0.6$) after applying all corrections. The data are divided into 3 subsamples by the time when they were collected.

Figure 2.7 shows the time dependence of the p_T balance. The first bin in this figure corresponds to the data used in Figure 2.5. We see that the balance is still within the 3% systematic uncertainty. One can say that the calorimeter response is stable over the time.
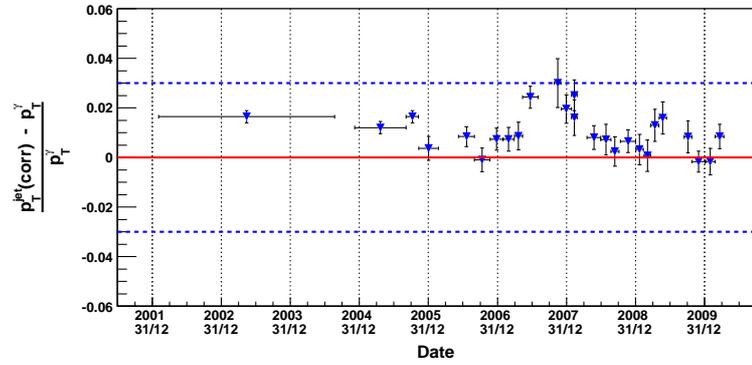
Figure 2.7: Time dependence of the p_T balance obtained from data for the central jets $(0.2 < |\eta| < 0.6)$ after applying all corrections. The blue dashed lines show the 3% systematic uncertainty, while the red line shows the perfect balance (equal to 0).

Chapter 3

Top quark physics

According to the SM, the top quark is the 3rd generation quark with the electric charge 2/3 and spin 1/2. Its mass is ≈ 35 times higher that the mass of the bottom quark and therefore, as we mentioned in Introduction, it is a very good object for testing of the SM and also for the search of a new physics.

3.1 Top quark production

At the Tevatron collider the top quarks can be produced by strong interactions or by electroweak ones. In the former case the top quarks are produced in pairs (as top-antitop pairs) and this mechanism is the main source of top quarks. In the electroweak mode single top quarks are produced and this production is suppressed in comparison with the main process about three times.

3.1.1 The top-antitop pairs production

The top-antitop $(t\bar{t})$ pairs come mainly from quark-antiquark $(q\bar{q})$ annihilation. At the Tevatron for Run II this happens 85% of the time. The rest - 15% of time the $t\bar{t}$ pairs are produced by gluon fusion (gg). The leading order diagrams for the production are shown in Figure 3.1.

Just for comparison at the LHC, CERN gluon fusion (90% [33]) dominates in the $t\bar{t}$ production, over the production by quark-antiquark annihilation (10%).

Using the factorization theorem [34] we can write an equation for the experimental differential cross-section of $t\bar{t}$ production in collision of two hadrons, proton, anti-proton
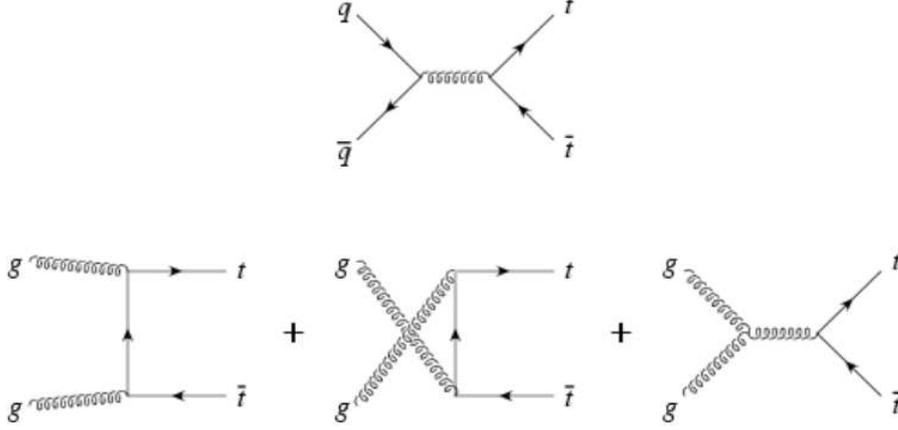
Figure 3.1: The leading order diagrams for $t\bar{t}$ production.

in the case of Tevatron experiments:

$$\sigma = \sum_{i,j} \int dx_1 dx_2 F_i^{(1)}(x_1, \mu_F) F_j^{(2)}(x_2, \mu_F) \hat{\sigma}_{ij}(s; \mu_F, \mu_R) \tag{3.1}$$

where $F_i(x_i, \mu_F)$ is the parton distribution function (PDF) for the parton i in the colliding hadron, μ_F and μ_R are the factorization and renormalization scales, respectively. x_i is the momentum fraction of interacting parton inside hadron, $\hat{\sigma}_{ij}$ is the parton cross-section and s is the square of partonic center-of-mass collision energy.

The factorization scale μ_F determine the hadron structure, which we can see by probing a hadron with a virtual photon having the squared momentum $q^2 = -\mu_F^2$. The renormalization scale μ_R determine $\alpha_S = \alpha_S(\mu_R^2)$, the running strong interaction coupling constant.

One can take scales $\mu_F = \mu_R = Q$, where Q^2 is characteristic scale of process, usually defined by mass or momenta of produced particle. In the case of top quark we use $Q^2 = m_t^2 + p_T^2$.

The PDF functions, that are not only functions of parton momentum fraction but also of factorization scale , are retrieved from a global fit of the deep inelastic scattering and Drell-Yan data at certain scale Q^2 and then the DGLAB evolution equation [35] is used to find the values of these functions at other (higher) scales.

The measured cross-section of the top-antitop ($t\bar{t}$) production (Figure 3.2) is $\sigma_{p\bar{p}}^{t\bar{t}} = 7.5 \pm 0.31$ (stat) ± 0.35 (syst) ± 0.15 (lumi) pb [39]. This result is in a good agreement with the theoretical cross-section for the top mass of 172.5 GeV/c^2 and for the current Tevatron center-of-mass energy $\sqrt{s} = 1.96$ TeV. Using the information in theoretical

references about the dependence of the prediction on the value of the top quark mass, the predictions were recalculated for the top quark mass of value of 172.5 GeV/c^2. The NNLO approximate calculation from Moch and Uwer [36] yields $\sigma^{t\bar{t}}_{p\bar{p}} = 7.46^{+0.66}_{-0.80}$ pb, the prediction from Cacciari et al. [37] is $\sigma^{t\bar{t}}_{p\bar{p}} = 7.26^{+0.78}_{-0.86}$ and from Kidonakis and Vogt [38] is $\sigma^{t\bar{t}}_{p\bar{p}} = 7.29^{+0.79}_{-0.85}$ pb.
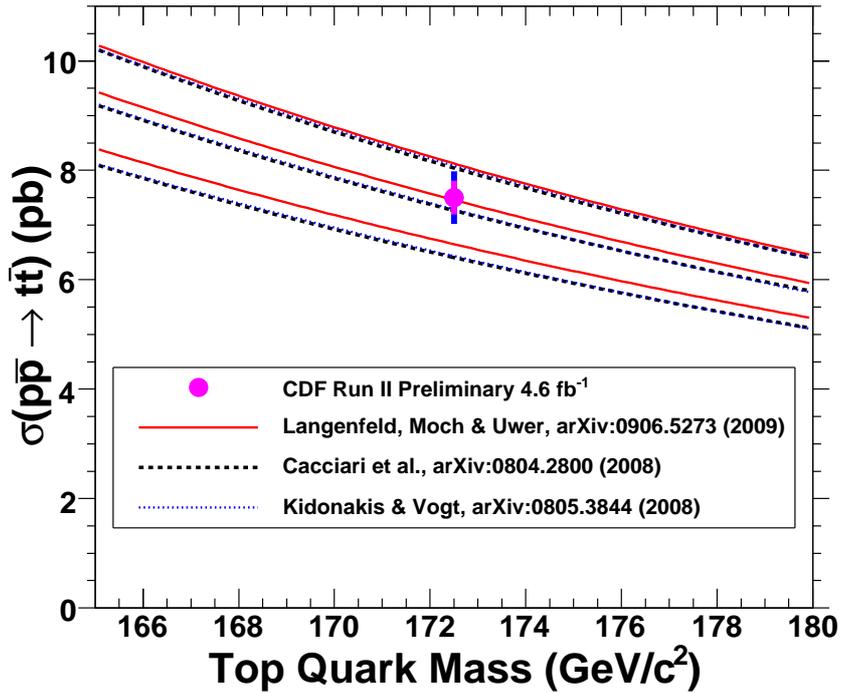
Figure 3.2: Combination of CDF measurements of the top quark pair production cross-section at 172.5 GeV/c^2 versus the theoretical predictions as a function of top quark mass.

3.1.2 Single top production

At the Tevatron energies, top quarks are predominantly produced in pairs via the strong interactions, but a significant amount of top quarks are produced singly via the electroweak interaction. there are three different production processes (see Figure 3.3):

- t-channel process: A virtual space-like ($q^2 \leq 0$) W boson interacts with a bottom quark from the "proton sea" and a top quark is created.

- s-channel process: A virtual time-like ($q^2 \geq 0$) W boson is created by the $q\bar{q}$ annihilation and decays into a (single) top quark and bottom quark.

- Associated production: A single top quark is produced with a real W boson. One of the initial partons is a bottom quark from the "proton sea".
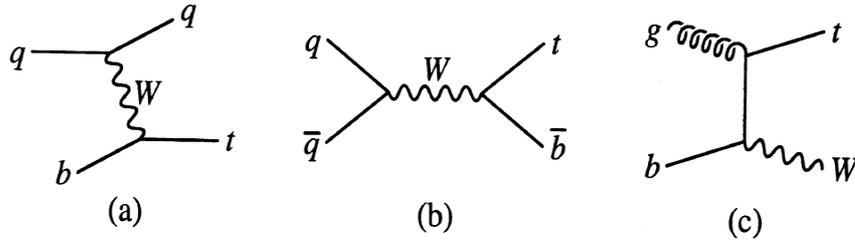
Figure 3.3: The leading order diagrams for single top production: a) t-channel process; b) s-channel process; c) associated production of top quark.

The cross-sections of those processes are smaller than that for the $t\bar{t}$ production. The predicted values by Harris and Sullivan's next-to-leading order (NLO) calculation [40] of the s- and t-channel cross-sections are $\sigma_{p\bar{p}}^t = 2.15 \pm 0.24$ pb and $\sigma_{p\bar{p}}^s = 0.99 \pm 0.07$ pb, respectively. But one can compare also to Kidonakis' NLO plus soft-gluon corrections calculation with cross-sections $\sigma_{p\bar{p}}^t = 2.34 \pm 0.13$ pb and $\sigma_{p\bar{p}}^s = 1.12 \pm 0.05$ pb [41]. In comparison with the $t\bar{t}$ pairs production, a fewer number of the signal events is expected with a higher background. It should be stressed that the single top production plays an important role. The reasons are the following:

- the cross-sections are proportional to the Cabbibo-Kobayashi-Maskawa (CKM) matrix element $|V_{tb}|^2$ (see Sec. 3.2) and these events are the only ones that enable to measure the element V_{tb} directly,

- events compose an important background for many Higgs and SuSy studies,

- the single top quarks are produced with almost 100% polarization and therefore could be used for a test of the V - A structure of the top quark charge weak current transitions.

In 2009 the CDF and D0 experiments published discovery of the single top quarks [42, 43]. The Tevatron combined s- and t-channel cross-section is $\sigma_{p\bar{p}}^{s+t} = 2.76_{-0.47}^{+0.58}$

(stat+sys) pb [44]. It is in agreement with Standard model expectations as is shown in Figure 3.4.
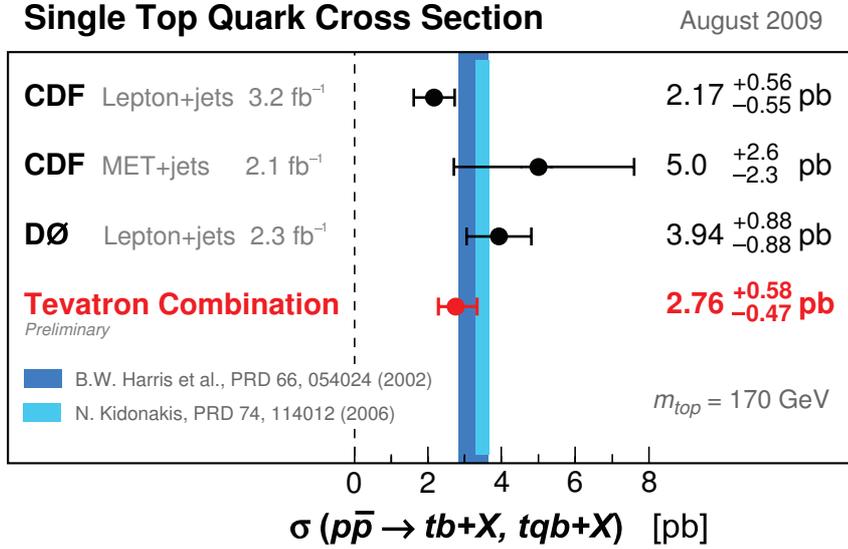
Figure 3.4: Tevatron single top cross-section measurements and their combination.

3.2 Top quark decay

The lifetime of the top quark is $\approx 5 \times 10^{-25}$ seconds, what is much shorter than the time needed for hadronization ($\tau_{hadr} \approx 10^{-23}$ seconds). Therefore we can study the top quark only via its decay products.

3.2.1 $|V_{tb}|$ matrix element

In the SM model top quark decays through the electroweak interaction predominately to b-quark, though decays to down (d) and strange (s) quarks are also possible. The probabilities of the transition to individual down quarks (d, s, b) are defined by the Cabbibo-Kobayashi-Maskawa (CKM) matrix V_{CKM} which is unitary. That means the sum of the transition probabilities is equal to one.

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}$$

The measured CKM matrix element for the top quark decay to bottom quark is $|V_{tb}| = 0.88 \pm 0.07$ and the limit was set $|V_{tb}| > 0.77$ at the 95% C.L. [44] So the top quark decays to down quark and strange quark are suppressed. Using the unitarity and assuming 3 generations of quarks the predicted values of V_{td}, V_{ts} and V_{tb} are [45]:

$$|V_{td}| = 0.00862^{+0.00026}_{-0.00020}$$
$$|V_{ts}| = 0.0403^{+0.0011}_{-0.0007}$$
$$|V_{tb}| = 0.999152^{+0.000030}_{-0.000045}$$

3.2.2 Decay width

The decay width for the dominant decay $t \to Wb$ can be theoretically expressed at next-to-leading-order as [46]:

$$\Gamma_{(t \to Wb)} = \frac{G_F m_t^3}{8\pi\sqrt{2}} \left(1 - \frac{M_W^2}{m_t^2}\right)^2 \left(1 + 2\frac{M_W^2}{m_t^2}\right) \left[1 - \frac{2\alpha_s}{3\pi}\left(\frac{2\pi^2}{3} - \frac{5}{2}\right)\right] \quad (3.2)$$

where G_F is Fermi constant and M_W is the W boson mass. The formula assumes that in the decay amplitudes are neglected terms of order m_b^2/m_t^2, α_S^2 and those of order $\alpha_S M_W^2/m_t^2$.

After applying QCD and electroweak corrections [47] the above equation predicts $\Gamma_{(t \to Wb)}$ of 1.3 GeV at the top quark mass of 172.5 GeV/c^2. By the direct measurement the CDF experiment established an upper limit at 95% confidence level of $\Gamma_{(t \to Wb)} < 7.6$ GeV [48], while D0 experiment indirectly measure value of $\Gamma_{(t \to Wb)} = 1.99^{+0.69}_{-0.55}$ GeV [49]. Both results are consistent with the Standard model prediction.

3.2.3 Decay modes

The W boson also lives very shortly and it is observable only via its decay products, too. It can decay either into two quarks ($W \to q\bar{q}$) or into lepton and neutrino ($W \to l\nu$).

The topology of $t\bar{t}$ events depends on the W boson decay. In each event we have two W bosons and so we distinguish the following three decay channels:

- **Lepton+jets channel**

  In these events one W boson decays to lepton (electron or muon) and neutrino and the second W boson decays to $q\bar{q}$ pair. Some missing transverse energy from

neutrino is present and the events contain at least four jets - two b-jets and two light jets from W decay.

Lepton+jets channel background comes from the generic QCD events with a fake W boson, W+multijet production, WW events and top events where only one top is produced. The main background is from the W+multijets production, but this is suppressed by the requirement on b-tagged jets.

This channel is called also golden channel due to background suppression and due to a relatively high gain (30% branching).

- **Dilepton channel**

  Both W bosons decay leptonically ($l\nu$), where by leptons we understand only electrons and muons. The case when W boson decays into τ lepton and τ neutrino is not included into the dilepton (neither into the lepton+jets one) top quark analysis, because it is hard to identify the hadronic decays of τ leptons. The W boson decays with tau are usually investigated as a special case.

  The signature of these events are two leptons (electrons or muons), a large amount of missing transverse energy (due to two neutrinos) and at least two jets.

  This channel is called cleanest, because there is small amount of background ($b\bar{b}$, WW, $Z \rightarrow \tau\tau$, Drell-Yan), but it occurs only in 5% of the time.

- **All hadronic channel**

  In this case both W bosons decay into $q\bar{q}$ pairs. So the event signature is at least six jets - two b-jets and four light jets from W boson decay.

  All hadronic channel occurs 44% of the time, but has a huge background from = mainly from the QCD multijets production processes.

The all possible $t\bar{t}$ decay modes are presented in Figure 3.5 where the fermions from W^- (W^+) are shown along the y(x) axis. The area of each region in the figure expresses the fraction of the shown decay mode.
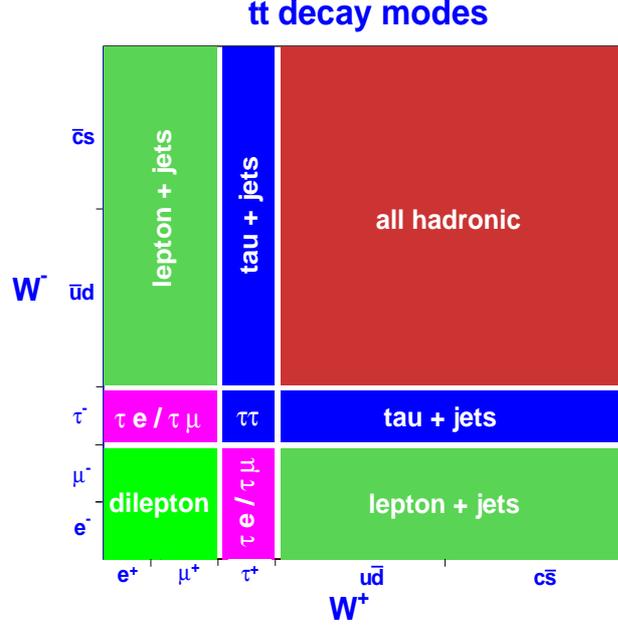
Figure 3.5: Representation of $t\bar{t}$ decay modes

3.3 Top quark mass

The top quark mass is a free parameter in Standard model. Together with W boson mass it is needed to limit the mass of Higgs boson (Figure 3.6). The masses of top quark, W boson and Higgs boson are bounded by the equation, which can be retrieved from the muon decay (or any other process going through W boson exchange) with the electroweak corrections included [50]:

$$M_W^2(1 - \frac{M_W^2}{M_Z^2}) = \frac{\pi\alpha}{\sqrt{2}G_F}(1 + \Delta r) \qquad (3.3)$$

where G_F is Fermi constant, α is the electromagnetic coupling constant, M_W and M_Z are the masses of W or Z boson, respectively and Δr accounts for radiative corrections, which are dependent on the top and Higgs masses. For the radiative corrections one can write [51]:

$$\Delta r = \Delta\alpha - \frac{cos^2\Theta_W}{sin^2\Theta_W}\Delta\rho + \Delta r_{rem} \qquad (3.4)$$

where Θ_W is the electroweak mixing angle, $\Delta\alpha$ contains the large logarithmic corrections from light fermions, $\Delta\rho \sim m_t^2$ and Δr_{rem} contains a dependence on the Higgs mass logarithm ($\Delta r_{rem} \propto lnM_H$).

The recent result of top quark mass is 173.3 GeV/c^2 measured with precision of 1.1 GeV/c^2. Figure 3.7 shows different measurements done by CDF and D0 experiments.
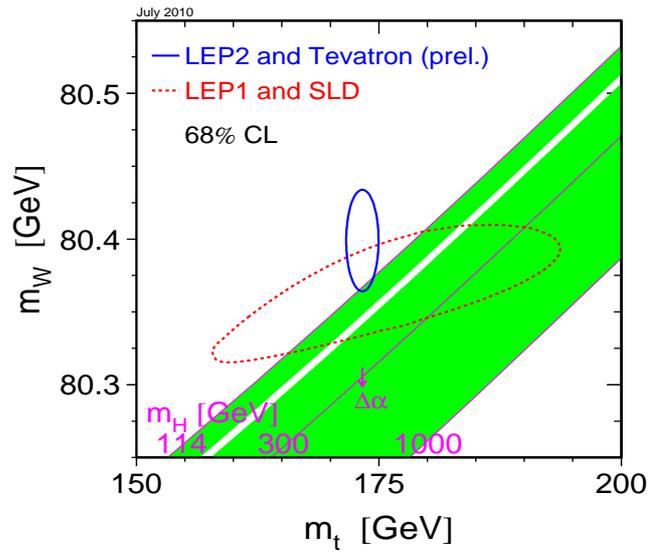
Figure 3.6: W boson vs. top quark mass with the lines of constant Higgs boson mass [52].
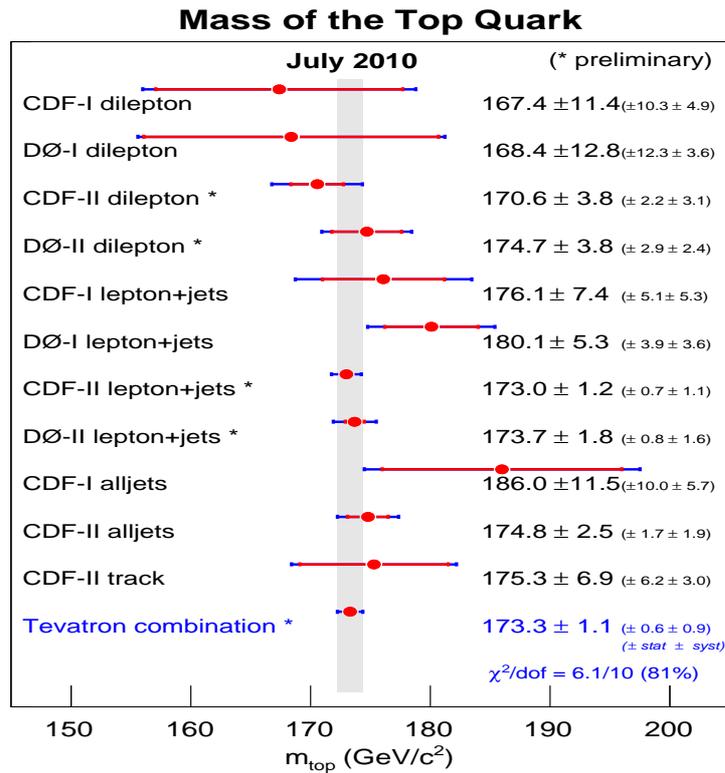
Figure 3.7: A summary of the input measurements resulting the Tevatron combination if top quark mass.

3.4 W boson polarization

Due to the short top quark's lifetime its spin characteristics are not diluted by hadronization and are transferred to its decay products. It enables to test the V-A structure of EW interactions by measuring the polarization amplitudes of W boson.

The W boson is a vector particle with the spin equal to 1. The projection of this spin to the direction of the motion (helicity) could be +1, 0, -1, what corresponds to the right-handed, longitudinal and left handed helicity states, respectively. Different helicity states of W bosons are reflected in the angular distribution of the decay products. The differential decay rate for top quarks is given by [53]:

$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\theta^\star} = f_- \cdot \frac{3}{8}(1 - \cos\theta^\star)^2 + f_0 \cdot \frac{3}{4}(1 - \cos^2\theta^\star) + f_+ \cdot \frac{3}{8}(1 + \cos\theta^\star)^2 \qquad (3.5)$$

where $\theta^\star$ is the angle between momentum of the charged lepton (or down type quark) in the W rest frame and the momentum of the W boson in the top quark rest frame; f_-, f_0 and f_+ are the fractions for left-handed, longitudinal and right-handed helicity states, respectively and $f_- + f_0 + f_+ = 1$. The Standard Model predicts the specific fractions for the cases when top quark decays into definite helicity states of W boson [54]:

$$f_0 \simeq 0.703 \pm 0.016$$
$$f_+ \simeq 3.6 \times 10^{-4} \pm 0.0045$$
$$f_- \simeq 0.297 \pm 0.016$$

A model independent simultaneous measurement of f_0 and f_+ done by CDF using 4.8 fb^{-1} of the data yields

$$f_0 = 0.78^{+0.19}_{-0.20}(\text{stat}) \pm 0.06(\text{syst})$$
$$f_+ = -0.12^{+0.11}_{-0.10}(\text{stat}) \pm 0.04(\text{syst})$$

what is consistent with the Standard model expectations [55].

3.5 Forward-backward asymmetry

One of the reason for a forward-backward asymmetry in top production is that the next-to-leading-order (NLO) QCD predicts a small but non-zero charge asymmetry.

Heavy flavor pair production via $q\bar{q}$ or gg does not discriminate between quark and anti-quark in the leading order. But the radiative corrections at the next-to-leading order involve a virtual (or real) gluon in $q\bar{q} \to Q\bar{Q}$ what leads to a difference in the Q and $\bar{Q}$ production, and therefore a charge asymmetry.

The asymmetry gets a positive contribution from interference of the tree-level and box diagrams, as in the upper diagrams in Figure 3.8 and a negative contribution from the interference of initial and final state radiation in $t\bar{t}$ + jet ($t\bar{t}j$) final states, as in the lower diagrams [57]. The overall charge asymmetry is positive and predicted to be about 6% [36].
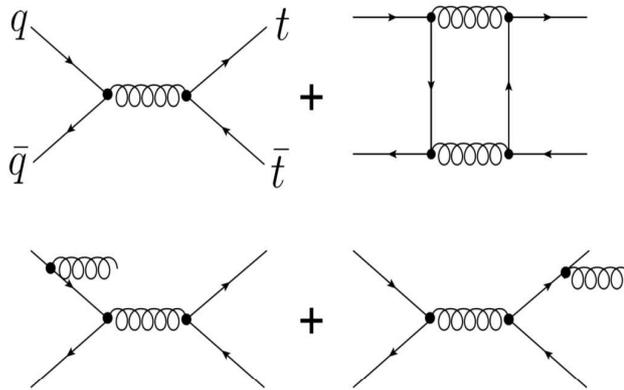
Figure 3.8: Interfering $q\bar{q} \to t\bar{t}$ (above) and $q\bar{q} \to t\bar{t}j$ (below) amplitudes.

Using the 5.1 fb^{-1} of data, the CDF measures the asymmetry value of $A_{FB} = 0.42 \pm 0.15$(stat) ± 0.05(syst), what means the the asymmetry is 2.3σ from the standard model prediction ($A_{FB}^{theory} = 0.06 \pm 0.01$). [56]

3.6 Spin correlation

The very short top quark life time (shorter than time needed for hadronization) leads to the top quark decays before it can form hadronic bound states. Thus the top quark spin information is not diluted, but is passed to the spins of the top quark decay products. The change of top quark spin-state by emitting a gluon before its decay is very unlikely due to the fact that the spin-flip time is much larger that top quark life time. Therefore the spin correlation can by determined by measuring the angular distributions of the t ($\bar{t}$) quark decay products. However the result depends on the choice of the quantization basis for the t and $\bar{t}$ quark spin. We can use:

- *helicity basis*, where we use the top quark momentum as the spin quantization axis,

- *beamline basis*, where the beam line direction is used as the spin quantization axis

- and *off-diagonal basis*, shown in Figure 3.9, is defined by axis with an angle ξ toward clockwise direction from the t or $\bar{t}$ quark momentum, respectively. The angle ξ is defined as:

$$\tan \xi = \sqrt{1 - \beta^2} \tan \theta^\star \tag{3.6}$$

where β is the top quark velocity and $\theta^\star$ is the angle of the top quark flight direction with respect to the proton direction in the $t\bar{t}$ center of mass frame [58].
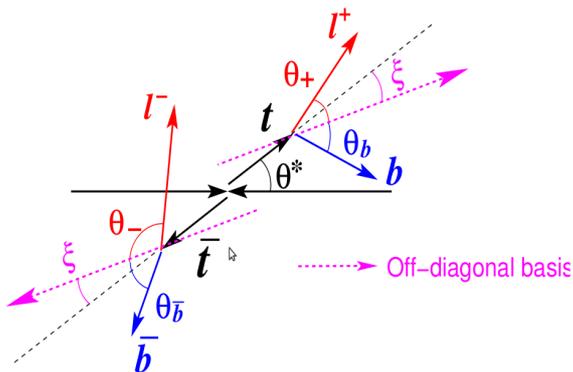
Figure 3.9: The definition of the angles in off-diagonal basis used in the top quark spin correlation measurement.

In the CDF analysis [58] done using the dilepton events (integrated luminosity of 2.8 fb^{-1}), the spin correlation, κ, can be expressed from the differential cross-section of the $t\bar{t}$ decay:

$$\frac{1}{\sigma} \frac{d^2\sigma}{d\cos\theta_+ d\cos\theta_-} = \frac{1 + \kappa \cos\theta_+ \cos\theta_-}{4} \tag{3.7}$$

where θ_+ (θ_-) is the angle of the positive (negative) lepton flight direction with respect to the off-diagonal basis quantization axis. The obtained value of spin correlation $\kappa = 0.320^{+0.545}_{-0.775}$ is consistent with the SM prediction of $\kappa \sim 0.8$ [59].

The other CDF analysis [60] done using the 5.3 fb^{-1} of lepton+jet events measures the spin correlation in the helicity and beamline basis. They use the four independent

helicity states: $\bar{t}_L t_R, \bar{t}_R t_L, \bar{t}_L t_L, \bar{t}_R t_R$ to describe the $t\bar{t}$ spin:

$$\kappa = \frac{[\sigma(\bar{t}_R t_L) + \sigma(\bar{t}_L t_R)] - [\sigma(\bar{t}_R t_R) + \sigma(\bar{t}_L t_L)]}{\sigma(\bar{t}_R t_R) + \sigma(\bar{t}_L t_L) + \sigma(\bar{t}_R t_L) + \sigma(\bar{t}_L t_R)} = \frac{N_o - N_s}{N_o + N_s} \qquad (3.8)$$

where the $\sigma(X)$ corresponds to the cross-section of the helicity state X, and N_o (N_s) is number of opposite (same) helicity fraction. Note that in the $t\bar{t}$ rest frame the quarks move back-to-back and the same spin states are those with opposite helicity $\bar{t}_L t_R, \bar{t}_R t_L$. By fitting the data with the same and opposite spin templates, the spin correlations in helicity and beamline basis were obtained:

$$\kappa_{helicity} = 0.48 \pm 0.48(\text{stat}) \pm 0.22(\text{syst})$$
$$\kappa_{beam} = 0.72 \pm 0.64(\text{stat}) \pm 0.26(\text{syst})$$

These values are consistent with the NLO SM predictions of $\kappa_{helicity} = 0.168$ in helicity basis and $\kappa_{beam} = 0.37$ in beamline basis [59].

3.7 Flavor changing neutral currents

The flavor changing neutral current (FCNC) couplings like tVc and tVu, where $V = g, \gamma, Z$, are highly suppressed in the SM (on the tree level they are forbidden). The observation of these processes would signal the existence of a new physics. In Table 3.1 we show the predicted values of the top quark branching rations in different models (e.g. SM, exotic quarks models, Supersymmetric models).

	SM [61]	SuSy models [62]	Exotic quarks models [63]
$B(t \to qg)$	5×10^{-11}	$\sim 10^{-3}$	$\sim 5 \times 10^{-4}$
$B(t \to q\gamma)$	5×10^{-13}	$\sim 10^{-5}$	$\sim 10^{-5}$
$B(t \to qZ)$	$\sim 10^{-13}$	$\sim 10^{-4}$	$\sim 10^{-2}$

Table 3.1: The predicted values of the branching ratio for the FCNC top quark decays in the SM, exotic model and Supersymmetry (SuSy) model.

The CDF measurements constrain the upper limits to the branching ratios: $B(t \to qZ) < 3.7\%$ at 95% confidence level [64], using the 1.9 fb^{-1} of the data, $B(t \to u + g) < 3.9 \times 10^{-4}$ and $B(t \to c + g) < 5.7 \times 10^{-3}$ [65], using the 2.2 fb^{-1} of the data.

Chapter 4

The top quark charge determination

The top quark lifetime is very sort, so before hadronization it decays predominantly to W boson and bottom quark. Its predicted charge is $+2/3$ of the elementary electric charge.

However, four years after the Tevatron observed the top quark [3, 4], the hypothesis that the observed particle can be a non-Standard model object with charge $-4/3$ was put forward [67]. The exotic scenario (XM - Exotic Model) assumes existence of additional heavy quarks Q_3 and Q_4 with charges $-1/3$ and $-4/3$, respectively. With the additional heavy quarks two physical quark doublets:

$$\begin{pmatrix} Q_3 \\ Q_4 \end{pmatrix}_L \qquad \begin{pmatrix} Q_3 \cos\Theta_b - b\sin\Theta_b \\ Q_4 \end{pmatrix}_R$$

and one singlet:

$$(Q_3 \cos\Theta_b + b\sin\Theta_b)_R$$

are added to the SM quark doublets and singlets. Hence, in addition to the SM, Q_3 mixes with b (mixing angle θ_b) and Q_4 decays into $b + W^-$.

In the XM, also the Q_4 and left-handed top quark masses were calculated. The Q_4 mass is approximately as large as the mass of the observed top quark (173 GeV/c^2) while left-handed top quark should have the mass around 274 GeV/c^2. The search for the forth generation of quarks done by CDF sets the lower limits of the t' and the b' quarks masses with 95% confidence level to the values 360 GeV/c^2 and 380 GeV/c^2,

respectively [15, 14], what exclude Chang's model. But the question about the top quark charge remains.

One possibility of the determination of the top charge is the direct measurement of photon radiation in $t\bar{t}$ events. One can look for the photon radiated by the top quark what gives access to its charge and electromagnetic coupling. Unfortunately, this type of measurement on the CDF is limited by the statistics [66].

The other possibility is to reconstruct the top quark charge of electric charges of its decay products. This approach consists of tree steps:

- determination of W boson charge,

- assignment of b-jet to the W boson

- and determination of b-jet charge.

As we already mentioned, in the lepton+jet (LJ) channel one W boson decays leptonically and the other one decays top light quarks. According to this we distinguish the leptonic and hadronic decay branch, respectively.

The W boson charge in leptonic branch is defined by the charge of the lepton while in the case of hadronic branch W boson its define as the opposite charge to the lepton charge, because we expect W bosons to have opposite charges.

The pairing of the W boson and b-jet is done by the kinematic fitter developed by CDF, which reconstruct the events and based on the χ^2 choose the combination with higher probability. More detailed description is in Section 4.3.

The last step is to determine the b-jet charge for what we use the Jet Charge algorithm, which is described in Section 4.4.

The study using the same methodology was done on the 1.5 fb^{-1} of the CDF II collected data (see [69]). However, from that time, there were several changes. The first is more precise measurement of top quark mass, what leads to the changing of kinematic fitter parameter - the top quark mass from 175 GeV/c^2 to 172.5 GeV/c^2. The second change is connected to the fact that the CDF collaboration started to use another tracking scripts to reconstruct the data collected from April 2008 to February 2010 what could affect our calculation of the jets' charges.

There is also another study of the top quark electric charge done on 3.2 fb^{-1} of the CDF II collected data, which uses the soft lepton tagger to determine the b-jet flavor. This study exclude the Exotic model with 95% confidence level [70].

4.1 Event selection

In Section 3.2 we divided events into three categories based on decay products. Our study is based on the lepton+jets channel so in this section we describe only events' selection for this channel.

4.1.1 One lepton requirement

We use three lepton datasets. Two of them contain events triggered by electrons and muons, which passed the central electron (CEM18) and central muon (CMUP18, CMX18) triggers, respectively. The last one is the non-triggered muons dataset, called "*extended muons*". In this datasets muons are reconstructed off-line. We don't use the τ lepton due to the hard identification because of its hadronization, but if the τ decays leptonically it can contribute to the electron or muon channel.

Electron

A candidate is selected by L1 trigger by requiring a track with $p_T > 8$ GeV/c, matched to a single tower in the CEM calorimeter having transverse energy $E_T > 8$ GeV. The hadronic to electromagnetic energy ratio (E_{had}/E_{em}) has to be less than 0.125.

At the L2 trigger, the cluster formed by adding the energy from neighboring towers in CEM calorimeter is required to have energy $E_T > 16$ GeV, while the E_{had}/E_{em} ratio has to remain the same (< 0.125).

Finally the L3 trigger requires: track with $p_T > 9$ GeV/c matched to a cluster of energy in CEM calorimeter with $E_T > 18$ GeV; ratio $E_{had}/E_{em} < 0.125$ [30].

Triggered muon

We can divide this type of dataset into two subsamples by the trigger which is used to select muons. Then we distinguish between the CMUP and the CMX muons which

51

were selected by the CMUP18 or CMX18 triggers, respectively.

The selection is similar, but trigger cuts are different. While for the CMUP muon the L1 trigger requires the COT track with $p_T > 4$ GeV/c, which is matched to a muon track segment with $p_T > 6$ GeV/c from the CMU and the CMP muon chambers, in case of the CMX muon the L1 trigger requires the COT track with $p_T > 8$ GeV/c, which is matched to a muon track segment with $p_T > 6$ GeV/c from the CMX muon chamber [30].

The L2 trigger automatically accepts events passing the L1 trigger [30].

To pass the L3 trigger a muon track has to have $p_T > 18$ GeV/c. In the case of the CMUP muon this track has to match the track segment in the CMU muon chamber within $|\Delta x| < 10$ cm and within $|\Delta x| < 20$ cm the track segment in the CMP muon chamber while in the case of the CMX muon the track has to match the track segment within $|\Delta x| < 10$ cm in the CMX muon chamber [30].

Non-triggered muons

To increase the muon acceptance we used non-triggered muons. These events have to pass trigger which requires the missing transverse energy (MET) of 25 GeV at the L1 trigger, adds a requirements to find at least two jets with $E_T > 10$ GeV at the L2 trigger and at the final L3 trigger fortifies the cut on MET to be higher than 35 GeV.

In addition, and to ensure full efficiency of the trigger, there events require at least two jets with $E_T > 25$ GeV, one of which should be central ($|\eta| < 0.9$) and separated by $\Delta R_{jj} > 1.0$.

We use several types of muons, which are described as follows:

1. Muon with a stub in the CMX muon chamber, which fails either the ρ cut or is in a CMX part that is bad for the trigger.

2. Central muon with track segment in only the CMU muon chamber and non-fiducial to the CMP muon chamber.

3. Central muon with track segment in only the CMP muon chamber and non-fiducial to the CMU muon chamber.

4. A muon that is not fiducial to any muon detector, has no stub in any detector, but passed the other standard muon selection cuts.

Further lepton cuts

All the above leptons types have to be isolated and are required to have E_T (in the case of muons p_T) above 20 GeV (GeV/c). To cut out some other possible processes which could lead to selection of wrong lepton, we use several vetos:

- *Dilepton veto.* If more than one lepton is present, the event is removed. It is done to avoid overlap with the dilepton channel.

- *Z veto.* If lepton with some another object have invariant mass with energy between 76 GeV and 106 GeV, event is vetoed to remove muons from Z boson decays.

- *Cosmic veto.* The cosmic muon which pass the detector close to the beam line may be reconstructed as a pair of charged particles. Using the timing capabilities of the COT we check if one of the tracks travel toward instead of away from the center of detector. If it is the case we reject such event [30].

- *Conversion veto.* If lepton is an electron coming from photon conversion, event is vetoed. The conversions are identified by a characteristic small opening angle between two oppositely charged tracks, that are parallel at their distance of closest approach to each other [30].

4.1.2 Jets and MET cuts

In addition to the one isolated lepton requirement, we select the events with at least four jets. Three of them have to be "tight" ($E_t > 20$ GeV, $|\eta| < 2.0$) and the fourth jet can be a tight or loose jet, respectively (loose jet: $E_t > 12$ GeV and $|\eta| < 2.4$). At least two of the jets has to be b-tagged by the loose SecVtx tagger [30].

Due to the presence of neutrino from the W boson decay we require to have the missing transverse energy $\not{E}_t > 20$ GeV.

These selection cuts are a result of a series of optimizations carried out using the corresponding Monte Carlo samples.

4.2 Definition of performance

The top quark charge analysis consists of several algorithms which contain some parameters. To optimize the parameters we need a quantitative criteria for picking the best options. We use a combination of two quantities - efficiency (ϵ) and purity (P):

- Efficiency is defined as fraction of events remaining after certain selection criteria from all events available before applying the criteria.

- Purity is define as the number of events that are correctly identified (based on MC information) over the number of events remaining after the selection. Denoting the N_{right} and N_{wrong} as the numbers of correctly and incorrectly identified events, the purity can be expressed as:

$$P = \frac{N_{right}}{N_{right} + N_{wrong}} \tag{4.1}$$

Let us now formally define N_+ as the number of events in favor of the SM hypothesis and N_- as the number of events in favor of the exotic model hypothesis. Then one can calculate the measured asymmetry as:

$$A_{meas} = \frac{N_+ - N_-}{N_+ + N_-} \tag{4.2}$$

while the true asymmetry is given by:

$$A = \frac{N_+^t - N_-^t}{N_+^t + N_-^t} \tag{4.3}$$

where N_+^t is number of true SM events and N_-^t is number of true exotic model events - here we formally assume that both hypotheses can occur in parallel.

One can define dilution factor related to the purity by $D = 2P - 1$ as follows:

$$D = \frac{N_{right} - N_{wrong}}{N_{right} + N_{wrong}} \tag{4.4}$$

where N_{right} and N_{wrong} are the number of rightly and wrongly assigned events.

That leads to the relation between the measured and true asymmetry:

$$A = \frac{A_{meas}}{D} \tag{4.5}$$

The dilution is equal to 1 if there are no wrongly assigned events and the measured asymmetry would be the same as the true asymmetry. If the dilution is equal to 0, we

can not distinguish between right and wrong events. So we want to have the dilution as high as possible.

It can be shown that the statistical uncertainty of the true asymmetry can be expressed as:

$$\sigma_A = \sqrt{\frac{1 - D^2 A}{\epsilon D^2 N}} \tag{4.6}$$

where N is the number of all events available for analysis and ϵ is the efficiency of the assignment method. Since the uncertainty scales with $1/\sqrt{\epsilon D^2 N}$ rather than $1/\sqrt{N}$, the ϵD^2 (performance) was chosen as the quantity to optimize for a given algorithm of interest.

4.3   W boson and b-jet pairing

The method is based on the full reconstruction of event topology by a kinematic fitter. Generally, there is one lepton and at least four jets, two of them tagged as b-jets. The b-jets are assumed to come directly from top quark decay, the non-b-jets come from the W boson decay and from gluons. The b-jet is called leptonic (hadronic) if it belongs to leptonic (hadronic) decay branch.

If there are more than four jets, we take into account only 4 of them with the highest transverse momenta, so we have 12 combinations how to pick up two b-jets. Due to unknown neutrino momentum z-component we finally get 24 combinations. For each of them, the full kinematic fit of the event is done using TopMassFitter (part of the CDF analysis tools package), which calculate a χ^2 value given by expression:

$$\chi^2 = \sum_{i=l,4jets} \frac{(\hat{p}_T^i - p_T^i)^2}{\sigma_i^2} + \sum_{j=x,y} \frac{(\hat{p}_j^{UE} - p_j^{UE})^2}{\sigma_j^2} + \frac{(m_{jj} - m_W)^2}{\Gamma_W^2}$$
$$+ \frac{(m_{l\nu} - m_W)^2}{\Gamma_W^2} + \frac{(m_{bjj} - m_t)^2}{\Gamma_t^2} + \frac{(m_{bl\nu} - m_t)^2}{\Gamma_t^2} \tag{4.7}$$

where the first term considers the difference in the transverse momenta between fitted $(\hat{p}_T^i)$ and reconstructed (p_T^i) transverse momentum values of the lepton and jets with σ_i (i=l,jets) to be lepton and jets p_T resolution. The second term considers the difference between the fitted and measured components of the unclustered energy (the energy out of the jet clusters) determined with resolutions $\sigma_{x,y}$. The following two terms calculate the mass difference between W (m_W) and its decay products – jets (m_{jj}) or leptons

($m_{l\nu}$). The remaining terms calculate the mass difference between top quark (m_t) and its decay products in the hadronic (m_{bjj}) and in leptonic ($m_{bl\nu}$) branches. Each of the terms is divided by the corresponding decay width Γ_W and Γ_t, respectively. The W boson mass, its decay widths Γ_W and Γ_t are fixed on their PDG values.

Combinations where leptonic and hadronic b-jets are b-tagged by the secondary vertex (SecVtx) tagger are called tagged combinations. To pick up the right kinematic topology, we require the combinations to be tagged and to have the lowest χ^2 value. To improve the purity of this method an additional requirement of $\chi^2 < 9$ is applied [68].

Fitter can work in two modes - constrained (using top mass $m_t = m_{\bar{t}} = 172.5$ GeV/c^2) or unconstrained (using the top mass as a free parameter). In Table 4.1 we compared these two modes by the performance defined in Section 4.2. The constrain fit gives us a better resolution. The distribution of the minimum χ^2 distribution for the constrained fit is shown in Figure 4.1.

Fitter type	efficiency ϵ (%)	purity P (%)	ϵD^2
constrained	53.2 ± 0.1	83.3 ± 0.1	0.236 ± 0.001
unconstrained	65.8 ± 0.1	72.6 ± 0.1	0.134 ± 0.001

Table 4.1: Efficiency, purity and performance (ϵD^2) for two different fitter modes.

As we have already mentioned in the previous study done on the 1.5 fb^{-1} of data, the top quark mass used in the kinematic fitter was constrained to 175 GeV/c^2. In this study we use the value of 172.5 GeV/c^2 which is closer to the measured one. We also use the new MC sample which was generated using this better value. To be sure that this change does not affect our analysis we checked the differences obtained from the old and new MC samples as well as that one from data.

In the MC samples we do not see big changes - the results are consistent with each other, if there is a differences it will be taken into account at the systematic errors evaluation.

However we are also wonder how it affects the data. We checked 1.9 fb^{-1} of the data using these two different values of the top quark mass. The results are present in Tables 4.2 and 4.3. We see that the changes are minor and are within statistical uncertainty. In these tables the SM like pairs in the last column "SM/XM like pairs"
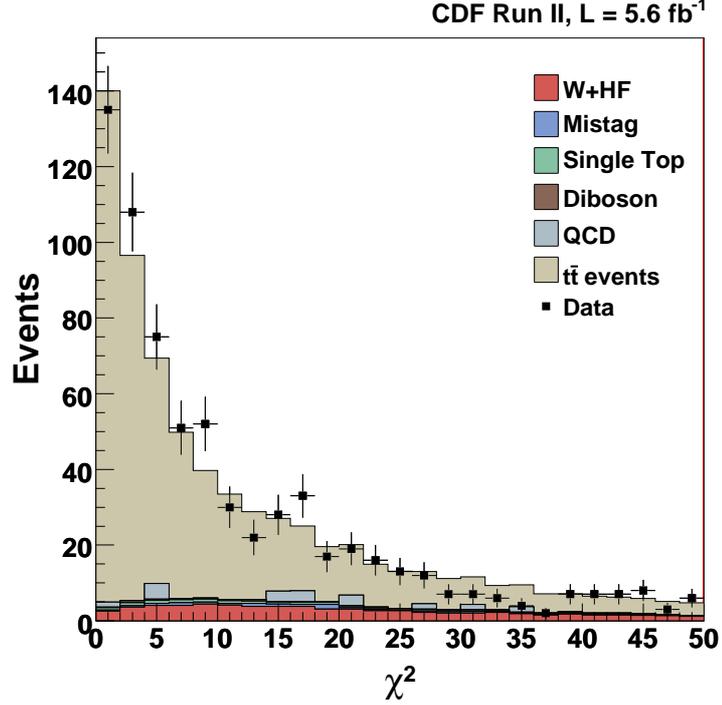
Figure 4.1: The minimum χ^2 distribution for the constrain fit. The MC distribution is normalized to the number of data. Different background samples are described in Section 6.2. For the analysis we use only events with $\chi^2 < 9$.

correspond to the W^+b ($W^-\bar{b}$) pairs, while the XM like pairs correspond to W^-b ($W^+\bar{b}$) pairs.

Lepton	# of events	after pairing cut	# of pairs	SM/XM like pairs
Electrons	136	68	133	82/51
CMUP muons	71	39	75	32/45
CMX muons	32	20	39	14/25
Non-triggered muons	60	25	48	25/23

Table 4.2: The data yields for the different lepton type obtained on 1.9 fb^{-1} of data using the top quark mass in the kinematic fitter constrained to 175 GeV/c^2.

Lepton	# of events	after pairing cut	# of pairs	SM/XM like pairs
Electrons	136	71	139	85/54
CMUP muons	71	40	77	34/43
CMX muons	32	20	39	16/23
Non-triggered muons	60	25	48	27/21

Table 4.3: The data yields for the different lepton type obtained on 1.9 fb^{-1} of data using the top quark mass in the kinematic fitter constrained to 172.5 GeV/c^2.

4.4   Jet Charge Algorithm

The last ingredient of the analysis is the flavor tagging of the b-jet. Combined with the pairing method, we will able to tell if the top quark candidate events have decayed into W^+b or W^-b. The determination of the average b-jet charge is done using a weighting technique based on finding a correlation between the b quark charge and the charges of the tracks belonging to the b-jet [71, 72]:

$$Q_{bjet} = \frac{\sum_i q_i |\vec{j}_i \cdot \vec{p}_i|^\kappa}{\sum_i |\vec{j}_i \cdot \vec{p}_i|^\kappa} \qquad (4.8)$$

where q_i and p_i are the charge and momentum of the ith track inside the jet, respectively, $\vec{j}$ is the unit vector along the jet axis and the κ is a parameter. We require tracks to pass the following criteria:

- have to be detected by Silicon Vertex detector,

- absolute value of the impact parameter is $|d_0| < 0.15$ cm,

- $p_T > 1.5$ GeV/c,

- have to be inside the jet, while the jet cone $\Delta R < 0.4$ is used,

- at least 2 tracks pass the above mentioned criteria.

Figure 4.2 shows the jet charge distribution obtained by using the equation 4.8 on the SecVtx loose tagged b jets in $t\bar{t}$ MC sample. A jet would correspond to a b if its charge is negative, and to a $\bar{b}$ if it is positive.

From the optimization by calculating the performance ϵD^2, the parameter κ was set to be equal to 0.5 [68].
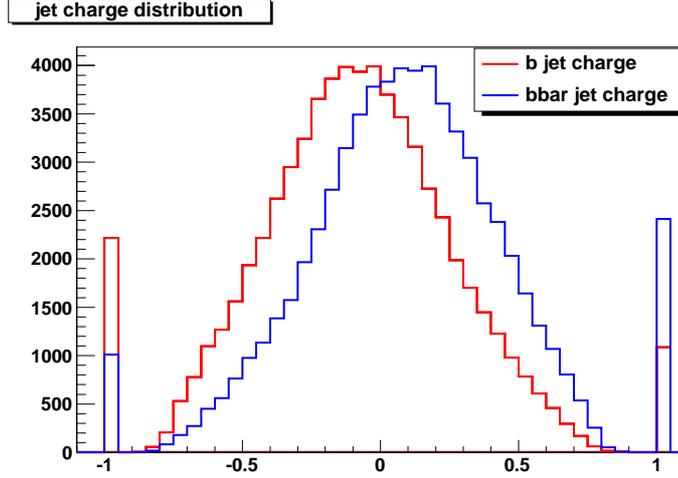
Figure 4.2: MC jet charge distribution for the loose tagged b and $\bar{b}$ jets. The parameter $\kappa = 0.5$ was used.

We did a couple of further checks or optimizations:

For the jet charge calculation we use the jet cone 0.4. Changing this cone we include or exclude some tracks, what can change the final value of the charge. We compared the difference between the jet charges associated with positive and negative leptons. As it is seen from Table 4.4, the jet cone $\Delta R = 0.4$, which is used by the CDF as the default one, gives the best result.

jet cone	$Q_- \pm \sigma_-$	$Q_+ \pm \sigma_+$	$r = \sqrt{2} \cdot \frac{\lvert Q_- - Q_+ \rvert}{\sqrt{\sigma_-^2 + \sigma_+^2}}$
0.30	0.0599 ± 0.0019	-0.0491 ± 0.0019	56.4
0.35	0.0597 ± 0.0019	-0.0494 ± 0.0019	57.5
0.40	0.0594 ± 0.0019	-0.0490 ± 0.0019	57.9
0.50	0.0588 ± 0.0019	-0.0488 ± 0.0019	57.7
0.60	0.0582 ± 0.0019	-0.0479 ± 0.0019	57.3
0.70	0.0567 ± 0.0018	-0.0465 ± 0.0018	56.5
0.80	0.0559 ± 0.0018	-0.0450 ± 0.0018	56.0

Table 4.4: Jet charge for the jets associated with positive (Q_+) and negative (Q_-) lepton.

An another check was how the jet charge depends on fraction of the energy carried by the positively or negatively charged tracks in the jet cone. We have expected that the purity of the jet charge calculation would increase with increasing value of the fraction. In the cases of the lower fractions, the most of the b-jet energy is carried by neutral particles and only a few tracks with lower p_T are used for the jet charge calculation. The disadvantage of the approach is cutting out more events. As one can see from Table 4.5 the purity is increasing with the increasing fraction of the energy carried by the charged tracks. Because of the performance distribution is flat in the region from 0.0 to 0.2, one can choose any value from this region. Our selection was to continue without this cut to keep the statistics as high as possible.

fraction	efficiency ϵ (%)	purity P (%)	ϵD^2
0.0	97.9 ± 0.04	60.8 ± 0.1	0.046 ± 0.001
0.1	96.8 ± 0.04	60.9 ± 0.1	0.046 ± 0.001
0.2	88.7 ± 0.1	61.5 ± 0.1	0.047 ± 0.001
0.3	72.8 ± 0.1	62.2 ± 0.2	0.043 ± 0.001
0.4	52.8 ± 0.1	63.3 ± 0.2	0.037 ± 0.001
0.5	33.5 ± 0.1	64.7 ± 0.3	0.029 ± 0.001
0.6	18.6 ± 0.1	66.5 ± 0.4	0.020 ± 0.001

Table 4.5: The jet charge algorithm optimization based on the fraction of jet energy carried by charged tracks.

Figure 4.3 shows the number of tracks used to the jet charge calculation. There is a small discrepancy between the data and MC distributions. The further study of this is present in Appendix A.

Since the used MC is not necessarily reliable in terms of jet fragmentation, we correct the purity for the jet charge algorithm obtained from the MC by using a dijet data sample enriched in heavy flavor.
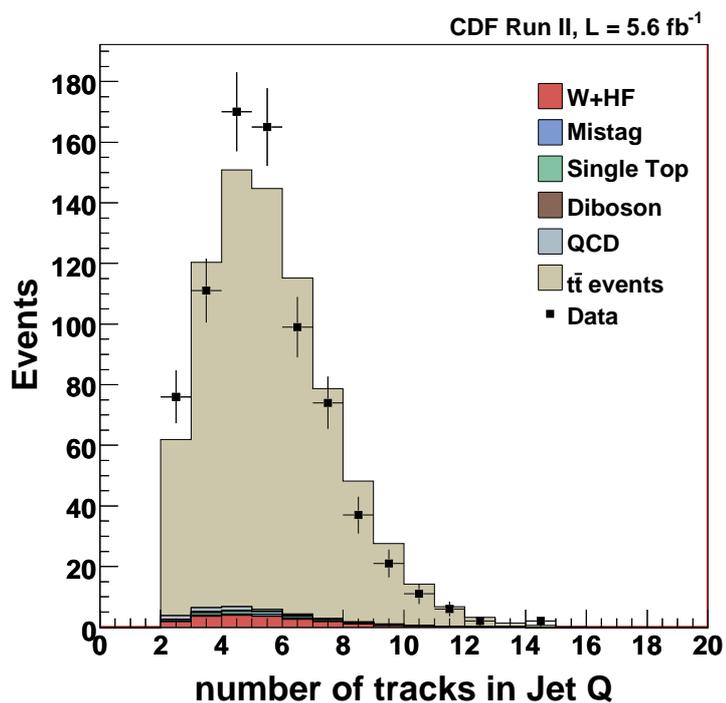
Figure 4.3: The number of tracks in the jet cone 0.4 used to the jet charge calculation. The MC distribution is normalized to the number of data events. Different background samples are described in Section 6.2.

Chapter 5

Calibration of the jet charge algorithm

The top charge measurement is most sensitive to the value and the uncertainty of the jet charge calculation's purity, P_{JQ}. It is define as the number of jets for which the algorithm makes the correct flavor assignment, over the total number of jets, where both cases, the numerator and denominator, are taking into account only jets matched to a b quarks at the parton level within the jet cone 0.4. Using the $t\bar{t}$ signal MC we found the purity 0.608 ± 0.01 (Table 4.5).

The dijet MC uses a B decay modeling and includes a minimum bias events. But due to the fact that the jet charge calculation is sensitive to a details of the fragmentation process, the purity should be obtained from data. A method that allows us to calibrate the jet charge (JQ) algorithm was developed by using a muon enriched dijet data sample, from which a sub-sample with $b\bar{b}$ pairs can be selected.

We use only pairs where it is possible to find a triggered muon inside one of jets, which comes from the semileptonic decay of the heavy flavor hadron and where the second jet travels back-to-back (in ϕ) to the first one. The jet which contains the muon is referred to as muon jet, while the second jet is referred to as away jet. Detailed events selection is described in Section 5.1.

The flavor of the away jet is determined by applying the jet charge algorithm described in Section 4.4. Knowing the charge of the muon jet from the muon charge, we can study the correlation between the jets' charges, which is present if the jets comes from $b\bar{b}$ pair. The correlation can be expressed by means of the purity, P_{obs},

defined as the number of pairs with opposite charge signs (OS) over the total number of pairs.

$$P_{obs} = \frac{N_{OS}}{N_{OS} + N_{SS}} \tag{5.1}$$

where N_{OS} and N_{SS} are the number of jet pairs with the opposite or the same charge signs, respectively.

The final result of the jet charge algorithm's calibration is presented in terms of the scale factor (SF_{JQ}), which is defined as a ratio of the purity obtained from data to that one from the MC. It allows us to convert the MC purity value to the appropriate value in data.

5.1 Event selection

For this study we used a muon enriched dijet data sample selected by the muon CMUP8 trigger, what means that the muon is required to leave a stub in the both CMU and CMX muon chamber and has to have $p_T > 8 GeV$. In the case of MC samples, we used several dijet samples which were generated by using different parton p_T thresholds (18, 40, 60 and 90 GeV/c) and the muon enriched dijet sample generated with jet $p_T > 20$ GeV and containing a muons with $p_T > 9$ GeV/c and $|\eta| < 0.6$.

From the above mentioned samples, by looking for the highest p_T muon or highest E_T jet, we select the pairs, which are composed from two jets with $E_T > 20$ GeV separated in ϕ by more than 2 radians. The muon jet is required to contain the reconstructed CMUP muon within a cone size 0.4 around its axis. Only one such pair per events is allowed. The list of all requirements is described in following:

- muon jet raw $E_T > 10$ GeV

- muon Jet $E_T > 20$ GeV

- muon track $p_T > 9$ GeV/c

- muon track $|z_0| < 60.0$ cm

- muon CMU stub $|\delta x| < 3.0$ cm

- muon CMP stub $|\delta x| < 5.0$ cm

- muon track distance to primary vertex $|z_0 - z_{vtx}| < 5.0$ cm

- Muon track isolation > 0.1

- Muon track must pass through every layer of the Silicon vertex detector

- Muon $|\eta| < 0.6$

- Away Jet $E_T > 20$ GeV

- Away Jet $|\eta| < 1.5$

- Away Jet must have at least two tracks passed the criteria for jet charge calculation

The similar selection criteria are used in the b-tag scale factor study in [73]. The only difference is in applying of a secondary vertex mass cut on the away jet. We do not apply it in order to keep the away jet as a generic one.

Due to the presence of the muon in the muon jet, its direction has to be corrected. To do so we use the equation:

$$\vec{P}_{corr} = \vec{P}_{jet} + \left(1 - \frac{2}{|\vec{p}_\mu|} \right) \vec{p}_\mu \qquad (5.2)$$

where the value of 2 GeV was chosen from [74] as the most probable energy deposition of a muon in the p_T range under consideration. No correction is applied to the muon jet E_T, since this study uses away jet E_T bins.

To increase the heavy flavor content of dijet pairs selected by the above mentioned criteria we require the away jet to be b-tagged by the loose SecVtx tagger, while the muon jet by the tight one.

5.2 Method and its correction

As we mentioned we use dijet $b\bar{b}$ pairs where we expect the correlation between the charges of muon and away jet. The purity which express this correlation is defined in the equation 5.1.

However the main issue in the analysis is to determine the real fraction of $b\bar{b}$ pairs in the events which passed the selection criteria as the muons can be also produced by

charm decays or it can be a fake muon or a light jet can be misidentified as an away b-jet.

Using the muon enriched MC sample, we divided the events into several categories depending on the result of jet-to-parton matching. The fractions of the treated categories are shown in Table 5.1.

	Cases	fraction (%) no MJ tag required	fraction (%) both jets tagged
1	MJ $= b$, AJ $= b$	77.2 ± 0.2	87.0 ± 0.3
2	MJ $= b$, AJ $= c$	2.2 ± 0.1	2.6 ± 0.1
3	MJ $= b$, AJ $= nonb, nonc$	5.3 ± 0.1	6.4 ± 0.2
4	MJ $= c$, AJ $= c$	8.8 ± 0.1	2.0 ± 0.1
5	MJ $= c$, AJ $= b$	2.2 ± 0.1	0.5 ± 0.1
6	MJ $= c$, AJ $= nonb, nonc$	3.2 ± 0.1	0.8 ± 0.1
7	MJ $= fakes$, AJ $= b/c/l$	1.0 ± 0.1	0.6 ± 0.1

Table 5.1: Classification of events in the muon enriched MC sample. The muon (MJ) and away jets (AJ) were matched to partons within a cone of 0.4 and classified accordingly in different cases. The *fakes* include those events where the reconstructed muon did not match, within a cone of 0.05, a generator level muon or those where although there is a matched muon, but the jet is not a b or c. The "no MJ tag required" column corresponds to the fraction of each case when only the away jet is tagged (loose) and the "both jets tagged" column to the case where also the muon jet is tagged (tight).

From the classification it can be seen that using the tight b-tagging requirement on the muon jet side increases the $b\bar{b}$ fraction in the selected events, while the fractions of the other cases are reduced. The special case of interest is the $c\bar{c}$, in which we also expect the the charge correlation between the muon and away jets. Due to this and the fact that the jet charge algorithm is sensitive to the charge of the charm jets, the result of P_{obs} could be biased. We do not expect the bias of the result from the other cases, since there is not expected any charge correlation between the muon and away jet. This includes the "fakes", which are underestimated in Table 5.1, since the fraction of $b\bar{b}$ was obtained on the muon enriched MC. However, we will obtain the $b\bar{b}$ fraction by fitting the dijet data.

In above we discussed the purity corrections caused by the background events, which passed the selection but were not the real $b\bar{b}$ pairs and so the charge correlation was not present (except the special case of $c\bar{c}$). But even if we have the real $b\bar{b}$ pairs some corrections are still needed. It is due to b quark semileptonic cascade decay and due to neutral B meson mixing.

If the secondary decay occurs in the muon jet according to the scheme $b \to c \to \mu$, the muon does not come from the b quark semileptonic decay, but from a c quark decay and its charge has opposite sign. What would mean that the muon and away jet should have the charges with the same sign. The same effect we expect if the b quark undergoes the mixing, i.e. due to the phenomena of the $B_0 \leftrightarrow \bar{B}_0$ oscillations it converts into $\bar{b}$ quark. Thus the fraction of $b\bar{b}$ case ($f_{b\bar{b}}$) should be corrected by the fraction of the secondary decay or B mixing cases, f_{secmix}, which contributes to the OS number of events with probability $(1 - P_{JQ})$. The case, when the mixing and secondary decay happen at the same time, contributes with the purity P_{JQ}. The case of mixing and a secondary decay like $b \to \bar{c} \to \mu$, which would contribute with $(1 - P_{JQ})$, is small but is also considered and included in the f_{secmix} calculation.

Taking into account all the corrections we can express the P_{obs} purity by the equation:

$$P_{obs} = f_{b\bar{b}}(1 - f_{secmix})P_{JQ} + f_{b\bar{b}}f_{secmix}(1 - P_{JQ}) + f_{case2,3,5-7}0.5 + f_{c\bar{c}}P_{c\bar{c}} \qquad (5.3)$$

where P_{JQ} is the purity we actually want to measure, that is, the performance of the JQ algorithm in b-jets. The fraction $f_{case2,3,5-7}$ corresponds to the background events with no jet charge correlation. The special case of $c\bar{c}$ events corresponds to the fraction $f_{c\bar{c}}$ which is very small (see Table 5.1) so we assign the $P_{c\bar{c}}$ probability also equal to 0.5. Thus the fraction of background f_{bckg} corresponds to $1 - f_{b\bar{b}}$ and the equation 5.3 can be expressed as:

$$P_{obs} = f_{b\bar{b}}(1 - f_{secmix})P_{JQ} + f_{b\bar{b}}f_{secmix}(1 - P_{JQ}) + 0.5(1 - f_{b\bar{b}}) \qquad (5.4)$$

5.2.1 $b\bar{b}$ fraction

To determine the $f_{b\bar{b}}$ fraction we used two variables, which are powerful to distinguish b quark jets from the light quark ones.

On the muon jet side we used the muon transverse momentum relative to the muon jet axis ($p_{T,rel}$), as it was done in the b-tag scale factor study [73]. We obtained the b and c jet templates from the jets matched by b or c quark on the parton level using the muon enriched MC. The template for the light or gluon jets (referred to as a light jet template) was obtained from dijet MC samples. In Figure 5.1 the templates are shown as a function of the away jet E_T. We see some dependence on away jet E_T in the heavy flavor case (b and c jets), while there is no dependence in the case of light jets. Such dependence is expected so it only confirms, that the used muon enriched MC well represents the E_T spectrum of the heavy flavor cases, for which it was tuned to.
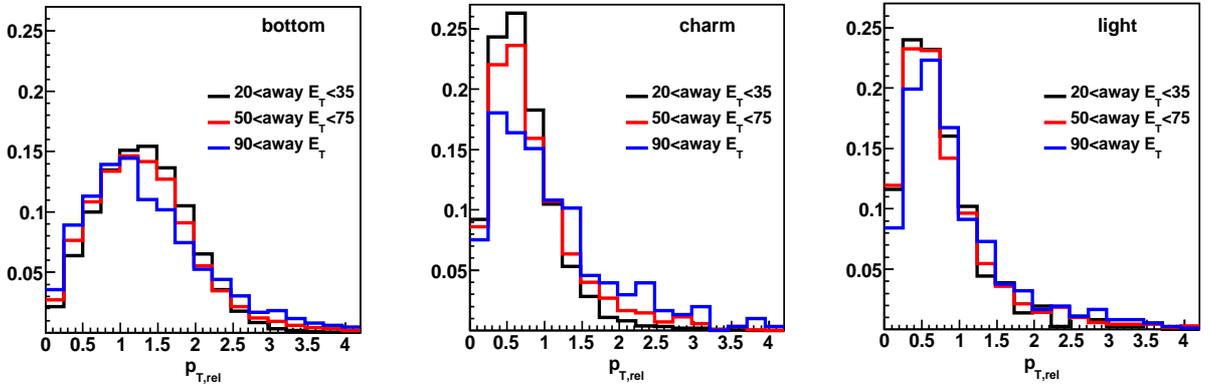
Figure 5.1: $p_{T,rel}$ templates for bottom and charm quarks obtained from the muon enriched MC and for light quarks/gluon obtained from the dijet MC for three intervals of away jet E_T.

For the away jet side, we used the secondary vertex invariant mass, M_{vtx}, calculated using the four momenta information of the jet, which originates from this vertex. The templates were obtained from the dijet MC, but to increase the statistics we did not require the muon presence in the muon jet. At the light jet template contribution we have used only the events where no heavy flavor quark on the parton level was present in both, muon and away jet. As we see in Figure 5.2 no strong dependence on the away jet E_T is observed.

As we have mentioned in Section 5.1 we have used the dijet MC samples generated with different parton (jet) p_T. Instead of reweighing the samples to obtain the correct jet E_T spectrum, we carried out the analysis in the form of the E_T function assuming that the templates are E_T independent in each away jet E_T bin.
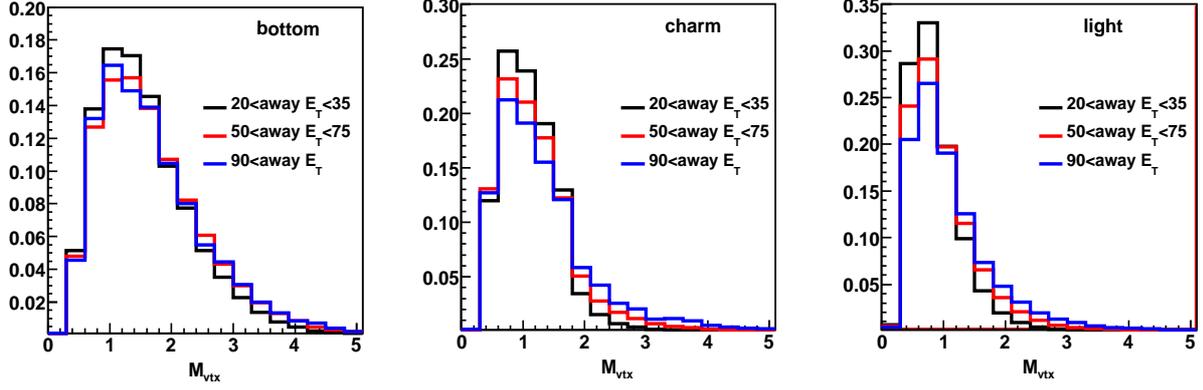
Figure 5.2: M_{vtx} templates for bottom, charm and light quarks/gluon jet obtained using the dijet MC for three intervals of away jet E_T.

Fitting procedure

Using the $p_{T,rel}$ and M_{vtx} templates, we fitted the dijet data in 9 different away jet E_T bins. To illustrate the procedure, Figure 5.3 shows the fits in one E_T bin. For the $p_{T,rel}$ case, the charm and light spectrum are similar and the fitter can not distinguish between them, so we used only b and c templates to fit the data distributions. For the M_{vtx} fit we used all 3 templates. The differences coming from the other choices of the templates which were used to the fits are applied as a systematic error (e.g. 3 templates fit of the $p_{T,rel}$, or 2 template fit of the M_{vtx}).
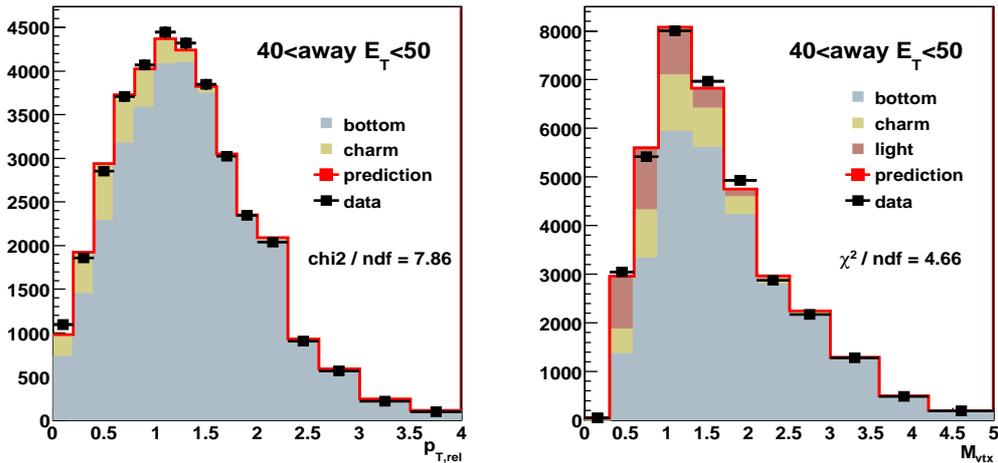
Figure 5.3: $p_{T,rel}$ and M_{vtx} fits using two and three templates, respectively, for one bin of away jet E_T.

In this particular E_T bin, the $p_{T,rel}$ fit result is $91.1 \pm 1.4\%$ what means that for

$\approx 9\%$ of the events the muon jet is not from a b jet. The M_{vtx} result is $76.9 \pm 1.2\%$. From these numbers we can obtain $f_{b\bar{b}}$ by setting the upper and lower limits. If we assume that in the muon jet case all of the 9% of non-b muon jets are paired with non-b away jets, the upper limit for the $b\bar{b}$ fraction would correspond to the M_{vtx} fit value, i.e. to 76.9%. On the other hand, if all of the 9% non-b muon jets are paired with b away jets, the lower limit for the fraction would then be found by subtracting the muon side contamination from the M_{vtx} value. The $b\bar{b}$ fraction is calculated as the average of those two limits. In this case it corresponds to:

- Upper limit = 76.9% (M_{vtx} fit result)

- Lower limit = 76.9% - (100 - 91.1)% = 68%

- $b\bar{b}$ fraction (average of the two limits) = $72.4 \pm 4.7\%$

The uncertainty of the $f_{b\bar{b}}$ is calculated in such a way that it covers the difference between the upper and lower limits.

5.2.2 f_{secmix} correction

We measured the fraction of secondary decay and mixing events in the muon enriched MC used for this analysis, which includes both effects. To cover all possible cases we defined f_{secmix} by the equation:

$$f_{secmix} = f_{secOS}(1 - f_{mix}) + (1 - f_{secOS})f_{mix} \tag{5.5}$$

where f_{mix} is the fraction of mixing events and f_{secOS} (measured only over no mixing event) corresponds to the fraction of the $b \to c \to \mu$ decays. The $b \to \bar{c} \to \mu$ decays are included in $(1 - f_{secOS})$ fraction.

To obtain the fractions f_{mix} and f_{secOS} we used the MC truth information. In forward case we checked if the muon comes from an oscillating bottom meson or baryon, while in the latter case we required the muon to originate of a charm meson or baryon. Note that both fractions are obtained on the sample of the true $b\bar{b}$ dijet events with opposite jet signs. Figure 5.4 shows the mixing and secondary decay fractions as a function of away jet E_T. The lower fraction of secondary decay events for the lower away jet E_T can be explained by applying the cut on muon p_T. The overall f_{secmix} is also shown.
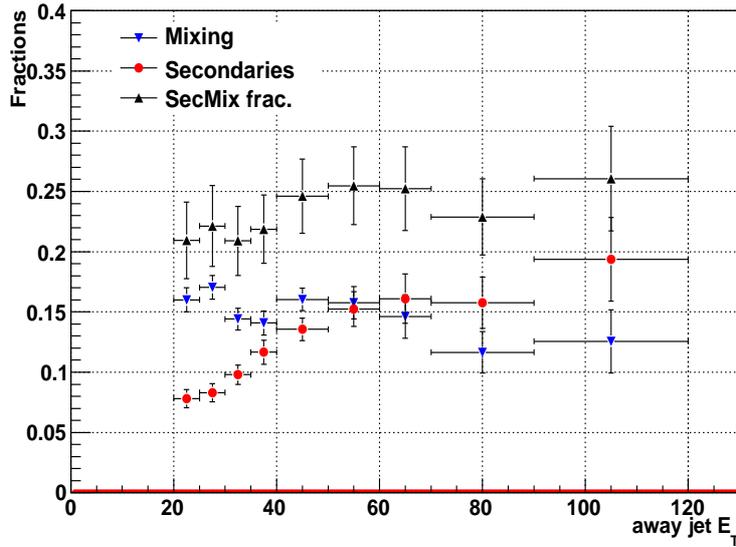
Figure 5.4: Fraction of the $b \to c \to \mu$ decays and mixing, $b \leftrightarrow \bar{b}$, as a function of away jet E_T

As the Event Generator values for B production rates and for semileptonic branching ratios are different from the values cited by the Particle Data Group (PDG), we scaled the fractions f_{mix} and f_{secOS} by the ratio between the PDG 2010 and non-bias MC values [75]. The final fractions which we used in equation 5.4 are:

$$\text{away jet } E_T < 40 \text{ GeV} \qquad f_{secmix} = (21.4 \pm 2.9)\%$$
$$\text{away jet } E_T > 40 \text{ GeV} \qquad f_{secmix} = (24.8 \pm 2.8)\%$$

5.3   Scale Factor

As we describe above, to obtain the the purity of the jet charge algorithm P_{JQ} from data, we need to determine the $f_{b\bar{b}}$. To do so we fitted the $p_{T,rel}$ data distribution by the MC templates. The results of the fits are shown in Figure 5.5 as a function of the away jet E_T. The b fraction decrease from 91% at low E_T to $\approx$88% at high E_T. This effect is caused by higher probability of finding a fake muon.

On the away jet side we used the M_{vtx} distributions. As it is shown in Figure 5.6, here the b fraction decrease almost by factor 2 (from 87% to 44%). The reason could be increasing of the mistag rate with the energy. Note that we use the loose SecVtx b-tagging for the away jet.
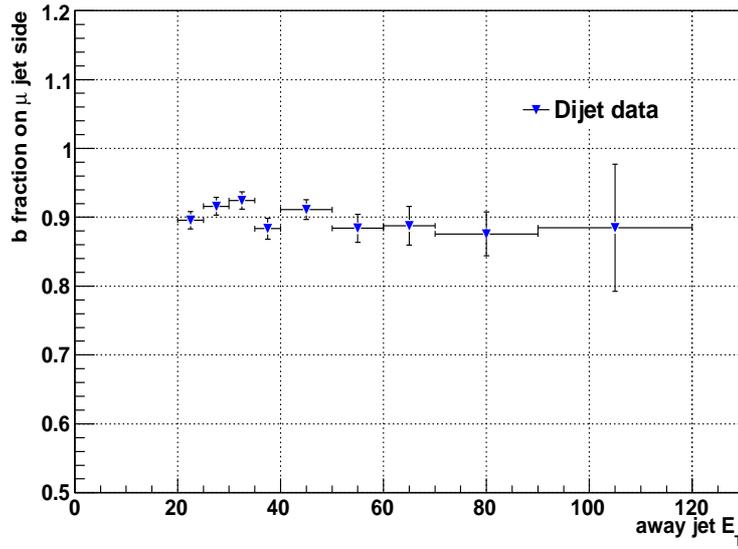
Figure 5.5: b fraction on muon side as a function of away jet E_T, obtained by fitting the $p_{T,rel}$ spectrum in a muon enriched data sample.
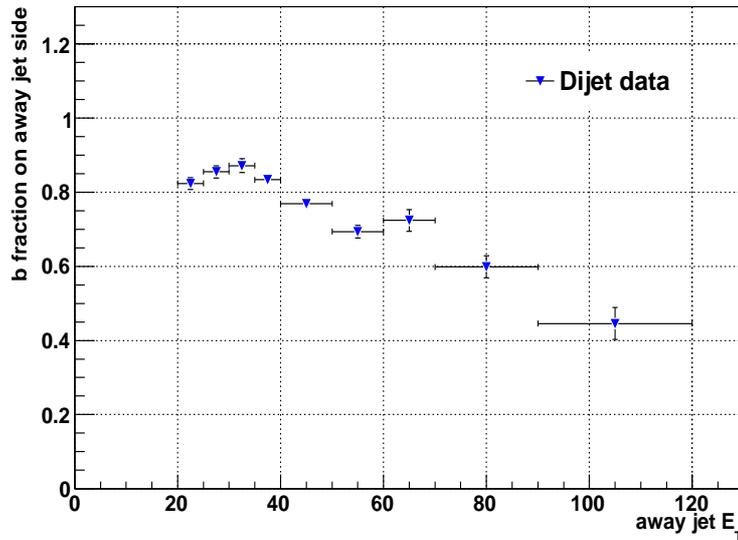
Figure 5.6: b content as a function of away jet E_T on away jets side in the muon calibration data sample, obtained by fitting its M_{vtx} spectrum.

Following the procedure from Section 5.2.1 we calculated the $f_{b\bar{b}}$ for each away jet E_T bin, as it is presented in Figure 5.7. We see the drop of the $b\bar{b}$ fraction at high away jet E_T for the dijet data sample. It is caused by the drop of the b fraction at the away jet side at high jet E_T. The MC dependence, which is also shown in this Figure, has also a decreasing tendency, but the drop is not so big.
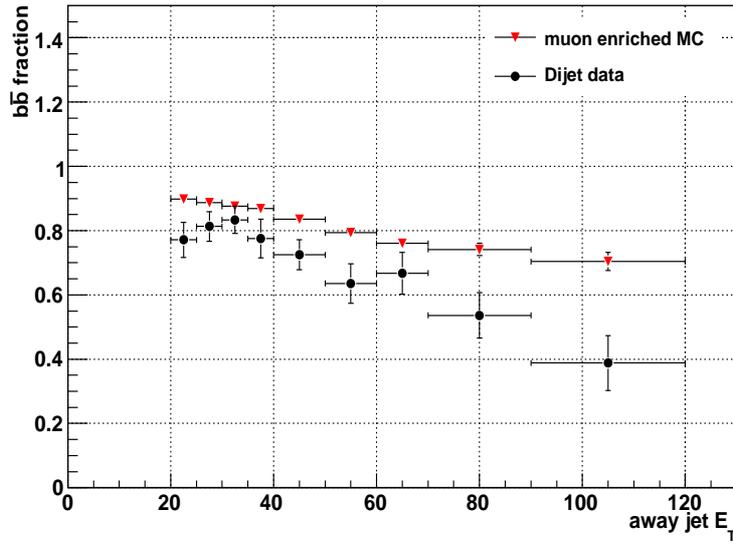
Figure 5.7: Fraction $b\bar{b}$ events in muon enriched data samples as a function of away jet E_T.

The last step to obtain P_{JQ} from the equation 5.4 is to calculate the observed purity, P_{obs}, using the dijet data. Figure 5.8 shows the dependence which is consistent with our expectations. As the $b\bar{b}$ fraction decreases with increasing away jet E_T, the background part with purity 0.5 is higher and therefore the P_{obs} is closer to 0.5.
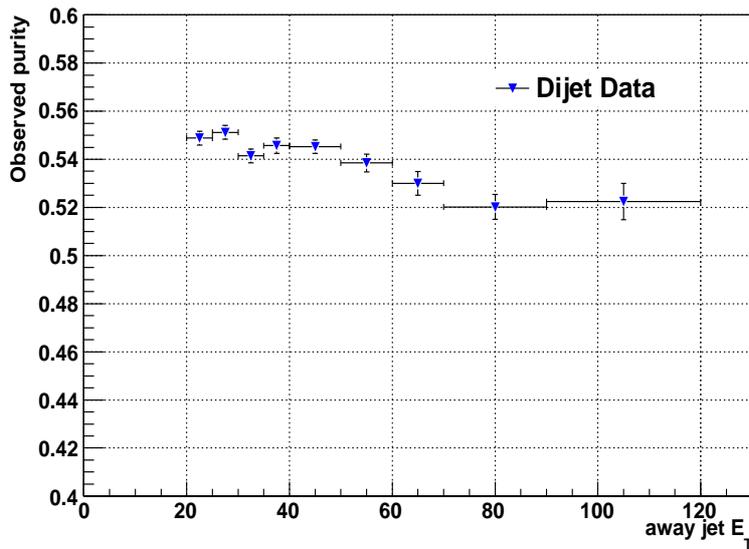
Figure 5.8: Measured purity, P_{obs} calculated as the ratio N_{OS}/N_{total} on jets using the muon enriched data sample as a function of away jet E_T.

SF final result

We present the final result as a scale factor SF_{JQ} defined as a ratio of the jet charge purity obtained from the data over the purity from MC. This enables to use the purity result in any sample, in particular for the high E_T b-jets in $t\bar{t}$ events.

The purity in MC was calculated as a weighted average of the results observed in the dijet and muon enriched MC samples, while the data purity P_{JQ} was extracted using the equation 5.4. Figure 5.9 shows the both, MC and data, distributions of the purity P_{JQ} as a function of the away jet E_T.
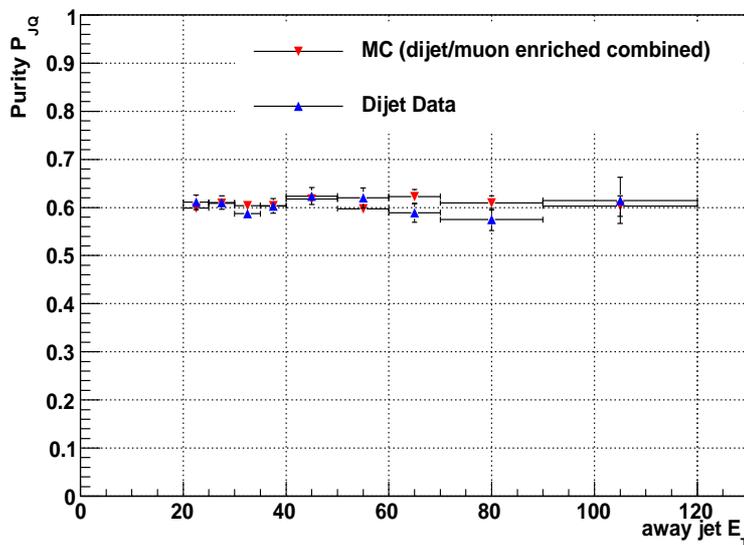
Figure 5.9: Corrected purity, P_{JQ}, as a function of away jet E_T. The red triangles correspond to the purity for b-matched jets of the used MC sample (weighed average between the purity in dijet MC and in muon enriched MC samples).

Figure 5.10 presents the result of the scale factor SF_{JQ} which was fitted by the constant function. A linear fit, used to obtain an uncertainty due to the E_T dependence, is also shown. From the constant fit one can see that:

$$SF_{JQ} = 0.99 \pm 0.01$$

5.4 Systematic uncertainties

In this section we present the systematic uncertainties which are related to the procedure used to find the b fraction on the muon and away jet sides. We also add the
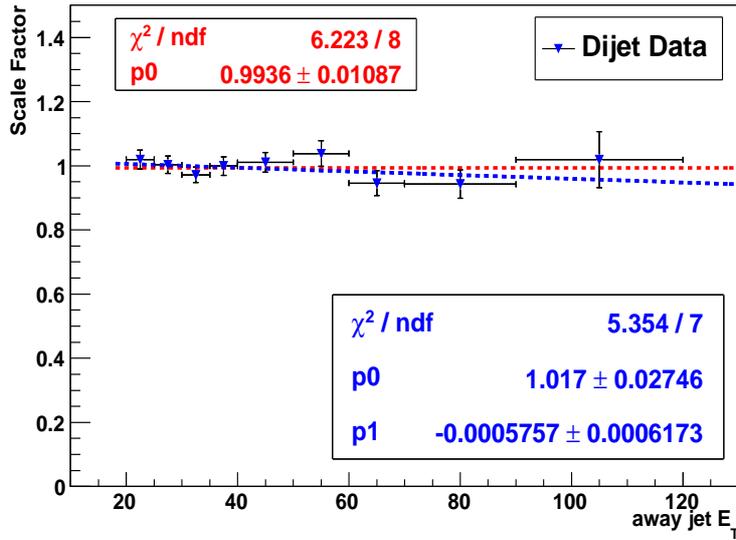
Figure 5.10: Scale factor as a function of E_T, for loose tagged jets, calculated from the ratio between the JQ purity in muon calibration data and in a weighted average between a dijet and muon enriched MC. The red line corresponds to a fit with a constant function while the blue one corresponds to a fitted line assuming a non-zero slope.

uncertainty caused by the E_T dependence.

5.4.1 $p_{T,rel}$ template tag bias

In the calculation of the SF_{JQ} we used the $p_{T,rel}$ templates obtained from the b-tagged jets, but there is a possibility of a bias due to the using of the b-tagging. Therefore we used the non-tagged jets to obtain the new b and c jet templates, which were used to determine the b-fraction on the muon jet side and finally to calculate the another scale factor value. The difference with respect to the nominal scale factor value was taken as the systematic uncertainty.

5.4.2 $p_{T,rel}$ non-b template

The fraction on the muon jet side was obtained by $p_{T,rel}$ fits using only the b and c templates. We calculated the scale factor by using the light template instead of the charm one and the difference with respect to the nominal value was added as the systematic uncertainty.

74

5.4.3 M_{vtx} template bias

The incorrect tracking efficiency in MC leads to obtaining the higher values of the secondary vertex invariant mass. By taking into account the uncertainty of the tracking efficiency, one could decrease the secondary vertex invariant mass by the 5%. We shifted the b, c and light templates by this value and obtained the scale factor. The difference with respect to the nominal value was added as the systematic uncertainty.

5.4.4 M_{vtx} fits

Extracting of the b fraction on the away jet side was done by the tree templates fits. We changed the fitting technique and instead of three templates we used only two templates with (bottom and charm or bottom and light). The shift with respect to the nominal scale factor was added as an addition systematic uncertainty.

5.4.5 E_T dependence

The scale factor was calculated by using dijet $b\bar{b}$ events, however it will be applied in $t\bar{t}$ events, where the average value of b jets E_T is higher. We assumed the scale factor to be constant, but an error which cover possible E_T dependence is added as a systematic uncertainty. The scale factor distribution was fitted by a line of non-zero slope and the systematic error was obtained by weighing the error of the linear fit by b-jet E_T distribution from $t\bar{t}$ events (see Figure 5.11).

5.4.6 Summary

Table 5.2 shows the systematic uncertainties from the sources mentioned above. The total systematic uncertainty is 0.03.

5.5 Dependences

We checked possible dependences of the SF on the number of vertices and jets' η. To be sure that the dependence does not come from the E_T dependence, we did the check in two different E_T regions - in the lower one for E_T between 20 and 35 GeV and in
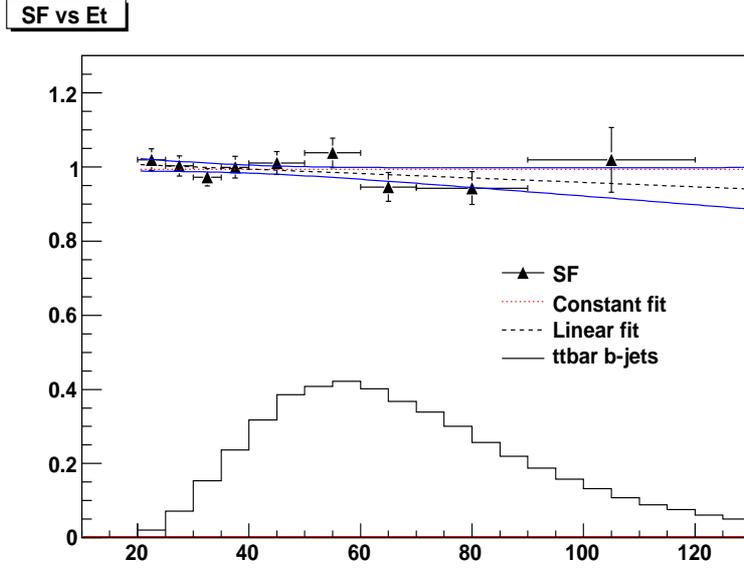
Figure 5.11: Obtaining the systematic error stemming from E_T dependence of the scale factor, SF_{JQ}. The linear fit is shown by dashed black line, blue lines present the 1σ deviation of the linear fit.

SF nominal value	0.9936 ± 0.0109	.	.
Systematic source	Scale Factor	Difference	Difference in %
$p_{T,rel}$ template tag bias	0.9885 ± 0.0099	0.0052	0.5
$p_{T,rel}$ non-b template bias	0.9897 ± 0.0101	0.0039	0.4
M_{vtx} 2 template fit	1.0116 ± 0.0120	0.0180	1.8
Track rec. ineff.	0.9718 ± 0.0096	0.0218	2.2
E_T dependence	.	0.0139	1.4
Total	.	0.0322	3.2

Table 5.2: Systematic uncertainty on the scale factor

the higher ranging from 50 to 75 GeV. As it is shown in Figures 5.12 and 5.13, no dependence was observed.

Due to the fact that one would like to use the tight SecVtx b-tagging instead of the loose one, we checked also the SF by using the tight b-tagging. The result shown in Figure 5.14 is consistent with the scale factor obtained by using the loose b-tagging. The difference is in a range of one standard deviation.

We mentioned in previous section that CDF collaboration used different tracking

scrips for data collected from April 2008 to February 2010. This could affect calculation of the jet charge. To be sure that there is no issue coming from the tracking we calculated jet charge scale factors SF_{JQ} for the data collected before and after April 2008 separately. The results, shown in Figure 5.15, are consistent within the statistics. (This check was done before obtaining the final result for all data together.)
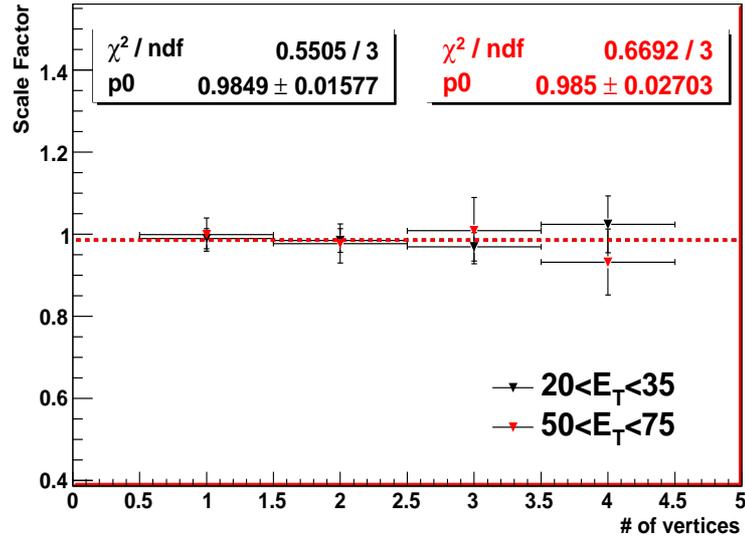
Figure 5.12: SF as a function of number of Z vertices, for two different E_T bins.
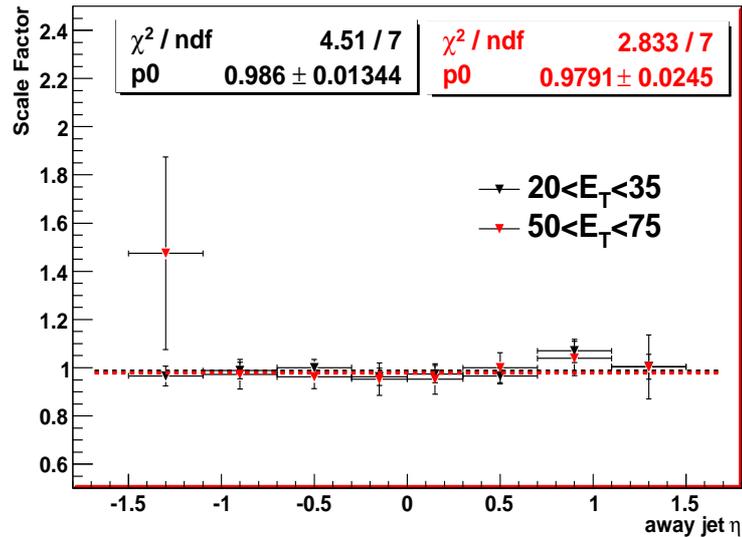
Figure 5.13: SF as a function of away jet η, for two different E_T bins.

77
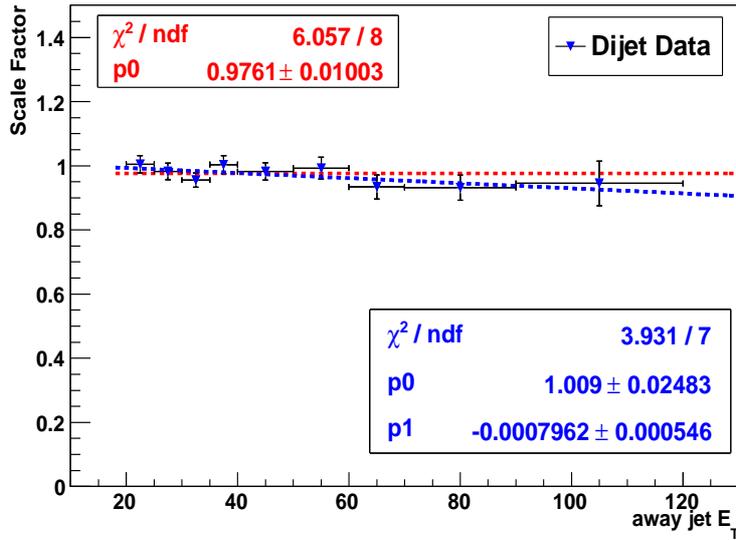
Figure 5.14: Scale factor as a function of E_T, obtained from the ratio between the JQ purity extracted from the tight tagged jets in the muon calibration data and in a weighted average MC (generic and a HF enriched samples). The red line corresponds to a fit with a constant function while the blue one was done by fitting to a line with non-zero slope.
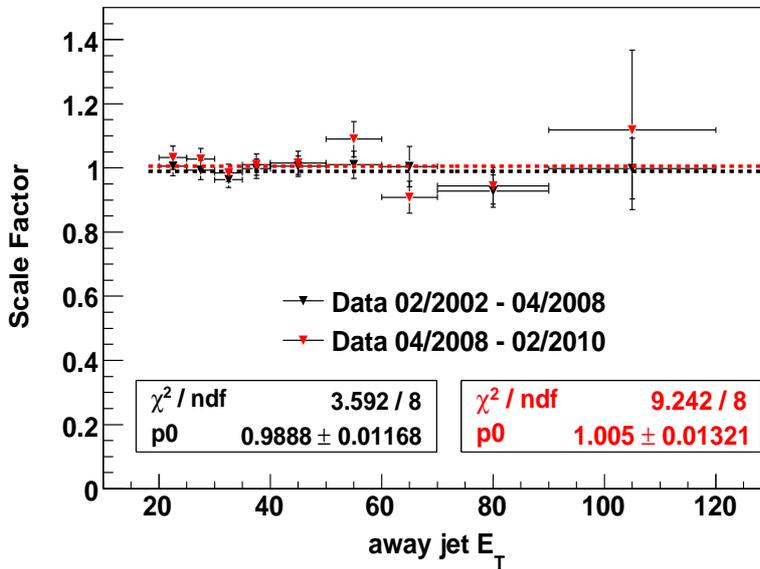
Figure 5.15: Scale factor calculated separately for data collected from February 2002 until April 2008 and data collected from April 2008 until February 2010. The dashed red (black) line corresponds to the constant fits of two data subsamples.

Chapter 6

MC expectations

This section contains the results obtained by means of MC studies. We describe contribution of the extended (non-triggered) muons and in detail we analyze the sources of the background events as well as the systematics uncertainties. The signal expectation is described at the end of the section.

6.1 Non-triggered Muons

The non-triggered muons increase our statistics by $\approx 20\%$. To select the events we use a trigger containing cuts on $\not{E}_T^{trig}$ and jets (see Section 4.1). However in the MC sample we do not have such a trigger. To correct the proportion of non-triggered muons we use the so-called trigger turn on curve, which sets a weight to each event. This weight depends on the $\not{E}_T^{trig}$ as it is shown in Figure 6.1.

As we use the fitter for the lepton - b-jet pairing and jet charge algorithm to determination the flavor of b-jets, we check on MC a possible dependence of the purities on the lepton type. From Table 6.1, one can see that there is no dependence of the jet charge purity on the lepton type, but the pairing purity of the non-triggered muons is slightly higher with respect to the other lepton types.

This is caused by the applying the *trigger turn on curve*. We found that the pairing purity depends on $\not{E}_T$ as it is shown in Table 6.2. This effect we see in all lepton types. The trigger turn on curve gives the higher weighs to the events with higher $\not{E}_T^{trig}$ which differs from $\not{E}_T$ only by the applied corrections. To obtain the $\not{E}_T$ we apply further corrections on $\not{E}_T^{trig}$.
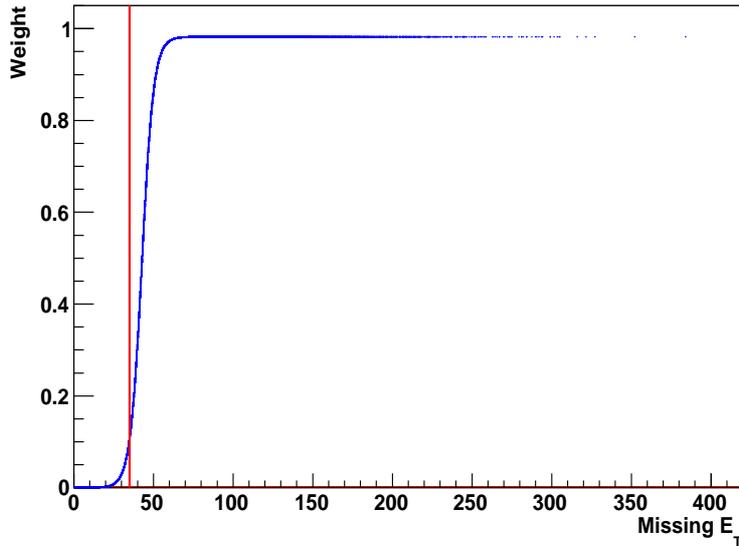
Figure 6.1: Trigger turn on curve used in MC to correct proportion of extended muons (blue line). Red line presents the trigger cut on $\not{E}_T^{trig}$.

Lepton type	Pairing purity (%)	jet charge purity (%)
Electrons	83.1 ± 0.2	60.7 ± 0.2
CMUP muons	83.3 ± 0.2	60.8 ± 0.3
CMX muons	82.8 ± 0.3	60.9 ± 0.4
Non-triggered muons	84.5 ± 0.2	61.0 ± 0.3

Table 6.1: Table of pairing and jet charge purity for different type of leptons.

Lepton type	$\not{E}_T < 50$ GeV	$\not{E}_T > 50$ GeV
Non-triggered muons no TTOC	81.5 ± 0.3	84.5 ± 0.3

Table 6.2: Table of pairing purity obtained for the non-triggered muons events without applying the trigger turn on curve (TTOC) for different $\not{E}_T$ regions.

Not applying the *trigger turn on curve* on the loose muons events in MC leads to the pairing purity $(83.0 \pm 0.2)\%$, what is consistent with other lepton types. If we combine all lepton types together, the difference in the total purity calculated with or without applying trigger turn on curve on loose muons events is only 0.3%. This has no impact to our final results.

6.2 Background

There are several sources of background in the lepton+jet channel. The main source is the QCD production of W boson plus multijets. If two jets comes from the heavy flavor quarks (b, c) we call this background as *W+heavy flavor*. In the case when all jets comes from light quarks (u, d, s) the source is named *Mistag*. The second biggest part of background is formed by QCD events without W boson and the other parts comes from single top and diboson events.

As we mentioned in Section 4.1, we use only events with at least two b jets, what suppresses the amount of background to $\approx 15\%$.

6.2.1 W+heavy flavor

This part of background is composed mainly from the W+gluon sample, where gluon splits to the $b\bar{b}$ or $c\bar{c}$ pair. The other source is the W+c subsample. The final contribution is obtained by weighing the particular inputs from these samples by the cross-section predictions. Due to the presence of heavy flavor jets and leptonically decaying W boson, this background passes the selection criteria easier than the events from other sources of background.

6.2.2 QCD

The QCD background contains the non-W multijets events that can contain $b\bar{b}$ or $c\bar{c}$ jet pairs. In this case the needed lepton can come from a jet that was mis-identified as a lepton and another possibility is a lepton coming from the semileptonic decays. To estimate this type of background we used data driven method. We ran over the special data sample suggested to treat the fake electrons, where all selection cuts were applied except of the lepton selection ones. The lepton candidate did not need to fulfill two of the identification criteria required for isolated high p_T lepton (see Section 4.1).

6.2.3 Mistag

The mistag events are produced in the same way as the W+heavy flavor events, but in this case the W boson is accompanied by jets from light quarks and gluons, e.g. a radiated gluon can split into light quarks. If we mis-identify these light jets as b jets,

event can pass our selection criteria. But in such case, estimation of the uncertainty is done in a different way than in the case of the W+heavy flavor, where the b tagged jets comes from the real b jets or c jets, which are supposed to pass the tagging criteria. That's the reason why we separated the samples.

6.2.4 Single Top

To treat this part of background we used two MC samples generated by the t-channel and s-channel single top production. We combined them by weighing them by their predicted values.

6.2.5 Diboson

Processes considered for this type of background are WW, ZZ and WZ events. The WW events can pass selection if one W boson decays leptonically, other one decays hadronically and there is a splitting gluon which was radiated in initial or final state. In the case of ZZ events one Z boson should decay hadronically to $b\bar{b}$ or $c\bar{c}$ pair and the second Z decays leptonically. Similarly in the case of WZ, the Z boson can decay hadronically, while the W boson should decay leptonically.

As in the other background cases considered, we weighed the samples by their predictions to obtain the final results.

6.2.6 Background Expectations

To predict the number of events for each type of background, we used the same method as was used by the top production cross-section measurement in L+J channel [76].

However we use some special cuts in our analysis like $\chi^2 < 9$ for the lepton - b-jet pairing and also we require at least two "good" tracks in jet cone for the jet charge calculation. Thus we apply the pairing efficiency to the predicted number of events. As we use two b-jet pairs per event, we multiply by two the number of events, which passed pairing cut, to obtain number of pairs. The jet charge efficiency is applied to number of pairs. In Table 6.3 you can found the efficiencies and corresponding values of the expected number of events/pairs for the each type of background.

Background	Prediction	Pairing efficiency	# of events after pairing	jet charge efficiency	N_b pairs
W+HF	66.27 ± 21.82	0.15 ± 0.00	10.4 ± 3.31	0.97 ± 0.00	19.47 ± 6.43
QCD fakes	17.97 ± 13.53	0.17 ± 0.08	3.05 ± 2.71	0.88 ± 0.12	5.35 ± 4.80
Diboson	4.67 ± 0.70	0.22 ± 0.02	1.01 ± 0.18	0.97 ± 0.01	1.96 ± 0.35
Mistag	9.68 ± 2.57	0.15 ± 0.02	1.45 ± 0.43	0.96 ± 0.02	2.79 ± 0.82
Single top	10.62 ± 1.28	0.21 ± 0.00	2.26 ± 0.28	0.97 ± 0.00	4.40 ± 0.54
Total	109.2 ± 25.9		17.8 ± 4.3		34.0 ± 8.1

Table 6.3: Table of background expectations, together with their measured efficiencies.

Our interest in the case of the background samples is to find out a possible correlation between the signal lepton charge and charge of the corresponding b-jet. One can express this correlation by the background purity defined as:

$$p_b = \frac{N_+}{N_+ + N_-} \tag{6.1}$$

where N_+ is the number of W^+b (or $W^-\bar{b}$) pairs and N_- is the number of W^-b (or $W^+\bar{b}$) pairs. In other words, one can say that it is the number of the SM like pairs (N_+) and the number of exotic model like pairs (N_-).

If there is no correlation between the lepton and b-jet charges, then $N_+ = N_-$ and the purity is equal to 0.5. We call such background also the symmetric background. If any correlation is present, the purity is not 0.5 and we say that the background is asymmetric.

We do not expect any correlation in background samples except two of them: single top and QCD background. In former case the real top quark was produced so we could see the asymmetry. In the latter case the asymmetry can comes from the $b\bar{b}$ events where the lepton from semileptonic decay is correlated to the charge of the b jet, from which it originates.

In Table 6.4 we present the background purity p_b for each type of background. The second column shows the measured purity, which we obtained by running over the MC samples, while column tree shows the used purity. In the cases when the measured purity is symmetric within the uncertainty and we do not expect any correlation, we use purity 0.5 ± 0.0 instead of the measured one. For cases where we expect some

correlation we used the measured purity. In this table the N_+ and N_- numbers for each type of background are shown.

Background	"meaured" purity	used purity	N_+ pairs	N_- pairs
W+HF	0.48 ± 0.01	0.5 ± 0.0	9.73 ± 3.21	9.73 ± 3.21
QCD fakes	0.48 ± 0.06		2.57 ± 2.33	2.78 ± 2.52
Diboson	0.50 ± 0.03	0.5 ± 0.0	0.98 ± 0.17	0.98 ± 0.17
Mistag	0.56 ± 0.04	0.5 ± 0.0	1.39 ± 0.41	1.39 ± 0.41
Single top	0.51 ± 0.01		2.25 ± 0.28	2.15 ± 0.27
Total	0.498 ± 0.010		16.92 ± 4.01	17.04 ± 4.13

Table 6.4: Table of background purities, together with the number of the SM (XM) like pairs N_+ (N_-).

6.3 Systematics

The predicted number of the signal and background expectations obtained from the cross-section analysis [76] already contains the systematic uncertainties. However the MC modeling of the geometrical and kinematic acceptance includes the effects of parton distribution functions (PDFs), initial and final state radiation (ISR and FSR) and jet energy scale which are sources of the systematic uncertainty. In addition uncertainty can comes from the choice of the generator (PYTHIA or HERWIG) or top mass used for generating the MC sample. All of these sources affect the purity and efficiency of lepton - b-jet pairing and also jet charge calculation efficiency and purity. The pairing can be also affected by the top mass which we used for generating of MC samples.

6.3.1 Jet Energy scale

The jet energy scale is used to correct the energy of jets, however it is also obtained with some uncertainty. We modify the value of the jet energy scale by its uncertainty of $\pm 1\sigma$ and correct the jet energy by these modified jet energy scale values. The average of the percentage difference between the shifted and nominal samples is taken as the systematic uncertainty.

6.3.2 Initial/final state radiation

This systematic uncertainty cover the cases when the jets picked up by our analysis originates from the gluon radiated before or after the hard collision. That corresponds to the initial (ISR) or final state radiation (FSR), respectively. Due to the fact that the amount of ISR/FSR in $t\bar{t}$ events is not well understood, we study the MC signal samples generated with various amount of these effects. To compare with nominal sample we use samples with more and less amount of ISR/FSR. The average percentage difference with respect to the nominal sample is considered as the systematic uncertainty.

6.3.3 Parton distribution function

Instead of generating many MC samples for each parton distribution function (PDF) set, we reweigh the nominal MC sample generated with the CTEQ PDF set [77]. For each event we look at the parton level information to figure out which of the partons interacted and what was the tranfer momentum q^2. Then for each PDF set we calculate the probability of such interaction which is later used as weight. We use the following PDF sets:

- CTEQ5L - the default set used for the nominal MC

- MRST72 - MRST PDF set [78] with the same value of strong interaction constant α_S as is used in CTEQ5L

- MRST75 - MRST PDF set with different value of α_S with respect to nominal CTEQ5L set

- 40 CTEQ PDF sets obtained by variating each of 20 CTEQ PDF parameters by $\pm 1.64\sigma$, what corresponds to the inclusion 90% of possible cases.

The total PDF systematic error is obtained in the following way:

We calculate the contribution of α_S uncertainty to the PDF systematic error by comparing the MRST72 and MRST75 PDF sets. Then we compare the uncertainty from 20 eigenvectors with the difference between the MRST72 and the nominal CTEQ5L sets. The greater value is then added to the α_S systematic uncertainty contribution - the sum is taken as the total systematic uncertainty caused by PDF.

6.3.4 Monte Carlo generator

Our nominal $t\bar{t}$ MC sample is generated by **PYTHIA**. We run over the **HERWIG** $t\bar{t}$ MC sample to compare the efficiencies and purities of our pairing and jet charge algorithms. The percentage difference is considered as the corresponding systematic uncertainty. However due to the fact that we calibrate the jet charge algorithm by using data, no MC generator systematic is assigned to jet charge purity.

6.3.5 Top Mass

The top mass of 172.5 GeV/c^2 was used as the input for generating our nominal MC sample. For kinematic reconstruction of the events we use the fitter (see Section 4.3) with top mass constrain to the same value (172.5 GeV/c^2). To treat this source of systematic, we run over other two MC samples generated with top mass of 170 GeV/c^2 and 175 GeV/c^2, while the fitter uses top mass constrained to 172.5 GeV/c^2 in both cases. The average percentage difference with respect to the nominal sample is considered as the corresponding systematic uncertainty.

6.3.6 b tagging

The errors on the b tagging efficiencies are included in the prediction numbers that are taken from the cross-section analysis. However the fraction of the b-tagged jets, that originate from c or light quarks is generally lower in MC than in data. To correct this discrepancy we obtain the new MC sample with a changed b-tagging method described in the following:

1. Each jet in the event has some probability to be tagged as b-jet (mistag probability). This probability depend on the jet properties. We increase the number of wrongly b-tagged jets by comparing a random number by the mistag probability. If the random number is lower than mistag probability we assign the jet as b-tagged. This is done only for jets, which did not pass SecVtx loose b-tagging.

2. For the SecVtx loose b-tagged jets (denote them as heavy flavor jets) we use another approach. As a probability we take the b-tagging scale factor [73], which is obtained by comparing b-tagging efficiency in data with that in MC. For the loose SecVtx tagging it is equal to $(98.1\pm6.9)\%$. If a random number is lower than

the b-tagging scale factor, we assign heavy flavor jet as b-tagged, in other case it is not assigned as b-tagged. This actually decrease the number for correctly b-tagged jets.

This new MC sample better describes the data and by comparing with the nominal MC we calculate the scale factor SF_{nonb} which corresponds to the non-b fraction f_{nonb} - the ratio of double tagged events with one or more b-tagged jets that do not originate from b quark. To obtain the uncertainty of the SF_{nonb} we variate the b-tag scale factor and mistag probabilities by their uncertainties. The final value of this non-b scale factor is $SF_{nonb} = 1.01 \pm 0.03$.

6.3.7   Final systematic error

The combined uncertainty on the efficiencies and purities are calculated by adding each individual uncertainty in quadrature. The b-tagging uncertainty is not included here, but we use the SF_{nonb} later in the analysis as is described in Section 6.4.

In Table 6.5 we show the systematic uncertainties on the pairing efficiency and purity and on the JetQ selection efficiency and purity.

Systematics (in %)	pairing ϵ	jet charge ϵ	pairing purity	jet charge purity
Jet Energy Scale	0.2	0.04	0.1	0.1
ISR/FSR	0.5	0.1	0.2	0.2
MC generator	0.2	0.1	0.1	(0.7)
top mass	0.4	0.2	0.9	0.5
PDF	0.7	0.02	0.1	0.02
Total	1.0	0.3	1.0	0.6

Table 6.5: Summary of systematic uncertainties (in %). The (0.7) number is given as information but not used in the total uncertainty since the JetQ purity is calibrated in data and so the scale factor uncertainty already includes the effect of different hadronisation models.

6.4 Signal expectations

We obtain the expected number of pairs in the same way as in the background case. The same method as was used by the top production cross-section measurement in L+J channel [76] is used to calculate the predicted number of events. After applying the pairing and jet charge calculation efficiencies, we express the expected number of signal pairs N_s (see Table 6.6).

Signal prediction	Pairing efficiency	# of events after pairing	jet charge efficiency	N_s pairs
671.3 ± 110.8	$0.532^{\pm 0.001(stat)}_{\pm 0.005(syst)}$	357.1 ± 59.1	$0.979^{\pm 0.000(stat)}_{\pm 0.002(syst)}$	669.6 ± 115.7

Table 6.6: Table of signal expectations, together with the measured efficiencies.

The calculation of the signal purity is more complicated as it is in case of the background purity. For the signal we have defined the pairing purity p_{pair} and jet charge purity p_{JQ} on the jets, which were matched with b quark at the parton level. To obtain the total purity we need to combine them, and to take into account that the sample also contains the mistag b-jets. The fraction of such cases is expressed as f_{nonb} and needs to be corrected by the non-b scale factor SF_{nonb} as we mentioned in Section 6.3.6. The purity of these "non-b" events, p_{nonb}, was obtained to be 0.50 ± 0.01, what confirms our expectations that there is no correlation between lepton and b-jet charges.

Taking into account all the above effects and also the fact that the jet charge scale factor, SF_{JQ}, is need to be apply to jet charge purity obtained from MC (as is discussed in Section 5), we can express the total signal purity, p_s, by the following expression:

$$p_s = f_{nonb} \cdot SF_{nonb} \cdot p_{nonb} +$$
$$(1 - f_{nonb} \cdot SF_{nonb})(p_{pair} \cdot p_{JQ} \cdot SF_{JQ} + (1 - p_{pair})(1 - p_{JQ} \cdot SF_{JQ})) \quad (6.2)$$

The values of all variables used in this equation are summarized in Table 6.7. The listed systematic uncertainties come from the propagation of the systematic uncertainties quoted in Table 6.5.

Using the signal purity, p_s, one can also express the expected number of SM like (N_+) or XM like (N_-) pairs for the signal MC. The corresponding values are shown in Table 6.8.

f_{nonb}	0.076 ± 0.001
SF_{nonb}	1.01 ± 0.03
p_{nonb}	0.5 ± 0.01
p_{pair}	$0.833 \pm 0.001(stat) \pm 0.008(syst)$
p_{JQ}	$0.608 \pm 0.001(stat) \pm 0.003(syst)$
SF_{JQ}	$0.99 \pm 0.01(stat) \pm 0.03(syst)$

Table 6.7: Elements needed to compute the combined purity, the description is in the text.

signal pairs N_s	signal purity p_s	N_+ pairs	N_- pairs
669.6 ± 115.7	$0.562^{\pm 0.004(stat)}_{\pm 0.011(syst)}$	393.5 ± 65.6	306.1 ± 51.3

Table 6.8: Table of the expected number of the SM (XM) like pairs N_+ (N_-) for the signal, together with the signal purity, p_s.

.

Chapter 7

Statistical treatment

In our analysis we want to decide if the measured data prefers the SM or exotic -4/3 quark hypothesis. To do so, we calculate the number of SM-like, N_+, and XM-like, N_-, pairs (as it is defined in Section 6.2.6) for data. The obtained values can be compared with the SM expectations and we can express the degree of evidence in favor of the SM over the XM or vice-versa.

However we need to use a parameter of interest which has different probability density functions for the SM hypothesis and the XM one. We choose the fraction of SM events, f_+, defined by the equation:

$$f_+ = \frac{N_+}{N_+ + N_-} \tag{7.1}$$

In the ideal case the f_+ should be equal to 1 in the case of the SM and equal to 0 in the case of the XM. Due to uncertainties of the pairing and jet charge calculation it is not always true, but in MC the f_+ value is still expected to be close to 1.

To compare our best estimate of f_+ (derived from the limited statistics of data sample) with that of our (MC) expectations we can use the frequentest or Bayesian approach. We express our results using both of them.

7.1 Profile likelihood

The basic idea of the profile likelihood is straightforward. Let us assume we have a probability model for our data which depends on k parameters $\pi = (\pi_1, ..., \pi_k)$ of interest to the researcher but also on additional nuisance parameters $\theta = (\theta_1, .., \theta_l)$.

If we denote the probability density function by $f(x|\pi, \theta)$ and we have independent observations $X = (X_1, .., X_n)$, then the full likelihood function is given by [79]:

$$L(\pi, \theta|X) = \prod_{i=0}^{n} f(X_i|\pi, \theta) \tag{7.2}$$

A standard technique for constructing confidence intervals is to find a corresponding hypothesis test. Here the hypothesis test is defined as H_0: $\pi = \pi_0$ versus H_1: $\pi \neq \pi_0$. The test can be based on the likelihood ratio given by:

$$\lambda(\pi_0|X) = \frac{max\{L(\pi_0, \theta|X); \theta\}}{max\{L(\pi, \theta|X); \pi, \theta\}} \tag{7.3}$$

where the maximum in the denominator is found over the full parameter space (we go through the all hypotheses as wall as through the whole nuisance parameter space), while the maximum in the numerator is found only over the subspace with $\pi = \pi_0$ [79].

Notice that λ is a function of π_0 (and the data) only, but does not depend on the nuisance parameters θ. In the context of nuisance parameters the function λ is also called the profile likelihood. One of the standard results known from statistics is that $-2log\lambda$ converges in distribution to a χ^2 random variable with k degrees of freedom (where k is number of parameters of interests) [79].

In the top charge analysis we define only one parameter of interest $\pi = f_+$ (see equation 7.1) and we use the following four nuisance parameters:

- N_s - the number expected signal events, which passed both - the pairing and jet charge calculation criteria

- N_b - the number expected background events, which passed both - the pairing and jet charge calculation criteria

- p_s - expected signal purity defined by the equation 6.2

- p_b - the expected background charge asymmetry, obtained by combining the all types of backgrounds inputs.

Using this we can express our likelihood by five terms:

$$L = L_s.L_{N_b}.L_{N_s}.L_{p_s}.L_{p_b} \tag{7.4}$$

where L_s is the Poisson distributed signal part, the next four Gaussian distributed terms are related to the uncertainties in number of signal (L_{N_s}) or background (L_{N_b}) events, respectively and purities for the signal (L_{p_s}) and background (L_{p_b}).

91

The signal term is defined:

$$L_s(x^+, x^-) = \frac{< N_+ >^{x^+} e^{(-<N_+>)}}{x^+!} \frac{< N_- >^{x^-} e^{(-<N_->)}}{x^-!} \tag{7.5}$$

where x^+ and x^- are the number of the SM-like and XM-like events, respectively, obtained from data. The $< N_+ >$ or $< N_- >$ are the expected mean values of Poisson distributions for N_+ or N_-, respectively and can be expressed (through the nuisance parameters) as:

$$< N_+ >= p_s N_s f_+ + (1 - p_s) N_s (1 - f_+) + p_b N_b \tag{7.6}$$

$$< N_- >= (1 - p_s) N_s f_+ + p_s N_s (1 - f_+) + (1 - p_b) N_b \tag{7.7}$$

The background uncertainty term is defined as:

$$L_{N_b}(y_b) = \frac{1}{\sigma_{N_b} \sqrt{2\pi}} e^{-\frac{(y_b - N_b)^2}{2\sigma_{N_b}^2}} \tag{7.8}$$

where N_b and σ_{N_b} are the number of background and its uncertainty and y_b is the random likelihood variable.

The signal uncertainty term is defined as:

$$L_{N_s}(y_s) = \frac{1}{\sigma_{N_s} \sqrt{2\pi}} e^{-\frac{(y_s - N_s)^2}{2\sigma_{N_s}^2}} \tag{7.9}$$

where N_s and σ_{N_s} are the number of background and its uncertainty and y_s is the random likelihood variable.

The signal purity term is defined as:

$$L_{p_s}(z_{p_s}) = \frac{1}{\sigma_{p_s} \sqrt{2\pi}} e^{-\frac{(z_{p_s} - p_s)^2}{2\sigma_{p_s}^2}} \tag{7.10}$$

where p_s and σ_{p_s} are the background asymmetry and its uncertainty and z_{p_s} is the random likelihood variable.

The last - background purity term is defined as:

$$L_{p_b}(z_{p_b}) = \frac{1}{\sigma_{p_b} \sqrt{2\pi}} e^{-\frac{(z_{p_b} - p_s)^2}{2\sigma_{p_b}^2}} \tag{7.11}$$

where p_b and σ_{p_b} are the background asymmetry and its uncertainty and z_{p_b} is the random likelihood variable.

To obtain the f_+ value corresponding to the observed data (x^+, x^-), we require the partial derivatives of the total likelihood with respect to each nuisance parameter to

be equal to 0. That leads to the system of nonlinear equation, which can not be solved analytically, so we use the MINUIT fitter to minimize the $-2log\lambda$ for the each value of f_+ from the region ([-1,2]) covering all possible values of f_+. The results is expressed as the $-2log\lambda$ function dependent only on f_+.

For the generating the pseudo-experiments (PE), based on MC prediction, we need to simulate, the x^+ (x^-) values. That is done by drawing a random numbers from corresponding Poisson distribution with the mean $< N_+ >$ ($< N_- >$). The $< N_+ >$ and $< N_- >$ values are obtained using Eq. 7.6 and 7.7 by fixing the f_+. For the other variables from the likelihood expression we draw a random number from the corresponding Gaussian distributions.

If we repeat the PEs 1,000,000 times and in each case we pick up the f_+ value corresponding to the likelihood minimum, we get the distribution (curve) of the f_+ values. Figure 7.1 shows two f_+ curves - in red if we assume that the SM hypothesis (H_0) is true and in black if we assume that XM hypothesis is true.
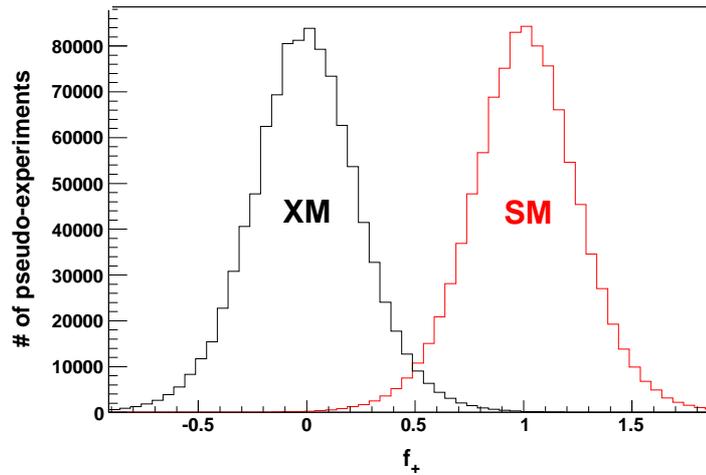
Figure 7.1: f_+ distributions for the SM and XM hypothesis obtained from 1,000,000 pseudo-experiment.

Hypotheses test

If we want to formulate the statement about the compatibility between the data and the various hypotheses in therms of a decision to accept or reject a given null hypothesis H_0, we need to start with introducing the test statistics X and the critical (rejection) region for X. If the value of X actually observed is in the critical region, we reject

the hypothesis H_0, otherwise, H_0 is accepted. The critical region is chosen such that the probability for X to be observed there, under assumption if the null hypothesis, is some value α, called the significance level of the test:

$$\alpha = \int_{-\infty}^{X_{cut}} f(X|H_0)dX \tag{7.12}$$

where $f(X|H_0)$ is a probability density and X_{cut} is a boundary which separates the critical region from the rest of the region. One would then accept the hypothesis H_0 if the value of X is higher than X_{cut}. There is thus a probability of α to reject H_0 if H_0 is true. This is called also the Type I error [80].

Type II error takes place if the hypothesis H_0 is accepted but the true hypothesis is H_1. The probability for this is:

$$\beta = \int_{X_{cut}}^{\infty} f(X|H_1)dX \tag{7.13}$$

where $1 - \beta$ is called the power of the test to discriminate against the alternative hypothesis H_1 [80].

p-value

P-value is one of the basic terms, which we use in statistical significance testing. We choose f_+ as our test statistics and define two p-values, one p_{SM} under the SM like distribution of f_+, the second one p_{XM} under the XM like distribution of f_+.

The p-value p_{SM} expresses the probability of obtaining a test statistics value at least as extreme as the one observed in data, provided that the null hypothesis (SM) is true:

$$p_{SM} = \int_{-\infty}^{f_+^{data}} f(X|H_0)dX \tag{7.14}$$

This definition is the similar to that for α, but in this case for the upper boundary of the integral, we use the measured f_+ value instead of a priori chosen one.

The p-value p_{XM} expresses the probability of obtaining a test statistic as extreme as the one observed, provided that the alternative hypothesis is right.

$$p_{XM} = \int_{f_+^{data}}^{\infty} f(X|H_1)dX \tag{7.15}$$

We compare the p-value with a priori defined Type I error α. If the p-value is lower

that α, we reject the hypothesis, otherwise we accept it. The quantity $1 - \alpha$ is reffered as the confidence level.

In our analysis we use two different Type I errors - one α_0 for the null hypothesis (SM) and the other α_1 for the H_1 hypothesis (XM):

- $\alpha_0 = 2.87 \times 10^{-7}$ or 1.3×10^{-3}, what corresponds to 5 or 3 sigma, respectively.

- $\alpha_1 = 0.05$

The a-priori values of the Type I errors were set based on the standard values utilized in high energy physics. In order to claim evidence or observation of a non SM behavior we use 3 or 5 sigma, respectively, while for the exclusion of a new physics hypothesis, like searches for exotic models such as SUSY, etc. we typically use a Type I error value equal to 5%. Therefore, for our analysis, the a-priori criteria are:

- $p_{SM} < 1.3 \times 10^{-4} \quad \Rightarrow 3\sigma$ evidence of the non SM effect

- $p_{SM} < 2.87 \times 10^{-7} \quad \Rightarrow 5\sigma$ observation of the non SM effect

- $p_{SM} > 1.3 \times 10^{-4} \quad \Rightarrow$ we do not exclude the SM

- $p_{XM} < 5\% \quad \Rightarrow$ we would exclude a XM effect with 95% CL

By performing the above two tests simultaneously it is possible to make one of the following four decisions:

1. To reject the XM and at the same time fail to reject the SM.

2. To reject the SM at the 3 or 5 sigma significance level and at the same time fail to reject the XM. This is evidence or observation of a new effect.

3. We fail to reject either the XM or SM. This is what statisticians call an outcome in the "no-decision region".

4. To reject both the XM and SM.

7.2 Bayes Factor

The second approach which is used in the high energy physics is the Bayesian treatment that expresses a degree of belief for a given hypothesis. In our case the difference with respect to the frequentest approach is that no minimization is done. In this case we evaluate the likelihood of the SM hypothesis (using $f_+ = 1$) as well as the likelihood of the XM hypothesis (using $f_+ = 0$) and integrate them over the nuisance parameters to include systematics uncertainties. Then by assuming that the probability of the SM hypothesis is equal to the probability of the XM hypothesis, the Bayes factor can be expressed by equation:

$$BF = \frac{P(x^+, x^-|f_+ = 1)}{P(x^+, x^-|f_+ = 0)} \qquad (7.16)$$

where the likelihood P can be expressed by the signal part L_s (as is defined by equation 7.5) and gausian distributions of the nuisance parameters as follows:

$$P = L_s . G_{N_s} . G_{N_b} . G_{p_s} . G_{p_b} \qquad (7.17)$$

The obtained value expresses how likely is the SM hypothesis in comparison with the XM one. By calculating $2.ln(BF)$ one can get the number similar to χ^2 which can be compared with the following scale [81]:

- 0-2: not worth than a bare mention

- 2-6: positive evidence

- 6-10: strong evidence

- > 10: very strong evidence.

Chapter 8

Results

This analysis is based on an integrated luminosity of 5.6 fb^{-1} collected with the CDF II detector between March 2002 and February 2010. To compare with the previous public result based on the same methodology [69] we increased the statistics by ≈ 4 times.

We have observed 815 events in lepton+jet channel. After applying the $\chi^2 < 9$ cut for the pairing and selection cuts for the jet charge calculation we have obtained 774 pairs, from which 416 pairs are SM-like and the rest 358 pairs are XM-like. In Table 8.1 we summarize the data yields for different lepton types.

Detector	events	events after pairing	JQ defined pairs	SM pairs	XM pairs
CEM electrons	378	183	361	206	155
CMUP muons	175	88	170	87	83
CMX muons	93	52	103	51	52
Non-triggered muons	169	74	140	72	68
Total	815	397	774	416	358

Table 8.1: The observed number of events before and after the pairing cut. The observed number of pairs with the jet charge defined and the observed SM-like and XM-like pairs according to lepton type.

The number of SM (XM) like pairs is obtained by counting the cases when a combine charge of the lepton - b-jet pair is negative (positive). The combined charge is defined

as:

$$Q_{comb} = Q(W) \cdot Q(b - jet) \tag{8.1}$$

where $Q(W)$ is the charge of W boson (or lepton) and Q(b-jet) is charge of the corresponding b-jet. The distribution of the combine charge is shown in Figure 8.1.
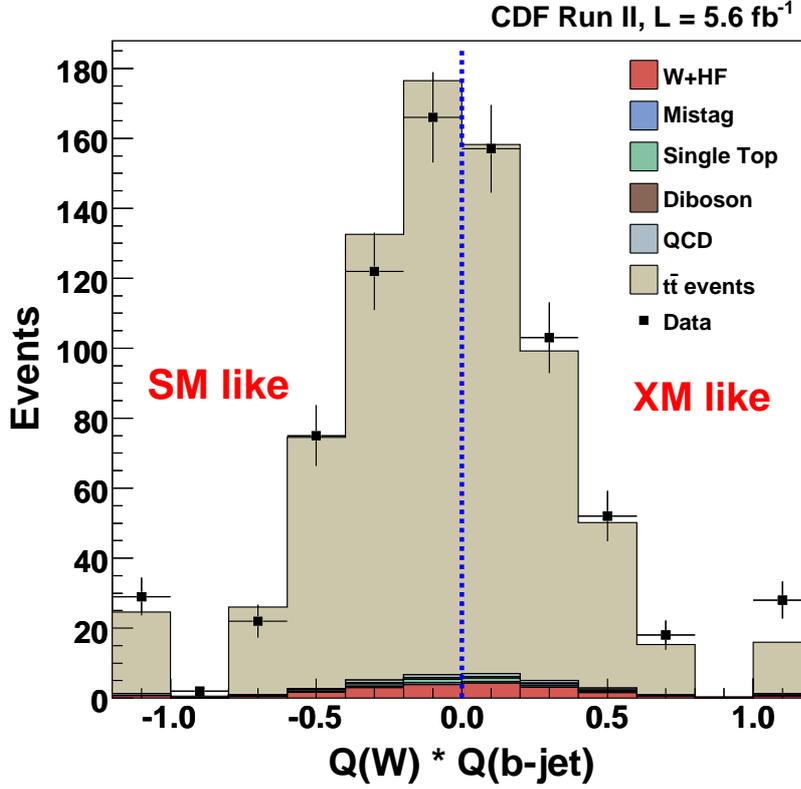
Figure 8.1: W charge $\times$ JetQ for the L+J channel, the SM-like pairs are on the negative side of the plot while the XM-like pairs are on the positive side.

N_s	699.6 ± 115.7
N_b	34.0 ± 8.1
p_s	$0.562 \pm 0.004(stat) \pm 0.011(syst)$
p_b	0.498 ± 0.010

Table 8.2: Expected number of the signal and background and pairs together with the corresponding purities.

Using the profile likelihood with the above number of the SM-like and XM-like pairs (see the last row of Table 8.1) and 4 nuisance parameters, which are summarized in

Table 8.2, we get the log likelihood curve shown in Figure 8.2. The minimum of the curve is at a value of $f_+ = 0.83$. This corresponds to a p-value of 13.4% under the SM hypothesis which we interpret as not excluding the SM hypothesis (when compared to the chosen *a priori* three standard deviation criterium for non-SM evidence). The p-value under the XM hypothesis is 0.014% which is less than 5% so we interpret this result as a 95% confidence level exclusion of the XM hypothesis. If we express the p_{XM} value in standard deviations we would get $\sim 3.5\sigma$. Figure 8.3 shows the probability distributions for SM and XM hypothesis as the function of f_+. The observed value of $f_+ = 0.83$ is shown.

Using the Bayesian treatment, we obtain a value of $2ln(BF) = 19.6$, and conclude that the data favors very strongly the SM over the XM hypothesis.
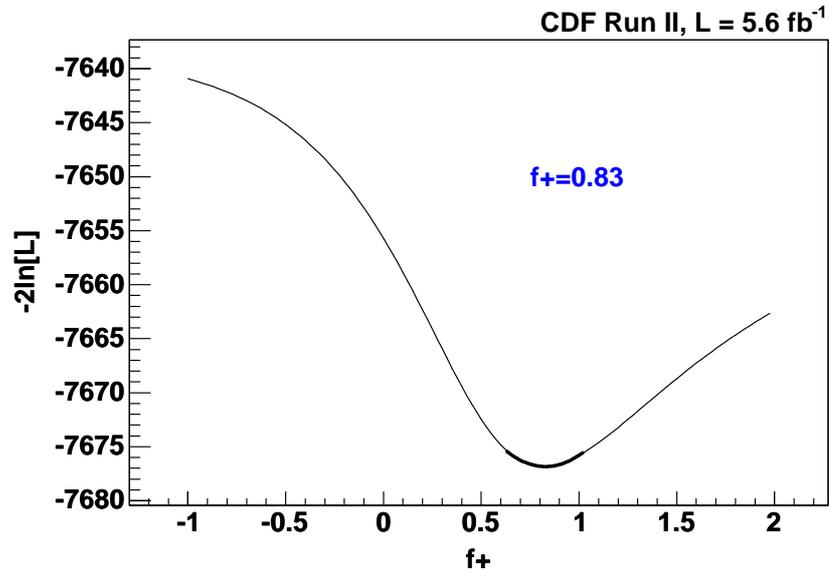
Figure 8.2: The -2LnL curve corresponding to the obtained results, the minimum is at the value of $f_+ = 0.83$.

We checked the results for electrons and muons separately. The summary of the expected number of signal or background pairs as well as corresponding purities can be found in Table 8.3. In Figures 8.4 and 8.5 we show the distribution of the best f_+ obtained using the pseudo-experiments based on either the SM hypothesis or the XM hypothesis for electrons and muons separately.

As we can see in Table 8.4 the p-value under the SM hypothesis for electrons is 0.671 and for muons 0.026. Both values are higher that a-priori criteria value 0.0013,
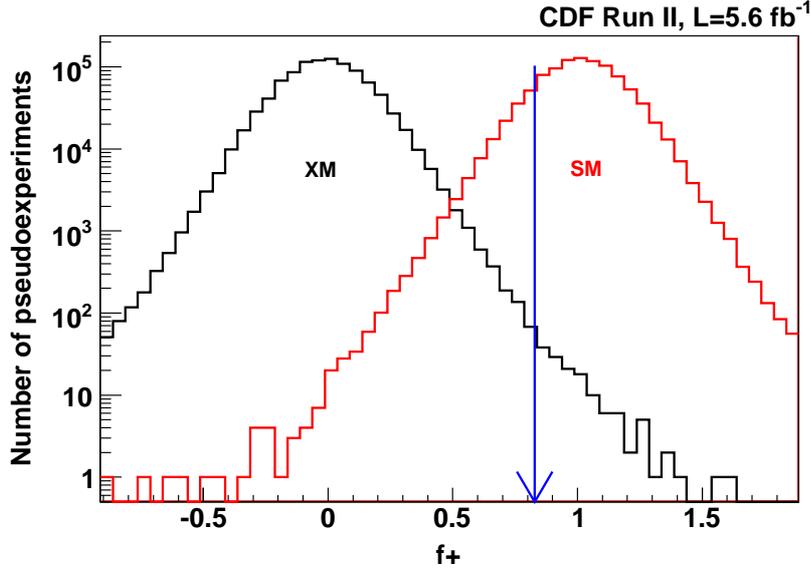
Figure 8.3: Distribution of best f_+ from the pseudo-experiments assuming the XM (black) and SM (red) hypothesis; the blue arrow corresponds to the observed value.

	electrons	muons
N_s	307.8 ± 50.8	391.7 ± 66.6
N_b	17.2 ± 4.6	16.8 ± 4.0
p_s	0.56 ± 0.01	0.56 ± 0.01
p_b	0.50 ± 0.02	0.50 ± 0.01

Table 8.3: Expected number of the signal and background pairs together with the corresponding purities for electrons and muons separately.

what means that we do not exclude the SM in either electrons or muons case. The both p-values under the XM hypothesis (for electron and for muons) are lower then 5%, what means that we can exclude XM hypothesis with 95% CL. One can say that we have the same conclusions as in the case of the combined $(e + \mu)$ result.

Based on the Bayes factor values electrons favors *very strongly* the SM over the XM hypothesis, while muons favors *positively* the SM over the XM hypothesis.
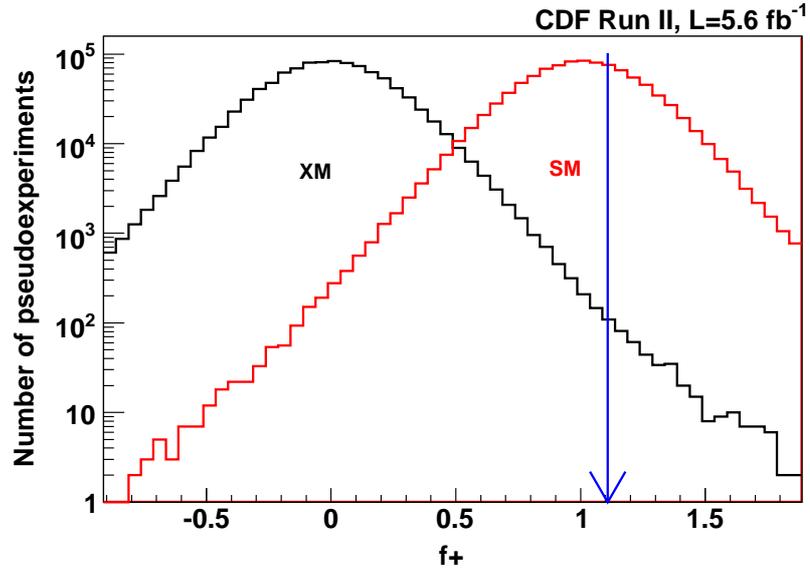
Figure 8.4: Distribution of best f_+ from the pseudo-experiments assuming the XM and the SM for electrons.
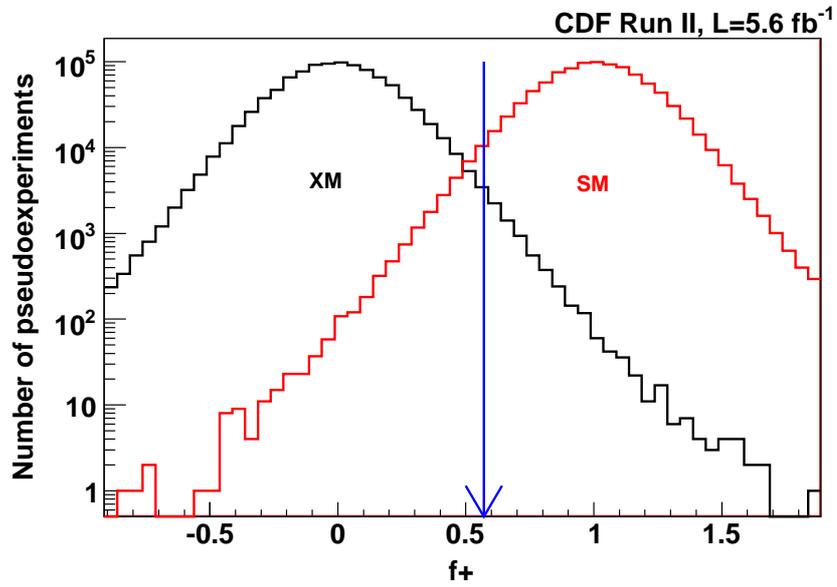
Figure 8.5: Distribution of best f_+ from the pseudo-experiments assuming the XM and the SM for muons.

	electrons	muons
pairs	206 SM like / 155 XM like	210 SM like / 203 XM like
$f+$	1.11	0.57
p_{SM}	0.671	0.026
p_{XM}	0.0004	0.007
$2ln(BF)$	20.3	2.7

Table 8.4: Results of the statistical treatment for electrons and muons separately.

Conclusions

In this thesis I present the measurement of the top quark electric charge via its decay products using the 5.6 fb^{-1} of the data collected by the CDF experiment. Our main goal was to decide if the measured top quark charge supports the Standard model hypothesis or the hypothesis of the exotic quark. In the former case the reconstructed top quark charge would be equal to $+2/3$, while in latter one the charge would be -4/3. The reconstruction of the top quark charge was done in tree steps:

- determination the W boson charge (via the lepton charge),

- correct pairing of the W boson with the b-jet - both should origin from the same top quark decay (using the CDF kinematic fitter, which reconstructs the event topology),

- determination of the b-jet charge (using the jet charge algorithm).

We tried to improve the jet charge algorithm and its purity by applying the further requirement (track to jet energy ratio) and also by studying the resolution of the algorithm as a function of used jet cone. However the results of these studies gave us no significant improvement, so we decided to use the original criteria to keep the statistics as high as possible.

Due to the fact that the MC is not necessarily reliable in term of the jet fragmentation, we calculated the scale factor needed to correct the purity of the jet charge algorithm obtained from the used MC sample. The final result of the scale factor is: $SF_{JQ} = 0.99 \pm 0.01(\text{stat}) \pm 0.03(\text{syst})$. We also checked a possible dependence of the scale factor on the different tracking algorithms used by the CDF (a new tracking algorithm was used for reconstruction of the data collected from April 2008 to February 2010). No strong dependence was observed.

After applying of the all selection criteria and optimization cuts, which were set in the previous analysis based on the 1.5 fb^{-1} of CDF data, we selected 774 W boson - b-jet pairs. Out of them the 140 pairs comes from the non-triggered muons, which we added to increase the muon acceptance. The non-triggered muons compose almost 20% of the whole data sample.

We recalculated the background contribution using also the high luminosity MC samples. These samples were generated with a higher (in comparison with the previous samples) instantaneous luminosity and have to be added to the original background MC samples due to the fact that the instantaneous luminosity of the data events increased. For the same reason we had to recalculate the systematics uncertainty.

The other reason, why we recalculated the background contribution and the systematics uncertainties, is that in the previous analysis the top mass of 175 GeV/c^2 was used to generate the $t\bar{t}$ MC samples and also to reconstruct the event topology by the CDF kinematic fitter. However, the more precise measurement of the top quark mass leaded to the general CDF collaboration's decision of using the top quark mass equal to 172.5 GeV/C^2 for both - generation of MC samples and top quark event reconstruction by the kinematic fitter.

By running over the whole analysis, we obtained 774 W boson - b-jet pairs, out of which the 416 pairs are SM like and 358 pairs are XM like. The statistical treatment of this result leads to the p-value under SM hypothesis of 13.4% and p-value under exotic quark hypothesis of 0.014%. The CDF general accepted approach compared the p-value under the SM with a priori set Type I error $\alpha_1 = 1.3 \times 10^{-4}$ (2.87×10^{-7}) and p-value under the XM model with a priori set Type I error $\alpha_2 = 5\%$.

Using these criteria we can say that we do not exclude the SM, while we exclude the XM with 95% confidence level.

Based on the Bayes scale we can say that the data favors strongly the SM hypothesis over the XM one.

Appendix A

In section 4 we show Figure 4.3, which present some discrepancy between the data and MC in the distribution of the number of tracks used for the jet charge calculation. We checked number of tracks in leptonic and hadronic decay branch separately, also we divided sample into subsamples with positive and negative tracks, but we did not see any difference in number of tracks neither in the tracks p_T distributions.

However this discrepancy could be covered by the jet charge calibration described in Section 5. We compared the distributions of the number of tracks used for the jet charge calculation in the MC samples which we used in our analysis. We compared the ttop25 - $t\bar{t}$ signal sample, dijet MC sample and muon enriched MC sample. As we already mentioned, to increase the statistics of the dijet MC sample we did not require the muon presence in jets.

Due to the fact that there are different selection criteria applied on the jets, we checked the number of tracks distributions only on central jets (the jets with $|\eta| < 0.6$). We also divided this central jets' sample into five subsamples by the jet p_T, because the $t\bar{t}$ events have b-jets with higher p_T with respect to the b-jets from dijet MC and muon enriched MC samples. In Figures 8.6 - 8.10 we present these distributions.

From the plots one can say that for three considered cases the number of tracks' distributions are slightly different. Hence we did further checks, how the selection of the MC sample (dijet or muon enriched one) can affect the final result of the scale factor calculation. There are two possibilities:

1. It can affect the b-fraction on the away jet side, because for the fitting the dijet data we used the templates from the dijet MC (the reason was higher statistics in comparison with muon enriched MC).

2. In the calculation of the jet charge scale factor we compare the jet charge purity
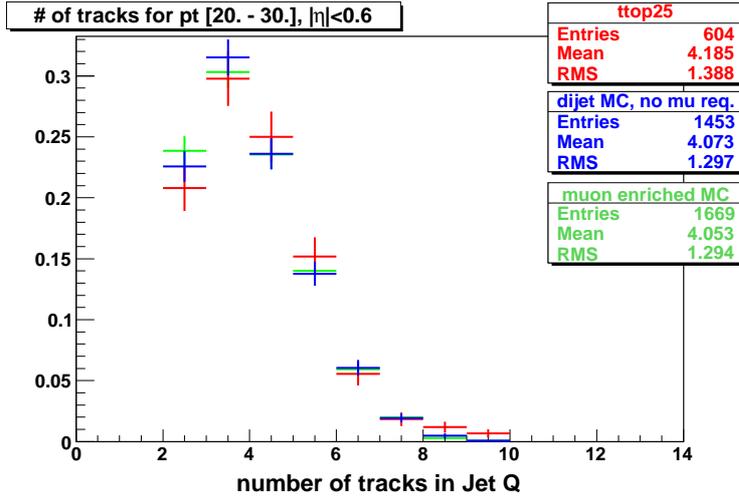
Figure 8.6: The number of tracks in the jet cone 0.4 used to the jet charge calculation for the central jets with p_T between 20 GeV/c and 30 GeV/c.
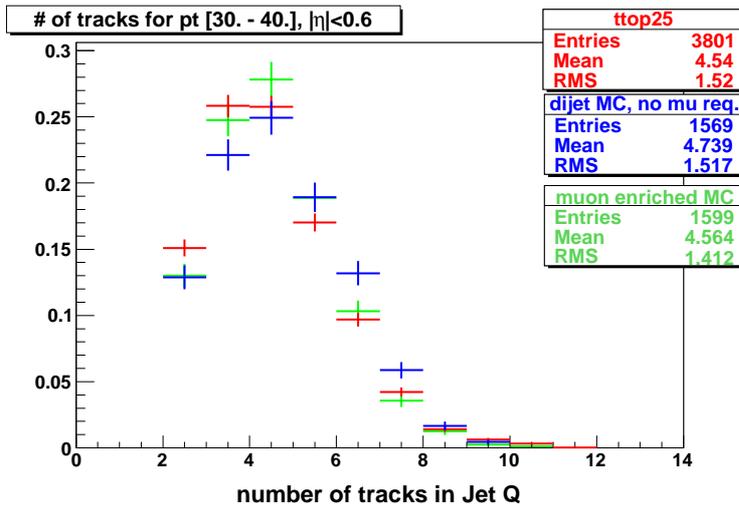
Figure 8.7: The number of tracks in the jet cone 0.4 used to the jet charge calculation for the central jets with p_T between 30 GeV/c and 40 GeV/c.

from the analyzed data with that from the MC. The value of the MC purity was obtained as a combination of the dijet and muon enriched MC samples.

For these reasons we did following:

First we used the muon enriched MC templates (instead of dijet the MC ones) to obtain the b-fraction on the away jet side. We obtained the the jet charge scale factor value of 98%. The difference with respect to the nominal value of $(99 \pm 1)\%$ is within one standard deviation. The result is shown in Figure 8.11.

106
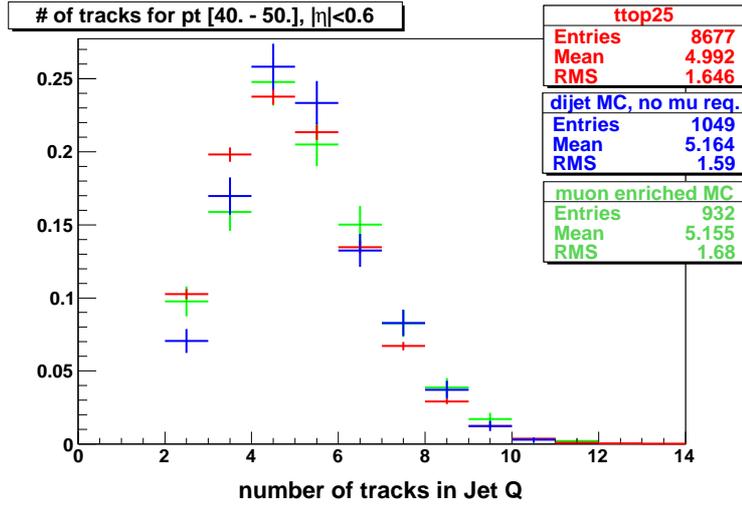
Figure 8.8: The number of tracks in the jet cone 0.4 used to the jet charge calculation for the central jets with p_T between 40 GeV/c and 50 GeV/c.
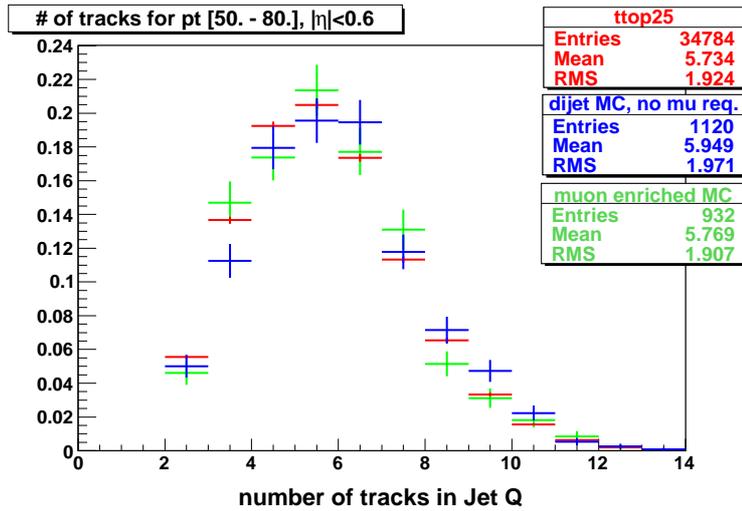
Figure 8.9: The number of tracks in the jet cone 0.4 used to the jet charge calculation for the central jets with p_T between 50 GeV/c and 80 GeV/c.

We also recalculated the scale factor values by using the jet charge purity obtained separately only from the dijet MC samples as well as only from the muon enriched MC sample. There was no big difference in the resulted values of the scale factors (see Figure 8.12).

From these checks we can conclude, that the difference between the data and MC, which we see in the distribution of the number of tracks used for the jet charge calculation is covered by the jet charge scale factor.
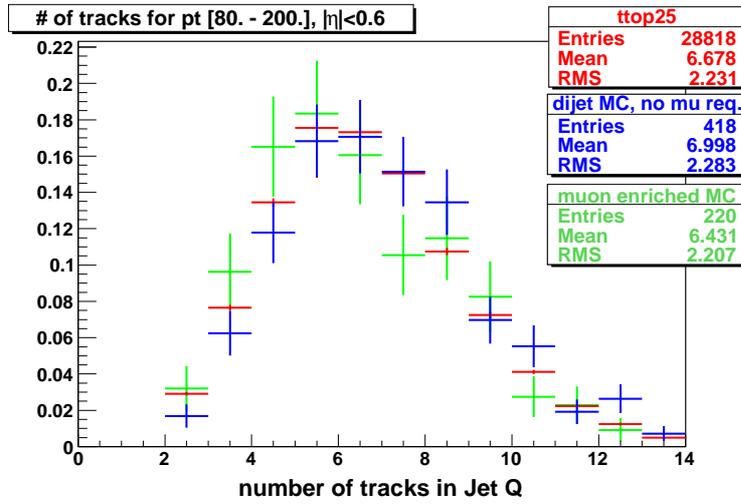
Figure 8.10: The number of tracks in the jet cone 0.4 used to the jet charge calculation for the central jets with p_T between 80 GeV/c and 200 GeV/c.
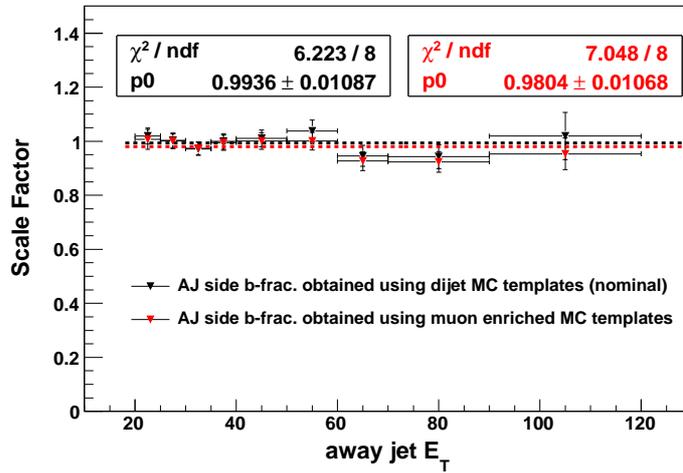
Figure 8.11: By using the muon enriched MC (dijet MC) templates to obtain the b-fraction on the away jet side, we calculated the jet charge scale factor which is shown in black (red) points.
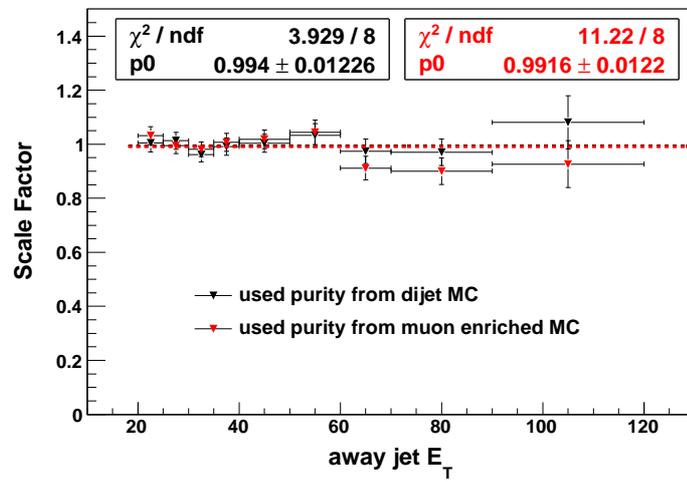
Figure 8.12: The jet charge scale factor calculated by using the purity obtained from the muon enriched (dijet) MC sample shown in black (red).

Bibliography

[1] Fermilab. 2010. *Accelerator Fermilab's Tevatron.* [online], [cited on 2011.03.15.] Available on the internet:
<http://www.fnal.gov/pub/science/accelerator/>

[2] GATTUSO, C. 2010. *Concepts Rookie Books* : Accelerator Concepts. [online], [cited on 2011.03.16.] Available on the internet:
<http://www-bdnew.fnal.gov/operations/rookie_books/Concepts_v3.6.pdf>

[3] ABE, F. et al. CDF collaboration. 1995. Observation of top quark production in $\bar{p}p$ collisions. In *Physical Review Letters.* ISSN 1079-7114, 1995, vol. 74, no. 14, p. 2626-2631.

[4] ABACHI, S. et al. D0 collaboration. 1995. Observation of the Top Quark. In *Physical Review Letters.* ISSN 1079-7114, 1995, vol. 74, no. 14, p. 2632-2637.

[5] AMIDEI, D. et al. CDF-II collaboration. 1996. *The CDF-II detector: Technical design report*, Fermilab, 1996. 325 p. FERMILAB-PUB-96-390-E.

[6] AALTONEN, A. et al. CDF Collaboration. 2010. Combination of CDF top quark mass measurements using up to 5.6 fb^{-1} of data, CDF Conference Note 10202.

[7] AALTONEN, A. et al. CDF Collaboration. 2011. Combination of CDF top quark mass measurements (Winter 2011), CDF Conference Note 10444.

[8] AALTONEN, A. et al. CDF Collaboration. 2008. First run II measurement of the W boson mass at the Fermilab Tevatron. In *Physical Review D.* ISSN 1089-4918, 2008, vol. 77, no. 11, 48 p.

[9] The Tevatron Electroweak Working Group. 2009. Updated Combination of CDF and D0 Results for the Mass of the W Boson, FERMILAB-TM-2439-E, arXiv:0908.1374v1 [hep-ex].

[10] AALTONEN, A. et al. CDF Collaboration. 2011. Search for New Dielectron Resonances and Randall-Sundrum Gravitons at CDF, CDF Note 10405.

[11] AALTONEN, A. et al. CDF Collaboration. 2009. Search for stop decaying into charm + neutralino in the MET+jets sample, CDF Note 9834.

[12] AALTONEN, A. et al. CDF Collaboration. 2009. Search for the Production of Narrow $t\bar{b}$ Resonances in 1.9 fb^{-1} of $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV. In *Physical Review Letters*. ISSN 1079-7114, 2009, vol. 103, no. 4, 7 p.

[13] AALTONEN, A. et al. CDF Collaboration. 2010. Search for high mass resonances decaying to muon pairs, CDF Note 10165.

[14] AALTONEN, A. et al. CDF Collaboration. 2011. Search for heavy bottom-like quarks decaying to an electron or muon and jets in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. In *Physical Review Letters*. ISSN 1079-7114, 2011, vol. 106, no. 14, 7 p.

[15] AALTONEN, A. et al. CDF Collaboration. 2011. Search for top quark pairs with invisible particles, CDF Conference Note 10374.

[16] SILL, A. et al. CDF Collaboration. 2000. CDF Run II silicon tracking projects. In *Nuclear Instruments and Methods in Physics Research Section A*. ISSN 0168-9002, 2000, vol. 447, no. 1-2, p. 1-8.

[17] AFFOLDER, T. et al. CDF Collaboration. 2004. COT Central Outer Tracker. In *Nuclear Instruments and Methods in Physics Research Section A*. ISSN 0168-9002, 2004, vol. 518, no. 1-2, p. 117-120.

[18] AALTONEN, A. et al. CDF Collaboration. 2011. Preliminary Study of Jet Substructure at CDF, CDF Note 10199.

[19] AALTONEN, A. et al. CDF Collaboration. 2010. Measurement of Z/$\gamma \to \mu\mu+$ jets Production Cross Section, CDF Note 10216.

[20] ABULENCIA, A. et al. CDF Collaboration. 2006. Observation of $B_s^0 - \bar{B}_s^0$ Oscillations. In *Physical Review Letters*. ISSN 1079-7114, 2006, vol. 97, no. 24, 8 p.

[21] The Tevatron New-Phenomena and Higgs Working Group. 2011, Combined CDF and D0 Upper Limits on Standard Model Higgs Boson Production with up to 8.2 fb^{-1} of Data, arXiv:1103.3233v2 [hep-ex], FERMILAB-CONF-11-044-E.

[22] AALTONEN, A. et al. CDF Collaboration. 2011. Search for H to WW production using 7.1 fb^{-1}, CDF Note 10432.

[23] ALBROW, M. G. et al. CDF Collaboration. 2002. The CDF plug upgrade electromagnetic calorimeter: test beam results. In *Nuclear Instruments and Methods in Physics Research Section A*. ISSN 0168-9002, 2002, vol. 480, p. 524-546.

[24] DOWNING, R. et al. CDF Collaboration. 2007. Track Extrapolation and Distribution for the CDF-II Trigger System. In *Nuclear Instruments and Methods in Physics Research Section A*. ISSN 0168-9002, 2007, vol. 570, p. 36-50.

[25] GOLDSTEIN, J. et al. 2002. Silicon Tracking for Plug Electrons, CDF Note 5970.

[26] BHATTTI, A. et al. 2006. Determination of the Jet Energy Scale at the Collider Detector at Fermilab. In *Nuclear Instruments and Methods in Physics Research Section A*. ISSN 0168-9002, 2006, vol. 566, p. 375-412.

[27] ABULENCIA, A. et al. CDF Collaboration. 2006. Measurement of the $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV using lepton+jet events with jet probability b-tagging. In *Physical Review D*. ISSN 1089-4918, 2006, vol. 74, no. 7, 38 p.

[28] AALTONEN, A. et al. CDF Collaboration. 2010. Measurement of the $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV using soft electron b-tagging. In *Physical Review D*. ISSN 1089-4918, 2010, vol. 81, no. 9, 18 p.

[29] ACOSTA, D. et al. CDF Collaboration. 2005 Measurement of the $t\bar{t}$ Production Cross Section in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV Using Lepton Plus Jets Events with Semileptonic B Decays to Muons. In *Physical Review D*. ISSN 1089-4918, 2005, vol. 72, no. 3, 25 p.

[30] ACOSTA, D. et al. CDF Collaboration. 2005. Measurement of the $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV using lepton+jets events wits secondary vertex b-tagging. In *Physical Review D.* ISSN 1089-4918, 2005, vol. 71, no. 5, 28 p.

[31] PALENCIA, E. - GRISTEIN, S. 2008. *High-Pt B-Tagging at CDF-II.* [online], [cited on 2001.03.18.] Available on the internet:
<http://www-cdf.fnal.gov/physics/new/top/2004/btag/>

[32] The Tevatron Electroweak Working Group. 2010. Combination of CDF and D0 Results on the Mass of the Top Quark, arXiv:1007.3178v1 [hep-ex], FERMILAB-TM-2466-E, TEVEWWG-TOP-2010-07.

[33] MANGANO, M. 1999. 1999 CERN Workshop on standard model physics (and more) at the LHC, CERN, Geneva, Switzerland, 25-26 May: Proceedings. In *Yellow report.* CERN-2000-004 (1999), 529 p.

[34] ELLIS, R. K. - STIRLING, W. - WEBBER, B. 1996. *QCD and Collider Physics.* 1st ed. Edinburgh: Cambridge university press, 1996. 435 p. ISBN 0 521 58189 3.

[35] LIPATOV, L. N. - GRIBOV, V. N. 1972. In *Soviet Journal of Nuclear Physics.* ISSN 0038-5506, 1972, vol. 15, p. 438.
LIPATOV, L. N. 1975. In *Soviet Journal of Nuclear Physics.* ISSN 0038-5506, 1975, vol. 20, p. 95.
DOKSHITZER, Y. L. 1977. In *Soviet Physics - JETP.* ISSN 0038-5646, vol. 45, p. 641.
ALTARELLI, G. - PARISI, G. 1977. Asymptotic freedom in parton language. In *Nuclear Physics B.* ISSN 0920-5632, 1977, vol. 126, no. 2, p. 298-318.

[36] MOCH, S. - UWER, P. 2008. Heavy-quark pair production at two loops in QCD. In *Nuclear Physics B - Proceedings Supplements.* ISSN 0920-5632, 2008, vol. 183, p. 75-80.

[37] CACCIARI, M. et al. 2008. Updated predictions for the total production cross sections of top and of heavier quark pairs at the Tevatron and at the LHC. In *Journal of High Energy Physics.* ISSN 1029-8479, 2008, vol. 2008, no. 9. p. 127.

[38] KIDONAKIS, N. - VOGT, R. 2003. Next-to-next-to-leading order soft-gluon corrections in top quark hadroproduction. In *Physical Review D*. ISSN 1089-4918, 2003, vol. 68, no. 11, 15 p.

[39] AALTONEN, A. et al. The CDF Collaboration. 2009. Combination of CDF top quark pair production cross section measurements with up to 4.6 fb^{s1}, CDF Conference Note 9913.

[40] HARRIS, B. W. at al. 2002. Fully differential single-top-quark cross section in next-to-leading order QCD. In *Physical Review D*. ISSN 1089-4918, 2002, vol. 66, no. 5, 19 p.
SULLIVAN, Z. 2004. Understanding single-top-quark production and jets at hadron colliders. In *Physical Review D*. ISSN 1089-4918, 2004, vol. 70, no. 11, 18 p.
CAMBELL, J. - ELLIS, R. K. - TRAMONTANO, F. 2004. Single top-quark production and decay at next-to-leading order. In *Physical Review D*. ISSN 1089-4918, 2004, vol. 70, no. 9, 17 p.

[41] KIDONAKIS, N. 2006. Single top quark production at the Fermilab Tevatron: Threshold resummation and finite-order soft gluon corrections. In *Physical Review D*. ISSN 1089-4918, 2006, vol. 74, no. 11, 12 p.

[42] AALTONEN, T. et al. CDF Collaboration. 2009. Observation of Electroweak Single Top-Quark Production. In *Physical Review Letters*. ISSN 1079-7114, 2009, vol. 103, no. 9, 8 p.

[43] ABAZOV, V. M. et al. D0 Collaboration. 2009. Observation of Single Top-Quark Production. In *Physical Review Letters*. ISSN 1079-7114, 2009, vol. 103, no. 9, 7 p.

[44] Tevatron Electroweak Working Group. CDF and D0 Collaborations. 2009. Combination of CDF and D0 Measurements of the Single Top Production Cross Section, arXiv:0908.2171v1 [hep-ex], FERMILAB-TM-2440-E.

[45] NAKAMURA, K. et al. Particle Data Group. 2010. Review of Particle Physics. In *Journal of Physics G: Nuclear and Particle Physics*. ISSN 1361-6471, 2010, vol. 37, no. 7A.

[46] JEZABEK, M. - KUHN, J. H. 1989. QCD corrections to semileptonic decays of heavy quarks. In *Nuclear Physics B.* ISSN 0920-5632, 1989, vol. 314, no. 1, p. 1-6.

[47] JEZABEK, M. - KUHN, J. H. 1993. Top quark width: Theoretical update. In *Physical Review D.* ISSN 1089-4918, 1993, vol. 48, no. 5.
JEZABEK, M. - KUHN, J. H. 1994. Erratum: Top quark width: Theoretical update. In *Physical Review D.* ISSN 1089-4918, 1994, vol. 49, no. 9.

[48] AALTONEN, A. et al. CDF Collaboration. 2010. Direct Top-Quark Width Measurement at CDF. In *Physical Review Letters.* ISSN 1079-7114, 2010, vol. 105, no. 23, 7 p.

[49] ABAZOV, V. M. et al. D0 Collaboration. 2011. Determination of the Width of the Top Quark. In *Physical Review Letters.* ISSN 1079-7114, 2011, vol. 106, no. 2, 7 p.

[50] GRŰNEWALD, M. et al. 2001. Present and Future Electroweak Precision Measurements and the Indirect Determination of the Mass of the Higgs Boson, arXiv:0202001v1 [hep-ph].

[51] ZEE, E. 1989. *CERN-JINR School of physics.* Geneva: CERN, 1991.

[52] CERN. 2010. *LEP/TEV EW WG Plots for Summer 2010.* [online], [cited on 2011.04.06.] Available on the internet:
<http://lepewwg.web.cern.ch/LEPEWWG/plots/summer2010/>

[53] KANE, G. L. - LADINSKY, G. A. - YUAN, C.-P. 1992. Using the top quark for testing standard-model polarization and CP predictions. In *Nuclear Physics B.* ISSN 0920-5632, 1992, vol. 45, no. 1, p. 124-141.

[54] AGUILAR-SAAVEDRA, J. et al. 2007. Probing anomalous Wtb couplings in top pair decays. In *The European Physical Journal C - Particles and Fields.* ISSN 1434-6052, 2007, vol. 50, no. 3, p. 519-533.

[55] AALTONEN, A. et al. CDF Collaboration. 2011. W boson helicity measurement in $t\bar{t}$ dilepton channel at CDF, CDF Conference Note 10333.

[56] AALTONEN, A. et al. CDF Collaboration. 2011. Measurement of the Forward Backward Asymmetry in Top Pair Production in the Dilepton Decay Channel using 5.1 fb^{-1}, CDF Conference Note 10436.

[57] AALTONEN, A. et al. CDF Collaboration. 2010. Evidence for a Mass Dependent Forward-Backward Asymmetry in Top Quark Pair Production, arXiv:1101.0034v1 [hep-ex], Fermilab-Pub-10-525-E.

[58] AALTONEN, A. et al. The CDF Collaboration. 2009. A Measurement of $t\bar{t}$ Spin Correlations Coefficient in 2.8 fb^{-1} Dilepton Candidates, CDF Conference Note 9824.

[59] BERNREUTHER, W. et al. 2004. Top quark pair production and decay at hadron colliders. In *Nuclear Physics B*. ISSN 0920-5632, 2004, vol. 690, no. 1-2, p. 81-137.

[60] AALTONEN, A. et al. The CDF Collaboration. 2010. Measurement of $t\bar{t}$ Helicity Fractions and Spin Correlation Using Reconstructed Lepton+Jets Events, CDF Conference Note 10211.

[61] GRZADKOWSKIA, B. - GUNIONB, J. F. - KRAWCZYKA, P. 1999. Neutral current flavor changing decays for the Z boson and the top quark in two-Higgs doublet models. In *Physics Letters B*. ISSN 0370-2693, 1999, vol. 268, no.1, p. 106-111.

[62] YANG, J. M. - YOUNG B-L. - ZHANG, X. 1998. Flavor-changing top quark decays in R-parity-violating supersymmetric models. In *Physical Review D*. ISSN 1089-4918, 1998, vol. 58, no. 5, 6 p.

[63] AGUILA, F. - AGUILAR-SAACEDRA, J. A. - MIQUEL, R. 1999. Constraints on Top Couplings in Models with Exotic Quarks. In *Physical Review Letters*. ISSN 1079-7114, 1999, vol. 82, no. 8, 9 p.

[64] AALTONEN, A. et al. The CDF Collaboration. 2008. Search for the Flavor Changing Neutral Current Decay $t \rightarrow Zq$ in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV. In *Physical Review Letters*. ISSN 1079-7114, 2008, vol. 101, no. 19, 7 p.

116

[65] AALTONEN, A. et al. The CDF Collaboration. 2009. Search for Top-Quark Production via Flavor-Changing Neutral Currents in W + 1 Jet Events at CDF. In *Physical Review Letters*. ISSN 1079-7114, 2009, vol. 102, no. 15, 7 p.

[66] BAUR, U. - BUICE, M. - ORR, L. H. 2001. Direct measurement of the top quark charge at hadron colliders. In *Physical Review D*. ISSN 1089-4918, 2001, vol. 64, no. 9, 12 p.

[67] CHANG, D. - CHANG, W. - MA, E. 1999. Alternative interpretation of the Fermilab Tevatron top events. In *Physical Review D*. ISSN 1089-4918, 1999, vol. 59, no. 9, 4 p.

[68] BEDNÁŘ, P. 2007. *Determination of Top Quark charge in CDF experiment* : PhD. Thesis. Bratislava : Comenius University, 2007. 106 p.

[69] AALTONEN, A. et al. CDF Collaboration. 2007. First CDF Measurement of the Top Quark Charge using the Top Decay Products, CDF Conference Note 8967.

[70] AALTONEN, A. et al. CDF Collaboration. 2010. Exclusion of Exotic Top-like Quark with $-4/3$ Electric Charge Using Soft Lepton Tags, CDF Conference Note 9939.

[71] FIELD, R. D. - FEYNMAN, R. P. 1978. A parametrization of the properties of quark jets. In *Nuclear Physics B*. ISSN 0920-5632, 1978, vol. 136, no. 1, p. 1-76.

[72] BARATE, R. et al. Collaboration ALEPH. 1998. Determination of A_{FB}^b using jet charge measurements in Z decays. In *Physics Letters B*. ISSN 0370-2693, 1998, vol. 426, no.1, p. 217-230.

[73] GARBERSON, F. - INCANDELA, J. - NEU, CH. 2007. SECVTX b-Tag Efficiency Measurement Using Muon Transverse Momentum for 1.2 fb^{-1} Analyses, CDF note 8640.

[74] BORTOLETTO, D. et al. 2005. ID Efficiency for Intermediate pT Muons, CDF note 7197.

[75] CHOU, J. P. - GUIMARAES DA COSTA, J. - SHERMAN, D. 2006. Studies of 6.1.4 Monte Carlo for Top Analyses, CDF note 8426.

[76] ABULENCIA, A. et al. CDF Collaboration. 2006. Measurement of the $t\bar{t}$ Production Cross Section in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV. In *Physical Review Letters*. ISSN 1079-7714, 2006, vol. 97, no. 8, 7 p.

[77] PUMPLIN, J. et al. 2002. New generation of parton distributions with uncertainties from global QCD analysis. In *Journal of High Energy Physics*. ISSN 1029-8479, 2002, vol. 2002, no. 7, 46 p.

[78] MARTIN, A. D. et al. 1995. The α_S dependence of parton distributions. In *Physics Letters B*. ISSN 0370-2693, 1995, vol. 356, no.1, p. 89-94.

[79] ROLKE, W. A. - LOPEZ, A. M. - CONRAD, J. 2005. Limits and confidence intervals in the presence of nuisance parameters. In *Nuclear Instruments and Methods in Physics Research Section A*. ISSN 0168-9002, 2005, vol. 551, no. 2-3, p. 493-503.

[80] COWAN, G. 1998. *Statistical Data Analysis*. 1st ed. Oxford : Oxford University Press, 1998. 200 p. ISBN 0-19-850156-0.

[81] KASS, R. E. - RAFTERY, A. E. 1995. Bayes Factors. In *Journal of the American Statistical Association*. ISSN 0162-1459,1995, vol. 90, no. 430, p. 773-795.

Resumé

Top kvark, predpovedaný Štandardným modelom, bol objavený experimentami CDF a D0 vo Fermilabe v roku 1995. Odvtedy oba experimenty merajú jeho vlastnosti, aby sa potvrdilo alebo vyvrátilo, že objavený top kvark sa správa podľa očakávaní Štandardného modelu (SM). Jednou z týchto vlastností je aj elektrický náboj top kvarku, ktorým sa zaoberáme v tejto práci.

Doba života top kvarku (5×10^{-25} sekundy) je kratšia ako čas potrebný na hadronizáciu ($\approx 10^{-23}$ sekundy), čo vedie k faktu, že štúdium top kvarkových vlastností je podmienené štúdiom jeho rozpadových produktov. Podľa SM sa top kvark (s elektrickým nábojom $+2/3$) rozpadá elektroslabou interakciou najmä na bottom kvark ($t \to W^+ b$) a anti-top kvark na anti-bottom kvark ($\bar{t} \to W^- \bar{b}$). Rozpady na *down* alebo strange kvark sú málo pravdepodobné.

V roku 1999 prišli Chang a kol. s alternatívnou interpretáciou top kvarkových eventov pozorovaných vo Fermilabe [67]. Podľa ich teórie, môžu tieto eventy obsahovať objekt rozpadajúci sa na W^- bozón a bottom kvark, čo by znamenalo, že elektrický náboj tohto objektu je -4/3. Takýmto objektom by podľa autorov mohol byť exotický kvark zo štvrtej generácie kvarkov a leptónov. Na základe ich výpočtov, by top kvark zo SM mal hmotnosť > 230 GeV/c² a hmotnosť exotického kvarku by bola ~ 170 GeV/c².

Táto hypotéza bola vylúčená určením dolných hraníc hmotností t', b' kvarkov ([15, 14]). Možnosť náboja -4/3 bola tiež vylúčená štúdiami [69, 70] avšak s menšou citlivosťou a použijúc menšiu vzorku nazbieraných dát v porovnaní s nami prezentovanými výsledkami.

Produkcia top kvarku

Top kvarkové eventy môžu byt produkované pomocou silnej alebo elektroslabej interakcie. V prípade silnej interakcie je top kvark produkovaný v $t\bar{t}$ pároch prostredníctvom

kvark-antikvarkovej anihilácie alebo gluónovej fúzie, zakiaľ čo v prípade elektroslabej interakcie vzniká v evente iba jeden top kvark. Produkcia prostredníctvom elektroslabej interakcii je v porovnaní s produkciou $t\bar{t}$ párov potlačená asi tri krát.

Dátová vzorka a selekcia eventov

Naša analýza je založená na dátovej vzorke o veľkosti 5.6 fb^{-1} $t\bar{t}$ eventov nameranej experimentom CDF v čase od Februára 2002 do Februára 2010.

Eventy sú najskôr selektované leptónovým triggerom, ktorý požaduje, aby v evente bol elektrón (mión) s priečnou energiu E_T (hybnosť p_T) väčšou ako 18 GeV (GeV/c). Z tejto vzorky potom vyberáme eventy rekonštruované *offline*, ktoré obsahujú jeden izolovaný elektrón s $E_T > 20$ GeV alebo izolovaný mión s $p_T > 20$ GeV/c, tri jety s $E_T > 20 GeV$ a $|\eta| < 2.0$, štvrtý jet, ktorý môže splniť aj slabšie jetové výberové kritériá: $E_T > 12 GeV$ a $|\eta| < 2.4$. Keďže pri leptónovom rozpade vzniká aj neutríno, požadujeme aby chýbajúca priečna energia, $\not{E}_T$, mala hodnotu aspoň 20 GeV.

Triggerované izolované elektróny sú identifikované na základe priradenia rekonštruovaného treku (stopy v trekových detektoroch) k energii zanechanej v elektromagnetickom kalorimetri. Aby sme odlíšili elektrón od hadrónov, požadujeme aby energia zanechaná v hadrónovom kalorimetri bola výrazne menšia oproti tej v elektromagnetickom kalorimetri. Izolačná podmienka elektrónu je splnená, ak energia v kónuse 0.4 okolo elektrónu (nepočítajúc energiu elektrónu) je menšia ako 10% energie elektrónu.

V prípade triggerovaných miónov požadujeme aby rekonštruovaný trek z trekových detektorov bol priradený k stope v oboch miónových komorách CMU a CMP (CMUP mión) alebo k stope v CMX miónovej komore (CMX mión). Izolovanosť miónu je definovaná podobne ako to bolo v prípade elektrónu.

Štatistiku miónových eventov sme rozšírili pridaním eventov s tzv. netriggerovanými miónami, ktoré neprešli triggerom, ale boli rekonštruované offline. Tieto eventy boli vybrané triggerom ktorý požaduje chýbajúcu priečnu energiu na triggerovej úrovni väčšiu ako 35 GeV a aspoň dva jety s $E_T > 10$ GeV. Selektované eventy musia obsahovať stopu v CMX miónovej komore, ktorá nie je "pokrytá" triggerom, alebo stopu v jednej z CMU alebo CMP miónovej komore. Podmienky na izolovanosť a p_T miónov sú aplikované aj v tomto prípade. Kvôli zabezpečeniu celkovej efektívnosti triggera požadujeme, aby eventy obsahovali aspoň dva jety s $E_T > 25$ GeV, vzdialené od seba

s $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} = 0.4$, z ktorých jeden je centrálnym jetom s $|\eta| < 0.9$. Tieto eventy musia spĺňať aj podmienky týkajúce sa chýbajúcej priečnej energie a jetov, ktoré boli uvedené vyššie.

Takto vybraté eventy voláme aj leptón-jetové (LJ) eventy, lebo predpokladáme, že jeden W bozón sa rozpadol leptónovo na leptón a zodpovedajúce neutríno a druhý W bozón sa rozpadol hadrónovo na dva ľahké jety.

Kvôli potlačeniu pozaďových eventov požadujeme, aby eventy vybraté danými kritériá, obsahovali aspoň dva b-taggované jety, pričom jet označíme ako b-taggovaný ak obsahuje sekundárny vertex, charakteristický pre rozpad B-hadrónov. Podrobnejší popis metódy b-taggovania na základe sekundárneho vertexu možno nájsť v [30].

Metóda určenia náboja top kvarku

Metóda rekonštrukcie elektrického náboja top kvarku, použitá v tejto práci, sa skladá z troch krokov:

- určenie náboja W bozónu,

- určenie náboja b-jetu

- a párovanie W bozónu s b-jetom, ktoré pochádzajú z rozpadu toho istého top kvarku.

Ako sme spomenuli v predchádzajúcej časti, vo vybraných eventoch sa jeden W bozón rozpadá leptónovo, zakiaľ čo druhý sa rozpadá hadrónovo na ľahké jety. Náboj leptónovo rozpadajúceho sa W bozónu je určený nábojom leptónu. V prípade druhého W bozónu, definujeme jeho náboj ako náboj opačný k náboju leptónu.

Na určenie náboja b-jetu používame metódu váhovania náboja trekov v jete, ktoré majú $p_T > 1.5$ GeV/c a ich zrážkový parameter je menší ako 0.15 cm. Náboj jetu môžme definovať vzťahom:

$$Q_{bjet} = \frac{\sum_i q_i |\vec{j}_i \cdot \vec{p}_i|^\kappa}{\sum_i |\vec{j}_i \cdot \vec{p}_i|^\kappa} \tag{1}$$

kde q_i a p_i sú náboj a hybnosť i-tého treku v jete, $\vec{j}$ je jednotkový vektor pozdĺž osi jetu a κ je parameter, ktorého hodnota bola optimalizáciou zvolená ako 0.5.

B-jet, ktorého takto určený náboj je kladný, označujeme ako pochádzajúci z b kvarku. Ak je hodnota určeného náboja záporná, hovoríme, že b-jet pochádza z $\bar{b}$

kvarku (čo v skutočnosti nie je vždy pravda). Na základe štúdií Monte Carlo (MC) eventov sme zistili efektívnosť tejto metódy, ktorá je $(97.9 \pm 0.1)\%$. Čistota (purity) metódy, definovaná ako pravdepodobnosť správneho určenia náboja b-jetu, je rovná hodnote $(60.8 \pm 0.1)\%$.

Posledným krokom určenia náboja top kvarku je priradenie W bozónu k správnemu b-jetu, na čo používame CDF kinematický fiter, ktorý počíta hodnotu χ^2 pomocou nasledujúceho vzťahu:

$$\chi^2 = \sum_{i=l,4jets} \frac{(\hat{p}_T^i - p_T^i)^2}{\sigma_i^2} + \sum_{j=x,y} \frac{(\hat{p}_j^{UE} - p_j^{UE})^2}{\sigma_j^2} + \frac{(m_{jj} - m_W)^2}{\Gamma_W^2}$$
$$+ \frac{(m_{l\nu} - m_W)^2}{\Gamma_W^2} + \frac{(m_{bjj} - m_t)^2}{\Gamma_t^2} + \frac{(m_{bl\nu} - m_t)^2}{\Gamma_t^2} \qquad (2)$$

kde prvý člen vyjadruje rozdiel hodnôt medzi rekonštruovanou (p_T^i) a fitovanou $(\hat{p}_T^i)$ priečnou hybnosťou, určenou s chybou σ_i, pre leptón a jety. Druhý člen zahŕňa rozdiel medzi nameranou a fitovanou hodnotou energie (určenej s chybou $\sigma_{x,y}$), ktorá nebola započítaná do jetových klasterov. Nasledujúce dva členy určujú rozdiel hmotnosti W bozónu a jeho rozpadových produktov - jetov (m_{jj}) a leptónov $(m_{l\nu})$. Ostatné členy uvažujú rozdiel medzi hmotnosťou top kvarku a jeho rozpadových produktov v hadrónovej (m_{bjj}) a leptónovej $(m_{bl\nu})$ vetve. Tieto členy sú predelené zodpovedajúcimi rozpadovými pološírkami Γ_W a Γ_t. Hmotnosť W bozónu a spomínané rozpadové pološírky sú fixované na ich hodnoty z Particle Data Group (PDG) [45]. V našom prípade sme fixovali aj hmotnosť top kvarku na hodnote 172.5 GeV/c^2.

Keďže eventy obsahujú 2 b-jety, máme iba dve možnosti priradenia W bozónov k b-jetom. Avšak kvôli neurčitosti z-ovej zložky neutrína, máme 4 možné výstupy z fitera, z ktorých vyberáme ten s najmenším χ^2. Touto metódou priraďujeme W bozón k b-jetu správne v 76% prípadov. Na základe optimalizácie sme zistili, že vyberaním iba takých eventov, ktorých minimálna hodnota χ^2 je menšia ako 9, sa zvýši čistota párovania (zlomok správnych priradení W bozónu k b-jetu) na $(83.3 \pm 0.1)\%$. Efektívnosť tohto výberového kritéria je $(53.2 \pm 0.1)\%$.

Kalibrácia čistoty určenia náboja b-jetov na dátach

Keďže MC simulácie neopisujú presne fragmentáciu jetov, je nutné urobiť korekciu čistoty metódy určovania náboja jetu získanej z MC. Používame na to dijetovú vzorku dát obohatenú o tažké kvarky (bottom a charm). Táto vzorka je vybraná použitím

triggera, ktorý požaduje centrálny mión s $p_T > 8$ GeV/c. Výberové kritériá ďalej požadujú prítomnosť miónového treku s $p_T > 9$ GeV/c, nachádzajúceho sa v jete s $E_T > 20$ GeV (tento jet voláme aj *miónovým jetom*) a prítomnosť druhého jetu s $E_T > 20$ GeV, nazývaného textitaway jet, ktorého smer je opačný ako smer miónového jetu ($\Delta\phi > 2$). Oba jety musia byť b-taggované - miónový jet na základe silnejších a away jet na základe slabších kritérií b-taggera používajúceho sekundárny vertex.

V takto vybratých dijetových eventoch je náboj miónového jetu určený nábojom miónu, zakiaľ čo náboj away jetu určujeme na základe metódy váhovania trekov ako sme uviedli vyššie (rovnica 1). Čistotu určenia náboj jetu metódou váhovania trekov môžeme potom určiť ako podiel počtu eventov, v ktorých majú jety opačné znamienko náboja a počtu všetkých eventov, ktoré prešli výberovými kritériami. Takto určená čistota však musí byť korigovaná, kvôli niekoľkým efektom. Ak mión pochádza zo sekundárneho rozpadu (napr. $b \to c \to \mu$) znamienko jeho náboja bude opačné ako v prípade, keď pochádza priamo z b rozpadu. Zmena znamienka môže nastať aj vtedy, ak dôjde v niektorom z jetov k B-mixingu. Ďalšou korekciou, ktorú treba zahrnúť je vplyv nesprávneho b-taggovania, kedy b-taggovaný jet nie je skutočným b-jetom. V takom prípade neočakávame nábojovú koreláciu medzi jetmi. Príspevky prvých dvoch vplyvov - sekundárnych rozpadov a B-mixingu, vyšetrujeme na základe MC, v prípade non-b jetov sme podiel eventov zistili použitím dát.

Na určenie podielu $b\bar{b}$ dijetových eventov (kedy oba jety pochádzajú z b kvarku) fitujeme dátové rozdelenia dvoch rôznych premenných zodpovedajúcimi rozdeleniami pre b, c a ľahké jety získanými z MC, použijúc *MC truth* (prípad gluónových jetov je zahrnutý v rozdelení pre ľahké jety). Použitie *MC truth* je založené na porovnávaní informácie generátora na partónovej úrovni s informáciou získanou z rekonštrukcie eventu.

Prvou premennou je $p_{T,rel}$ (priečna komponenta miónovej hybnosti určená vzhľadom na os jetu), ktorej stredná hodnota je vyššia pre mióny pochádzajúce z b-kvarkového jetu v porovnaní s miónami pochádzajúcimi z c-kvarkového jetu alebo jetov pochádzajúcich z ľahkých kvarkov. Na fitovanie dátového rozdelenia tejto premennej používame iba MC rozdelenia b-kvarkového a c-kvarkového jetu, nakoľko rozdelenia pre c-kvarkový jet a ľahké jety sú podobné a pre fiter ťažko rozlíšiteľné. Rozdiel vzniknutý rozdielnym fitovaním (použitie troch MC rozdelení namiesto uvedených dvoch) je započítaný do

systematickej chyby.

Druhou premennou je invariantná hmotnosť sekundárneho vertexu M_{vtx}, ktorú používame na určenie podielu b-jetov v away jetoch. Táto hmotnosť je priamoúmerná hmotnosti kvarku vytvárajúceho jet.

Obrázok 1 zobrazuje rozdelenia $p_{T,rel}$ a M_{vtx} pre jety pochádzajúce z b, c kvarkových jetov a ľahkých jetov.
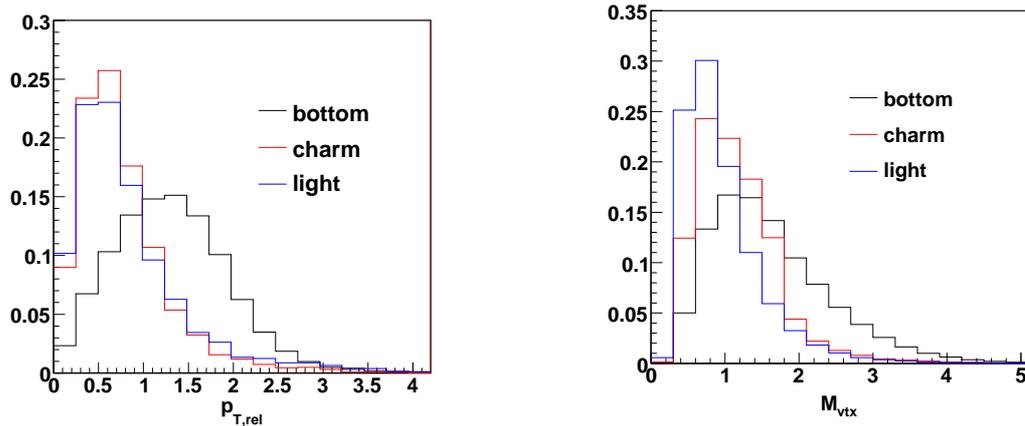
Obr. 1: Rozdelenia $p_{T,rel}$ (vľavo) a M_{vtx} (vpravo) pre jety pochádzajúce z b, c kvarkových jetov a ľahkých jetov.

Nakoľko sme zistili, že MC rozdelenia daných premenných sú závislé na energii away jetu, rozdelili sme danú vzorku do 9-tich binov na základe away jet E_T. Pre každý bin sme spočítali podiel $b\bar{b}$ dijetových eventov, $f_{b\bar{b}}$, ako priemernú hodnotu z maximálneho a minimálneho podielu b-kvarkových jetov. Maximálna hodnota podielu b-kvarkových jetov je určená fitom dátového rozdelenia M_{vtx} pre away jety, zakiaľ čo minimálna hodnota je určená odčítaním podielu nie-b-kvarkových jetov v miónovom jete (určenej fitom dátového rozdelenia $p_{T,rel}$) od maximálnej hodnoty. Chyba podielu $f_{b\bar{b}}$ je určená tak, aby pokrývala rozdiel medzi maximálnou a minimálnou hodnotou.

Kombináciou podielu $b\bar{b}$ dijetových eventov, $f_{b\bar{b}}$, s podielmi eventov, v ktorých došlo k sekundárnemu rozpadu, alebo B-mixingu, môžeme zo získanej čistoty metódy určenia náboja jetu vypočítať reálnu čistotu tejto metódy. Konečný výsledok kalibrácie potom môžeme vyjadriť ako škálovací faktor definovaný podielom reálnej čistoty metódy určenej z dát a čistoty získanej z MC vzorky. Ako sme spomenuli, analýzu sme urobili pre 9 rôznych away jet E_T binov. Z Obrázku 2 je vidno, že škálovací faktor nie je závislý na E_T away jetu.
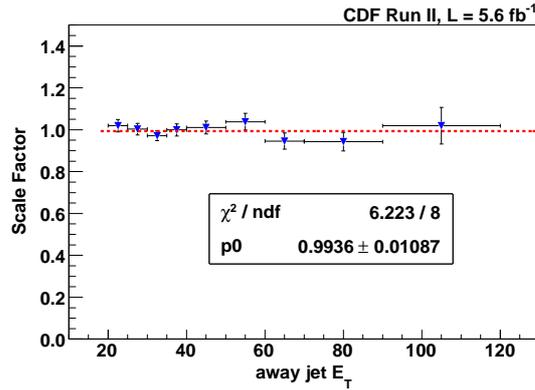
Obr. 2: Škálovací faktor ako funkcia away jet E_T.

Systematická chyba určenia škálovacieho faktora zahŕňa tri rôzne efekty. Prvým je neurčitosť v MC rozdeleniach pre b, c a ľahké kvarkové jety spôsobená napr. efektívnosťou určovania trekov, ďalším je výber počtu (dvoch alebo troch) MC rozdelení použitých pri fitovaní a posledný zahŕňa možnú závislosť na away jet E_T.

Finálny výsledok kalibrácie - škálovací faktor má hodnotu:

$$SF_{JQ} = 0.99 \pm 0.01(\text{stat}) \pm 0.03(\text{syst}).$$

Pozadie pre top kvarkové eventy

Pre leptón-jetové eventy tvorí dominantnú časť pozadia QCD produkcia W bozónov s multi-jetmi. Tieto eventy môžu prejsť výberovými kritériami ak jeden z jetov je b-jet (W+HF), lebo je ľahký jet nesprávne označený ako b-jet (mistag). Medzi ďalšie zdroje pozadia patria QCD multi-jetové eventy, v ktorých dva jety boli nesprávne označené ako b-jety a jeden z jetov bol interpretovaný ako leptón, produkcia eventov s jedným top kvarkom a dibozónové eventy. Príspevok pozaďových eventov je veľmi malý ($\approx 15\%$) vďaka požiadavke aspoň dvoch b-jetov v evente.

Predpokladané počty eventov pre jednotlivé typy pozadia sme získali rovnakou metódou, aká bola použitá v meraní účinného prierezu $t\bar{t}$ eventov v leptón-jetovom kanále (viď [76]). Aplikovaním efektívnosti párovania W bozónu s b-jetom a efektívnosti počítania náboja jetu, určených pre jednotlivé typy pozadia zvlášť, sme získali konečné počty leptón - b-jetových párov pre pozaďové vzorky, ktoré vstupujú do analýzy.

Našim cieľom pri štúdiu pozaďových eventov je zistiť, či je v danom pozadí možné pozorovať koreláciu medzi nábojmi leptónu a k nemu priradenemu b-jetu. Takúto koreláciu očakávame iba v dvoch prípadoch: eventoch obsahujúcich iba jeden top kvark

Pozadie	N_b	Čistota (korelácia)	N^+	N^-
W+HF	19.5 ± 6.4	0.5 ± 0.0	9.7 ± 3.2	9.7 ± 3.2
QCD fakes	5.4 ± 4.8	0.48 ± 0.06	2.6 ± 2.3	2.8 ± 2.5
Dibozóny	2.0 ± 0.4	0.5 ± 0.0	1.0 ± 0.2	1.0 ± 0.2
Mistag	2.8 ± 0.8	0.5 ± 0.0	1.4 ± 0.4	1.4 ± 0.4
Single top	4.4 ± 0.5	0.51 ± 0.01	2.25 ± 0.3	2.15 ± 0.3
Celkom	$\mathbf{34.0 \pm 8.1}$	$\mathbf{0.50 \pm 0.01}$	$\mathbf{16.9 \pm 4.0}$	$\mathbf{17.0 \pm 4.1}$
Signál	$\mathbf{699.6 \pm 115.7}$	$\mathbf{0.562^{\pm 0.004(stat)}_{\pm 0.011(syst)}}$	$\mathbf{393.5 \pm 65.6}$	$\mathbf{306.1 \pm 51.3}$

Tabuľka 1: Hodnoty korelácie pre jednotlivé type pozadia spolu s očakávaným počtom párov zodpovedajúcich SM hypotéze (N^+) a XM hypotéze (N^-). Metóda určenia čistoty pre signálové eventy je popísaná neskôr.

a QCD $b\bar{b}$ eventoch, ktoré prešli výberovými kritériami vďaka tomu, že leptón zo semileptónového rozpadu bol nesprávne určený ako signálny leptón. V prvom prípade sme určili koreláciu použitím zodpovedajúcej MC vzorky, zakiaľ čo v druhom prípade sme použili špeciálnu QCD vzorku určenú na štúdium fake (falošných) elektrónov (elektrónov, ktoré nesplnili aspoň dve kritéria požadované pre signálny elektrón).

V Tabuľke 1 uvádzame hodnoty príspevkov jednotlivých typov pozadia, ako aj hodnoty korelácie (hodnoty 0.5 ± 0.0 sú uvedené v prípadoch, kedy sme neočakávali ani v MC nepozorovali koreláciu). V poslednom riadku tejto tabuľky možno nájsť aj informáciu týkajúcu sa predpokladaného počtu signálnych $t\bar{t}$ eventov.

Určenie systematickej chyby

Systematická chyba tejto analýzy pochádza z modelovania geometrickej a kinematickej akceptancie metódou MC, efektívnosti b-taggovania jetov na základe sekundárneho vertexu, z neurčitostí jetovej energetickej škály, predpokladaného počtu pozaďových eventov a luminozity.

Monte Carlo modelovanie geometrickej a kinematickej akceptancie zahŕňa efekty partónovej rozdeľovacej funkcie (PDF), radiácie (gluónu alebo fotónu) pred (ISR) a po (FSR) zrážke partónov, ale aj jetovú energetickú škálu. Systematická chyba spôsobená týmito efektami je počítaná porovnávaním použitia rôznych PDF setov a variovaním

Systematika (in %)	efekt. párovania	efekt. JetQ ϵ	čistota párovania	čistota JetQ
Jetová energetická škála	0.2	0.04	0.1	0.1
ISR/FSR	0.5	0.1	0.2	0.2
MC generátor	0.2	0.1	0.1	(0.7)
Hmotnosť top kvarku	0.4	0.2	0.9	0.5
PDF	0.7	0.02	0.1	0.02
Celkom	1.0	0.3	1.0	0.6

Tabuľka 2: Systematická chyba (v %). Hodnota "(0.7)" je uvedená iba informatívne, nie je zahrnutá do celkovej systematickej chyby čistoty určenia náboja jetu, nakoľko táto čistota je kalibrovaná na dátach a jej chyba zahŕňa aj neurčitosť spôsobenú rozdielnymi modelmi hadronizácie.

ISF, FSR efektov ako aj variovaním jetovej energetickej škály.

Ďalším zdrojom systematickej chyby je výber MC generátora. Pre jej odhadnutie sme porovnali HERWIG a PYTHIA MC generátory.

Posledný príspevok systematickej chyby je spôsobený fixovaním hmotnosti top kvarku v CDF kinematickom fiteri. Jeho hodnotu sme určili použitím viacerých MC vzoriek generovaných s rôznymi hmotnosťami top kvarku, pričom vo fiteri sme hmotnosť top kvarku fixovali stále na 172.5 GeV/c².

V Tabuľke 2 uvádzame ako jednotlivé zdroje systematických chýb prispievajú k neurčitosti efektívnosti a čistoty párovania W bozónu s b-jetom, ale aj neurčitosti efektívnosti a čistoty určenia náboja jetu (jetQ).

Určenie čistoty signálu

Na určenie čistoty signálu nestačí iba vynásobiť čistotu párovania, $p_{pairing}$, a čistotu určenia náboja b-jetu, p_{JetQ}. Do celkovej čistoty signálu musíme zahrnúť aj prípady, kedy sme nesprávne priradili W bozón k b-jetu, ale zároveň sme aj nesprávne určili náboj b-jetu. Treba uvážiť aj eventy, v ktorých b-tagger nesprávne označil ľahké jety ako b-jety a teda v nich neočakávame koreláciu medzi nábojmi leptónu a prislúchajúceho b-jetu. Ak zhrnieme tieto prípady do jednej rovnice, môžeme pre celkovú čistotu signálu písať:

$$p_s = f_{nonb} \cdot SF_{nonb} \cdot p_{nonb} +$$
$$(1 - f_{nonb} \cdot SF_{nonb})(p_{pair} \cdot p_{JQ} \cdot SF_{JQ} + (1 - p_{pair})(1 - p_{JQ} \cdot SF_{JQ})) \quad (3)$$

f_{nonb}	0.076 ± 0.001
SF_{nonb}	1.01 ± 0.03
p_{nonb}	0.5 ± 0.01
p_{pair}	$0.833 \pm 0.001(stat) \pm 0.008(syst)$
p_{JQ}	$0.608 \pm 0.001(stat) \pm 0.003(syst)$
SF_{JQ}	$0.99 \pm 0.01(stat) \pm 0.03(syst)$

(a)

N_s	699.6 ± 115.7
N_b	34.0 ± 8.1
p_s	$0.562 \pm 0.004(stat) \pm 0.011(syst)$
p_b	0.498 ± 0.010

(b)

Tabuľka 3: Hodnoty premenných použité pre určenie celkovej čistoty signálu (vľavo) a hodnoty premenných, ktoré budú použité pri štatistickom vyšetrovaní výsledkov (vpravo).

kde f_{nonb} predstavuje podiel eventov, v ktorých b-jet nesprávne b-taggovaný, škálovací faktor, SF_{nonb}, koriguje tento podiel, f_{nonb}, nakoľko je v MC vzorkách podhodnotený a p_{nonb} vyjadruje nábojovú koreláciu medzi nesprávne b-taggovaným b-jetom zodpovedajúcim leptónom. Škálovací faktor SF_{JQ} bol definovaný v predchádzajúcom texte.

V Tabuľke 3.a sme zhrnuli hodnoty premenných použité pre výpočet čistoty signálu. V Tabuľke 3.b sumarizujeme hodnoty veličín, ktoré budú použité pri štatistickom vyšetrovaní výsledkov.

Štatistická analýza

Po priradení W bozónu k b-jetu a určení náboja b-jetu, môžme každý takýto pár označiť ako pár Štandardného modelu alebo pár exotického modelu. Na určenie stupňa vierohodnosti daných hypotéz používame metódu pravdepodobnostného profilu popísanú v [79]. Naša pravdepodobnostná funkcia závisí od podielu párov Štandardného modelu f_+ určeného zo signálnej vzorky a štyroch tzv. *nuisance* parametrov - počtu signálnych (N_s) a pozaďových (N_b) párov a čistoty signálu (p_s) a pozadia (p_b). Je vyjadrená ako suma člena reprezentujúceho kombinovanú čistotu signálu a pozadia (Poissonovské rozdelenie), a štyroch členov reprezentujúcich *nuisance* parametre (Gaussovské rozdelenia):

$$L_s = \frac{(N_+)^{x^+} e^{-N_+}}{x^+!} \frac{(N_-)^{x^-} e^{-N_-}}{x^-!} \qquad L_Z = \frac{1}{\sigma_Z \sqrt{2\pi}} e^{\frac{(t-Z)^2}{2\sigma_Z^2}} \qquad (4)$$

kde L_Z zastupuje Gaussovské rozdelenie pre *nuisance* parameter Z, t je parameter fitu a N_+ a N_- sú predpokladané počty párov Štandardného a exotického modelu,

určené rovnicami:

$$N_+ = p_s N_s f_+ + (1 - p_s)N_s(1 - f_+) + p_b N_b \tag{5}$$

$$N_- = (1 - p_s)N_s f_+ + p_s N_s(1 - f_+) + (1 - p_b)N_b \tag{6}$$

Členy x^+ a x^- reprezentujú počet párov Štandardného a exotického modelu v dátach,

Skenovaním hodnoty f_+ v intervale od -1 po 2 a minimalizáciou pravdepodobnostnej funkcie cez *nuisance* parametre dostávame krivku závislosti pravdepodobnosti na hodnote f_+. Minimum tejto krivky je hľadanou hodnotou podielu f_+ (viď Obrázok 3). Obrázok 4 zobrazuje rozdelenie hodnôt f_+ určených minimalizáciou pravdepodobnostnej funkcie, získaných z pseudo-experimentov predpokladajúc SM (červená čiara) ako aj XM(čierna čiara). Určením hodnoty f_+ z dát, môžeme na základe týchto rozdelení určiť p-value pre hypotézu SM, p_{SM} a p-value pre hypotézu XM, p_{XM}, ktoré potom porovnávame s *a priórne* stanovenými Type I chybami α. V našom prípade porovnávame p_{SM} s $\alpha_1 = 1.3 \times 10^{-4}$ (2.87×10^{-7}), čo zodpovedá 3σ (5σ), zatiaľ čo p_{XM} porovnávame s $\alpha_2 = 5\%$. Rozdielnosť týchto kritérií je daná tým, čo je všeobecne používané vo fyzike vysokých energií, a teda aj experimentom CDF.

Okrem hodnôt p-value, využívame aj prístup založený na Bayes faktore (BF), ktorý vyjadruje pomer pravdepodobnosti SM hypotézy v porovnaní s XM hypotézou. Určením hodnoty výrazu $2ln(BF)$ a použitím škály definovanej v [81] môžeme povedať, nakoľko je SM hypotéza pravdepodobnejšia ako XM hypotéza.
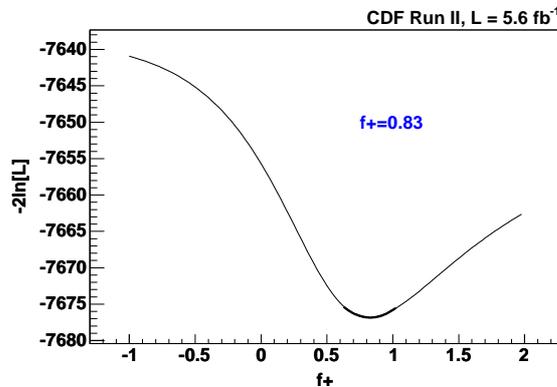
Obr. 3: Krivka závislosti pravdepodobnostnej funkcie od hodnoty f_+. Funkcia nadobúda minimum v hodnote 0.83.
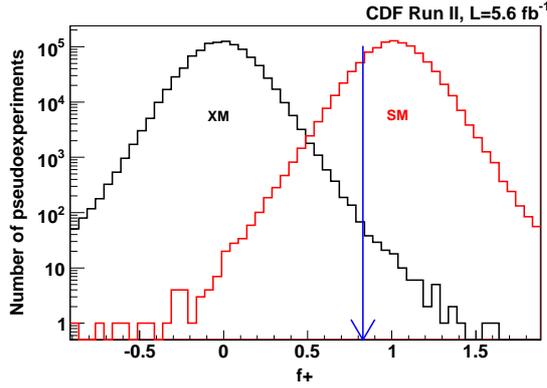
Obr. 4: Rozdelenie hodnôt f_+ určených minimalizáciou pravdepodobnostnej funkcie, získaných z pseudo-experimentov predpokladajúc SM (červená čiara) ako aj XM (čierna čiara). Modrá šípka zobrazuje hodnotu f_+ určenú z dát.

Počet eventov	po párovaní	počet párov	SM	XM
815	397	774 pairs	416	358

Tabuľka 4: Počet leptón-jetových eventov získaných z dát pred a po aplikovaní párovacích kritérií. Počet párov, v ktorých b-jet splnil kritériá na určenie jeho náboja a počty párov zodpovedajúcich Štandardnému modelu (SM) a exotickému modelu (XM).

Výsledky a diskusia

V Tabuľke 4 prezentujeme počet eventov a párov, ktoré prešli leptón-jetovými selekčnými kritériami ako aj požiadavkami aplikovanými pri párovaní W bozónu s b-jetom a pri počítaní náboja b-jetu. Uvádzame tiež počty párov zodpovedajúcich hypotéze SM a XM.

Na základe týchto hodnôt dostávame krivku závislosti pravdepodobnostnej funkcie od f_+, ktorej minimum nachádzame v hodnote 0.83 (viď Obrázok 3). Táto hodnota zodpovedá p-value pre hypotézu SM $p_{SM} = 13.4\%$ (viď Obrázok 4), čo po porovnaní s *a priórne* stanovenou hodnotou α_1, zodpovedajúcou 3σ (5σ) kritériu vylúčenia tejto hypotézy, interpretujeme ako nevylúčenie hypotézy SM. P-value pre hypotézu XM je $p_{XM} = 0.014\%$, čo po porovnaní s *a priórne* stanovenou hodnotou $\alpha_2 = 5\%$, prezentujeme ako vylúčenie hypotézy XM s 95% stupňom vierohodnosti. Treba však povedať, že p_{XM} zodpovedá približne 3.5σ.

Na základe Bayes faktora sme určili hodnotu $2ln(BF) = 19.6$, čo na základe

spomínanej škály vedie k záveru, že dáta *veľmi silno* preferujú SM oproti XM. Na Obrázku 5 zobrazujeme grafickú reprezentáciu nášho výsledku.

Pre kontrolu sme porovnali výsledky určené separátne pre elektróny a mióny. V oboch prípadoch môžeme povedať, že nevylučujeme SM, zakiaľ čo XM vylučujeme s 95% stupňom vierohodnosti. Na základe Bayes faktora môžeme skonštatovať, že elektróny *veľmi silno* preferujú SM oproti XM. V prípade miónov je *pozitívne* preferovaný SM oproti XM.
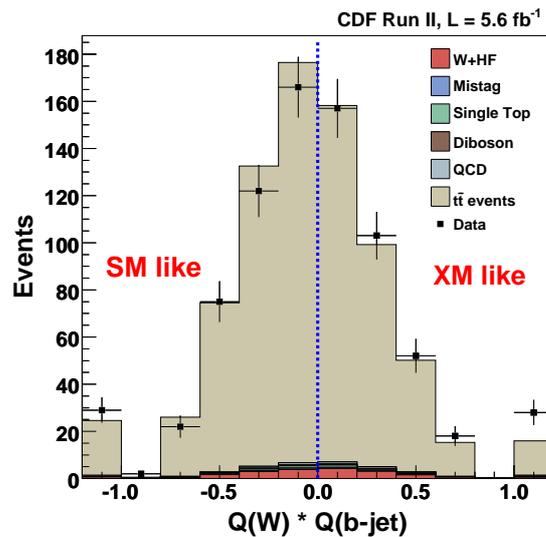
Obr. 5: Náboj W bozónu násobený nábojom zodpovedajúceho b-jetu. Páry reprezentujúce hypotézu SM nadobúdajú záporné hodnoty tejto veličiny, v prípade párov reprezentujúcich XM má daná veličina kladné hodnoty.