Dottorato di Ricerca in Fisica – Ciclo XIX

Università degli studi di Bologna *"Alma Mater Studiorum"*
Dipartimento di Fisica "Augusto Righi"

Measurement of the top quark mass in the all-hadronic channel at proton antiproton collisions at $\sqrt{s} = 1.96\,\mathbf{TeV}$ with CDF II

Coordinatore:
Ch.mo Prof. Fabio Ortolani

Relatore:
Ch.mo Prof. Andrea Castro

Dottorando:
Dott. Fabrizio Margaroli

Contents

List of Figures

List of Tables

Preface

We report a measurement of the top quark mass in the all-hadronic channel with the upgraded Collider Detector at Fermilab using an integrated lumininosity of 1.02 fb^{-1}. Top quarks are produced mostly in pairs at the Tevatron Collider; they subsequently decay almost 100% of the time in a W boson and a b quark each. Final states are classified according to the decays of the W bosons. Here we study only those events in which both W's decay into a quark pair. This channel has the advantage of a large branching ratio and of being fully reconstructed. On the other hand the signal is overwhelmed by a background which surpasses it by three orders of magnitude even after the requirement of a specific trigger. A neural network is thus used to discriminate signal from background in order to achieve a reasonable signal over background ratio. We look then for a variable which is strongly correlated with the top quark mass, and compare the corresponding distribution in the data to signal and background templates in order to measure the top quark mass. Here is an outline of the following of this work: we discuss in Chapter 2 the need for a 6th quark in the Standard Model, the phenomenology of the top quark and a summary of the current experimental knowledgeof its properties, and motivate the need for a precise measurement of its mass. In Chapter 3 we describe the experimental apparatus which is needed in order to produce top quarks, i.e. the Tevatron Collider, and the detector which collects the data analyzed in this work, CDF II. Chapter 4 describes how CDF II interprets the output from the various subdetectors and translate them into the physics objects which are needed for the measurement. In Chapter 5 we present how we overcome the problem of the huge background which overwhelms the top production by using for the first time for the event selection a neural network approach. Finally, in Chapter 6 we present and discuss the technique used to extract the top quark mass measurement, how we control this technique, evaluate the systematics uncertainties, and finally apply the method to the data to obtain the top quark mass measurement.

Chapter 1

Introduction

In this chapter we present a brief history of how the quest for the intimate properties of nature led to the development of one of the most predictive (and beatiful) theories of human history: the Standard Model of fundamental interactions.

1.1 A historical framework

The idea that all matter is composed of elementary objects dates back at least the to 6th century BC. The philosophical doctrine was proposed by ancient Greek philosophers such as Leucippus, Democritus, and Epicurus who believed that nature is composed of small blocks of matter, the $\alpha\tau o\mu o\xi$ (meaning "indivisible"). The idea proved fruitful for scientific research only in relatively recent times: in the 19th century J. Dalton, through his work on stoichiometry, concluded that each element of nature was composed of a single, unique type of particle which has been promptly named atom. In the second half of the century, the Russian chemist D. Mendeleev noticed the periodicity of the properties of the known elements and proposed the well-known periodic table; in addition, he predicted the properties of many elements yet to be discovered. However, near the end of the century, physicists discovered that atoms were not, in fact, the fundamental particles of nature, but conglomerates of even smaller particles. It was in fact the work of J. J. Thomson that led to the discovery of a particle which is still believed today to be a fundamental one, the

electron. The discovery of the electron opened the way to the understanding of the atom structure; the discovery of the proton and of the neutron singed the birth of nuclear physics. Another major step is the discovery of the positron soon after the Dirac quanto-relativistic theory of spin 1/2 particles predicted the existence of an electron with a positive charge. At that point new particles started to appear in cosmic rays and then at the first colliding machines, leading to a proliferation which in the 50's came to be known as the "particle zoo". Here Gell-Mann made a theoretical work which can be compared historically to the Mendeelev table of elements, in putting back order in what looked like a chaos. Gell-Mann hypothesized that all the hadrons known at that time could be explained in terms of a combination of new mathematical objects, which he named quarks. Particle colliders proved to be a crucial tool in the development of particle physics in providing the mean to confirm the existence of quarks as real objects. A decisive confirmation of the quark theory came with the discovery of the J/Ψ, which was soon interpreted as a quark-antiquark resonant state. The experimental evidence for the existence of quarks soon accumulated. We show in Table 1.1 some of the most important ingredients of the Standard Model. Not all the particles shown are fundamental: the pions and kaons are presented because their discovery shed light the building of the theory, the pions because they are the first mesons produced and long believed to be the carrier of the strong interactions; the kaons because they led to the discovery of the s quark, the parity violation, as well as the CP violation in the weak sector. The result is that the fermion sector of the fundamental particles of nature are particles of two different kinds, the leptons, which interact via gravitational, electromagnetic and weak forces, and the quarks, which participate in all kind of interactions. Quarks and leptons and some of their basic properties are shown in Table 1.2, 1.3.

At the same time, the understanding of the fundamental interactions of nature progressed: after the unification of electric and magnetic interactions with the Maxwell theory, the 20th century saw the systematization of the former into a quanto-relativstic version, the quantum-electrodynamics (QED). The QED, as a gauge theory, i.e. a theory which is invariant underl local phase transformation, soon became a mathematical model which could explain the newly discovered weak and strong interactions. The adaptation of the gauge theory formalism to weak

Particle	***Year***	***Nobel Prize***	***Method***
e^-	1897	Thomson	Discharge in gases
p	1919	Rutherford	Natural radioactivity
n	1932	Chadwick	Natural radioactivity
e^+	1933	Anderson	Cosmic Rays
$\mu^\pm$	1937	Neddermeyer, Anderson	Cosmic Rays
$\pi^\pm$	1947	Powell, Occhialini	Cosmic Rays
$K^\pm$	1949	Powell	Cosmic Rays
π^0	1949	Bjorklund	Accelerator
K^0	1951	Armenteros	Cosmic Rays
ν_e	1956	Cowan, Reines,	Nuclear reactor
ν_μ	1962	Lederman	Accelerator
J/Ψ (c quark)	1975	Ting, Richter	Accelerator
τ	1974-1977	Perl	Accelerator
Υ (b quark)	1977	none (E288 collaboration)	Accelerator
$W^\pm$ & Z^0	1984	Rubbia, van der Meer	Accelerator
t quark	1995	none (CDF & D0 collaboration)	Accelerator
ν_τ	2000	none (DONUT (E872) collaboration)	Accelerator

Table 1.1: Here is a list of some of the most important particles whose discovery led to deeper insight in the understanding of nature. We added to the discovery of the fundamental particles, leptons and quarks, the discovery of pions and kaons, the first because they were long believed to be the carriers of the strong interactions, the second because their studies allowed the discovery of the parity violation and the CP violation in the weak interaction, and the s quark.

Quark	***Electric charge (e)***	***Mass (MeV/c^2)***
up u	$+2/3$	$1.4 - 4$
down d	$-1/3$	$4 - 8$
strange s	$-1/3$	$80 - 130$
charm c	$+2/3$	$1.15 - 1.35 \times 10^3$
beauty b	$-1/3$	$4.1 - 4.9 \times 10^3$
top t	$+2/3$	1.714×10^5

Table 1.2: The three generations of quarks.

Lepton	***Electric charge (e)***	***Mass (MeV/c^2)***
e	-1	0.51099892
ν_e	0	$< 3 \times 10^{-6}$
μ	-1	105.658369
ν_μ	0	< 0.19
τ	-1	1777
$\mathrm{t}\nu + \tau$	0	< 18.2

Table 1.3: The three generations of leptons.

interactions led to an enormous effort by Glashow, Weinberg and Salam which culminated in the unification of electromagnetic and weak interactions, i.e. with the $SU(2) \times U(1)$ theory of electroweak interactions. QCD (quantum-cromodynamics) came out as a consequence of the discovery of an additional quantum number, the color, which is regarded as the charge of the strong interactions. The Standard Model emerge at the end as a $SU(3) \times SU(2) \times U(1)$ gauge theory of the fundamental interactions[1]. The fundamental particles which mediates the 4 known interactions are shown in Table 1.4.

There are three main ideas which led the physicist community to this notable achievement: that "the language of God is mathematics" presented by G. Galilei at the birth of modern science; the related quest for "beauty", which often in mathematics translates into the display of the theory of some kind of symmetry; and Reductionism, the idea that complex structures can be explained in terms of more fundamental constituents[2].

There is only one missing item in the Standard Model, which is unfortunately a very fundamental one. The Higgs boson has been proposed as a result of a mechanism called spontaneous symmetry breaking to give mass to particles which otherwise would be required by the theory to be massless. It is important to stress that while direct searches of this hypothetical particle goes on, the Standard Model allows us to predict its mass within a certain range using the measured parameters

[1]We are ignoring here the presence of gravitation which led to additional problems once quantized.

[2]It is worth noting that holism - to say with its proposer, Aristotle, that "the whole is more than the sum of its parts" - is an idea which lies at the basis of other fields in physics, that of chaotic systems as a primary example.

Boson	***Spin ($\hbar$)***	***Mass (GeV/c^2)***
γ (photon)	1	0
g (gluon)	1	0
$W^{\pm}$	1	80.40
Z^0	1	91.188
graviton	2	?
H^0 (Higgs)	1	???

Table 1.4: The "force carriers" in the Standard Model.

of the theory. The Higgs mass is particularly sensitive to the top quark mass value; this motivates a precise measurement of the latter. We will investigate further this crucial issue in the next Chapter.

Chapter 2

The top quark

The top quark is naturally accounted for in the Standard Model of fundamental interactions. We will address the reasons for this occurrence, and describe the typical features of the top quark production and decays at a hadron collider. In the last section we will finally give a brief review of the measured properties of the top quark and stress the importance of the measurement of its mass.

2.1 The third generation of quarks: the Υ discovery

The first particle of the third generation to be observed was the τ lepton in 1975. Short time later, in 1977, the Υ was discovered at Fermilab as a resonance in the $\mu^+\mu^-$ invariant mass spectrum in the reaction $p + nucleon \rightarrow \mu^+\mu^- + X$. This resonance was soon interpreted as a $b\bar{b}$ bound state. The b quark pole mass has been extrapolated from the masses of the Υ and B mesons. Both its charge and isospin properties ($Q_b = -1/3$ and $I_3 = -1/2$) have then been measured and found to be in agreement with the Standard Model expectations.

The value of the charge of the b quark has been inferred from measurements of the Υ leptonic width at the DORIS e^+e^- storage ring: the width is in fact proportional

to the square of the charge of the b quark and can be estimated quantitatively from heavy quark-antiquark potential models. The charge assignment was subsequently confirmed by the measurement of the ratio R of hadron production cross section over $\mu^+\mu^-$ cross section, $R = \sigma(e^+e^- \to hadrons)/\sigma(e^+e^- \to \mu^+\mu^-)$. At lowest order and ignoring resonance effects, $R = \sum_{quarks} 3Q_q^2$, where the factor of three arises from the fact that quarks come in three colors. The sum is over all quarks that can be produced with mass below one-half the center-of-mass energy of the e^+e^- system.

The weak isospin of the b quark was first extracted from the forward-backward asymmetry (A_{FB}) in $e^+e^- \to b\bar{b}$. This asymmetry is defined in terms of the b quark production cross-section $\sigma(b)$ as

$$A_{FB} = \frac{\sigma(b, \theta > 90^o) - \sigma(b, \theta < 90^o)}{\sigma(b, \theta > 90^o) + \sigma(b, \theta < 90^o)} \tag{2.1}$$

where θ is the polar angle of the b quark in the e^+e^- center of mass as measured from the direction of flight of the e^-. The asymmetry originates from the coupling of the Z to fermions, which in the Standard Model depends on the weak isospin through a term in the Lagrangian of the form $\bar{f}\gamma_\mu(g_V - g_A\gamma_5)Z^\mu f$, where f is the fermion spinor, and the vector and axial couplings $g_{V,A}$ are given by:

$$g_V = \frac{I_3 - 2Qsin^2\theta_W}{2sin\theta_W cos\theta_W} \tag{2.2}$$

$$g_A = \frac{I_3}{2sin\theta_W cos\theta_W} \tag{2.3}$$

where θ_W is the Weinberg angle. This measurement was found to be consistent with the Standard Model prediction.

2.2 Indirect evidence for top quark existence

Beyond the aesthetics lying in the symmetry between lepton and quark families, there are many reasons to advocate the existence of an isospin partner of the b quark.

2.2.1 Absence of FCNC decays

If the b quark were to form a weak-isospin singlet and if there were only 5 quarks then it can be shown that the leptonic branching ratio (BR) is $BR(B \to Xl^+l^-) \geq 0.12$. This was found to be inconsistent with the first upper limits placed on flavour-changing neutral currents in b decays, $BR(B \to Xl^+l^-) < 0.008$ at 90% confidence level.

2.2.2 Absence of triangle anomalies

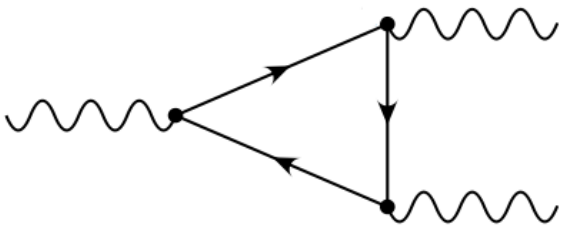


Figure 2.1: Example of triangle Feynman diagram giving rise to anomalies.

Triangle diagrams such as in Fig. 2.1 give rise to divergencies in the theory, thus spoiling the renormalizability of the Standard Model. These divergencies are canceled thanks to the GIM mechanism only if there are the same number of lepton and quark families. In fact, if we are to consider only isospin singlets and doublets, the only anomaly corresponding to triangle diagrams is proportional to $c^f Q_f^2$, where Q_f is the charge and c^f is the axial coupling to the weak neutral current, i.e. the isospin. The three leptonic doublets contribute as in

$$\sum_{leptonic\ doublets} c^f Q_f^2 = \frac{1}{2}(0)^2 - \frac{1}{2}(-1)^2 \tag{2.4}$$

and the quarks as in

$$\sum_{quarks} N_{colors} \times c^f Q_f^2 = 3 \times \left(\frac{1}{2} N_c \left(+\frac{2}{3} \right)^2 - \frac{1}{2} N_c \left(-\frac{1}{3} \right)^2 \right) \tag{2.5}$$

thus summing to zero. These two arguments make the case for an up-type isospin partner of the third generation b quark.

On the other hand, measurements of the Z boson width at the LEP and SLC colliders rule out the existence of the 4th generation neutrino with mass $M_\nu < 4M_Z/2$. The top quark is thus considered to be the last fermion expected in the Standard Model of the fundamental interactions.

2.3 The heaviest fermion

The search for the top quark started right after the discovery of the b quark and lasted almost 20 years when it was finally discovered at the Tevatron collider in 1995 by the CDF [1] and D0 [2] collaborations: its mass proved to be surprisingly large, having been measured at the end of the 1992-1995 data taking period to be equal to 178 ± 4.8 GeV/c^2, i.e. about the mass of a gold atom. With the top quark, the fermion sector of the Standard Model is now believed to be complete. Here we describe how it is produced in a hadron collider and what are the characteristic signatures of its decay that will be exploited in the next chapters.

2.3.1 Top quark production and decays

There are two mechanisms for top quark production in $p\bar{p}$ collisions:

- **pair production** $p\bar{p} \rightarrow t\bar{t} + X$. This process proceeds from two different initial states: at the Tevatron energies, most of the time the process originates from quark-antiquark annihilation (85%), and the rest through gluon-gluon fusion (15%)[1];
- **single production** proceeds mainly through W-gluon fusion and quark-antiquark annihilation;

where the relative Feynman diagrams are shown in Fig. 2.2.

[1]The $t\bar{t}$ pairs can also be produced through a Z or a photon leading to much smaller cross section.

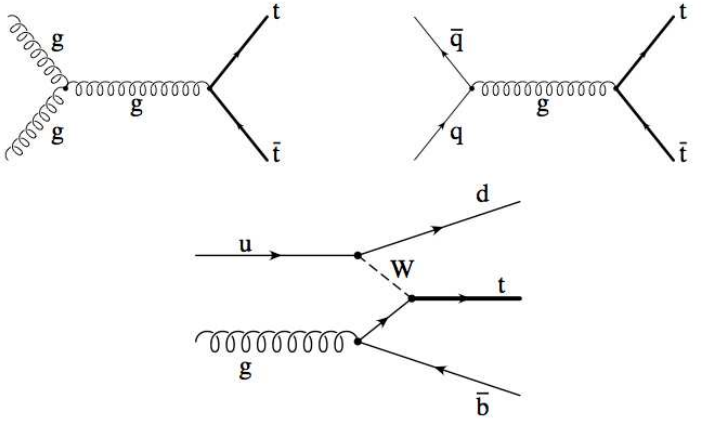


Figure 2.2: Main production mechanisms at the Tevatron.

The dominant production processes are the one where top quarks are produced in pairs, their cross section being twice as large as the single top production; pair production is also the mechanism studied in this work so we will only mention briefly in the following the single top production.

Quarks are not observed as free particles but are confined to form hadronic bound states. The top quark is special with respect to other quarks: being its lifetime $\tau \simeq 10^{-24}\,\mathrm{s}$ shorter than the typical hadronization time, which is estimated to be $\Lambda_{QCD}^{-1} \simeq \mathrm{O}(100\,\mathrm{MeV})^{-1} \simeq \mathrm{O}(10^{-23}\,\mathrm{s})$, the top quark decays before hadronizing. As a consequence we can not observe a $t\bar{t}$ resonant state as it was possible with c and b quarks, but only detect it through its decay products. On the other hand the daughter particles will retain information on the quantum numbers of the parent top quark.

According to the Standard Model, the top quark decays to Wb almost 100% of the times. The decays $t \to Ws$ and $t \to Wd$ are allowed too; but the former is suppressed with respect to the $t \to Wb$ by a factor $V_{ts}^2/V_{tb}^2 \simeq 10^{-3}$, and the latter by a factor $V_{td}^2/V_{tb}^2 \simeq 10^{-4}$. For this reason these rare decays are not considered in this analysis. The W boson decays 1/3 of the times into a lepton and a neutrino and 2/3 of the times into quarks. The final states are thus determined by the decays of the two W bosons; using the calculated branching ratios, (see Table 2.1), we expect

Signature	***Branching ratio***
$q_1\bar{q}'_2 q_3\bar{q}'_4$	36/81
$e + q\bar{q}'$	12/81
$\mu + q\bar{q}'$	12/81
$\tau + q\bar{q}'$	12/81
ee	2/81
$e\mu$	2/81
$e\tau$	2/81
$\mu\tau$	1/81
$\mu\mu$	1/81
$\tau\tau$	1/81

Table 2.1: Branching ratios for the various top pair decay modes.

that the fully hadronic and the semileptonic decays of top quark pairs make up the majority of the events.

Here we describe briefly the typical signatures which pair produced top quarks will leave in the detector:

- **The "dilepton" channel**

  The nominal signature for this channel has two high-P_T leptons, missing transverse energy from the two neutrinos and two jets from the b quarks. The yield is pretty small, mostly due to the low branching ratio in this channel, 5%. Moreover, the presence of two high energy neutrinos complicates the event reconstruction. On the other hand, the background is very small, mainly coming from Drell-Yan events, so we expect a very clean sample.

- **The "lepton+jets" channel**

  The nominal signature has a high-P_T lepton, missing transverse energy from the neutrino, and four jets out of which two are expected to contain B mesons from the hadronization of the b quarks. Without considering the events with τ's, we expect $BR \sim 30\%$. The main background contribution comes from W+multijet production and is large. The difference between light-quark jets and b-jets is exploited here to reduce this contamination. This has been done with two different techniques, the first one looking for highly displaced secondary vertices due to the long lifetime of the B mesons and the second looking

for semileptonic decays of the B mesons. This topology is where we obtain the most precise measurement of the top quark properties, as we will see in the following section.

- **The "all-hadronic" channel**

  In this decay mode we expect six final state jets, four of which come from the hadronic decays of the two W's and two from the b quarks. Approximately 44% of the $t\bar{t}$ events have this decay signature. The major challenge here is to overcome the huge background coming from QCD multijet production. In order to isolate a signal and maintain high efficiency, we require at least six well-separated jets, a specific kinematical selection, and finally ask for the presence of a secondary vertex. This channel is the one used for the measurement in this work.; its characteristics will be explained in more details in Chapter 5.

2.4 Meaurement of the top quark properties

As done for the b quark at that time, a number of measurements have to be performed to estabilish the nature of the newly discovered particle. Due to its high mass, the only existing facility where top quarks can be produced is the Tevatron Collider[2]. Here we present some of these measurements which both confirm the Standard Model nature of the top quark and serve as tests of the Standard Model itself:

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

  An accurate measurement of the $t\bar{t}$ production cross section is a precision test of the Standard Model. A cross section significantly higher than the theoretical expectation would be a sign of non-Standard Model mechanisms, for example the decay of a heavy resonant state into $t\bar{t}$ pairs. The excess observed with respect to the background in the various channels translates immediately into measurements of the pair production cross section. The value

[2]We remind the reader that two detectors, CDF and D0, sit at the two Tevatron collision points.

obtained from the combination of the measurements in the various channel led to the 1992-1995 (Run I) measurement $\sigma_{t\bar{t}} = 7.5^{+1.9}_{-1.6}$ pb to be compared with the theoretical cross section calculation by Laenen *et al.* [3] of 4.8 pb at 176 GeV/c^2. The apparent disagreement between theoretical and experimental values has

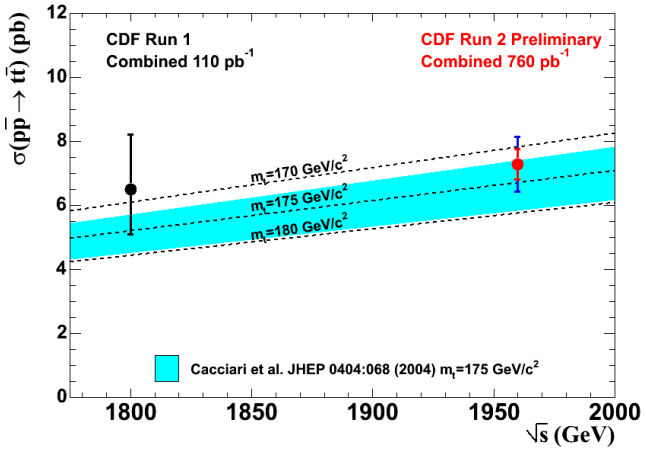


Figure 2.3: $t\bar{t}$ production cross section for the two Tevatron data-taking periods.

faded away with the most recent theoretical calculation[4] and the 2002-2006 measurement (Run II) performed combining the various decay modes as can be seen in Figure 2.3.

- **Single top production cross section**

  The D0 Collaboration recently presented the first evidence for the production of single top quarks at the Tevatron Collider, measuring $\sigma(p\bar{p} \to tb + X, tqb + X) = 4.9 \pm 1.4$ pb. The probability to measure a cross section at this value or higher in the absence of signal is 0.035%, corresponding to 3.4 standard deviations. The cross section measurement is used to directly determine the CKM matrix element that describes the Wtb coupling, setting a limit $0.68 < |V_{tb}| \leq 1$ at 95% C.L.

- **Study of $t\bar{t}$ production mechanism**

  A measurement of $\sigma(gg \to t\bar{t})/\sigma(p\bar{p} \to t\bar{t})$ using the low-P_T track multiplicity in the lepton+jet channel is performed to separate out gg initial states. It is

demonstrated that the average number of low-P_T tracks scales with the gluon content of the sample. Taking advantage of the fact that the gluon composition of the gluon-rich fraction of the Standard Model $t\bar{t}$ process is close to that of the gluon-rich fraction of dijet samples with relatively high leading jet E_T values, and that the $W+0$ jet sample is dominated by $q\bar{q}$ initial states, the gluon-rich fraction is extracted and the ratio $\sigma(gg \to t\bar{t})/\sigma(p\bar{p} \to t\bar{t})$ is measured, finding the value of $0.01 \pm 0.16\,(stat.) \pm 0.07\,(syst.)$.

- **Determination of the top charge**

  The first determination of the electric charge of the top quark is performed, using top quark pairs produced in $p\bar{p}$ collisions using data collected by the D0 experiment and selecting events with at least one high-P_T electron or muon, high transverse energy imbalance, and four or more jets. Discriminating between b- and $\bar{b}$-quark jets by using the charge and momenta of tracks within the jet cones, the data is found to be consistent with the expected electric charge, $+2/3\,e$. A charge of $+4/3\,e$ is also excluded at the 92% C.L.

- **Measurement of $t \to Wq$ branching ratio**

  A measurement of the ratio of top quark branching fractions $R = BR(t \to Wb)/BR(t \to Wq)$ using lepton+jets and dilepton data sets has been performed by the CDF collaboration. This measurement is derived from the relative numbers of $t\bar{t}$ events with different multiplicity of identified secondary vertices, setting a lower limit of $R > 0.61$ at 95% confidence level.

- **Measurement of W helicity in top decays**

  The top quark is predicted by the Standard Model to decay into a bottom quark and a W boson by the electroweak interaction. The spin-one W boson can have 3 different helicities or orientations of its spin relative to its direction of motion. Because of the heavy top quark mass relative to the W boson mass, the top decays most of the time to a longitudinal W boson (70%). However, because the weak interaction violates parity maximally, the remaining time it decays to left-handed W bosons (30%). The purpose of this analysis is to test whether the $V-A$ rule holds true for the top quark. By taking advantage of

a relationship between the helicity of the W boson, and the invariant mass of the charged lepton from the W decay and the b quark from the top decay, we are able to distinguish between left-handed and right-handed W bosons. The measured helicity fraction is consistent with the Standard Model.

All of the above measurements test our knowledge of the production and decay mechanism in the framework of the Standard Model theory of fundamental interactions. No discrepancy has yet been observed in the top quark sector.

Interestingly, the top quark can be used as a mean to explore physics at the highest achievable mass scale. Looking at its production and decay mechanism, one can find hints of new physics to appear, for example, in possible resonant production, or in branching ratios different from the expectectations, and so on. A number of measurements have been performed setting limits to new particle production:

- **Measurement of $t \to H^+b$ branching ratio**

  Extensions of the Standard Model predict the existence of charged Higgs bosons ($H^{\pm}$). In such models the branching ratio $BR(t \to H^+b)$ can be large thus competing with the Standard Model decay $t \to W^+b$. This search is based on the cross section measurements of $t\bar{t}$ production in three exclusive decay channels: the dilepton, lepton+jets and lepton+τ channels. Assuming the charged Higgs decays into $c\bar{s}$, $\tau\nu$, t^*b and Wh^0, limits to the $t \to H^+b$ branching ratio are obtained.

- **Search for resonant $t\bar{t}$ production**

  The CDF collaboration performed a search for non-Standard Model resonant production of $t\bar{t}$ pairs in the lepton+jets channel. The $t\bar{t}$ invariant mass ($M_{t\bar{t}}$) is reconstructed and the resulting distribution is then tested for possible resonant production (X^o) The CDF collaboration tested for vector resonances with masses in the range $450 - 900\,\mathrm{GeV/c^2}$ and natural widths equal to 1.2 % of their mass, and limits are set on the product of the cross section times the branching ratio to $t\bar{t}$ pairs, $\sigma_{X^0} \times BR(X^0 \to t\bar{t})$.

- **Search for a massive t'**

The CDF collaboration searched for the heavy top quark (t') pair production decaying to Wq final states in lepton+jets events. Masses below 258 $\mathrm{GeV/c^2}$ have been excluded for Standard Model fourth-generation t' quark at 95% C.L.

- **Search for a W' in single top production**

  The D0 collaboration searched for the production of a new heavy gauge boson W' that decays to a top quark and a bottom quark. No significant excess of events is found in any region of the final state invariant mass distribution, thus setting upper limits on the production cross section of W' bosons at the 95% confidence level for several different W' boson masses.

2.5 Motivation for a precise top quark mass measurement

Having confirmed the nature of the top quark, it is important to estimate its mass which is a free parameter in the Standard Model[3]. The top quark mass, M_{top}, is in fact one of the most important electroweak measurements to be made at the Tevatron Collider.

The Standard Model has undergone extensive tests along the decades from its orginal formulation; all but one of the particles predicted have been observed and many predictions which can be inferred from the theory have been confirmed by present and past experiments. However, further precision in both theoretical and experimental determination of observables, to be ran together with direct searches for new phenomena, is desirable. To many, the Standard Model has to be regarded as a theory which proved to be effective up to the GeV scale available at current machines, but which hides a more general theory; as long as we are unable to investigate the TeV scale, new physics could emerge in the form of radiative corrections to

[3]The free parameters in the Standard Model are the strong coupling constant α_s, the weak coupling constant G_F, the electromagnetic coupling constant α_{em}, the Weinberg angle θ_W, the mass of the Higgs boson, the masses of the six quarks and of the six leptons and the four quark and four neutrino mixing parameters which determine the Cabibbo-Kobayashi-Maskawa matrix.

Measurement	$2\ \mathrm{fb}^{-1}$	Comment
Yields		
N_{3jet*b}	990	*identified* events
$N_{4jet*2b}$	240	clean m_t sample
δm_t	3	total precision GeV/c^2
Production		
$\delta\sigma_{t\bar{t}}$	9%	test top QCD couplings
$\delta\sigma_{ll}/\sigma_{l+j}$	12%	test non W decay
$\delta\sigma_{t\bar{b}X+b\bar{t}X}$	24%	isolate "single top"
$\delta\sigma \cdot \mathrm{B}(Z' \to t\bar{t})$	90 fb	"topcolor" $M_{Z'} = 1\ \mathrm{TeV}/c^2$
Decay		
$\delta B(t \to W(b))$	2.8%	from N(bb)/N(bX)
$\delta B(t \to b(W))$	9%	from N(ll)/N(lX)
$\delta\mathrm{B}(W_{V+A})$	2.7%	$W \to l\,\nu$ helicity
$\delta\mathrm{B}(W_{long})$	5.5%	$\frac{W_{long}}{W_{left}} = \frac{1}{2}(\frac{m_{top}}{m_W})^2$
$\delta\Gamma(t \to Wb)$	26%	using single top
δV_{tb}	13%	from above
Rare Decays		
$\mathrm{B}(c\,\gamma)$	$\leq 2.8 \times 10^{-3}$	(95% CL)
$\mathrm{B}(cZ)$	$\leq 1.3 \times 10^{-2}$	(95% CL)
$\mathrm{B}(Hb)$	$\leq 12\%$	from σ_{ll}/σ_{l+j}

Figure 2.4: Yields and sensitivities expected for $2\,\mathrm{fb}^{-1}$ of data to be collected by CDF II. We would like the reader to note that the current measurement (with only $1\,\mathrm{fb}^{-1}$) on the top quark mass has already surpassed the expectation. Table taken from [**?**].

precision observables. The top quark plays a central role in the predictions of many Standard Model observables by contributing to their radiative corrections. Good examples are the W and Z boson propagators, in which loops involving top quarks are expected to strongly contribute, as illustrated in Fig. 2.5. These diagrams can exist for any type of quark or lepton, but the very large value of M_{top} makes the top quark contribution dominant. To illustrate the effect of the top quark presence, we consider the theoretical calculation of the W boson mass:

$$M_W^2 = \frac{\pi\alpha}{\sqrt{2}G_F sin^2\theta_W}\frac{1}{1-\Delta r} \tag{2.6}$$

where α_{em} is the fine structure constant, θ_W is the Weinberg angle and Δr contains the radiative corrections and is approximately given by

$$\Delta r \sim \Delta\rho_0 - \frac{\Delta\rho}{tan^2\theta_W} \tag{2.7}$$

The term $\Delta\rho_0$ is due to the running of α and the term $\Delta\rho$ is due to the one-loop top quark correction to W boson propagators illustrated in Fig. 2.5 and is given by

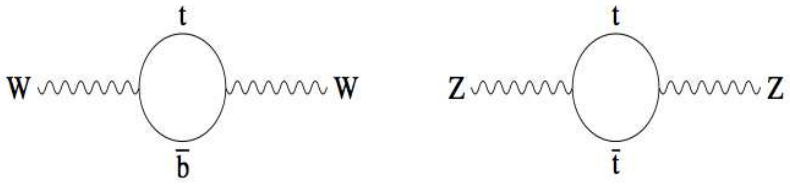


Figure 2.5: Contribution of the top quark to radiative corrections to the W and Z bosons propagators.

$$\Delta\rho = \frac{3G_F M_{top}^2}{8\sqrt{2}\pi^2} \tag{2.8}$$

The uncertainty on the Fermi constant, G_F, is completely negligible with respect to the one on the top quark mass. The terms $\Delta\rho_0$ and $sin\theta_W$ are known to a precision of 0.2%, while the top quark mass is known to a precision of one order of magnitude worse, thus being the limiting factor in the theoretical computation of the W mass.

In the same way, through the $\Delta\rho$ parameter, the top quark mass enters into many electroweak parameters.

The most precise measurement of the top quark mass have been performed in the lepton + jets channel thanks to the high statistics of the sample and the low background contamination. Nonetheless, it is important to measure the same parameter in the complementary channels as well, both as a consistency check, and in order to reduce the overall uncertainty on the top quark mass combining informations from measurements which have been performed in different data samples, with different techniques, and which are affected from partly or wholly different systematic uncertainties. It turns out that the all-hadronic channel provides the second most precise measurement of the top quark mass and helps significantly to reduce the overall uncertainty on this fundamental parameter.

One consistency check is to compare the measured value of M_{top} with the prediction from Standard Model observables. The indirect constraint from the theory yields $M_{top} = 181^{+12}_{-9}\,\mathrm{GeV/c^2}$ [5] which is in remarkable agreement with the current Tevatron averaged value of $M_{top} = 171.4 \pm 2.1\,\mathrm{GeV/c^2}$ [6]. The goodness of the Standard Model prediction on the top quark mass and its quantum numbers, add confidence to the analogous predictions obtained using the electroweak parameters on the Higgs mass, M_H. The only direct information on M_H is a lower bound obtained from searches at LEP: $M_H > 114\,\mathrm{GeV/c^2}$ at 95% C.L. [7]. Indirect constraint on M_H can be obtained with precise measurements of M_W and M_{top}. Indeed, the correction to the W mass given in previous equation contains additional terms due to the Higgs boson loops. These corrections depend only logarithmically on M_H and have thus a weaker dependence on M_H than on M_{top}. Still, precise determination of M_{top} and M_W can be used to obtain meaningful constraints on M_H as illustrated in Fig. 2.6. The χ^2 from a global electroweak fit [5] is shown in Figure 2.7, together with the LEP excluded region. Numerically, the constraints amount to

$$M_H = 85^{+39}_{-28}\,\mathrm{GeV/c^2} \tag{2.9}$$

$$M_H < 199\,\mathrm{GeV/c^2}\,\mathrm{at}\,95\%\,\mathrm{C.L.} \tag{2.10}$$

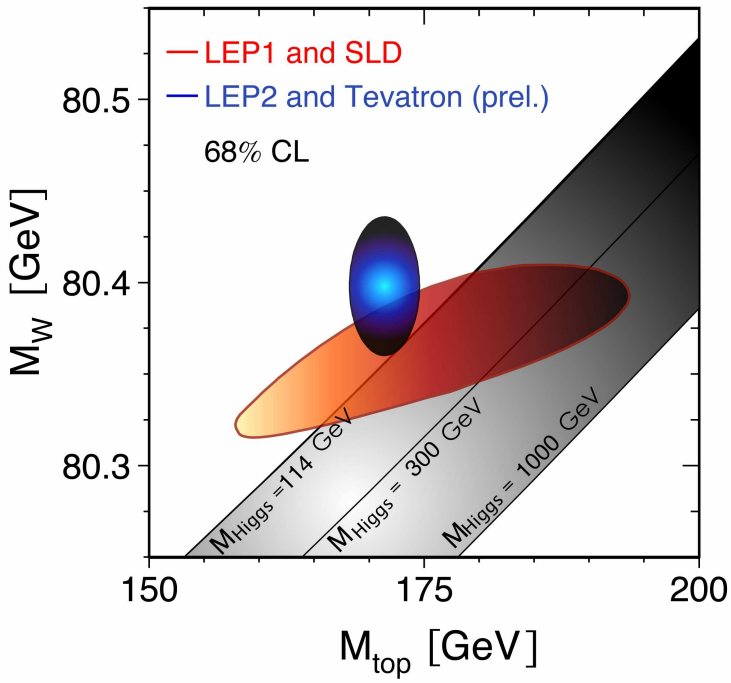


Figure 2.6: Constraint on the Higgs boson mass due to the most recent W mass and top quark mass measurements.

thus guiding current and future searches at the Tevatron and LHC colliders. Moreover, this result does not contain the latest measurement of the W mass as performed by the CDF collaboration [8], which hints again at a low mass Higgs, (pushing down the forementioned upper limit by $10\,\mathrm{GeV/c^2}$) which is potentially observable at the Tevatron once the full expected luminosity will be integrated (see Fig. 2.8).

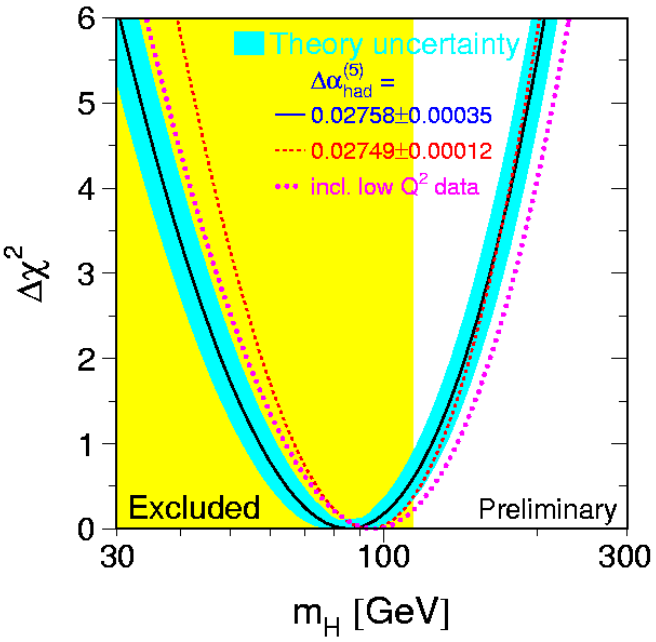


Figure 2.7: $\Delta\chi^2 = \chi^2 - \chi^2_{min}$ *vs.* M_H curve. The line is a result of the fit using electroweak data; the band represents an estimate of the theoretical error. The vertical band shows the 95% C.L. exclusion limit in M_H from the LEP direct search.

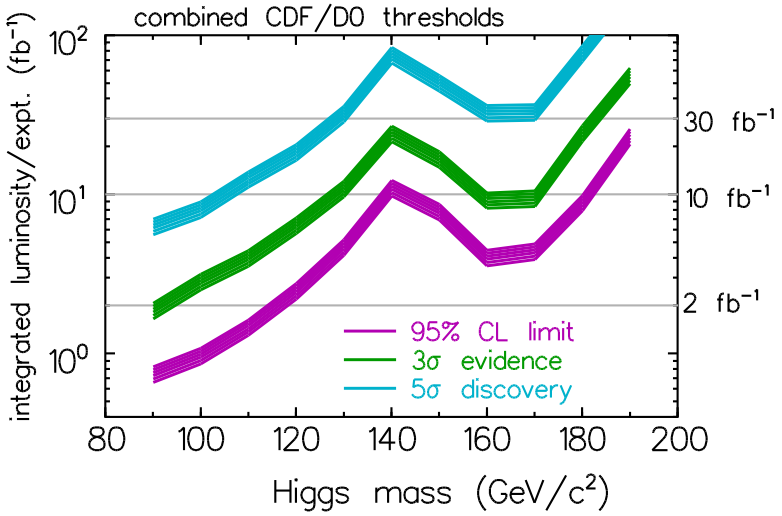


Figure 2.8: Potential for putting limits/finding evidence (3σ) or observation (5σ) of the Higgs particle as a function of its mass for different integrated luminosity at the Tevatron. The current prediction is to integrate a luminosity of 6-8 fb^{-1} before shutting down the complex.

Chapter 3

The experimental apparatus

The measurement of the top quark mass, which is the subject of this thesis, is performed identifying and reconstructing $t\bar{t}$ pairs decaying to a final state entirely made of quarks: $b\,\bar{b}\,q_1\,\bar{q}_2\,q_3\,\bar{q}_4$. These events are produced by the Tevatron, which is a $p\bar{p}$ collider operating at $\sqrt{s} = 2\ TeV$, located at the Fermi National Accelerator Laboratory in Batavia, Illinois. The data are collected by the CDF II multi-purpose detector and processed by a computing farm. In this chapter we describe the experimental apparatus needed to carry out the measurement.

3.1 The accelerator chain

Fermilab is a world class laboratory for high energy physics research; it is located in Batavia, in the suburbs of Chicago, Illinois. Inside the complex is located the accelerator which produces the data to be exploited in this analysis. The accelerator is actually an accelerating chain composed mainly by four subsystems: the proton source, the main injector, the antiproton source and the Tevatron. A pictorial representation of the accelerating complex can be seen in Fig. 3.1.

3.1.1 Proton source

The first device in the chain is the Cockroft-Walton electrostatic accelerator. Here a gas of hydrogen is ionized by the addition of an electron. The resulting ions are

accelerated up to an energy of 750 keV. The ions enter then into a 130 m long linear accelerator (Linac) in which a series of radiofrequency cavities accelerate the ions up to 400 MeV. At the outer end of the Linac the ions smash through carbon foils to strip the atoms of their electrons. The resulting protons are then inserted into the Booster, a synchroton accelerator with a circumference of about 0.5 km. There, the protons reach the energy of 8 GeV, and subsequentely enter the Main Injector.

3.1.2 Main injector

The main injector is a synchroton accelerator with a circumference of about 3 km. Its purpose is to:

- accept protons coming from the Booster or antiprotons from the Accumulator;
- accelerate protons up to 120 GeV and send them either to the Target Station, or to the fixed target aerea for other experiment's use, or in alternative to a neutrino facility;
- accelerate protons and antiprotons from 8 GeV to 150 GeV and send them to the Tevatron collider;
- accept antiprotons from the Tevatron, decelerate them to 8 GeV.

3.1.3 Antiproton source

Protons coming out of the main injector at 120 GeV are smashed on a target in the Target Station. The number of antiprotons collected is very low: about 20 $\bar{p}$ per millions of protons. The resulting particles are focused into a beam using magnet quadrupoles and a 8 GeV beam of antiprotons is obtained; thereafter the beam is sent to the Accumulator, a storage ring used to collect antiprotons until about $1.3 \times 10^{12}\,\bar{p}$ are produced. The accumulation rate is about $7 \times 10^{10}\,\bar{p}$ per hour so about 15 hours are needed to collect an adequate amount. The beam is finally transferred to the Main Injector where it is accelerated up to 150 GeV before being injected into the Tevatron.

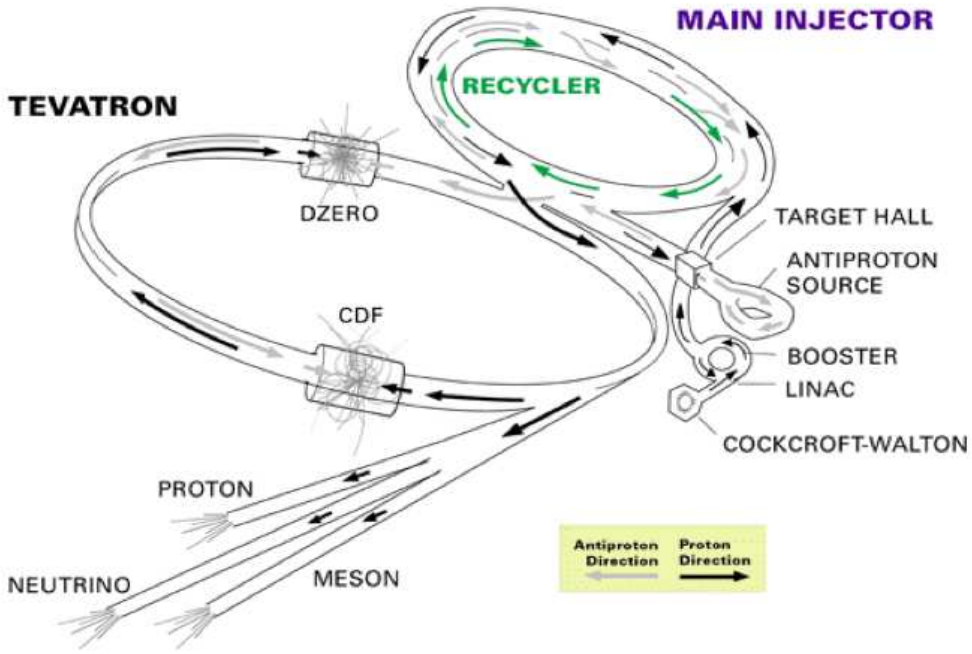


Figure 3.1: Schematic view of the Tevatron accelerator complex.

3.1.4 The Tevatron

The Tevatron is a synchroton with 1 km radius. Particles are bent thanks to superconducting magnets with $B \simeq 5.7\,\mathrm{T}$, and energies reach 980 GeV per beam. Beams are subdivided in 36 bunches each, which meet at 72 interaction points along the ring. Most of them are parasitical, while in the two regions where the CDF II and D0 detectors are located the beams are further focused to increase the luminosity. The latter is defined (for a machine in which beams collide head-on) as

$$\mathcal{L} = \frac{N_p\, N_{\bar{p}}\, f_R\, N_B}{2\pi(\sigma_p^2 + \sigma_{\bar{p}}^2)} \tag{3.1}$$

where $N_p(N_{\bar{p}})$ is the number of $p(\bar{p})$ circulating, f_R is the revolution frequency, N_B is the number of circulating bunches and $\sigma_{p(\bar{p})}$ is the width of the spatial distribution of $p(\bar{p})$ on the transverse plane at the interaction point. The number of multiple interactions is a Poisson variable whose mean is shown in Fig. 3.2 as a function of the number of circulating bunches. To date, the maximum luminosity achieved with the

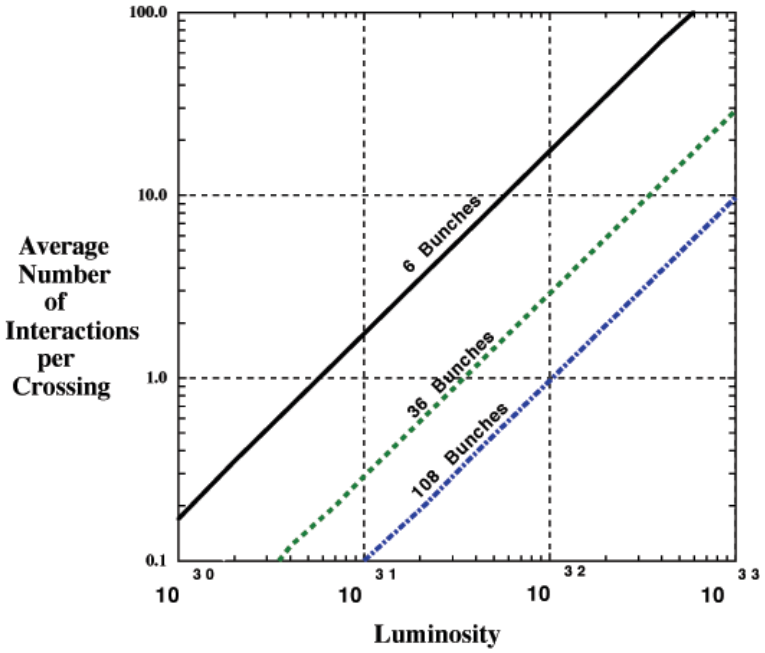
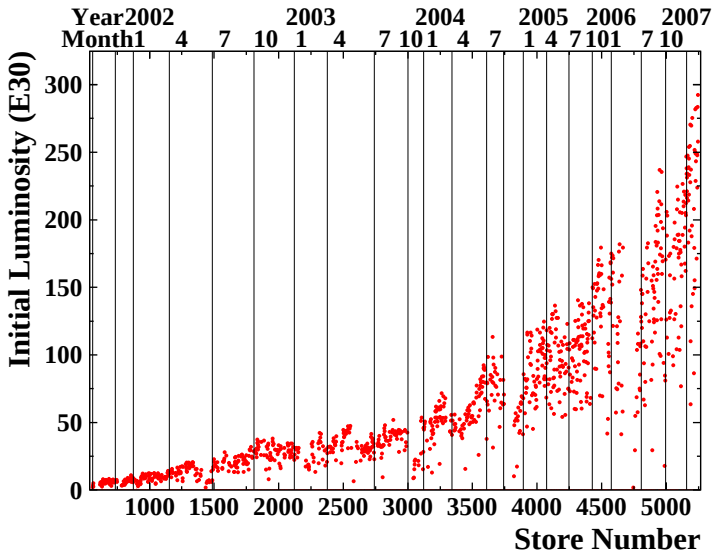


Figure 3.2: Average number of $p\bar{p}$ interactions per bunch crossing as a function of the instantaneous luminosity and of the number of circulating bunches. The current configuration is set to 32×32 bunches.

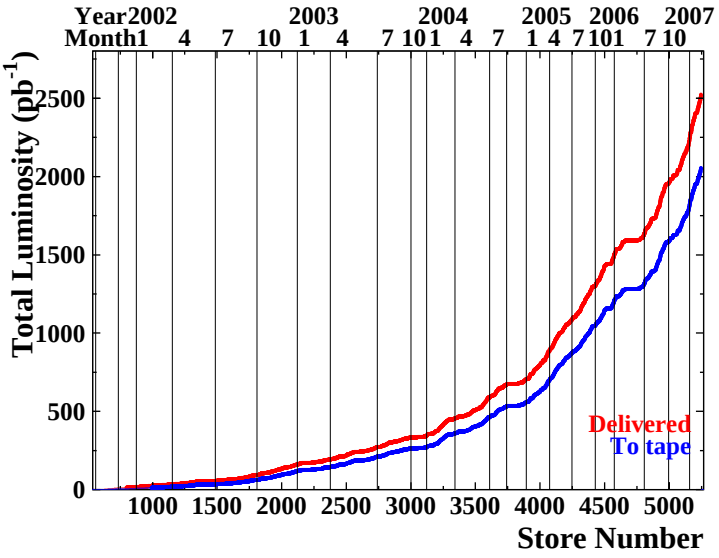
Fermilab accelerator complex reached $2.7 \cdot 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and the overall integrated luminosity during Run II amounts so far to $2.5{\cdot}10^{39}\,\mathrm{cm}^{-2} = 2.5\,\mathrm{fb}^{-1}$, as can be seen in Fig. 3.3. Between 6 and 8 fb^{-1} are expected to be integrated before shutting down the complex in 2009. The interaction region has a spatial distribution of about 30 cm along the beam direction ($\sigma_z \simeq 30\,\mathrm{cm}$) while on the transverse plane the beam is approximately circular with $\sigma_T^{beam} \simeq 25\,\mu\mathrm{m}$. The Tevatron most important parameters are summarized in Table 3.1.

3.2 The CDF II detector

The Collider Detector at Fermilab (CDF II) (Fig. 3.4) is a multi-purpose detector; in fact it is designed to study a wide range of physics processes produced at proton-antiproton interactions characterized by final states with high transverse momenta particles. Since the two beams collide head-on, the detector exhibits a forward-backward symmetry, and a cylindrical symmetry around the beam-pipe. Starting



(a)



(b)

Figure 3.3: a) Tevatron initial instantaneous luminosity and b) integrated luminosity as a function of store number (Run II data taking period, 2002-2007).

Parameters	***Value***
$p \times \bar{p}$ bunches	36×36
Number of p per bunch	$3.3 \cdot 10^{11}$
Number of $\bar{p}$ per bunch	$3.6 \cdot 10^{10}$
Total number of $\bar{p}$	$1.1 \cdot 10^{12}$
p emittance (mm mrad)	30
$\bar{p}$ emittance (mm mrad)	20
Energy $(p + \bar{p})$ (GeV)	980+980
Bunch spacing (ns)	396
$\mathcal{L}$ ($\mathrm{cm}^{-2}\mathrm{s}^{-1}$) (peak)	$2.7 \cdot 10^{32}$
Number of interactions/collisions (peak)	8

Table 3.1: Summary of the most important Tevatron parameters.

from the beam-pipe and proceeding radially outwards, we first encounter the silicon vertex detector, surrounded by a drift chamber, both contained in a superconducting magnetic field of 1.4 T. Outside that is located the time of flight detector, made out of many scintillating bars; the energy of photons, electrons and hadrons is measured by electromagnetic and hadronic sampling calorimeters. Finally muons are identified and their four-momenta measured in the proportional chambers located outside the calorimeters. In the forward region we also have some additional detectors, out of which a very important one is a Cherenkov luminosity counter. There are too many collisions to be recorded, but luckily most of them are of little interest. A trigger system made of three levels decide whether or not to record the outcome of the collisions. We will now describe the CDF II reference system and define some quantities which will be used in the following; a detailed description of the subdetectors listed above follows.

Reference systems

Given that the detector has a cylindrical symmetry around the beam axis, a convenient choice for the coordinate system is the cylindrical one, where the z axis coincides with the proton beam direction, and ρ and ϕ are measured on the orthogonal plane intersecting the nominal interaction vertex. We define also a cartesian reference system using x and y where the former points outside the accelerator plane

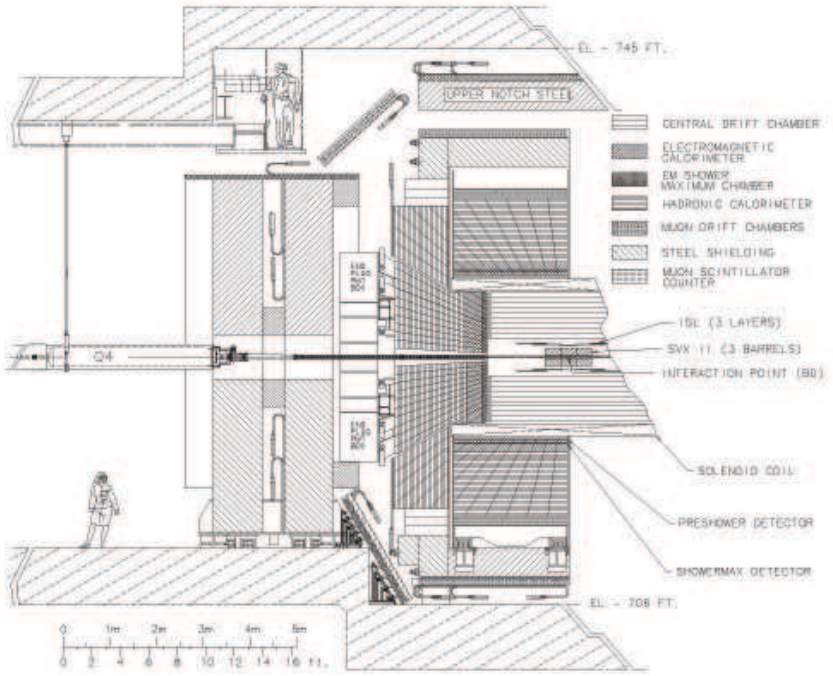


Figure 3.4: Side view of the CDF II detector.

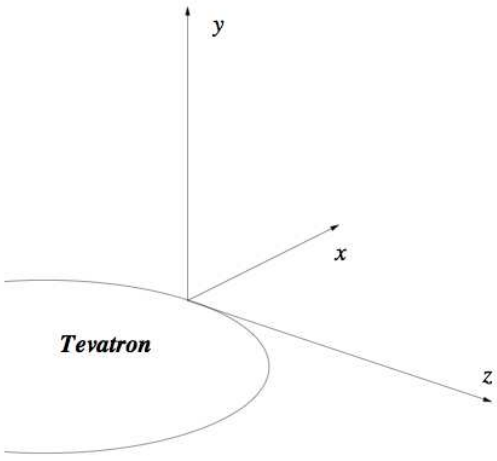


Figure 3.5: CDF II (x, y, z) reference system.

and the latter perpendicularly to it, as shown in Fig. 3.5. The ϕ angle is measured with respect to the positive direction of the x axis. It is also useful to define θ as the angle with respect to the positive z direction. Since we do not know the Lorentz boost of the proton-antiproton center of mass with respect to the laboratory, we introduce a quantity which is invariant under such transformations, to describe the forward direction, while the transverse component of the momentum automatically satisfy this condition.

Let's consider a particle with energy E and momentum p. The quantity called rapidity y is defined as:

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right).$$

The rapidity transforms as:

$$y \rightarrow y + \text{const} = y + \tanh^{-1} \beta_z$$

(being $v_z = \beta_z c$ the velocity in the reference frame of the partons). The difference in rapidity between two particles is thus unaltered by a Lorentz boost along the beam axis. The rapidity is particularly useful in the limit $p \gg m$ (which is generally correct at the Tevatron energies):

$$y(p \gg m) = \frac{1}{2} \ln \left(\frac{p + p_z}{p - p_z} \right) = - \ln \left(\tan \frac{\theta}{2} \right) = \eta. \tag{3.2}$$

The quantity η, called pseudorapidity, is a function of the θ angle only; from now on we will use η to describe the direction of particles along the (y, z) plane.
A related quantity, $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. is typically used to estabilish criteria of isolation, labeling as close those particles contained inside a circumference of arbitrary radius ΔR on a (η, ϕ) plane.

3.2.1 Tracking and vertexing systems

The tracking system occupies the inner volume of the solenoid (see Fig. 3.6). It is made up of the following detectors (proceeding from the beam pipe radially out-

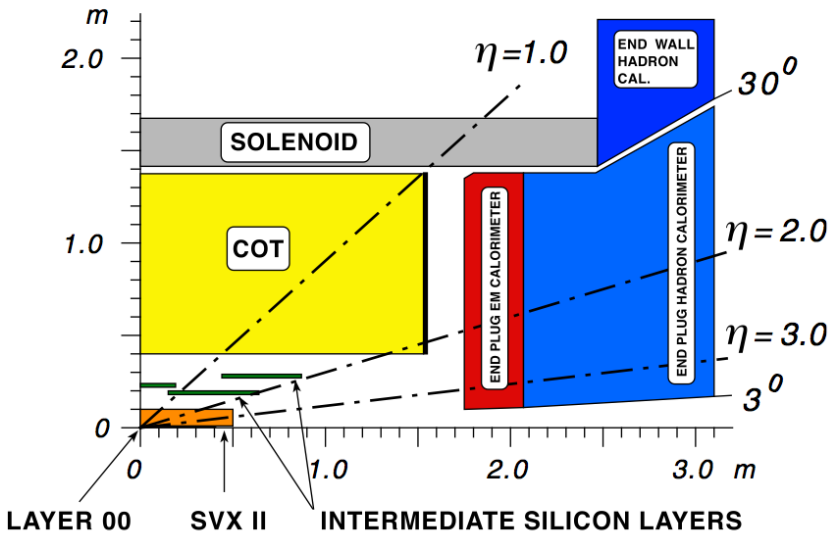


Figure 3.6: Cross section view of the CDF II tracking and calorimetry subsystems.

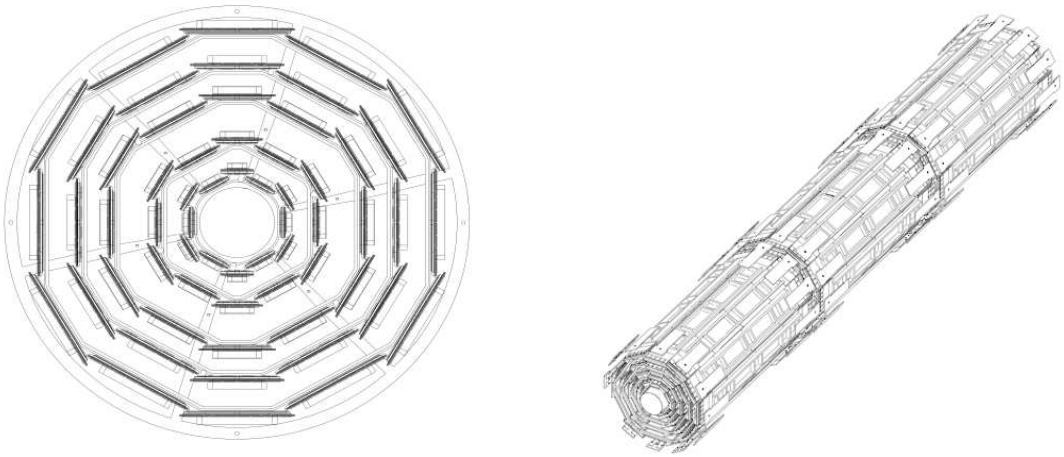
ward): the "Layer 00", the "SVX II" (*Silicon VerteX detector*), the "ISL" (*Intermediate Silicon Layer*) and the "COT" (*Central OuTer chamber*). All these detectors have cylindrical simmetry. In CDF II the silicon detectors are comprised between the beam pipe and the COT and constitute a stand-alone tracking system that has a pseudorapidity coverage which extends up to $|\eta| \leq 2$.

Layer 00

It is constituted from a single layer of silicon sensors with microstrips aligned along the axis of the beam, with a distance of 25 μm. The spatial resolution on the impact point of charged particles is approximately 6 μm. Layer 00 lies on the external surface of the beam pipe, therefore at an average distance $r \simeq 1.6$ cm from the nominal beam axis. Longitudinally it covers the region $|z| < 40$ cm.

SVX II

At a distance from the axis comprised between 2.4 cm and 10.7 cm, 5 silicon layers are arranged radially on 3 barrels, each subdivided in 12 wedges (Fig. 3.7). Each layer has microstrips on both sides: on one side they are aligned along the z-axis, on the other orthogonally (3 layers) or to a stereo angle $\pm 1.2^o$ (2 layers). Such a



(a) (b)

Figure 3.7: (a) SVX II cross sectional view. (b) The three SVX II barrels.

Parameter	***Value***
Distance from beam pipe (cm)	$2.4 \leftrightarrow 10.7$
Pseudorapidity range	$\|\eta\| < 2.0$
Barrels/layers/wedges	3 / 5 / 12
Length/active length (cm)	$96 \,/\, 29 \times 3$
$r-\phi$ pitch (μm)	60, 62, 60, 60, 65
$r-z$ pitch (μm)	141, 125.5, 60, 141, 65
Stereo angle (degrees)	90, 90, 1.2, 90, 1.2
Channels	$211968\,(r-\phi) + 193536\,(stereo)$

Table 3.2: Characteristic SVX II parameters.

geometry of the strips allows an optimal reconstruction of the tracks in the transverse plane and in the $r-z$ plane. The detector extends longitudinally for approximately 96 cm covering a fraction of the luminous region corresponding to $2.5\,\sigma$.

The microstrips are separated by approximately $60\,\mu$m on the $r-\phi$ plane. The resolution on the position of the single hit is approximately $16\,\mu$m for axial strip and for the small stereo angle layer ($38\,\mu$m for the stereo orthogonal layer). The main informations are summarized in Table 3.2. Approximately 406000 strips (or channels) of SVX II are read using the fast chips SVX3D in less than $10\,\mu$s. The information on the $r-\phi$ coordinates is made available first to the trigger system. The

electronic devices and the sensors are designed in order to resist to large radiation doses ($\sim 0.5\,\mathrm{MRad/fb^{-1}}$). The Layer 00 guarantees five layers of sensors in the case of damaging of the first silicon layer. The thickness of the detector in term of the radiation length (X_0) is $0.3\,\mathrm{X_0}$ ($\theta = 90^o$) that becomes $0.6\,\mathrm{X_0}$ for the regions where is installed the readout electronics.

ISL

The Intermediate Silicon Layer is located between the vertex detector and the drift chamber. It consists of three layers of silicon with axial microstrips on one side, separated by $110\,\mu$m, and stereo, with an angle of $\pm 1.2^o$, on the other ($146\,\mu$m). The resolution on the position of the single hit is about $16\,\mu$m for axial strips and $23\,\mu$m for the stereo ones. Of the three layers, the central has an average distance of 22 cm from the beam axis and covers the pseudorapidity range $|\eta| < 1$ while the remaining two are respectively at 20 cm and 29 cm, both covering $1 < |\eta| < 2$. The total length is approximately 174 cm. The layers are partially overlapped on the $r - z$ plane (Fig. 3.6) and on the $r - \phi$ plane. The thickness of a detector layer is, in average, $0.5\,\mathrm{X_0}$. The tracking system composed from SVX II and ISL covers the entire luminous region and has 6 layers available overall in the central region ($|\eta| < 1$) and 7 in the plug ($1 < |\eta| < 2$). SVX II, because of its insufficient radial extension, does not allow a good resolution on the transverse momentum P_T (and therefore on the impact parameter (d_0) and on the variables relative to the $r-z$ plane ($\cot\theta$ and z_0). The information supplied from ISL is useful for three-dimensional reconstruction in the central region (where the COT measurement is also available), and in the plug region where the presence of an additional silicon layer compensates the reduced coverage of the drift chamber.

COT

The Central Outer Tracker is an open-cell drift chamber located at radii betrween 40 cm and 132 cm. The wires are subdivided in 8 superlayers, divided in 4 axial superlayers for the measurement on the transverse plane, and 4 stereo superlayers (stereo angle $\pm 2^o$), for the measurement of the z coordinate. Each superlayer contains 12 wires for the collection of signal for a total of 96 measurement points of

Parameter	***Value***
Radius (cm)	$40 \leftrightarrow 132$
Length (cm)	310
Gas	Ar-Et-CF_4 (50:35:15)
Max drift distance (cm)	0.88
Max drift time (ns)	100
Drift field (kV·cm^{-1})	2.6
Lorentz angle	35^o
Superlayer × axial wires	4×12
Stereo angle	$\pm 3^o$
Superlayers × stereo wires	4×12
Total wire number	63000
Thickness (X_0)	1.7%

Table 3.3: Main COT parameters.

each charge particle trajectory. The maximum drift time is approximately 100 ns. This allows a correct operation of the chamber with a bunch spacing of 396 ns and to use the information of the COT at the first level of the trigger. The expected resolution on the position measurement of a single hit is of approximately 180 μm[1]. The material of the COT is equivalent to approximately 1.7% of a radiation length ($\theta = 90^o$). The main informations are given in Table 3.3.

The COT resolution on the charged particles transverse momentum is $\sigma_{P_T}/P_T^2 \sim 1.7 \cdot 10^{-3}$ (GeV/c)$^{-1}$. If we consider altogether the three tracking systems (SVX II + ISL + COT), then the resolution becomes $\sigma_{P_T}/P_T^2 \sim 1 \cdot 10^{-3}$ (GeV/c)$^{-1}$.

3.2.2 Calorimetry

Calorimetry is the main component used for the measurement of the energy of hadrons, jets, electrons and photons. The system employed in CDF II consists of a sampling calorimeter shaped in a projective tower geometry which provide a full azimuthal coverage (2π) and pseudorapidity coverage up to $\eta = 3.6$. The central tower segmentation in η is $\Delta\eta = 0.1$ and in ϕ is $\Delta\phi = 7.5^o - 15^o$. Each tower is actually made of two separate devices, the electromagnetic calorimeter and the

[1]The efficiency is estimated to be approximately 95% in a wide range of luminosity.

hadronic calorimeter. Both are sampling devices, the first being made of alternating layers of iron and scintillator and the second of lead and scintillator. Overall, the CDF II calorimeter is composed of three main parts, which correspond to two different regions in η:

- the central and wall calorimeter ($|\eta| < 1$);
- the plug calorimeter ($1 < |\eta| < 3.6$).

We deal with the first two together in the next section (since the wall calorimeter actually is a subset of the central hadronic calorimeter). The description of the plug calorimeter will follow.

Central calorimeter

The central calorimeter is a collection of different devices. Proceeding radially outwards we meet first a preshower detector, then the electromagnetic calorimeter which is equipped with an electromagnetic shower maximum detector, and finally the hadronic calorimeter. Let's now describe the subdetectors in some detail.

Right outside of the solenoid we find the proportional chamber CPR (*Central PRereadiator detector*), which supplies the information on the position of electromagnetic showers that are produced in the solenoid.

Then the electromagnetic calorimeter (or CEM – Central ElectroMagnetic) [9] follows. It is divided in two halves symmetrical with respect to the plane $z = 0$. The segmentation $\Delta\eta \times \Delta\phi = 0.10 \times 15^o$ corresponds to having 24 wedges along the azimuthal direction, where each of the two halves is subdivided in 10 wedges. The CEM consists of 31 layers of plastic scintillator 5 mm thick alternated with 30 layers of lead 3 mm thick each, for a total of 18 X_0. Light guides collect from two sides of the tower the photons coming from the scintillators and carry them to two photomultipliers located in the external region of the calorimeter. The space for the light guides constitutes a dead zone in the regions of separation in ϕ between the towers, called ϕ-cracks, that amounts to the 4.5% of the entire coverage in ϕ. Not-instrumented zones are also present along the η direction (η-crack) because of

the aluminum sheets of $\sim 0.4\,\mathrm{mm}$ of thickness that coat absorber and scintillators. Another η crack is constituted from the region $\eta = 0$ corresponding to the mechanical separation between the two halves. The electromagnetic calorimeter is calibrated with electrons from test beam and monitored with LED, xenon, and with radiation from ^{60}Co and ^{137}Cs sources. The energy resolution amounts to $\frac{13.5\,\%}{\sqrt{E_T}} \oplus (1.5\,\%)$.

Another proportional chamber, the CES (*Central the Electromagnetic Strip detector*), is located within the calorimeter at a distance of approximately $6\,X_0$) from the lower face of the CEM, that corresponds to the distance in which on average the development of the electromagnetic shower is maximum. Its purpose is to facilitate the identification of $e^{\pm}$ and γ through the observation of the electromagnetic shower that characterizes their interactions in the scintillator, and to allow the separation of π^0 from γ through the different shape of the shower[2].

The central hadronic calorimeter is divided in two separated mechanical parts: Central HAdronic and Wall HAdronic, respectively CHA and WHA [10]. The first one contains 8 towers for side, each made of 32 layers of alternating scintillator and absorber, where the scintillator layers are 1.0 cm thick and the absorber is constituted from sheets of 2.5 cm of iron; the second is made of 6 towers for each side with 15 layers, of which half of scintillators (1.0 cm thick) and the rest of iron absorber of 5.0 cm. Both cover in total 5 interaction lengths (λ_i). Every tower is read by two photomultipliers. The calibration is done with π's, while the monitoring with mixed techniques that include light laser, γ emissions from ^{137}Cs, sources of β radiation. The energy resolution is $75.0\%/\sqrt{E_T} \oplus 3\%$ for the CHA and $80.0\%/\sqrt{E_t} \oplus 4\%$ for the WHA.

Plug Calorimeters

The plug calorimeters covers the pseudorapidity range $1.1 < |\eta| < 3.6$. The scintillator is composed of a mixture of 50% argon e 50% ethane. Segmentation in $\eta \times \phi$

[2]The CES is made of anodic wires that run along the z axis of the CDF II detector and measure the x coordinate, and orthogonal cathodic strips that measure the z coordinate. The CES fiducial coverage corresponds to $|x| < 22\,\mathrm{cm}$ and $14\,\mathrm{cm} < z < 217\,\mathrm{cm}$. A clustering algorithm identifies the strip or wire clusters, that will be associates to photons or π^0, or to electrons in the case there are the tracks that point to the cluster.

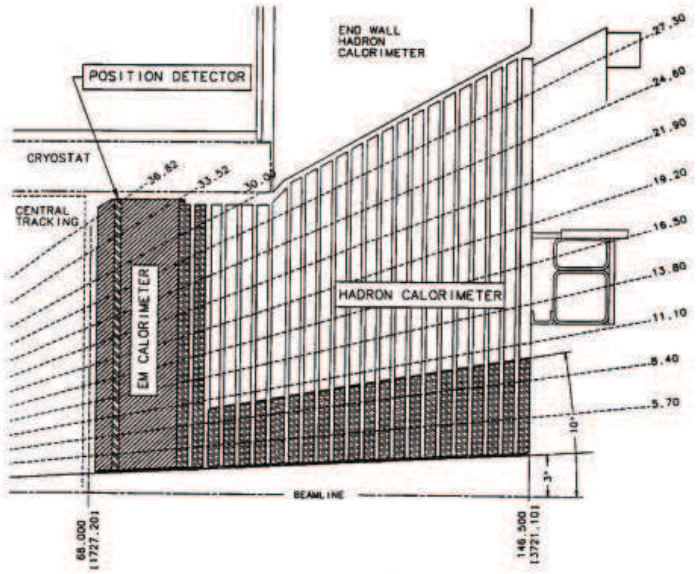


Figure 3.8: Quadrant view of the plug calorimeter CDF II.

varies from 0.11×15^o to 0.11×5^o depending on the region. The system is subdivided, analogously to the other calorimeters, into an electromagnetic device and an hadronic device (Fig. 3.8). The electromagnetic calorimeter [11] is composed of 23 alternated layers of 4.5 mm of lead absorber and 4 mm of scintillator for a total of $\sim 21\,X_0$. The first layer is a scintillator 10 mm thick. The hadronic calorimeter is also a sampling device with 23 layers where sheets of 5 mm of iron are interleaved with sheets of 6 mm of scintillator. The segmentation $\Delta\eta \times \Delta\phi$ varies from 0.1×7.5^o to 0.6×15^o. The resolution for the measurement of energy is approximately $\frac{14\,\%}{\sqrt{E}} \oplus 1\,\%$ and $\frac{80\,\%}{\sqrt{E}} \oplus 5\%$ for the electromagnetic and hadronic calorimeters respectively. As was the case for the central calorimeter, the plug calorimeter is equipped with a shower maximum detector [12]. It is made of scintillator strips at approximately 180 cm from the origin of the reference system, corresponding to approximately $6\,X_0$ from the base of the calorimeter, as to say at the depth where in average the extension of the electromagnetic shower is maximum. The longitudinal coverage of hadronic showers amounts to 8 interaction lengths λ_i. A summary of the main informations on the CDF II calorimetry is available in Table 3.4.

Electromagnetic calorimeter					
Subsystem	*η Region*	*Type*	$\sigma_{E_t}/E_t\,(\%)$	*Thickness*	$\Delta\eta \times \Delta\phi$
CEM	$\lvert\eta\rvert < 1.0$	*Pb-scint.*	$13.5/\sqrt{E_t} \oplus 1.5$	$18\,X_0$	0.1×15^o
PEM	$1.1 < \lvert\eta\rvert < 3.6$	*Pb-scint.*	$14.0/\sqrt{E_t} \oplus 1$	$18 - 21\,X_0$	0.1×5^o
Hadronic calorimeter					
Subsystem	*Region*	*Type*	$\sigma_{E_t}/E_t\,(\%)$	*Thickness*	$\Delta\eta \times \Delta\phi$
CHA	$\lvert\eta\rvert < 0.9$	*Fe-scint.*	$75.0/\sqrt{E_t} \oplus 3$	$5.5\,\lambda_i$	0.1×15^o
WHA	$0.6 < \lvert\eta\rvert < 1.3$	*Fe-scint.*	$80.0/\sqrt{E_t} \oplus 4$	$5.5\,\lambda_i$	0.1×15^o
PHA	$1.1 < \lvert\eta\rvert < 3.6$	*Fe-scint.*	$80.0/\sqrt{E_t} \oplus 5$	$8.0\,\lambda_i$	0.1×15^o

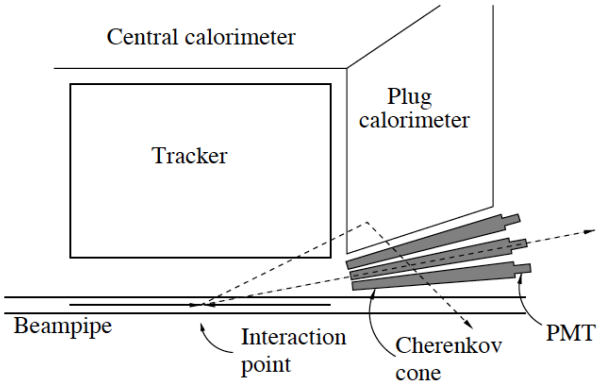
Table 3.4: Main characteristics of the CDF II calorimeter. The resolutions for the electromagnetic calorimeters (hadronic) are relative to isolated photons and electrons (pions). $E_t = E \cdot \sin\theta$ (GeV). $A \oplus B \equiv \sqrt{A^2 + B^2}$. The thickness, for particles incidents normally, are indicate in radiation lengths (X_0) and interaction lengths (λ_i) ($21\,X_0 \simeq 1\,\lambda_i$). These represent the average distance traveled such that, respectively, an electron loses $1/e$ of its initial energy for emission of radiation (*bremsstrahlung*) and a pion gives rise to an inelastic interaction

3.2.3 Cherenkov Luminosity Counter

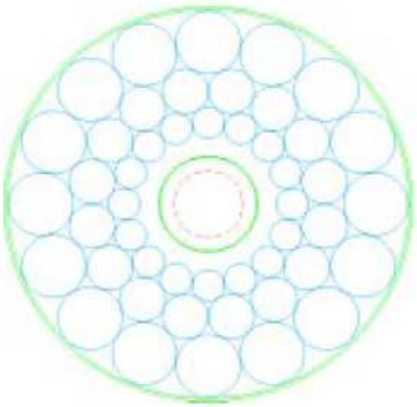
The Cherenkov Luminosity Counter (CLC) measures the average number of interactions per bunch crossing, μ. The instantaneous luminosity $\mathcal{L}$ is extracted using the equation

$$\mu \cdot f_{bc} = \sigma_{p\bar{p}} \cdot \mathcal{L} \tag{3.3}$$

where $\sigma_{p\bar{p}}$ is the relatively well known total $p\bar{p}$ cross section at $\sqrt{s} = 1.96$ TeV, and f_{bc} is the rate of bunch crossings at the Tevatron. The CLC uses the effect known as the Cherenkov radiation, where particles traversing a medium at a speed higher than the speed of light in that medium radiate light into a cone around the particle direction; the cone's opening angle depends on the ratio of the two speeds and on the refraction index of the medium. The idea is to use an assembly of long gas Cherenkov counters positioned in the plug calorimeter 3^o gap so that they point toward the interaction point as schematically shown in Fig. 3.9. This arrangement allows to make the detector much more sensitive to the particles coming directly from the interaction point because their flight path in the gas of the counter is the longest and therefore the amount of light produced the largest. Excellent time resolution (~ 50 ps) and clever design allow the CLC to discern multiple interactions within the



(a)



(b)

Figure 3.9: a) Quadrant view and b) cross sectional view of the Cherenkov Luminosity Counters at CDF II.

same bunch crossing and achieve an overall accuracy of the luminosity measurement better than 6%. In depth information on the CLC design and performance is given in reference [13].

3.2.4 Other parts

This are subdetector systems which are not used in the analysis but which we describe for completeness.

The Time Of Flight detector (*TOF)*

Distinguishing π's from K's is particularly important, in the studies of b physics, in order to increase the efficiency of some algorithms of tagging[3] and in the analysis of the totally hadronic B decays. For this reason a device is supplied to identify pions, kaons and protons through the measurement of their time of flight [14][4]. CDF takes advantage of the existing interstice between the external cylindrical surface of the COT that is about 140 cm from the beam axis, and the cryostat for the solenoid cooling. In this empty space (4.7 cm in the radial direction) are located 216 scintillator bars[5] 279 cm long (so as to to cover the longitudinal extension of the drift chamber) and square section (4 cm of side). To both the extremities PMT's are applied for the collection of signal. This apparatus measures the time that elapses between the collision and the signal emitted from the interactions of particles

[3] In b physics the word "tagging" refers to a tool that allows to determine if a meson of type B, at the moment of its production, contains a quark b or a quark $\bar{b}$ flavour-tagging).

[4] A particle, of known impulse p, in the two hypotheses that has mass m_1 or m_2, covers the distance L respectively, in times $t_1 = L/\beta_1$ and $t_2 = L/\beta_2$ (being $\beta_{1,2} = pc/\sqrt{p^2 + m_{1,2}^2}$). With the condition $m_{1,2}^2 c^2 \ll p^2$ it turns out

$$\Delta t = t_1 - t_2 = \frac{L}{c}\left(\frac{1}{\beta_1} - \frac{1}{\beta_2}\right) = \frac{L}{c}\left(\sqrt{1 + \frac{m_1^2 c^2}{p^2}} - \sqrt{1 + \frac{m_2^2 c^2}{p^2}}\right) \simeq \frac{Lc}{2p^2}(m_1^2 - m_2^2). \qquad (3.4)$$

The time needed to cover a certain distance can be used in order to establish which of the two probable hypotheses is more likely. However, for this purpose the time has to be measured with a resolution higher than Δt.

[5] The number of bars that have been inserted is the result of the average number of particles that, in every collisons, hit the same scintillators. The presence of more overlapped hits, in fact, remarkablly compromises the performances of the detector. Currently we expect in average 0.2 (0.4) charged particles cross one bar in the case of 2 (10) interactions $p\bar{p}$ for *bunch crossing*.

generated in the collision in the scintillator. The detector should determine the moment in which a particle crosses the scintillators with a resolution of $\sigma = 100\,\text{ps}$ to establish if it generated by a π or of K with a statistic separation higher than 2 standard deviation (for impulses smaller than 1.6 GeV). The TOF thus provides informations complementary to the dE/dx for particle identification.

Muon detection

Muons are characterized by their high penetrating power through matter. They interact only electromagnetically and weakly, and loose a small amount of energy through brehmsstrahlung as opposed to electrons, thanks to their higher mass. As a result they usually pass undetected in the calorimeter and reach the dedicated drift chambers located at the outermost part of the detector. In fact, outside the hadronic calorimeter CHA, almost 3.5 m away from the beam line ($\sim 5.4\,\lambda_i$), is located the CMU (*Central MUon*) made of 4 layers of drift chambers (4 hits). A similar device, CMP (*Central Muon uPgrade*), is located behind an iron layer 60 cm thick ($\sim 3\,\lambda_i$). A muon has to have at least 1.4 GeV/c of transverse momentum to reach the CMU and higher than 2.8 GeV/c to reach the CMP.

The CMU can contain also an hadronic signal with a 1% probability, while for the CMP this is negligible. CMU and CMP cover the region $|\eta| < 0.6$. Coverage of the region $0.6 < |\eta| < 1$ is assured from the CMX chambers (*Central Muon eXtension*). These are made of 4 layers, and are located between two scintillating layers to be used for triggering (CSX). The single hit resolution is about $250\,\mu$m on the $r-\phi$ plane and 1.2 mm along the z axis. Since outside the solenoid the magnetic field is absent, we can reconstruct only stubs of tracks with the muon chamber. This information has to be integrated with the COT tracks to define a muon. Another constraint is the presence of energy in the hadronic calorimeter compatible with a release from a μ (~ 0.5 GeV).

3.2.5 Trigger and data acquisition systems

In collisions between two bunches, a proton and an antiproton interact through an inelastic collisions. The outcome is the production of tens of particles that fly away

from the interaction zone. A fraction of these go through the detector in one or more active regions (the silicon of the vertex detector, the gas in the drift chamber, some towers of the calorimeter etc.) where they will interact with the detector material and produce some kind of release. The data acquisition system (DAQ - Data AcQuisition) is a collection of devices that have the task to manage the informations coming from the detector from the phase of reading until the phase of writing on a support of permanent memory. The acquisition begins by codifing the analogic signal supplied from the single detectors in digital data. Then they are collected, elaborated in real time and, finally, recorded on a magnetic tape where they remain available for the offline analyis. The acquisition of the data relative to the collision between two bunches is labeled "event acquisition".

In the current accelerator configuration bunches collide every 396 ns, that is equivalent to more than $7.5 \cdot 10^6$ hits per second. To acquire the event for every bunch crossing would demand electronic devices of prohibitive speed and storage volume. However, to collect all the events indiscriminately is not necessary: the interesting processes (therefore less known) constitute only a small fraction of the total, in the proportion between their production cross sections and the total $p\bar{p}$ cross section. It is indispensable, therefore, a selection of the events that works while the system of data acquisition is on, so as to to record on tape only those events that appear to be interesting at the moment of their acquisition, analyzing the information that comes out in real time from the various detectors ("online analysis"). The task of selecting the events is performed by the trigger system. The DAQ works, therefore, in synchronism with the trigger and is based on the instructions received from the latter. The basic requirement is is that an inelastic interaction between a proton and an antiproton has undergone. This kind of process is nearly always accompanied from a large particle multiplicity, produced in the partonic interactions at low transferred moment, emitted with a small angle θ.

The trigger system is structured in three levels. Every level selects a fraction of the events that have survived the previous one, allowing, before the following, a more detailed analysis. The electronic devices who make part of the trigger give a

measurement of the quantity of the event that is wanted to be recorded, i.e. the signal. When the measured quantity exceeds the threshold value, it triggers the execution of the following level, otherwise the information is lost and the DAQ goes on to examine the following event.

The first level takes into account all the events after every bunch crossing. The events that survive the third level are written on tape. Generally a trigger does not succeed to decide if to keep or not an event before the following collision. The dead-time is defined as the fraction of time in which the trigger is busy examining an event and can not accept others. The utility of having a trigger organized on more levels is that of being able to apply sufficiently effective criteria of selection without introducing a significant dead-time.

The effectiveness of a certain trigger is quantified through the rejection factor and its efficiency. The rejection factor is the ratio between the total number of events analyzed and the number that survive the selection criteria. The fraction of events passing the trigger requirement is called trigger rate. The rejection factor must be set to limit the trigger rate to a value that does not to introduce a dead time higher than a fixed threshold (approximately 10%). The efficiency, instead, refers to a specific signal and is defined as the fraction of the events of signal that survives the action of the trigger. The choice of the selection criteria has the scope of increasing the efficiency maintaining the rejection factor within the limits.

For this purpose two new processors have been built: XFT (*eXtremely Fast Tracker*) that allows to reconstruct charged tracks in the transverse plane, in a time short enough for the first level, and SVT (*Silicon Vertex Tracker*) to be able to measure the impact parameter of the reconstructed tracks and sending the information to the second level of the trigger system. The fundamental importance of these devices resides in the possibility to introduce selection criteria based, at Level 1, exclusively on the charged tracks and, at Level 2, on the presence of secondary vertices requiring tracks with high impact parameter. Such criterion is suited in order to select the events in which a b quark is produced, being these characteristic of the presence of a hadron with long lifettime. The time that the first level employs in order to take one decision is approximately $5.5\,\mu$s. In order to avoid to introduce dead-time every detector must have the possibility to store up to 42 data of suc-

cessive collisions. For this reason all the systems of reading of CDF II are equipped with sliding registers with 42 cells (pipeline structure) where to keep the information while waiting for the trigger decision. The diagram of Fig. 3.10 describes the structure of the system of DAQ and trigger of CDF II. We now describe in some detail the three trigger level (Fig. 3.11).

Level 1 (*L1*) Utilizes the information of the COT, the calorimeters and the muon chambers. The XFT processor reconstructs the charged tracks using the hits from the axial layers of the COT with reduced resolution ("XFT tracks"). The measured parameters are p_t and ϕ. The average time needed is $2.7\,\mu$s. The trajectories are then extrapolated until intersecting the lower face of the calorimeters and of the muon chambers (XTRP of Fig. 3.11). The possibility to formulate the selection criteria is based on the presence of tracks in the COT that satisfy specific kinematic requirements. Electrons, photons and jet candidates are identified imposing the presence of energy in the single towers of the calorimeters above the threshold values. The value of the sum of the energy released on all the towers is used for the selections based on the total transverse energy and the missing transverse energy. The selection of events containing muons is based on the location of stubs in the muon chamber in coincidence with signal from the scintillators. The availability of the parameters of the charged tracks allows to improve the identification of electromagnetic particles, hadronic and muons by checking the contiguity between the XFT tracks and the calorimetric towers and the stubs in the muon chamber. On charged tracks it is possible, moreover, to impose kinematic cuts (as an example the presence of one pair of tracks with $p_t > 3\,\mathrm{GeV/c}$).

At this point the trigger system is able to examine simultaneously 64 sets of requests (each specific to various physical process) in order to decide to whether to accept an event or not. The maximum trigger rate expected for Level 1 is approximately 50 kHz (rejection factor: $\sim$ 150).

Level 2 (*L2*) Accepts events from L1 and records them on a buffer waiting to be examined. There are 4 available buffers. The XFT tracks with 4 hits in

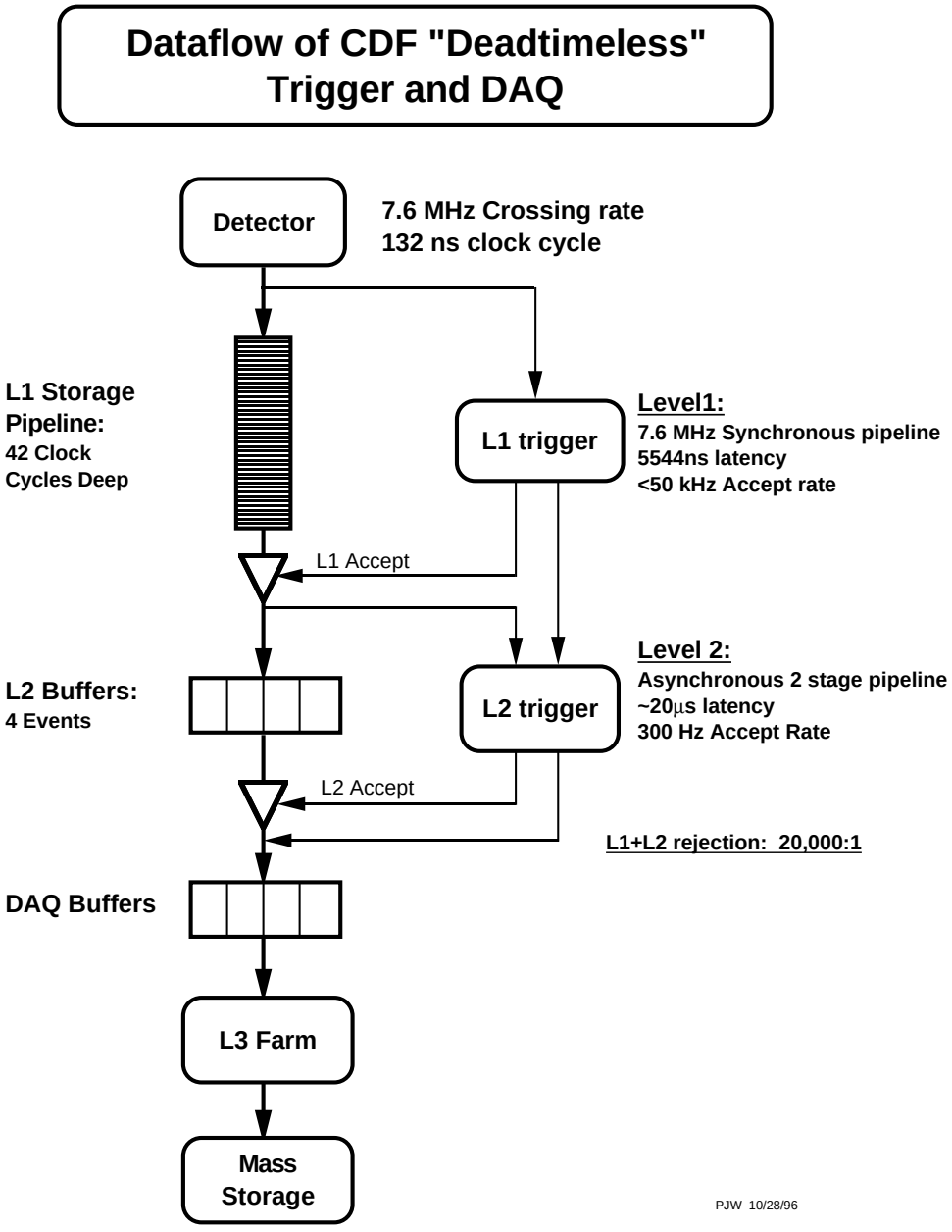


Figure 3.10: Data acquisition and trigger system at CDF II.

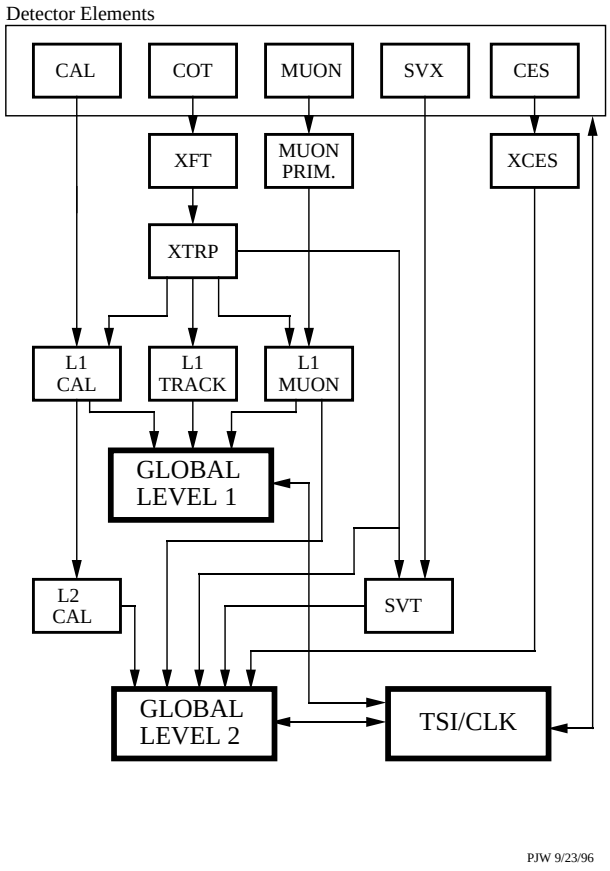


Figure 3.11: Trigger system at CDF II.

SVX II and $p_t > 2\,\mathrm{GeV/c}$ are reconstructed from SVT taking advantage of the additional information of the silicon detector: the measured parameters are p_t, ϕ and d with resolution comparable to that achievable with offline algorithms. Particles produced from secondary vertices are selectable demanding tracks with high impact parameter. Algorithms are applied to trigger electrons photons and jets: the energy for clusters of adjacent towers, and the information from the detectors of maximum expansion of the shower in the electromagnetic calorimeters (module XCES of the Figure 3.11) is now available. The time of execution is approximately $20\,\mu$s. The frequency of L2 trigger is limited to 300 Hz (rejection factor of ~ 150).

Level 3 (*L3*) At this level complete informations supplied from the various detectors are available. The events selected from Level 2 are therefore elaborated with algorithms analogue to those of the offline analysis. In particular, the tracking is completed executing the three-dimensional reconstruction of the trajectories in the volume $|\eta| < 2$, and more detailed algorithms reconstruct the energy in the calorimeters.

Add something

Chapter 4

Physical objects at CDF

This analysis searches for a $t\bar{t}$ pair decaying to $bW^+\bar{b}W^-$, where the W's subsequently decay into quarks. In this chapter all the tools needed to select a $t\bar{t}$ sample are described. The quarks are identified by the presence of highly energetic sprays of particles which release energy in a wide region of the calorimeter. The trigger identify these releases with a fast algorithm, then the full jet reconstruction is performed. Primary and secondary vertex finding algorithms are described, the former to select collisions where a hard interaction occurred, the latter to identify jets which contain heavy flavour quarks.

4.1 Introduction

The topology under study here is characterized by a very striking signature, that of at least 6 quarks in the final state. Quarks and gluons are not free particles in the Standard Model. They hadronize almost immediately giving rise to a spray of hadronic particles around the parent parton motion. The experimental signature of a jet is thus a cluster of approximately collinear hadrons which leave hits in the tracking system and spread out energy deposit in the calorimeter. Traditionally, only the calorimetric information is used to reconstruct the four-momenta of the jets. At first level of the trigger, time constraints allow only to collect information from single towers, then a quick reconstruction is performed at the second level of the trigger; finally the full reconstruction is performed at the third level once the time

constraint is released. The reconstruction algorithm used by CDF is a cone-based algorithm. After a jet four-momentum is defined, it has to be corrected for both calorimetry and physics effects in order to best reproduce the parton energy; this is done trough many steps each accounting for different effects. Finally, a systematic uncertainty on the jet energy scale is derived.

If a hard interaction occurred, we expect a large number of tracks to point to a certain region along the beam axis z. It is thus useful to exploit the z coordinate of all tracks at the point of closest approach to the beam pipe to look for a primary vertex as a signature of the occurrence of the hard interaction itself. Finally, the production of heavy flavour quarks is a feature which we will find very useful to get rid of the background, since it is characterized by long-lived mesons which produce secondary vertices displaced with respect to the vertex due to the primary interaction. We will describe the algorithm which identify jets containing such signatures and label them consequently as b-jets (b-tagging algorithm).

4.2 Trigger selection

The inelastic cross section for $p\bar{p}$ scattering has been measured at $\sqrt{s} = 1.8\,\text{TeV}$ and extrapolated at the actual configuration, $\sqrt{s} = 1.96\,\text{TeV}$, where it amounts to 61 mb [15]. The most recent theoretical calculations for the $t\bar{t}$ production cross section is about 7 pb, that is to say about 1 $t\bar{t}$ event produced every 10^7 inelastic collisions. Thus we need to exploit the kinematical and topological features of $t\bar{t}$ events to retain high efficiency while maintaining a low rate. Here we describe how objects are reconstructed at the various trigger levels and the requirements we make to select multijet events suitable to be good candidates for our analysis.

Level 1

At this level towers are not considered individually, but merged in pairs along η. The primitives are called "trigger towers" and are $\Delta\eta \times \Delta\phi \simeq 0.2 \times 15^o$ wide; the result is a 24×24 $\eta - \phi$ map of the calorimeter region extending to $|\eta| < 3.6$. The energy measured in each tower is the energy deposit weighted by the sine of

the centroid of the tower, $sin\theta$. The whole energy released in the calorimeter is also available at this time. All-hadronic $t\bar{t}$ events are characterized by large parton multiplicity, each of which has energy of the order of tens of GeV. What we want to detect are sprays of particles which are not expected to be fully contained in a single trigger tower; the energy threshold for Level 1 trigger requirement has to be set at a value smaller than the required spray energy in order not to loose efficiency. We make a very basic requirement asking at least one trigger tower with energy above 10 GeV.

Level 2

The Level 1 of the trigger implies rates which are unacceptable at Level 3. Level 2 of the trigger achieve the necessary rate reduction by considering in this case the energy associated to clusters of contiguous trigger towers. Let's first describe how these clusters are identified at Level 2. Two energy thresholds are needed to the cluster finder: the *seed* threshold defines which trigger towers have to be used as a starting point to the algorithm, while the *shoulder* threshold, on the other hand, essentially establishes the length of the recursive procedure which makes the third step. To reconstruct jets, the input values are $E_T(seed) = 3.0\,\mathrm{GeV}$, $E_T(shoulder) = 1.0\,\mathrm{GeV}$. The definition of clusters within an event requires a cluster finder algorithm which is explained in the following:

- **Step 1** All trigger towers with energy content above the seed threshold are identified and recorded as "seed towers". When all seed towers have been found, a second loop is performed on the remaining trigger towers in order to tag the "shoulder towers", that is, the trigger towers whose energy exceeds the shoulder threshold.

- **Step 2** Here, starting from one seed tower, the algorithm moves to the adjacent towers in η and ϕ and, if the adjacent tower is a seed or a shoulder tower, it is flagged as "found"; then the procedure is repeated starting from the "found" tower to its adjacents ones. The procedure is repeated until no more contiguous shoulder towers are found.

- **Step 3** The previous step is repeated for all seed towers not yet flagged as "found" in the previous iteration.

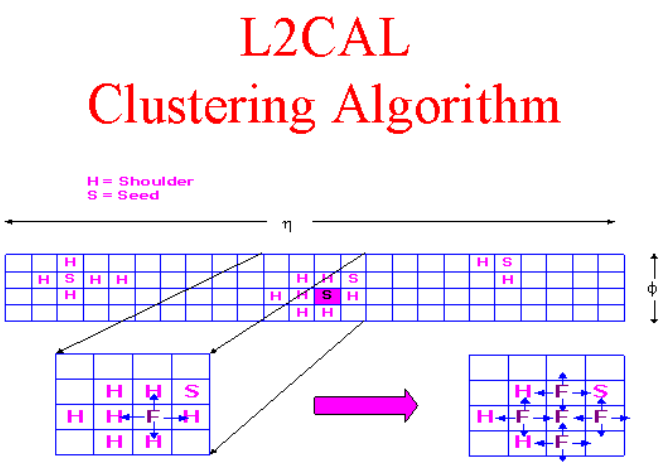


Figure 4.1: Schematic summary of the Level 2 cluster finder algorithm.

A summary of the Level 2 cluster finder algorithm is pictured in Fig. 4.2. When the procedure comes to an end, clusters are identified as groups of contiguous trigger towers that have been found around a seed tower. Each cluster is assigned an energy equal to the sum of the energies of all trigger towers belonging to the same group. The $\eta - \phi$ position of each cluster is then identified with the $\eta - \phi$ position of the seed tower which started the procedure. Beside the spatial position and the energy content, at Level 2 the informations on the number of clusters and the total clustered energy are available. We require for at least 4 clusters with transverse energy of 15 GeV each, and a sum of the individual clusters of at least 175 GeV[1].

[1] At the end of February 2005 there has been a change in Level 2 trigger in order to cope with the increasing instantaneous luminosity. The level 2 cut on the total transverse energy of all clusters has been raised from 125 to 175 GeV. Such a cut reduced the effective cross section of the multijet trigger by about a factor of 4, with high efficiency ($\approx 90\%$ on $t\bar{t}$ events with ≥ 6 jets). Given the high efficiency we decided to simplify the analysis by requiring $\sum E_T(L2) \geq 175$ GeV also on the data taken with the old trigger.

Level 3

The information provided by the cluster finding algorithm at the trigger level can be considered as a first-order jet reconstruction. At Level 3, a looser time constraint enables to exploit the full detector segmentation for a better jet energy and direction determination. The standard CDF jet clustering algorithm [19] is an *iterative cone algorithm* based on a fixed cone-radius (R) in $\eta - \phi$ space.

The jet-finding algorithm begins by searching for seed towers among all the calorimeter towers with a transverse energy above 1 GeV. In the plug calorimeter region, towers are grouped in ϕ into sets of three to mimic the central segmentation. The list of seed towers, ordered in decreasing transverse energy, is used to create preclusters. A precluster is formed by merging together an unbroken chain of contiguous seed towers. If a tower is outside a window of 7×7 towers surrounding the starting seed, it is used to form a new precluster. The precluster centroid is calculated taking the E_T-weighted average of the tower positions.

The clustering starts using the most energetic precluster, grouping all the towers with an E_T above 100 MeV which lie inside a cone of a radius R around the precluster seed centroid. Then the E_T-weighted centroid of the new cluster is calculated from the set of towers within the cluster cone, and a new cone is defined around this direction. Towers might be added or deleted from the new cone cluster during this procedure. The iterative process stops when the tower assignment to clusters remains unchanged.

The algorithm also provides a prescription for treating overlapping clusters. If the towers of one cluster are completely contained within another, the less energetic cluster is dropped. If two clusters have some common towers, the shared energy is computed by summing the E_T of the overlapping towers. If this shared energy is above the 75% of the total E_T of the less energetic cluster the two clusters are combined. If the shared energy is smaller than this threshold, the clusters remain separate, and each tower in the overlap region is assigned to the closest cluster in $\eta - \phi$ space. After all towers are uniquely assigned to clusters, the centroids are recalculated. This process of centroid computation and tower reshuffling is iterated until the tower lists remain fixed.

Trigger Level	***Requirement***	E_T (GeV)	$\sum E_T$ (GeV)
Level 1	≥ 1 trigger tower	≥ 10	-
Level 2	≥ 4 clusters	≥ 15	≥ 175
Level 3	≥ 4 jet	≥ 10	-

Table 4.1: Summary of the multijet trigger requirements as stated in the CDF II trigger table [18].

A cluster four-vector ($p_x,\ p_y,\ p_z,\ E$) is defined once a stable configuration is reached. A massless four-vector is assigned to each electromagnetic and hadronic tower having a magnitude equal to the energy deposit in the tower, and a direction defined by the unit vector pointing from the nominal detector origin to the center of the tower (calculated at the depth that corresponds to the shower maximum). Then all the tower four-vectors belonging to the same cluster are summed:

$$p_x = \Sigma_i p_x^i \qquad p_y = \Sigma_i p_y^i \qquad p_z = \Sigma_i p_z^i \qquad E = \Sigma_i E^i \tag{4.1}$$

Several jet variables are calculated subsequently from these quantities:

$$E_T = E \cdot \sin\theta \qquad \eta = -\ln\tan\left(\frac{\theta}{2}\right) \qquad \phi = \arctan\left(\frac{p_y}{p_x}\right) \tag{4.2}$$

where

$$\theta = \arcsin\left[\frac{\sqrt{p_x^2 + p_y^2}}{\sqrt{p_x^2 + p_y^2 + p_z^2}}\right]. \tag{4.3}$$

4.2.1 Jet Corrections

The transverse energy and momentum in the above definition are based on the raw calorimeter energy. These uncorrected quantities differ from the true partonic

four-momentum for a variety of reasons. Some of these originate from detector performance limitations:

- The calorimeter response to low-energy charged particles shows a non-linearity for momenta below 10 GeV.
- Charged particles with a transverse momentum below 400 MeV/c are confined inside the COT volume by the magnetic field and thus never reach the calorimeter. At slightly higher transverse momenta, the magnetic field can spread particles outside the jet cone.
- Particles showering in a poorly instrumented region of the calorimeter, like boundary regions between calorimeter modules or junctions between central, plug and forward subsystems, have on average a smaller energy response.

Others result from algorithm limitations and physics processes:

- Energy coming from soft-scattering processes or extra $p\bar{p}$ interactions in the same bunch crossing (underlying event), might be collected into the jet cone.
- Soft energy radiation and fragmentation effects might cause energy leakage due to particle loss outside the jet cone.
- Neutrinos and muons, embedded into jets, escape the calorimeter without depositing a substantial energy, resulting in a mismeasured jet energy.

The standard CDF jet correction procedure (JetCorr [20]) takes care of all these effects. The raw detector energy is studied as a function of:

- jet cone radius (R)
- the raw jet transverse momentum (p_T)
- the jet detector pseudorapidity (η_d)
- the electromagnetic fraction (em)
- the number of reconstructed vertices (N_V)

	P_0	P_1	P_2
$p_T \leq 100$ GeV	1.37	1.1946	-0.0008724
$p_T > 100$ GeV	-4.03	1.1794	-0.0001805

Table 4.2: Parameters used for the absolute jet energy scale corrections for cone size 0.4 (taken from Run I).

This jet correction consists of many steps, some of which are devoted to obtaining a flat response in η, while others account for absolute energy mismeasurements, the underlying event and out-of-cone effects. In this analysis the jets are corrected for all these effects to get the full "parent parton" transverse energy. The functional expression of the jet correction procedure is described by the following equation:

$$P_T(R) = (P_T^{raw}(R) \times f_{rel} - UEM(R)) \times f_{abs}(R) - UE(R) + OC(R) \qquad (4.4)$$

where R is the cone radius. Each correction term appearing in the above equation is detailed in the following.

f_{rel} : Relative Jet Energy Corrections

The first step in jet correction is to adjust the non uniform energy response of the calorimeter making it independent on η (see Fig. 4.2). The reason for this nonuniformity is the presence of uninstrumented regions (cracks) and differing tower segmentation.

The central calorimeters CEM/CHA are the best understood calorimeters in CDF II and they cover the regions far away from cracks. Their response, measured with test beam data and checked with the COT information during the data taking, is found to be linear.

The correction procedure is based on di-jet p_T balancing. Assuming the di-jet production to be a $2 \rightarrow 2$ process, the transverse energy of the two jets has to be equal. Confining one jet to lie in the central region, $0.2 < |\eta| < 0.6$ (*trigger jet*), and the other spanning the whole calorimeter (*probe jet*), a scale factor can be derived as a function of η and p_T. This correction is parametrized for three different cone

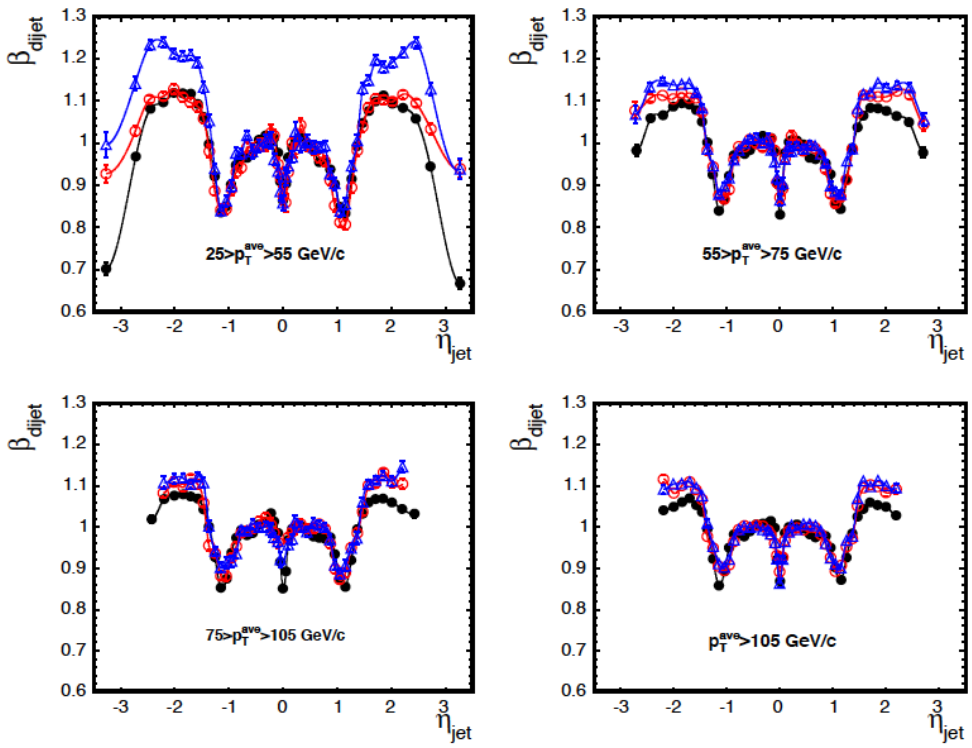


Figure 4.2: Relative calorimeter response of simulated data compared to di-jet data. Jets are reconstructed with a cone radius of 0.4.

sizes: 0.4, 0.7 and 1.0.

The energy response also changes with time, mostly due to the ageing of the phototubes. Data are corrected for this time variation.

f_{abs} : Absolute Jet Energy Corrections

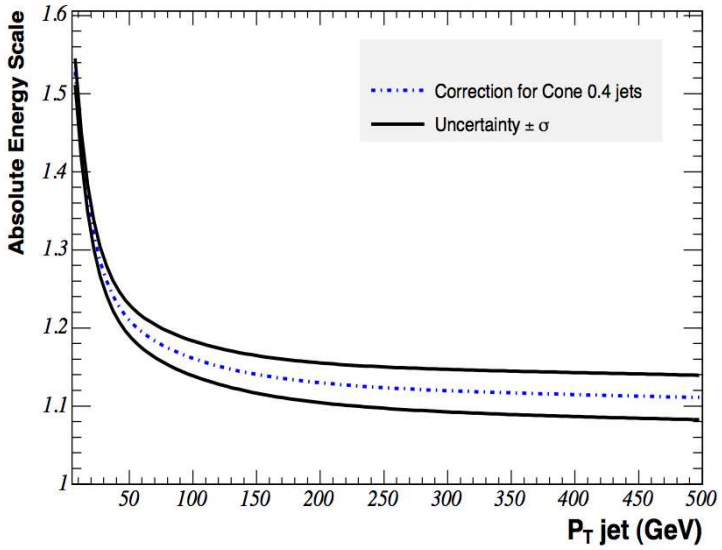


Figure 4.3: Absolute jet energy scale correction factor as a function of the jet P_T for a cone size $R = 0.4$., together with its uncertainty.

The conversion factor from calorimeter energy to particle-level energy is called absolute energy correction. This factor depends on the jet fragmentation properties and on the non-linearity of the calorimeter response due to the difference in the response to π^0 and $\pi^{\pm}$.

In Run I the absolute energy corrections were determined using jets generated with tuned Monte Carlo. In this context the reconstructed jet energy is compared to the sum of the four momenta of all the MC particles lying inside the jet cone. A quadratic fit is used to parametrize the mean jet response as a function of the transverse momentum for different cone sizes:

$$f_{abs} = (P_0 + P_1 \cdot p_T + P_2 \cdot p_T^2)/100. \tag{4.5}$$

Here P_i are the fit parameters, and their values are shown in Table 4.2. In Figure 4.3 the correction function is plotted.

UEM : Multiple Interaction Corrections

With the current instantaneous luminosity and with 36 proton and anti-proton bunches, several interactions per bunch crossing are expected. Moreover, since the number of interactions per bunch crossing follows a Poisson distribution, some events have more overlapping minimum bias interactions. The energy released in these collisions could fall inside the clustering cone, and should be subtracted from the jet energy.

This correction is a linear function of the number of primary vertices in the event, and its value is derived from minimum bias events measuring the transverse energy in a random cone.

UE : Underlying Event Correction

The underlying event is the result of soft collisions between the spectator partons inside the proton and antiproton. These interactions produce background energy in the whole calorimeter which contributes to the jet energy although it is not associated to the hard scattering. This energy is estimated using single vertex minimum bias events, measuring the transverse energy density (ΣE_T divided by the whole calorimeter surface) in the central calorimeter. The jets are then corrected according to their cone radius by subtracting a definite amount of energy proportional to the energy density mentioned before.

OC : Out-of-Cone Corrections

To scale the jet energy to the parent parton value the out-of-cone correction is applied. This factor is determined using Monte Carlo events. This correction is independent of the calorimeter performance while it is a function of the jet transverse momentum and the cone radius. The additional energy is parametrized as:

$$p_T^{ooc} = A[conesize] \cdot (1.0 - B[conesize] \cdot e^{-C[conesize] \cdot p_T}) \tag{4.6}$$

	A	B	C
Cone radius 0.4	22.999	0.915	0.00740

Table 4.3: Out-of-cone energy parameters (taken form Run I).

The values of the parameters describing the out-of-cone correction are listed in Table 4.3.

4.2.2 Jet Corrections Systematics

Each correction described above has associated systematic uncertainties [20] which are shown in Fig. 4.5 as a function of the jet E_T. The uncertainties are strongly

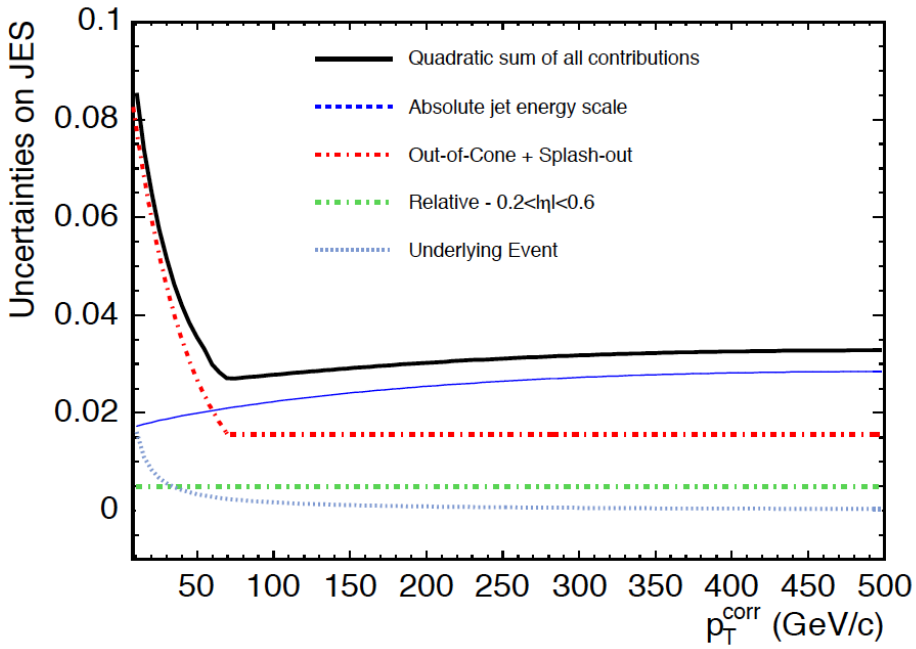


Figure 4.4: Systematic uncertainty on the jet energy scale as a function of the corrected jet E_T.

dependent on the jet transverse energy. For low-E_T jets ($15 - 30$ GeV) the total systematic uncertainty is about 8%. The dominant contribution is coming from the out-of-cone correction. For high-E_T jets (≥ 50 GeV) the uncertainty is at the level of 4%.

It has been observed that the Monte Carlo simulation of the calorimeter does not exactly reproduce the CDF II calorimeter energy response (see Fig. 4.2) as a function of η. Therefore an extra contribution to the relative correction factor has been introduced to account for the difference between data and Monte Carlo.

4.3 Primary vertex Reconstruction

The first step in the event building is the primary vertex reconstruction where the $p\bar{p}$ interaction occurred. It proceeds through an iterative algorithm that extrapolates the track direction to the beam line. The algorithm starts extracting a primary vertex candidate with a p_T-weighted fit using all the reconstructed tracks. Subsequently, tracks with a high impact parameter with respect to the fitted vertex are discarded. This procedure is repeated until a stable χ^2 is reached.

The resulting resolution on the primary vertex position in the transverse plane ranges between 6 and 26 μm, depending on the event track multiplicity. We require the events to have at least one well-reconstructed primary vertex and to lie in the detector fiducial region ($|z_{vertex}| \leq 60\,\mathrm{cm}$).

4.4 Secondary vertex Reconstruction

Particles with long lifetime can produce in their decays vertices that are displaced with respect to the primary vertex. We are interested in particular to the events in which b quarks are produced, since we expect to find them in our decay channel. High energy b quark, while hadronizing, form B mesons and a shower of other hadronic particles. These jets are thus characterized by the presence of long-lifetime B mesons, $\tau_B \sim \mathcal{O}(1\,\mathrm{ps})$. In order to reconstruct these vertices, we need the high precision of the silicon detector in the spatial determination of the hits. The algorithm which looks for a secondary vertex, called SecVTX, needs the tracks from the COT, and the SVX II and ISL. At least two tracks which pass a minimum threshold on transverse momentum (typically $0.5\,\mathrm{GeV/c}$) and whose closest approach to the

z axis is sufficiently displaced[2]are needed. Now, in order to tag a jet, the algorithm needs to look if one of these vertices lies inside a jet. We define as L_{xy} the projection of the vector connecting the primary to the secondary vertex, $\vec{x}_{sec} - \vec{x}_{prim}$, on the transverse jet direction:

$$L_{xy} = (x_{sec} - x_{prim}) \cdot cos\phi_{jet} + (y_{sec} - y_{prim}) \cdot sin\phi_{jet} \tag{4.7}$$

SecVTX considers as generated by heavy quarks the jets with L_{xy} positive and whose significance ($L_{xy}/\sigma_{L_{xy}}$) is greater than 7.5. In terms of tagging performance, some difference is found between data and Monte Carlo events. A correction factor (scale factor) is then introduced to compensate for such difference. The scale factor (SF) has been calculated by comparing the per-jet b-tagging efficiency (ϵ_{btag}) in QCD $b\bar{b}$ simulated events to the experimental efficiency evaluated in heavy flavour enriched events, that is events where one electron is embedded within a jet. Such factor amounts to 0.89 ± 0.07. A similar determination has been performed for c-jets, finding 0.89 ± 0.14.

4.5 Lepton reconstruction

In this work we analyze events in which we expect to find hadrons only. In order to be orthogonal to other analysis where the presence of a lepton is required, we want to identify them in order to veto thei presence and obtain a sample which is free of events where the top quark decays leptonically. Clearly, both electrons and muons need tracks to be reconstructed by the tracking system. Here we briefly describe how leptons are identified at CDF.

Electrons

[2]Which means

- significance on d_0: $S(d_0) = d_0/\sigma_{d_0} > xxx$;
- significance on z_0: $S(z_0) = (z - z_0)/\sigma_{z_0} > xxx$.
- overall significance $\sqrt{S(d_0)^2 + S(z_0)^2} > xxx$.

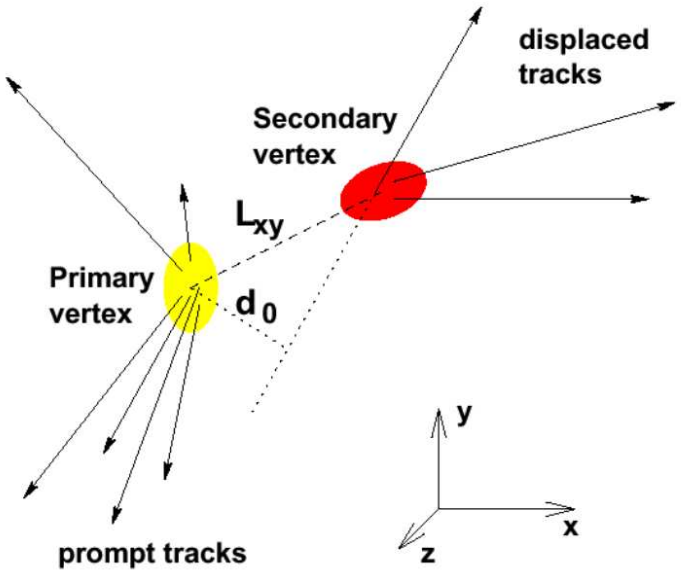


Figure 4.5: Schematic representation of the seconday vertex finding algorithm.

In addition to the track, in order to define an electron, we need a deposit in the electromagnetic calorimeter. This deposit has to be coherent with the hypothesis that it comes from an electron, i.e., the extrapolation of the tracks to the calorimeter must point to the calorimetric tower in issue, and the ratio of hadronic to electromagnetic energy has to be below a threshold, $E_{had}/E_{em} < 0.055 + 0.00045 \cdot E$ where E is the total energy of the cluster in GeV, to reject hadrons faking an electron. The amount of energy released has to be compatible with the momentum measured in the tracking system. Moreover, a CES cluster compatible with an electromagnetic shower has to be present and matched to the candidate track. Finally, we require the calorimetric release to be isolated, i.e. that the energy in the radius $R = 0.4$ around that tower be below 10% of the cluster E_T.

Muons

We form muon candidates by matching a track to "stubs" in the muon chambers. The stubs are reconstructed as line segments formed by hits in the CMU, CMP and CMX detectors. Furthermore, the energy deposited in the calorimeter by the candidate track is required to be consistent with that of minimum ionizing particles to remove fake hadronic particles. Finally, the same isolation requirement as for

the electron candidates is applied, with the exception that the energy in the cone is compared to the track P_T instead.

4.6 Missing E_T

The missing transverse energy $\vec{\not\!\!E}_T$ is calculated as the vector sum of the energy in each calorimeter tower multiplied by a unit vector pointing to the azimuthal direction of the tower. If isolated high momentum muons are found in the event, the $\vec{\not\!\!E}_T$ is corrected by subtracting the muon energy in the calorimeter and adding the muon $\vec{p}_T$ to the vector sum. The $\not\!\!E_T$ is defined as the magnitude of $\vec{\not\!\!E}_T$.

4.7 Run Requirements

Each run stored by the CDF experiment is associated to a database entry which specifies the amount of integrated luminosity written to tape and the status of all the sub-detectors. The runs used in this analysis are selected from those taken during the data taking period from March 2002 until June 2006. Collider runs which are known to contain problems are removed. The runs are required to have the calorimetry and the silicon tracker on and in good condition, and to contain at least $10\,\mathrm{nb}^{-1}$ of data written to tape.

4.8 Luminosity Measurements

The total integrated luminosity is calculated from the rate $R_{p\bar{p}}$ of inelastic $p\bar{p}$ events measured by the luminosity monitor, Cerenkov Luminosity Counters (CLC). Defining ϵ_{CLC} as the CLC acceptance, and σ_{in} the inelastic cross section, the luminosity is expressed by:

$$L = \frac{R_{p\bar{p}}}{\epsilon_{CLC} \cdot \sigma_{in}} \tag{4.8}$$

The CLC acceptance is estimated to be $(60.2 \pm 2.4)\%$ [**?**]. The inelastic cross section is derived by scaling the CDF measurement of $\sigma_{in} = 60.4 \pm 2.3$ mb at $\sqrt{s} = 1.8$ TeV to 1.96 TeV, resulting in $\sigma_{in} = 61.7$ nb.

Using these numbers and requiring the goodness of the run as described in Section 4.7, the data used in this analysis correspond to an estimated integrated luminosity of

$$(1.02 \pm 0.06)\,\mathrm{fb}^{-1}. \tag{4.9}$$

The 6% quoted uncertainty is dominated by the measurements of the absolute normalization of the CLC acceptance for a single $p\bar{p}$ inelastic collision.

Chapter 5

The event selection

The identification of $t\bar{t}$ candidate events in the all-hadronic decay mode is made difficult by the presence of a large background of QCD multijet production. In this chapter we explain how this problem has been overcome by the use, for the first time in this channel, of a neural network approach. The various processes which make up the background are modeled through a data-driven technique.

5.1 The multijet dataset

We described in the previous chapter the requirements we set at the trigger level to collect a large amount of hadronically decaying top pairs. The resulting multijet dataset amounts to an integrated luminosity of about $1.02\,\mathrm{fb}^{-1}$. The trigger has high efficiency on the signal, but on the other hand the background still dominates by 3 orders of magnitude before any event selection. In order to model the signal in this analysis we use the PYTHIA v. 6.2 [23] and HERWIG v. 6.4 [24] leading-order Monte Carlo generators with parton showering followed by a simulation of the CDF II detector. The reference top quark mass chosen for the event selection studies is $M_{\mathrm{top}} = 175$[1] $\mathrm{GeV/c^2}$. The background is expected to come mostly from QCD multijet production of light and heavy flavour partons. The cross section of such processes are poorly known; moreover, we would need a huge number of Monte

[1]Since this value is close to the most recent Tevatron average of the previous measurements.

Carlo simulated events to characterize the background. For this reason, we model the background from the multijet data events themselves: this is possible since the signal fraction at the initial stage is expected to be very small, $\approx 0.3\%$.

5.1.1 Preliminary requirements

We apply a set of preliminary cuts to our sample in order to model the basic features of the kinematics and topology of $t\bar{t}$ events. These prerequisites require the run to be a good one (i.e. all detector parts needed for the analysis are in working conditions), to have at least one well-reconstructed primary vertex (as defined in the previous chapter) which lies inside the luminous region ($|Z_V| \leq 60\,\mathrm{cm}$). In the following, each jet is required to have $E_T \geq 15$ GeV and $|\eta| \leq 2$ after all jet corrections have been applied. In order to remove the events from the $t\bar{t}$ leptonic channels, we veto events containing any well identified high-P_T electrons and muons as defined in [25]. We require that $\frac{\not\!\!E_T}{\sqrt{\sum E_T}}$ be $< 3\sqrt{\mathrm{GeV}}$ where the missing transverse energy, $\not\!\!E_T$, defined in the previous chapter, is corrected for both the momentum of any identified muons and the position of the $p\bar{p}$ collision point, while $\sum E_T$ is obtained by summing the E_T's of all the selected jets[2].

The background after the b-tagging application will be evaluated using a data-driven estimate based on the parametrization of the probability of tagging a jet as a b-jet according to jet and event features. This probability is derived in a control sample of events with 4 jets and is then applied to events with different jet multiplicity. This technique works because of the parametrization of the probability in terms of the most important quantities like jet-E_T, jet track multiplicity, and number of vertices. We notice however that in going from 4 to 6 and more jets, jets tend to be closer in direction, and this might affect the tag rate because of partial overlap between jets and their tracks. The current statistics does not allow us to introduce additional parameters in the parametrization so we decided to cut instead on the minimum separation between jets (ΔR_{min}), requiring events to have

[2]Such cut is needed also to be orthogonal to another signature used in CDF to search for top semileptonic events characterized by high $\not\!\!E_T$ [26].

$\Delta R_{min} \geq 0.5$. About 3 481 000 events pass these preselection requirements (S/B $\sim 1/1000$). Finally, events with 4 or 5 jets are used as *control regions*, while the *signal region* is obtained by selecting events with a number of jets $6 \leq N_{\text{jets}} \leq 8$ to optimize the signal fraction. A total of 506 567 events pass this additional requirement with an expected S/B of about 1/370.

5.2 A neural network for kinematical selection

The main features of events from $t\bar{t}$ production, where both top quarks decay hadronically, are high jet multiplicity with b-jets with transverse energy around 60 GeV and light flavour jets with energy around 50 GeV; as a consequence we expect to have a large energy deposit in the calorimeter. These events are produced mostly centrally in the detector, another feature which we will want to exploit. The light flavour jets are expected to contain informations on the W bosons, and once combined with the b-jets, to contain informations on the top presence. Finally, as stated in the previous chapter, b-jets are expected to contain B mesons whose long life-time can be exploited to identify secondary vertices within a jet and thus tag it as a originated by a b quark.

We introduce here a neural network approach to recognize in more detail the features of signal and background events, including correlations between the kinematical variables. We thus expect a better separation between signal and background relative to a "cascade cut" approach where correlations are not fully considered. The relevant kinematical quantities enter as input nodes in a neural network, and the output of the network will become the variable we cut on. The performance of a neural network depends on the choice of the input variables, which should exhibit good discrimination between signal and background. The number of variables used should allow the best possible description of the event properties but, at the same time, too many input variables might worsen the performance. As a guideline we study different neural network configurations, in terms of inputs and hidden nodes, adding a few variables at a time, finally looking for the best performance in terms of S/B. The first quantities considered are those used in the previous analysis at

CDF in this channel [27, 45]: the total transverse energy of the jets, $\sum E_T$; the quantity $\sum E_{T,3} \equiv \sum E_T - E_T^1 - E_T^2$, obtained by removing the contribution of the two jets with the highest E_T; the centrality, defined as $C = \frac{\sum E_T}{\sqrt{\hat{s}}}$, where $\sqrt{\hat{s}}$ is the invariant mass of the multijet system; the aplanarity A, defined as $A = \frac{3}{2}\mathcal{Q}_1$, where $\mathcal{Q}_1$ is the smallest of the three normalized eigenvalues of the sphericity tensor, $M^{ab} = \sum_j P_j^a P_j^b$, calculated in the center-of-mass system of all jets, where the indices a and b refer to the spatial components of the jet four-momentum P_j. In addition we consider the dynamical properties of dijet and trijet systems through the use of the minimum and maximum value of the invariant mass among all possible jet permutations: M_{2j}^{min}, M_{2j}^{max}, M_{3j}^{min} and M_{3j}^{max}. Finally we obtain another set of discriminating variables combining the transverse energy of the jets with their emission direction in the center-of-mass frame of all jets, represented by the angle $\theta^\star$ between the jet and the proton beam axis. We define the quantity $E_T^\star = E_T \sin^2\theta^\star$ which tends to have larger values in the signal in comparison to the background events; this effect is enhanced for the jets with higher E_T. The variables we choose as additional inputs to the neural network are then $E_T^{\star,1}$ and $E_T^{\star,2}$ for the two highest-E_T jets, and $\langle E_T^\star \rangle$ defined as the geometric mean over the remaining $(N_{\text{jets}} - 2)$ jets. The 11 variables used as inputs to the neural network are summarized in Table 5.1.

Variable	**Description**
$\sum E_T$	Scalar sum of the transverse energies of all jets
$\sum_3 E_T$	As above, except the two leading jets
C	Centrality
A	Aplanarity
M_{2j}^{min}	Minimum dijet invariant mass
M_{2j}^{max}	Maximum dijet invariant mass
M_{3j}^{min}	Minimum trijet invariant mass
M_{3j}^{max}	Maximum trijet invariant mass
$E_T^{\star,1}$	$E_T \sin^2\theta^\star$ for the leading jet
$E_T^{\star,2}$	$E_T \sin^2\theta^\star$ for the next-to-leading jet
$\langle E_T^\star \rangle$	Geometric mean over the remaining jets

Table 5.1: Input variables to the neural network.

Comparisons of the background-dominated data and Monte Carlo generated sig-

nal events for the 11 kinematic variables are shown in Figs. 5.1, 5.2 and 5.3.

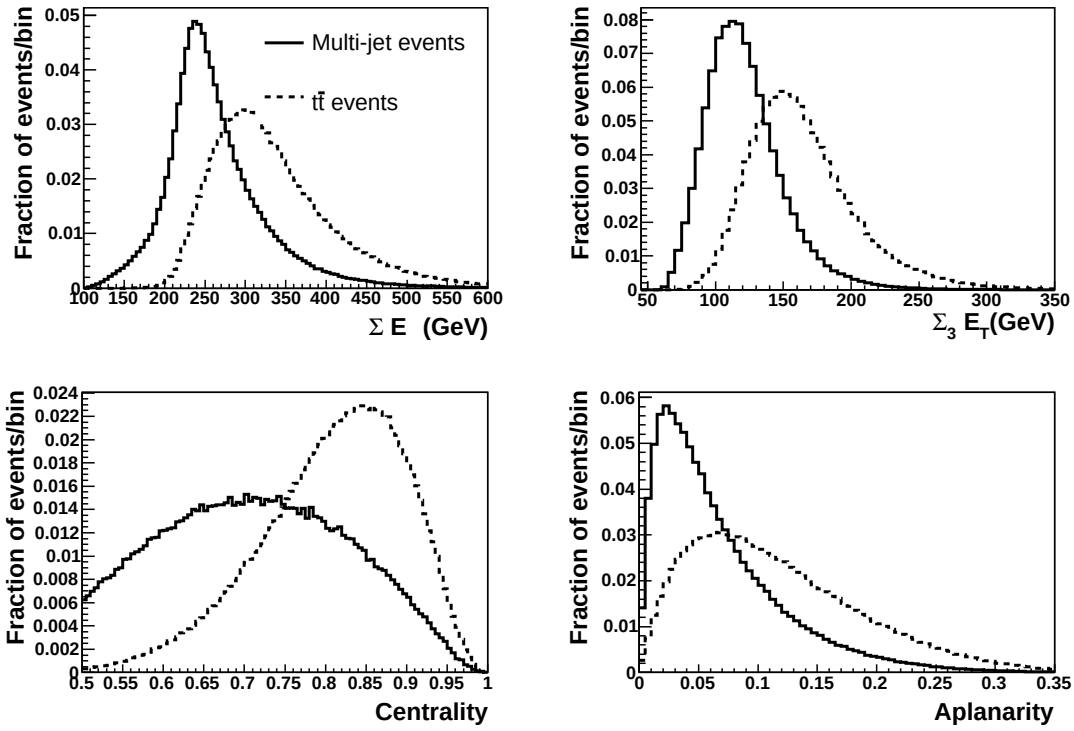


Figure 5.1: Kinematic distributions in multijet (solid histogram) and $t\bar{t}$ Monte Carlo (dashed histogram) events with $6 \leq N_{\text{jets}} \leq 8$. Clock-wise from top-left: $\sum E_T$, $\sum_3 E_T$, aplanarity, centrality. All histograms are normalized to unity.

5.2.1 Training the neural network

The neural network chosen here is the Multilayer perceptron (MLP [28]) as implemented inside ROOT [29] through the class *TMultiLayerPerceptron*. This is a simple feed-forward network with an input layer, some hidden layers and an output layer. Among the configurations investigated, the one which provides the largest expected S/B has two hidden layers with 20 and 10 hidden nodes respectively, and 1 output node. As for the output, we choose one single output node which represents the value, N_{out}, we will be cutting on in our selection. We give in Fig. 5.4 a pictorial representation on the neural network layout. The aim of the training/learning process is to minimize the total error on a set of weighted examples, where the error is defined as the quadrature sum (divided by 2) of the errors on each individual

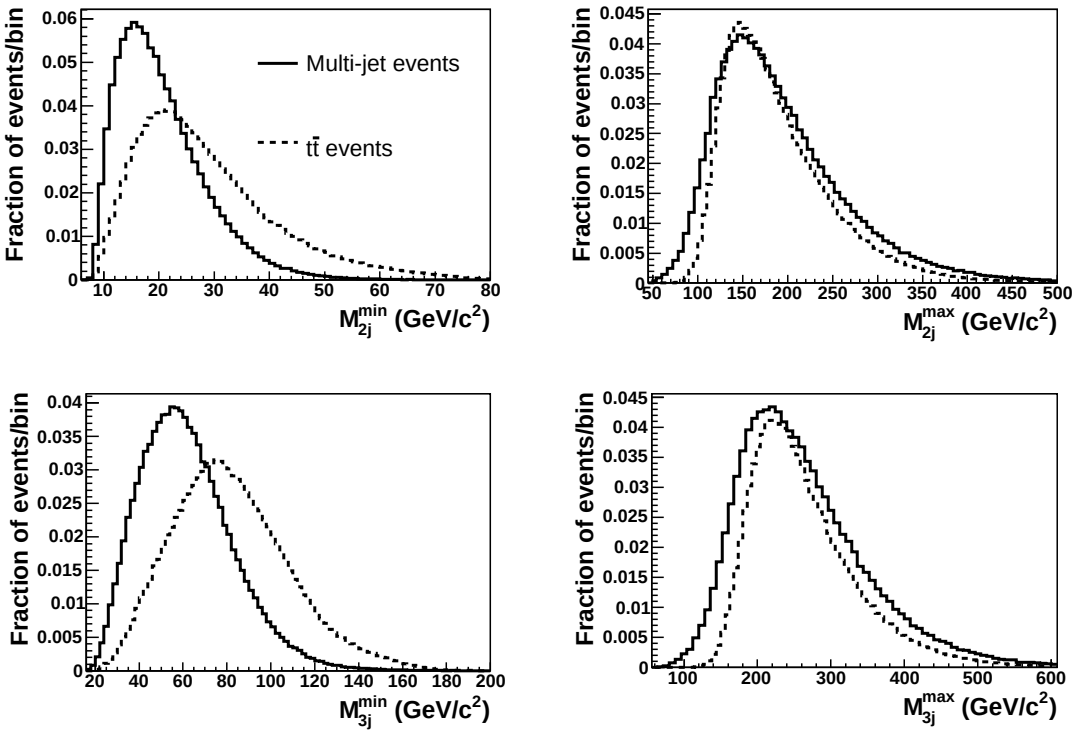


Figure 5.2: Kinematic distributions in multijet (solid histogram) and $t\bar{t}$ Monte Carlo (dashed histogram) events with $6 \leq N_{\text{jets}} \leq 8$. Clock-wise from top-left: M_{2j}^{min}, M_{2j}^{max}, M_{3j}^{max} and M_{3j}^{min}. All histograms are normalized to unity.

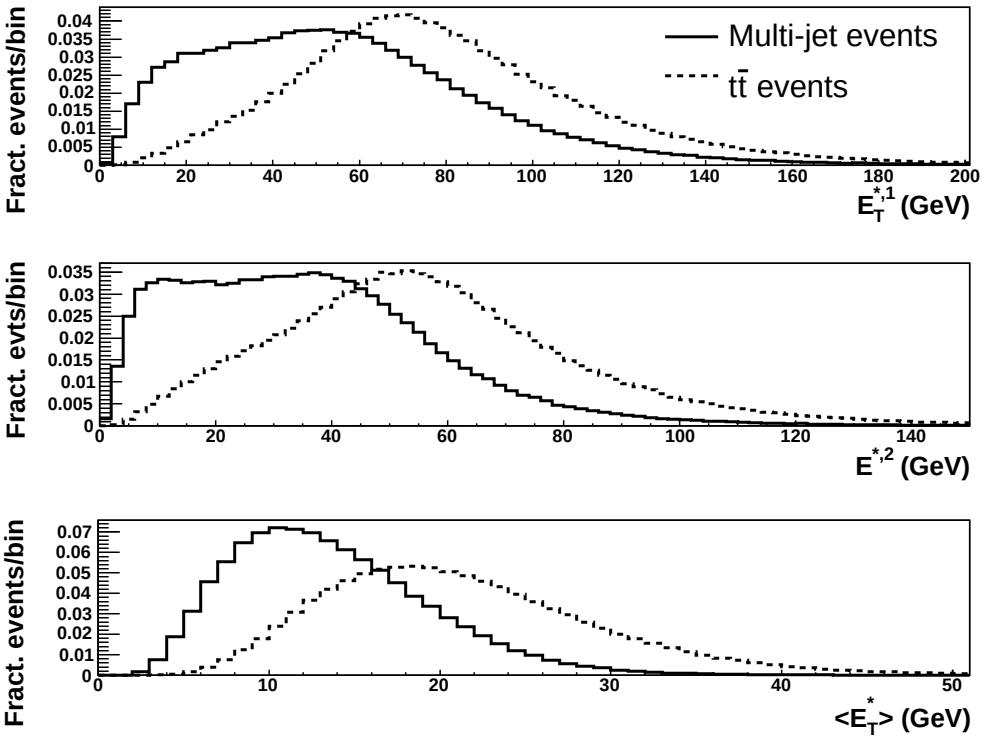


Figure 5.3: Kinematic distributions in multijet (solid histogram) and $t\bar{t}$ Monte Carlo (dashed histogram) events with $6 \leq N_{\text{jets}} \leq 8$. From top: $E_T^{\star,1}$, $E_T^{\star,2}$ and $\langle E_T^{\star} \rangle$. All histograms are normalized to unity.

output node (only 1 node in our case). The *TMultiLayerPerceptron* C++ class implements a variety of learning methods, but we consider here only the Broyden, Fletcher, Goldfarb, Shanno (BFGS) method [30] which implies the computation of a $N_{weights} \times N_{weights}$ matrix and is considered (by the MLP authors for instance) to be more powerful. This method is also much faster than the stochastic one. In the BFGS, weights are set to the minimum along the line defined by the conjugate gradient. Parameters are τ and *Reset*, which defines the epochs where the direction is reset to the steepest descent. The values we will be using are the default ones, $\tau = 3$ and $Reset = 50$. We do the training on background/signal samples of equal amount: i.e. 507 thousand background events and 507 thousand signal events with $6 \leq N_{jets} \leq 8$ and $\Delta R_{min} \geq 0.5$. Half of the sample will be used indeed for the training while the other half is used as test sample. The training is repeated for 100 epochs (defined as reading and testing over the whole training sample).

5.3 Applying the neural network

The learning of the neural network is quite fast: Fig. 5.5 shows the error as we proceed with the training and testing; in less than about 20 epochs the errors are stable. The output of the neural network, N_{out}, after the training is shown in Fig. 5.6 as evaluated on the test set (of half the events). The training took about 1000 cpu minutes on a standard PC. The outcome of the neural network training is exported as a standalone C++ class, containing all the weights for the network. When we apply back the neural network to the whole sample of 507 thousand multijet events and to the $t\bar{t}$ events normalized to the expectation at this level, i.e. $\simeq 1300$ events, we obtain the distributions of Fig. 5.7. The region $N_{out} \geq 0.8$ is where we can gain quite much in S/B, because of the fast rise in the signal distribution.

The behavior of this neural network is described in Fig. 5.8 where we show the S/B as a function of the efficiency. We quantify the performance in comparison with the "cascade cuts" selection used in earlier analysis [27] ($\epsilon_{old} = 6.43\%$, $S/B_{old} = 1/24$ assuming $\sigma_{t\bar{t}} = 6.7$ pb). We see that requiring for instance $N_{out} \geq 0.91$ gives a comparable efficiency ($\epsilon_N = 6.79 \pm 0.01(stat)\%$) but a much better $S/B = 1/16$, i.e. a 60% relative increase.

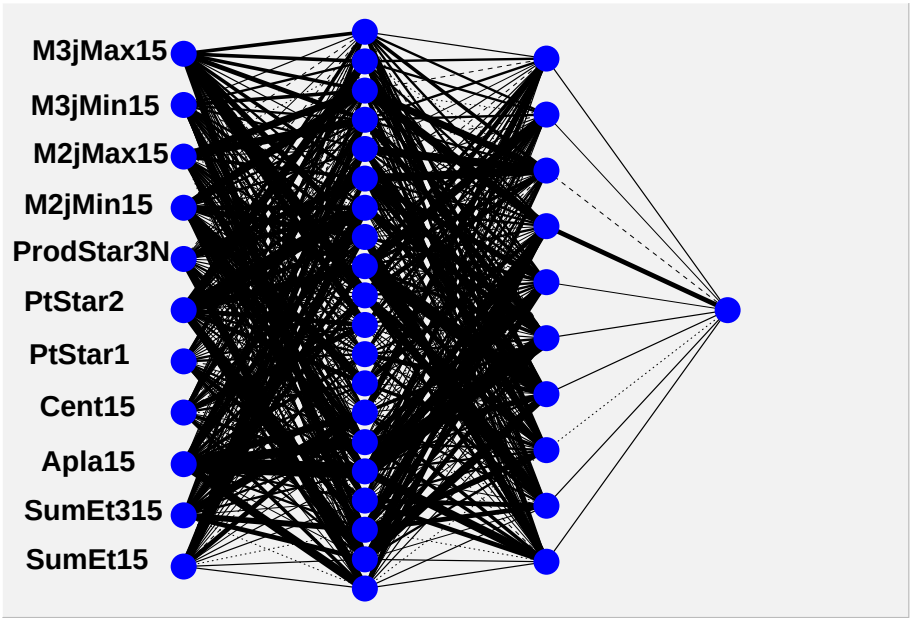


Figure 5.4: Neural network pictorial description. The thickness of the lines is proportional to the weight.

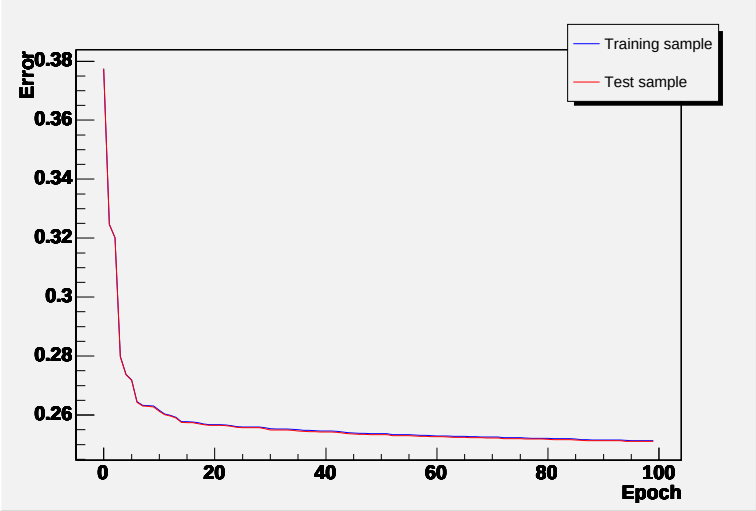


Figure 5.5: Variation of training and testing error as a function of the training epoch. (Note: the two curves are indistinguishable)

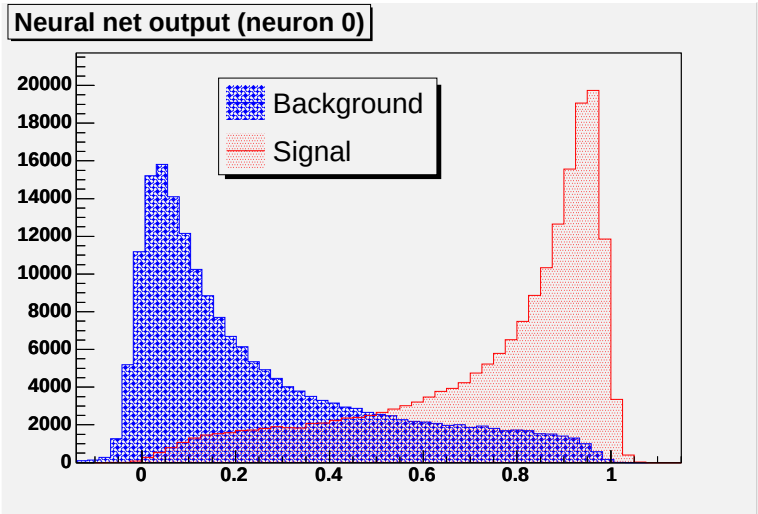


Figure 5.6: N_{out} distribution for the "test sample".

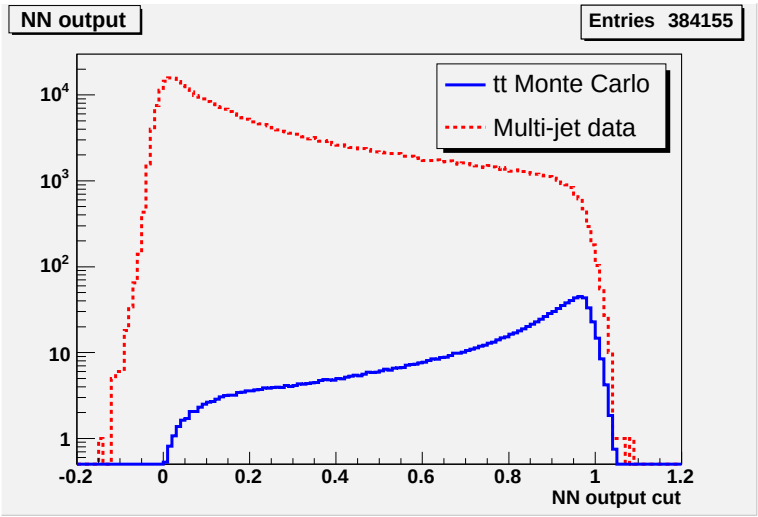


Figure 5.7: N_{out} distribution for the whole sample (the $t\bar{t}$ distribution is normalized to the expectation we get from the theoretical cross section and the calculated acceptance with $6 \leq N_{jet} \leq 8$ and $\Delta R_{min} \geq 0.5$).

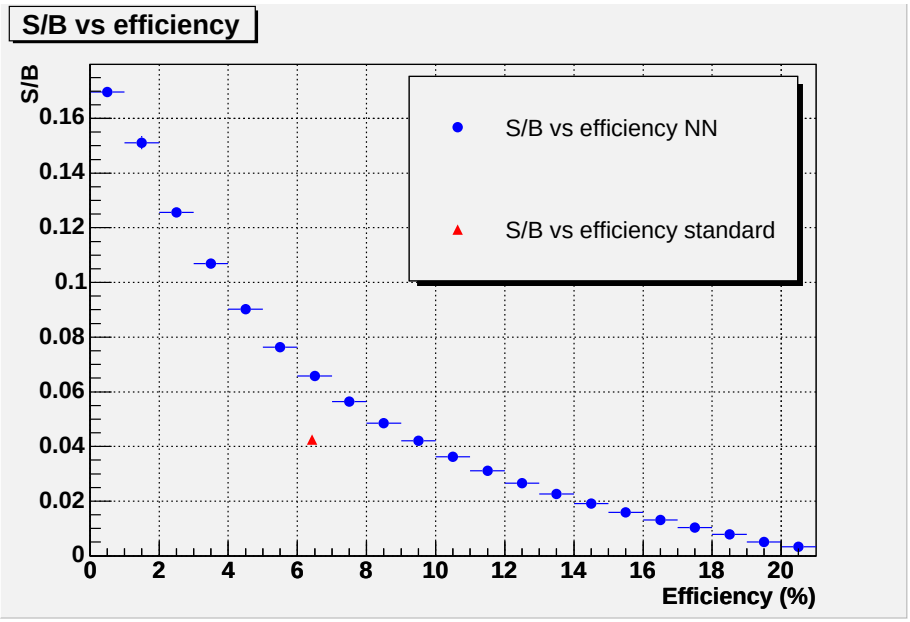


Figure 5.8: Shown here is the signal over background ratio S/B as a function of the selection efficiency. As a comparison is shown (red triangle) the value for the "old" selection.

It is interesting to see the effect of the neural network cut on each kinematical distribution, and compare the multijet kinematical distributions with the expectation for $t\bar{t}$ events. Fig. 5.9 is for the first 4 kinematical variables, Fig. 5.10 is for $E_T^{1\star}$, $E_T^{2\star}$, and $\langle E_T^* \rangle_{3N}$. Fig. 5.11 is for the dijet and trijet masses. All the plots show a similarity between the distributions for the selected multijet events and $t\bar{t}$ events.

5.4 Neural network selection

If one wants to consider the neural network kinematical selection before and independently of the b-tagging, then the optimization of the cut can be based on the maximum of the signal statistical significance, $S/\sqrt{B+S}$. As shown in Fig. 5.12 such a maximum is reached around 0.91. The cut ≥ 0.91 happens also to be the cut which gives similar efficiency as the previous selection (6.79% vs 6.43%), so we know already that such a cuts improves the S/B from 1/24 to 1/16. We will motivate further this choice in the next chapter, where we will see that not only it maximizes the statistical significance of the sample, but also allows the smallest sta-

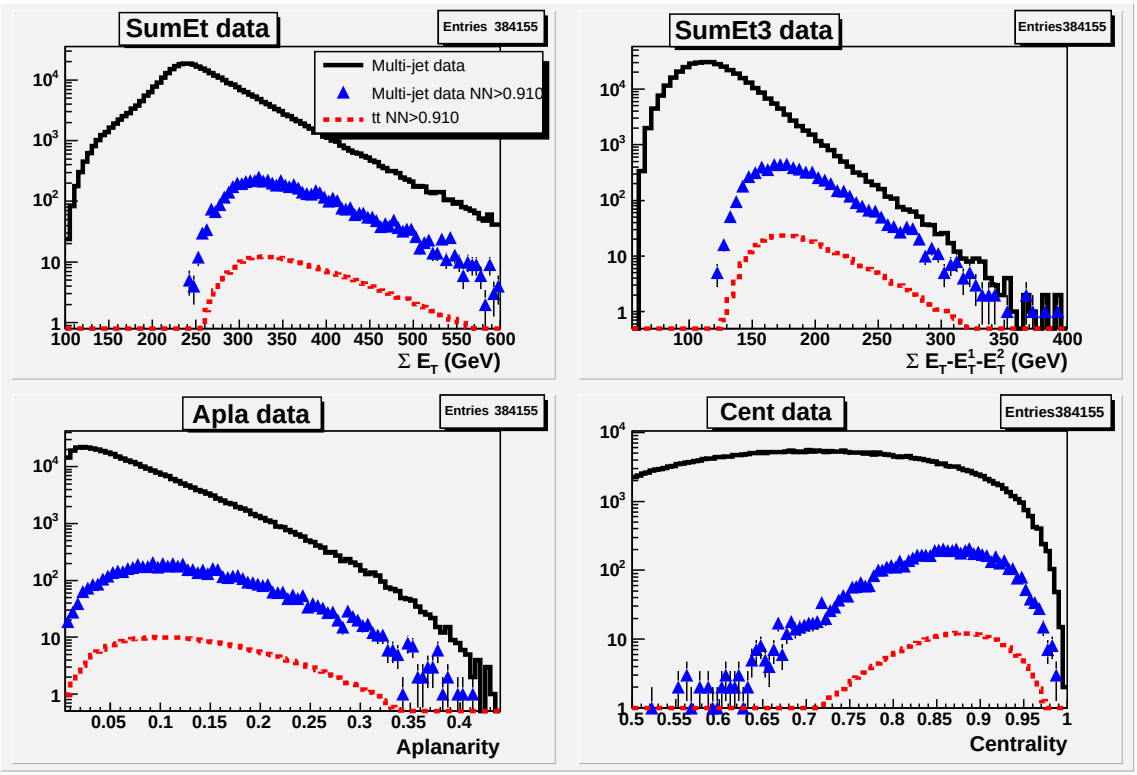


Figure 5.9: $\sum E_T$ (top-left), $\sum_3 E_T$ (top-right), aplanarity (bottom-left), and centrality (bottom-right) distributions for multijet events with $6 \leq N_{jet} \leq 8$ and $\Delta R_{min} \geq 0.5$ (black histogram) compared to events surviving the neural network cut at 0.91 (blue triangles). Also shown with a dashed red line is the $t\bar{t}$ expectation ($\sigma_{t\bar{t}} = 6.7$ pb). Plots based on a subset of 384 155 out of 506 567 events.

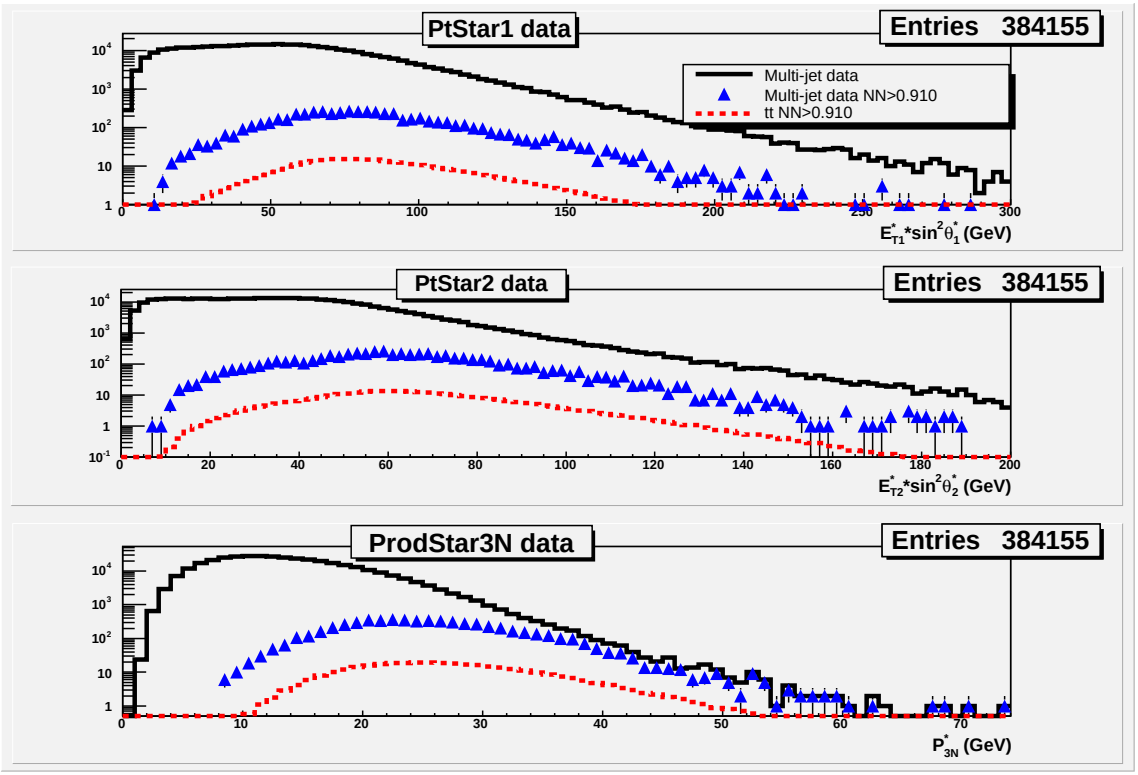


Figure 5.10: $E_T^{1\star}$ (top), $E_T^{2\star}$ (middle) and $\langle E_T^{\star}\rangle_{3N}$ (bottom) distributions for multijet events with $6 \leq N_{jet} \leq 8$ and $\Delta R_{min} \geq 0.5$ (black histogram) compared to events surviving the neural network cut at 0.91 (blue triangles). Also shown with a dashed red line is the $t\bar{t}$ expectation ($\sigma_{t\bar{t}} = 6.7$ pb). Plots based on a subset of 384 155 out of 506 567 events.

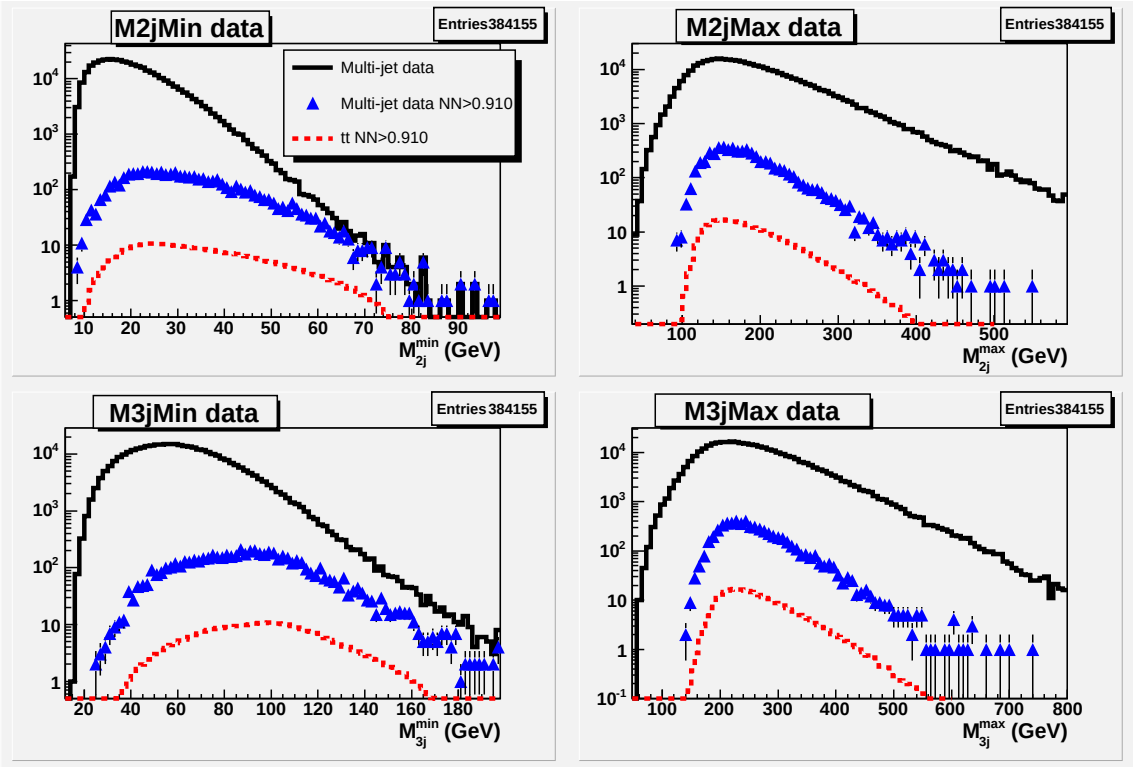


Figure 5.11: M_{2j}^{min} (top-left), M_{2j}^{max} (top-right), M_{3j}^{min} (borrom-left) and M_{3j}^{max} (bottom-right) distributions for multijet events with $6 \leq N_{jet} \leq 8$ and $\Delta R_{min} \geq 0.5$ (black histogram) compared to events surviving the neural network cut at 0.91(blue triangles). Also shown with a dashed red line is the $t\bar{t}$ expectation ($\sigma_{t\bar{t}} = 6.7$ pb). Plots based on a subset of 384 155 out of 506 567 events.

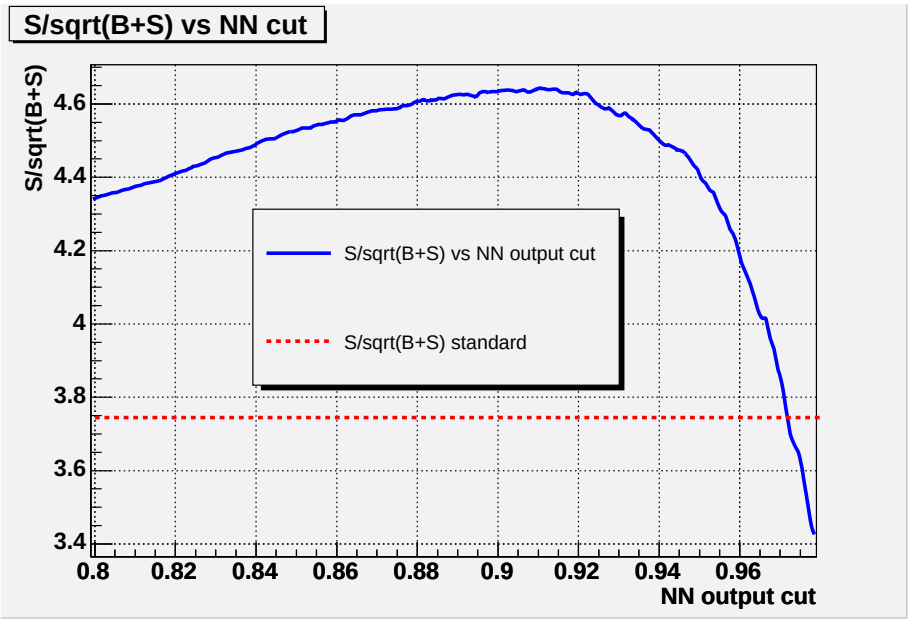


Figure 5.12: Statistical significance $S/\sqrt{B+S}$ vs neural network cut on N_{out}. Also shown as dashed red line is the value for the kinematical selection used in the previous analysis.

tistical uncertainty on the top quark mass measurement. Up to this point we have

Mass	***Efficiency*** ($N_{out} \geq 0.91$) ***(%)***
165 GeV/c^2	5.78 ± 0.05
170 GeV/c^2	6.38 ± 0.05
175 GeV/c^2	6.96 ± 0.01
180 GeV/c^2	7.26 ± 0.06

Table 5.2: Neural network selection efficiency calculated for different top quark masses, for the cut $N_{out} \geq 0.91$.

not exploited yet the presence of heavy flavour jets in the event. We already know from Chapter 2 that these jets can be identified (tagged) using the silicon vertex detector with a high efficiency. We thus want require at least one of the jet in the event to be tagged as a b jet.

5.5 Tagging Efficiency

The SecVtx tagging efficiency depends strongly on the cut to be chosen on N_{out} because events with large values of N_{out} are typically more energetic and more central in the detector, hence more taggable. So, we study the efficiency as a function of the neural network cut. Since the background modeling is based on tag characteristics and not on event features, we end up counting tags and not events; we thus prefer here to use the name of "average number of tags" (n_{ave}^{tag}) to indicate the average number of positive tags per event, as expected from a $t\bar{t}$ event. We use here the "counting" method: since for Monte Carlo events we tend to overestimate the number of tags, we count tagged jets after we randomly degrade their efficiency with the appropriate scale factor[3], i.e. we discard a fraction $1 - SF$ of the tagged jets.

Fig. 5.13 shows how n_{ave}^{tag} varies as a function of the cut on the neural network output. For reference we show also the tagging efficiency, ϵ_{tag}, (i.e. the probability for a $t\bar{t}$ event to have at least one positive tag).

5.6 The background modeling

The method used for the background estimate in the all-hadronic channel is based on the parametrization of the jet tagging probability, $\mathcal{P}$, as a function of the jet-E_T, the number of SVX tracks within the jet, N_{trk}, and the number of vertices, N_{vert}. The number of tags expected from background processes, N_{exp}, is then defined as:

$$N_{exp} = \sum_{i}^{N_{events}} \sum_{j=1}^{N_{taggable\ jets}} \mathcal{P}(E_T^j, N_{trk}^j, N_{vert})$$

This method has been proven to work quite well for multijet events (see [32]) predicting the background with an uncertainty of 4.5%. We apply here the same method after having reevaluated the tagging probability based on a sample of multijet events with exactly 4 jets and with $\Delta R_{min} \geq 0.5$. The agreement between observed (N_{obs})

[3]We use a scale factor of $SF_b = 0.89 \pm 0.07$ for b-jets, $SF_c = 0.89 \pm 0.14$ for c-jets and $SF = 1$ for the other jets. See [31].

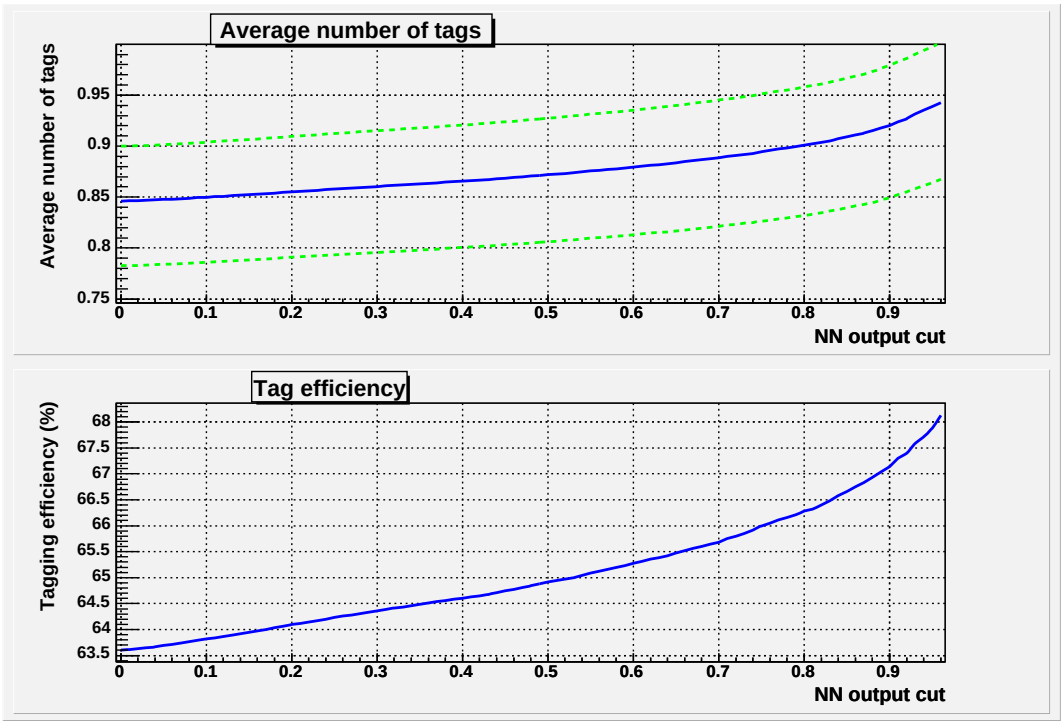


Figure 5.13: Top: Average number of tags, n_{ave}^{tag}, as a function of the cut on N_{out}; also shown is the variation if we vary SF by $\pm\delta SF$. Bottom: event tagging efficiency, ϵ_{tag}, as a function of the cut on N_{out}. For $t\bar{t}$ events with $6 \leq N_{jet} \leq 8$ and $\Delta R_{min} \geq 0.5$.

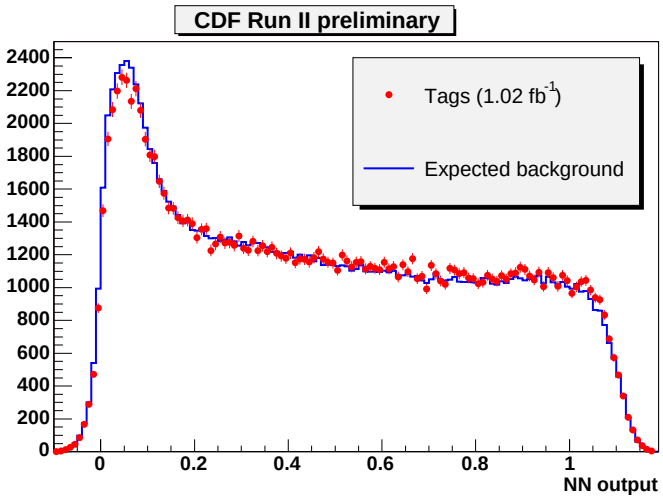


Figure 5.14: N_{out} distributions for tagged events and for the expected background. Events with 4 jets and $\Delta R_{min} \geq 0.5$.

and expected tags (which is by construction perfect in the 4 jet bin) is indeed good also for the 5 and more jets. At most we have a $-1.3 \pm 0.5\%$ difference for 6 or more jets (more observed tags than expected), but this is due to the presence of $t\bar{t}$ signal which becomes relevant for 5 and more jets bin[4]. Once accounted for this effect the difference becomes $0.1 \pm 0.7\%$.

Having established that this technique works quite well at different jet multiplicities we need to study possible biases due to the neural network selection. Figs. 5.14, 5.15, and 5.16 show the neural network output distribution for tagged events with 4[5], 5 and 6 or more jets (where we have multiple entries for multiple tags) compared to the background modeling expectation.

[4]We need to avoid double counting in the background, i.e. the fact that this method is applied also to $t\bar{t}$ events in the pretag sample. To account for this we subtract from the observed tags from $t\bar{t}$, the expectation coming from the application of this technique to $t\bar{t}$ events themselves. A more appropriate (but substantially similar) iterative method is used when one calculates the cross section.

[5]The N_{out} distribution for events with 4 jets is very peculiar in the sense that there is a big fraction of events around $N_{out} = 1$. This is due to the fact that, given the smaller number of jets, some variables like M_{2j}^{max} and M_{3j}^{max} are larger than what usually happens for events with 6 or more jets; hence these events are considered more $t\bar{t}$-like than regular background events with 6 or more jets.

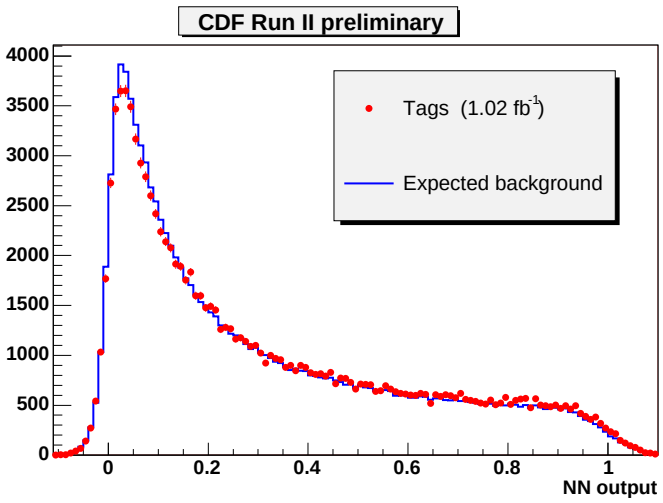


Figure 5.15: N_{out} distributions for tagged events and for the expected background. Events with 5 jets and $\Delta R_{min} \geq 0.5$.

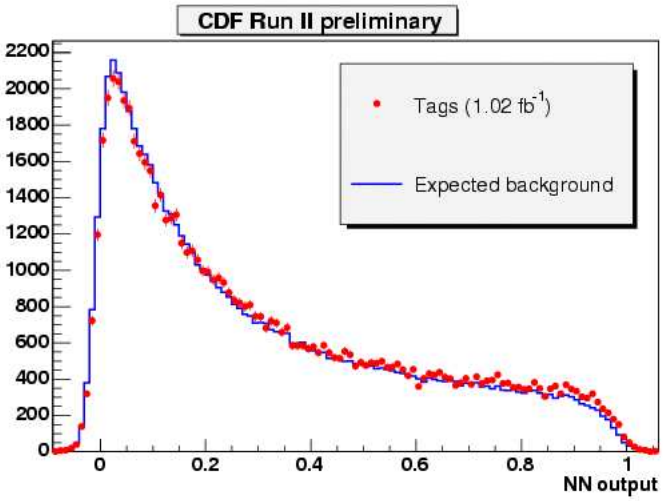


Figure 5.16: N_{out} distributions for tagged events and for the expected background. Events with 6 to 8 jets and $\Delta R_{min} \geq 0.5$.

Apart from a minor disagreement below 0.1 units, the background is well represented for values of the neural network output larger than 0.1 for all jet multiplicities.

To be more detailed and in order to evaluate the systematic uncertainty to assign to the background estimate we consider possible biases in the estimate due to:

1. the use of the tagging probability at different jet multiplicities;

2. the bias introduced by requiring a neural network output larger than a certain value.

The first issue is studied by considering the agreement between observed and expected tags in a "control region" defined as $0.1 \leq N_{out} < 0.8$[6]. In the "control region" the largest disagreement for different jet multiplicities is $-1.1 \pm 0.7\%$ for the 6 to 8 jets case.

We consider then second issue, i.e. the bias introduced by cutting on N_{out}, by looking at the "signal region" $N_{out} \geq 0.8$ for 4 jets. The background tends to be slightly underestimated in the "signal region" $N_{out} \geq 0.8$ and to be more quantitative we plot in Fig. 5.17 the relative difference $(N_{exp} - N_{obs})/N_{exp}$ as a function of the neural network cut. The relative difference in this region amount to an average of about 2.6% and is quite stable for different values of the N_{out} cut. If we account for the $t\bar{t}$ presence then this relative difference would be about 2.1%.

Considering now both effects (the use of a tagging matrix derived on 4 jet events to events with 6 or more jets, and the cut on neural network output in the vicinity of $N_{out} \geq 0.9$) and allowing for the presence of $t\bar{t}$ events, we assign to the background estimate a relative uncertainty of 2.5% (rounding off the quadrature sum 2.4% of the uncertainties for the two effects).

[6]The choice of $0.1 \leq N_{out} < 0.8$ might seem arbitrary because it is the region where the distributions agree the best. However if we were to consider just $N_{out} < 0.8$ then the disagreement between N_{exp} and N_{obs} would be even smaller: $(-0.4 \pm 0.6)\%$ for $6 \leq N_{jets} \leq 8$ instead of $(-2.0 \pm 0.6)\%$.

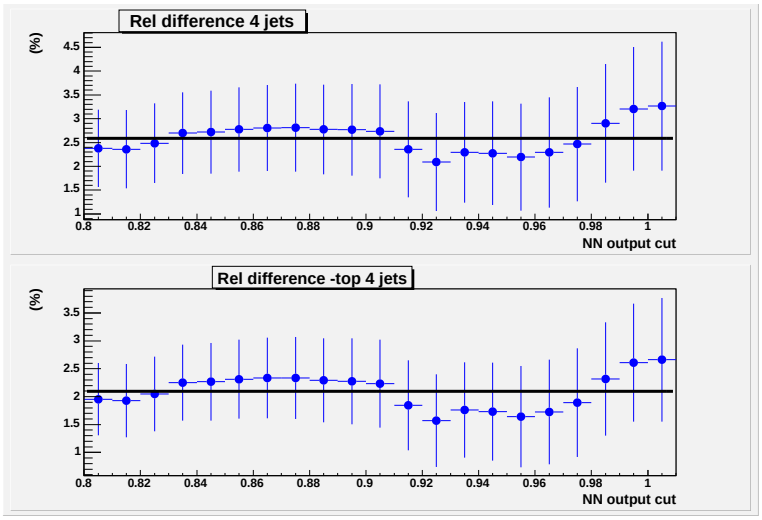


Figure 5.17: Top: Relative difference $(N_{exp} - N_{obs})/N_{exp}$ as a function of the cut on N_{out}, in the range of interest $0.8 \leq N_{out} \leq 1.0$. Bottom: $(N_{exp+t\bar{t}} - N_{obs})/N_{exp}$. Events with 4 jets and $\Delta R_{min} \geq 0.5$

Chapter 6

The mass measurement

In this chapter we present the technique which we will use for the measurement of the top quark mass. We use simulated events to find a variable which is closely correlated with the true top quark mass, and build a distribution of this variable, which we call template, that is characteristic of the signal and will be used to discriminate it from the background. The mass measurement is obtained fitting the data to a sum of signal and background templates; after that we run simulated experiments to estimate possible biases and the statistical accuracy of the measurement. All known sources of systematic uncertainties are studied and their effects on the mass measurement is discussed.

6.1 Kinematic fitter

The $t\bar{t}$ events under study in this work are characterized by the presence of 6 partons in the final states, and nominally 6 reconstructed jets in the detector; two of them are originated by b quarks so we expect to have tagged jets in the event. We want to fully reconstruct the event kinematics and exploit the presence of the W and top quark and their invariant masses to constrain the event topology. In order to do so, we use all events passing the kinematical selection described in the previous chapter (where the value for the N_{out} cut will be decided in the following) but consider only

the 6 leading jets (in E_T) in the event[1].

We have 16 equations connecting the four-momenta of the two top quarks and the six final state particles according to the $t\bar{t} \to b\bar{b}\,W^+W^- \to b\bar{b}\,q_1\bar{q}_2\,q_3\bar{q}_4$ hypothesis:

$$\begin{aligned} p_t^\mu &= p_{W^+}^\mu + p_b^\mu \\ p_{\bar{t}}^\mu &= p_{W^-}^\mu + p_{\bar{b}}^\mu \\ p_{W^-}^\mu &= p_{q_1}^\mu + p_{\bar{q}_2}^\mu \\ p_{W^+}^\mu &= p_{q_3}^\mu + p_{\bar{q}_4}^\mu \\ (\mu &= 0, 1, 2, 3) \end{aligned}$$

There are 13 unknown quantities, i.e. the three-momenta of the top quarks and of the W's, plus the unknown top quark mass, so the kinematics of the event is overconstrained. There are 90 possible permutations of jet-to-parton association with two jet doublets giving a W and of two jet triplets giving the top quarks. To reduce the number of permutations, only events with at least 1 tagged jet are used in this analysis, with the association of the b-tagged jet to a b quark; in this way the number of permutations is lowered to 30. We construct the χ^2 function:

$$\chi^2 = \frac{(m_{jj_1} - m_W)^2}{\Gamma_W^2} + \frac{(m_{jj_2} - m_W)^2}{\Gamma_W^2} + \frac{(m_{jjj_1} - m_t^{reco})^2}{\Gamma_t^2} +$$

$$+ \frac{(m_{jjj_2} - m_t^{reco})^2}{\Gamma_t^2} + \Sigma_{i=1}^{6} \frac{(p_{T,i}^{fit} - p_{T,i}^{meas})^2}{\sigma_i^2}$$

where $m_{jj_{1,2}}$ are the invariant masses of the dijet systems, $m_{jjj_{1,2}}$ are the invariant masses of the trijet systems, $\Gamma_W = 2.1$ GeV/c^2 is the measured natural width of the W [34], and Γ_t, fixed to 1.5 GeV/c^2, is the assumed natural width of the top quark. The measured jet transverse energies, $p_{T,i}^{meas}$, are free to vary within their known resolution, σ_i. The χ^2 is minimized with respect to the 7 free parameters, i.e.

[1]This is necessary in order to limit the number of ways in which we can combine the jets to reconstruct the events. In fact a high number of combinations will translate in a lower probability of choosing the correct association. On the other hand, less energetic jets are likely to be originated by gluon radiation.

the reconstructed top quark mass, m_t^{reco}, and the 6 jets transverse momenta, $p_{T,i}^{fit}$. Only the combination which gives the lowest χ^2 value is selected and the respective reconstructed top quark mass enters a mass distribution.

The reconstructed mass distribution is not expected to be a "real" top quark invariant mass distribution, in fact it will differ sensibly from the distribution we would get at the parton level: the lowest χ^2 choice will leave us with the correct jet-to-parton association only a relatively small fraction of the time[2]. Jet resolution will spread out the correct mass distribution, as well as give rise to incorrect associations. The possible presence of initial and final state radiation jets in the 6 leading jets would also populate the tails of the distribution. Nonetheless, the reconstructed mass distribution is strongly correlated with the input top quark mass and will be used as a template for the signal and the background.

6.1.1 Background validation

We already mentioned in Chapter 4 that our background is modeled using the data through a tagging probability parametrization. In order to obtain a data-driven background template we apply the kinematic fitter to the sample of events passing the neural network selection, but before the requirement of identified b-jets (pre-tag sample). Within an event the fitter is applied once for each fiducial jet, assuming it is a b-quark. The resulting invariant mass enters an histogram with a weight given by the tagging probability associated with the fiducial jet. This procedure does not allow a separation between the background expected for events with 1 or 2 tags, but treats them together. We follow the same approach for the data and the signal, so the fit is performed for each association of tagged jets with one or the other of the two b quarks; events with more than 1 tag contribute then with multiple entries. In order to check how properly this background modeling describes the multijet data, we subdivide the data in 4 non-overlapping control regions with respect to the neural network output N_{out}, in a range where the signal presence is very low (at or below

[2]This fraction has been measured in a subset of Monte Carlo events with exactly 6 jets, and associated in direction with the partons, to be about 40%.

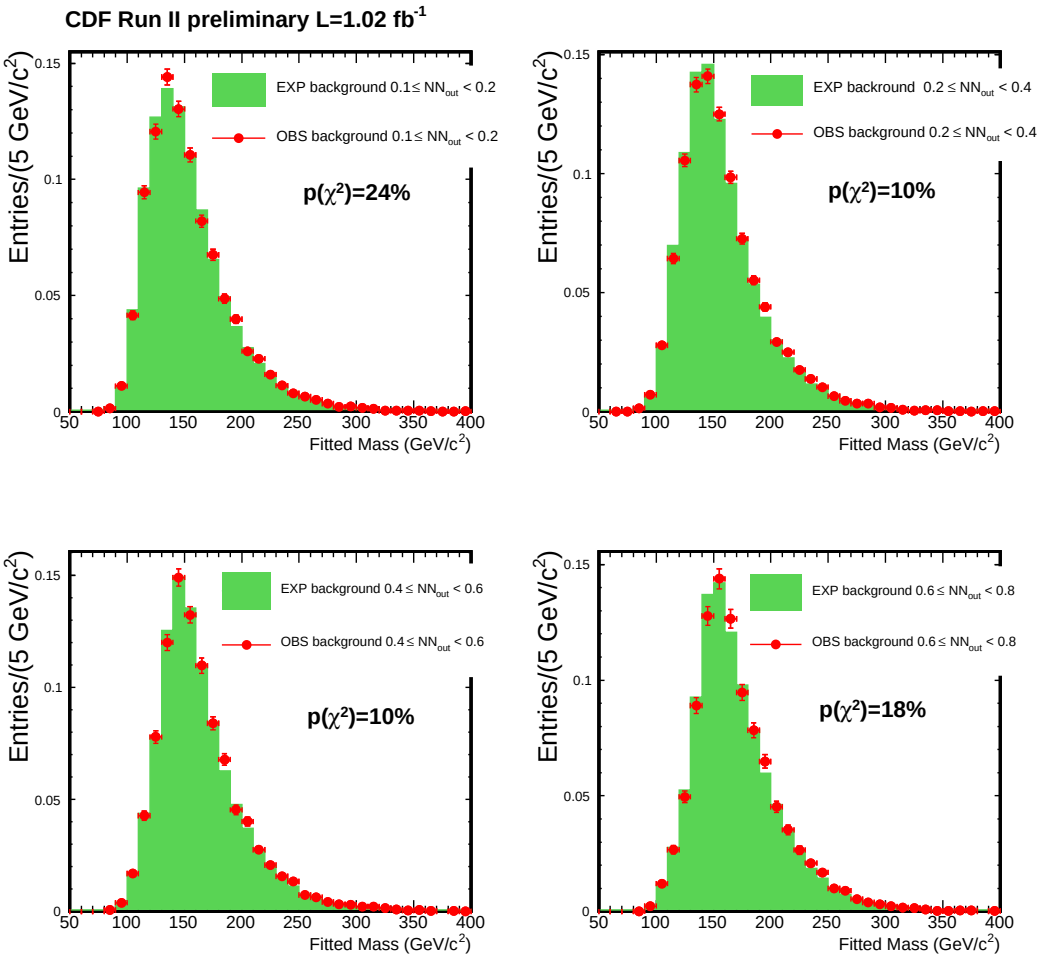


Figure 6.1: Comparison of the reconstructed top quark mass for the multijet tagged events (red dots) and the expectation from the background modeling (green histogram) in 4 non-overlapping control samples.

the % level).

- $0.1 \leq N_{out} < 0.2$;
- $0.2 \leq N_{out} < 0.4$;
- $0.4 \leq N_{out} < 0.6$;
- $0.6 \leq N_{out} < 0.8$.

For each of these subsamples, we plot the quantity which is essential to the measurement, i.e. the m_t^{reco}. We can see in Fig. 6.1 that the agreement is good over the

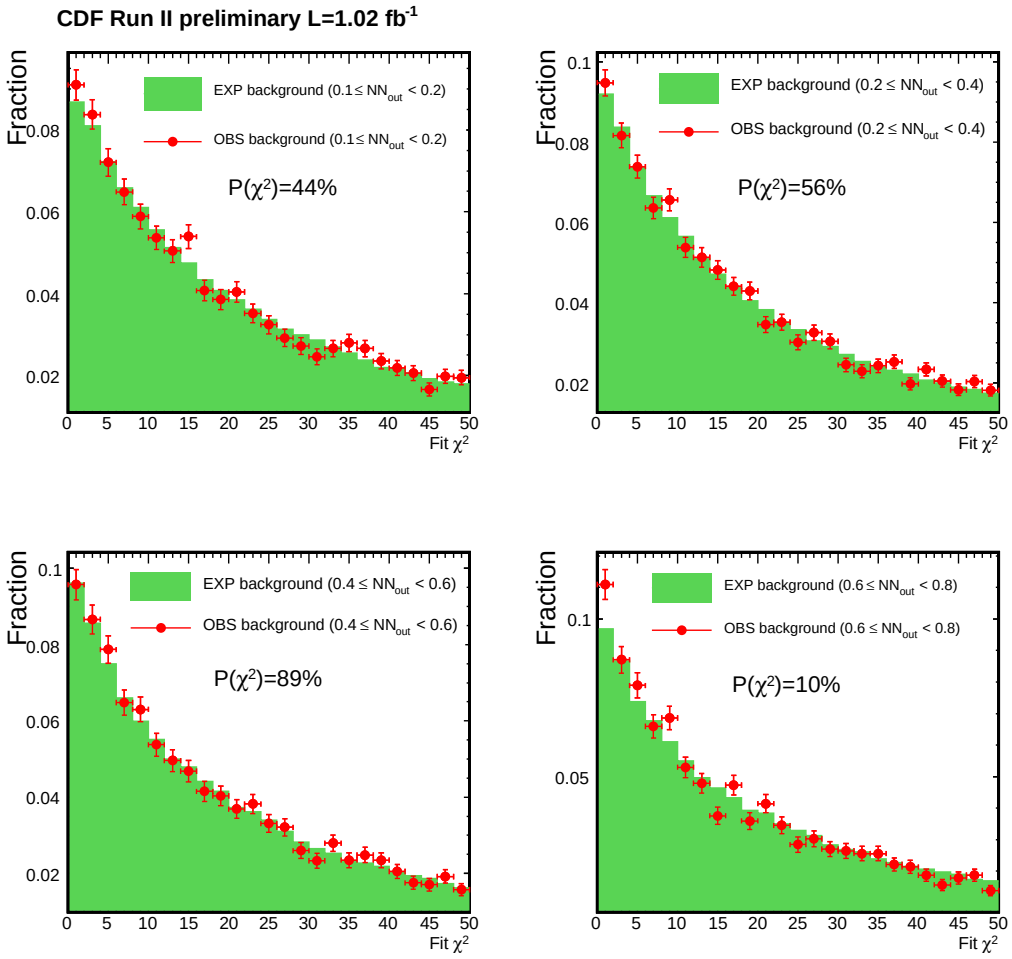
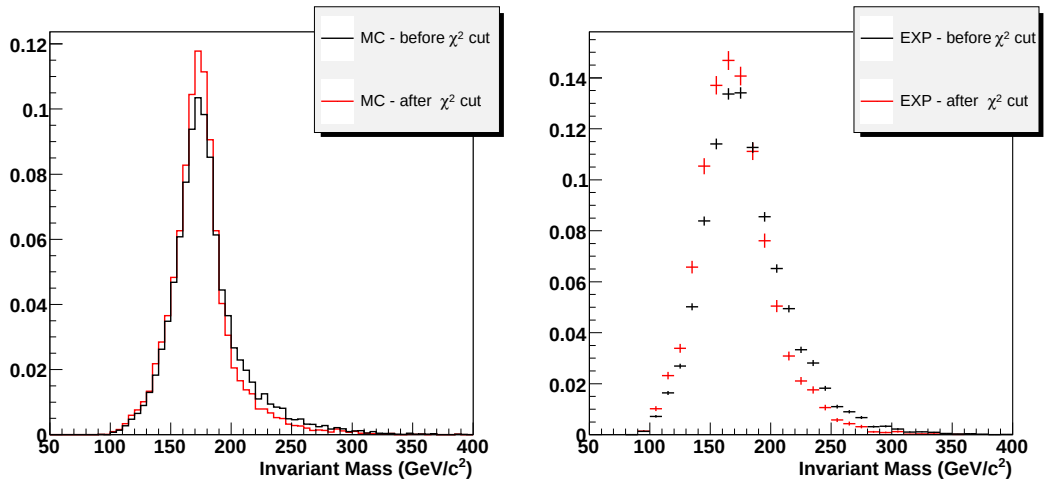


Figure 6.2: Comparison of χ^2 distribution for the multijet tagged events (red dots) and the expectation from the background modeling (green histogram) in 4 non-overlapping control samples.



(a) (b)

Figure 6.3: Effect of the χ^2 cut on the m_t^{reco} (a) $t\bar{t}$ distribution and for (b) background events. We see that as we cut on χ^2 the signal distribution gets narrower, while the background shifts towards lower mass values.

whole N_{out} range.

Additionally, we want to validate the χ^2 distribution associated with the mass; as for the reconstructed mass distribution, we show a comparison between the pre-tag sample derived background and data in Fig. 6.2. Again, the agreement is good over the whole N_{out} range.

Having shown that the background models well the χ^2 distribution in the data in the control region, we are able to introduce a cut on the χ^2 value. As an effect, we expect to increase our S/B ratio, and to reduce the number of uncorrectly associated jets in our kinematical fitter. In Fig. 6.3 we show the effect of the χ^2 cut on the distribution of signal and background templates. We see that indeed the signal distribution becomes narrower, as expected since we are choosing events which adhere more to the hypothesis of a $t\bar{t}$ events decaying hadronically. The background shape is shifted slightly towards lower mass values, another feature which will increase the statistical precision of this technique. We will study this effect in more detail in Sec. 6.3.1.

6.2 Likelihood fit

The technique described above gives a distribution of reconstructed masses (one for each tag in the event) in the corresponding data sample, which is a mixture of signal and background. In order to measure the top quark mass, we compare the m_t^{reco} distribution from the data to the signal and background templates. From the templates we first derive probability density functions (p.d.f.'s) and then perform an unbinned likelihood fit to determine the value of true top quark mass, $M_{\rm top}$, that best describes the data.

6.2.1 The Likelihood Function

The p.d.f.'s distributions are parametrized in order to have a functional form which varies smoothly with respect to $M_{\rm top}$. For the signal, we use $t\bar{t}$ Monte Carlo events generated with HERWIG v 6.508 [24] with top quark masses ranging from 150 to 200 GeV/c^2 in 2.5 GeV/c^2 increments. The function for the signal p.d.f., $P_{sig}(m|M_{\rm top})$, represents the probability to obtain a value m for m_t^{reco}, given a true top quark mass M_{top} in a $t\bar{t}$ event. The form used is a sum of a Gamma distribution, choosen to describe the invariant mass due to incorrect jet-parton assignments, and two Gaussian distributions, which model the core of the distribution. Its explicit expression is:

$$P_{sig}(m|M_{\rm top}) = \delta_7 \cdot \frac{\delta_2^{1+\delta_1}}{\Gamma(1+\delta_1)} \cdot (m-\delta_0)^{\delta_1} \cdot e^{-\delta_2(m-\delta_0)} +$$

$$+\delta_8 \cdot \frac{1}{\sqrt{2\pi}\delta_4} \cdot e^{\frac{-(m-\delta_3)^2}{2\delta_4^2}} +$$

$$+(1-\delta_7-\delta_8) \cdot \frac{1}{\sqrt{2\pi}\delta_6} \cdot e^{\frac{-(m-\delta_5)^2}{2\delta_6^2}}$$

where each parameter δ_i is linearly dependent on $M_{\rm top}$

$$\delta_i = \alpha_i + \beta_i \cdot (M_{\rm top} - 175), \qquad i = 0, 1, ...8;$$

so the total number of parameters used is 18. Fig. 6.4. shows some of the signal

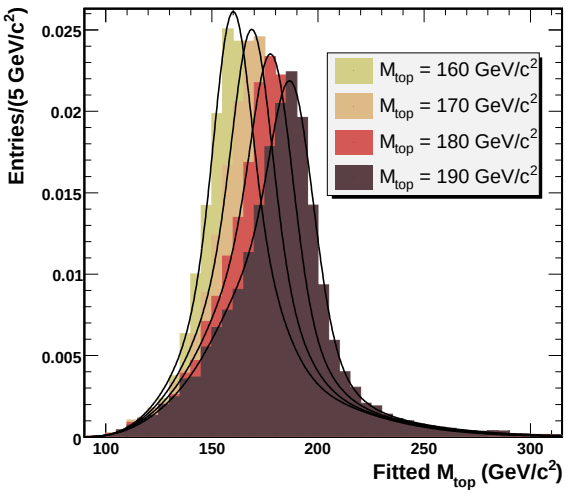


Figure 6.4: Signal m_t^{reco} templates with their p.d.f.'s superimposed. The plot is obtained from the Herwig $t\bar{t}$ samples after cutting at $N_{out} \geq 0.91$ and $\chi^2 \leq 16$.

templates along with the p.d.f.'s superimposed.

The background reconstructed top quark mass is computed as described in the previous section. The integral of this distribution is the predicted amount of background, which needs to be corrected for the presence of signal with the same iterative technique described in the previous chapter. The correction needs to be implemented also in terms of shape, so we subtract from the background mass distribution the distribution expected for the signal. The systematic uncertainty associated to this procedure is estimated and reported in the dedicated section. The background template is parametrized with two Gamma distributions and one Gaussian distribution. Clearly, there is no need here for a dependence of the parameters on the top quark mass. The resulting background p.d.f., $P_{bkd}(m)$, is as follows:

$$P_{bkd}(m) = \delta_8 \cdot \frac{\delta_2^{1+\delta_1}}{\Gamma(1+\delta_1)} \cdot (m-\delta_0)^{\delta_1} \cdot e^{-\delta_2(m-\delta_0)} +$$

$$+\delta_9 \cdot \frac{\delta_5^{1+\delta_4}}{\Gamma(1+\delta_4)} \cdot (m-\delta_3)^{\delta_4} \cdot e^{-\delta_5(m-\delta_3)} +$$

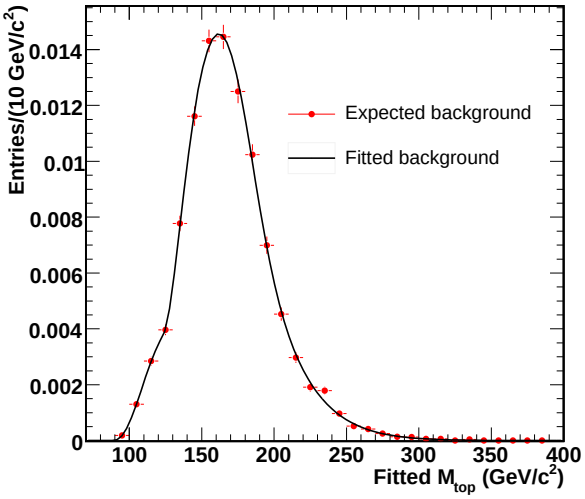


Figure 6.5: Background m_t^{reco} template with its p.d.f superimposed. The plot is obtained from the pre-tag data sample after cutting at $N_{out} \geq 0.91$ and $\chi^2 \leq 16$.

$$+(1-\delta_8-\delta_9)\cdot\frac{1}{\sqrt{2\pi}\delta_7}e^{\frac{-(m-\delta_6)^2}{2\delta_7^2}}$$

and is plotted in Fig. 6.5.

We introduce now a likelihood function which gives the probability that our data are described by an admixture of background and all-hadronic $t\bar{t}$ events with a certain top quark mass. The function is divided in two parts. The former constrains the number of background tags, n_b, to the expectation, n_b^{exp}, and the number of signal and background tags, $n_s + n_b$, to be equal to the number observed in the data, N. In the latter the signal and background probabilities are assigned by comparing the reconstructed top quark mass values m_i measured in the data with the parametrized signal and background p.d.f.'s P_{sig} and P_{bck}. Namely, our likelihood is defined as:

$$\mathcal{L} = e^{-\frac{(n_s+n_b-N)^2}{2\sigma_N^2}} \times e^{-\frac{(n_b-n_b^{exp})^2}{2\sigma_{n_b}^2}} \times$$

$$\times \prod_{i=1}^{N} \frac{n_s \cdot P_{sig}(m_i|M_{\rm top}) + n_b \cdot P_{bkd}(m_i)}{n_s + n_b}$$

where σ_{n_b} is the uncertainty expected on the corrected background after the iterative

correction described above, and $\sigma_N = \sqrt{N}$ is the uncertainty expected on the total number of observed tags, N. In order to facilitate the computation, we minimize the negative logarithm of the likelihood, $-ln\mathcal{L}$, instead of maximizing the likelihood itself. The minimization is performed using MINUIT [35] with respect to the 3 free parameters n_s, n_b and $M_{\rm top}$. The statistical uncertainty on the top quark mass is taken from the points where the $-ln\mathcal{L}$ changes by 0.5 units from its maximum. Correlations between the parameters are taken into account.

6.3 Pseudo-experiments procedure

We want to quantify the statistical power of the method, as well as investigate the possible presence of biases in the top quark mass measurement before performing the actual measurement on the data sample. To study both effects, we use a very useful tool: the pseudo-experiments. The procedure is described in the following: for each of the 21 input top quark masses, ranging from 150 GeV/c^2 to 200 GeV/c^2, we fix the total amount of tags to be the same as observed in the data. Then we take our background expectation of tags passing the cuts, fluctuate it according to Gaussian statistics, and get the number of signal tags as the difference between N and the fluctuated background. Finally we extract a pseudo-data mass distribution from the respective templates with the predicted amount. In such a way we can study the technique performances in more detail before looking at the real data.

6.3.1 Optimization of the kinematical selection

The neural network kinematical selection for $t\bar{t}$ all-hadronic events has been developed as a generic tool [36, 37] for the analysis in this channel. We are now left with the task of deciding what is the best cut on N_{out} for this specific analysis. We have already shown that cutting at $N_{out} \geq 0.91$ provides the highest statistical significance, $S/\sqrt{S+B}$. Cutting harder would increase our signal over background ratio but decrease our efficiency. The goal of our neural network cut selection is to provide the best mass measurement, where the assumption is that the systematic uncertainties will not differ sensibly in amount in the region $N_{out} \approx 0.9$.

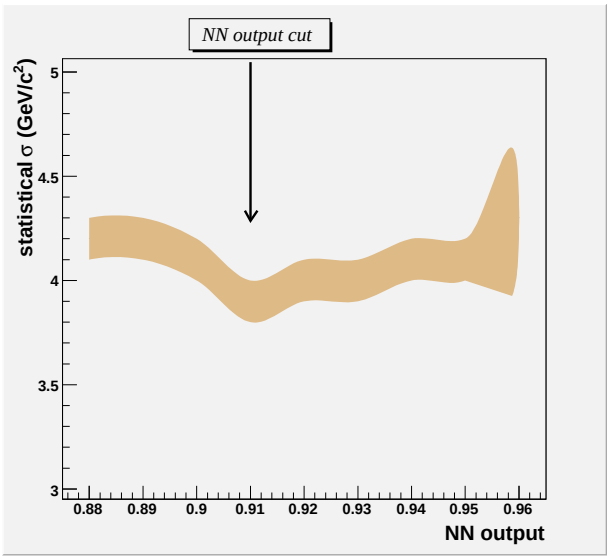


Figure 6.6: Statistical uncertainty on the top quark mass measurement as a function of the neural network output as calculated through pseudo-experiments. The integrated luminosity used in this optimization amounts to $800\,\mathrm{pb}^{-1}$.

This optimization has been done using as input top quark mass $M_{top} = 175\,\mathrm{GeV/c^2}$, as signal tags the number extrapolated from a theoretical cross section of 6.7 pb and the calculated efficiency, and as the background normalization the one we get from the data[3].

Ww build signal and background templates for N_{out} cuts from 0.88 to 0.96, in steps of 0.01. We use as input pseudo-data drawn from our templates, which are redone for each cut choice. The result is shown in Fig. 6.6, where we see that the cut at 0.91 is the optimal one, giving the lowest expected statistical uncertainty.

As we did for the N_{out} variable, we scan the χ^2 region to look for statistical uncertainty reduction. We start cutting at $\chi^2 \leq 30$ and go down to $\chi^2 \leq 6$, in steps of two units. For each step we reevaluate all signal and the background templates, then we proceed to evaluate the statistical uncertainty using pseudo-experiments. The result is shown in Fig. 6.7, where we plot the symmetrized statistical uncertainty as a function of the χ^2 cut. Errors are pretty large, due mostly to the limited amount

[3]The optimization has been performed on a subset of $800\,\mathrm{pb}^{-1}$, what was available at that time, but the same behavior has been observed in the whole dataset.

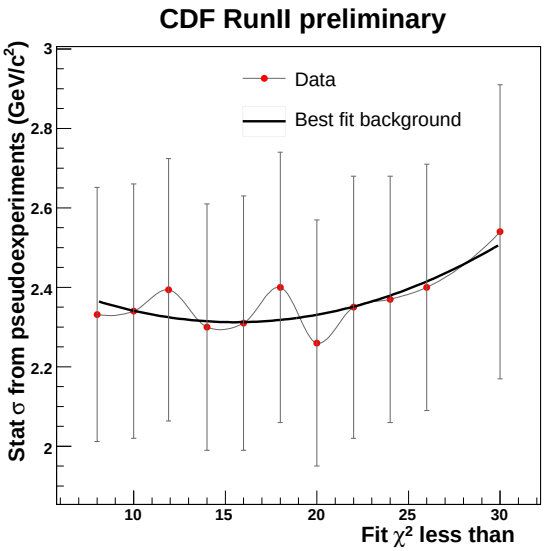


Figure 6.7: Statistical uncertainty on the top quark mass measurement as a function of the χ^2 cut as calculated through pseudo-experiments. The integrated luminosity used in this optimization amounts to $1.02\,\mathrm{fb}^{-1}$. Overlaid is a parabolic fit of the points, whose minimum corresponds to $\chi^2 = 16$.

of pre-tag data used to construct our background. We interpolate the points with a cubic fit to find a minimum corresponding to $\chi^2 = 16$; this is the cut we will use from now on.

6.3.2 Validation of the method

We perform the measurement on many different sets of pseudo-events, and plot the fitted top quark mass with respect to the input mass in Fig. 6.8, seeing no systematic bias. The deviation of the fitted mass from the input value, divided by the fitted measurement uncertainty is called "pull". The pull distribution is fitted with a Gaussian and its width ("pull width") indicates how accurate are the uncertainties obtained from the fit. We see from Figs. 6.8 and 6.9 that there is no need to add any correction for residual mismeasurements.

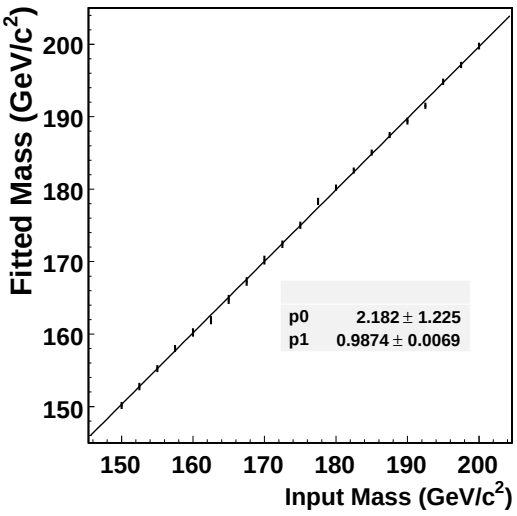
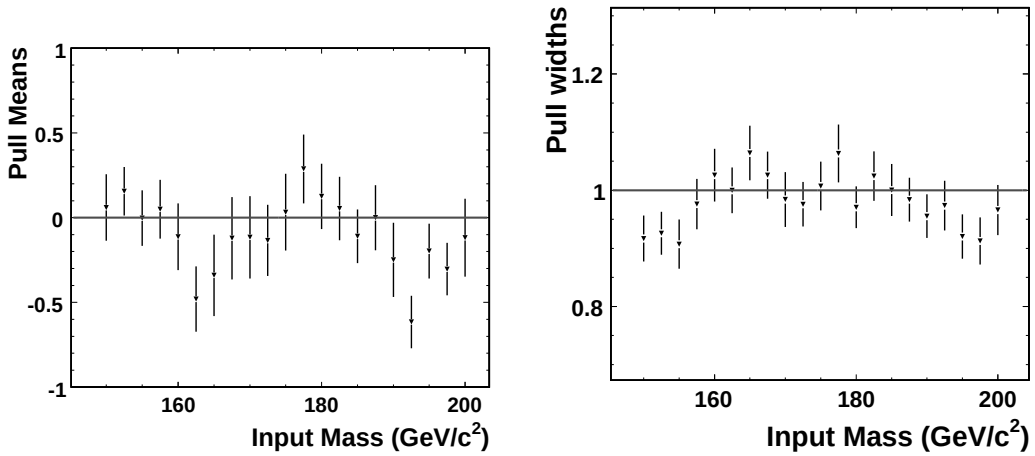


Figure 6.8: Fitted mass as a function of the input top quark mass using pseudo-experiments. The fit slope is consistent with 1.



(a) (b)

Figure 6.9: Mean (a) and width (b) of the pull distributions, as a function of the top quark input masses. We see that the pull means are compatible with 0, showing no bias, and that the pull widths are in good agreement with the $y = 1$ line, showing that the statistical uncertainty estimate is indeed accurate.

6.4 Systematic uncertainties

Various sources of systematic uncertainty affect the mass measurement, and they are presented and their effect studied in this section. Systematic effects arise from mismodeling in the simulation of the detector response to jets, and from uncertainties in the simulation of the $t\bar{t}$ signal. The evaluation of the mass shift due to each source of systematic uncertainty consists in varying the effects from each possible source by $\pm 1\sigma$, building invariant mass templates out of the varied sample, drawing from that template the expected amount of pseudo-events and finally compare it with the signal and background unchanged templates. We take as estimate of the respective uncertainty the resulting half-difference between the extreme values of the measured mass. If the shift is smaller than the statistical uncertainty on the estimate itself, we assign the latter as systematic uncertainty.

The main contribution to the systematic uncertainty stems from the residual uncertainty on the jet energy after it is corrected for the known effects [20]. These include calorimeter non-linearity in response to single hadronic particles, energy loss due to non-instrumented regions, additional energy coming from secondary interactions and from the underlying event. We calculate the systematic uncertainty originating from each of these sources varying each corrected jet energy in the simulation by the corresponding uncertainty, performing pseudo-experiments with the modified resulting templates, and finally add them in quadrature to quote a total systematic due to the jet energy scale. Since the jet energy corrections are derived on data samples deprived of heavy quarks, an additional uncertainty comes from considering the different fragmentation properties of b quarks [38, 39]. As for the generic jet energy scale uncertainties, we perform pseudo-experiments where we vary the b-jet energy scale by $\pm 1\sigma$ and pick the half-difference as the b-jet energy scale uncertainty. Many sources of systematic effects arise from uncertainty in the Monte Carlo modeling of the hard interaction. PYTHIA and HERWIG generators differ in their hadronization schemes and in their description of the underlying event and multiple interactions. A corresponding systematic uncertainty is evaluated drawing pseudo-events from PYTHIA-generated events and comparing the resulting mass distributions with the template constructed using HERWIG. Additional jets coming from initial and final

state radiation (ISR and FSR) might fall among the six leading jets and populate the tails in the top quark invariant mass distribution. These effects are studied building pseudo-experiments where we extract invariant masses from templates generated with different values of Λ_{QCD} and K factor (see [40]), and comparing them with the standard templates. Since the shift is very small, we assign as ISR/FSR systematic uncertainty the statistical uncertainty on the shift. The choice of parton distribution functions (PDF) inside the proton can affect the kinematics of $t\bar{t}$ events and thus the top quark mass measurement. We estimate the uncertainty resulting from the use of Monte Carlo samples based on the default CTEQ5L [41] PDF and the one calculated from the MRST group [42], MRST72 and MRST75, which differ by the value of Λ_{QCD} used to compute the PDF. The background normalization is known within 5% error due to the tagging probability parametrization technique. We vary the background contribution using pseudo-experiments where we increase or decrease the expected background amount by 1σ. We consider also the uncertainty associated with the small presence of signal in the data-driven background. To do so we build two background templates, where we subtract from the background mass distribution the expected signal mass distribution assuming the two values 172.5 GeV/c^2 and 177.5 GeV/c^2 for the top quark mass. A systematic uncertainty due to the finite template statistics is determined by Poisson fluctuating the templates 100 times and using the fluctuated, reparametrized templates to run pseudo-experiments. The width of the resulting mass distribution is used as systematic uncertainty. A bias in the measurement can arise if an inadequate functional form is used for the invariant mass templates. This is checked performing pseudo-experiments where we extract pseudo-events directly from the top quark invariant mass histograms and compare them with the parametrized p.d.f.'s. The average of the difference between the measured top quark masses and the input masses is chosen as a systematic uncertainty on the functional parametrization. The b-tagging efficiency agrees well between data and simulation; still a possible dependence on jet kinematic properties could lead to a shift in the measured mass. We evaluate here a systematic due to E_T dependence of b-tagging scale factor by varying this factor by $\pm 1\sigma$ with respect to the corresponding slope in E_T and reobtaining the signal template from pseudo-experiments. Since the background estimate is data-driven,

the analysis is sensitive to an overall uncertainty in the b-tagging scale factor only through signal shapes.

In Table 6.1 is shown a summary of all the systematic uncertainties and the total systematic uncertainty, which amounts to 4.8 GeV/c^2.

Source	***Systematics (GeV/c^2)***
Jet energy scale	4.5
Generator	1.0
b-jet energy scale	0.5
Parton Distribution Function	0.5
Background shape	0.5
Background fraction	0.5
ISR	0.5
FSR	0.5
b-tag	0.5
MC statistics	0.1
Template parametrization	0.1
Total	**4.8**

Table 6.1: Breakdown of systematic uncertainties from different sources and their respective and total amount.

6.5 The top quark mass measurement

After the kinematic selection with $N_{out} \geq 0.91$, the b-tagging requirement and the cut on the goodness of the fit, $\chi^2 \leq 16$, we find 926 tags in 772 events. The background, corrected for the contribution due to $t\bar{t}$ events (see Section 6) amounts to 567 ± 28 tags.

The likelihood procedure is applied to the data sample to derive $M_{\rm top} = 174.0 \pm 2.2(stat.) \pm 4.8(syst.)$ GeV/c^2, value which maximizes the likelihood shown in Fig. 6.10. The same plot in Fig. 6.10 shows the reconstructed top quark mass distribution for the data compared to the expected background and the signal for a top quark mass of 174.0 GeV/c^2. The plot in Fig. 6.11 compares the measured statistical uncertainty with the expected distribution from pseudo-experiments using as input mass $M_{\rm top} = 174.0$ GeV/c^2. We find that the probability of achieving a better sensitivity

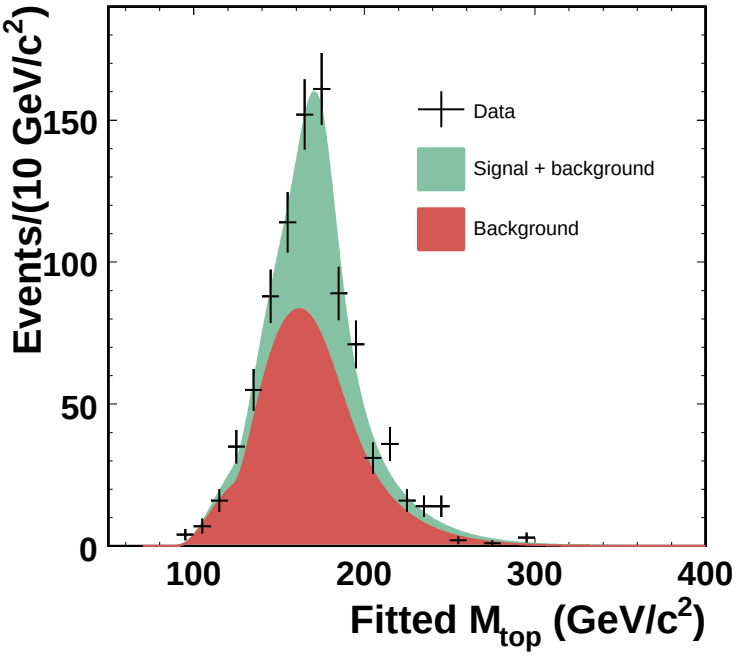


Figure 6.10: The reconstructed top quark mass distribution for events with $N_{out} \geq 0.91$, $\chi^2 \leq 16$ and at least 1 b-tagged jet. Superimposed are the background and the $t\bar{t}$ signal expected for $M_{\rm top} = 174.0$ GeV/c^2.

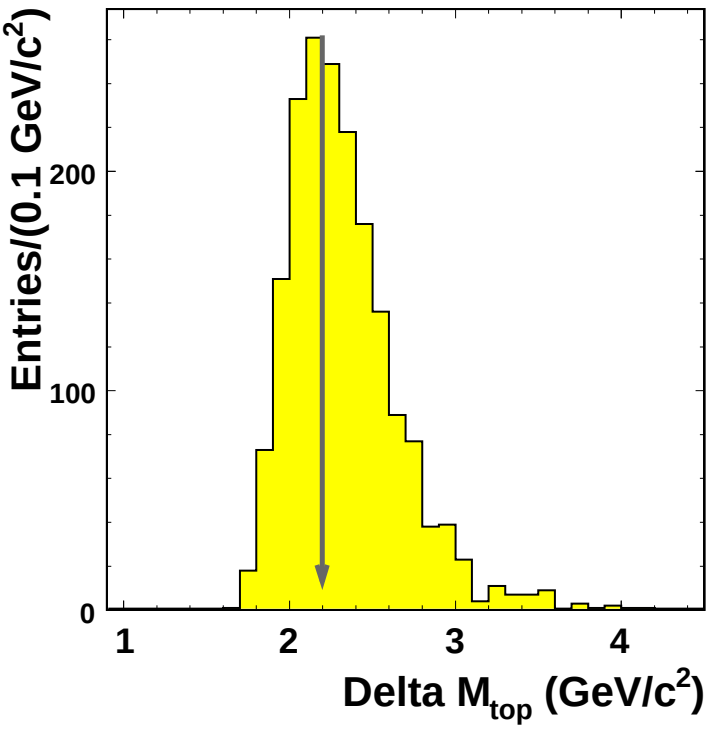


Figure 6.11: On the top-right corner is shown the negative log-likelihood, where the scale is changed to have the minimum equal to zero. The expected statistical uncertainty using pseudo-experiments with input mass equal to 174.0 GeV/c^2 for the $t\bar{t}$ signal. The arrow represents the measured statistical uncertainty.

is 40%. As a last check, we perform the measurement removing in the likelihood definition the Gaussian term which constrains the number of background tags to be the one predicted via the tagging probability parametrization, and obtain nearly the same value, $M_{\rm top} = 174.1 \pm 2.2(stat.)$ GeV/c^2.

At the moment of the presentation of this measurement [43] at the CDF collaboration and then at the Summer 2006 conferences, this was the Tevatron most precise measurement of the top quark mass in the all-hadronic decay channel. It was also the second best measurement performed by the CDF collaboration. The result is consistent with the measurements obtained in the same channel in Run I [44] and with the measurement obtained in Run II with $L = 311$ pb^{-1} of data with a different technique [45], as well as with the results obtained in the lepton+jets channels [46]

and dilepton channel [47] by CDF with $\mathcal{L} \simeq 1.0\,\mathrm{fb}^{-1}$. The measurements performed at CDF II in Run I and in Run II, together with the measurement presented in this work, are used to perform the Summer 2006 combination [51] which yields the top quark mass of $170.9 \pm 1.4(stat.) \pm 1.9(syst.)\,\mathrm{GeV/c^2}$ (see Fig. 6.12). The measurement has also been combined [52] with the most precise D0 measurements in the data samples collected in Run I and in Run II to obtain the Summer 2006 world average on top quark mass of $171.4 \pm 2.1\,\mathrm{GeV/c^2}$ (see Fig. 6.13).

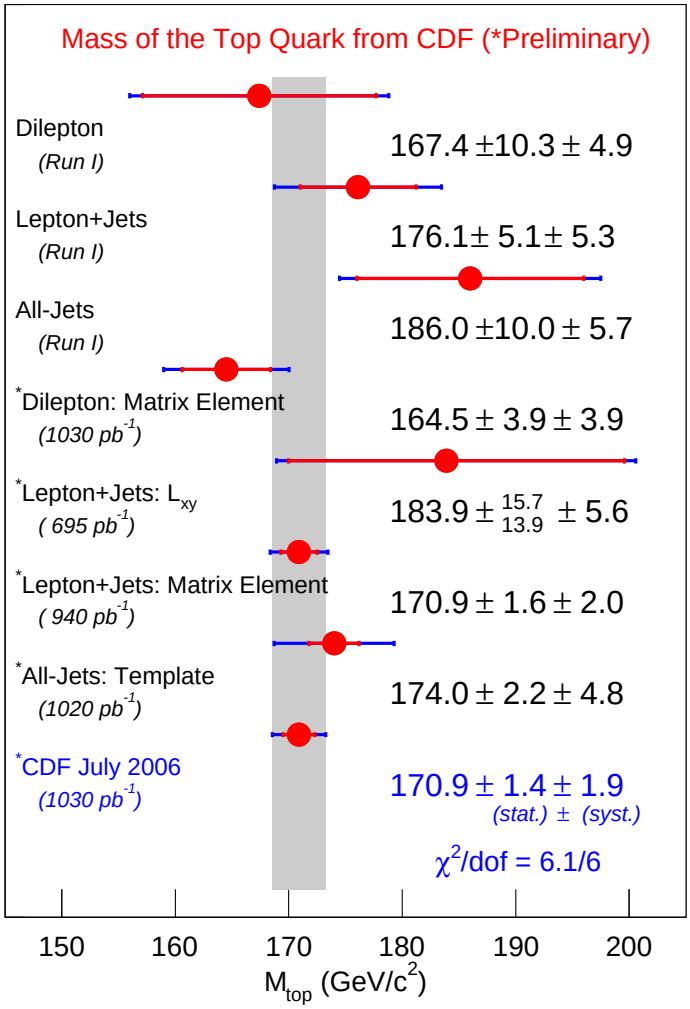


Figure 6.12: CDF average of the most precise measurements in each channel in Run I and in Run II. The measurement labeled "L_{xy}" is added since its uncertainties are mostly uncorrelated with the other measurements.

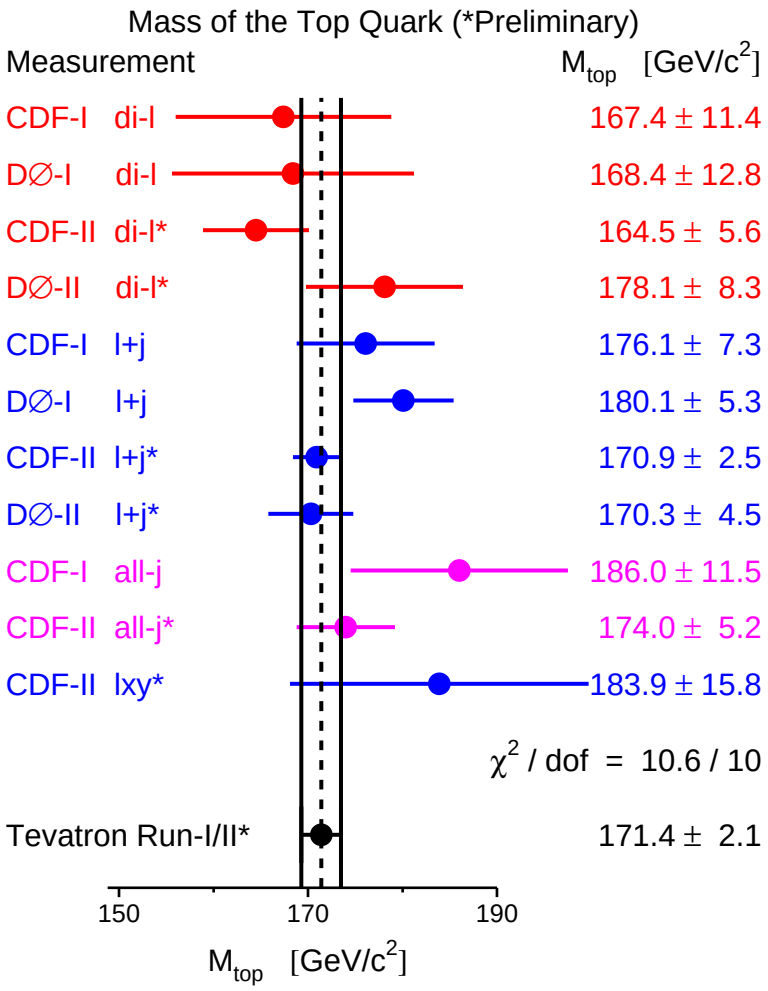


Figure 6.13: Combination of the most precise measurements performed by the CDF and D0 collaborations in each channel in Run I and in Run II. The CDF measurement labeled "L_{xy}" is added since its uncertainties are mostly uncorrelated with the other measurements.

Conclusions

We present here a measurement of the top quark mass in the all-hadronic channel. Using an optimized neural network based kinematic selection and a b-jet identification technique, we are able to improve the S/B of the initial multijet sample obtained with a dedicated trigger from 1/1100 to $\approx$ 1/2. With the selected sample we meaure the top quark invariant mass for the overconstrained kinematic system and compare it to templates representing the signal and background. After evaluating the systematic uncertainties which affect the measurement, we finally apply the technique to the the data to find 773 events in which we measure $M_{top} = 174.0 \pm 2.2(stat.) \pm 4.8(syst.)\,\mathrm{GeV/c^2}$. This measurement has been combined with the same measurements in the leptonic channels from CDF and D0 experiments to obtain the Summer 2006 world average top quark mass.

Bibliography

[1] F. Abe *et al.* [CDF Collaboration], Phys. Rev. Lett. **74** (1995) 2626 [arXiv:hep-ex/9503002].

[2] S. Abachi *et al.* [D0 Collaboration], Phys. Rev. Lett. **74**, 2632 (1995) [arXiv:hep-ex/9503003].

[3] E. Laenen, J. Smith and W. L. van Neerven, Phys. Lett. B **321**, 254 (1994) [arXiv:hep-ph/9310233].

[4] M. Cacciari *et al.*, JHEP 0404:068 (2004); N. Kidonakis and R. Vogt, Phys. Rev. D **68**, 114014 (2003)

[5] T. L. E. Group [OPAL Collaboration], arXiv:hep-ex/0612034.

[6] E. Brubaker *et al.* [Tevatron Electroweak Working Group], arXiv:hep-ex/0608032.

[7] R. Barate *et al.* [LEP Working Group for Higgs boson searches], Phys. Lett. B **565**, 61 (2003) [arXiv:hep-ex/0306033].

[8] C. Hays, A. Kotwal, L. Nodulman, O. Stelzer-Chilton, W. Trischuk, I. Vollrath CDF/PUB/ELECTROWEAK/PUBLIC/8665

[9] L. Balka *et al.* [CDF Collaboration], Nucl. Instrum. Meth. A **267**, 272 (1988).

[10] S. Bertolucci *et al.* [CDF Collaboration], Nucl. Instrum. Meth. A **267** (1988) 301.

[11] S. Aota *et al.*, Nucl. Instrum. Meth. A **352**, 557 (1995).

[12] G. Apollinari, K. Goulianos, P. Melese and M. Lindgren, Nucl. Instrum. Meth. A **412** (1998) 515.

[13] J. Elias *et al.*, Nucl. Instrum. Meth. A **441**, 366 (2000).

[14] S. Cabrera *et al.* [CDF Collaboration], Nucl. Instrum. Meth. A **494**, 416 (2002).

[15] S. Klimenko, J. Konigsberg and T. M. Liss, *"Averaging of the inelastic cross sections measured by the CDF and the E811 experiments"*, FERMILAB-FN-0741, December 2003.

[16] H. Ray, *"Level 2 and the L1/L2 Trigger System for Non-Experts"*, CDF internal note 5860, December 2002.

[17] A. Bhatti *et. al.*, *"Proposal for Level-2 Calorimeter Trigger Upgrade"*, CDF internal note 8415, July 2006.

[18] M. Albrow *et. al.*, *"CDF Run-II Trigger Table and Datasets Plan"*, CDF internal note 4718, December 2001.

[19] F. Abe *et al.* [CDF Collaboration], Phys. Rev. D **45**, 1448 (1992).

[20] A. Bhatti *et al.*, Nucl. Instrum. Meth. A **566**, 375 (2006) [arXiv:hep-ex/0510047].

[21] T. Muller *et al.*, *"Vxprim in Run II"*, CDF internal note 6047, July 2002.

[22] D. Acosta *et al.* [CDF Collaboration], Phys. Rev. D **71**, 052003 (2005) [arXiv:hep-ex/0410041].

[23] T. Sjostrand *et al.*, Comput. Phys. Commun. **135**, 238 (2001).

[24] G. Marchesini *et al.*, Comput. Phys. Commun. **67**, 465 (1992); G. Corcella *et al.*, J. High Energy Phys. **0101**, 010 (2001).

[25] BOH

[26] A. Abulencia *et al.* [CDF Collaboration], Phys. Rev. Lett. **96**, 202002 (2006) [arXiv:hep-ex/0603043].

[27] A. Abulencia *et al.* [CDF Collaboration], Phys. Rev. D **74**, 072005 (2006) [arXiv:hep-ex/0607095].

[28] MLPFIT: A Tool for Multi-Layer Perceptrons, http://schwind.home.cern.ch/schwind/MLPfit.html.

[29] R. Brun and F. Rademaker, ROOT Version 4.00/08, http://root.cern.ch/.

[30] S. McLoone *et al.*, IEEE Trans. Neural Networks **9** (1998) 669.

[31] See http://www-cdf.fnal.gov/internal/physics/top/RunIIBtag/wi06.html.

[32] P. Azzi, A. Castro, A. Gresele, J. Konigsberg, G. Lungu, A. Sukhanov, *"B-tagging efficiency and background estimate in the Run II multijet dataset (with 311 pb^{-1})"*, CDF internal note 7723, July 2005.

[33] A. Abulencia *et al.* [CDF Collaboration], Phys. Rev. D **73**, 032003 (2006).

[34] S. Eidelman *et al.* (Particle Data Group), Phys. Rev. Lett. B **592**, 1 (2004).

[35] F. James, *"MINUIT Function Minimization and Error Analysis - Reference Manual Version 94.1"*, CERN Program Library Long Writup D506 (1994).

[36] A. Castro, *"A neural net based kinematical selection for all-hadronic $t\bar{t}$ events in the Run II multijet dataset"*, CDF internal note 7972, December 2005.

[37] A. Castro, *"A Neural Net based kinematical selection for All-hadronic $t\bar{t}$ events (890 pb^{-1})"*, CDF internal note 8122, February 2006.

[38] J.F. Arguin, P.K. Sinervo, *"b-jets Energy Scale Uncertainty From Existing Experimental Constraints"*, CDF internal note 7252, September 2004.

[39] S. Xie, P.K. Sinervo, *"The Effects of Colour Flow on B Jet Energies in Top Quark Events"*, CDF internal note 7238, September 2004.

[40] Y.K. Kim, U.K. Yang *"Title: Initial state gluon radiation studies on Drell-Yan data for top-pair production in hadron collider"*, CDF internal note 6804, May 2004.

[41] H.L. Lai *et al.*, Eur. Phys. J. **C12**, 375 (2000).

[42] A.D. Martin, R.G. Roberts, W.J. Stirling and R.S. Thorne, Eur. Phys. J. C **4**, 463 (1998).

[43] A. Abulencia *et al.* [CDF Collaboration], *"Measurement of the top mass in the all-hadronic channel using the TMT method (*$1.02\,fb^{-1}$*)"*, CDF public note 8420, July 2006.

[44] F. Abe *et al.* [CDF Collaboration], Phys. Rev. Lett. **79**, 1992 (1997).

[45] A. Abulencia *et al.* [CDF Collaboration], *"Measurement of the top quark mass in the all-hadronic channel using Ideogram technique"*, FERMILAB-PUB-06-468-E. Submitted to Phys. Rev. Lett.

[46] A. Abulencia *et al.* [CDF Collaboration], *"Precise Measurement of the top quark mass in the lepton+jets topology at CDF II"*, CDF public note 8642, December 2006 (To be submitted to Phys. Rev. Lett.).

[47] A. Abulencia *et al.* [CDF Collaboration], *"Measurement of the top quark mass in the dilepton channel using a matrix element method with* $1\,fb^{-1}$*"*, CDF public note 8369, July 2006.

[48] A. Abulencia *et al.* [CDF Collaboration], arXiv:hep-ex/0612061.

[49] V.M. Abazov *et al.* [D0 Collaboration], Phys. Rev. D **74**, 092005 (2006) [arXiv:hep-ex/0609053].

[50] V.M. Abazov *et al.* [D0 Collaboration], arXiv:hep-ex/0609056.

[51] A. Abulencia *et al.* [CDF Collaboration], *"Combination of CDF Top Mass Results Using up to* $1\,fb^{-1}$ *of Data"*, CDF public note 8459.

[52] E. Brubaker *et al.* [Tevatron Electroweak Working Group], arXiv:hep-ex/0608032.