FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS
Comenius University in Bratislava

Department of Nuclear Physics and Biophysics

Determination of Top Quark charge in CDF experiment

PhD Thesis

Mgr. Peter Bednář

Supervisor: doc. RNDr. Stanislav Tokár, CSc.   Bratislava 2007

Contents

Chapter 1

Introduction

The present particle physics is concentrating mainly on the Standard Model theory [1] and search for new physics beyond.

The Standard Model (SM) is a theoretical conception based on calibration symmetries, that consist of the following parts:

- Theory of electro-weak interaction - it is a calibration theory with Lagrangean based on $SU(2){\otimes}U(1)$ calibration symmetry. The EW theory describes electromagnetic and weak processes.

- Theory of strong interaction - Quantum chromodynamic (QCD). It is based on $SU(3)$ calibration symmetry that lead to 3 color state of quarks, i.e. to 3 charges of strong interactions.

In SM there are three sectors of fundamental particles (matter constituents), for details see Table 1.1:

- fundamental fermions - 3 generations of leptons and quarks. They are spin 1/2 particles (i.e. obey Pauli principle) and are the basic constituents that create the discrete structures like nucleon, atom, etc.

- intermedial bosons - quanta of force field (γ - EM force, $W^{\pm}, Z$ - weak force, 8 gluons - strong force). They are spin 1 particles that mediate particle interactions.

- Higgs sector - in the minimal SM there is one neutral particle. Interaction with Higgs field gives mass to fermions and W and Z bosons.

From the all possible combinations of the SM quarks only the SU(3) color singlet mesons (combinations of $q\bar{q}$) and baryons (qqq-combinations) occur in nature (have been observed up to now). The quark family include the

Particles	Mass [MeV/c^2]	Spin	Charge/e	Color states
down (d)	4-8	1/2	-1/3	3
up (u)	1.5-4	1/2	+2/3	3
e^-	0.511	1/2	-1	0
ν_e	< 0.000003	1/2	0	0
strange (s)	80-130	1/2	-1/3	3
charm (c)	1150-1350	1/2	+2/3	3
μ^-	105.66	1/2	-1	0
ν_μ	< 0.19	1/2	0	0
bottom (b)	4100-4900	1/2	-1/3	3
top (t)	174300	1/2	+2/3	3
τ^-	1777	1/2	-1	0
ν_τ	< 18.2	1/2	0	0
γ	0	1	0	0
$W^\pm$	80419	1	± 1	0
Z	91188	1	0	0
gluon (g)	0	1	0	8
Higgs (H)	> 114000	0	0	0

Table 1.1: The experimental status of the elementary particles of the Standard Model

up and down quarks that make up the neutron and proton. The 6 leptons include the electron and its partner, the electron neutrino. With the recent observation of the tau neutrino at Fermilab, all 12 fermions and all 4 gauge bosons have been observed. Seven of these 16 particles (charm, bottom, top, tau neutrino, W, Z, gluon) were predicted by the SM before they were observed experimentally.

There is one additional particle predicted by the SM called the Higgs boson, which has not yet been observed. The Higgs field has non-zero vacuum expectation value, which gives masses to W and Z bosons as well as fermions. While photons and gluons have no mass, the W and Z are quite heavy. The W weighs 80.3 GeV ($\approx$ 80 times as much as the proton) and the Z weighs 91.2 GeV. The Higgs is expected to be heavy as well.

It should be noted that the SM does not include gravitational force as up to now theory failed to create a quantum theory of gravitation. This poses one of the main drawbacks of SM. It also fails to explain the hierarchy of particle masses and number of particle generations, does not explain satisfactorily the origin of CP violation and it has many free parameters (25 in full version of SM).

Several conceptions going beyond the frame of SM have been created, like the GUT - theories, theory of supersymmetry, technicolor or string theory. In experiments a big effort is devoted now to the search for a new physics. From this point of view the top quark physics appears to be very perspective. Therefor it is important to measure all its properties and to understand the mechanism of its production and decay.

1.1   Goals of the thesis

This thesis deals with the problematic of top quark charge measurement in CDF experiment at Fermilab. The goal is to determine if the top quark observed on Tevatron experiments is the Standard Model particle with the predicted charge 2/3 or it is some exotic 4th generation quark with the charge of -4/3 as suggested by some alternative theories.

The CDF top quark charge measurement uses two decay channels of $t\bar{t}$ pair for the charge estimation. The author's contribution was in so called lepton + jets channel (LJ), but to have the complete picture, also the work done in dilepton channel (DIL) is mentioned.

In the lepton + jets channel the author was responsible for following tasks:

- optimization study for χ^2 cut, κ coefficient (sec. 4.3.3) and use of the b-tagging algorithms and Top Mass Fitter modes (sec. 4.3.2)

- obtaining the fraction of mixing and secondary events in calibration of jet flavor tagging algorithm (sec. 4.4.3)

- background study for all L+J types of background except for QCD (sec. 4.5)

- systematic uncertainties estimation for Initial/Final state radiation, MC modeling and top mass (sec. 4.6)

- 1-jet bin test (sec. 4.7.2)

Chapter 2

Experimental Setup

2.1 Tevatron

2.1.1 Accelerator Chain

The Tevatron proton-antiproton ($p\bar{p}$) collider is located at Fermi National Accelerator Laboratory near Chicago/Illinois. The Main accelerator ring has a circumference of 6 km and operates at a center of mass energy of 1.96 TeV. The collisions occur at two interaction regions where the detectors D0 and CDF II are placed. For this thesis, the data collected by CDF II detector are used.

The figure 2.1 shows the accelerator chain used in Fermilab. For initial particle acceleration a chain of successive preaccelerators is used. In the first stage, the ionized hydrogen gas is accelerated to a kinetic energy of 750 keV inside the *Cockroft-Walton* device, from where the ions enter a linear accelerator (*Linac*) which uses oscillating electric fields to accelerate the ions to 400 MeV. Due to the geometry of the oscillating fields, the ions are grouped into bunches at the end of the Linac.

In the next phase, the ions travel into the *Booster* - a circular synchrotron 74.5 m in diameter - where they are directed onto a carbon foil target, stripping the hydrogen off of its electrons and leaving bare protons. After this, they are accelerated them to 8 GeV. The intensity of the proton beam is increased here by repeatedly injecting further protons into the same orbit as the already circulating ones. Protons are then extracted from the *Booster* into the *Main Injector*, a 3 km circumference synchrotron, where protons can be accelerated up to 120 GeV for antiproton production or up to 150 GeV for injection into the Tevatron.

Antiprotons are produced by directing the 120 GeV proton beam from *Main Injector* onto a nickel target where approximately 20 antiprotons with
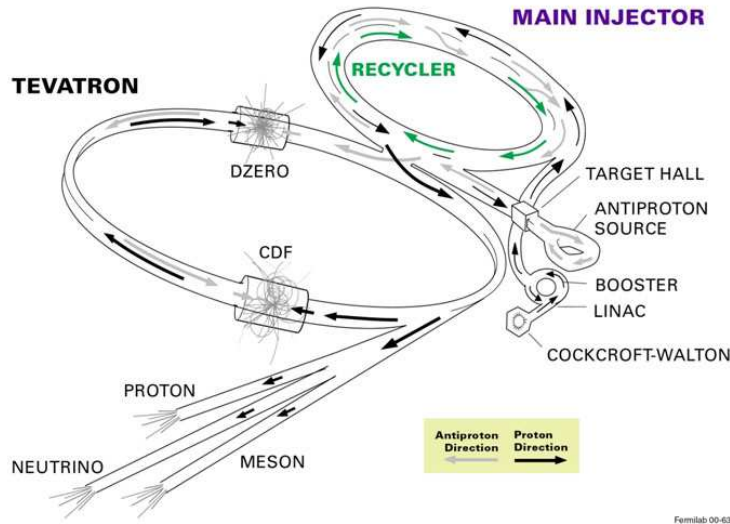
Figure 2.1: The Fermilab accelerator complex.

a kinetic energy of 8 GeV are produced per 10^6 protons. The antiprotons are separated by a pulsed magnet and focused by a lithium lens and accumulated in the 8 Gev *Accumulator Synchrotron* where the stochastic cooling is used to reduce the spread in the antiprotons kinetic energy spectrum before directing them to the *Main Injector*.

The Production rate of antiprotons is the limiting factor for operating at high luminosities. It takes about 15-20 hours to build up a stack of antiprotons that will be injected into the *Tevatron*. Approximately once per day the 36 bunches of stacked antiprotons ($\sim 90 \times 10^{10}$ particles) and protons ($\sim 3 \times 10^{11}$ particles) are injected from the *Main Injector* into *Tevatron* where the both beams are accelerated to the final energy of 980 GeV. The last part of the accelerator chain is *Antiproton Recycler*, housed in *Main Injector* ring. As the large part of the produced antiprotons survive the store, the original idea was to collect them, cool them down and keep in the *Antiproton Recycler* for the next store. The process of cooling has proofed to be very inefficient, so the idea of recycling was abandoned. Now, when there are $\approx 50^{10}$ antiprotons produced in the accumulator, they are tranfered to recycler and kept there, while the production continues. The reason behind this is, that the production efficiency decrease with increasing stack of antiprotons.

2.1.2 Performance

Since winter 2004 the *Antiproton Recycler* is fully commissioned and integrated into the Tevatron operation. As the production rate of antiprotons is the limiting factor to achieve high luminosities, the use of the Antiproton Recycler as a container for produced antiprotons resulted in a significant improvement of the performance in terms of peak luminosity and the Run 2 design luminosity of $\mathcal{L} = 10^{32} cm^{-2} s^{-1}$ could be reached. The luminosity of collisions can be written as:

$$\mathcal{L} = fF \frac{N_B N_p N_{\bar{p}}}{2\pi(\sigma_p^2 + \sigma_{\bar{p}}^2)} \qquad (2.1)$$

where f is the revolution frequency, F is a form factor describing the geometric properties of a bunch, N_B the number of bunches, $N_{p,\bar{p}}$ the number of protons/antiprotons per bunch and $\sigma_{p,\bar{p}}$ the RMS beam size at the interaction point.

The integrated luminosity over the time of a data taking period multiplied with the cross section σ of a given process yields an estimate for the predicted number of events $N = \sigma \times \int \mathcal{L}dt$.

Table 2.1 summarize the values of given parameters for Run 2. Figures 2.2 and 2.3 summarize the development of peak and integrated luminosity at the CDF II detector since the beginning of Run 2 in 2002.

Parameter	Value for Run 2
N_B	36
bunch length [m]	0.37
bunch spacing [ns]	396
protons/bunch (N_p)	$\sim 3 \times 10^{11}$
antiproton/bunch ($N_{\bar{p}}$)	$\sim 90 \times 10^{10}$
highest peak $\mathcal{L}$ $[cm^{-2}s^{-1}]$	292×10^{32}

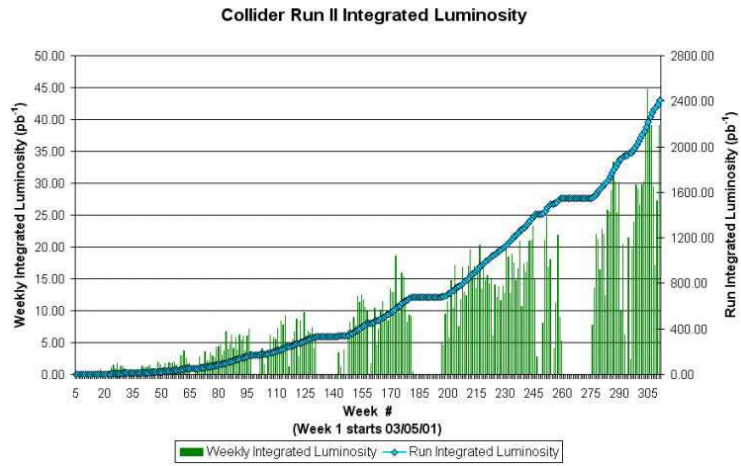Table 2.1: Accelerator parameters of the Tevatron.

Figure 2.2: Tevatron Run 2 integrated luminosity.
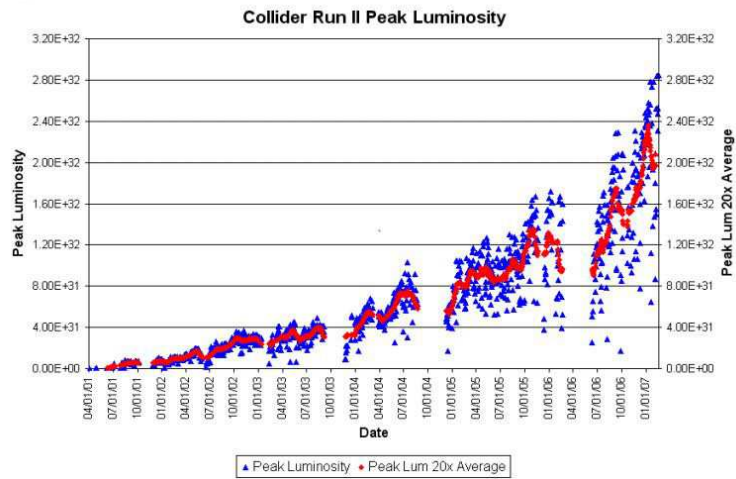

Figure 2.3: Tevatron Run 2 peak luminosity

2.2   CDF experiment

2.2.1   Physical goals of the CDF experiment in Run II

The analysis results of experimental data measured by CDF detector until 1996 (the Run I) can be found in nearly two hundred publications and it is considered an invaluable contribution to particle physics. The top quark was experimentally observed in 1995 with mass $m_t = 176.1 \pm 6.6 GeV/c^2$ []. The next interesting result is the precise electroweak measurement of W boson mass $m_W = 80.422 \pm 0.079 GeV/c^2$ []. An important part of the analysis studied the processes with b hadrons, such as mixing, CP violation etc. [].

The physical program in the first phase of Run II picks up where the previous analysis left off, with the additional benefit of a new detector (greater acceptance, better track reconstruction, more effective trigger etc.). The amount of data expected for the Run II is $4 - 8 fb^{-1}$. The main goals for $2 fb^{-1}$ of already collected data in the first period of Run II are []:

- determination of top quark properties

- precise electroweak measurements

- direct measurement of new processes

- testing of perturbative QCD

- determination of CKM matrix elements (especially V_{ts}) using high statistics of b hadron decays

- CP violation study

- b meson spectroscopy

The part of the CDF Run II results should be the first complete study of top quark using sample ≈ 1400 $t\bar{t}$ events. The top mass is expected to be measured with a precision of 3.5 GeV/c^2 and the cross section with a precision $\delta\sigma_{t\bar{t}}/\sigma_{t\bar{t}} < 9\%$. It is expected to isolate the electroweak production of top quark with determination of $\Gamma(t \to Wb)$ in order of 26% and determination of $|V_{tb}|$ with precision 13%.

One of the main goals of the "electroweak program" is the W boson mass measurement with a precision of ± 40 MeV/c^2. This is important in regard to our knowledge of top quark mass for precise electroweak tests of the Standard Model (SM) and to obtain the indirect bounds for the Higgs boson mass.

For the verification of SM predictions, the measurement of charge asymmetry A_{FB} for e^+e^- pair production in $p\bar{p}$ collisions can be used, because the

value of asymmetry is sensitive to possible contribution from processes not predicted by SM. The advantage of Tevatron experiment is the measurement of A_{FB} up to the invariant mass of e^+e^- pair $600 GeV/c^2$ [].

The study of processes with b hadrons should help to better specify five parameters of CKM matrix and the study of CP violation. The advantage of CDF experiment is the sufficient energy of collision for B_s^0 and $\bar{B}_s^0$ production. Using the measurement of mixing parameters of $B_s^0 - \bar{B}_s^0$ the elements $|\frac{V_{td}}{V_{ts}}|$ can be determined in the scale of SM prediction.

2.2.2 CDF detector - overview

The CDF II detector is a multipurpose collider detector [2] designed to detect and measure properties of particles being produced in $p\bar{p}-$collisions. It features a vertexing and tracking system, particle identification, a superconducting solenoid with length of 4.8 m diameter of 1.5 m generating a 1.4 T magnetic field, calorimetry and muon chambers. The components are arranged in the cylindrical symmetry typical for collider detectors. Fig 2.4 shows the side view of the CDF II detector. The z axis is placed in the center of the detector parallel to the beam. Polar angle θ is measured from the beam with the beginning in the positive half of the z axis (the direction of protons) and the pseudorapidity η is defined as

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right) \tag{2.2}$$

The last upgrade of the detector was performed in 1996 - 2001 [2]. The main goal was to upgrade the tracking system to improve the performance at higher luminosities and to increase the detector acceptance.

After the upgrade, the detector allows to reconstruct the tracks of charged particles with high efficiency for higher pseudorapidity (up to $|\eta| < 2.8$). The improved trigger system allows:

- trigger the events of interest better than in Run I

- better identify and reconstruct kinematic parameters of photons, electrons, muons and b-hadrons

- more efficiently determine the type of particle by measuring the energy losses and the time of flight for the area of pseudorapidity $|\eta| < 1.0$.
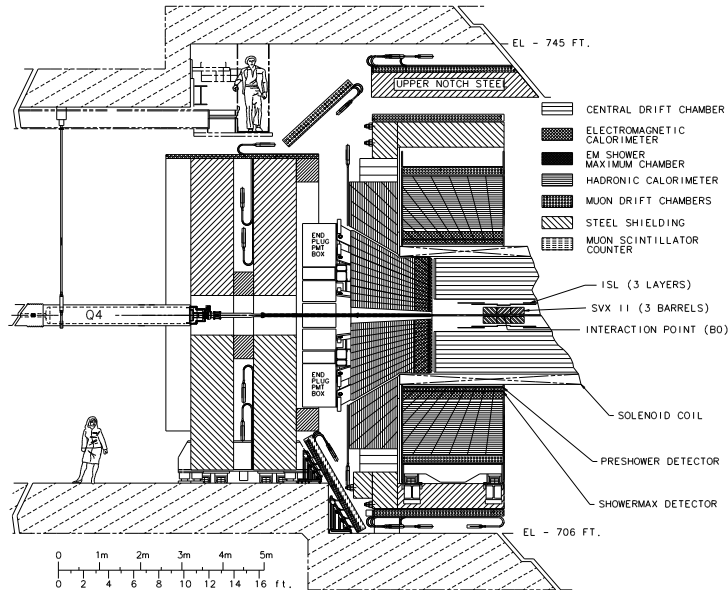
Figure 2.4: Sideview of the CDF detector

2.2.3 Tracking system

The CDF II tracking system consists of several subdetectors that can be used to detect charged particles and measure momenta and displacements with respect to the collision point (primary vertex). It consists of two detectors, the *Central Outer Tracker* (COT) and the *Silicon Vertex Detector* with its subsystems SVXII, ISL (*Intermediate Silicon Layer*) and L00 (*Layer00*). The tracking system is built cylindrically around the beam pipe and contained within the 1.4 T magnetic field created by a superconducting solenoid. Figure 2.5 shows the view of the tracking system.

The Silicon Vertex Detector of CDF II covers a radial range from r = 1.6 cm to r = 28 cm, provides an acceptance up to $|\eta| < 2.0$ and consists of 7 - 8 layers of double and single sided silicon sensors. The innermost single sided layer closest to the beam line at a radius of 1.6 cm is called Layer00. Outside of L00 from r = 2.1 cm to r = 17.3 cm follows SVXII, a five-layered double-sided silicon microstrip detector. The strips are aligned axially on one side, with 90-degree stereo on the other side for layers 0, 1 and 3 and small angle stereo (1.2 degrees) for layers 2 and 4. Between SVXII and the Central Outer Tracker (COT), the *Intermediate Silicon Layer* enhances linking of tracks measured in both detectors. The ISL central layer is located at r = 22 cm and forward/backward layers are at r = 20 cm and r = 28 cm. The central layer covers $|\eta| < 1.0$ and the forward/backward layers cover
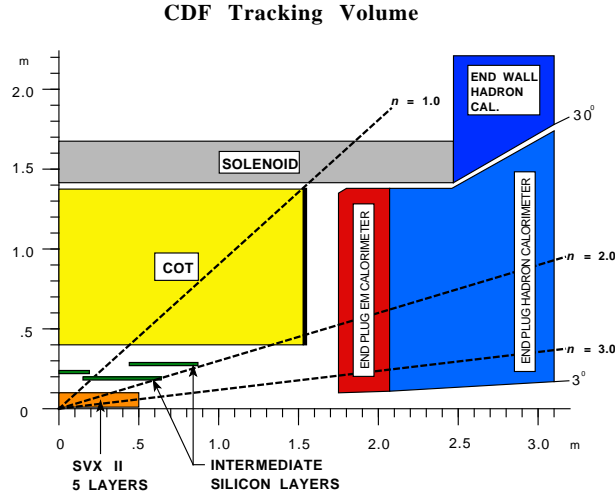
13

Figure 2.5: CDF tracking system - layout

$1.0 < |\eta| < 2.0$. The ISL is a double-sided silicon microstrip detector (axial on one side, small-angle stereo on the other side). The layout of SVX detector system is illustrated in Figure 2.6.

The ensemble of silicon vertex detector devices is particularly suited to measure the impact parameter d_0 and the azimuthal angle ϕ of tracks with high precision and can therefore determine the decay vertex of long lived unstable particles with high spacial resolution.

The *Central Outer Tracker* (COT) is a cylindrical drift chamber covering the radial range from r = 33 cm to r = 143 cm. The COT contains 96 wire layers grouped into eight superlayers, where in four superlayers the wires run in z-direction ("axial superlayers") and in the other four are tilted by 2 degree with respect to the z-direction ("stereo superlayers"). The chamber operates with an Ar-Ethane-CF_4 gas mixture. Particles passing through the COT in the range $|\eta| < 1$ pass through all eight superlayers, particles up to $|\eta| < 1.3$ pass through at least four superlayers and ionize the gas mixture. Electrons drifting toward the sense wire induce an electronic signal that is read out by an ASDQ (amplifier, shaper, discriminator and charge encoding) chip [3]. Due to geometrical reasons the track density in the COT tracking volume is significantly smaller compared to the SVX, in addition the COT produces up to 96 hit measurements per tracks, so it provides an accurate measurement of a particle's transversal momentum (in the r-ϕ plane) and less precise information of the particle's momentum in the r-z plane. Figure 2.7 shows the geometry of the COT endplate, Figure 2.8 shows the cell layout for superlayer 2.

14
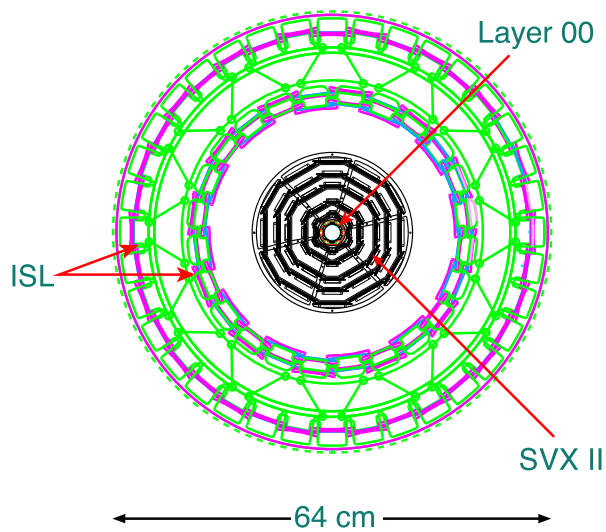
Figure 2.6: Silicon Vertex Detector - endview
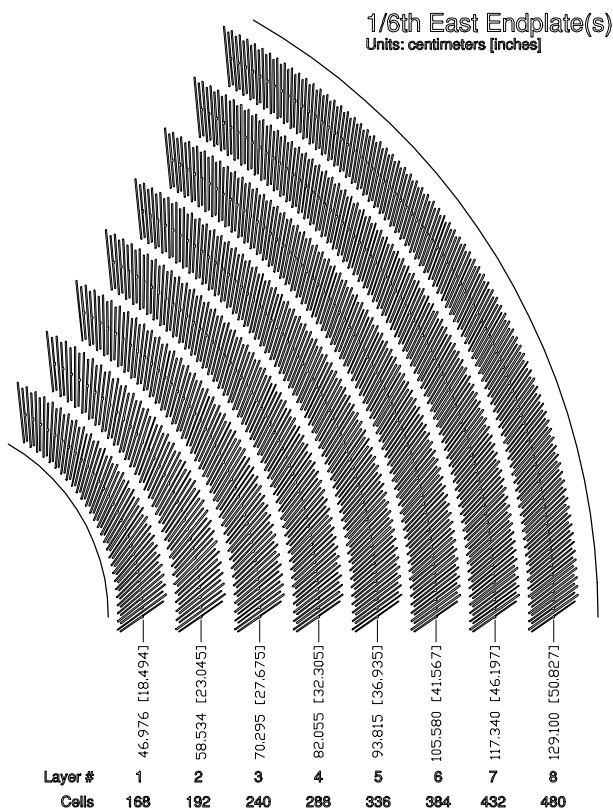
Figure 2.7: 1/6 of COT endplate
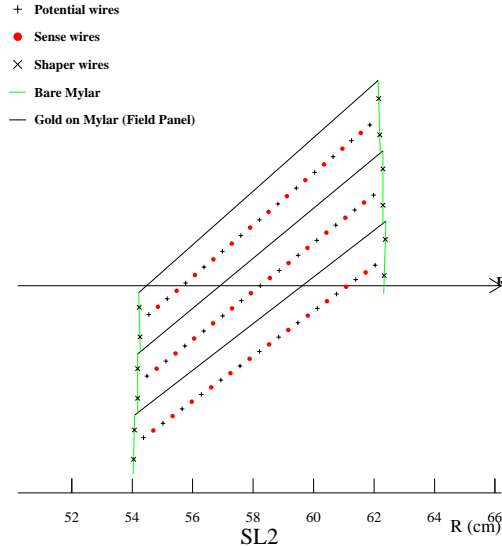
Figure 2.8: Cell layout for superlayer 2

2.2.4 Calorimeters

The CDF II calorimetry system consists of several independent calorimeters: the central and end-plug electromagnetic and hadronic calorimeters and the hadronic end-wall calorimeters. The calorimetry system covers the pseudo-rapidity region $|\eta| < 3.6$ and is used to measure the energy of high energy electrons, photons and hadronic jets.

Central calorimeter consists of towers constructed alternating passive lead and active scintillating parts, where the signal is being detected. Central electromagnetic calorimeter with thickness of $19X_0$ (radiation length) has resolution

$$\frac{\sigma(E)}{E} = \frac{14.0\%}{\sqrt{E}} \oplus 2\% \tag{2.3}$$

There are wire chambers in the calorimeter located the approximate position of maximum of electromagnetic shower, which give the information about the placement of the shower and helps to suppress the cases of wrong identification of electrons. The other chamber placed in front of the calorimeter in the role of preshower detector, helps in distinguishing between electrons and photons. Figure 2.9 shows one segment of the central electromagnetic calorimeter.

Hadronic part of the central calorimeter consists of 23 layers of metal and
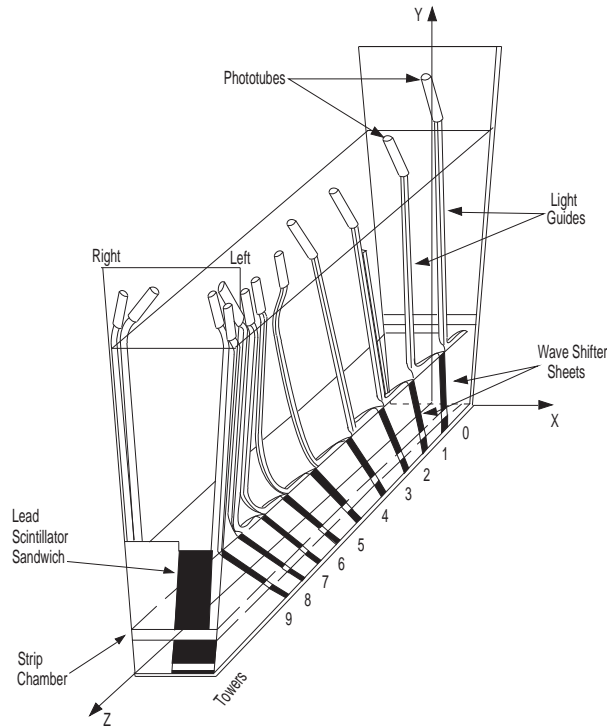
Figure 2.9: One wedge of the central electromagnetic calorimeter
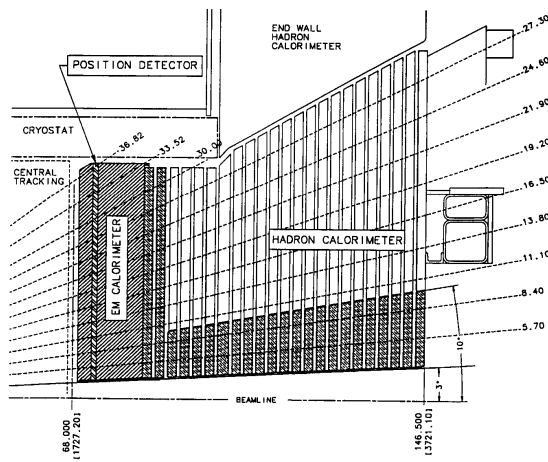
Figure 2.10: The upper part of the forward calorimeter

17

scintillating plates. Projective geometry is the same as in the electromagnetic part. The hadronic calorimeter with thickness of $\sim 4.5\lambda_I$ (interaction length) has the resolution

$$\frac{\sigma(E)}{E} = \frac{50.0\%}{\sqrt{E}} \oplus 3\% \tag{2.4}$$

In the forward area $1.1 < |\eta| < 3.64$ the same technology as in central area was used, for the homogeneity of data and the response.

Forward calorimeter is divided in r-ϕ plane in modules with $\delta\psi = 15$ deg. In respect to pseudorapidity, the module consists of 12 layers. Up to pseudorapidity $|\eta| < 2.11$ it is segmented into 16 towers, in area of $2.11 < |\eta| < 3.6$ into four as can be seen in figure 2.10.

As an absorber the lead is used, alternating with scintillators with thickness of 4 mm. The first active layer functioning as preshower detector is thicker (10mm) and clearer. It is read separately from the other layers of calorimeter.

The forward electromagnetic calorimeter with thickness of $21X_0$ has the resolution

$$\frac{\sigma(E)}{E} = \frac{16.0\%}{\sqrt{E}} \oplus 1\% \tag{2.5}$$

Also in this calorimeter, the detector for measuring the coordinates of the shower maximum is placed in about six radiation lengths (behind fourth absorber plate). It consists of two layers labeled as U and V. Each layer is made of 5 mm thick scintillating strips covering the area of ϕ angle of 45 deg. The layers are displaced in respect to module for $+22.5$ deg and -22.5 deg and give two-dimensional information about the shower coordinates with precision of 1 mm.

The forward hadronic calorimeter covers the area of pseudorapidity $1.30 < |\eta| < 3.64$ with thickness of $7\lambda_I$ and its resolution

$$\frac{\sigma(E)}{E} = \frac{80.0\%}{\sqrt{E}} \oplus 5\% \tag{2.6}$$

is limited mainly by sampling fluctuation in metal absorber. The main attention by the construction of calorimeter was paid to the homogeneity of signal.

2.2.5 Muon detectors

The CDF II muon system consists of detectors mounted radially outside the calorimetry system. The parts of the central system are *Central Muon De-*

tector (CMU), Central Muon Upgrade (CMP), Central Scintillator Upgrade (CSP), Central Muon Extension (CMX) and Central Scintillator Detector (CSX). The CMU consists of 144 modules, each module of 16 square cells with dimensions $6.35 \times 2.86 \times 225 cm^3$ [4]. Each cell is a proportional chamber. CMU is placed directly after hadronic calorimeter, which absorbs 99% of charged hadrons. Around this detector (covering the area up to $|\eta| < 0.6$) the cells of CMP detector are placed with additional shielding of steel plates 60 cm thick. The cell of CMP detector are one-wire proportional chambers arrayed in four layers. On the edge of CMP (with respect to point of interaction) is CPS detector made of scintillating plates. The CMX and CSX detectors extend the coverage from $|\eta| < 0.6$ to $|\eta| < 1.0$ with the gap at $\phi = 30 \deg$ because of cryogenic system and the gap at $\phi = 90 \deg$ because of the floor of the hall in which the CDF detector is placed. The CMX consists of conically shaped drifting chambers. The coverage of muon detectors in $\eta - \phi$ plane is shown in Figure 2.11.

In the forward region, the *Intermediate Muon Detector (IMU)* is used. For its shielding against hadrons, the toroidal magnet from previous construction of CDF without energetization is used with additional steel plates. For this detector again the combination of drifting chambers and scintillating plates is used. The schematic geometry of IMU barrel is show in Figure 2.12. Most of the particles that reach the muon system without producing a shower and being absorbed in the calorimeters are actually muons, since they interact with matter mainly via ionization. Kaons and pions surviving the calorimeter passage produce muon misidentification background at the level of a few percent.
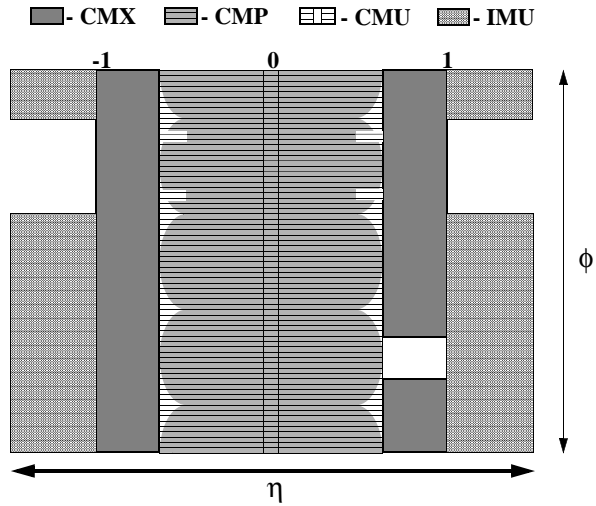
Figure 2.11: $\eta - \phi$ coverage of muon detectors
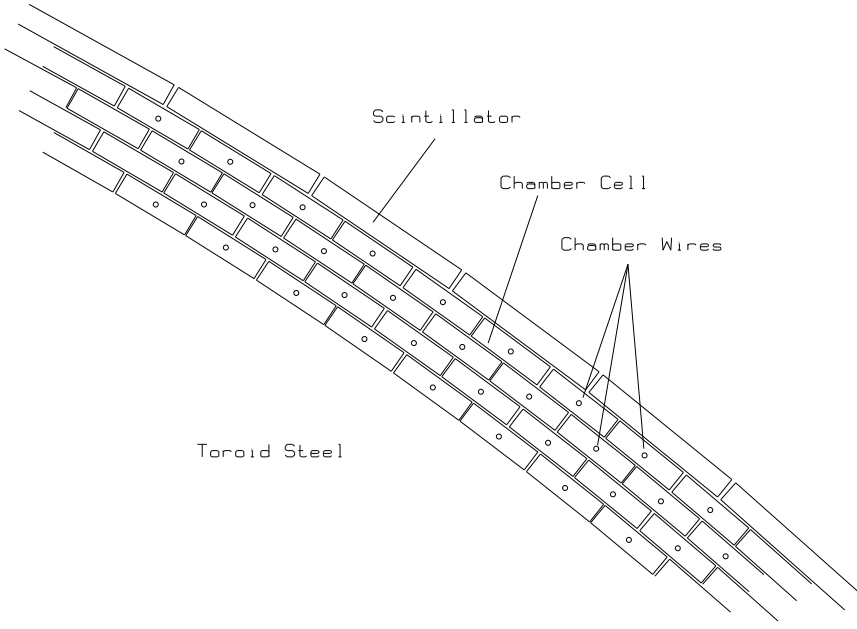
Figure 2.12: Geometry of IMU barrel with detail of drifting chambers and scintilators

2.3 Particle identification

The collision of a proton and an antiproton can result in creation of a large variety of final states particles. This section describes the requirements imposed on the information gathered from different sub-detectors to efficiently identify the different particles from $t\bar{t}$ decay in the high energy range.

2.3.1 Electron identification

High energy electrons can be identified by a high-p_T track in the drift chamber and large energy depositions in electromagnetic calorimeters. The geometry of the CDF Run 2 detector imposes the categorization of electrons in two types according to the calorimeter in which the deposition takes place.

Central electrons

Central electrons transverse the central part of the detector, $|\eta| < 1.1$ leaving a high-p_T track in COT and depositing their energy in the CEM calorimeter. To identify central electrons the following requirements are applied:

- $E_T = E \cdot \sin\theta > 20$ Gev

  where E is the total electromagnetic energy deposited by the electron in the CEM cluster, θ is the angle of the COT track pointing to the seed tower of the cluster. The electron cluster is formed of a seed EM tower - defined as the tower where most of the energy is deposited - and a number of shoulder towers which are added to the seed tower until the maximum cluster size is reached. The largest cluster spans two towers in pseudorapidity and on tower in azimuth. The energy E is corrected to account for non-linear effects and known differences in response between the different towers, as measured from $Z \to e^+e^-$ candidates.

- $p_t > 10$ GeV/c

  where p_t is the transverse momentum of the COT track. The resolution of the track is improved by constraining the track to originate from the beam line position.

- $E_{had}/E < 0.055 + 0.00045 \cdot \text{GeV}^{-1} \cdot E_{total}$

  The ratio of the energy deposited in the hadronic calorimeter cluster to what is deposited in the electromagnetic cluster. The second therm

21

in the right side compensates for effect of electron showers leaking into the hadronic calorimeter at very high energies.

- $E/P < 2$

  Ratio of the electromagnetic calorimeter energy to the momentum of the track as measured from the COT track. This cut helps reducing the number of jets that fake electrons. These jets typically contain a π^0 that deposits its energy in the electromagnetic calorimeter. This cut also helps discard those electrons that radiated a high energy photon. The photon is typically collinear with the electron track and generally deposits its energy in the same calorimeter tower and therefore the cluster energy does not differ much from the original energy of the electron. The p_t of the track however does significantly changes, and the ratio E/P cam be much larger than one.

- $L_{shr} < 0.2$

  This quantity uses the lateral shower profile and compares it to what is expected for electromagnetic showers. It is defined as

  $$L_{shr} = 0.14 \sum \frac{E_i^{measured} - E_i^{expected}}{sqrt(0.14 sqrtE)^2 + \sigma^2_{E_i^{expected}}} \qquad (2.7)$$

  where the i runs over towers, $E_i^{measured}$ is the energy measured in tower i and $E_i^{expected}$ is the energy expected from test beam data. The error in the energy measurement is represented by $0.14 sqrtE$ and $\sigma^2_{E_i^{expected}}$ is the uncertainty in the energy estimate. Typically L_{shr} is a two-tower sum. Any extra particles accompanying the one responsible for the main EM shower will tend to add to the energy in adjacted towers and make L_{shr} a larger number.

- -3.0 cm $< Q \cdot \Delta x < 1.5$ cm; $|\Delta z| < 3$ cm

  The distance between the COT track extrapolated to the CES and the best matching CES cluster are represented by Δx and Δz. This requirement imposes a close match between the two. The cut on Δx has been multiplied by the charge of the track Q and it is asymmetric in $r - \phi$ plane to account for possible photon bremsstrahlung in the direction of the outside of the track.

- $\chi^2_{strip} < 10$

  This quantity compares the shower profile in the shower maximum detector, CES, with the shower profile obtained from test beam measurements.

- $|z_{vertex}| < 60$ cm

  The position in z-coordinate of the primary vertex interaction. The z_{vertex} is determined by the intersection of the track with the beam axis. The longitudinal spread of the event vertex about the nominal interaction point $z = 0$ is a Gaussian with $\sigma = 26$ cm. The vertex position is required to be within 2σ to avoid tracks to pass through uninstrumented regions of the detector.

- Track quality cuts

  A well reconstructed track should have at least 7 hits in each of at least three axial and three stereo superlayers of the COT.

- Fiduciality

  This variable requires that the electron is reconstructed in a region of the CDF detector that is well instrumented.

- Conversion veto

  The interaction of photons with the detector material can result in the conversion of electron-positron pairs. These conversions can be identified by the presence of another electron candidate with opposite charge near the electron candidate. If both tracks are close on ϕ at the point of conversion the electron candidate is flagged as conversion.

- Calorimeter isolation: $E_t^{iso}/E_t^{cluster} < 0.1$

  The ratio of the transverse energy in the electromagnetic and hadronic calorimeter in a cone of radius $\Delta R = sqrt(\Delta\eta)^2 + (\Delta\phi)^2 < 0.4$ excluding the electron energy, to the electron energy. This cut rejects electrons that are not isolated from the hadronic activity, including those coming from the semi-leptonic quark decays.

The efficiency of the central electron identification cuts is determined from $\gamma/Z^* \to e^+e^-$ data sample. Events were required to have opposite charged electrons, with an invariant mass of the electron pair in a ±15 GeV/c^2 range from the nominal Z mass. One electron was required to pass all the cuts, while the other was used to obtains the efficiency of the cuts. The identification efficiency was found to be $\epsilon_{data} = (82.5\pm0.2)\%$. To compensate for differences in the reconstruction in data and MC a scale factor of $\epsilon_{data}/\epsilon_{MC} = 0.965\pm0.006$ is applied to correct acceptances obtained from MC.

Plug electron

Electron candidates depositing energy in the PEM calorimeter are called *plug electrons*. The following requirements are applied to identify a plug electron :

- $1.2 < |\eta| < 2.0$

  The PEM allows the identification of electromagnetic clusters of energy at pseudorapidities up to $|\eta| < 2.5$. Electron candidates with $|\eta| > 2$ have, however, a large charge misidentification rate and the cut $|\eta| < 2.0$ is used. This requirement has a small effect on the acceptance of $t\bar{t}$ events since these are mostly central.

  Plug electrons identified solely based on a cluster in the forward calorimeter are called PEM electrons. The charge misidentification rate of PEM can be very large as forward electrons may not traverse a large part of the COT, and track information is limited. The track efficiency is improved using an algorithm named the Phoenix algorithm [5].

  The algorithm starts by associating the primary vertex and the center of the energy cluster observed in the PEM. Two tracks are constructed based on the two possible charges of the electron. The algorithm next looks for hits in the silicon that matches the tracks. If hits are matched, the track (called 'the Phoenix track') is reconstructed. If both tracks are reconstructed, the one with higher quality of fit is taken. Electron candidates with a cluster in the PEM and a Phoenix track are called PHX electrons.

- $E_t > 20$ GeV

  The transverse energy of the PEM cluster, which is limited to two towers in pseudorapidity and two towers in azimuth, must be greater than 20 GeV.

- $E_{had}/E < 0.05$

  Similar to the central electron requirement, the ratio of energy deposited in the hadronic calorimeter cluster to what is deposited in the electromagnetic cluster.

- $U_{5 \times 9} \geq 0.65$ and $V_{5 \times 9} \geq 0.65$

  The variables $U_{5 \times 9}$ and $V_{5 \times 9}$ are isolation variables for the shower maximum detector. They are independently applied to both the U and V layers. The clustering in each layer is performed by ordering strips in decreasing energy with the highest energy strips used as seeds. The

PES cluster has a fixed width of nine strips. The quantities $U_{5\times9}$ and $V_{5\times9}$ represent the ratios of energy sum in the central 5 strips to the total energy in all nine strips.

- $\chi^2_{3\times3} < 10$

  This variable compares the energy distribution in the 3×3 PEM towers around the seed tower to what was obtained from test beam electrons by means of a χ^2 test statistics.

- $|\Delta R_{PES}| < 3$ cm

  This variable compares the position of the shower obtained from the $\chi^2_{3\times3}$ fit to the intersection of the centroids in the layers U and V. The difference in pseudorapidity $\Delta\eta$ and in azimuth $\Delta\phi$ determines $(\Delta R_{PES})^2 = (\Delta\eta)^2 + (\Delta\phi)^2$.

- $N^{Si}_{hits} \leq 3$

  The number of hits in the silicon detector.

- $|z_{vertex}| < 60$ cm

  Same as the central electron requirement.

- Calorimeter isolation: $E^{iso}_t / E^{cluster}_t < 0.1$

  Same as the central electron requirement, the ratio of the transverse energy in the electromagnetic and hadronic calorimeter in a cone of $\Delta R < 0.4$ excluding the electron energy, to the electron cluster energy.

The plug electron identification efficiency is determined from $\gamma/Z^* \to e^+e^-$ data sample. A central electron and a plug electron are required with the invariant mass of the two-electron system in a 40 GeV/c^2 window around the Z mass. The plug electron identification efficiency is found to be $\epsilon_{data} = (65.1 \pm 0.8)\%$. The same calculation in MC yields $\epsilon_{MC} = (74.9 \pm 0.2)\%$. To compensate for differences in the reconstruction in data and MC a scale factor of $\epsilon_{data}/\epsilon_{MC} = 0.87 \pm 0.01$ is applied to correct acceptance obtained from MC.

2.3.2 Muon identification

Muons are minimum ionizing particles which penetrate large amounts of material with little energy loss. Their passage through the detector is characterized by a track in the COT, small energy deposition in the calorimeter and hits in the muon chambers.

Muons at CDF are categorized by the muon subdetectors through witch they pass. Those reconstructed in the CMU detector are called CMU muons etc. Muons reconstructed in both the CMU and CMP detectors are called CMUP muons. In particular central tracks that are not expected to pass through any of the muon detectors are called Central Minimum Ionizing Objects or CMIO muons. They are also called stubless muons, to indicate that no stub was reconstructed in the muon detector.

For muon identification, the following requirements are applied:

- $P_t > 20$ GeV/c

  The transverse momentum of the COT beam-constraint track.

- $|z_0| < 60$ cm

  The z-position of the track.

- $d_0 < 0.2$ cm for tracks with no silicon hits.

  $d_0 < 0.02$ cm for tracks with silicon hits.

  The impact parameter d_0 is the distance from primary vertex to the the track in transverse plane. This selection is of course used for the default muon track, i.e. without the beam constrain requirement. This cut forces the muon to originate form the nominal interaction point and substantially reduces cosmic muons that entered the detector in the time window of the collision.

- Track quality cuts

  Identical to the electron quality tracks. A well reconstructed track should leave a signal in at least three axial and three stereo superlayers of the COT, at least 7 hits in each one of them.

- $E_{em} \leq 2$ GeV $+max(0, 0.115(P - 100$ GeV/$c))$

  It requires the energy deposited in the electromagnetic calorimeter to be very small. A small change is introduced for muons with $P > 100$ GeV/c.

- $E_{had} \leq 6$ GeV $+max(0, 0.0280(P - 100$ GeV/$c))$

  Requires the energy deposited in the hadronic calorimeter to be smaller than that of strongly interacting jets.

- $E_{em} + E_{had} > 0.1$ GeV for stubless muons only

Stubless muons are required to have non-zero energy deposition in the calorimeter to limit background from electrons escaping the detector through non-instrumented regions of calorimeter.

- $|\Delta X|_{CMU} \leq 3$ cm, $|\Delta X|_{CMP} \leq 5$ cm and $|\Delta X|_{CMX} \leq 6$ cm

  The distance in the $r - \phi$ plane between the COT track extrapolated to the stub segment and the position of the reconstructed stub for the muons in different sub-detectors. The requirements are lower for CMP and CMX since the muons traverse more material and the effect of multiple scattering is larger.

- Fiduciality

  This variable requires that the track of the muon candidate extrapolates to the proper muon chamber. For stubless muons it requires the track to pass through well instrumented regions of the CDF detector and through no muon sub-detector.

- $\rho_{COT} > 140$ cm

  The radius at which the track leaves the COT. This track is used for CMX only, as tracks may leave the COT from the side. In general, data and MC simulation do not agree for $\rho_{COT} \leq 140$ cm due to the data bias introduced by the XFT trigger which requires the track to have hits in at least 4 COT superlayers. Data and MC agree for $\rho_{COT} > 140$ cm.

- Isolation: $E_t^{iso}/P_t < 0.1$

  The quantity E^{iso} is energy of calorimeter towers in the cone of $\Delta R = 0.4$ without including the tower associated with the track.

- Not a cosmic muon

  Muons coming from cosmic background are identified as dimuon events with an angular separation close to $180\,\text{deg}$ in ϕ. Cosmic rays enter the detector at random times and random locations. As cosmic muons traverse the detector, they leave energy in the calorimeter, traverse through the COT leaving a single track that is reconstructed as two and leave energy in the calorimeter as it go out. The time difference between the energy deposited in both ends of the hadronic calorimeter can be analyzed to identify cosmic muons. In addition a cut in the impact parameter of the track efficiently identify cosmic muons. Events in which a cosmic muon is tagged are rejected.

27

As with electrons the muon identification efficiencies are measured using the $\gamma/Z^* \to \mu^+\mu^-$ data sample. One muon is required to have strict cuts an is associated with the Level-1 trigger. The other muon is chosen to be fiducial (i.e. the track points to the proper sub-detector) and examined to see if passes all the identification cuts.

Leptons satisfying all the criteria mentioned above are called tight leptons. Leptons satisfying all criteria but isolation are called loose leptons.

2.3.3 Jet reconstruction

In a typical $p\bar{p}$ collision quarks and gluons are created. These carry color charge and are therefor subjected to the hadronization process, where original parton results in a stream of colorless particles. These particles are collimated along the direction of the original parton and after traversing the tracks deposit their energy in a cluster of towers in the calorimeter detector. This stream of particles is called a *jet*.

Jets are reconstructed from the energy deposited in the calorimeter towers with an algorithm called *jet clustering*. This algorithm starts by identifying the seed tower as the one with the largest calorimeter energy. The cluster of towers within the cone $\Delta R = sqrt(\Delta \eta)^2 + (\Delta \phi)^2 = 0.4$ from the center of the seed tower are identified. The cone size $\Delta R = 0.4$ is chosen to include most of the jet energy without including a large contribution from other event activity.

After the cluster is formed, the shower center in the plane (η, ϕ) of the calorimeter is determined as follows:

$$\eta_{centroid} = \frac{\sum_i E_t^i \eta^i}{\sum_i E_t^i} \qquad \phi_{centroid} = \frac{\sum_i E_t^i \phi^i}{\sum_i E_t^i} \tag{2.8}$$

where the sum runs over the towers in the cluster and η^i, ϕ^i represent the η, ϕ coordinate of the tower. After the position of the shower is calculated a new cone is defined centered around the new axis and the calculation is repeated for the new cluster. This process is then iterated until the cluster remains unchanged.

The transverse energy of the jet is determined as

$$E_t^{raw} = \sqrt{(\sum_i E_i \sin \theta_i \cos \phi_i)^2 + (\sum_i E_i \sin \theta_i \cos \phi_i)^2} \tag{2.9}$$

where E_i is the total energy in tower i. This quantity represents the energy deposited in the cluster and does not include corrections to account for detector effects or other physics processes and is thus referred to as *raw*.

The corrected value of the jet E_t is derived from the raw value using a set of multiplicative and additive correction factors resulting in corrected values that are close to the energy of the initial parton. The corrections are given by the following relation:

$$E_t = (E_t^{raw} \times f_{rel} \times f_{time} \times f_{scale} - E_T^{MI}) \times f_{abs} - E_t^{UE} + E_t^{OC} \qquad (2.10)$$

The seven factors are briefly explained bellow.

- Relative correction f_{rel} - factor accounting for the relative tower-to-tower differences in calorimeter response.

- Time dependent correction f_{time} - factor accounting for the calorimeter response declination with time. It depends on the date the event was recorded.

- Energy scale correction f_{scale} - factor accounting for the non-linear response of the calorimeter

- Multiple interaction E_t^{MI} - with current luminosities one interaction per bunch crossing is expected on average. The distribution follows Poisson statistics and in a particular event this number can be larger and the physics process under study may have an overlapping minimum bias event. The energy of the minimum bias event may fall in the jet clustering cone and must be subtracted. This is calculated on an event-by-event basis.

- Absolute energy correction f_{abs} - this factor corrects the energy of the jet to that of the original parton that generated it. It includes nuclear absorption and particle leakage effects that can reduce the total energy measured in the jet cluster.

- Underlying event correction E_t^{UE} - extra energy can be deposited in the jet cluster that is not coming from the original parton. This energy can originate from spectator partons (beam remnants) and bunch crossing.

- Out of cone corrections E_t^{OC} - part of the energy of the initial parton may fall outside the cone resulting in an underestimation of its energy. This quantity is obtained from MC simulation of the physics process in study.

Each level of correction has its own systematic uncertainty contributing to the total systematic uncertainty of the jet E_t measurement.

Jets with $E_T > 15$ GeV and $|\eta| < 2.0$ are called thight jets. Jets with $E_T > 8$ GeV and $|\eta| < 2.4$ are called loose jets.

2.3.4 Missing transverse energy

Although with an unknown magnitude, the momentum of the initial partons is known to be in the beam direction. Momentum conservation thus requires that the energy of all the final states balances in the plane transverse to the beam line.

Many interesting physics processes contain neutrinos in their final state. Neutrinos escape the detector, some of them carrying significant amounts of energy. The identification of the presence of a neutrino is possible by the detection of imbalance in the transverse energy. The transverse energy E_t is a two-dimensional vector defined as:

$$\vec{E}_t^{\,raw} = \sum_i E_i^{raw} \vec{n_i} \sin \theta_i \qquad (2.11)$$

where the sum is over all the calorimeter towers, E_i^{raw} is the raw energy of each tower, θ_i is the polar angle to the center of the tower and $\vec{n_i}$ is a transverse unit vector pointing to the center of the tower.

The value of $\vec{E}_t^{\,raw}$ defined above is corrected on an event-by-event basis. The first correction is related to the muons in the event. The amount of energy deposited by muons in the calorimeter is significantly smaller than the energy they carry. The estimation based on the calorimeter is then corrected to include the energy of the muon as measured by its track information. A second correction is related to the jets in the event. Since the jet cluster energy is corrected as explained above, in the calculation of the $\vec{E}_t^{\,raw}$ the raw energy of the towers in the cluster has to be replaced by their corrected energy. The corrected resulting transverse energy is referred to as $\vec{E}_t$.

This procedure is done for technical reasons, as there are no jet energy correction included in standard format of data. The jet energy corrections are applied later up to certain level based on the needs of particular analysis.

Since the energy is assumed to be balanced, it is often more useful to talk in term of the "missing energy" defined simply as $\vec{\not{E}}_t = -\vec{E}_t$. If neutrinos are present in the final state, the $\vec{\not{E}}_t$ indicates the direction and magnitude of the vector sum of the neutrinos transverse energies in the event.

Chapter 3

Top Quark physics

3.1 Status of Top Quark physics

The top quark was first experimentally observed in the experiments CDF and D0 on Tevatron accelerator at Fermilab in 1995 [6, 7]. Top quark is produced mainly by strong interactions in $t\bar{t}$ pairs created by quark-antiquark ($q\bar{q}$) annihilation or gluon fusion (gg) (fig 3.1). In addition to that, the top quark can be (under the SM) also produced through electroweak interactions (fig 3.6).

$$q\bar{q} \to t\bar{t} \qquad \qquad gg \to t\bar{t}$$
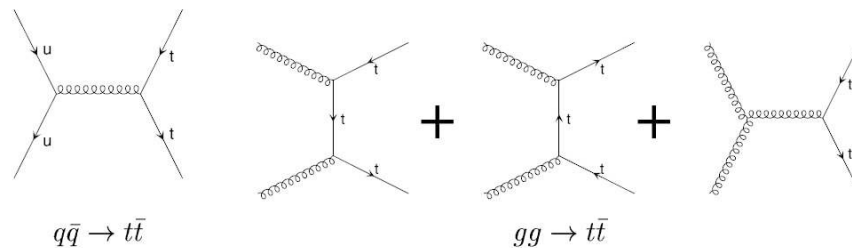
Figure 3.1: Feynman diagrams for $t\bar{t}$ production via strong interaction

At the time of the discovery, the top mass was determined to be $m_t \approx 174$ GeV/c^2. For this mass the SM predicts the decay width $\Gamma \approx 1.4$ GeV [8] and the half-life $\tau \approx 10^{-24}s$.

The top quark is an interesting object for testing of SM and looking for a new physics beyond SM. Its main features are

- Top quark mass is very high - it is near the electroweak symmetry breaking scale, therefor top quark can play an important role at EW symmetry breaking.

31

- Its great mass in comparison either to the rest of the quarks or to the W and Z bosons makes top quark a very good perturbative object - it is produced at small distances, the effective coupling constant is small and therefor perturbative expansion converges rapidly, which can be used for tests of QCD theory.

- It decays before hadronization - its spin characteristics is not diluted by hadronization. By studying the spin correlations and polarization of top decay products (W-polarization) the V-A structure of electroweak interactions can be tested and possible manifestation of new physics can be looked for.

- Under the SM it decays almost exclusively to b-quark and W-boson. Through determination of the V_{tb} matrix element, the unitarity of CKM matrix can be tested with a possible consequence on the 4th generation of fundamental fermions.

- Precise knowledge of top mass leads, through higher order EW corrections, to a restriction on the mass of Higgs boson.

- The top production processes are an important background for Higgs boson production.

3.2   Top quark production

In experimental investigation of the top quark production it should be taken into account that partons (quarks and gluons) cannot be experimentally observed directly – they can manifest themselves only through hadrons. To retrieve information about the partons the QCD factorization theorem is used [9, 10]. The basic idea of the factorization is schematically shown in Fig. 3.2.

The factorization theorem enables to express experimentally measured cross section as a convolution of parton distribution functions (PDF) of colliding hadrons and theoreticaly calculated parton cross sections.

$$\sigma_X = \sum_{ij} \int dx_1 dx_2 f_i^{(1)}(x_1, \mu_F) f_j^{(2)}(x_2, \mu_F) \hat{\sigma}_{ij \to X}(\hat{s}; \mu_F, \mu_R) \qquad (3.1)$$

where $f_i^{1,2}$ are parton distribution functions of parton of type i in colliding hadrons 1 and 2; $\hat{\sigma}_{ij \to X}$ – partonic cross section for the measured output
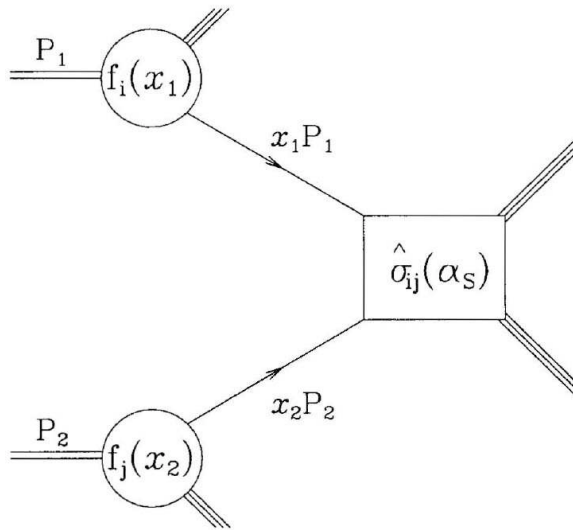
Figure 3.2: Parton structure of hadron-hadron interaction: the colliding hadrons with momenta P_1 and P_2 interact through their partons i and j carrying momenta x_1P_1 and x_1P_1.

channel X; μ_F and μ_R are the factorisation and renormalization scales, respectively; x_1 and x_2 are momentum fractions of the interacting partons in hadrons 1 and 2, respectively and $\hat{s}$ ($= x_1x_2s$) is the square of partonic CMS collision energy.

The PDF functions, that are not only functions of parton momentum fraction but also of factorization scale (see below), are retrieved from a global fit of the deep inelastic scattering and Drell-Yan data and the DGLAB evolution equation [11] is used to find the values of these functions at other (higher) scales. The parton splitting function that are used in the DGLAB equation are now calculated in the NNLO approximation. The partonic cross section, $\hat{\sigma}$, can be expanded in the running strong coupling costant $\alpha_S(\mu_R)$:

$$\hat{\sigma} = \alpha_S^k \sum_{m=0}^{n} \alpha_S^m c^{(m)} \tag{3.2}$$

where the $c^{(m)}$ are functions of the partonic process kinematic variables and the factorization scale μ_F. In Eq. 3.2 $n = 0$ coresponds to the leading order (LO), $n = 1$ to the next-to leading order (NLO), $n = 2$ to the next-to-next-to leading order (NNLO), etc. The leading power coefficient k depends on the type of hard process, e.g. for the di-jet production $k = 2$. At the expansion a problem of so called big logaritmic contributions arise (an increased pro-

duction in a certain phase space region) and the resummation procedures are used to take them correctly into account.

The factorization scale, μ_F, and the renormalization one, μ_R, are in principle arbitrary parameters in the expansion 3.2, but a rapid convergence of the expansion requires proper choice of their values for investigated process. The former one (μ_F) can be thought of as the scale that separates the short and long-distance interaction and determines the parton structure of colliding hadrons, while the latter one (μ_R) determines the value of the running coupling constant.

3.2.1 $t\bar{t}$ production cross-section

Thanks to the great mass of the top quark, the precise calculation of $t\bar{t}$ production cross-section in frame of perturbative QCD is possible. The calculations of cross-section includes all the leading order (LO) and next to leading order (NLO) terms along with inclusion of the resumation of next to leading logarithmic order (softly emitted gluons). [12]. Attempts are made to include the next to next leading order (NNLO) with higher logarithmic terms (NNLL, NNNLL) [13, 14, 15].

The predicted value of $t\bar{t}$ production cross-section (in 1.96 TeV $p\bar{p}$ collisions for $m_t = 175$ GeV/c^2) is $6.70^{+0.71}_{-0.88}$ pb [16]. The average value measured by CDF collaboration in Run2 is $\sigma = 7.3 \pm 0.5 \pm 0.6 \pm 0.4$ pb, which was obtained as a weighted average of all the analysis in all top decay channels (Figure 3.3). This measurement is consistent with the expectation and doesn't suggests any non-considered effects. The Figure 3.4 shows the theoretical cross section dependence on top mass compared to CDF measurement.

As mentioned before, $t\bar{t}$ pairs can be produced in two ways: quark antiquark annihilation ($q\bar{q} \rightarrow t\bar{t}$) and gluon fusion ($gg \rightarrow t\bar{t}$). At Tevatron, the $q\bar{q}$ annihilation is dominant process ($\approx 85\%$ [17]). At LHC accelerator at CERN, the gluon fussion will dominate ($\approx 90\%$ [18]). Figure 3.5 shows the top mass dependence of the predicted $t\bar{t}$ production cross section for Tevatron and LHC.

3.2.2 Electroweak top quark production

In SM top quark is produced also individually via electroweak interaction (so called single-top production). The leading-order Feynman diagrams are shown in Figure 3.6. The cross section of this kind of production is lower than for $t\bar{t}$. At Run 2 of Tevatron, the SM expectations for single-top production are 0.9 pb via the s-channel process and 2.0 pb via the t-channel. Compared to $t\bar{t}$ production, fewer events are expected on the higher background.
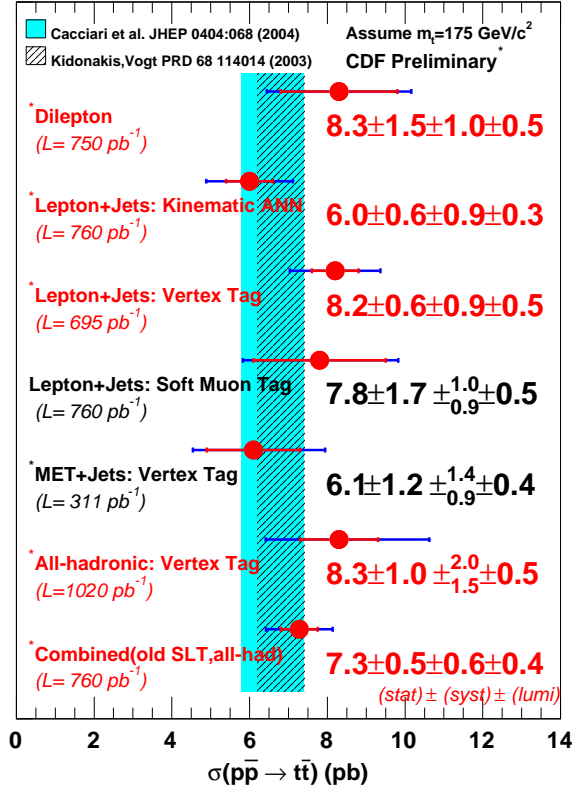
Figure 3.3: The CDF cross section measurements
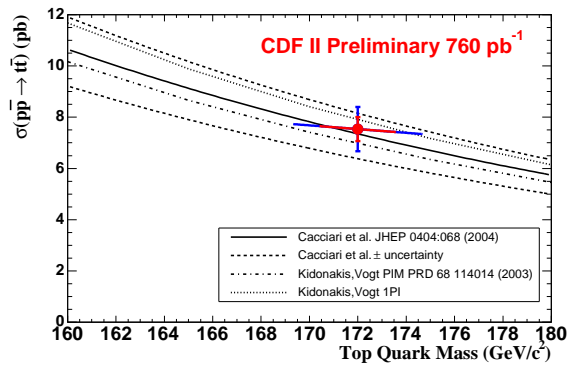
Figure 3.4: The production cross section dependence on top mass
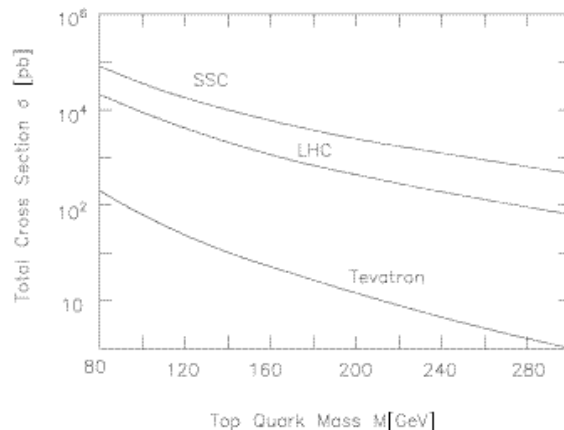
Figure 3.5: Comparison of the predicted $t\bar{t}$ production cross section for Tevatron and LHC
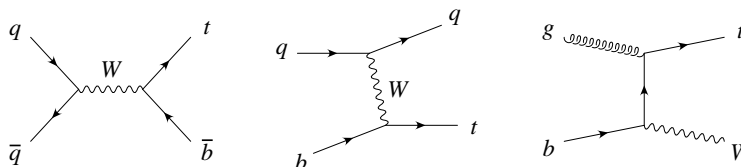
Figure 3.6: Some of the leading-order Feynman diagrams for single-top production

Although electroweak top quark production cross section has not been determined in the Tevatron experiments yet, both CDF and D0 observe the EW production at more than 3σ level [19].

3.3 Top quark decay

3.3.1 Top quark decay modes

Because of the short life time of the top quark, it decays before it can hadronize. The top events (events with top quark) can be categorized by the final decay product configuration. The dominant is the electroweak decay to b quark and W boson, the electroweak decays to s or d quark are significantly suppressed (the corresponding elements of CMK matrix are small). The W boson also lives very short time and it is observable only via its decay products. It can decay either into two quarks ($W \to q\bar{q}$) or leptonically ($W \to l\nu$). From experimental point of view it is convenient to introduce three different

36

categories of the $t\bar{t}$ pair decay (so called decay modes), accordingly to the W bosons decays [20]:

1. Hadronic mode (44% of all $t\bar{t}$ decays) - both of the W bosons from $t\bar{t}$ pair decay hadronicaly

2. Semileptonic mode (30% of all $t\bar{t}$ decays) - one of the W bosons decay hadronicaly, the other one leptonicaly (only decay to e, μ taken into account)

3. Dilepton mode (5% of all $t\bar{t}$ decays) - both of the W bosons decay leptonicaly (only decay to e, μ taken into account)

- The W boson decays into τ ($W \rightarrow \tau\nu$) are not taken into account because the τ can decay both hadronicaly and leptonicaly. The techniques to identify such decays are rather complex, therefore this decay mode is not used for most top quark analyses.

Hadronic mode

Although this decay mode is the most common, its disadvantage is the big QCD background [17]. To select the signal, high transverse momentum p_t of jets is requested. The signal to background ration can be improved with increase of required number of jets and by use of b-tagging (procedure to identify jets coming from b quarks).

Semileptonic mode

This mode is also called the lepton plus jets mode. In this mode, an isolated lepton with high transverse momentum p_t and missing transverse energy (indicating neutrino) is required. The signature of this event is also characteristic by high transverse momentum jets (two coming from W decay and two from b quark decay). The b-tagging requirement is not necessary, on the other hand it helps to significantly reduce the background. The background processes include mainly QCD, W production with many jets and diboson (WW, WZ, ZZ) production.

Dilepton mode

For selection of this mode, two isolated leptons with high transverse momentum and large missing transverse energy (from two neutrinos) are required. Although this decay channel is the rarest, its advantage is a significantly reduced QCD background. The other background processes for this mode include Drell-Yan ($Z^*/\gamma \to e^+e^-, \mu^+\mu^-$), $Z \to \tau\tau$ and diboson production.

3.3.2 Top quark decay width

The top quark decay width was not measured by experiment, but it can be theoretically calculated from the SM. As the decay $t \to Wb$ is by far the most dominant (because of structure of CKM matrix), in the first order of perturbative theory we obtain [21]

$$\Gamma_t = \frac{\alpha_{em}}{4}\frac{m_t}{2\theta_W^2}|V_{tb}|^2\frac{\kappa}{m^2t}\Big[\frac{m_t^2+m_b^2}{2m_t^2} + \frac{(m_t^2-m_b^2)^2}{2m_t^2m_W^2} - \frac{m_W^2}{m_t^2}\Big] \qquad (3.3)$$

where

$$\kappa = \sqrt{(m_t^2 - (m_W^2 + m_b^2)^2)(m_t^2 - (m_W^2 - m_b^2)^2)} \qquad (3.4)$$

α_{em} is the electromagnetic coupling constant, θ_W is electroweak mixing parameter, V_{tb} is the CMK matrix element, m_t, m_W, m_b are the top quark, W boson and b quark masses.

3.3.3 Determination of $|V_{tb}|$

In CDF experiment, there was an attempt to estimate the ratio

$$B = \frac{\Gamma(t \to Wb)}{\Gamma(t \to Wq)} \approx \frac{|V_{tb}|^2}{|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2} \qquad (3.5)$$

Using the b-tagging method, the value of $B = 0.94^{+0.31}_{-0.24}$ [22] was obtained. If assuming three generations of quarks and CKM matrix unitarity, this value is consistent with $V_{tb} = 0.999 \pm 0.003$ [22]. Another possibility to determine the value of V_{tb} is to measure the electroweak production cross-section of top quark (single top production).

The Direct determination of $V_{b\bar{b}}$ is important as it allows to test the unitarity of CKM matrix and thereby it is an instrument of the search for the manifestation of the fourth generation of fundamental fermions.

3.3.4 *W* polarization in top decay

In top decay the longitudinally polarized W boson are produced. In approximation of massless b-quarks and W-bosons, the fraction of longitudinal polarization would be $f_0 = 100\%$. However, because of the non-zero mass of the W-boson, SM which is based on left handed electroweak currents, gives for the unpolarized top quark decay the relative ratio of W-boson polarization in the top quark system $f_0 = 0.7$, $f_- = 0.3$ and $f_+ = 0$ [23]. The results of Tevatron experiments are:

- CDF: $f_0 = 0.85^{+0.15}_{-0.12}(stat) \pm 0.06(syst)$; $f_+ = 0.05^{+0.11}_{-0.05}(stat) \pm 0.03(syst)$ [24].

- D0: $f_0 = 0.56 \pm 0.31$; $f_+ = 0.0 \pm 0.13$ [25]

Both results are compatible with SM, but more precise results are needed as the W-polarization is a very important tool for searching for traces of a new physics.

3.4 Mass measurements

The mass of the top quark is a free parameter of SM. Its precise measurement along with the W-boson mass set also restrictions for the mass of Higgs boson through electroweak high order corrections. Figure 3.7 shows the theoretical predictions for Higgs mass together with the current experimental results from W-boson and top quark mass measurements.

As in the case of cross-section, also here all three top decay channels were used. Published Run 1 measurements were combined with available preliminary Run 2 measurements. The final combined result is $m_t = 171.4 \pm 1.2(stat) \pm 1.8(sys)$ GeV$/c^2$. The individual contributions from both D0 and CDF measurements are shown in Figure 3.8, while all CDF Run 2 top mass measurements available are shown in Figure 3.9.

Before the discovery of top quark at Tevatron, its mass estimates from electroweak data from LEP were done using higher order calculation for precisely measured EW observables (M_W, Γ_Z, $\sin\Theta_W$). The Z boson decay width was calculated with 1-loop corrections from the virtual correction of diagrams shown in Figure 3.11. From Z boson decay width measurement (Figure 3.10) the top mass $m_t = 160 \pm 15 \pm 15$ GeV$/c^2$ was calculated [26]. The first uncertainty comes from experimental uncertainty of Γ_Z measurement, the second includes the uncertainty of strong coupling constant and of Higgs boson mass which was considered to be somewhere in the region of 60 GeV$/c^2 < m_H < 1000$ Gev$/c^2$.
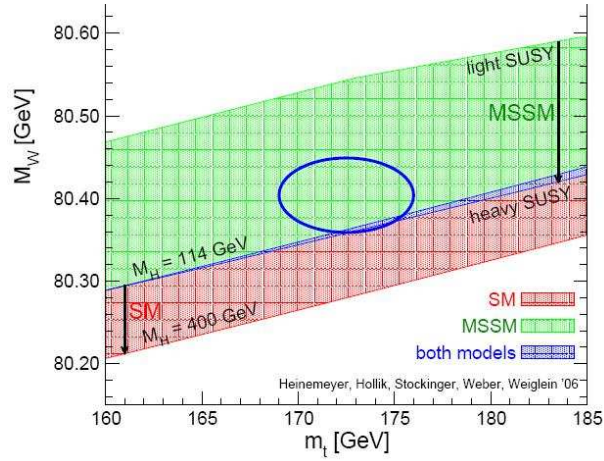
Figure 3.7: Higgs mass dependence on W-boson and top quark mass
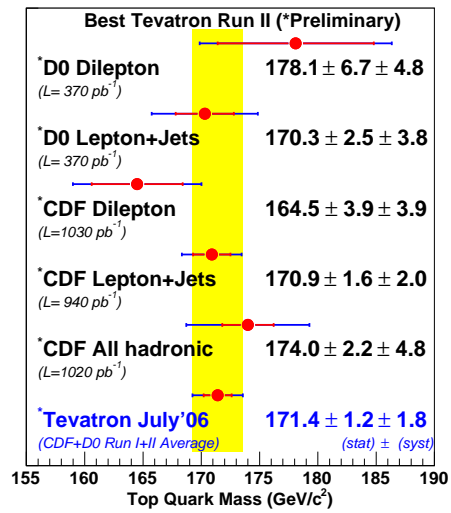
Figure 3.8: Combined result from Tevatron top mass measurements

40
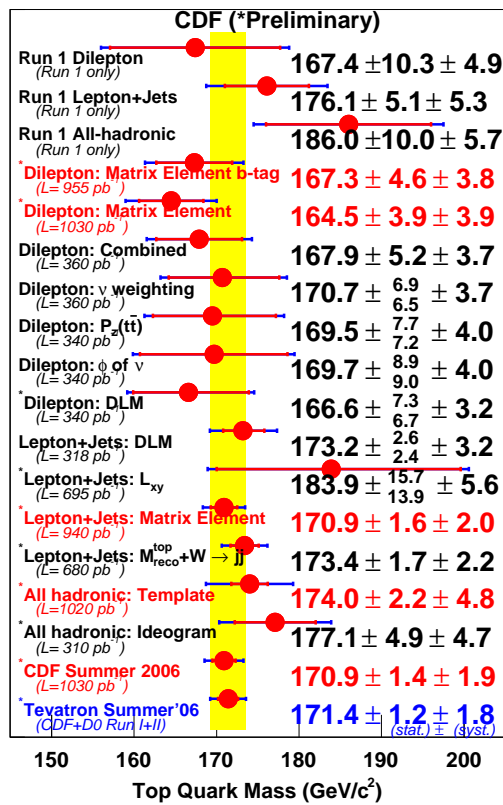
Figure 3.9: The combined result from the CDF Run 2 top mass measurements
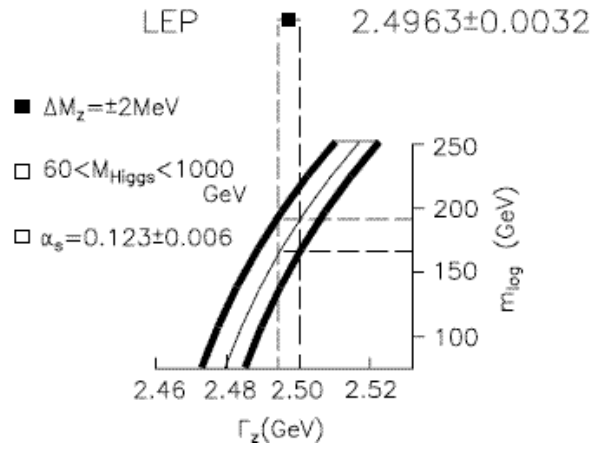
Figure 3.10: The top mass determination from Γ_Z
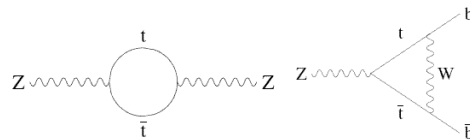
Figure 3.11: Some of the Feynman diagrams contributing to Z boson decay width

The consistency of m_t values from LEP data using EW approach and Tevatron experiments using QCD indicates that if there is any effect coming from new particles, it is relatively small .

42

Chapter 4

The determination of Top Quark charge

4.1 Introduction

Since the discovery of the top quark, CDF has made several analyses to confirm that it has the properties expected in the SM - decay fraction into different final states, Lorentz structure of the $t \to bW$ vertex, mass, etc. One of the properties still pending is the electric charge.

The electromagnetic couplings can be measured using $t\bar{t}\gamma$ events [27], although this measurement needs more data than is available on CDF. Alternatively, one can reconstruct the top charge from the charge of its decay products.

The SM top quark is expected to have the charge $+2/3$. However, there is a theory suggesting another explanation for the CDF Run 1 experimental data [28]. This theory states that the quark with the mass $m \approx 170$ GeV/c^2 discovered at Tevatron is not the expected SM top quark, but an exotic quark of a doublet $(Q_1, Q_4)_R$, where Q_4 has the charge $-4/3$ and Q_1 that mixes with the right component of b-quark, has the charge -1/3. The SM top quark with the charge 2/3 should have the mass $m_t = 274$ GeV/c^2 in this scheme. Using such an exotic quark qives better results in fitting electroweak data from precision measurements at LEP, SLAC and Tevatron colliders.

To reconstruct the top charge from the charges of its decay products in $t\bar{t}$ events, the following ingredients are needed:

- correctly associating the W boson with the right b-jet from the same top quark decay

- determining the charge of the W boson (from leptonic decay $W \to l\mu$)

- determining the flavor of the b-jet (i.e. to find out if it is initiated by b- or $\bar{b}$-quark)

The necessary steps for obtaining the needed informations are described in this chapter.

4.2 Event selection

For the top charge analysis we have used only the di-lepton and lepton + jets samples. The fully hadronic sample has not been used due to a huge QCD background expected in this case.

4.2.1 Dilepton event selection

The event selection for dilepton channel follows the selection used in the top cross section analysis based on the data coresponding to integrated luminosity $1.2 fb^{-1}$ [29].

- Two isolated tight leptons with $E_T > 20$ GeV with opposite sign.

- At least two jets with high transversal energy, $E_T > 15$ GeV

- high missing transversal energy, $\vec{\not{E}}_t > 25$ GeV

- $H_T = p_{Tlep} + E_{Tjet} + \vec{\not{E}}_t > 200$ GeV

- Z veto. The event is vetoed if the lepton and another object have an invariant mass with energy between 76 Gev and 106 GeV. If the lepton is an electron, the other object can be an electromagnetic object, a jet or a track of an opposite charge particle. If the lepton is muon, the other object can be a minimum-ionizing track of opposite charge.

- One jet tagged as b-jet using Tight SecVtx tagging algorithm [31].

4.2.2 Lepton + jets event selection

The event selection in the lepton + jets channel follows the selection used in the top cross section analysis based on the data coresponding to integrated luminostity $1.12 fb^{-1}$ [30].

- One tight lepton - electron or muon with $p_T > 20$ GeV.

- $\vec{\not{E}}_t > 20$ GeV.

- Dilepton veto. The event is removed if additional loose lepton is found.

- Z veto. The same as in case of dilepton event selection.

- Cosmic veto. The event is vetoed if the lepton is a muon identified as cosmic.

- Conversion veto. The event is vetoed if the lepton is an electron identified as coming from a photon conversion.

- At least 4 tight jets or 3 tight jets and one loose jet .

- Two jets tagged as b-jets using Loose SecVtx tagging algorithm [31].

4.3   Optimization of parameters

4.3.1   Definition of performance factors

The analysis of the top charge consists of several algorithms (lepton and b-jet pairing, b-jet flavor tagging, etc.) each of them containing parameters that need to be optimized. In order to be able to do this, there is a need for quantitative criteria for the best options. The variables that were chosen for this analysis are:

- efficiency - the number of events remaining after certain selection criteria over the number of events available before applying the cut

- purity - the number of events that are correctly identified (based on MC information) over the number of events remaining after the cut

There is a trade-off between efficiency and purity. It is good to have as many events as possible to have a statistically significant measurement, but at the same time, to have a many of wrongly assigned events will dilute the measurement.

Formally, if the N_+ is defined as a number of events in favor of SM hypothesis and N_- as number of events in favor of the exotic hypothesis, then the measured asymmetry is given by:

$$A_{meas} = \frac{N_+ - N_-}{N_+ + N_-} \tag{4.1}$$

while the true asymmetry is given by:

$$A = \frac{N_+^t - N_-^t}{N_+^t + N_-^t} \tag{4.2}$$

where N_+^t is the number of true SM events and N_-^t is the number of true exotic model events. The assumption is that both the SM hypothesis and the exotic one can occur in paralel. If the dilution is defined to be:

$$D = \frac{N_{right} - N_{wrong}}{N_{right} + N_{wrong}} \tag{4.3}$$

where N_{right} and N_{wrong} are the number of rightly and wrongly assigned events, then the true purity can be expressed as

$$A = \frac{A_{meas}}{D} \tag{4.4}$$

If there were no wrongly assigned events, the dilution would be 1 and the measured asymmetry would be the same as the true asymmetry. In real experiments the dilution is between 0 and 1. The $D = 0$ means that the analysis does not have the power to distinguish between two hypothesis. The dilution is related to the purity $P = N_{right}/(N_{right} + N_{wrong})$ by $D = 2P - 1$. It can be shown that the statistical uncertainty of the true asymmetry is given by:

$$\sigma_A = \sqrt{\frac{1 - D^2 A}{\epsilon D^2 N}} \qquad (4.5)$$

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 was chosen as the quantitative criterium for parameter optimization.

4.3.2   Pairing between lepton and b-jet

As was mentioned before, for the reconstruction of top charge, there is a need to correctly associate lepton coming from W boson decay with the b-quark from the same top quark decay. As the event characteristics is not the same for dilepton and lepton + jets channels, different methods for lepton b-jet pairing are used. In dilepton channel, the criteria for correct pair selection is based on invariant mass of the lepton - b-jet system. In lepton + jets channel, the selection is based on complete kinematic fit of the event. Each method will be described in more details.

Lepton b-jet pairing in dilepton channel

For the selection of the right lepton-jet pair in dilepton channel the squared value of invariant mass of the pair M_{lb}^2 was used. Based on the event selection criteria, in the final state of $t\bar{t}$ decay there are two leptons and two or more jets, at least one of them tagged as b-jet. When ordered according to transverse energy, the first two jets with highest E_T are considered as b-jets where one of them is b-tagged. If the b-tagged jet is not among this two jets, the event is not used.

In dilepton events, with 2 b-jets and 2 leptons, there are two possible combinations of lepton b-jet pairing and for each combination two lepton-jet invariant masses can be reconstructed. Figure 4.1 shows the M_{lb}^2 distribution for correct (lepton and b-jet come from the same top quark) and incorrect (lepton and b-jet come from different top quarks) lepton b-jet pair. The distribution for the correct pair is relatively narrow with very small tail toward
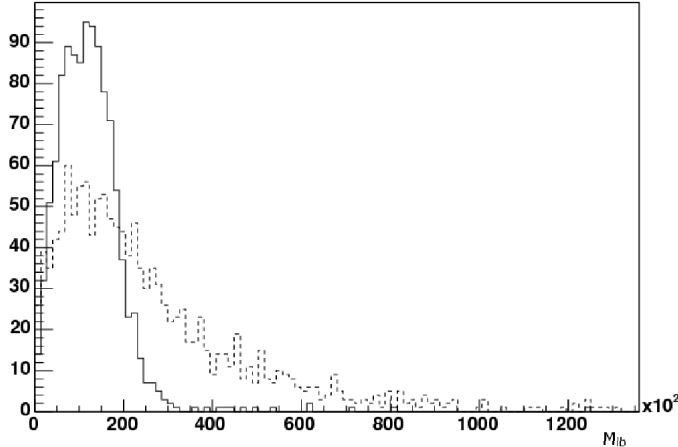
Figure 4.1: M_{lb}^2 distribution for correct lepton b-jet pair (solid) and incorrect (dashed) one

the high values and with the maximal possible value of M_{lb}^2 to be around $25000 \ GeV^2/c^4$. On the other hand, the distribution for the incorrect pair is broader with large tail.

To optimize the selection method from the purity point of view, the four values of M_{lb}^2 existing in this case are ordered, and the combination that does not contain the highest of the 4 M_{lb}^2 values is chosen as the right one. With this prescription the efficiency of the selection is 100% and the purity is 69%, which gives $\epsilon D^2 = 0.14$.

The way to increase the purity is to set a lower limit for the M_{lb-max}^2. By the M_{lb-max} is denoted the maximal of the 4 lepton b-jet invariant masses in dilepton event. The events where the M_{lb-max}^2 is less than the selected limit are refused, since for these events it is hard to select the right combination. Table 4.1 summarizes the values of ϵD^2 for various cuts on M_{lb-max}^2 where b-tagging is used. Table 4.2 shows the same situation without considering b-tagging information. The numbers in this table were obtained using Monte-Carlo. Conclusion from this optimization study is to set the cut and keep only events where $M_{lb-max}^2 > 21,000 \ \mathrm{GeV}^2/c^4$.

One can observe that the ϵD^2 values are slightly higher when there is no b-tagging. However the impact of background must be considered to justify the decision for using the tagging information. When considering background, the full purity is given by

$$p_{full} = f_{sig}p_{sig} + (1 - f_{sig})p_{bg} \qquad (4.6)$$

49

where f_{sig} is the fraction of signal events and p_{sig}, p_{bg} are the purities of signal and background, respectively. Later it will be shown that the f_{sig} is about 56% when b-tagging information is not used and is 1 when it is used, and that p_{bg} is 0.5. The $\epsilon D^2 = 0.099$ without b-tagging and $\epsilon D^2 = 0.185$ with b-tagging.

M^2_{lb-max} cut	ϵ	P	D	ϵD^2
no cut	1	0.686 ± 0.004	0.374	0.140
5,000	0.990 ± 0.001	0.688 ± 0.004	0.376	0.140
10,000	0.859 ± 0.003	0.722 ± 0.004	0.444	0.169
15,000	0.627 ± 0.004	0.812 ± 0.004	0.624	0.244
21,000	0.369 ± 0.004	0.959 ± 0.003	0.918	0.311
30,000	0.226 ± 0.003	0.994 ± 0.001	0.988	0.221

Table 4.1: Performance of M^2_{lb-max} cut for dilepton events with b-tagging

M^2_{lb-max} cut	ϵ	P	D	ϵD^2
no cut	1	0.693 ± 0.003	0.386	0.149
5,000	0.992 ± 0.001	0.692 ± 0.003	0.384	0.146
10,000	0.860 ± 0.002	0.727 ± 0.003	0.454	0.177
15,000	0.628 ± 0.003	0.816 ± 0.003	0.632	0.251
21,000	0.374 ± 0.003	0.960 ± 0.002	0.920	0.317
30,000	0.232 ± 0.002	0.993 ± 0.001	0.986	0.226

Table 4.2: Performance of M^2_{lb-max} cut for dilepton events without b-tagging

Lepton b-jet pairing in lepton + jets channel

In the lepton + jets channel a different strategy for the pairing is used. It is based on the event by event kinematic fit that is used in the top quark mass measurements. The essence of this approach is in full reconstruction of event topology.

Generaly in lepton + jets channel there is one lepton and at least four jets, two of them tagged as b-jets in the final state of $t\bar{t}$ decay. The b-jets are assumed to come directly from top quark decay, non-b-jets come from W boson decay and gluons (from initial and final state radiations). In this channel one of the W bosons decays into lepton and neutrino, the other into two light quarks. The top quark, which the leptonicaly decaying W comes from, and its decay products are called leptonic branch of the $t\bar{t}$ decay. The top quark, which the hadronicaly decaying W comes from, and

its decay products are called hadronic branch. The *b*-jets are called leptonic or hadronic according to which branch they belong to.

If there are more than 4 jets in a candidate event, only 4 of them with the highest transverse momenta are taken into account for finding the event topology by the kinematical fit. Ignoring the experimental inormation about b-tagging, there are 12 ways how to choose 2 *b*-jets from the 4 jets of event. On top of this, the sign of neutrino momentum z-component is not known therefore there are altogether 24 possible topological combinations for arangement of jets (assigning each of the jets the position of either leptonic or hadronic *b*-jet or one of the light quark jets coming from *W* decay) and neutrino z-direction. For each of these combinations, the full kinematic fit of the event is done using TopMassFitter (part of the CDF analysis tools package) yielding the value of χ^2 given by the expression

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

where the first term considers the difference in the transverse momenta between fitted ($\hat{p}_T^i$) and measured (p_T^i) transverse momentum values of the jets and lepton with σ_i (i=l,jet) to be the lepton and 4 leading jets p_T resolutions, the second term considers the difference between the fitted and measured components of the unclustered energy 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 ($m_{l\nu}$) and the remaining terms calculate the mass difference between top quark (m_t) and its decay products in hadronic branch (m_{bjj}) and leptonic one ($m_{bl\nu}$), respectively. Each of the terms is divided by the according decay width Γ_W and Γ_t respectively. At the event kinimatic fit the values of W mass and decay widths Γ_W and Γ_t were fixed at their PDG values The top mass, in addition to the constraint $m_t = m_{\bar{t}}$, was usualy restricted to its nominal value but sometimes (for the study of systematics) was left to be unconstrained.

As the information about *b*-tagging is part of the Top Mass Fitter input, in the end each of the 24 combination does have a flag, which tell if the combination satisfies the tagging requirement - meaning that if the jets tagged as *b*-jets (using SecVtx tagging algorithm) are in position of either leptonic or hadronic *b*-jet in the final configuration, the combination is labeled as a "tagged" one. In case of this analysis, where two *b*-tagged jets are required, there are only four tagged combinations. As the right one of them, the com-

bination with lowest value of χ^2 is taken. Using this selection, the Top Mass Fitter (TMF) gives the correct answer in $\approx 60\%$ of cases.

Table 4.3 shows the efficiency of lepton b-jet association for the double tagged events using the above mentioned TMF method. For comparison, Table 4.4 shows the same quantities for the events with sigle b-tag and Table 4.5 shows the efficiency for the events without considering the b-tagging information and with the righ combination selected as the one with the lowest χ^2 from all 24 possible combinations. The selection efficiency is retrieved by using the MC truth (the parton level information from MC).

The jet-to-parton association is done by looking for the MC generated parton inside the cone of $\Delta R < 0.4$ around the reconstructed jet axis. The correct association expects the jets originating from b-quark to be associated with positively charged lepton (l^+) and the jets originating from $\bar{b}$-quark to be associated with negatively charged lepton (l^-). In case of the b-jets from the hadronic branch of $t\bar{t}$ decay the association with lepton of opposite charge to the one from the leptonic branch is assumed.

decay branch	% of b-tagged jets with parton association	% of b-tagged jets with correct association	% of b-tagged jets with incorrect association
all	92.6	65.6	26.0
leptonic	95.2	65.1	25.6
hadronic	93.7	65.1	26.4

Table 4.3: Efficiency of lepton - b-jet association for the 2 b-tagged events

decay branch	% of b-tagged jets with parton association	% of b-tagged jets with correct association	% of b-tagged jets with incorrect association
all	94.4	61.1	33.3
leptonic	95.2	62.4	32.8
hadronic	93.7	60.0	33.7

Table 4.4: Efficiency of lepton - b-jet association for the 1 b-tagged events

To optimize the lepton b-jet association, additional studies were done examining following effects:

- using cut on value of χ^2 (accepting only events, where the χ^2 of selected combination is below a given value),

decay branch	% of b-fitter jets with parton association	% of fitter jets with correct association	% of fitter jets with incorrect association
all	53.9	37.2	16.7
leptonic	55.0	41.0	14.0
hadronic	52.9	33.5	19.4

Table 4.5: Efficiency of lepton - b-jet association for the untagged events

- running Top Mass Fitter in different modes (unconstrained fit with top mass as a free parameter, or constraining the top mass to $m_t = 175 \pm 3$ GeV/c^2),

- using different tagging algorithm (SecVtxLoose versus SecVtxTight tagging [31])

- using also single tagged events.

Figure 4.2 shows the performance of Top Mass Fitter (in sense of ϵD^2) as a function of χ^2 cut obtained on a MC generated signal sample. As a result of this study, the optimal value for χ^2 cut of 9 was selected.

Table 4.6 summarizes the effects of using Top Mass Fitter in the constrained or unconstrained mode and using tight or loose tagging algorithm. As can be seen, by using the constrained mode, the efficiency decreases about 10%, but at the same tame, there is the 10% increase of purity. That is why the overall performance is better when using Top Mass Fitter in the constrained mode.

In case of the tagging algorithm, while there is no significant change of purity, when using the loose tagger compared to the tight one, there is a gain of $\approx 40\%$ in the tagging efficiency when the loose tagger is used. The final decision based on this study was to use the loose tagging algorithm and Top Mass Fitter in the constrained mode.

When requiring at least one tagged jet instead of at least two, the event selection efficiency increases from 5% to 15%, however the purity of pairing decreases considerably from 82% to 62%. Therefor overall performance is lower than in case of using double tagged events.

4.3.3   Jet flavor tagging algorithm

The next important part of the analysis is the determination of the flavor of b-jet, i.e. to determine whether the jet is from a b-quark or from an $\bar{b}$-quark.
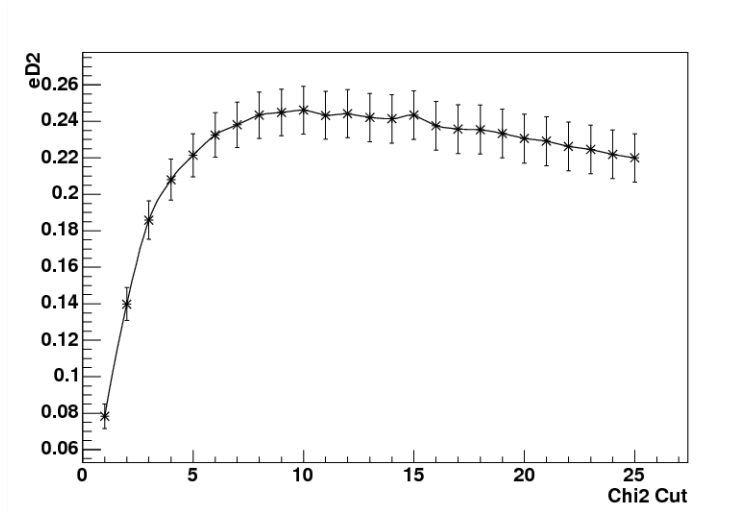
Figure 4.2: Performance of Top Mass Fitter as a function of χ^2 cut

TMF mode	tagger	ϵ	P	D	ϵD^2
unconstrained	tight	0.671 ± 0.004	0.752 ± 0.005	0.504	0.171
unconstrained	loose	0.660 ± 0.004	0.744 ± 0.004	0.487	0.157
constrained	tight	0.565 ± 0.005	0.843 ± 0.005	0.685	0.265
constrained	loose	0.553 ± 0.004	0.851 ± 0.004	0.702	0.272

Table 4.6: Effect of different Top Mass Fitter modes and different tagger versions

There are several methods of obtaining the information on the b-quark flavor.

One of the possibilities is to use the semileptonic decay of B-meson. The b-quark in meson decays into lepton, its neutrino and most often to c-quark. The charge of the lepton is in this case correlated (has the same sign) with the flavor of original B-meson. The advantage of this approach is a relatively high purity. The main sources affecting the purity are the $B^0_{d(s)}$ oscillations and cases where the B-meson decays into D-meson (contains c-quark) which decays leptonicaly. In this two cases, the charge of the final lepton will have the opposite sign as that of the original b-quark. The bigest disadvantage of this method is the low efficiency (B-meson decays semileptonicaly in $\approx 10\%$ of cases).

Another approach is to calculate the weighted charge of the tracks coming from b-quark hadronization, as their charges are sensitive to certain degree to the charge of the original b-quark. The purity of this method is lower than that of the first approach, but the efficiency is close to 100%.

Track selection for jet charge calculation

To use the weighting method for the jet charge calculation, the tracks belonging to a given jet should be identified. For the track-to-jet association, the relative angle of track and jet axis was used.

Originaly, the tracks used for the jet charge calculation were selected from all tracks in the event. For a track to be selected it was required to have the transverse momentum $p_T > 0.5$ GeV. If it passed this criteria, the track momentum was compared to the b-jet axis. The relative angle ΔR between the track and b-jet axis was calculated and if the $\Delta R < 0.4$ then the track was associated to this b-jet.

It was discovered that this simple selection introduces significant bias between positive and negative flavour of the b-jet. It was found to be caused by the difference of the interaction with the detector material for positive and negative charged particles. Figure 4.3 shows the p_T distribution for negative and positive tracks in the event. The excess of positive tracks at low values of p_T can be observed. Figure 4.4 shows the MC information about the x,y coordinates of the track creation vertex for the negative and positive tracks. Again here can be observed the higher number of the positive tracks created inside the detector.

To eliminate this effect, the track selection was modified, to cut off the low p_T tracks from material interaction. For the jet charge calculation only the tracks from the group of tracks used for the secondary vertex fit (SecVtx tracks) were selected. The p_T cut was raised to the value of 1.5 GeV and
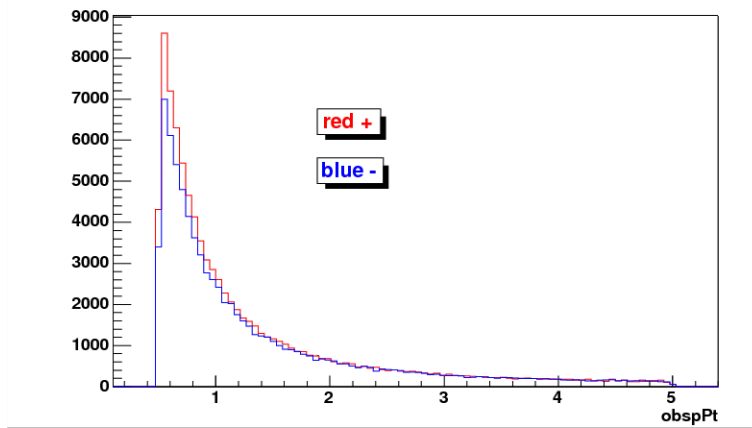
Figure 4.3: The p_T distribution of positive (red) and negative (blue) tracks in the MC sample
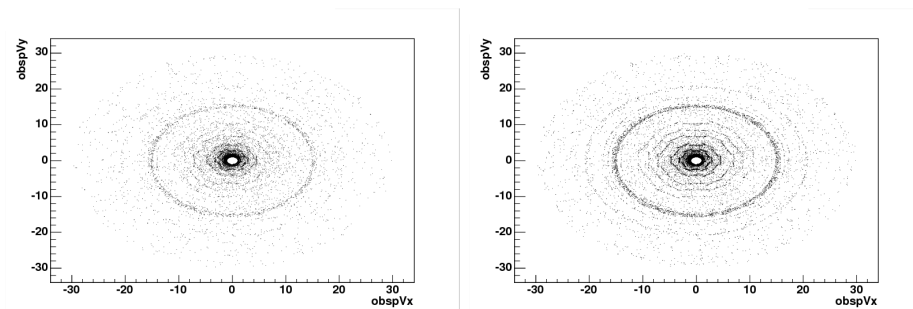
Figure 4.4: The distribution of the creation vertex coordinates from MC in x-y plane for negative (left) and positive (right) tracks

56

an additional cut on the impact parameter $|d_0| < 0.15$ cm with respect to secondary vertex was introduced. Figure 4.5 shows that the p_T distributions of the negative and positive tracks after the new selection are almost identical.
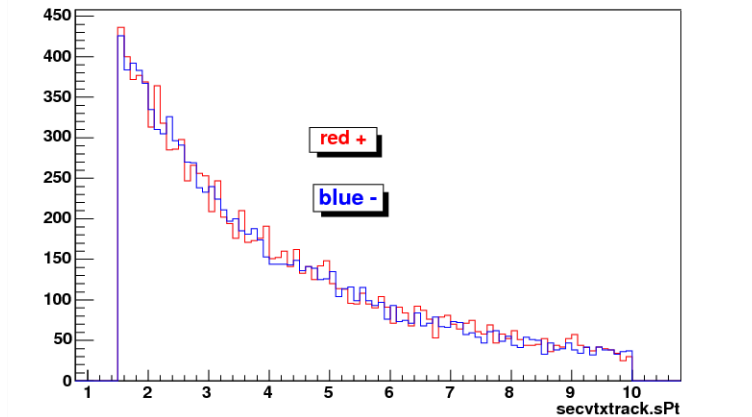
Figure 4.5: The p_T distributions for positive (red) and negative (blue) tracks after new selection

Optimization of track weighting

Due to the low event statistics expected to be available for the analysis, the algorithm using weighted charge of the tracks inside b-jet cone was chosen. Using this method, the charge of the jet is calculated as

$$q_{jet} = \frac{\sum_i w_i^\kappa \cdot q_i}{\sum_i w_i^\kappa} \tag{4.8}$$

where q_i is the charge of the i-th track, w_i is the appropriate weight assigned to this track and κ is an optimization coefficient. It is also required for each jet to have at least two tracks, satisfying the selection criteria mentioned above, associated with it. For the purpose of the analysis, the weight was chosen in the form of $w_i = \vec{j} \cdot \vec{p_i}$.

In the formula for the jet charge calculation 4.8, there is also an optimization coeficient κ. In the optimization study, the value of κ was varied from 0.1 to 1.0 observing the performance of the weighting metod. The dependence of the purity of the jet charge algorithm on this coeficient is shown in figure 4.6. Considering the errors on purity, the plot is nearly flat between the values $0.4 - 0.8$. As the result of this study, the value of $\kappa = 0.5$ was chosen for the final analysis.
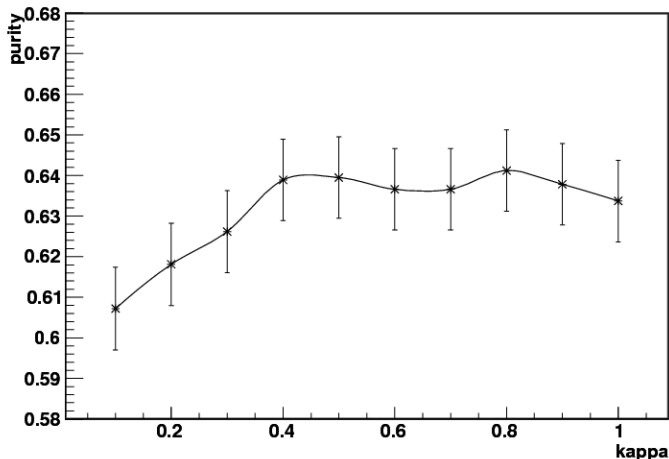
Figure 4.6: Dependence of the jet charge purity on optimization coeficient κ

Additional optimization

There are two different methods that can be used for charge analysis of $t\bar{t}$ events. One of the methods treats t and $\bar{t}$ decays sepparately while the other one not.

In dilepton channel, there are two leptons and two b-jets. After paring, two lepton - jet pairs are obtained. In lepton plus jets, there is only one lepton, with which the b-jet from leptonic branch of $t\bar{t}$ decay is associated. However, as the W from leptonic branch should have opposite charge as the W from hadronic branch, the hadronic b-jet can be assumed to be associated with lepton of opposite charge to the real lepton present in the event.

For both channels, there are two lepton – b-jet pairs and thus two independent measurements per event.

The alternative approach is based on the requirement that the two b-jets must have the opposite sign charges (OS) after applying the jet flavour tagging algorithm. Using this method, there is only one measurement per event, as the measurements from leptonic and hadronic brach of $t\bar{t}$ decay are correlated.

Using OS reduces the efficiency of the event selection significantly, but increases the purity of jet flavor tagging. However, since the number of events availible for the top quark charge analysis is low and the purity of pairing is high enough, it was decided to use the first method (lepton - b-jet pairs). For the lepton plus jets sample, a study was done to estimate the performances

of both approaches. Table 4.7 shows the performance for each of them.

method	ϵ % (events)	ϵ (pairs)	P %	ϵD^2
pairs	98.00 ± 0.07	98.00 ± 0.07	60.77 ± 0.27	0.056
OS	50.02 ± 0.36	25.02 ± 0.36	70.78 ± 0.44	0.045

Table 4.7: Performance of jet flavor tagging algorithm using opposite sign and lepton - b-jet pairs.

In the first aproach (using pairs of the b-charge measurements per event), the events are accepted when the jet tagging algorithm for both b-jets can be applied (meaning each of the two b-jets does have at least two tracks associated with it). Using this requirement the efficiency of event selection is 98%. When using the method with OS, the efficiency of event selection is 50%. This includes only events, where the jet flavor tagging algoritm can be applied to both b-jets and gives opposite signs of the calculated charge for them.

However, to be able to compare the performances of the two methods, the efficiency should be considered in terms of pair selection. In OS method only one pair per event can be used for charge measurement. The efficiency in respect to the pair selection is therefor reduced by factor of two. For the other method, both pairs are taken from each event, so the efficiency for pairs is not changed.

Comparing the two methods, the performance is slightly better when using pairs. Also, because there are two pairs for each event, the number of measurements is doubled. This should help increase the statistical significance of the final results.

4.4 Calibration of jet flavor tagging algorithm

4.4.1 Jet charge calibration method

One of the biggest challenges in the top quark charge analysis is to tag correctly the flavor of b-jet. The flavor tagging method was described in the previous section. However, the method is sensitive to details of the fragmentation process, where the MC cannot be fully trusted to reproduce the real situation. Therefore the purity of the algorithm should be obtained from a b-jet containing experimental data. It is needed especially because the value of purity and its uncertainty are the parameters for which the top quark charge measurement is the most sensitive. The value of purity serves as one of the input parameters in estimating the statistical significance of observed experimental results.

To be able to calibrate the jet flavor tagging algorithm, the muon enriched di-jet data sample was used. Using an appropriate event selection, a sub-sample with a high population of $b\bar{b}$ pairs can be obtained. In these events, one of the jets is identified by the heavy flavor hadron that decays semileptonicaly, producing the triggered muon. This jet is referred to as *muon* jet. The other jet, opposite to the muon one is referred to as the *away* jet.

If the muon and away jet are indeed from $b\bar{b}$, the charge of the muon should be correlated with the flavor of the away jet (if there is no mixing the μ charge has the opposite sign compared to the b-quark initiating the away jet), which is obtained by applying the flavor tagging algorithm. The purity (P_{obs}) can be determined as the number of muon - away jet pairs with opposite sign of charge (OS) over the total number of muon - away jet pairs.

$$P_{obs} = \frac{N_{OS}}{N_{OS} + N_{SS}} \qquad (4.9)$$

where N_{OS} and N_{SS} correspond to the number of events with the opposite and the same sign of charge respectively.

A main drawback of this method is in determination of the $b\bar{b}$ dijet fraction ($f_{b\bar{b}}$) from the pairs passing the selection cuts, since muon can be also produced by a charm meson decay or can be a jet misidentified as muon, paired with a light jet wrongly assigned as an away jet, for example in gluon splitting events.

The equation for the P_{obs} needs a few corrections. First, the muon in the muon jet can come not directly form a semileptonic decay of b, but from a secondary decay ($b \rightarrow c \rightarrow \mu$). This would change the charge of the muon and thus it will lead to an incorrect flavor assignment. Instead of an OS

pair we will have a muon – away jet pair with the same sign (SS) one and vice versa. The same effect occurs in case of B meson mixing – in this case produced b quark ($\bar{b}$ quark) converts due to the weak interactions to its anti-quark. Lastly, the correlation will be diluted by non-$b\bar{b}$, non $c\bar{c}$ events, since no charge correlation is expected for them. Taking all this into consideration the equation 4.9 is changed to:

$$P_{obs} = f_{b\bar{b}}(1 - f_{secmix})P_{JQ} + f_{b\bar{b}}f_{secmix}(1 - P_{JQ}) + 0.5 \cdot f_{bkgd} \qquad (4.10)$$

where P_{JQ} is the true purity corresponding to the case when the flavour of the away b-jet at production is known – it presents the maximal performance of the jet flavor tagging algorithm, $f_{b\bar{b}}$ is the fraction of $b\bar{b}$ events in the sample of the selected dijet events, f_{secmix} corresponds to the fraction of the events where either mixing occured in the away jet branch or the cascade decay ($b \rightarrow c \rightarrow \mu$) or mixing occured in the muon jet and $f_{bkgd} = 1 - f_{b\bar{b}}$ is the fraction of background. The relation 4.10 follows from the fact that the observed number of OS events (N_{OS}) can be expressed as follows:

$$N_{OS} = f_{b\bar{b}}(1 - f_{secmix})N_{OS}^t + f_{b\bar{b}}f_{secmix}N_{SS}^t + 0.5N_{bkgd} \qquad (4.11)$$

where N_{OS}^t (N_{SS}^t) is the number of the true OS (SS) events and N_{bckg} is the number of background events.

In order to determine the fraction of true $b\bar{b}$ events, the transverse momentum of the muon relative to the muon jet axis ($p_{T,rel}$) and the invariant mass of the secondary vertex on tagged away jet (M_{vtx}) [32] were used. Figure 4.7 show their distribution for b,c or light quark jets (tight tagged for the muon jet and loose tagged for the away jet) obtained from the di-jet MC.

The $b\bar{b}$ fraction is obtained by fitting the corresponding data spectra on the muon and away jets with templates from MC. The P_{obs} is corrected for non-$b\bar{b}$ contribution. Then the purity of jet flavour tagging algorithm (P_{JQ}) for signal can be calculated, which is later used in the statistical treatment.

4.4.2   Event selection

To select the sub-sample enriched in $b\bar{b}$ events, the following selection criteria were applied

- two high E_T jets, one of them containing a reconstructed CMUP muon within the cone $\Delta R < 0.4$ around the jet axis (muon jet)

- away jet is separated from muon jet, $\Delta\phi > 2$ rad
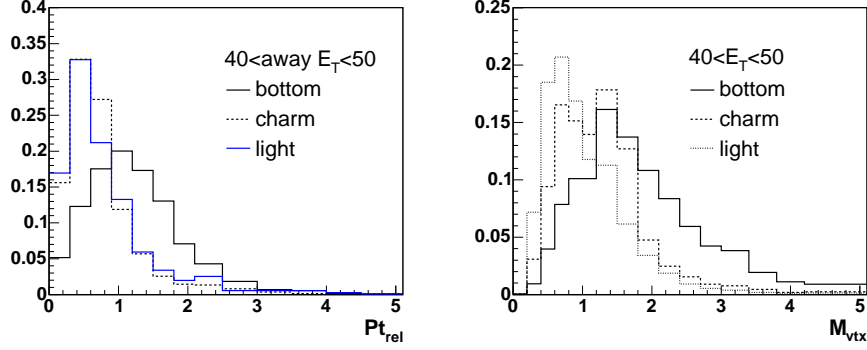
Figure 4.7: The distribution of muon transverse momentum relative to the muon jet axis and invariant mass of the secondary vertex on tagged away jet obtained from di-jet MC

- muon jet $E_T > 20$ GeV/c^2

- muon track $p_T > 9$ GeV/c

- muon track $|z_0| < 60$ cm

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

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

- distance of muon track to primary vertex $|z_0 - z_{vtx}| < 5$ cm

- muon track isolation > 0.1

- muon track have to pass through every layer of SVXII (no hit required)

- away jet $E_T > 20$ GeV/c^2

- away jet $|\eta| < 1.5$

- at least two good secondary vertex tracks used for loose tagger

Only one pair of muon jet - away jet per event is allowed, always looking for the highest p_T muon or the highest E_T jet. This selection is very similar to the one used to obtain the b-tag scale factor [32, 33]. The difference is that to keep the away jet as a generic one, the invariant secondary vertex mass cut was not applied.

The muon jet direction is corrected by the shift caused by the presence of muon. For this correction the following equation is used:

$$\vec{P}_{corr} = \vec{P}_{jet} + (1 - \frac{2GeV/c}{|\vec{p}_\mu|})\vec{p}_\mu \qquad (4.12)$$

where the value of 2 GeV was chosen from [34] as the most probable energy deposition of a muon in the p_T range under consideration. There is no correction applied to the muon jet E_T, since the study uses away jet E_T bins.

To enhance the heavy flavor content of the sub-sample, the away jet is required to have a loose secondary vertex tag. For the muon jet, tight tag is required. Table 4.8 shows the composition of the sub-sample after all the selection criteria have been applied (obtained from Monte-Carlo).

	Cases		pretagged (%)	tagged (%)
1	$\mu = b$	$AJ = b$	77	86.8
2	$\mu = b$	$AJ = c$	2.4	2.8
3	$\mu = b$	$AJ = light$	5.4	6.6
4	$\mu = c$	$AJ = b$	8.4	1.7
5	$\mu = c$	$AJ = c$	2.2	0.4
6	$\mu = c$	$AJ = light$	3.2	0.8
7	$\mu = fakes$	$AJ = b/c/light$	1.0	1.0

Table 4.8: Classification of events in the enriched MC sample. Muon (μ) and away jet (AJ) are matched to the partons within a cone of $\Delta R < 0.4$. *Fakes* include the events where the reconstructed muon did not match a generator level muon within a cone of $\Delta R < 0.05$, or those where the muon is matched but the jet is not coming from b or c quark. "Pretagged" column corresponds to the case, where only away jet is loose tagged, "tagged"' column to the case where also the muon jet is tight tagged

From this table, it can be seen that the first case, which is the case of interest, is enhanced by tight tagging of the muon jet, while the other contributions are reduced. Special case is the $c\bar{c}$, because its presence could bias the result, as these events also have a correlation between the muon charge and the away jet flavor. The "fakes" (hadrons misidentified as muons) in this table is underestimated, since the heavy flavor enriched MC that requires presence of a muon on generator level, is used. However, the $b\bar{b}$ fraction will be obtained by fitting the experimental data.

4.4.3 Measuring the $b\bar{b}$ fraction

The problem now is to actually measure the $b\bar{b}$ fraction ($f_{b\bar{b}}$) in the data sample. In order to extract the b content on the muon side, the $p_{T,rel}$ distribution of the muon transverse energy with respect to the jet axis, was

used. To obtain the shapes of muon $p_{T,rel}$ distributions used as templates, the heavy flavor enriched MC sample (subsample of di-jet MC, where the presence of muon on generator level was required) was used for the parton matched b and c jets, and the di-jet MC for the light/gluon jets. The $p_{T,rel}$ templates are shown in Figure 4.8 for three selected intervals of away jet E_T. Although some dependence on away jet E_T can be seen for the heavy flavor jets, there is no such a dependence for the light/gluon jets. This is good, since MC can be trusted to represent well the E_T spectrum for the heavy flavor jets in data, but not for the light/gluon jets.

For the away jet side, the invariant mass, M_{vtx}, of the secondary vertex tracks belonging to the away jet is reconstructed and used. The M_{vtx} templates were obtained using the di-jet MC sample, where the jets were required to pass the same selection cuts listed for the enriched sub-sample. In the case of light jets only events with no heavy flavor quarks were used (based on MC generator information). Figure 4.9 shows the M_{vtx} templates for three selected intervals of away jet E_T.

Figure 4.8: The $p_{T,rel}$ templates for the bottom, charm and light quark/gluon jets obtained from the MC for three selected intervals of away jet E_T
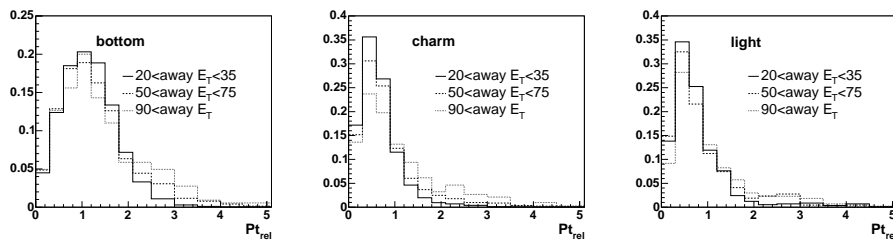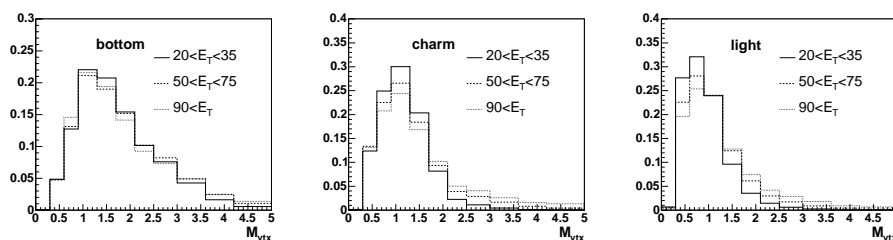

Figure 4.9: The M_{vtx} templates for the bottom, charm and light quark/gluon jets obtained from the MC for three selected intervals of away jet E_T

Fitting procedure

Using the $p_{T,rel}$ and M_{vtx} templates, the selected experimental di-jet data were fitted in nine E_T away jet bins. To illustrate the procedure, Figure 4.10 shows the fits in one selected away jet E_T bin ($40 < E_T < 50$ GeV) where the contributions from different sources are summed up. For the $p_{T,rel}$ case, the charm and light jet spectrum are so similar that the fitting procedure cannot distinguish between them. Therefore a two template fit was used. For the M_{vtx} spectra a three template fit was applied, however the fit was evaluated using only two of them (either superposition of b-c or b-light templates) and the difference of the two fits was treated as a systematic uncertainty.
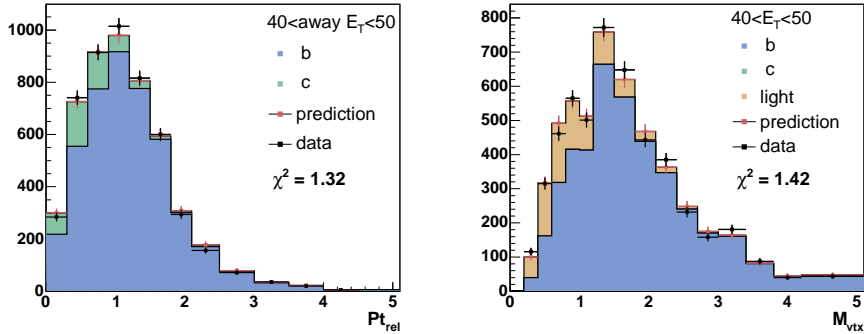
Figure 4.10: The $p_{T,rel}$ and M_{vtx} fit using two and three templates respectively for one bin of away jet E_T

In this particular E_T bin, the $p_{T,rel}$ fit result is $89,5\pm2.7\%$ for the presence of the muon b-jet in the sample. This implies that for $\sim 10\%$ of events, the muon jet is not from a b-jet. The M_{vtx} result is $83.0 \pm 2.6\%$. In the case that all of the 10% of non-b muon jets are paired with the away jets not containing b-quark, the upper limit for the $b\bar{b}$ fraction can be obtained and it would correspond to the M_{vtx} fit value. On the other hand, if all of these 10% non-b muon jets were paired with the b away jets, the lower limit for the $b\bar{b}$ fraction would then be found by substracting from the $b\bar{b}$ fraction determined by M_{vtx} approach the non-b "contamination" from the muon side. The $b\bar{b}$ fraction is calculated as the average of those two limits.

For this E_T bin ($40 < E_T < 50$ GeV) it corresponds to:

- upper limit = 83% (M_{vtx} fit result)

- lower limit = $83\% - (100 - 89.5\%) = 72.5\%$

- $b\bar{b}$ fraction $= 77.8 \pm 5.9\%$

This process is repeated for all E_T bins, obtaining the fraction of $b\bar{b}$ as a function of away jet E_T.

Mixing and secondaries corrections

As mentioned before, muons in the data sample can be produced not only directly by b decay ($b \rightarrow \mu$), but also by mechanism like $b \rightarrow c \rightarrow \mu$. In the later case the expected muon charge will be changed and therefore, the correlation with the away jet charge would be of the same sign, instead of opposite. The same effect is caused by the occurence of B meson mixing. To take this into account, the fraction of cascade and mixing events were calculated in the enriched MC used for this calibration.

To include all possible cases, the fraction of events with incorrect muon sign, f_{secmix}, was defined as

$$f_{secmix} = f_{secOS}(1 - f_{mix}) + (1 - f_{secOS})f_{mix} \qquad (4.13)$$

where f_{secOS} includes only the $b \rightarrow c \rightarrow \mu$ case, and is measured in sample without any mixing events in muon branch and f_{mix} is the fraction of events, in the heavy flavor sample, with the B-oscillation occurred.

The values of these fractions obtained from MC are

away jet $E_T < 50$ GeV	$f_{secOS} = 0.109 \pm 0.005$	$f_{mix} = 0.154 \pm 0.005$
away jet $E_T > 50$ GeV	$f_{secOS} = 0.165 \pm 0.017$	$f_{mix} = 0.147 \pm 0.015$

In order to take into account the differences in B production rates and semileptonic branching ratios between values used in the MC generator and cited by the Particle Data Group (PDG), the fractions were scaled by the ratios between PDG and unbiased MC [35]. The final result f_{secmix} is 0.229 ± 0.029 for low away jet E_T region and 0.263 ± 0.031 for high away jet E_T region.

4.4.4   Scale factor

Following the above described procedure the muon calibration data of integrated luminosity 1.5 fb^{-1} were fitted and the b fraction of muon jets was found. The fraction is close to 92% for low away jet transverse energies (E_T), but decreases to $\sim 85\%$ at high away jet E_T values, where the probability of finding fake muons is increased. The muon side b fraction dependence on away jet E_T is shown in Figure 4.11.

In Figure 4.12 is shown a dependence of the b-fraction on the away jet side as a function of away jet E_T. The result has been obtained by fitting the secondary vertex invariant mass on the away jet side. In this case, the b fraction decreases from 90% to 40%. This drop could be the result of the increasing b-mistag rate with increasing energy.

Combining the two results, the $b\bar{b}$ fraction in the experimental sample is obtained for each E_T bin. The distribution follows the trend of the b fraction on the muon and away jet side and has a steep drop, which is shown in Figure 4.13. In this figure, also the $b\bar{b}$ fraction from the comparative MC sample is shown. Here, the drop is not as obvious as in the data, but is still present.

The last ingredient for the calculation of the purity P_{JQ}, is the observed purity P_{obs}. The value of P_{obs} is found as a fraction of the OS events in each bin of the data sample. The dependence of the P_{obs} on the away jet E_T is shown in Figure 4.14. The E_T dependence is consistent with the $b\bar{b}$ fraction distribution, since the more background is present, the larger is the dilution and the closer is the purity to the value of 0.5.
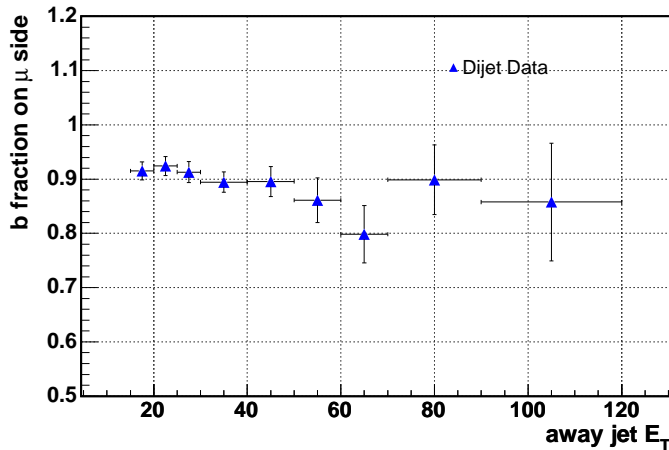
Figure 4.11: Dependence of the b fraction on muon side on the E_T of the away jet

To be able to get the jet flavor tagging purity (P_{JQ}) for any sample, in particular for the high E_T b-jets in top events, the scale factor (SF) was calculated as the ratio of the purity P_{obs} in muon calibration data and P_{obs} in a weighted average between a generic and an eriched MC samples (taking into accout the size of the samples). In top quark charge analysis, the dilepton channel requires one tight tagged b-jet, the lepton plus jets channel requires two loose tagged b-jets. To prove that the SF can be applied to any generic
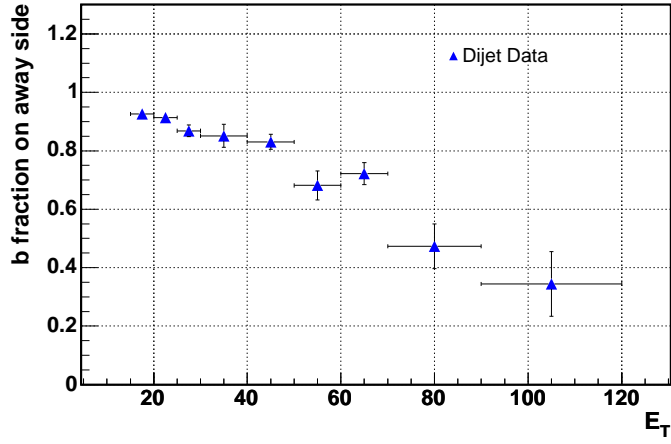
Figure 4.12: Dependence of the b fraction on away jet side on the E_T of the away jet
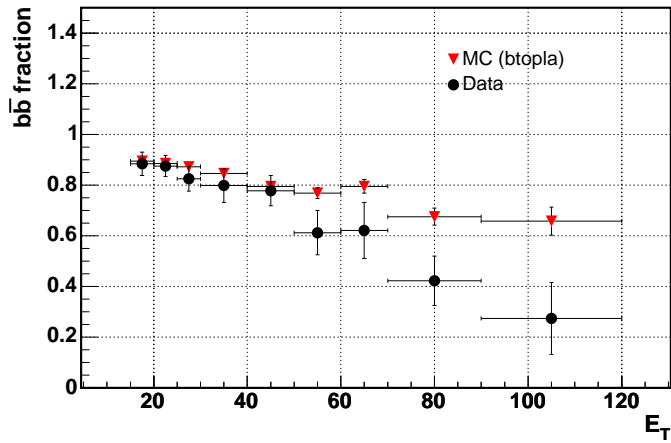
Figure 4.13: Dependence of the $b\bar{b}$ fraction on the E_T of the away jet

68
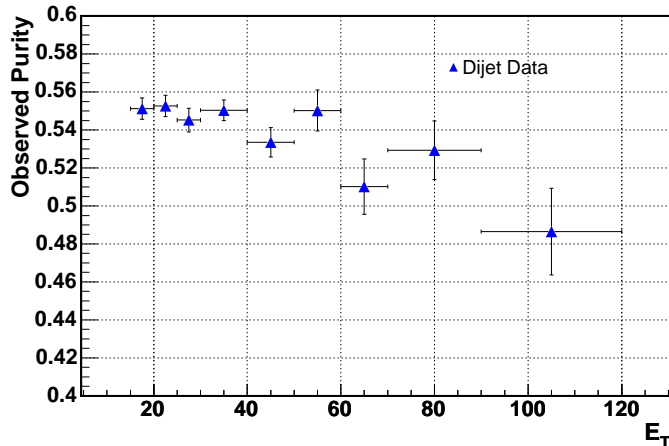
Figure 4.14: Dependence of the observed purity P_{obs} on the E_T of the away jet

jet, tagged or not, and without the concern about the choice of the tagging algorithm, the calibration study was done also on the tight tagged away jets.

As the observed purity (P_{obs}) in the MC samples, the weighted average between the jet charge purity observed in the heavy flavor enriched MC samples and in the generic (di-jet) MC one was used. Figure 4.15 shows the purity observed on loose tagged jets matched to b-quark using di-jet MC samples, heavy flavor enriched MC (in both cases parton matching to b-quark was used to define correct flavor assignment) and again using heavy flavor enriched MC, but where the same cuts as in data were applied to the sample and the flavor of away jet was done using correlation with charge of the muon in muon jet. The later was done to look for bias in the method, which was not observed.

The purity of the jet flavor tagging algorithm in the data is obtained by inserting the P_{obs} values, shown in the previous section, and the $b\bar{b}$ fractions into equation 4.10. The dependence of the purity (P_{JQ}) on the away jet E_T is shown in the Figure 4.16, together with the corresponding values obtained from the weighted average of MC purities. It will be shown later that the final result does not depend on the selection of MC, however the uncertainty does.

Figure 4.17 presents the result for the SF, for the loose tagger, with a fit of the ratio between data and MC with a constant function. A linear fit, used to obtaind an uncertainty due to the E_T dependence is also shown.

The study was repeated using the tight tagger and the resulting SF is

69

shown in Figure 4.18. There is no significant effect observed in respect to using different tagger.
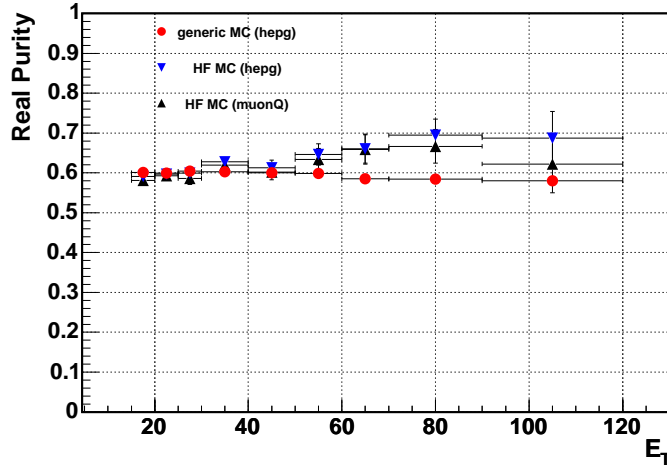
Figure 4.15: Jet flavor tagging purity P_{JQ} from jets matched to b in generic and heavy flavor enriched MC. The black triangles correspond to the purity from the enriched MC when using muon charge correlation to decide on correct jet flavor assignment.

4.4.5 Scale factor dependence on E_T, $|\eta|$ and number of vertices

E_T exptrapolation

The study of the flavor tagging purity (P_{JQ}) was performed on the di-jet $b\bar{b}$ enriched data sample, however the result is to be used in $t\bar{t}$ events, where the b-jets do have a different kinematics, mainly higher E_T.

Although the SF is assumed to be a constant, an error was assigned to account for a possible dependence on E_T. To estimate this error, the ratio between data and MC purities (SF) was fitted as a function of jet E_T (blue line in figure 4.17, with a line of non-zero slope and the result was weighted by the b-jet E_T distribution in $t\bar{t}$ events (Figure 4.19). The difference between this SF and the nominal one (the one obtained by fitting with a constant function) was added as a systematic error. When this uncertainty is estimated using the heavy flavor enriched MC instead of the di-jet one, the error

70
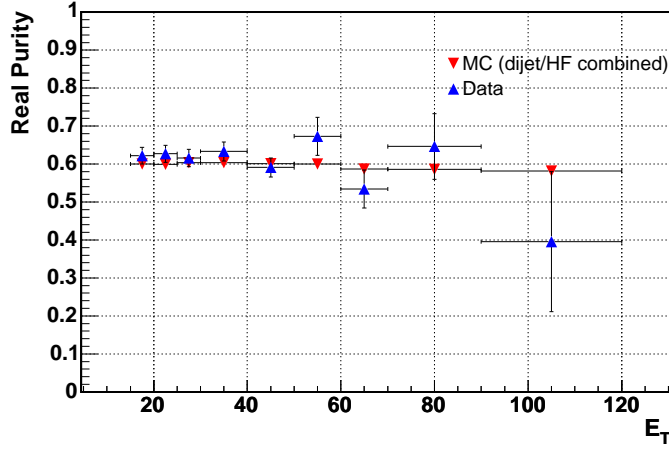
Figure 4.16: Jet flavor tagging purity P_{JQ} as a function of away jet E_T. Red triangles corespond to the purity for jets matched to b on MC sample (weighted average between purity in dijet MC and in enriched sample).
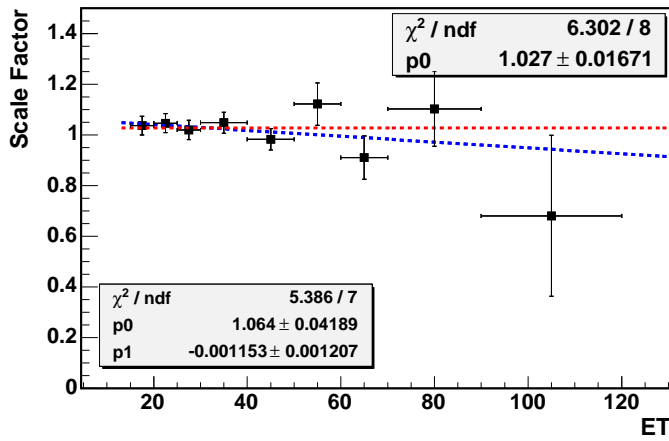
Figure 4.17: Scale factor as a function of E_T, for the loose tagged jets, calculated from the ratio between the JQ purity in muon calibration data and in a weighted average between a generic and an eriched MC. The red line corresponds to a fit with a constant function, while the blue is the fit with linear function with non-zero slope.
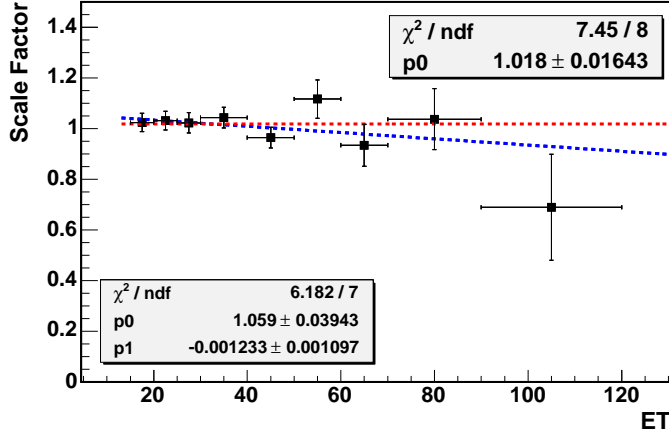
Figure 4.18: Scale factor as a function of E_T, for the tight tagged jets, calculated from the ratio between the JQ purity in muon calibration data and in a weighted average between a generic and an eriched MC. The red line corresponds to a fit with a constant function, while the blue is the fit with linear function with non-zero slope.

increases from 2% to 8% because of the low statistics of the enriched MC sample.

Dependence on η and number of vertices.

Except of E_T dependence, also dependence on the η of the away jet and the number of vertices was studied. Both were done in two different E_T bins to ensure that any dependence seen is not from E_T.

The $p_{T,rel}$ and M_{vtx} templates for b and light jets are shown in the Figure 4.20 for three selected intervals of η. No dependence is seen. Figures 4.21 and 4.22 show the SF for two E_T bins for selected intervals of η and number of vertices, respectively. Again, no significant dependence is observed.

4.4.6 Systematic uncertainties

There are several sources of systematic uncertainties related to the procedure used to find the b fraction on the muon jet side and also related to the b fraction on the away jet side. Also an uncertainty due to the E_T dependence is present.
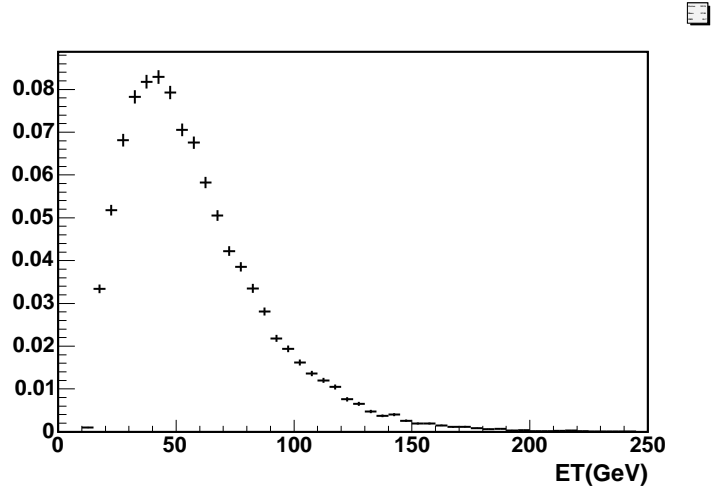
Figure 4.19: E_T distribution of b-jets in $t\bar{t}$ events.
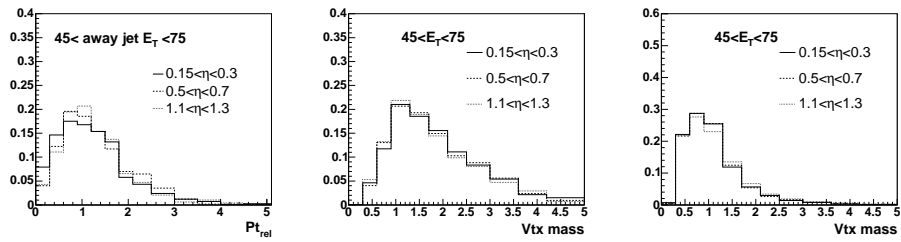
Figure 4.20: $p_{T,rel}$ templates for bottom (left) and M_{vtx} templates for bottom (middle) and light (right) jets as a function of η for one E_T bin.

73
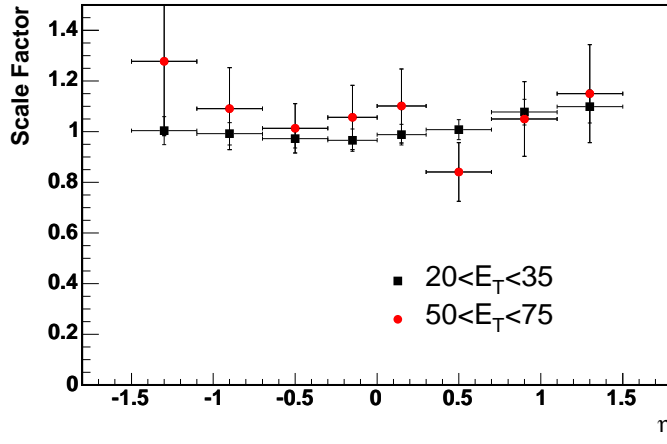
Figure 4.21: SF as a function of number of η, for two different E_T bins.
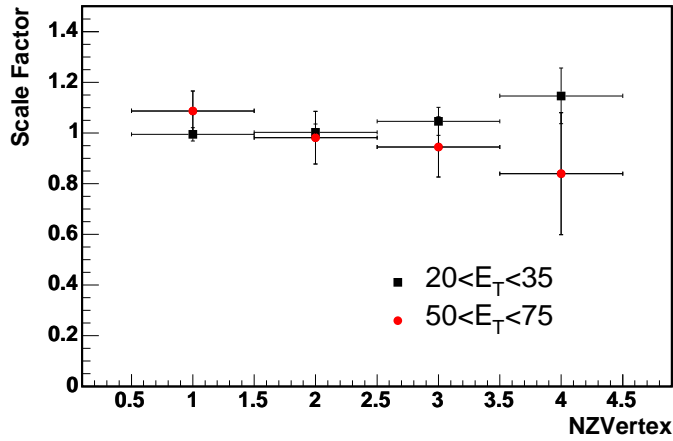
Figure 4.22: SF as a function of number of No. of vertices, for two different E_T bins.

$p_{T,rel}$ template tag bias

For the nominal values of SF, the $p_{T,rel}$ templates obtained on the tagged b-jets have been used. An uncertainty was introduced, to take into account a possible bias due to the tagging. The SF was recalculated with the $p_{T,rel}$ templates extracted from the non-tagged b-jets and the difference with respect to the nominal value was assigned as a systematic uncertainty.

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

The $p_{T,rel}$ fits were done using only two templates, obtained from b and c jets. The SF was also calculated using the light quark templates instead of the charm ones. The full shift with respect to the nominal SF value was added as a systematic uncertainty.

M_{vtx} template bias

Due to the incorrect tracking efficiency in MC (tracking efficiency and therefor the invariant mass is higher in the MC than in the data), the templates were shifted by 5% and the SF was recalculated. The difference with respect to the nominal value was assighned as a systematic uncertainty.

M_{vtx} fits

Although the SF was obtained using a M_{vtx} three template fit, the study was repeated with only two templates, combining either b-c or b-light. The shifts with respect to the nominal value of SF were used as a systematic uncertainty

Combined systematics

Table 4.9 shows the systematic uncertainties assigned to SF, from the above sources. The uncertainty due to E_T dependence (described before) is also included. The total systematic uncertainty is estimated to be 2.3%.

Systematic source	Relative Systematic Uncertainty (%)
tag bias	1.0
non-b	0.1
M_{vtx} fit (bc)	1.
M_{vtx} fit (bl)	0.1
track rec. ineff.	1.5
E_T dependence	1.
total	2.3

Table 4.9: Relative systematic uncertainties on the scale factor

4.5 Background

For the top charge analysis, the expected background composition after standard selection criteria is the same as in other top quark analyses for both dilepton and lepton + jets channel. Therefor this analysis directly relies on the top cross-section background estimation published in [29, 30]. On top of this, however, there is a need to examine each of the backgrounds for top charge analysis specific characteristics. It means that the efficiencies of the top charge analysis specific cuts for each type of background are needed. In addition, there is also a need to investigate whether the individual background types are more likely to mimic the SM-like $t\bar{t}$ events (the charge of b-jet is anticorellated with the charge of the lepton) or the exotic-model-like $t\bar{t}$ events (the charge of b-jet is corellated with the charge of the lepton).

To be able to do this investigation, the purity of background is defined as

$$p_b = \frac{N_+}{N_+ + N_-} \tag{4.14}$$

where N_+ is the number of SM like events and N_- the number of exotic-model-like events. The value of $p_b = 0.5$ therefore means that the treated background proces will not mimic neither the SM nor the exotic model. This type of background is also called the symmetric background (leads to a symmetric b-jet charge spectrum with zero mean value). A value of $p_b > 0.5$ means that the background resembles the SM and a value of $p_b < 0.5$ means the background resembles the exotic model events. For those two cases, the background is called the asymmetric background.

For each type of background, the value of p_b is obtained from the MC generated samples.

4.5.1 Dilepton background

In the dilepton channel there are three types of background that should to be taken into account: Drell-Yan processes, diboson production and fakes (jet is misidentified as a lepton). To obtain the efficiencies for these types of background, the MC generated samples were used, applying the full set of the selection criteria described in the section 4.3. To obtain purities, however, the b-tagging criteria were omited to increase the number of events available.

Drell-Yan

One of the dominant background processes in dilepton channel are Drell-Yan processes: (Z/γ^*) decaying to e^+e^- or $\mu^+\mu^-$ final states with associated jets. Since there are no neutrinos in these events, in order to pass the $\vec{E}_T$ cut, they must have substantial $\vec{E}_T$ resulting from mismeasurements of jet energies.

Fakes

This type of background consists of events where one of more jets are misidentified as leptons in the detector. It is the second largest source of dilepton background. These events are mostly the result of $W \to l\nu$ decay with additional jets where one of the jets is reconstructed as a lepton.

Diboson

The remaining source of background taken into consideration is the diboson (WW and WZ) production. Both of the bosons have to decay leptonicaly in order to get two high p_t leptons and there is two or more additional jets from gluon radiation must be present.

In all of these tree types of background, jets are created from gluon radiation. There is no correlation between the charges of the leptons and the flavors of the associated jets. All three backgrounds are expected to be symmetric in terms of charge analysis, i.e the charge spectra of the associated jets are symmetric with the zero mean value.

4.5.2 Lepton + jets background

In the lepton plus jets channel, the W production accompanied by heavy flavor quarks (b,c) has the greatest contribution to the background. The next

is QCD background, W production accompanied by light quarks (u,d,s) also called mistag, followed by singletop production and the smallest contribution is from diboson production.

W + heavy flavor

The W production accompanied by heavy flavor quarks is the largest of the backgrounds in lepton plus jets channel. However because the Top Mass Fitter is applied, which one of the $b\bar{b}$, $c\bar{c}$ jets will get paired with lepton is still a random occurrence. There is therefore no correlation between charge of the lepton and flavor of the associated jet. The background is expected to be symmetric with regard to charge analysis.

Mistag

W production accompanied by light flavor quarks is essentially the same type of background as the W + heavy flavor. The reason why it is treated separately lies in the b-tagging procedure. In case of the W + heavy flavor the tagged jets are regular b-jets or c-jets which are supposed to pass the tagging criteria. In case of the W + light flavor, the tagged jets come from light quarks, and are incorrectly identified as b-jets. therefore the estimation of uncertainty for this background is done in a little bit different way than in the case of the W + heavy flavor. For this reason the two backgrounds are separated. For the top quark charge analysis, however, the behaviour of both of them is the same. The W + light flavor is also expected to be symmetric.

QCD

The QCD background is the most dificult type of the backgrounds as it can not be really well simulated by MC. The events in this background can be divided into two cases. In the first case one of the jets is misidentified as a lepton, in the second case, the lepton is a real soft lepton from semileptonic b or c decay. In the former case, there is no correlation between the charge of the jet misidentified as lepton and the flavor of the jet that Top Mass Fitter will associate to this "lepton". In the latter case, the b,c quarks come from dijet production of $b\bar{b}$, $c\bar{c}$ pairs. The flavor of the second jet can be correlated with the charge of the soft lepton.

Using a traditional approach based on employing the correlation between $\vec{\not{E}}_T$ and lepton isolation, the so-called Met versus Iso method, for estimation of charge asymmetry in QCD events, the events are divided into regions

according to the $\vec{\not{E}}_T$ and isolation of lepton (see section 2.3). The cuts for each region are:

region A $iso > 0.2$ $\vec{\not{E}}_T < 15$ GeV/c^2
region B $iso < 0.1$ $\vec{\not{E}}_T < 15$ GeV/c^2
region C $iso > 0.2$ $\vec{\not{E}}_T > 15$ GeV/c^2
region D $iso < 0.1$ $\vec{\not{E}}_T > 15$ GeV/c^2

But after applying all the selection cuts, there was not enough statistics in any region to estimate the purity for this background, only the efficiency was obtained. Therefore, there is no direct result for purity for this type of background.   The asymmetry in QCD background is expected to arise from radiative $b\bar{b}$ or $c\bar{c}$ events with additional jets obtained from gluon radiation. To get the upper limit for the value of purity, the di-jet MC was used enhanced on generator level for soft lepton decay in the $b\bar{b}$ production. The soft lepton and the second b-jet were required to be back to back in azimuth. For this case, the purity 0.545 ± 0.005 was obtained. This method was similar to the one used to obtain the purity from calibration data, which was described earlier.

The purity obtained from di-jet MC events is diluted by all other types of QCD background, which are symmetric. For actual purity of the whole QCD background, the fraction of $b\bar{b}$ and $c\bar{c}$ events contributing to it needs to be known. In an attempt to find this fraction, the fake electron QCD data sample was used (specialized QCD jet data sample, used in CDF analyses for fake electron studies), where the criteria for jet to be identified as fake electron are

- $E_T > 20$ GeV/c^2

- $|\eta| < 1.1$

- $0.8 < emFraction < 0.95$

- number of tracks in the jet ≥ 4

This study was not conclusive due to low statistics. However, running on di-jet MC using this fake electron definition, the purity was found to be 0.497 ± 0.008, which veryfied the assumption that the purity coming from $b\bar{b}$, $c\bar{c}$ gets diluted by other types of QCD backgrounds. Finally, it was decided to use conservative value for purity $0.545^{+0.005}_{-0.045}$, using the upper limit mean value, but assigning an asymmetric error to account for purity dilution down to value of 0.5.

Diboson

The diboson production accounts for a small fraction of background events in lepton plus jets channel. As in the case of dilepton channel, there is no mechanism that could cause a bias toward +2/3 or -4/3 events. This type of background is expected to be symmetric.

Singletop

For most of the analysis with top quark, the singletop production is considered as a background proces. For the top quark charge analysis it will contribute to the signal, but still it is treated as a background. As there is the top quark in this background, it is expected to show the charge asymmetry similar to the signal. In this case, the top quark is always produced with a b-quark in both s-channel and t-channel. However, the Top Mass Fitter treats the event as standard $t\bar{t}$ decay. It applies the constrains for W mass in hadronic branch of the decay by χ^2 calculation. As the light jets come from gluon radiation in this case, the pairing is less accurate than in case of true signal events, and the asymmetry is expected to be lower than in case of $t\bar{t}$ signal.

4.6 Systematics

The uncertainties on number of signal and background events due to systematics are already included on number obtained from cross-section predictions. To get the full systematic uncertainty of the analysis, the systematic uncertainties on the efficiencies and purities of pairing and jet flavor tagging need to be evaluated. Those uncertainties include:

- Jet energy scale (JES)

- Initial/Final state radiation (ISR/FSR)

- Top mass uncertainty

- b-tagging procedure

- parton distribution functions (PDF)

- MC modeling

- W polarization

How each of the contributions is dealt with is described in this section.

4.6.1 Jet energy scale

A detailed study of the jet energy scale (JES) was carried out in [36]. It was shown that the JES uncertainties depend on p_T of the jet and are in range of 1.8% (low p_T) to 2.8% (high p_T).

To determine the systematic uncertainties due to the JES, jet energies are shifted by the value of $\pm\sigma$ (JES error) in the signal MC samples. The efficiencies and purities of pairing and jet flavour tagging are recalculated with these shifted samples and the average of the percentage difference between the shifted ($\pm\sigma$) samples and the nominal one is taken as the systematics uncertainty.

4.6.2 Initial/final state radiation

In $t\bar{t}$ events, the jets can be initiated not only by partons comming from top decay (b-quark or light quarks from W-boson), but also by gluons radiated from initial (before hard collision occured) or final (after hard collision occured) state of interaction. There is a possibility to identify such a jet coming from initial or final state radiation (ISR/FSR) as the one comming directly from top decay.

The amount of ISR/FSR in the $t\bar{t}$ events is not well understood. In CDF analyses for estimation of the systematics uncertainties the signal MC samples are used, where the amount of ISR/FSR is varied (there are more or less ISR/FSR than in the nominal MC). The percentage difference between each of the shifted samples and the nominal one is taken. The ISR/FSR systematics uncertainty is then calculated as the quadrature sum of the largest shifts for ISR and FSR.

4.6.3   Top mass uncertainty

For the lepton + jets channel, the top mass is used as a constrain in the calculation of χ^2 by the Top Mass Fitter. The nominal value corresponds to a constrain of $m_t = 175$ GeV/c^2 on a MC input sample generated with the same mass of top quark.

The behaviour of efficiencies and purities was investigated, when the generated m_t in the MC input samples changed. The samples with generated $m_t = 170$ GeV/c^2 and $m_t = 180$ GeV/c^2 were used as input, with constrain on top mass in the fitter set to $m_t = 175$ GeV/c^2. The average of the percentage difference with respect to the nominal values was taken as the systematic uncertainty due to top mass.

4.6.4   *b*-tagging procedure

The uncertainties due to b-tagging efficiencies are already included in the signal and background composition from top quark cross-section analysis, which are used in the presented analysis. However the fact that the mistag rate in MC is lower than that in the data, will affect the number of events where the assigned b-jets are actually not true b-jets.

To take this fact into consideration, the mistag matrix on the non heavy flavor jets was applied. It means that for the light jets the b-tag is set using the experimental mistag probability (contents of the mistag matrix) and not by the reconstruction procedure itself. The same was done for the heavy flavor jets, using the b-tag scale factor as a sampling probability. The scale factor between the new double tagged sample (tagged by explicit mistag probability) and the nominal MC sample (tagged by the reconstruction procedure) is calculated. This scale factor coresponds to the fraction of double tagged events where one or more tagged jets do not come from b-quark. The measured value of the scale factor is $SF = 1.05 \pm 0.05$. The uncertainty on it is determined by varying the b-tag scale factor within its uncertainties and the mistag probability by 20% simultaneously.

4.6.5 PDF

For estimation of PDF systematics the standard procedure used in CDF analysis is applied [37]. The sets of 40 CTEQ [38] eigenvectors with "90% CL" variations are used. The difference between 20 eigenvectors pairs is added up in quadrature. In next step, the difference between MRST [39] PDF set generated with the same α_s value (MRST72) and default CTEQ PDF set is looked at. If the difference is greater than the 20 eigenvectors pairs uncertainty, it's added up in quadrature. At last, the difference between MRST PDF set generated with the same α_s value (MRST72) and MRST PDF set generated with different α_s value (MRST75) is added up in quadrature.

4.6.6 Monte Carlo modeling

To account for different $t\bar{t}$ modeling in different MC generators, the efficiencies and pairing purity obtained from the default sample (generated by Pythia MC) were compared with the efficiencies and pairing purity obtained from Herwig MC. The percentage difference is taken as the systematic uncertainty due to MC generator. Because the jet flavor tagging algorithm was calibrated on data, no systematics due to MC generator is assigned to the jet charge purity.

4.6.7 *W* polarization

The polarization of W boson affects the kinematic distribution of the lepton coming from W-decay. This could have an impact on variables, which contain the kinematic parameters of events. In dilepton channel where the M_{lb-max} is used the effect of polarization needs to be investigated. In lepton + jets channel, the polarization doesn't affect the event selection and doesn't need to be taken into account.

For dilepton channel, to estimate the systematics uncertainties due to W polarization, the method of reweighting events of nominal MC sample was used. In the nominal MC, in agreement with SM, the fraction of longitudinaly polarized W is $f_0 = 0.7$ and fraction of left polarized W is $f_- = 0.3$. The f_0 (and f_- accordingly) was changed first to 0.65 then to 0.75. For each of these values, the distribution of $\cos\theta$ was made for the lepton angle in the W rest frame with respect to reversed top quark direction. The difference in $\cos\theta$ distribution between the shifted and nominal MC samples was used as a weight in event-by-event reweighting. The plots on Figure 4.23 show the distribution of $\cos\theta$ for each of the two leptons. Figure 4.24 shows the distribution of M_{lb-max} for nominal and each of reweighted samples. Finally,

the difference in efficiencies and purities between nominal and each of the shifted samples were summed up in quadrature and taken as a systematic uncertainty.

However the effect of W-boson polarization proved to be small and was not included in final combined systematic uncertainty.

4.6.8 Combined systematics

The combined systematics uncertainty on the efficiencies and purities are calculated by summing each individual contribution in quadrature. Table 4.10 summarizes the sources and amount of systematics for the lepton + jets and dilepton channels respectively. The systematics uncertainties are shown for the total event selection efficiency, purity of pairing and jet flavor tagging purity.

all numbers are in %	pairing eff	pairing pur	jetQ eff	jetQ purity
L+J				
ISR/FSR	2.8	0	0	0.4
MC generator	0.6	0	0.1	(1.64)
JES	0.3	0	0	0
PDF	1	0.3	0	0
top mass	1.3	3.3	0.1	0.54
total	3.3	3.4	0.1	0.7
DIL				
ISR/FSR	3.1	0.5	0.3	0.7
MC generator	0	0	1.0	(2.0)
JES	4.4	1.1	0.4	0
PDF	4.0	0.4	0	0
top mass	3	1	0	0
total	7.3	1.6	1.1	1.7

Table 4.10: Systematics uncertainties for the lepton + jets and dilepton channels.

cos1	
Entries	37342
Mean	-0.1144
RMS	0.4663

cos1 w65	
Entries	33962
Mean	-0.1344
RMS	0.4685

cos1 w70	
Entries	33962
Mean	-0.114
RMS	0.4666

cos1 w75	
Entries	33962
Mean	-0.0938
RMS	0.4638

cos2	
Entries	37324
Mean	-0.1083
RMS	0.4655

cos2 w65	
Entries	33962
Mean	-0.1275
RMS	0.468

cos2 w70	
Entries	33962
Mean	-0.1072
RMS	0.466

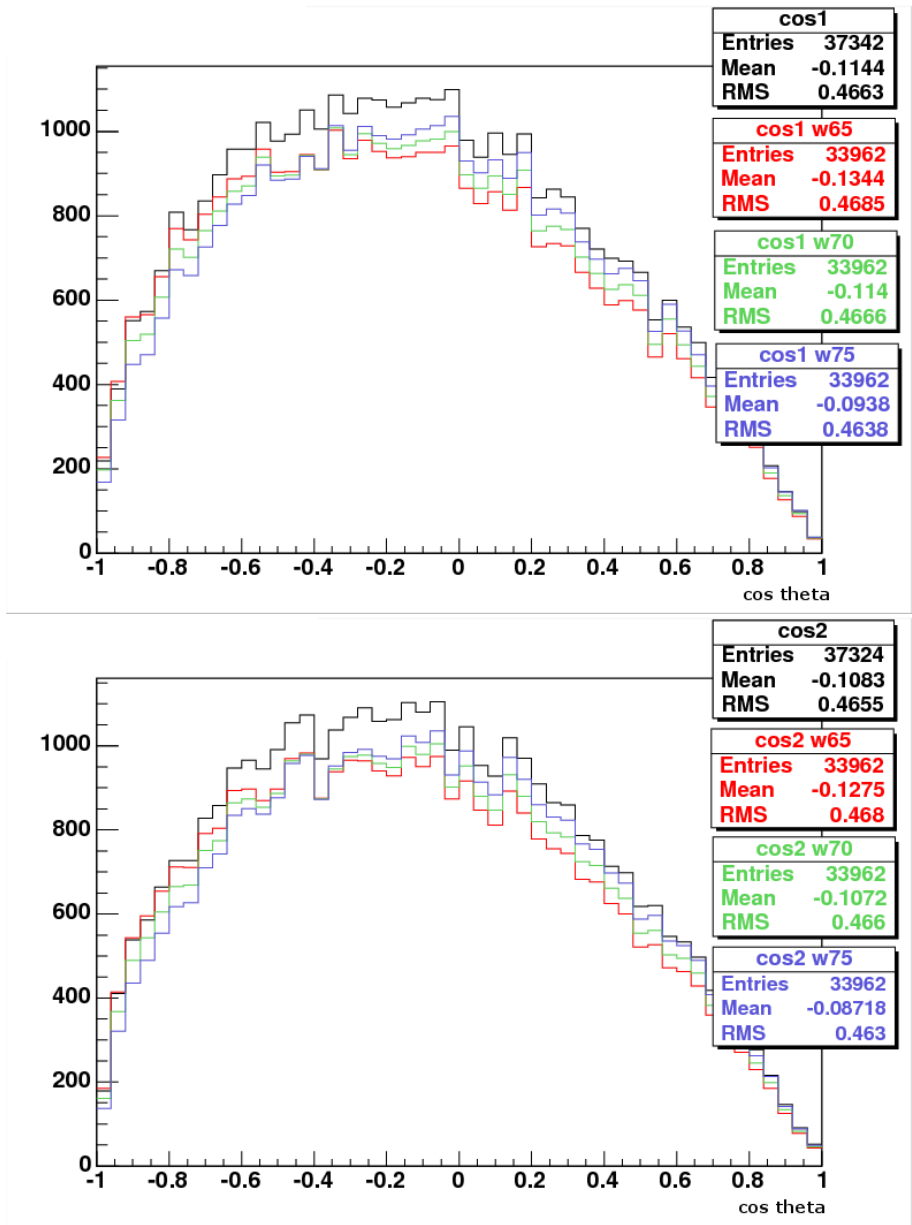| cos2 w75	
Entries	33962
Mean	-0.08718
RMS	0.463

Figure 4.23: Difference between the nominal sample and samples with shifted polarization in distribution of $\cos\theta$ for both leptons in DIL events.

85

MlbsMax

Entries	30970
Mean	2.051e+04
RMS	1.169e+04

MlbsMaxW65

Entries	26639
Mean	2.057e+04
RMS	1.166e+04

MlbsMaxW70

Entries	26639
Mean	2.075e+04
RMS	1.167e+04

MlbsMaxW75

Entries	26639
Mean	2.094e+04
RMS	1.168e+04
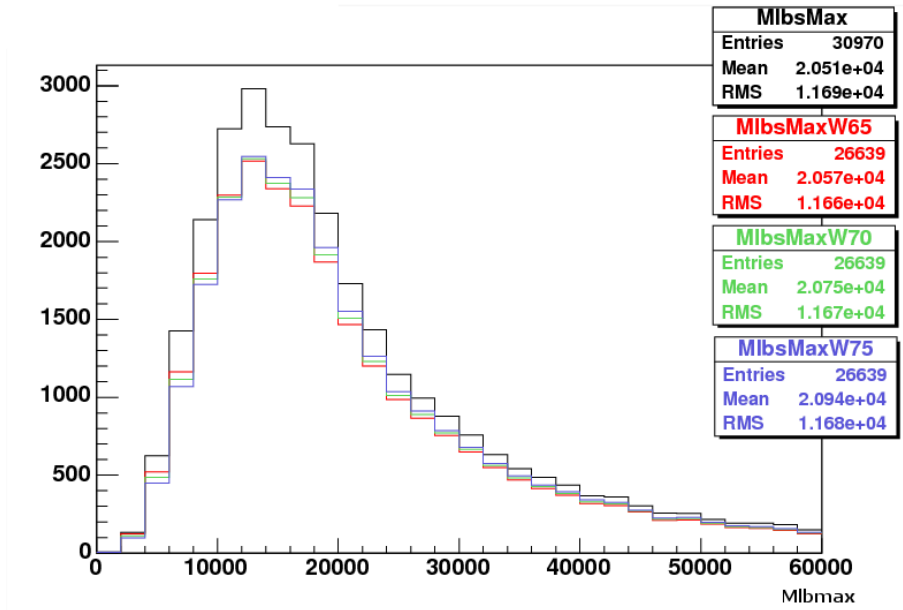
Figure 4.24: Difference between the nominal sample and samples with shifted polarization in distribution of M^2_{lb-max}

4.7 Expectations

4.7.1 Signal and background expectations

After all the optimizations and background and systematic studies, the final expectation numbers can be obtained. These numbers include the total efficiencies and purities for the signal and background and the final number of the signal and background events. These numbers will be used to determine the statistical sensitivity of the analysis.

To get this numbers, the background and signal predictions are taken from the cross-section studies for both the dilepton [29] and lepton + jets channel [30]. The number of events for the signal and each background is multiplied by the efficiency obtained after applying the all top quark charge specific analysis cuts. These numbers with the corresponding efficiencies are shown in Table 4.11. The uncertainties on N_b and N_s are propagated from the prediction numbers and efficiencies.

After getting the expected number of events for the signal and each backgroud, the expected number of events with the SM signature and those with the XM signature are estimated using the charge asymmetries obtained from the signal and background studies. For the background where no charge asymmetry is expected, the value 0.5 ± 0.0 is used. For the backgrounds

where the charge asymmetry is expected, the actual value from the MC studies is used. The final number of expected SM like and XM like events for each background are shown in Table 4.12.

background	prediction	efficiency	N_b or N_s
L + J			
W+HF	10.23 ± 4.31	0.14 ± 0.01	1.47 ± 0.62
QCD fakes	4.06 ± 4.94	0.15 ± 0.03	0.61 ± 0.75
Diboson	0.95 ± 0.15	0.20 ± 0.03	0.19 ± 0.04
Mistag	2.29 ± 0.68	0.15 ± 0.01	0.33 ± 0.10
Singletop	2.64 ± 0.38	0.21 ± 0.01	0.55 ± 0.08
Total	20.17 ± 6.61	-	3.15 ± 0.99
Signal	138.56 ± 24.02	$0.52^{\pm 0.002(stat)}_{\pm 0.02(sys)}$	72.09 ± 12.73

DIL			
Drell-Yan	$0.51^{+1.02}_{-0.51}$	0.30 ± 0.05	$0.15^{+0.31}_{-0.15}$
Fakes	$2.82^{+5.24}_{-2.82}$	0.25 ± 0.04	$0.71^{+1.39}_{-0.69}$
Diboson	$0.19^{+0.38}_{-0.19}$	0.5 ± 0.08	$0.9^{+0.19}_{-0.09}$
Total	$3.52^{+5.75}_{-2.87}$	-	$0.96^{+1.47}_{-0.73}$
Signal	41.09 ± 3.8	$0.33^{\pm 0.003(stat)}_{\pm 0.02(sys)}$	13.44 ± 1.60

Total Background	$1.404^{+0.961}_{-0.336}$
Total Signal	$45.652^{\pm 0.571(stat)}_{\pm 8.062(sys)}$

Table 4.11: The background and signal predictions with the top quark charge specific efficiencies for the sample 1.5 fb^{-1}

When all the purities for each background are collected, they need to be combined into total background purity. In dilepton channel, all the backgrounds are symmetric, so the value 0.5 ± 0.0 is used. However, in lepton plus jets channel, there are two backgrounds that show an asymmetry - QCD and singletop. The total purity is calculated considering the fraction of both backgrounds in respect to the symetric rest of the background and using their measured purities.

For the signal, the efficiency shown in Table 4.11 is the combined efficiency of pairing and flavor tagging. The estimation of purity is not so straightforward, since the purity is different for cases when b-tagged jets really correspond to b-quarks, and when they do not. The purities for both of the cases can be found by examining b-jet charges event by event in MC, and after that the difference of flavor tagging algorithm performance between

background	N_b or N_s	purity	N^+	N^-
L + J (1.5 fb^{-1})				
W+HF	1.47 ± 0.62	0.5 ± 0.0	0.74 ± 0.31	0.74 ± 0.31
QCD fakes	0.61 ± 0.75	$0.504^{+0.001}_{-0.004}$	0.31 ± 0.38	0.30 ± 0.37
Diboson	0.19 ± 0.04	0.5 ± 0.0	0.09 ± 0.02	0.09 ± 0.02
Mistag	0.33 ± 0.10	0.5 ± 0.0	0.17 ± 0.05	0.17 ± 0.05
Singletop	0.55 ± 0.08	0.51 ± 0.01	0.28 ± 0.04	0.27 ± 0.04
Total	3.15 ± 0.99	0.503 ± 0.002	1.59 ± 0.50	1.57 ± 0.49
Signal	72.09 ± 12.73	$0.569^{\pm 0.004(stat)}_{\pm 0.010(sys)}$	41.02 ± 7.28	31.07 ± 5.54

DIL (1.5 fb^{-1})				
Drell-Yan	$0.15^{+0.31}_{-0.15}$	0.5 ± 0.0	$0.08^{+0.15}_{-0.08}$	$0.08^{+0.15}_{-0.08}$
Fakes	$0.71^{+1.43}_{-0.71}$	0.52 ± 0.02	$0.37^{+0.74}_{-0.37}$	$0.34^{+0.69}_{-0.34}$
Diboson	$0.09^{+0.9}_{-0.19}$	0.5 ± 0.0	$0.05^{+0.09}_{-0.05}$	$0.05^{+0.09}_{-0.05}$
Total	$0.96^{+1.47}_{-0.73}$	$0.513^{+0.016}_{-0.014}$	$0.49^{+0.76}_{-0.38}$	$0.47^{+0.71}_{-0.35}$
Signal	13.44 ± 1.60	$0.587^{\pm 0.006(stat)}_{\pm 0.013(sys)}$	7.89 ± 0.96	5.55 ± 0.69

Total Bckg.	$4.11^{+1.77}_{-1.23}$	0.505 ± 0.005	$2.08^{+0.91}_{-0.63}$	$2.04^{+0.86}_{-0.61}$
Total Signal	85.54 ± 12.83	$0.572^{\pm 0.003(stat)}_{\pm 0.008(sys)}$	48.91 ± 7.35	36.62 ± 5.58

Table 4.12: Purities and number of expecten SM like (N^+) and XM like (N^-) events for each background and signal for the sample of 1.5 fb^{-1}

MC and data needs to be taken into account. This study was explained previously, and as a result of that the scale factor $SF_{JQ} = 1.03 \pm 0.01 \pm 0.04$ was obtained. Apart of this, there is another scale factor (SF_{nonb}) to take into consideration, which corresponds to the mistags (see section 4.6.4). To include all the effects, each term that is contributing to SM like case was extracted and the expression for the combined purity (final signal purity) reads:

$$
\begin{aligned}
p_{comb} &= f_{nonb} \cdot SF_{nonb} \cdot p_{nonb} + (1 - f_{nonb} \cdot SF_{nonb}) \cdot \\
&\quad (p_{pair} \cdot p_{JQ} \cdot SF_{JQ} + (1 - p_{pair}) \cdot (1 - p_{JQ} \cdot SF_{JQ})) \quad (4.15)
\end{aligned}
$$

where f_{nonb} is the fraction of b-tagged jets which do not really come from b-quark, SF_{nonb} is the scale factor accounting for mistag rate, p_{nonb} is the purity of the non-b events, p_{pair} is the purity of pairing, p_{JQ} is the purity of jet flavor tagging and SF_{JQ} is the scale factor accounting for difference in jet flavor tagging performance between MC and data.

The statistical and systematic uncertainties for the final signal purity were calculated using a standard error propagation based on the same formula (4.15). Table 4.13 lists each variable contributing to the P_{obs} calculation with its asigned uncertainties.

	DIL	LJ
f_{nonb}	0.078 ± 0.002	0.077 ± 0.001
SF_{nonb}	1.05 ± 0.05	1.05 ± 0.05
p_{nonb}	0.5 ± 0.01	0.5 ± 0.01
p_{pair}	$0.930 \pm 0.002(stat) \pm 0.015(sys)$	$0.831 \pm 0.001(stat) \pm 0.028(sys)$
p_{JQ}	$0.604 \pm 0.004(stat) \pm 0.010(sys)$	$0.607 \pm 0.002(stat) \pm 0.004(sys)$
SF_{JQ}	$1.01 \pm 0.01(stat) \pm 0.02(sys)$	$1.01 \pm 0.01(stat) \pm 0.02(sys)$

Table 4.13: Variables used for calculation of signal purity

Finaly, after all the pieces fit togheter, the total amount of background and signal events for both channels can be calculated considering two measurements of b-jet charge per event. The total purities are calculated from the values for each channel, where the amount of each background is taken into account. These values, which are summarized in Table 4.14 are used as input to the statistical treatment of the analysis.

N_s	171.07 ± 25.66
N_b	8.23 ± 3.55
p_s	$0.572 \pm 0.003(stat) \pm 0.008(sys)$
p_b	0.505 ± 0.005

Table 4.14: Expected total number of signal and background events and their total purities.

4.7.2   1-jet bin test

A so called 1-jet bin test in lepton plus jets channel was run to cross-check the method selected for the top quark charge analysis. Some modifications to the applied selection criteria have been made. Instead of four jets, there is only one jet required to be present and this jet must be tagged as b-jet using Tight SecVtx algorithm (to eliminate mistag contribution). Top Mass Fitter is not used and this jet is directly paired with the lepton in the event. The total number of events, eficiency and purity is calculated from MC samples and compared to 1 fb^{-1} of experimental data. The expected number of signal and background events is taken from the cross section analysis for 695 pb^{-1} [40] and scaled to 1 fb^{-1}.

This test was used as a final check of all the procedures used in top quark charge analysis, to make sure there are no inconsistencies between MC and experimental data, before proceeding with the main analysis.

Table 4.15 shows the expected number of background and signal in the 1-jet bin lepton plus jets channel, the purities for each background and signal and the expected number of SM like and XM like events. The last line in the table represents the numbers obtained from experimental data.

From all the backgrounds, the asymmetry in this jet bin is expected to come only from W + c. The asymmetry is diluted due to other symmetric parts of the background. Even for the singletop production and signal $t\bar{t}$ events, there is no asymmetry expected. In this events, there are always two b-jets present. As the selection cut requires only one b-jet, it means that the second b-jet was not reconstructed for some reason. Which of the two jets is picked up is a random event in this case and thus the corelation is lost.

More events observed than was expected with regard to data. This difference is probably caused by using the background signal and background cross-section prediction numbers based on smaller sample and scaled to account for the data of total luminosity 1 pb^{-1}. But still the numbers are consistent within the errors. Also the total expected purity for signal plus background is consistent with the purity observed in the data.

background	N_b	purity	N_+	N_-
Wbb + Wcc	132.26 ± 33.9	0.5 ± 0.0	66.13 ± 16.9	66.13 ± 16.9
Wc	99.23 ± 26.61	0.615 ± 0.012	61.03 ± 16.41	32.2 ± 10.31
QCD fakes	42.25 ± 10.05	0.5 ± 0.0	21.18 ± 5.03	21.18 ± 5.03
Diboson	4.47 ± 0.69	0.5 ± 0.0	2.23 ± 0.35	2.23 ± 0.35
Mistag	94.05 ± 19.28	0.5 ± 0.0	47.03 ± 9.64	47.03 ± 9.64
Singletop	6.86 ± 1.71	0.5 ± 0.0	3.43 ± 0.85	3.43 ± 0.85
Signal	2.51 ± 0.32	0.5 ± 0.0	1.26 ± 0.16	1.26 ± 0.16
Total	381.63 ± 48.31	0.53 ± 0.007	202.29 ± 25.96	179.46 ± 22.61
Data	434	0.546 ± 0.024	237	197

Table 4.15: Comparison of MC prediction of purities to data in 1-jet bin

4.8 Statistical treatment

4.8.1 Profile likelihood method

The statistical treatment used for top quark charge measurement follows the method outlined in [41]. The frequentist approach is chosen. The measurement contains several nuisance parameters that parametrize the size of signal and background and their purities, therefor the profile likelihood method was adopted.

The basic assumption is that there is a hypothesis or probability model for the experimental data which depends on a set of parameters of interest $\pi = (\pi_1, \cdots \pi_k)$ and also on a set of additional nuisance parameters $\theta = (\theta_1, \cdots \theta_l)$. If the probability density function for the model is denoted by $f(x|\pi, \theta)$ and there is a set of independent measurements $X = (X_1, \cdots X_n)$ then the likelihood function can be writen as

$$L(\pi, \theta|X) = \prod_i f(X_i|\pi, \theta) \qquad (4.16)$$

A standard technique for constructing confidence intervals is to find a corresponding hypothesis test. The hypothesis test can be expressed as verification of the hypothesis formulated as: $\pi = \pi_0$ versus the hypothesis $\pi \neq \pi_0$ and a test can be based on the profile likelihood ratio:

$$\lambda(\pi_0|X) = \frac{max(L(\pi_0, \theta|X); \theta)}{max(L(\pi, \theta|X); \pi, \theta)} \qquad (4.17)$$

The maximum in the denominator is found over the full space of parameters (over all values of the nusisance parameters and parameters of interest), while the maximum in the numerator is found only over the subspace with

$\pi = \pi_0$. Now the ratio λ is a funcion of π_0 and the data X only and not of the nuisance parameters. Acording to the theory $-2\log\lambda$ has an approximate χ^2 distribution with k degrees of freedom.

In the particular case of the top quark charge analysis there is only one parameter of interest $\pi = f_+$ – fracton of events with the signature of the SM hypothesis. The end result of analysis will be a χ^2 distribution as a function of f_+. In the analysis there are four nuisance parameters:

- N_s - expected number of signal events after lepton - jet pairing and flavor tagging

- N_b - expected number of all the background events added together after paring and flavor tagging

- p_s - expected purity of the pairing and flavor tagging methods on the signal

- p_b - expected charge asymmetery for all the backgrounds combined together

These parameters are known with certain uncertainties and this fact should be taken into account at statistical treatment – in particular at finding the profile likelihood 4.17 the values of nuisance parameters are generated using the corresponding Nuisance parameter distributions.

Concerning the independent measurement in the analysis the data vector has two componets (x^+, x^-), where x^+ is the number of observations (signal + background) with the SM signature, while x^- is that with the XM signature. They can be looked at as a realisation of a Poissonic random vector (X_+, X_-). At the same time using the above mentioned nusisance parameters and the parameter of interest (f_+), the vector component mean values can be found as follows:

$$
\begin{aligned}
< N_+ > &= p_s N_s f_+ + (1 - p_s) N_s (1 - f_+) + p_b N_b \\
< N_- > &= (1 - p_s) N_s f_+ + p_s N_s (1 - f_+) + (1 - p_b) N_b
\end{aligned}
\qquad (4.18)
$$

The likelihood L used in the top quark charge measurement is made out of 5 parts:

$$
L = L_s \cdot L_b \cdot L_{su} \cdot L_{p_b} \cdot L_{p_s}
\qquad (4.19)
$$

These parts correspond to:

- L_s - the observed signal which is Poisson distributed (and includes both the true signal and the background),

- L_b - the background which is Gaussian distributed,

- L_{su} - the part related with the uncertainties of the number of signal events which is also Gaussian distributed,

- L_{p_s} - the purities for signal, which are also Gaussian distributed.

- L_{p_b} - the background charge asymmetry (background purity).

The signal part can be expressed as:

$$L_s = \frac{<N_+>^{x^+} e^{(-<N_+>)}}{x^+!} \frac{<N_->^{x^-} e^{(-<N_->)}}{x^-!} \qquad (4.20)$$

Using the above mentioned definitions the likelihood for the signal can be rewriten

$$L_{s_+} = \frac{(p_s N_s f_+ + (1-p_s) N_s (1-f_+) + p_b N_b)^{x^+}}{x^+!} \cdot \\ \frac{e^{-(p_s N_s f_+ + (1-p_s) N_s (1-f_+) + p_b N_b)}}{x^+!} \qquad (4.21)$$

and

$$L_{s_-} = \frac{((1-p_s) N_s f_+ + p_s N_s (1-f_+) + (1-p_b) N_b)^{x^-}}{x^-!} \cdot \\ \frac{e^{-((1-p_s) N_s f_+ + p_s N_s (1-f_+) + (1-p_b) N_b)}}{x^-!} \qquad (4.22)$$

The second part of the likelihood is the background term:

$$L_b = \frac{1}{\sigma_{N_b} \sqrt{2\pi}} \cdot e^{-\frac{(y_b - N_b)^2}{2\sigma_{N_b}^2}} \qquad (4.23)$$

The third part is the signal uncertainty term:

$$L_{su} = \frac{1}{\sigma_{N_s} \sqrt{2\pi}} \cdot e^{-\frac{(y_s - N_s)^2}{2\sigma_{N_s}^2}} \qquad (4.24)$$

The fourth part is the signal purity term:

$$L_{p_s} = \frac{1}{\sigma_{p_s}\sqrt{2\pi}} \cdot e^{-\frac{(z_{p_s}-p_s)^2}{2\sigma_{p_s}^2}} \tag{4.25}$$

The fifth part is the background charge asymmetry term:

$$L_{p_b} = \frac{1}{\sigma_{p_b}\sqrt{2\pi}} \cdot e^{-\frac{(z_{p_b}-p_b)^2}{2\sigma_{p_b}^2}} \tag{4.26}$$

where

- N_s, N_b, p_s, p_b will be returned by the fit

- x^+ - number of events following the SM hypotesis in the data

- x^- - number of events following the XM hypotesis in the data

- y_b, σ_{y_b} - number of background events and its error from background prediction studies

- y_s, σ_{y_s} - number of signal events and its error from signal prediction studies

- $z_{p_s}, \sigma_{z_{p_s}}$ - purity of the signal and its error from the purity studies in MC and data

- $z_{p_b}, \sigma_{z_{p_b}}$ - charge asymmetry and its error for the backgrounds from background studies

To obtain the value of f_+ that corresponds to the observed data (x^+, x^-), the partial derivates of the total profile likelihood

$$\lambda = \lambda(f_+; N_s, N_b, p_s, p_b | (x^+, x^-)) \tag{4.27}$$

with respect to each nuisance parameter should be taken and required to be equal to 0. The system of nonlinear equations is obtained. Since it cannot be solved analytically, the likelihood $-2\log L$ is minimized using MINUIT fitter over the whole space of nuisance parameters for a number of values of f_+ selected to cover the whole range of possible f_+ values. As the result, the $-2\log\lambda$ is obtained, which is only a function of f_+.

4.8.2 The analysis deliverables

An importand part of any statistical treatment is a definiton of what the analysis is trying to answer. In case of the top quark charge measurement, the sought answer is if the data is composed of SM top like events or XM like events. There is a need to make sure that the statistical treatment is saying whether the analysis has enough sensitivity to exclude either hypotesis. The end result of the profile likelihood method described above is $-2\ln L$ distribution as a function of f_+ for the data. If there was enough statistics, the value of f_+ at the minimum of $-2\ln L$ distribution could be read off and asigned an error assuming that the shape of $-2\ln L$ is indeed a χ^2 shape.

The practical goal is to decide which of the two hypothesis (the SM hypothesis versus XM one), treated exclusively, is true. The test static of the analysis, i.e. the random variable used to make this decision, is variable f_+. Taking the SM hypothesis as the null hypothesis of the analysis, its acceptance and rejection regions need to be found, expressed in terms of the test static intervals. The approach is following:

Make a distribution of f_+ (the test static) at the minimum $-2\ln L$ for each pseudo-experiment represented by a generated pair (x^+, x^-). The result are one distribution for pseudo-experiment generated accordingly to SM $(g(f_+|SM))$ and one distribution for pseudo-experiment generated accordingly to XM $(g(f_+|XM))$.

Practically it means that in the pseudoexperiments the couples (x^+, x^-) were sampled from the distribution 4.20 at the same time the mean values for this distribution were found using 4.18 assuming the SM fypothesis $(f_+ = 1)$ or the exotic one $(f_+ = 0)$. After that minimizing $-2\ln L$ the value of f_+ was extracted for each generated couple (x^+, x^-).

The decission on which of the two hypothesis is true can be expressed in term of significance level (α) and power of the test $(1 - \beta)$.

- $\alpha = P(rejectSM|SM)$ – probability of rejecting SM hypothesis, if SM hypothesis is true.

  As the f_+ is the test statistic for the top quark charge analysis, the cut f_{+cut} is chosen such that the probability of observing $f_+ < f_{+cut}$, provided the SM hypothesis is true, is equal to alpha.

$$\alpha = \int_{-\infty}^{f_{+cut}} g(f_+|SM) \tag{4.28}$$

- $\beta = P(acceptSM|XM)$ – probability of accepting SM hypothesis, if XM hypothesis is true.

For each given value of α (each f_{+cut}) the probability of SM theory being accepted when XM theory is true, can be calculated.

$$\beta = \int_{f_{+cut}}^{\infty} g(f_+|XM) \tag{4.29}$$

The value of $1 - \beta$ is thus called the power of the test to discriminate against the XM hypothesis.

P-value.

One of the basic terms used in statistical testing of a hypothesis is the p-value. It expresses the probability of obtaining a result at least as extreme as a given data point, provided that the null hypothesis is true.

For any result, i.e. data pair (x^+, x^-) (measured or sampled) we can calculate the p-value assuming that a certain hypothesis is valid. Though the p-value can be calculated for any of the possible hypotheses ($f_+ \in\, <0, 1>$), for practical reasons it has a sense to calculate it for the SM hypothesis ($f_+ = 1$) and exotic one ($f_+ = 0$)). At the p-value calculation for a given realisation (x^+, x^-) of our observables we calculate the probability that, assuming a given hypothesis ($f_+ = 1$ or $f_+ = 0$), a realisation (x_i^+, x_i^-) of the random data vector $\vec{X}$ will occur with probability less or equal to the probability of (x^+, x^-), i.e. $L_s(x_i^+, x_i^-|f_+ = c) \leq L_s(x^+, x^-|f_+ = c)$, where $c = 1$ ($c = 0$) correcsponds to SM (XM) hypothesis.

As for any data pair (x^+, x^-) the value of the used test static $(f_+)_{rec}$ can be unambiguously reconstructed. Then the p-value can be expressed in terms of reconstructed f_+ instead of (x^+, x^-). The reconstructed f_+ distribution for the SM $g(f_+|SM)$ and XM $g(f_+|XM)$ are shown in 4.25.

If the $(f_+)_{observ}$ is the value of thest static corresponding to the observed data $(x^+, x^-)_{observ}$ then the p-value can be expressed as:

$$p_{value} = \int_{-\infty}^{(f_+)_{observ}} g(f_+|SM) \tag{4.30}$$

As only values of $f_+ < 1$ are taken into account, the so called one-tail p-value is used. It should be noted that the ultimate goal of the analysis is to calculate the p-value for the experimentaly observed pair (x^+, x^-).

With the distributions of f_+ for XM and SM, the pseudo-experiments generated accordingly to the XM can be used, and for each one of them the area under the SM distribution is looked at. The distribution of p-values according to the SM distribution is obtained. After this, a value for α is chosen - the probability of wrongly rejecting SM. The alpha value is wanted

to be small, for analysis it was chosen to be 1%. Then, for the chosen value of $\alpha = 0.01$ the value of β is calculated.

Once the value of f_+ for the actual data is obtained, its p-value using the SM distribution is extracted. If the p-value is greater than 0.01, it can be said that the analysis excludes the XM at the confidence level equal to $1 - \beta$ already determined from sensitivity study. If the p-value is less than 0.01, the analysis excludes the SM at the 99% confidence level.

Bayes factor.

Another observable is to look at how likely the SM hypothesis is compared to XM and build the likelihood ratio. In this Bayesian approach the systematic uncertainties are integrated over. No minimization is done. The likelihoods at $f_+ = 0$ and $f_+ = 1$ are evaluated and integrated over the nuisance parameters. Then the ratio, called the Bayes Factor (BF), is taken. Looking at the SM and exotic hypotheses as mutually exclusive (either SM or exotics is true) the Bayes factor for these hypotheses can be calculated as follows:

$$F_{Bayes} = \frac{P(x^+, x^- | f_+ = 1)}{P(x^+, x^- | f_+ = 0)} \tag{4.31}$$

where $P = L_s \cdot G_{N_s} G_{N_b} G_{p_s} G_{p_b}$ (G_X is Gaussian distribution of the nuisance parameter X).

The value obtained can be interpreted as how likely the SM (XM) is, compared to the XM (SM). By taking $2\ln(BF)$ the number similar to a χ^2 is obtained and compared to following scale:

- 0-2: not worth mentioning,

- 2-6: positive evidence,

- 6-10: strong evidence,

- 10-: very strong evidence.

Figure 4.25 shows the f_+ distribution for the SM and the XM based on the expectation. Figure 4.26 shows the distribution of p-values for events generated acording to XM but reconstructed assuming SM is true. The area under the p-value curve at $\alpha = 0.01$ is $1 - \beta = 0.87$ which is the sensitivity of the analysis.
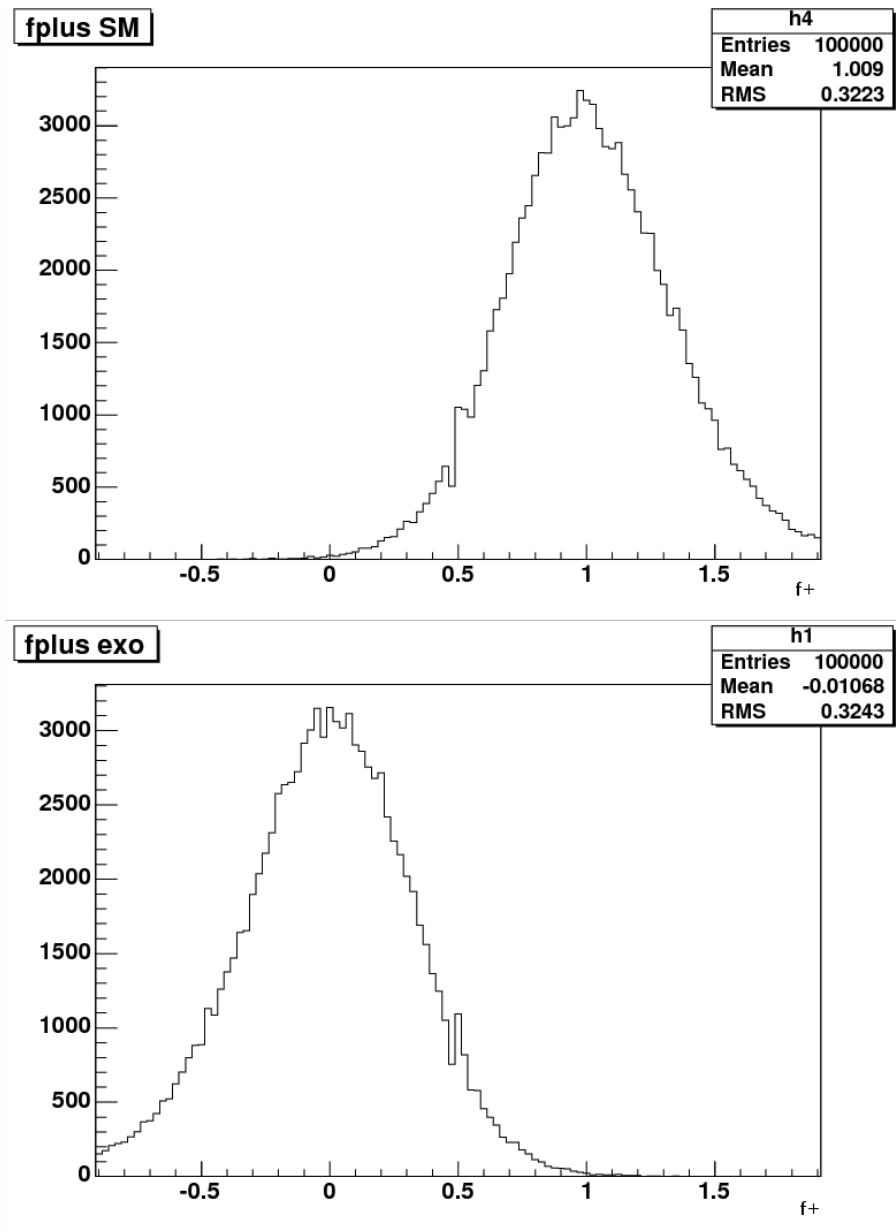
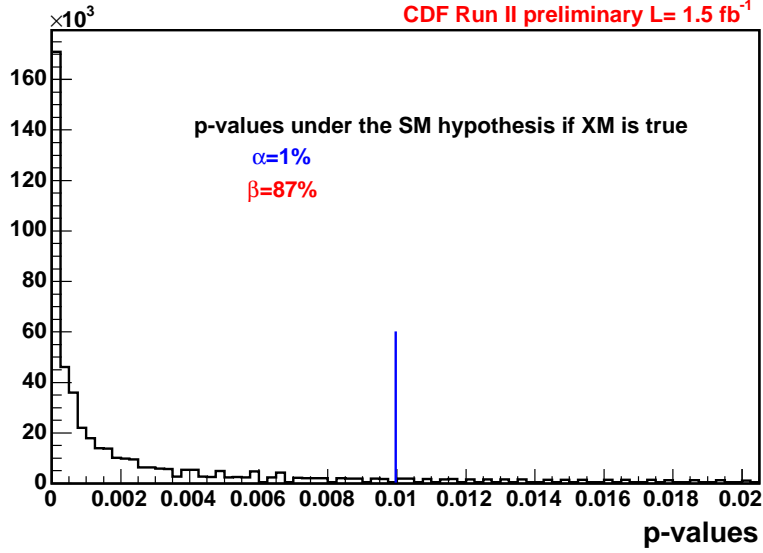Figure 4.25: Distribution of f_+ assuming the SM (top) and XM (bottom) using parameters from the expectation

Figure 4.26: p-value distribution according to the SM if XM is true.

4.9 Results

After finishing all the optimization and calibration studies the experimental data from CDF detector were analyzed and the statistical significance of observed results was estimated.

In the table 4.16 the final numbers of observed pairs for dilepton and lepton + jets channels are listed. In dilepton channel, after applying all the selection criteria, there are 26 lepton - b-jet pairs left. 13 of them look like SM and 13 look like XM. In lepton + jets the number of lepton - b-jet pairs after selection criteria have been applied is 199, 111 pairs are SM like and 88 are XM like. Overall there are 225 pairs, from which 124 look like SM and 101 look like XM.

Figure 4.27 shows the charge distribution of $Q_W * Q_b$ for the MC and experimental data events. In case of SM like pairs, the b-jet charge is anticorrelated with the W charge, in case of XM like pairs, the charges are correlated.

Figure 4.28 shows the profile likelihood function calculated for the actual amount of expected signal and background and final purities. For the observed number of SM and XM like pairs in experimental data, the minimum is found at the value of $f_+ = 0.87$. The corresponding p-value under SM hypothesis is $p = 0.31$. As this value is greater than the value of $\alpha = 0.01$ the XM hypothesis is excluded with the confidence of $1 - \beta = 87\%$.

Figure 4.29 shows the probability distributions for SM and XM hypothesis as the function of f_+. The observed value of $f_+ = 0.87$ is shown.

Using the bayesian approach, the Bayes factor is calculated using the final observed numbers. From the value of $2\log(BF) = 12.01$ using the scale mentioned in previous section, ti can be said that the experimental data *very strongly favors* standard model over exotic one.

Yield	Observed	After pairing	JQ defined	SM	XM
L + J	193	102	199 pairs	111	88
DIL	44	14	26 pairs	13	13
Total	237	116	225 pairs	124	101

Table 4.16: The final number of events obtained from data
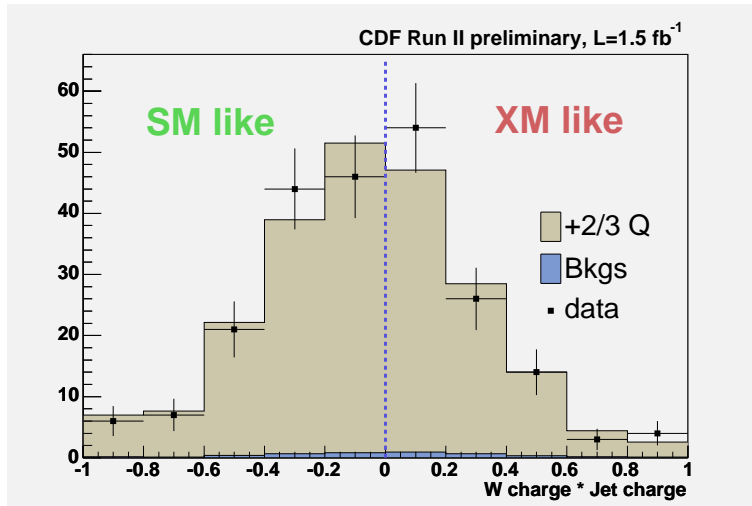

Figure 4.27: W-charge * Jet charge for $t\bar{t}$ events
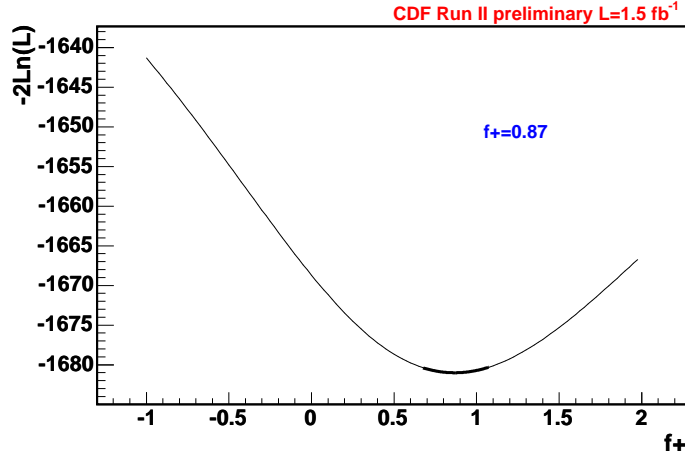
Figure 4.28: Profile Likelihood function for f_+. For the observed number of SM and XM like pairs $(124, 101)$, the profile likelihood function was calculated for selected values of f_+ covering the whole range of possible f_+ values. The minimum was found for the $f_+ = 0.87$.
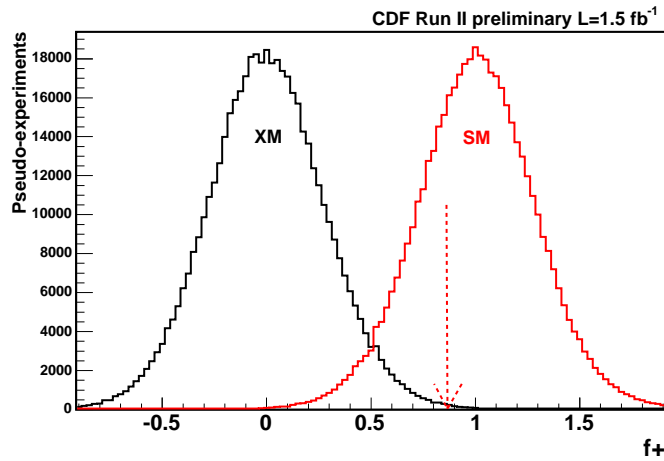

Figure 4.29: Probability distribution for XM and SM hypothesis. The value of $f_+ = 0.87$ is show, for which the minimum of the profile likelihood function was found. The area under SM curve up to the this value of f_+ is the $p - value = 0.31$.

Chapter 5

Final Summary

This thesis has presented the first CDF top quark charge measurement. The goal was to decide if the top quark discovered on the Tevatron experiments is really the particle predicted by SM with the charge 2/3 or if it is, as suggested by some alternative theories, an exotic quark with the charge -4/3.

The basic idea of the measurement was based on the analysis of the electric charges of the $t\bar{t}$ decay products ($t \to Wb$). The procedure consists of the following basic steps:

- correct association of W-boson and b-quark from the same top decay (lepton - b-jet invariant mass criteria for the dilepton and kinematic fit for the lepton + jets channel);

- determination of W-boson charge (the sign of the lepton from leptonic decay);

- determination of the charge of b-jet using weighting method.

As the final output of the procedure there are lepton - b-jet pairs with corresponding b-jet charges. Under the SM hypothesis, the charges of associated lepton and b-quark should have the opposite signs while under the XM hypothesis, the signs of the charges should be the same.

In case of the b-jets, the spectrum of charges evaluated by the weighting procedure has a wide spread and the anticorrelation (correlation) of the paired lepton and b-jet's charges in SM (XM) can be violated, but it should be valid for the mean value of the charge spectrum of the paired jet.

After applying the whole procedure, there are 225 pairs, out of which 124 look like SM and 101 look like exotic model.

After the statistical treatment, taking into account all the sensitivity and systematic studies, the result is that the data exclude XM with confidence

level of 87%. Using Bayesian approach, it can be said that the experimental data favors strongly SM over XM.

To improve this results, different methods of pairing and b-jet charge determination can be investigated. For pairing one could combine the information from the Top Mass Fitter with the lepton - b-jet invariant mass calculation, to increase the purity of the pairing. For b-jet charge determination, the template method, neural network or semileptonic B-meson decay could be used. However all these methods require higher statistics of experimental data than it is available now. Some of them would become possible when the CDF collects the planed $4 - 8$ fb$^-1$ amount of data. The others could be applied in the future experiments on LHC, where the amount of $t\bar{t}$ events will be much higher.

Bibliography

[1] F. Halzen, A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*, 1984.

[2] The CDF II Collaboration. *The CDF II Detector Technical Design Report*, 1996. FERMILAB-Pub-96/390-E.

[3] A. Mukherjee and R. Wagner. *COT Central Outer Tracker*, 2003. CDF/PUB/TRACKING/PUBLIC/6267

[4] F. Ascoli et al., *The CDF Central Muon Detector*, Nucl. Instrum. Meth. A 268 (1998) 33.

[5] J. Goldstein et al., *Silicon Tracking for Plug Electrons*, 2002. CDF Note 5970

[6] Abe et al.(CDF collaboration),Phys.Rev.Lett. 74, 2626(1995).

[7] S. Abachi et al.(D0 collaboration),Phys.Rev.Lett. 74, 2632(1995).

[8] A. Denner and T. Sack, Z. Phys. C 46 (1990) 653.

[9] R.K.Ellis, W.J.Stirling, B.R.Webber, QCD and Collider Physics, Cambridge Univ. Press, Cambridge CB2RU, 1996

[10] M. Mangano, Introduction to QCDin Hadronic Collissions, 2005 CERN-CLAF School of HEP,CERN-2006-015, Dec 2006, CERN

[11] L.N.Lipatov, Sov. J. Nuc. Phys. 20(1975) 95; V.N. Gribov and L.N.Lipatov, Sov. J. Nuc. Phys. 15(1972) 438; G. Altarelli and G. Parisi, Nucl. Phys. B126(1977)298; Yu.L. Dokshitzer, Sov. Phys. JETP 46(1977)641.

[12] E. Laen, J. Smith and W. L. van Neerven, Phys. Lett. B321 (1994) 254.

[13] N. Kidonakis and R. Vogt, Eur. Phys. J. C33, 466 (2004).

[14] N. Kidonakis, Phys. Rev. D64,014009 (2001).

[15] N. Kidonakis et al., Phys. Rev. D64, 114001 (2001).

[16] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason and G. Ridolfi, *The t anti-t cross-section at 1.8 TeV and 1.96 TeV: A study of the systematics due to parton densities and scale dependence*, JHEP 0404, 068 (2004).

[17] G. Chiarelli, Int. J. Mod. Phys. A 13 (1998) 2883.

[18] M. Mangano, Yellow report, CERN-2000-004 (1999), QCDF section, p.445.

[19] V. M. Abazov et. al., Phys. Rev. Lett. 98 (2007) 181802.

[20] T. Affolder et al. (CDF Collaboration), Phys. Rev. D63 (2001) 032003.

[21] A. Denner and T. Sack., Nucl. Phys. B358 (1991) 46.

[22] K. Hagiwara et al. (Particle Data Group Collaboration), Phys. Rev. D66 (2002) 010001.

[23] C. Quigg, hep-ph/9802320.

[24] A. Abulencia et al. (The CDF Collaboration), Phys. Rev. D75 (2007) 052001.

[25] N. M. Abazov et al., Phys Lett. B617,1 (2005).

[26] B. Lampe, hep-ph/9512276.

[27] M. Ciljak, M. Jurcovicova, S. Tokar, U. Bauer, *Top charge measurement at ATLAS detector*, ATLAS Note, PHYS-2003-35 (2003).

[28] D. Chang, W.F.Chang, E. Ma, Phys.Rev.D59, 091503(1999) and Phys.Rev.D61, 037301(2000).

[29] T. Maki, J. Antos, A. Beretvas, Y.C. Chen, R. Lysak, *Top Mass Measurement id Dileptonic Channel using Template Method in 1.2 fb^{-1}*, CDF Note 8727 (2007).

[30] Franklin, Grinstein, Guimaraes da Costa, Sherman, *Measurement of the Top Pair Cross Section in Lepton+Jets in 1.12 fb^{-1}*, CDF Note 8767 (2007).

[31] D. Acosta et al. (CDF Collaboration), Pys. Rev D71, 052003 (2005).

[32] T. Wright, *SecVtx B-tag Efficiency using Muon Transverse Momentum*, CDF Note 7448 (2005).

[33] F. Garberson, C. Hill, J. Incandela, C. Neu, *Determination of Gen 6 SecVtx Btag Efficiency Scale-Factor Using the Muon Relative Pt Method*, CDF Note 8014 (2006).

[34] D. Bortoletto et al., *ID Efficiency for Intermediate pT Muons*, CDF Note 7197 (2005).

[35] J.P. Chou, J. Guimaraes da Costa, D. Sherman, *Studies of 6.1.4 Monte Carlo for Top Analyses*, CDF Note 8426 (2006).

[36] A. Bhati et.all, *Jet Energy Corrections at CDF*, CDF Note 7543 (2005).

[37] CDF procedure for estimation of systematic uncertainties due to PDF. http://www-cdf.fnal.gov/internal/physics/joint_physics /instructions/PDFUncertainties/pdf.html

[38] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, J. High Energy Phys., 007 (2002).

[39] A. D. Martin et al., *The α_s dependence of parton distributions*, Phys. Lett. B356 (1995).

[40] J. Guimaraes da Costa, D. Sherman, *Measurement of the Top Pairl Production Cross Section in 700 pb^{-1} with Loose SecVtx*, CDF Note 8120 (2006).

[41] W. A. Rolke, A. M. Lopez, J. Conrad, Nucl. Instrum.Meth. A551 (2005) 493-503.