Universita' degli Studi di Trento

Dipartimento di Fisica

DOTTORATO DI RICERCA IN FISICA

CICLO XVIII

Study of $t\bar{t}$ production in tau + jets channel at CDFII using neural networks

Supervisore: Prof. Ignazio Lazzizzera

Dottoranda: Silvia Amerio

1 dicembre 2005

Contents

Introduction **1**

1 Top and τ physics **3**

1.1 The top quark 4

1.1.1 Discovery 4

1.1.2 Top production and decays 6

1.1.3 Top properties 11

1.2 The τ lepton 15

1.2.1 Discovery 16

1.2.2 τ production and decays 18

1.2.3 Physics with taus 19

2 The accelerator complex and CDF detector **27**

2.1 The accelerator complex 28

2.1.1 Instantaneous and integrated luminosity 28

2.1.2 The accelerator complex 30

2.2 CDF detector 34

3 Trigger and DAQ system **49**

3.1 Level 1 49

3.2 Level 2 51

3.3 Level 3 53

4 Reconstruction of physical objects **57**

4.1 Tracks 57

4.1.1 Standalone tracking 59
4.1.2 Outside-In tracking 61
4.1.3 Vertex reconstruction 61
4.2 Jets 63
4.2.1 Reconstruction algorithm 63
4.2.2 Jet corrections 64
4.3 Leptons 72
4.3.1 electron 72
4.3.2 muon 72
4.3.3 tau 72
4.3.4 neutrino 73
4.4 Photons 73
4.5 b quark tagging 74
4.5.1 SecVtx algorithm 74
4.5.2 JetProbability algorithm 75
4.5.3 Soft Lepton Tag algorithm 78

5 Neural Networks 81
5.1 Brief introduction to the main concepts 81
5.2 Reactive Tabu Search 86

6 $t\bar{t} \to \tau + jets$ selection 91
6.1 τ candidate reconstruction 91
6.2 τ identification variables 93
6.2.1 Input variables: global 93
6.2.2 Input variables: tau specific 100
6.3 $t\bar{t} \to \tau + jets$ selection using NNs 107
6.3.1 Neural Network analysis 110
6.3.2 Training samples 111
6.3.3 Single NN 116
6.3.4 Two NN cascade 120
6.3.5 Comparing the two techniques 124

7 $t\bar{t}$ cross section measurement **129**

7.1 The method 129

7.2 Efficiencies 131

7.3 Validation on MC samples 136

7.4 Systematics 139

7.4.1 Systematic error introduced by NNs 139

7.4.2 Other sources of systematic errors 141

7.5 Measurement on data 146

Conclusions **151**

Acknowledgements **153**

Introduction

CDF (Collider Detector at Fermilab) is a particle detector located at Fermi National Laboratories, near Chicago. It allows to study decay products of $p\bar{p}$ collisions at center-of-mass energy of 1.96 TeV. During its first period of data taking (*RunI*), *CDF* observed for the first time the top quark (1995). The current period of data taking (*RunII*) is devoted to precise measurements of top properties and to search for new physics.

This thesis work is about the top decay channel named τ + *jets*. A $t\bar{t}$ pair decays in two W bosons and two b quarks. In a $\tau + jets$ event, one out of the two W decays into two jets of hadrons, while the other produces a τ lepton and a neutrino; the τ decays semileptonically in one or more charged and neutral pions while b quarks hadronize producing two jets of particles. Thus the final state of a $\tau + jets$ event has this specific signature: five jets, one τ - *like*, i.e. narrow and with low track multiplicity, two from b quarks, two from a W boson and a large amount of missing energy from two τ neutrinos. We search for this signal in 311 pb^{-1} of data collected with *TOP_MULTIJET* trigger. We use neural networks to separate signal from background and on the selected sample we perform a $t\bar{t}$ production cross section measurement.

The thesis is structured as follows: in **Chapter 1** we outline the physics of top and τ, concentrating on their discovery, production mechanisms and current physics results involving them. **Chapter 2** is devoted to the description of the experimental setup: the accelerator complex first and *CDF* detector then. The trigger system is described in **Chapter 3**, while **Chapter 4** shows how particles are reconstructed exploiting information from different *CDF* subdetectors. With **Chapter 5** we begin to present our analysis: we use a feed forward neural network

based on a minimization algorithm developed in Trento University, called *Reactive Taboo Search* (*RTS*), especially designed to rapidly escape from local minima. Using this neural network, we explore two techniques to select $t\bar{t} \rightarrow \tau + jets$ events, the first based on a single net, the second on two neural networks in cascade; both techniques are described in **Chapter 6**, together with the variables used as inputs for the nets. Finally, in **Chapter 7** we present a method to measure cross section on the sample of events selected by neural networks.

Chapter 1

Top and τ physics

The Standard Model (SM) is a gauge theory, based on the group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, which describes strong, weak and electromagnetic interactions via the exchange of corresponding spin-1 gauge bosons: eight massless gluons and one massless photon for the strong and the electromagnetic forces respectively, three massive bosons, $W^{\pm}$ and Z, for the weak interactions [1][2][3][4].

According to the Standard Model, the fermionic matter content is organized in a 3-fold family structure:

$$\begin{bmatrix} \nu_e & u \\ e^- & d \end{bmatrix}, \qquad \begin{bmatrix} \nu_\mu & c \\ \mu^- & s \end{bmatrix}, \qquad \begin{bmatrix} \nu_\tau & t \\ \tau^- & b \end{bmatrix}, \tag{1.1}$$

where (each quark appears in three different colours)

$$\begin{bmatrix} \nu_l & q_u \\ l^- & q_d \end{bmatrix} \equiv \begin{pmatrix} \nu_l \\ l^- \end{pmatrix}_L, \qquad \begin{pmatrix} q_u \\ q_d \end{pmatrix}_L, \qquad l^-_R, \qquad (q_u)_R, \qquad (q_d)_R, \tag{1.2}$$

plus the corresponding antiparticles.

The gauge symmetry is broken by the vacuum, through the Spontaneous Symmetry Breaking mechanism (SSB) [5][6], which generates the masses of the weak gauge bosons and gives rise to the appearance of a physical scalar particle, the *Higgs* boson. The fermion masses are also generated through the SSB mechanism.

The Standard Model is one of the more succesful achievements in modern physics. It provides a very elegant theoretical framework, which is able to describe all known

experimental facts in particle physics. But a lot of open questions still remain. Why fermions are replicated in three (and only three) nearly identical copies? Are the masses the only difference among the three families? What is the origin of the SM flavour structure? These unanswered questions and more make difficult to consider SM as a complete description of the fundamental forces.

Top quark and τ lepton are both members of the third generation. Studying them we are at the frontier of experimental known physics. For all these reasons, the study of these two particles is more and more challenging.

In the following we will briefly review how top and τ were discovered, their production and decay mechanisms and the main studies involving them.

1.1 The top quark

1.1.1 Discovery

Top quark is the weak isospin partner of bottom quark. Top quark was first observed at Fermilab in 1995 by the *CDF* and *DØ* collaborations during the so called *RunI* at $\sqrt{s} = 1.8$ TeV. Before its discovery, indirect evidence for its existence came from a lot of sources [7]. The main theoretical motivation for its existence is the request of consistency for Standard Model. In particular, renormalizability demands the absence of triangle anomalies. Such a requirement can be translated into a condition on the electric charges of all left-handed fermions, condition which is met in a complete standard family where the electric charged of the leptons plus those of all color components of the quarks add up to zero. This condition would be violated for the third family if top quark was absent.

Turning to experimental evidences, observing $e^+e^- \to b\bar{b}$ at *LEP*, an asymmetry in the scattering of the b quark relative to the incoming electron direction was observed: this asymmetry was found in excellent agreement with the Standard Model expectation, assuming that the b quark is a member of an $SU(2)$ doublet.

Moreover, the experimentally determined absence of *flavour-changing neutral currents* such as $b \to \mu^+\mu^- X$ or $b \to sX$, where X is a state with no net flavour quantum numbers, implies that the b quark is a member of an $SU(2)$ doublet too.

How could *CDF* claim the discovery of top quark in 1995? *CDF* experimenters looked for $t\bar{t}$ production with subsequent decay $t\bar{t} \rightarrow WbW\bar{b}$. The search was performed in those channels where one or both Ws decay into electron or muon. These channels, especially the one with just one lepton in the final state, have good branching ratios, as we will see later, and a clean signature, due to the presence of one or two leptons. A fundamental step of the analysis was the identification of jets originating from b quarks: two algorithm were used, *SecVtx* and *SLT*, which will be described in Chapter 4. The *observation* was claimed [8] when it was seen a large excess in the signal, inconsistent with the background prediction by 4.8σ (fig. 1.1). The interpretation of the excess as $t\bar{t}$ production was supported by a peak in the mass distribution for fully reconstructed events (fig. 1.2).

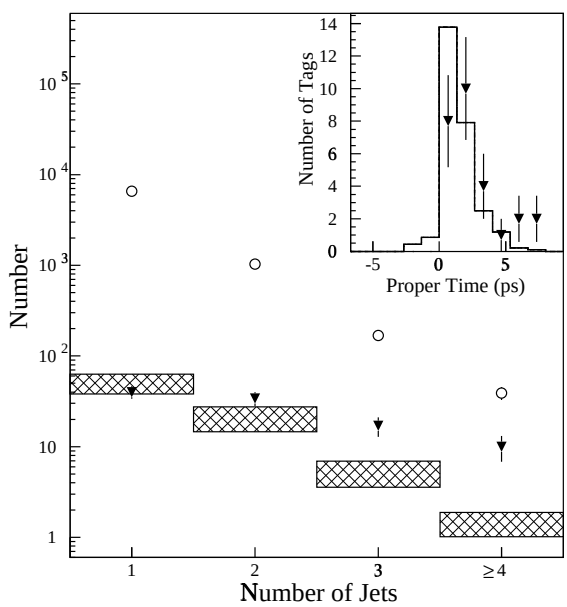

Figure 1.1: Top quark first observation: number of events before b quark tagging (circles), after quark tagging (triangles) and expected number of background events(hatched).

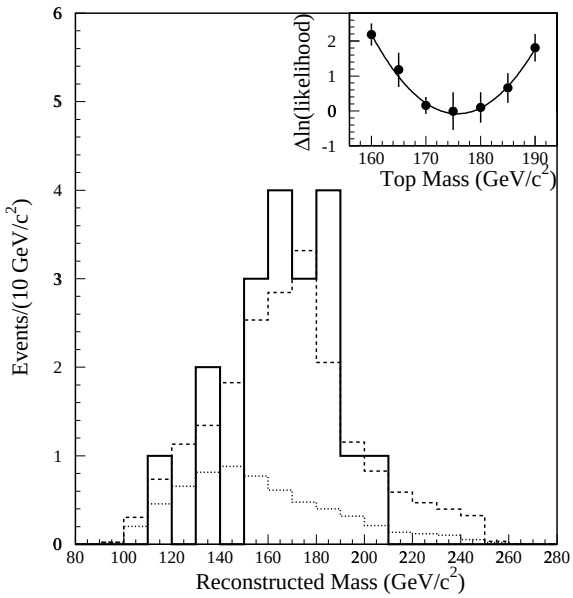

Figure 1.2: Top quark first observation: reconstructed mass distribution (solid line), background shape (dotted) and the sum of background plus Monte Carlo for $M_{top} = 175$ GeV (dashed).

1.1.2 Top production and decays

At hadron colliders, two distinct Standard Model production mechanisms for top are possible, $t\bar{t}$ production via the strong interaction and single-top production via the electroweak interaction [9][10].

Pair production

$t\bar{t}$ pairs are produced via strong interactions as a result of quark-antiquark annihilation or quark-gluon fusion. Fig. 1.3 shows the corresponding leading order Feynman diagrams. The total cross section in terms of the parton-parton process $ij \to t\bar{t}$ is

$$\sigma(p\bar{p} \to t\bar{t}) = \sum_{i,j} \int dx_1 dx_2 f_i^p(x_1, \mu^2) f_j^{\bar{p}}(x_2, \mu^2) \hat{\sigma}(ij \to t\bar{t}; \hat{s}; \alpha_s(\mu^2), M_t) \qquad (1.3)$$

where f_i^p ($f_j^{\bar{p}}$) is the probability density of finding parton i (j) with a given fraction of the proton (antiproton) momentum between x_1 and $x_1 + dx_1$ (x_2 and $x_2 + dx_2$),

$\hat{\sigma}$ is the cross section for the parton-parton subprocesses, $\hat{s}$ is the center-of-mass energy of the i-j parton system, α_s the strong coupling constant function of the renormalization scale μ and M_t is top mass.

In *RunII*, with a center-of-mass energy of 1.96 TeV, 85% of $t\bar{t}$ pairs are produced via $q\bar{q}$ annihilation against 15% produced via quark-gluon fusion.

$t\bar{t}$ production cross section[1] is small compared with other processes which take place at the Tevatron. Roughly one in 10^{10} collisions ($\sim 7 \cdot 10^{-4}$ Hz rate) produces top quark pairs.

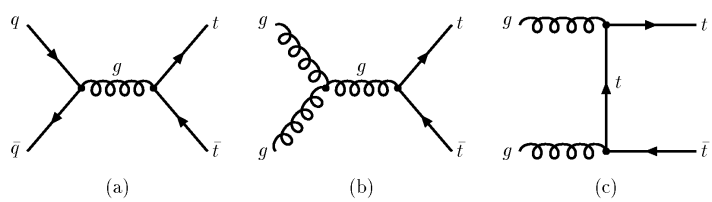

Figure 1.3: Leading order Feynman diagrams for $t\bar{t}$ production via the strong interaction: (a) $q\bar{q}$ annihilation, (b,c) gg fusion

Single-top production

A single top can be produced via electroweak interaction through the following processes (fig 1.4):

- t-channel: a space-like W boson ($q^2 \leq 0$) strikes a b quark in the proton sea, promoting it to a top quark; this channel is often referred to as $W - gluon$ fusion, because the b quark arises from a gluon splitting to $b\bar{b}$;

- s-channel: rotating the $t-$channel diagram such that the W boson becomes time-like ($q^2 \geq (m_t + m_b^2)$), one has another process that produces a single top, through a $q\bar{q}$ annihilation;

- associated production: a single top quark may also be produced via weak interaction in association with a real W boson ($q^2 = M_w^2$); one of the initial

[1]For $M_{top} = 175$ GeV and $\sqrt{s} = 1.96$ TeV, $\sigma_{t\bar{t}} = 6.7$ pb, while $\sigma_{total} \sim 10^{11}$ pb.

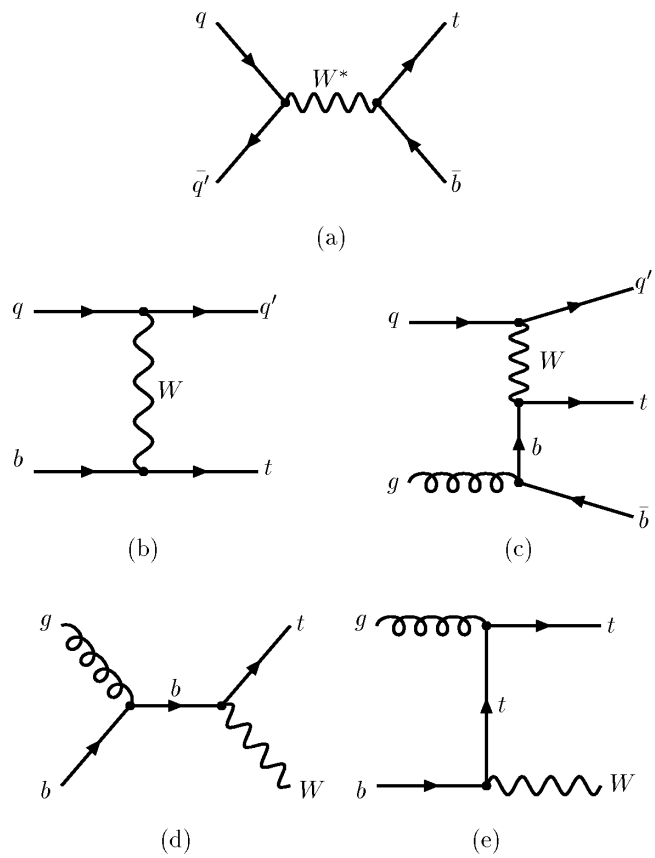

Figure 1.4: Leading order Feynman diagrams for electroweak production of single top quarks: (a) s-channel, (b,c) t-channel and (d,e) associated production with a W.

partons is a b quark in the proton sea, as in the t-channel.

The cross sections for all three processes are proportional to the matrix element $|V_{tb}|^2$ of the Cabibbo-Kobayashi-Maskawa matrix (CKM). Therefore, measuring the single top quark production cross section provides a direct probe of $|V_{tb}|$ and the weak tbW vertex in general. In table 1.1 cross section for the different channels are reported.

cross section (pb)	cross section (pb)
$\sigma^{NLO}_{s-channel}$	0.447 ± 0.002
$\sigma^{NLO}_{t-channel}$	0.959 ± 0.002
σ^{LL}_{assoc}	0.093 ± 0.024

Table 1.1: Single top quark production cross sections at Tevatron *RunII*

Top decays

Standard Model predicts that top decays into W boson and a b quark with a branching ratio greater than 0.998. Other decays, such as $t \to Ws$ and $t \to Wd$ are also allowed but suppressed by factor 10^{-3} - 10^{-4} by the square of the CKM matrix elements V_{ts} and V_{td}, whose values can be estimated, under the assumption of unitarity of the three-generation CKM matrix, to be less than 0.043 and 0.014 respectively [11].

Top decay width is then [12]

$$\Gamma(t \to Wb) = \frac{G_F M_t^3}{8\pi\sqrt{2}}\Big(1 - \frac{M_W^2}{M_t^2}\Big)^2\Big(1 + 2\frac{M_W^2}{M_t^2}\Big)^2\Big[1 - \frac{2\alpha_s}{3\pi}\Big(\frac{2\pi^2}{3} - \frac{5}{2}\Big)\Big] \tag{1.4}$$

For $M_t = 175$ GeV/c^2 we have

$$\Gamma(t \to Wb) \approx 1.55 GeV \to \tau_t = \Big(\frac{1}{\Gamma_t}\Big) \sim 4 \cdot 10^{-25} s \tag{1.5}$$

Top decay width is smaller than the characteristic hadronization time of QCD ($\tau_{had} \approx 28 \cdot 10^{-25}$) so top decays before hadronizing. However the distance between its production and decay verteces is order 10^{-16} m, well below the spatial resolution of any detector by many orders of magnitude. Thus, to detect top quark we need to identify and reconstruct its decay products. Due to top large mass, its decay products have good angular separations and high momenta in the laboratory frame. Most of them travel in the central region of the detector and have transverse momentum exceeding 20 GeV.

W's lifetime is so short ($\sim 3 \cdot 10^{-25}$ s) that it instantaneously decays into lepton-neutrino pair or quark-antiquark pair. The first decay type manifests itself as an

electron, a muon or a tau. Moreover, a large amount of missing energy from neutrinos is present. The hadronic decay of W, on the other hand, manifests itself as two jets of particles. b quark decays in a shower of particles too and due to its large lifetime, compared to typical hadronization time, it travels a few millimeters before decaying.

Depending on the decay of W, we can divide top decays into three different channels:

- **dilepton**
  Both Ws decay into leptons (electrons or muons). This channel has the lowest branching ratio ($\sim 5\%$) but it is the *cleanest* among the three, due to its two leptons that give a clean signature. Main background sources are di-boson and Drell-Yan events.

- **lepton plus jets**
  One W decays in electron or muon, the other into hadrons. Branching ratio is greater ($\sim 30\%$) but background too (W + jets events).

- **all-hadronic**
  The channel with the greatest branching ratio ($\sim 45\%$) but the most difficult to detect; in this channel, in fact, both Ws decay into hadrons, thus giving a final state with six jets, difficult to separate from the huge background from QCD multijet events that arise from a $2 \rightarrow 2$ low generation parton process producing two energetic (*hard*) leading jets and less energetic (*soft*) radiated gluon jets.

Channels involving a τ lepton are treated separately. All three channels described above can have a τ (or two in the dilepton case) in the final state, but only taus decaying into hadrons can be revealed: if τ decays into leptons, the electron or muon coming from it is not distinguishable from an electron/muon from W.

In fig. 1.5 $t\bar{t}$ production and decay are schematized; in table 1.2 branching ratios for all possible W decays are reported.

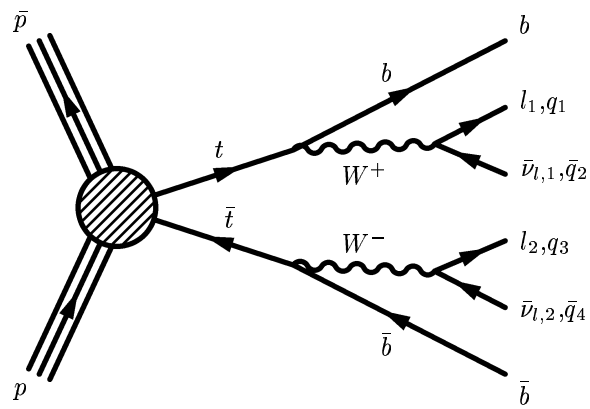

Figure 1.5: $p\bar{p}$ collision and top decay

W^- ↓ W^+→	e^+	μ^+	τ^+	$u\bar{d}$ (x3)	$c\bar{s}$ (x3)
e^-	$\frac{1}{81}$	$\frac{1}{81}$	$\frac{1}{81}$	$\frac{3}{81}$	$\frac{3}{81}$
μ^-	$\frac{1}{81}$	$\frac{1}{81}$	$\frac{1}{81}$	$\frac{3}{81}$	$\frac{3}{81}$
τ^-	$\frac{1}{81}$	$\frac{1}{81}$	$\frac{1}{81}$	$\frac{3}{81}$	$\frac{3}{81}$
$\bar{u}d$ (x3)	$\frac{3}{81}$	$\frac{3}{81}$	$\frac{3}{81}$	$\frac{9}{81}$	$\frac{9}{81}$
$\bar{c}s$ (x3)	$\frac{3}{81}$	$\frac{3}{81}$	$\frac{3}{81}$	$\frac{9}{81}$	$\frac{9}{81}$

Table 1.2: Single top quark production cross sections.

1.1.3 Top properties

Mass

Top quark mass is a parameter of fundamental impact in different areas of particle physics. With a mass[2] of 172.7±2.9 GeV/c^2 it is 35 times heavier than the next heaviest quark, the bottom, and it is the heaviest elementary particle known. A precise measurement of top mass, combined with improved measurements of other electroweak parameters, determine important constraints on the value of Higgs boson mass. W mass theoretical calculation, in fact, is subject to radiative corrections that arise from creation and absorption of virtual quarks and bosons. Quark correc-

[2]Tevatron summer 2005 combination of *CDF* and *DØ* *RunI* and *RunII* results [13]

tion depends on top mass while boson correction on $log(M_H)$, where M_H is Higgs boson mass. Measuring with high precision W and top masses we can deduce boson correction and thus obtain a constraint on Higgs mass (see fig. 1.6) [14]. In B and K physics, moreover, many observables have terms depending on top mass and for precision Standard Model electroweak fits, M_t enters quadratically in many places as well.

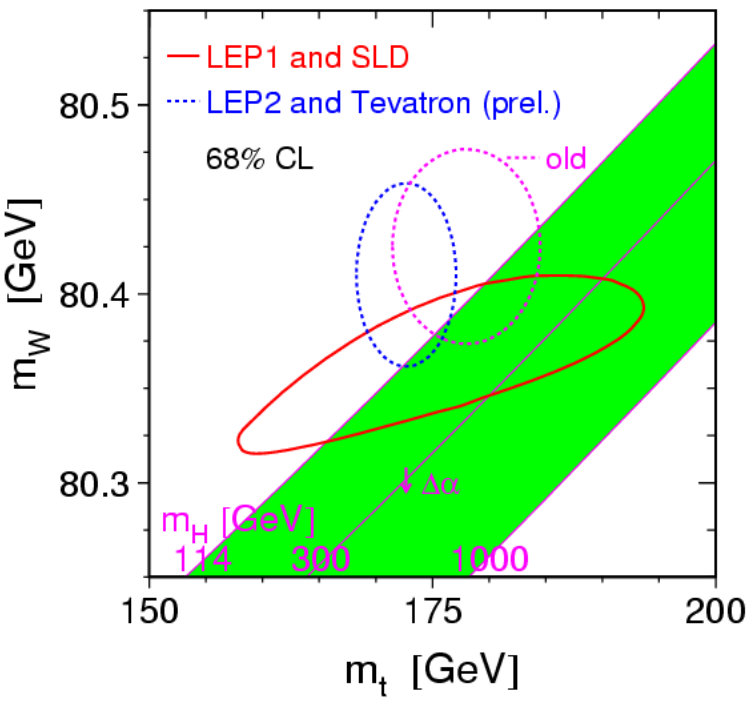

Figure 1.6: Constraint on higgs mass from top and W mass measurements.Green bands show theoretical predictions as a function of Higgs mass.

At Tevatron top mass has been measured using different methods. The *template method* uses fits to the kinematic variables to determine the most probable jet-parton assignment for each event. In this way, mass templates can be reconstructed for different Monte Carlo samples. By comparing the mass distribution observed in data to the expected templates for Monte Carlo generated signal and background,

top quark mass can be determined [15]. The best top mass measure obtained so far is based on a more sophisticated version of the template method applied to the lepton + jets channel. In this method mass templates are built both for top quark and W boson; the latter are function of both jet energies systematics (JES) [3] and top mass. Data are fitted simultaneously for top mass and JES, decreasing in this way the uncertainty due to the JES themselves [16]. Template method has been also applied to dilepton channel, with the difference that the presence of two neutrinos implies that one needs extra information to reconstruct the final state. In the *neutrino weighting approach* a probability distribution is computed for each event as a function of top mass and considering all possible lepton-jet combinations, weighted by their probability. The top mass maximizing the probability is used to build the mass template [17].

Another method to measure top mass is the *matrix element method* [18]. The idea is to calculate for each event a likelihood as a function of m_{top} according to matrix elements:

$$\mathcal{L}(m_{top}) = \sum_{jet-parton} \int \frac{2\pi^4}{Flux} |\mathcal{M}|^2 F(z_1, z_2) f(p_T) w(\mathbf{x}, \mathbf{y}; m_{top}) d\mathbf{x} \tag{1.6}$$

where $\mathcal{M}$ is the matrix element, $F(z_1, z_2)$ the parton distribution function, $f(p_T)$ the probability for the transverse momentum of $t\bar{t}$ system and $w(\mathbf{x}, \mathbf{y}; m_{top})$ a transfer function which describes correlations between final state parton variables $\mathbf{x}$ and observed quantities $\mathbf{y}$. The sum is over all jet-parton combinations. An event joint likelihood $L = \prod_i \mathcal{L}(m_t)$ can be finally built and its maximum provides an estimate of the top quark mass.

In figs. 1.7 and 1.8 CDF and $D\emptyset$ top mass estimates are summarized.

Cross section

The measurement of the production cross section for $t\bar{t}$ pairs can be a test of QCD. A significant deviation of the measured cross section from the predicted value can signal a non-Standard Model production mechanism. The theoretical value of

[3] Jet energy corrections will be treated in more detail in section 4.2.2

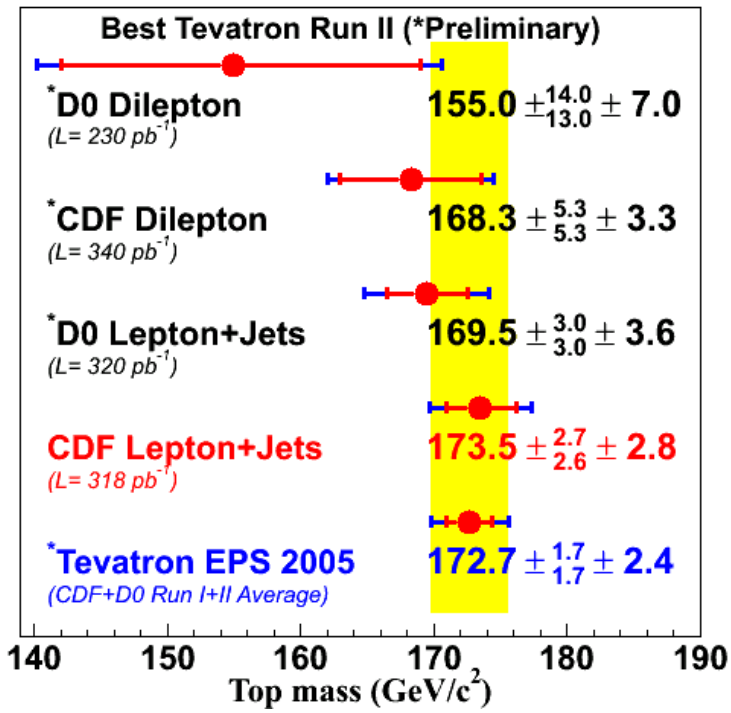

Figure 1.7: $D\emptyset$ and CDF best top mass measurements.

cross section has been calculated with relatively little uncertainty and for M_{top} = 175 GeV and $\sqrt{s}$ = 1.96 TeV it is 6.7 ± 0.5 pb [19].

In $RunI$ the combined estimate of all measurements was 6.5 + 1.7 - 1.4 pb, for M_{top} =175 GeV, obtained with 110 pb^{-1} of data, to be compared with the value found by $D\emptyset$, which is 5.7 ± 1.6 pb[4].

In $RunII$, with increasing luminosity we expect to obtain a more precise measurement: CDF preliminary combined measurement made with 350 pb^{-1} is 7.1±1.0 pb [20].

In figure 1.9 theoretical values of cross section as a function of top mass are shown together with the current best estimate obtained at CDF. In fig. 1.10 all cross section estimates made at Tevatron during $RunII$ are reported.

[4] The theoretical value from [19] is 5.19 ± 0.33 for M_{top} = 175 GeV and $\sqrt{s}$ = 1.8 TeV

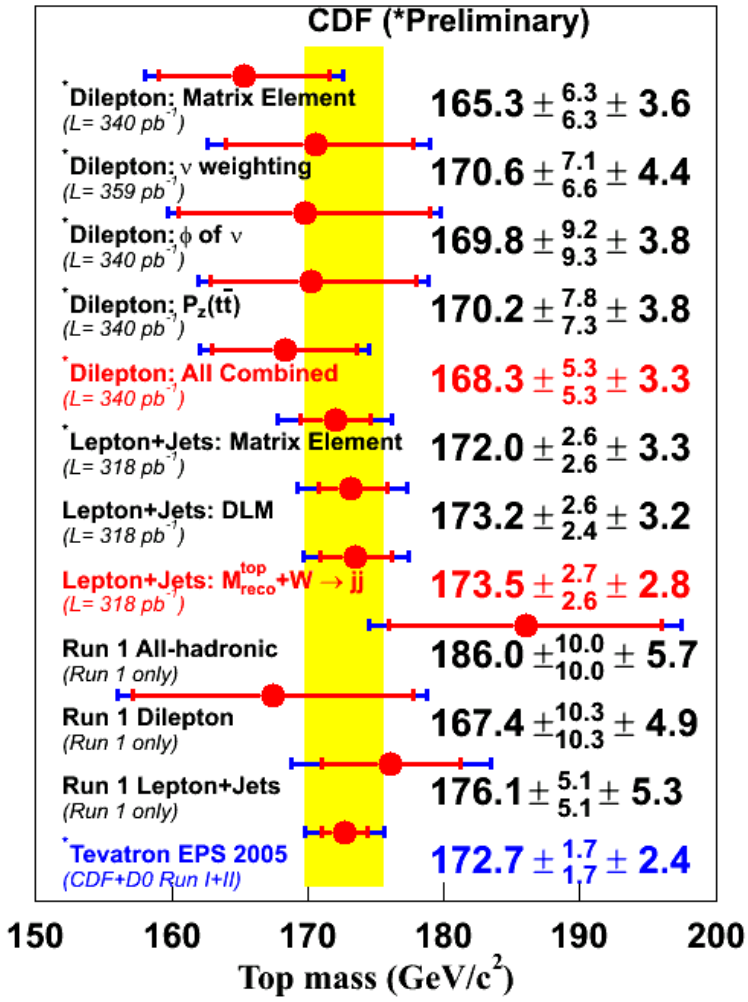

Figure 1.8: *CDF* top mass measurement from different analysis.

1.2 τ physics

According to Standard Model, τ lepton is a fermion, the heaviest of the leptons, member of the third generation which decays into particles belonging to the first and second ones. In the following we will give a short description of τ discovery and then review its possible production mechanisms and decays and the main studies involving it.

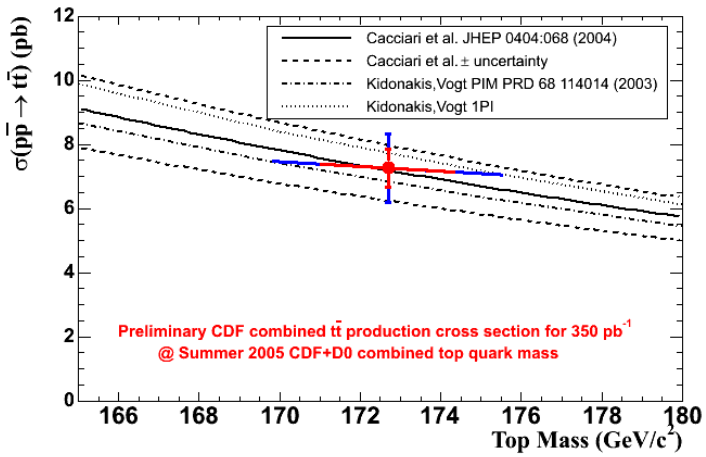

Figure 1.9: Theoretical cross section as a function of top mass; red dot is *CDF* latest estimate of cross section.

1.2.1 Discovery

"*Since muons exist in nature for no apparent reason, it is possible that other heavy leptons may exist in nature*". These words are taken from the famous paper by Tsai, *Decay correlations of heavy leptons in* $e^+ + e^- \to l^+ + l^-$, published in 1971 [21]. In this paper the existence of a lepton heavier than muon is supposed and its properties are studied.

τ lepton was then discovered in 1975 at *SPEAR* e^+e^- collider ring. The idea was to look for $e^+e^- \to L^+L^-$, with $L^+ \to e^+/\mu^+ + \nu$ and $L^- \to e^-/\mu^- + \nu$. *SPEAR* had its first collision in April 1972 and took a sizeable amount of data from the spring of 1973 trough the spring of 1974. In 1974, Martin Perl started looking at the data and found events of the form $e^+ + e^- \to e^\pm + \mu^\mp +$ missing energy, which couldn't be explained by conventional backgrounds [22]. Most of those events were detected at or above a center of mass energy of 4 GeV. The missing energy and missing momentum spectra required that at least two additional particles were produced in each event. What were the possible sources of these events? Two sources were

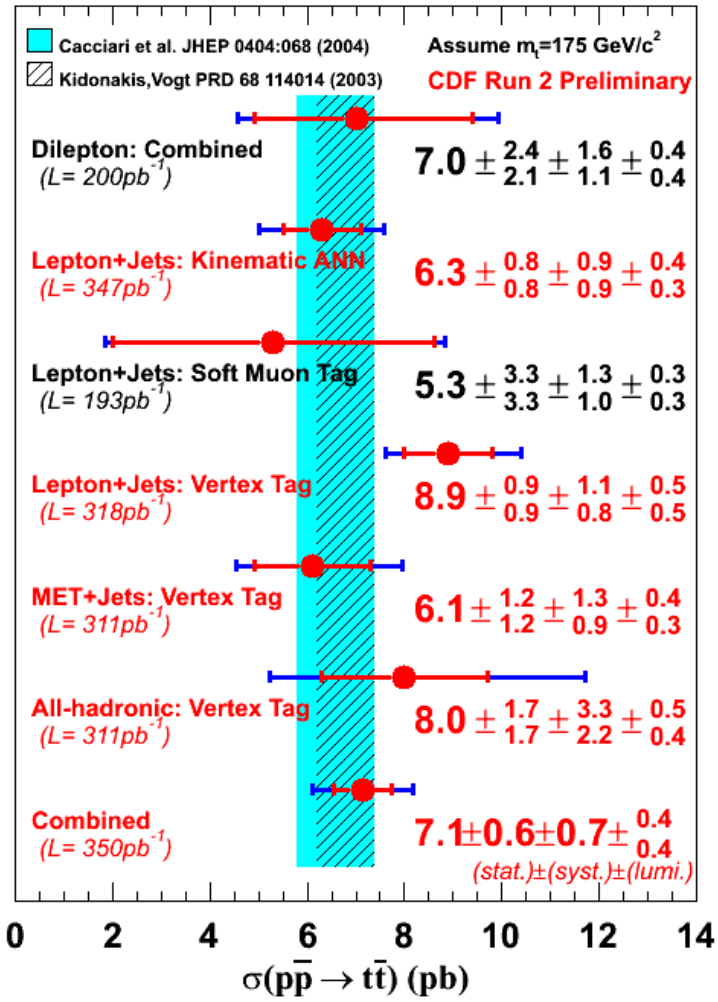

Figure 1.10: All *CDF* and *DØ* cross section measurement

possible: a meson decaying by a two-body decay,

$$e^+e^- \to M^+M^-, \quad M^+ \to e^+\nu, \quad M^- \to \mu^-\bar{\nu}$$

or a lepton decaying by a three body decay,

$$e^+e^- \to L^+L^-, \quad L^+ \to e^+\nu\bar{\nu}, \quad L^- \to \mu^-\bar{\nu}\nu$$

After detector improvements and the growing of data, looking at the momentum spectrum of the events, it was possible to determine that it was compatible with a

spectrum of a three-body decay but in total disagreement with a two-body decay. The next step was to determine the nature of the two missing particles. Neutron was eliminated because it was set an upper limit of 0.7 GeV/c^2 at 95% CL on the mass possessed by any of the undetected particles and an upper limit of 39% at 90% CL was set on the percentage of $e^{\pm}\mu^{\mp}$ events containing kaons, photons and charged particle escaping detection. Thus missing particles had to be neutrinos and each decay had to have a lepton and two missing neutrinos. The only particle with this signature was a heavy lepton [23]. It was time to give a proper name to this new lepton, named U particle - *unknown* particle - up to then. A greek letter was chosen, τ, standing for *triton*, that means *third*. This name was used for the first time in [24] where a detailed study of the new lepton just discovered was presented: τ mass was measured to be 1.9 $\pm$ 0.1 GeV/c^2, 1.2 standard deviations above today accepted value[5] and an upper limit of 600 MeV/c^2 was set on τ neutrino.

1.2.2 τ production and decays

We have seen that τ was first produced in e^+e^- collision. Depending on the energy of colliding electron and positron, from threshold to about 10 GeV, τ production is dominated by γ exchange; from above 10 GeV to Z^0 resonance, the Z^0 exchange contributes trough interference with the γ exchange, and obvioulsy it is the only exchange at Z^0 resonance. At higher energies, γ exchange once again dominates.

τ can also be produced by photoproduction:

$$\gamma + N \rightarrow \tau^+ + \tau^- + N'$$

where N is a target proton or nucleus and N' represents the final hadronic state. Moreover, τ can come from decay of other particles, such as W and D and B. Finally, virtual photons emitted in the collision of a pair of heavy ions can produce a $\tau^+\tau^-$ pair when the ions are at energies much greater than the τ mass [25].

τ can decay leptonically into electrons and muons or semileptonically into hadrons. Both decay types are accompanied by neutrinos. In table 1.3 all τ decays are listed

[5] $1777.03^{+0.30}_{-0.26}$ MeV

with their branching fractions.

Semileptonic decays are preferably into one charged hadron (*one-prong event*) or three charged hadrons (*three-prong event*). Hadrons are mostly pions; decays into other charged hadrons, such as kaons, are very rare. In about 73% of one-prong events and 41% of three-prong events, one or more neutral pion is present.

1.2.3 Physics with taus

τ lepton production and decay mechanisms are know with precision. τ is heavy enough to have a large variety of decay modes, thus being more sensitive than electrons or muons to new physics related to flavour and mass generation problems. **Tests of lepton universality** have been carried out at *LEP* and at *CDF* too, as we will see later: finding out if the leptonic coupling constants for weak currents are identical for all generations is an important test of Standard Model. Moreover, **Lorentz structure of leptonic τ decay** can be investigated to determine τ decay parameters and, exploiting τ decay into hadrons one can make **QCD tests** and **precise determination of the strong coupling constant**. In the minimal Standard Model with massless neutrinos, there is a separately conserved additive lepton number for each generation. All present data are consistent with this conservation law. However, there are no strong theroretical reasons forbidding a mixing among different leptons, in the same way as happens in the quark sector. Many models in fact predict **lepton-flavour** or even **lepton-number violation** at some level. Experimental searches for these processes can provide information on the scale at which new physics begins to play a significant role. A detailed description of the studies mentioned above can be found in [26]; now we want to concentrate on the analysis involving τ carried out at *CDF*.

Lepton universality

At *CDF* a test of lepton universality has been done measuring the ratio $\frac{BR(W\to\tau\nu)}{BR(W\to e\nu)}$ [27]. The measurement was performed using 72 pb^{-1} of data collected with a trigger requiring a τ lepton and missing energy. As the branching ratio is proportional to the

Decay mode	Branching ratio (%)
Leptonic decay modes	
$\tau^- \to e\bar{\nu}_e\nu_\tau$	17.84 ± 0.06
$\tau^- \to \mu\bar{\nu}_\mu\nu_\tau$	17.36 ± 0.06
One prong semileptonic decay modes	
0 neutrals:	
$\tau^- \to \pi^-\nu_\tau$	11.06 ± 0.11
$\tau^- \to K^-\nu_\tau$	0.69 ± 0.02
$\tau^- \to K^*(892)^-\nu_\tau$	1.29 ± 0.18
total	13.04 ± 0.21
≥ 1 neutrals:	
$\tau^- \to \rho\nu_\tau \to \pi^-\pi^0\nu_\tau$	25.42 ± 0.14
$\tau^- \to K^-\pi^0\nu_\tau$	0.45 ± 0.03
$\tau^- \to a_1\nu_\tau \to \pi^-2\pi^0\nu_\tau$	9.17 ± 0.14
$\tau^- \to \pi^-3\pi^0\nu_\tau$	1.08 ± 0.10
total	36.12 ± 0.22
Three prong semileptonic decay modes	
0 neutrals:	
$\tau^- \to a_1\nu_\tau \to \pi^-\pi^-\pi^+\nu_\tau$	9.47 ± 0.10
≥ 1 neutrals:	
$\tau^- \to ... \to \pi^-\pi^-\pi^+ \geq 1\pi^0\nu_\tau$	5.63 ± 0.22
$\tau^- \to (a_1\pi)^-\nu_\tau$	
$\tau^- \to \rho^0\pi^-\pi^0\nu_\tau$	
$\tau^- \to \rho^-\pi^-\pi^+\nu_\tau$	
$\tau^- \to \rho^+\pi^-\pi^-\nu_\tau$	
$\tau^- \to \omega\pi^-\nu_\tau$	1.6 ± 0.4

Table 1.3: Branching ratios of τ decay modes.

square of the coupling constant we obtain

$$\frac{g_\tau}{g_e} = \sqrt{\frac{BR(W \to \tau\nu)}{BR(W \to e\nu)}} = 0.99 \pm 0.02(stat) \pm 0.04(syst) \tag{1.7}$$

in excellent agreement with the theoretical value 1.021 ± 0.015 [28].

$Z^0 \to \tau\tau$ cross section measurement

A well measured $Z^0 \to \tau\tau$ cross section is very important in all analysis involving taus. Being τ the heaviest lepton, various new physics processes can lead to final states with one or more taus. Standard Model $Z^0 \to \tau\tau$ is thus the main background in most analysis looking for new particles decaying in τ pairs.

$Z^0 \to \tau\tau$ cross section has been measured on 350 pb^{-1} of data in the channel were one τ decays into electron and the other into hadrons. Major backgrounds are QCD and γ + jets events, electroweak events with two real leptons and W + jets events. The measured cross section is

$$\sigma(p\bar{p} \to Z^0)BR(Z^0 \to \tau\tau) = 265 \pm 20(stat) \pm 21(syst) \pm 15(lumi)pb \tag{1.8}$$

in agreement with the NLO prediction of approximately 260 pb [29].

Search for new physics with high mass τ pairs

Various new physics processes can lead to very high mass τ pairs. Examples are $Z' \to \tau\tau$ and MSSM Higgs $A \to \tau\tau$. Known backgrounds are from the high mass tail of Drell-Yan processes $Z/\gamma^* \to \tau\tau$ and jets faking τ's from W + jets, QCD dijet and multijet events.

One strategy used to search for such processes is by performing a counting experiment. Having selected events with $e+\tau_h$, $\mu+\tau_h$ and $\tau+\tau_h$, where index h denotes a decay into hadrons and using the four-momentum sum of lepton and τ and the missing energy, an invariant mass m is defined. Applying a cut at 120 GeV/c^2 the large peak from $Z^0 \to \tau\tau$ is removed. The region below this cut is used as a control sample while in the signal region a blind analysis is performed. No excess over background was seen: a 95% CL upper limit on $\sigma(p\bar{p} \to X)BR(X \to \tau\tau)$ vector and scalar bosons was set [30](fig. 1.11).

As far as MSSM $A \to \tau\tau$ is concerned, a search was made for $90 < m_A < 250\,\mathrm{GeV/c^2}$. One τ is detected in the decay to e or μ and neutrinos, the other in the decay to hadrons and neutrino. No signal evidence for the mass range considered wa seen, thus an exclusion limit at 95% CL was set on $\sigma(p\bar{p} \to X)BR(A \to \tau\tau)$ [31] (see fig. 1.12).

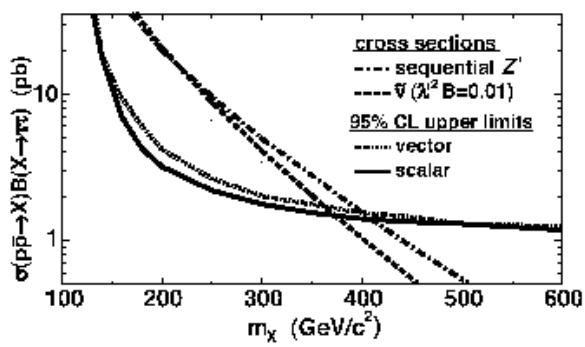

Figure 1.11: Upper limits at 95% CL on the production cross section times branching ratio to tau pairs of scalar and vector particles, as a function of particle mass. The figure also shows the cross section times tau pair branching ratio for scalar neutrinos and sequential Z' bosons.

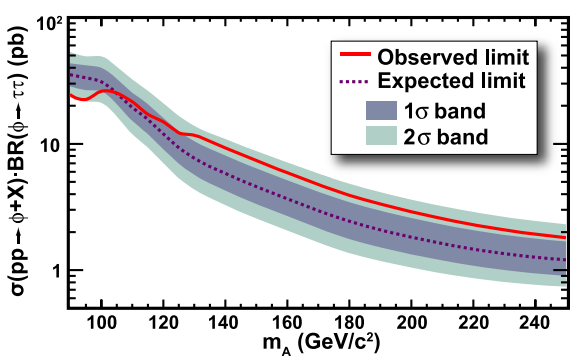

Figure 1.12: Upper limits at 95% CL on Higgs production cross section times branching fraction to τ pairs. The expected limits from pseudo-experiments are also shown.

Bibliography

[1] S.Weimberg, **A Model of Leptons**, *Phys.Rev.Lett.19,1264 (1967)*

[2] S.L.Glashow, J.Iliopoulos, L.Maiani, **Weak Interactions with Lepton-Hadron Symmetry**, *Phys.Rev.D 2,1985 (1970)*

[3] H.Fritzsch, M.Gell-Mann, H.Leutwyler, **Advantages of the Color Octet Gluon Picture**, *Phys.Lett.B 47, 365 (1973)*

[4] E.Leader, E.Predazzi, **An introduction to gauge theories and modern particle physics**, *Volume 1 and 2, Cambridge University Press (1996)*

[5] J.Goldstone, A.Salam, S.Weimberg, **Broken Symmetries**, *Phys.Rev.127, 965-970 (1962)*

[6] P.W.Higgs, **Broken Symmetries and the Masses of Gauge Bosons**, *Phys.Rev.Lett.13, 508-509 (1964)*

[7] J.H.Kühn, **Theory of top quark production and decay**, *TTP96-18*, 1996

[8] F.Abe et al.The *CDF* Collaboration, **Observation of Top Quark Production in p anti-p Collisions with the Collider Detector at Fermilab**, *Phys. Rev. Lett. 74, 2626*, (1995)

[9] M.Beneke, I.Efthymiopoulos, M.L.Mangano, J.Womersley, **Top quark physics**, *hep-ph/0003033*, 2000

[10] D.Chakraborty, J.Konigsberg, D.Rainwater, **Review of Top Quark Physics**, *Ann.Rev.Nucl.Part.Sci (2003) 301-351*

[11] S.Eidelman et al. (Particle Data Group), **The Cabibbo-Kobayashi-Maskawa quark-mixing matrix**, *Phys.Rev.Lett. B 592, 1 (2004)* (`http://pdg.lbl.gov`)

[12] S.Eidelman et al.(Particle Data Group), **The Top Quark**, *Phys.Rev.Lett. B 592, 1 (2004)* (`http://pdg.lbl.gov`)

[13] *CDF* and *DØ* collaborations, **Combination of *CDF* and *D* results on the top-quark mass**, *hep-ex/0507091*, 2005

[14] G.D'Agostini, G.Degrassi, **Constraints on the Higgs Boson Mass From Direct Searches and Precision Measurements** *DFPD-99-TH-02, hep-ph/9902226 (1999)*

[15] F.Abe et al. (CDF Collaboration), *Phys.Rev.D50, 2966 (1994)*

[16] The CDF Collaboration, **Precision Top Quark Mass Measurement in the Lepton + Jets Topology in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV**, *CDF preprint hep-ex/0510049, Submitted to Phys.Rev.Lett. October 18, 2005*

[17] The CDF Collaboration, **Measurement of the Top Quark Mass using the Neutrino Weighting Algorithm on Dilepton Events at CDF**, *CDF/PHYS/TOP/PUBLIC/7303*

[18] K.Kondo, *J.Phys.Soc.57, 4126 (1988)*

[19] M.Cacciari et al., **The $t\bar{t}$ cross section at 1.8 and 1.96 TeV: a study of the systematics due to parton densities and scale dependence**, *JHEP 0404 (2004) 068*

[20] *CDF* collaborations, **Combination of *CDF* top pair production cross section measurement**, *CDF/PHYS/TOP/PUBLIC/7749 (2005)*

[21] Y.-S.Tsai, **Decay Correlation of Heavy Leptons in $e^+ + e^- \rightarrow l^+l^-$**, *Phys.Rev.D 4 2821-2837 (1971)*

[22] M.L.Perl et al., **Evidence for anomalous lepton production in $e^+ - e^-$ annihilation**, *Phys.Rev.Lett 35, 1489 (1975)*

[23] M.L.Perl et al., **Properties of anomalous $e\mu$ events produced in $e^+ - e^-$ annihilation**, *Phys.Rev.Lett 63B, 466 (1975)*

[24] M.L.Perl et al., **Properties of the proposed τ lepton**, *Phys.Rev.Lett 70B, 487 (1975)*

[25] M.L.Perl, **Tau Physics**, *Presented at the 20th SLAC Summer Institute on Particle Physics: The Third Family and the Physics of Flavor Published in SLAC Summer Inst. 1992:47-80*

[26] A.Pich, **Tau Physics**, *Int.J.Mod.Phys.A15S1 (2000) 157-173*

[27] P.Murat, **Measurement of the $\sigma(p\bar{p} \rightarrow W)BR(W \rightarrow \tau\nu)$ and $BR(W \rightarrow \tau\nu)/BR(W \rightarrow e\nu)$ with RunII data**, *CDF/PUB/ELECTROWEAK/PUBLIC/6010 (2003)*

[28] A.Pich, **Tau Physics: Theoretichal Perspective**, *Nucl.Phys.Proc.Suppl.98 (2001) 385-396*

[29] S.Baroiant et al., **$Z^0 \rightarrow \tau\tau$ Cross-Section Measurement**, *CDF/ANAL/EXOTIC/CDFR/7589 (2005)*

[30] D.Acosta et al. (CDF Collaboration), **Search for New Physics Using High Mass Tau Pairs from 1.96 TeV $p\bar{p}$ Collisions**, *Phys.Rev.Lett.95 131801 (2005)*

[31] A.Abulencia et al.(CDF Collaboration), **Search for neutral MSSM Higgs boson decaying to τ pairs in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV**, *hep-ex/0508051 (2005)*

Chapter 2

The accelerator complex and CDF detector

The Fermi National Laboratories (*Fermilab*) host the most powerful particle accelerator working at present in the world. While we are speaking a new and more powerful one, the Large Hadron Collider (*LHC*), is being built at *CERN* laboratories, near Geneva, but for the next few years Fermilab will be the only laboratory capable of taking us up to a dimension of less than 10^{-15} m, looking at hadron constituents, the quarks.

Fermilab, originally named the National Accelerator Laboratory is located in Batavia, near Chicago (Illinois); it was commissioned in 1967 and in 1974 was renamed in honor of Enrico Fermi. Its accelerator complex accelerates protons and antiprotons and makes them collide in correspondence of two detectors, *CDF* and *DØ*. Two major components of the Standard Model were observed at Fermilab: the bottom quark (May-June 1977) and the top quark (February 1995). In July 2000, Fermilab experimenters announced the first direct observation of the tau neutrino.

In the following sections, we will describe Fermilab accelerating complex and how protons and antiprotons are produced, but first we will spend a few words on a very important accelerator parameter, luminosity.

2.1 The accelerator complex

2.1.1 Instantaneous and integrated luminosity

While building an accelerator, a fundamental construction parameter is the luminosity we want to achieve. Luminosity is a parameter that directly rises from the computation of the probability $W_{i\to f}$ for a generic process $i \to f$, where i and f are the initial and final states, respectively. In our case, the initial state is made up by two particles, a proton and an antiproton, while the final state by a generic number N of particles. The probability amplitude, emphasizing the overall four-momentum conservation, has the general structure

$$\langle f\,|T|\,p;\,\bar{p}\rangle \;=\; (2\pi)^4\;\delta^{(4)}(P_f - p_p - p_{\bar{p}})\,\langle P_f\,|M|\,p_p;p_{\bar{p}}\rangle \tag{2.1}$$

where it is supposed that each particle a in the initial and final states is described by a narrow wave packet that obeys, as obvious, on-shell mass condition, the Klein-Gordon equation, while, moreover, it is peaked around a four-momentum p_a:

$$\mathcal{F}_{\rm p}{}^{a}(x) \;\equiv\; \langle x\,|\,a\rangle \;=\; \frac{1}{(2\pi)^{3/2}}\int d^4q\;\theta(q_0)\;\;\delta(q^2 - m_a{}^2)\;\hat{f}_{\rm p}(q)\;e^{-i\,qx} \tag{2.2}$$

(hiding all remaining quantum numbers). Integrating the square modulus of 2.1 over its space dependences and after other manipulations, that use approximations permitted by the narrowness of the wave packets, assuming that protons and antiprotons are grouped in bunches, we end up with the transition probability $W_{i\to f}$ given by

$$W_{i\to f} \cong (2\pi)^4\;\delta^4(P_f - p_p - p_{\bar{p}})\;|\langle f\;|M|\;p_p;p_{\bar{p}}\rangle|^2\;\frac{1}{4\,\omega_p\,\omega_{\bar{p}}}\int d^4x\;\rho_p(x)\,\rho_{\bar{p}}(x) \tag{2.3}$$

where the ω's are energies, ν is the frequency of the crossing of the proton and antiproton bunches and the ρ's, that have the meaning of probability density of particle location, are the time component of conserved four-currents given by

$$i\;(\mathcal{F}^*\,\partial_\mu\mathcal{F}\;-\;\mathcal{F}\,\partial_\mu\mathcal{F}^*) \tag{2.4}$$

The square amplitude in 2.3 is what could be computed by a theory (in fact the Standard Model). What appears in the integral depends on the experimental setup;

the integral itself has dimension of inverse cross section and is a measure of the chance the incoming protons and antiprotons have to come in interaction.
We can assume that the densities ρ are Gaussian near the collision points and that, for simplicity, the collisions themselves are head-on; then, parametrizing the bunches path by s and calling $x(s)y(s)$ a plane orthogonal to the path in s, we can write approximately

$$\rho^{\pm}\,(\,x(s),\,y(s),\,s\pm vt\,) = \frac{N^{\pm}}{(2\pi)^{(3/2)}\,\sigma_x\,\sigma_y\,\sigma_s}\;e^{-\frac{x^2}{2\,\sigma_x^2}-\frac{y^2}{2\,\sigma_y^2}-\frac{(s\pm vt)^2}{2\,\sigma_s^2}} \tag{2.5}$$

where $\pm$ refers to *proton/antiproton*, N is the number of particles in a bunch, v is the speed of the bunches and the σ's are radii of the portion of the the crossing bunches that effectively overlap. In 2.3 we have consequently

$$\int d^4x\;\rho_p(x)\,\rho_{\bar{p}}(x) \;\equiv\; \nu\,\Delta\int dx\,dy\,ds\,dt\;\rho_p(x,\,y,\,s+vt\,)\,\rho_{\bar{p}}(x,\,y,\,s-vt\,) \tag{2.6}$$

$$= \;\nu\frac{N_p\,N_{\bar{p}}}{4\pi\,\sigma_x\,\sigma_y}\,\frac{\Delta}{2v} \tag{2.7}$$

$$\equiv\;\mathcal{L}\,\frac{\Delta}{2v} \tag{2.8}$$

where Δ is the whole lasting of the data taking, long with respect to the duration of each effective crossing of the colliding bunches, and (the lab reference frame is also the center of mass frame in our case)

$$v=\frac{|\vec{p}\,|}{\omega}\,,\qquad |\vec{p}\,|=|\vec{p}_p|=|\vec{p}_{\bar{p}}|,\qquad \sqrt{m_p^2+\vec{p}^{\,2}}=\omega=\omega_p=\omega_{\bar{p}} \tag{2.9}$$

Thus we have

$$\frac{d\,W_{i\to f}}{d\,t}\;\cong\;\delta^4(P_f-p_p-p_{\bar{p}})\,\frac{(2\pi)^4}{2\,\omega\,|\vec{p}\,|}\;|\langle f\;|M|\;p_p;p_{\bar{p}}\rangle|^2\;\mathcal{L} \tag{2.10}$$

$$=\;\frac{(2\pi)^4\;\delta^4(P_f-p_p-p_{\bar{p}})}{\sqrt{(\,p_p\cdot p_{\bar{p}})^2-m_p^2m_{\bar{p}}^2}}\;|\langle f\;|M|\;p_p;p_{\bar{p}}\rangle|^2\;\mathcal{L} \tag{2.11}$$

$$\equiv\;\sigma_{\mathrm{int}}\;\mathcal{L} \tag{2.12}$$

$\mathcal{L}$ is usually called (*instantaneous*) *luminosity*, while its integral over time L is called *integrated luminosity*. Greater the luminosity, greater the chance to observe an interaction. For this reason the Tevatron has undergone a series of improvements

during its life in order to increase this fundamental parameter. In the first period of data taking, from 1992 to 1996 (*RunI*) the accelerator, operating at a center-of-mass energy of 1.8 TeV, delivered to each experiment an integrated luminosity in excess of 160 pb^{-1}. *RunI* was followed by a long shutdown period begun in 1996; during this period both the accelerator complex and *CDF* detector were improved in order to increase the number of expected events. After detector and accelerator upgrades, since 2001 a second period of data taking has begun (*RunII*).

Detector improvements will be treated later; as far as luminosity is concerned, the number of bunches per beam was increased from 6 to 36 and finally to 108, while maintaining the average number of particles roughly the same as in *RunI*. In fact increasing the number of particles per beam, the main problem one has to face with is the superposition of multiple $p\bar{p}$ interactions within the same bunch crossing. If this happens, event reconstruction is much more complicated. So the best thing we can do is leaving the number of particles per bunch unchanged while increasing the number of bunches (see figure 2.1).

But a limiting factor in increasing the number of particles per beam is antiproton production: its rate is much smaller than proton one, so during the upgrade period a lot of effort was spent to improve antiproton storage and recycling. In figure 2.2 *RunII* peak luminosity course is shown. The best value obtained so far is $1.40 \cdot 10^{32}$ $cm{-}2s^{-1}$ [1].

2.1.2 The accelerator complex

The accelerator chain is drawn in figure 2.3. We give a brief description of its components and then describe in more detail proton and antiproton production [2].

Cockroft-Walton accelerator

A Cockroft-Walton accelerator is used as the first stage of acceleration. It consists of a multi-step voltage divider which accelerates negative hydrogen ions up to 750 keV through constant voltage steps.

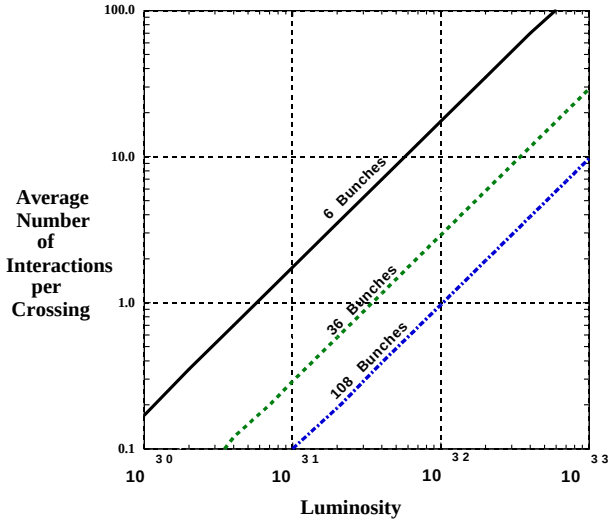

Figure 2.1: Average number of interactions per crossing for different beam conditions: with 108 bunches per beam a greater luminosity can be achieved without increasing the number of interactions per crossing.

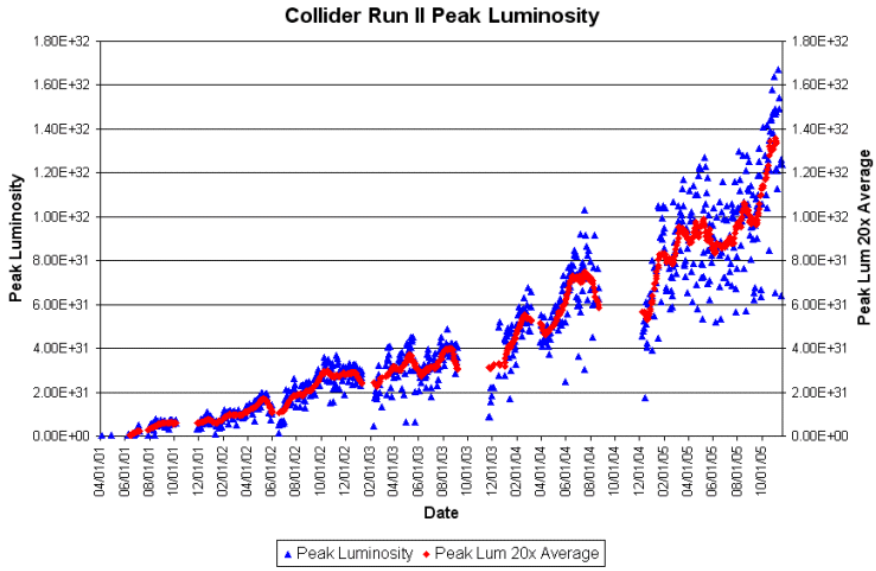

Figure 2.2: *RunII* peak luminosity since April 2001.

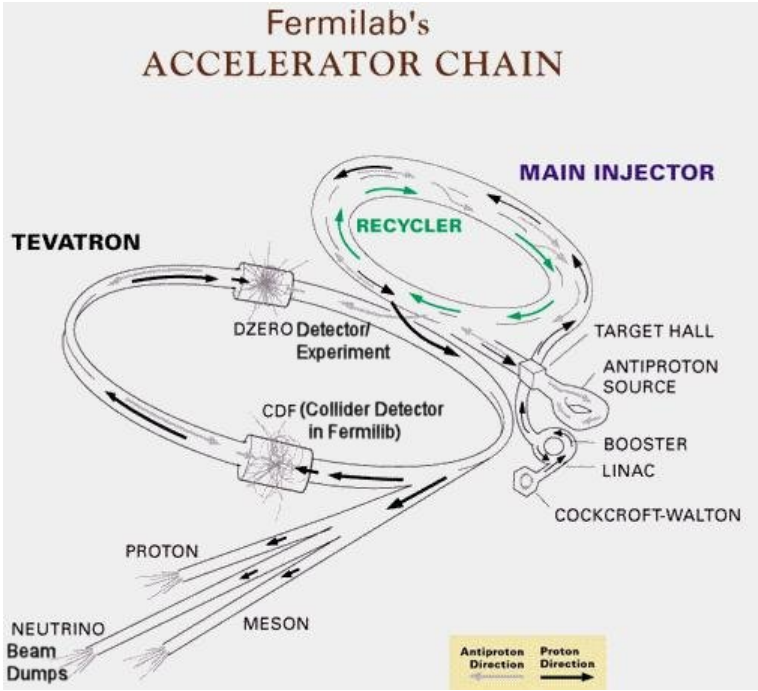

Figure 2.3: Schematic view of the accelerator complex.

Linac

The second stage of the accelerator chain is a linear accelerator, 150 meters long. This device induces an oscillating electric field between a series of electrodes which accelerates up to 400 MeV negative hydrogen ions coming from Cockroft-Walton accelerator.

The Booster

The Booster is the first synchrotron in the chain of Fermilab accelerators. It consists of a series of magnets around a 150 meter circumference. It takes 400 MeV protons obtained from the Linac - after electron stripping - and accelerates them up to 8 GeV.

The Main Injector

The Main Injector is a 3 km circumference synchrotron. It replaces the *RunI Main Ring*, with a higher repetition rate for stacking antiprotons. It accelerates 8 GeV protons from the Booster to either 120 GeV or 150 GeV, depending on their destination: 120 GeV if protons are used to produce antiprotons, 150 GeV if they have to be injected in Tevatron. It is also capable of decelerating antiprotons, a very useful feature to recover unused antiprotons when Tevatron is not in colliding mode.

The Recycler Ring

The Recycler is a storage ring constructed primarily by permanent magnets located above the Main Injector magnets. It receives 8 GeV antiprotons from Main Injector and stores them for an indefinite period of time, next sending them back to the Main Injector.

This device, a *RunII* innovation, allows to recycle $\frac{2}{3}$ of antiprotons that still circulate in the Tevatron at the end of a store[1] (75 % of initial antiprotons), thus increasing by a factor two the average luminosity. Moreover, due to its permanent magnets, it avoids wasting antiprotons in case of power losses.

The Tevatron

The Tevatron is the largest of Fermilab accelerators, with a circumference of about 6 km. It accepts protons and antiprotons from Main Injector and accelerates them from 150 GeV to 980 GeV, using cryogenically cooled magnets. Magnets are cooled by liquid helium (T = 4 K) and provide a 4.5 T field. 36 bunches of protons and antiprotons, with 392 ns separation, circulate in the same beam pipe, kept apart by electrostatic separators. They collide in two intersection points, in correspondence of *CDF* and *DØ* detectors.

[1] A store is when the beam is kept circulating continuously in the Tevatron; can be a HEP store (with protons and antiprotons) or proton-only store for studies and maintenance.

Proton production and acceleration

The first step of protons production is ionization of gaseous hydrogen in H^- ions; ion energy is increased up to 750 KeV by the Cockcroft-Walton accelerator and then up to 400 MeV through Linac. The ion beam so-obtained is focalized on a thin graphite layer in order to strip the electrons off and obtain protons, which are accelerated up to 8 GeV by the Booster. Using radio-frequency cavities protons are collected into bunches which are then injected in the Main Injector.

If protons are needed to produce antiprotons, they are sent to a nickel target with an energy of 120 GeV, while if they are needed for a collision, they are injected in Tevatron with an energy of 150 GeV.

antiproton production and acceleration

Antiprotons production starts in the **Target**, where more than $6 \cdot 10^{12}$ protons strike a nickel target every 2-3 seconds. A lithium lens collects and focuses all particles emerging from the target and a bend magnet separate antiprotons from all other negatively-charged particles. antiprotons produced in this way, with an energy of about 8 GeV, are then subject to stochastic cooling in the **Debuncher Ring** and stored in a storage ring, the **Accumulator**. In the Accumulator, a system of radio-frequency cavities collects antiprotons in bunches that are then injected in the Main Injector and finally in the Tevatron, where they reach their final energy of 980 GeV. antiproton production rate is much smaller than proton one: an antiproton is produced every $\approx 7 \cdot 10^4$ protons striking the nickel target.

2.2 CDF detector

CDF history can be divided in different periods:

- Run 0 (1988 - 1989), ~ 4.5 pb^{-1} of integrated luminosity,
- Run Ia (1992 - 1993), ~ 19 pb^{-1},
- Run Ib (1994 - 1996), ~ 90 pb^{-1},

- Run II (2001 - today) $\sim$ 1 fb^{-1}.

These different periods reflect the improvements that *CDF* has undergone in its life, in order to increase the number of collected events. In fact the number of events for a process, with cross section σ, we expect to detect is

$$N_{events} = \sigma\epsilon \int \mathcal{L}$$

where ϵ is the detector efficiency for that process, $\mathcal{L}$ is the instantaneous luminosity and the integral is over the period of data taking. To increase σ we have to increase the center-of-mass energy, while, as we have seen in the previous section, to increase luminosity we have to act on bunch frequency and antiproton storage and recycling. Increasing ϵ means improving the detector acceptance; from *RunI* to *RunII* there have been several improvements on *CDF* to achieve this goal as we will point out describing the detector.

In the following, after a brief remark on *CDF* coordinate system, we will describe the detector from inside to outside, following the trajectory of a particle [3] [4] [5].

Starting from the interacting point, we immediately find the tracking system that, being immersed in a magnetic field, allows us to measure particle momentum in addition to its trajectory. Then, outside the solenoid coil, we find calorimeters and further on muon detectors. A time-of-flight detector, Cherenkov luminosity counters and diffractive events detectors complete the *CDF* design.

CDF coordinate system

CDF detector with particular attention to coordinate system is shown in figure 2.4: z axis is along the beam line; $+z$ is proton travel direction (east) while $-z$ is is antiproton one (west); the interaction point is at $z=0$. x axis is north-south direction ($+x$ = north, $-x$ = south), y axis is up-down direction ($+y$ = up, $-y$ = down).

r and θ are the radial distance and the polar angle from the beam line; $\theta = 0°$ is in the +z direction, $\theta = 90°$ is straight up and $\theta = 180°$ in the -z direction. Usually instead of θ, a function of θ is used, called *pseudo-rapidity*, defined in the following

way:

$$\eta = -ln(tan\frac{\theta}{2}) \tag{2.13}$$

Particles perpendicular to the beam line have $\eta = 0$. η is preferred to θ because differences in η are Lorentz invariant and after the collision, z-component of the system is not exactly 0, so particles are subject to boosts on z-direction. ϕ is the azimuthal angle around the beam line; $\phi = 0°$ corresponds to north, $\phi = 90°$ to up and $\phi = 180°$ to south.

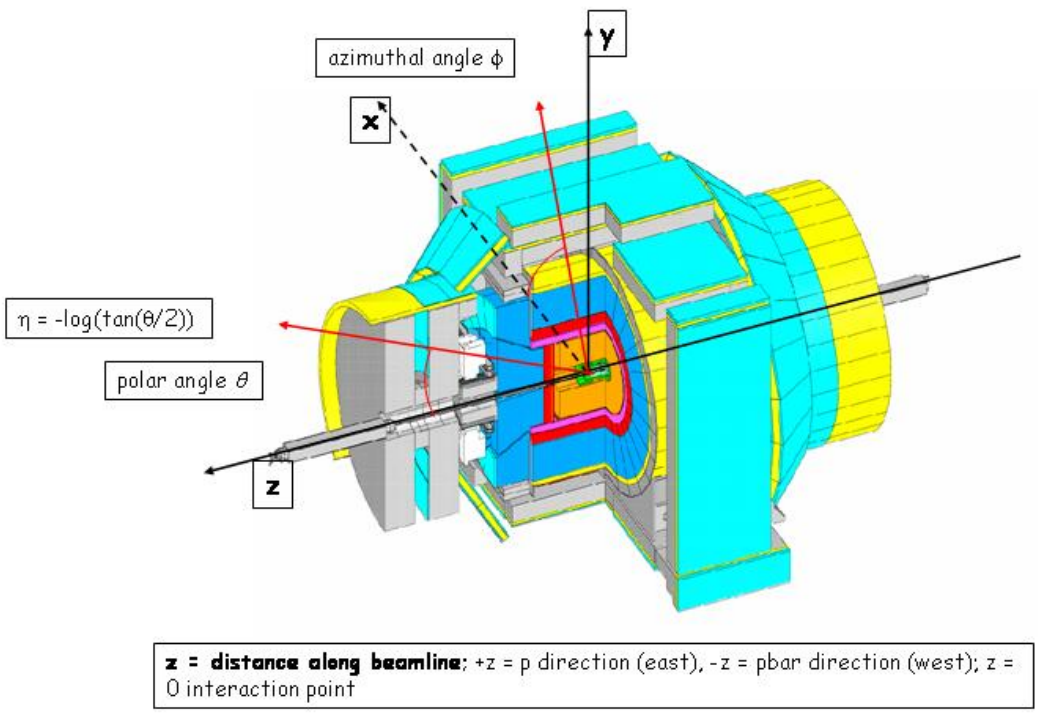

Figure 2.4: View of *CDF* detector and its coordinate system.

Tracking System

CDF inner tracking system consists of three silicon detectors responsible for high precision measurements and a drift chamber devoted to add further information for

track reconstruction; going in more detail and starting from the interaction point we find:

- **Layer 00 (L00)**
  A single sided silicon micro-strip detector located immediately outside the beam pipe, at a radius of approximately 1.6 cm and covering $|\eta| \leq 4.0$ (figure 2.5) [6]. New for *RunII*, it provides more precise track measurements and better b quark tagging efficiency compared to the other tracking devices.

- **Silicon Vertex Detector (SVXII)**
  A double sided silicon micro-strip detector, located outside *L00*, extending from $r = 2.1$ cm to $r = 17.3$ cm and covering $|\eta| \leq 2.0$ (figure 2.6). It consists of 5 layers and its strips are aligned axially to the beam on one side, while in the other side are aligned with a small (1.2°) angle stereo (layers 2 and 4) and with a 90-degree stereo (layers 0, 1 and 3). It provides high precision tracking and secondary vertex detection.

- **Intermediate Silicon Layer (ISL)**
  A double sided silicon micro-strip detector, with axial strips on one side and small angle stereo strips on the other side. It consists of three layers, positioned at different radii (figure 2.7): central layer is at $r = 22$ cm while forward and backward layers are respectively at $r = 20$ cm and $r = 28$ cm. They also have different η coverage: $|\eta| \leq 1.0$ for central layer and $1.0 \leq |\eta| \leq 2.0$ for the others.

- **Central Outer Tracker (COT)**
  An open cell drift chamber with argon-ethane gas in a 50/50 mixture. It's located outside *SVX* from $r = 40$ cm to $r = 137$ cm, covering $|\eta| \leq 1.0$, thus providing tracking in the central regions of the detector (figure 2.8). Its 2520 cells are divided into 8 super-layers, each containing twelve layers of sense wires. The odd super-layers have wires parallel to the beam (*axial* super-layers) while the even have wires at a small (2.0°) stereo angle (*stereo* super-layers). *COT* replaces the *Run I Central Tracking Chamber (CTC)*, featuring

a greater number of sense wires, enhanced stereo coverage and faster drift times.

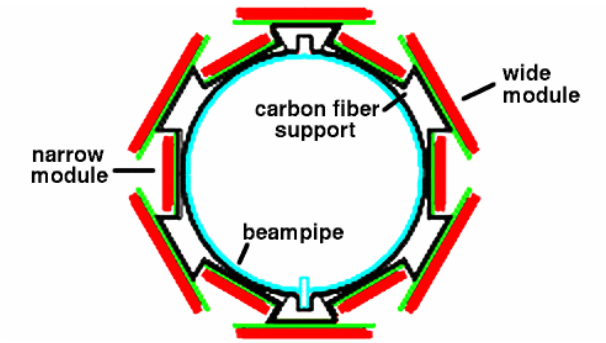

Figure 2.5: Schematic view of *L00*

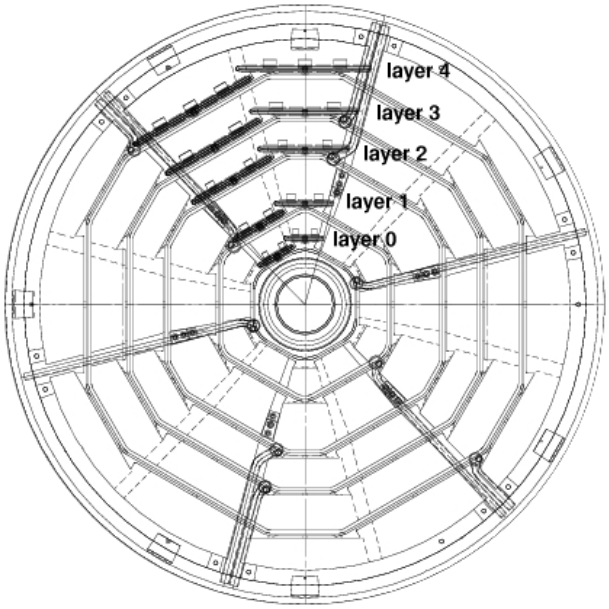

Figure 2.6: Schematic view of *SVXII*

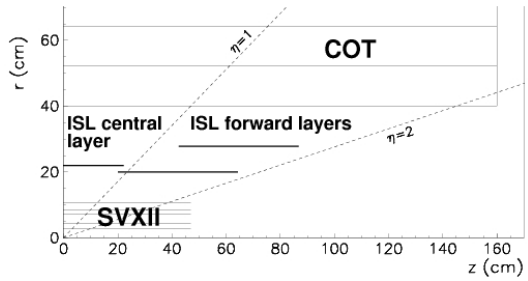

Figure 2.7: *ISL* position.

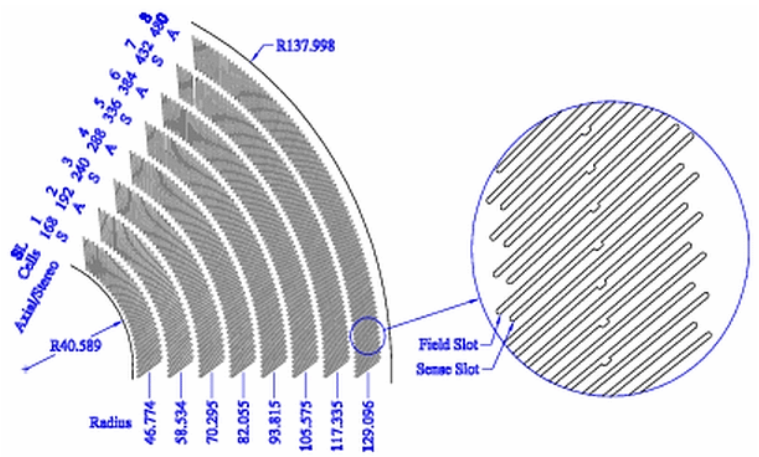

Figure 2.8: COT section: the eight superlayers (left) and the alternation of field plates and wire planes (right)

Time-of-flight detector

Located outside *COT* (at $r = 140$ cm), it provides time-of-flight information to improve particle identification capabilities in the central detector (especially for K-π discrimination) [7]. New for *RunII*, it consists of 216 scintillator bars, each running the length of *COT* (figure 2.9) and arranged cilindrically around it, with a PMT at each end of each bar.

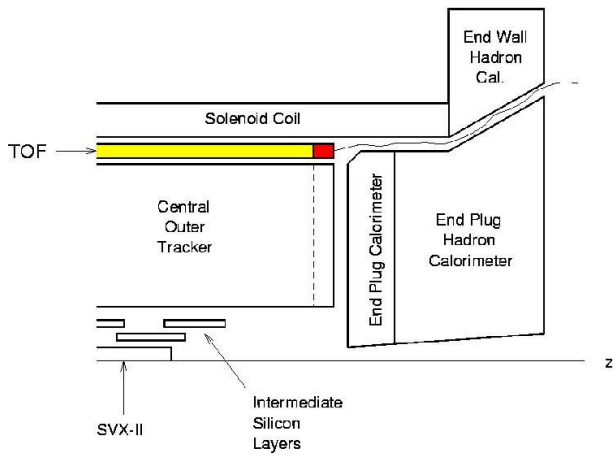

Figure 2.9: Time-of-flight detector position.

Calorimeters

CDF calorimeter system consists of an electromagnetic section followed by a hadronic one. Both sections are made up of sampling calorimeters, with alternate layers of absorber and scintillator material. Scintillator tiles are read out by wavelength shifting fibers (*WLS*) embedded in the scintillators. *WLS* fibers carry the light out to photomultipliers tubes located in the backplane of the calorimeters.

Electron and photon energies are measured in the electromagnetic sections: particles interact in the absorber producing an electromagnetic cascade; the readout system (scintillator and PMT) measures the excitation caused by the cascade and if the calorimeter is thick enough to contain the whole cascade, readout response is proportional to the energy of the particle that produced the cascade. In *CDF*

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

Table 2.1: Calorimeter segmentation

electromagnetic calorimeters are made up of 4 mm lead layers alternate to 4 mm scintillator layers. Energy resolution is about $\frac{14\%}{\sqrt{E}}$ for central calorimeter and $\frac{16\%}{\sqrt{E}}$ for plug one. Electromagnetic calorimeters are organized in *towers* pointing to the interaction point. Calorimeter segmentation is different for different η ranges (see table 2.1): $\Delta\phi$ is 7.5° or 15°, while $\Delta\eta$ goes from 0.1 to 0.6.

Hadron energies are measured in hadronic calorimeters: these have the same geometry and segmentation as electromagnetic ones but differ in the absorber material (iron instead of lead) and absorber layers thickness (3-5 cm). Hadrons interacting with absorber material produce nuclear interaction cascades: these are much longer and broader than electromagnetic and contain also *invisible* energy from neutrons, neutrinos and soft particles, so due to larger energy fluctuations with respect to electromagnetic sections, energy resolution is lower ($\frac{75\%}{\sqrt{E}}$ for central calorimeter and $\frac{80\%}{\sqrt{E}}$ for plug one).

According to the position in the detector (Central, Endwall and Plug) and their specific function (revealing hadronic or electromagnetic energy), *CDF* calorimeters are:

- **Central Electromagnetic Calorimeter (CEM)** A Pb/scintillator sampling calorimeter, 31 layers deep; it is located outside the solenoid in the central part of the detector and provides energy measurements of electromagnetic showers in the central region of the detector ($|\eta| \leq 1.1$).

- **Central Hadronic Calorimeter (CHA)** A Fe/scintillator sampling calorimeter, 32 layers deep. It covers $|\eta| \leq 0.9$

- **Endwall Hadronic Calorimeter (WHA)** A calorimeter similar to CHA, but 15 layers deep and covering $0.8 \leq |\eta| \leq 1.2$.

- **Plug Electromagnetic Calorimeter (PEM)** An electromagnetic calorimeter (Pb/ scintillator) covering the plug region of the detector ($1.1 \leq |\eta| \leq 3.6$). It is located outside the barrel end of COT, one plug on each side.

- **Plug Hadronic Calorimeter (PHA)** Located beyond PEM, it's a Fe/scintillator calorimeter, 23 layers deep.

Plug and endwall calorimeters are visible in figure 2.10; for central ones see figure 2.11.

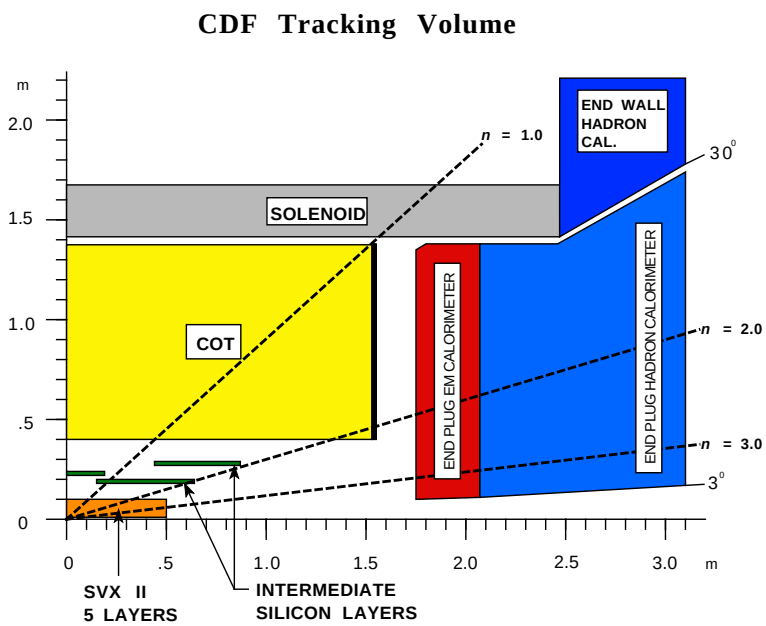

Figure 2.10: Longitudinal view of the tracking system with plug and endwall calorimeters.

Shower and pre-shower detectors

A pre-shower detector, the *Central Pre-Radiate Chamber* (*CPR*) is placed immediately in front of *CEM* (figure 2.11). It consists of scintillator tiles coupled

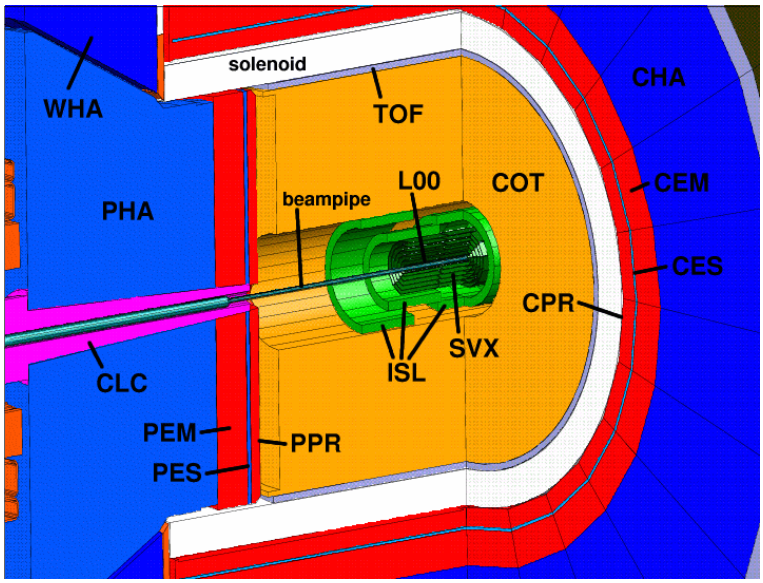

Figure 2.11: Schematic view of the inner parts of *CDF* detector.

to wavelength shifter fibers read out by photomultipliers. With its 3072 channels, *CPR* collects charge deposited by showers originated by interaction of particles with tracking system and solenoid material. It can help in discriminating pions from electrons and photons, because the latter deposite in the chamber a greater amount of energy.

A shower maximum detector, the *Central Electromagnetic Strip* (*CES*), is located in the *CEM*, at a distance r = 184 cm (figure 2.11). It is a drift chamber with strips and wires placed perpendicularly to each other. Strips are placed orthogonally to z-axis and their spacing is 2 cm while distance between wires is 2.5 cm. *CES* gives information about electromagnetic shower position and thus allows a more precise measurement of electromagnetic objets, because it provides a more refined azimuthal segmentation with respect to calorimetric towers.

Muon detectors

Muon detectors are divided in *muon chambers* and *muon scintillators* (figure 2.12). The first are single wire chambers operating in proportional mode and depending

on their location are divided in

- **Central Muon Chambers (CMU)** Located at the outside edge of *CHA*, it covers $|\eta| < 0.6$. It is divided in 2304 cells, organized in 144 modules.

- **Central Muon Upgrade (CMP)** Located along the walls, floor and top of *CDF* detector, it serves as a confirmation of *CMU* tracks. It covers $|\eta| < 0.6$ and consist of 1068 cells, arranged in 4 layers.

- **Central Muon Extension (CMX)** An extension of *CMU* in order to cover $0.6 < |\eta| < 1.0$.

- **Barrel Muon Chambers (BMU)** Located on the outside of the toriods, they provide muon detection in the forward region, covering $1.0 < |\eta| < 1.5$

Muon scintillators provide fast timing and trigger counters for muon chambers. Each muon chamber has its own scintillator:

- **Central Scintillator Upgrade (CSP) and CSP Wall Scintillators (CSW)** Both located outside of *CMP* chambers, *CSP*, new for RunII, covers the top and the bottom of the detector, while *CSW* covers the north and south walls

- **CMX Miniskirt Scintillators (MSX)** Scintillator tiles located in the inner surface of *CMX* miniskirt [2] chambers

- **Barrel Scintillator Upgrade (BSU)** Scintillator tiles located on the outer surface of *BMU* chambers

- **Toroid Scintillator Upgrade (TSU)** Scintillator tiles located on the inner face of the toroids, adding additionale triggering power for forward muons

BSU together with *TSU* are sometimes referred to as *ISU* (*Intermediate Scintillator Upgrade*), while *BMU+BSU+TSU* make the *Intermediate Muon Detector* (*IMU*).

[2] The lower part of *CMX*, due to the floor, has a different geometry; that part is called *miniskirt*

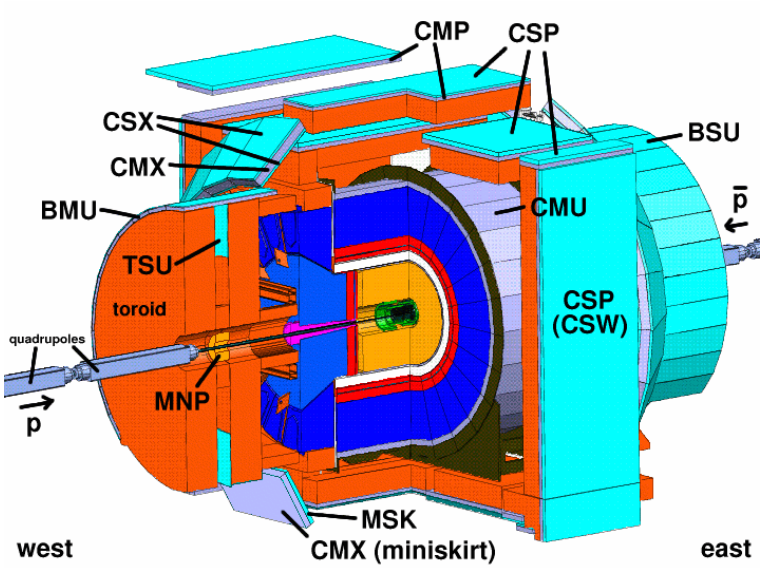

Figure 2.12: Schematic view of the whole *CDF* detector.

Cherenkov Luminosity Counters

To have an estimate of luminosity one has to count the number of $p\bar{p}$ interactions per bunch crossing. This goal is achieved through two *Cherenkov Luminosity Counters (CLC)* [8]. They are located inside the endplug calorimeters, in the forward and backward regions ($3.7 < |\eta| < 4.7$). Each module consists of 48 thin, long, conical, gas filled Cherenkov counters. These counters are arranged around the beam pipe in three concentric layers with 16 counters each and pointing to the center of the interaction region (see figure 2.13). Prompt particles coming from the $p\bar{p}$ interactions traverse the full length of the counter and generate a large amplitude signal in photomultipliers. On the contrary, particles originating from beam halo interactions or from secondary interactions of prompt particles in the detector and beam pipe material, are softer, traverse the counters at large angles with shorter path lenghts and their light suffers a large number of reflections, giving thus a much smaller signal than prompt particles.

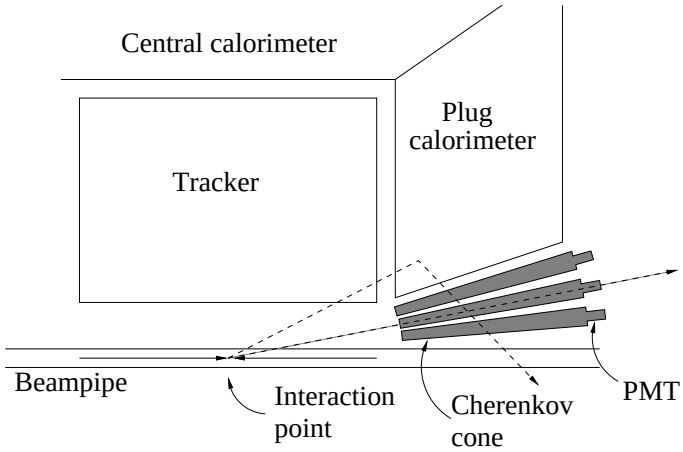

Figure 2.13: Schematic view of the luminosity monitor inside a quadrant of *CDF*.

Forward detectors

In the forward region three detectors provide measurements of diffractive events an forward particles (figure 2.14):

- **Miniplug Calorimeter (MNP)** A Pb/liquid scintillator sampling calorimeter, located inside the central hole of the toroids, and covering $3.6 < |\eta| < 5.2$.

- **Beam Shower Counters (BSC)** Set of scintillator counters located along the beampipe at various distances from the interaction point (four on the west side and three on the east side).

- **Roman Pot Spectrometer (RPS)** It consists of three scintillator fibers pots located along the beampipe on the west side, beyond the last BSC and about 57 m far from the interaction point.

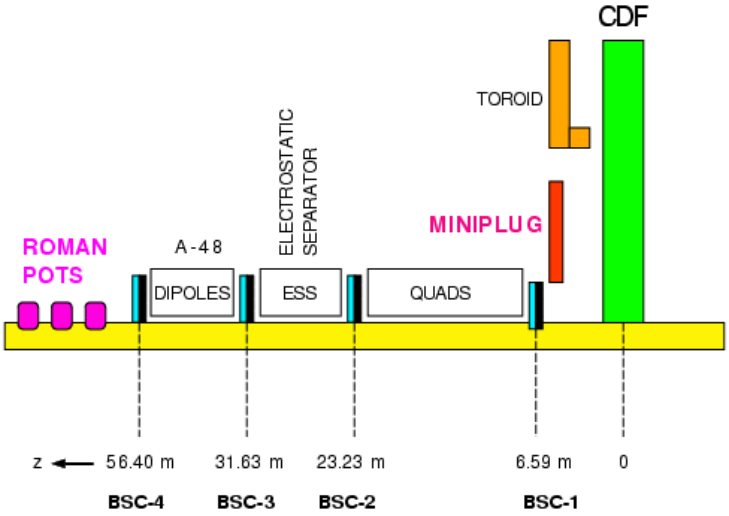

Figure 2.14: Location of *CDF* forward detectors.

Bibliography

[1] ***CDF*'s official luminosity web page**, `http://www-cdf.fnal.gov/~konigsb/lum_official_page.html`

[2] **Fermilab Accelerator Division**, `http://www-bd.fnal.gov`

[3] The *CDF* collaboration, **The *CDFII* techinical design report**, *FERMILAB-Pub-96/390-E*, 1996

[4] **An ace's guide to *CDF* detector**, `http://www-cdfonline.fnal.gov/ace2help/detector_guide/detector_guide.html`

[5] **Lecture series on *CDF* detectors and offline**, `http://www-cdf-fnal.gov/internal/spokes/lecture_detector.html`

[6] Timothy K.Nelson, **The *CDF L00* detector**, *CDF/PUB/SEC_VTX/PUBLIC/5780*, 2001

[7] I.Vila et al.,**A Time-of-flight Detector for CDF**, *International Journal of Modern Physics A, Vol.16,Suppl.1C (2001), 1119-1121*

[8] D.Acosta et al., **The CDF Cherenkov Luminosity Monitor**, *CDF/ANAL/TOP/CDFR/5559*, 2001

Chapter 3

Trigger and DAQ system

At *CDF*, a 3-level trigger architecture was chosen (fig. 3.1). *Level 1* (*L1*) always occurs at a fixed time ($\sim 5\ \mu s$) after the collision: using a custom designed hardware it makes a raw reconstruction of physics objects and takes a decision just counting them. Its input rate is 1.7 MHz while its accept rate is $\sim$ 20 KHz, limited by the following trigger level. *Level 2* (*L2*) is a combination of hardware and software trigger making a decision on a partially reconstructed event. Its average processing time is $\sim 30\ \mu s$ and its accept rate $\sim$ 350 Hz, limited by event building. Full detector resolution is finally used by *Level 3* (*L3*), a purely software trigger that exploits fully reconstructed events, giving a final accept rate of $\sim$ 75 Hz, limited by tape writing.

A trigger unit can be schematized as a decision unit whose inputs are synchronous streams processing info from the different parts of the detector. In the following, we will describe the three levels in more detail [1].

3.1 Level 1

L1 decision unit [2] is fed with inputs from three different streams, one finding calorimeter based objects (*L1CAL*), the second muons (*L1MUON*) and the third *COT* tracks (*L1TRACK*).

- *L1CAL* exploits information from all calorimetric towers, both electromagnetic and hadronic [3]. Calorimeter triggers are divided in *object triggers* (electrons,

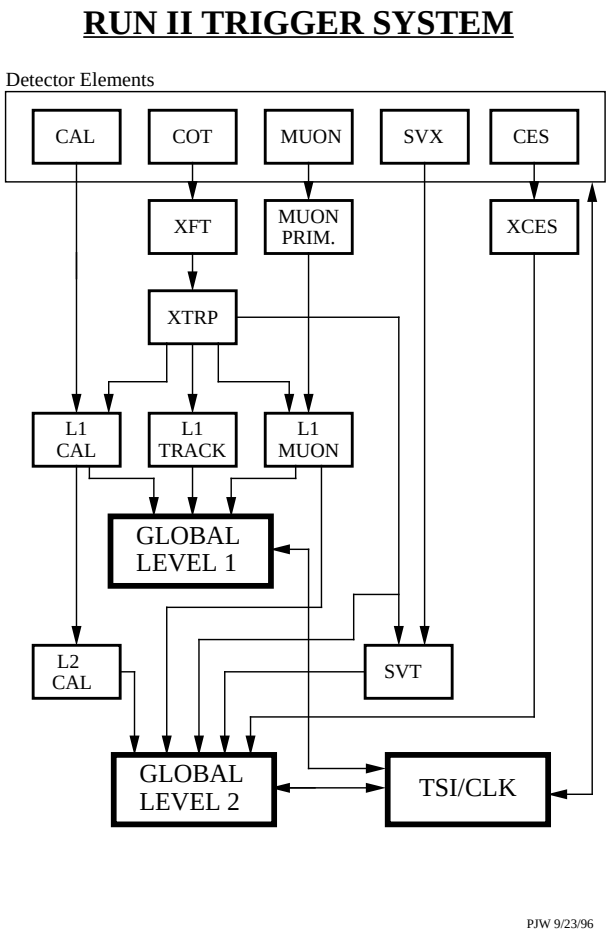

Figure 3.1: The *RunII* trigger system block diagram.

photons and jets) and *global* ones (total and missing transverse energies). The first are obtained applying thresholds to individual towers, electromagnetic only for electrons and photons, also hadronic for jets, whereas threshold for global triggers are applied after summing on all towers.

- *L1MUON* finds single and di-muon objects. It exploits information from muon detectors - wire chambers and scintillators - and from tracking system. Trigger primitives are derived from single hits or coincidence of hits for the scintillators and pairs or patterns of hits (*muon stubs*) on projective wires from wire chambers. Hits and stubs are matched to tracks from *L1TRACK*.

- *L1TRACK* finds *COT* tracks reconstructed by a online track processor, the eXtremely Fast Tracker (*XFT*). The algorithm first digitizes *COT* hits and classifies them as prompt or delayed, depending on their drift time, thus providing a raw timing information in the hits. Then hits are analyzed by *Finder* modules, each working independently from the others and reading the 48 wires contained in four adjacent *COT* cells (fig. 3.2). Each *Finder* determines if its cells contain a segment comparing classified hits with a list of prompt and delayed hits that actually correspond to a line segment. Once the *Finder* has completed its task, the *Linker* looks for combinations of three or four segments that could belong to the same track.

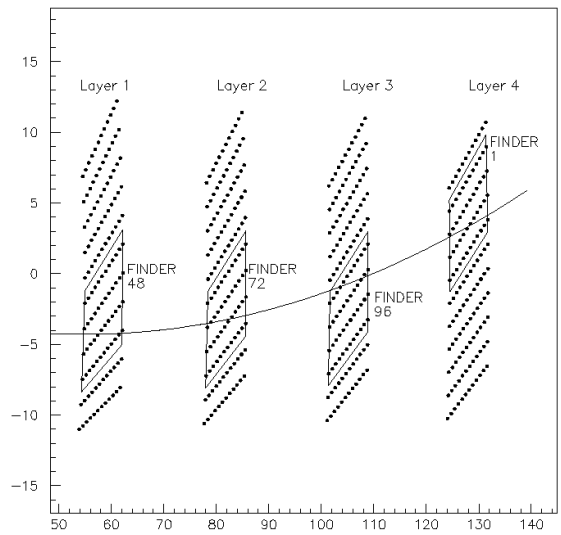

Figure 3.2: A close up view of a track in the COT. All cells in all four axial layers are shown. The relevant Finder in each layer is highlighted

3.2 Level 2

L2 trigger takes a decision on a partially reconstructed event, exploiting data collected from *L1* and from the calorimeter shower maximum detectors (*XCES* and *XPES* triggers, receiving data from central and plug shower maximum detectors

respectively). Simultaneously a hardware cluster finder (*L2CAL*) processes data from calorimeters and a track processor (*SVT*) finds tracks in the silicon vertex detector. After all this information is stored in the processors, the event is examined to find out if trigger criteria is satisfied.

- *L2CAL* is a cluster finding algorithm: it combines calorimeter towers to form clusters, thus allowing a more precise trigger on calorimetric objects (fig. 3.3). Each cluster starts with a tower whose energy is above a *seed* threshold[1]. Once a seed tower is found, all adjacent towers above a second threshold -*shoulder* threshold, lower than the seed one - are added to the cluster. This procedure goes on until no towers adjacent to the cluster have energy above the shoulder threshold. After a complete cluster is found, the total electromagnetic and hadronic energies are calculated and recorded along with the number of towers and the η and ϕ coordinates of the seed tower.

- *XCES and XPES* can help reducing trigger rates for electrons and photon triggers because requiring a cluster above threshold in a shower maximum detector can eliminate background from single-phototube discharge. Moreover, since the spatial resolution in shower maximum detector is much smaller than a calorimeter wedge, matching tracks from *COT* to *XCES* allows a better identification of electrons.

- *SVT* is a track processor designed to detect secondary vertices (fig. 3.4)[4] [5]. Its inputs are the list of axial *COT* tracks found by *XFT* and the data from *SVXII*. First *SVXII* hits are found by a hit finder algorithm and stored in hit buffers; then association between *XFT* and *SVXII* tracks is performed by *Associative Memory* (*AM*), a massive parallel mechanism based on the search of *roads* among the list of *SVXII* hits and *XFT* tracks; a road is a coincidence between hits on four of the silicon layers and *XFT* tracks. Upon receiving a list of hits and tracks, each *AM* chip checks to see if all of the components of one of its roads are present in the list of hits and *XFT* tracks. When *AM* has determined that a road might contain a track, the road's hits are retrieved

[1] typically a few GeV.

from the hit buffer and passed to a track fitter to compute track parameters.

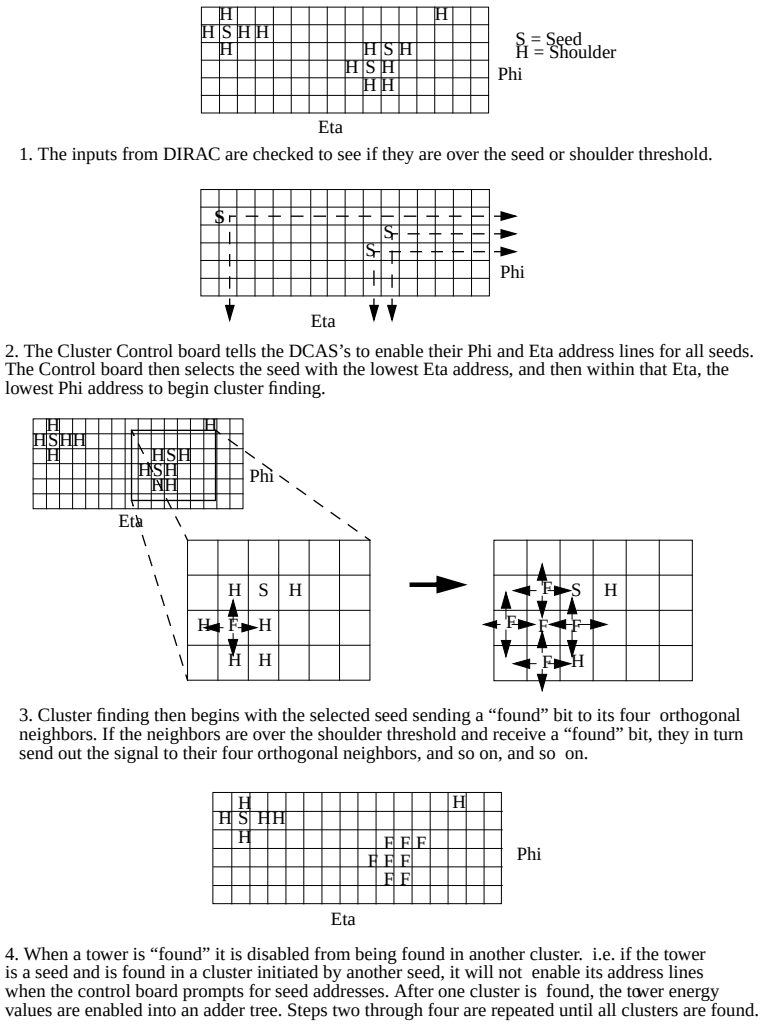

Figure 3.3: *Level2* calorimeter cluster finding.

3.3 Level 3

After being accepted by *L2*, all event data are available to the *Event Builder* which fully reconstructs the event and sends it to the *Level 3* trigger where the final event filtering is done. Events surviving *L3* cuts are written to tape for offline

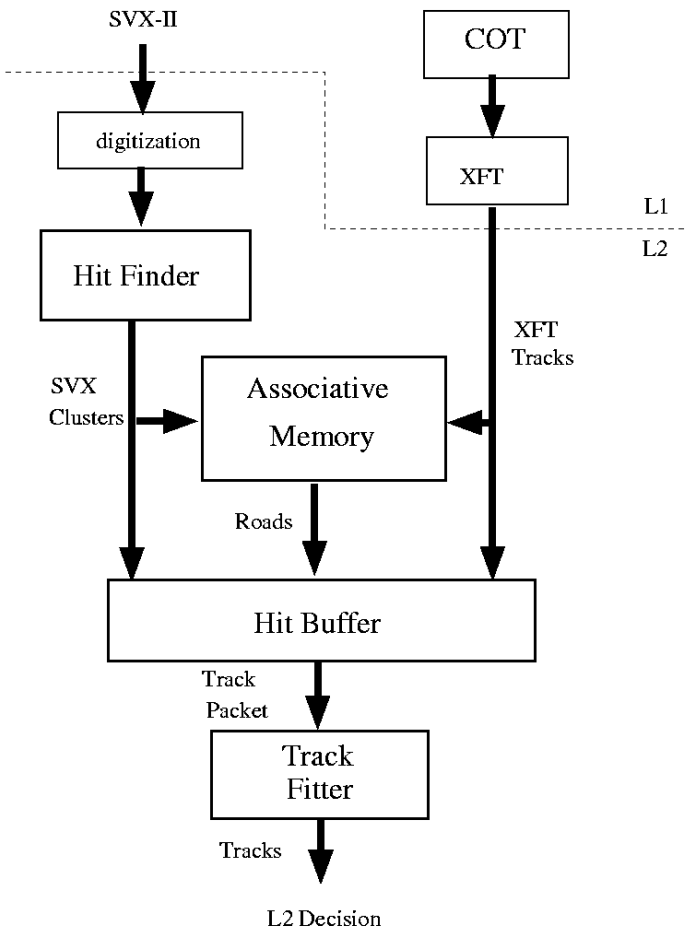

Figure 3.4: Architecture of SVT trigger.

analysis. Event reconstruction and filtering is made on a farm of Linux PC's; in order to ensure a high level of agreement between results in *L3* and offline analysis, the *L3* code is based on the offline production reconstruction code. Offline reconstruction will be described in next chapter.

Bibliography

[1] The *CDF* collaboration, **The *CDFII* techinical design report**, *FERMILAB-Pub-96/390-E*, 1996

[2] Cheng-Ju Lin, **Introduction to *CDF L1/L2* trigger**, `http://www-cdf.fnal.gov/internal/WebTalks/Archived/0411/041118_detector_lecture_cdf_trigger`

[3] H.Frisch et al., **Conceptual design of the L1 calorimeter trigger for the Run II upgrade**, *CDF/DOC/TRIGGER/CDFR/2909*,1996

[4] Roberto Carosi, **The *CDF* online Silicon Vertex Tracker**, `http://www-cdf.fnal.gov/internal/WebTalks/Archived/0411/041118_detector_lecture_cdf_trigger`

[5] SVT Upgrade Group, **Silicon Vertex Trigger Upgrade**, *CDF/DOC/TRIGGER/CDFR/7064*

Chapter 4

Reconstruction of physical objects

In this chapter we will describe how we can identify the particles produced in a $p\bar{p}$ collision starting from the raw outputs of the different parts of the detector. First we will see how information from silicon detectors and *COT* are used to reconstruct charged particle trajectories. Then we will move to the reconstruction of jets of hadronic particles, based on calorimeters. A section will be devoted to the correction of jet energies for different error sources introduced by calorimeters and reconstruction algorithms. After a brief description of the identification of leptons and photons, we will end with the different methods used at *CDF* to identify a jet of particles originated from a b quark.

4.1 Tracks

Track reconstruction is performed using data from silicon tracking system and *COT*. The reconstruction is based on the position of the hits leaved by charged particles on detector components. Combining these hits one can reconstruct particle trajectories.

Which information do we obtain from the different parts of the tracking system? *L00* has microstrips parallel to z axis, in order to measure the ϕ coordinate of a hit. *SVXII* is composed of 5 layers with microstrips on each side of the layers. Each layer has strips parallel to z axis on one side in order to measure ϕ coordinate. Layers 1, 3 and 5 have on the other side strips perpendicular to z axis thus providing a

measure of z position. Layers 2 and 4 have strips with a small angle with respect to ϕ strips. These strips are used to combine the r/ϕ measurement of a track with its z measurement. *ISL* provides r/ϕ measurements of a hit, because its layers have axial strips on one side and small stereo angle on the other. *COT* provides r/ϕ measurement too, because half of its layers have wires parallel to z axis and the others have wires at a small stereo angle.

Different algorithms are devoted to track reconstruction: those used for trigger purposes - *XFT* and *SVT* - are described in section 3; at the end of this section we will describe those developed for offline tracking. But before going in more detail on tracking algorithms, it is worth to spend a few words on particles trajectories.

The whole tracking system is immersed in a 1.4 T magnetic field. Charged particles moving in a homogeneous magnetic field describe a helix trajectory. The helix axis is parallel to the magnetic field [1]. Measuring the radius of curvature of the helix, one can obtain the particle's transverse momentum, while the longitudinal momentum is related to the helix pitch. To describe a helix five parameters are needed, three to parametrize the circle in r/ϕ projection and two to parametrize the trajectory in z. At *CDF* the following parameters are used to describe the helix of a charged particle:

$$\vec{\alpha} = (\cot\theta, C, z_0, D, \phi_0)$$

where

- $\cot\theta$ is the cotangent of the polar angle at minimum approach to the origin,
- C is the half curvature, whose sign, by definition, is the same as the particle,
- z_0 is the position on z axis at point of minimum approach to the helix origin,
- D is the signed impact parameter i.e. the distance between the helix and the origin at minimum approach; if (x_0, y_0) is the center of the circle, then the impact parameter is calculated as

  $$D = Q(\sqrt{x_0^2 + y_0^2} - \rho)$$

  where $\rho = \frac{1}{|2C|} = \frac{1}{2QC}$ is the radius of the circle and Q the particle charge,

[1] At CDF the magnetic field direction is -z.

- ϕ_0 is the direction in r/ϕ of the helix at the point of minimum approach.

In figures 4.1 and 4.2 *CDF* track parameters and the relation between particle charge sign and impact parameter is showed.

Having described the parametrization of a particle trajectory, now we'll turn on the main tracking algorithms developed for offline analysis, the Standalone and the Outside-In algorithms.

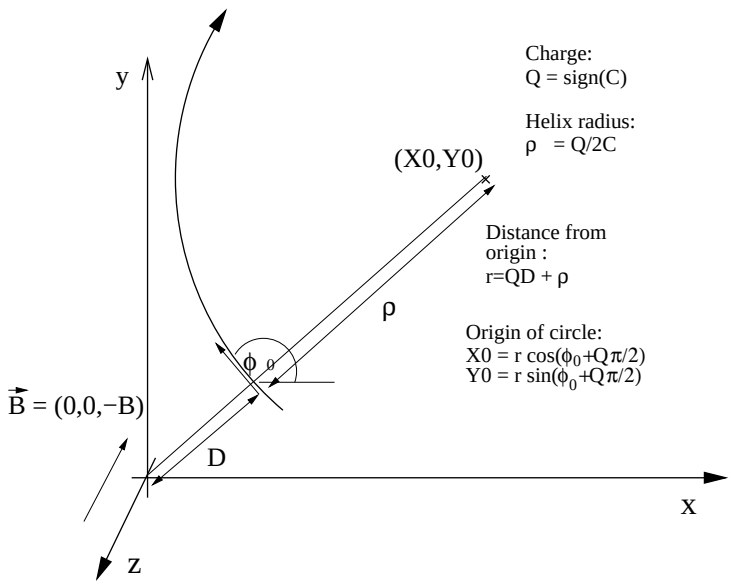

Figure 4.1: Parameters used to describe a track at CDF.

4.1.1 Standalone tracking

Standalone tracking [1] is a strategy to reconstruct tracks in the silicon detector. It consists in finding triplets of aligned 3D hits, extrapolating them and adding matching 3D hits on other layers. This technique is called *standalone* because it doesn't require any input from outside: it performs tracking completely inside the silicon detector.

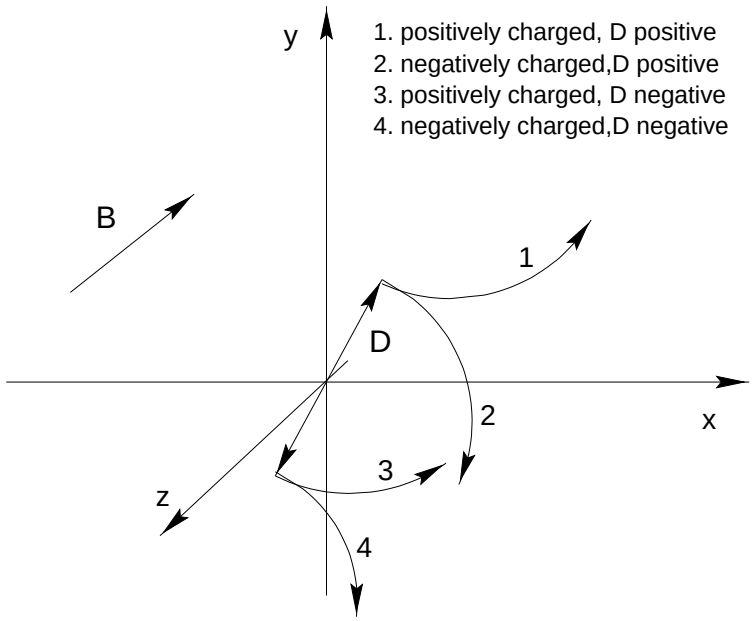

Figure 4.2: Relation between particle charge sign and impact parameter.

First the algorithm builds 3D hits from all possible couples of intersecting axial and stereo[2] strips on each layer. Once a list of such hits is available, the algorithm searches for triplets of aligned hits. This search is performed fixing a layer and doing a loop on all hits in the inner and outer layer with respect to the fixed one. For each hit pair - one in the inner and one in the outer layer - a straight line in the r-z plane is drawn. Next step consists in examining the layer in the middle: each of its hits is used to build a helix together with the two hits of the inner and outer layers.

The triplets found so far are track candidates. Once the list of candidates is complete, each of them is extrapolated to all silicon layers looking for new hits in the proximity of the intersection between candidate and layer. If there is more than one hit, the candidate is cloned and a different hit is attached to each clone. Full helix fits are performed on all the candidates. The best candidate in a clone group

[2]We recall that *axial* strips are parallel to z axis, while *stereo* strips have a small angle with respect to z axis.

is kept, the others are rejected.

4.1.2 Outside-In tracking

The *Outside-In* algorithm [2] exploits information from both *COT* and silicon. The first step is tracking in the *COT*, which starts translating the measured drift times in hits positions; once all *COT* hit candidates in the event are known, the eight superlayers are scanned looking for line segments. A line segment is defined as a triplet of aligned hits which belong to consecutive layers. A list of candidate segments is formed and ordered by increasing slope of the segment with respect to the radial direction so that high momentum tracks will be given precedence. Once segments are available, the tracking algorithm tries to assemble them into tracks. At first, axial segments are joined in a 2D track and then stereo segments and individual stereo hits are attached to each axial track. *Outside-In* algorithm takes *COT* tracks and extrapolates them into the silicon detectors, adding hits via a progressive fit. As each layer of silicon is encountered (going from the outside in), a road[3] size is established based on the error matrix of the track; currently, it is four standard deviations big. Hits that are within the road are added to the track, and the track parameters and error matrix are refit with this new information. A new track candidate is generated for each hit in the road, and each of these new candidates are then extrapolated to the next layer in, where the process is repeated. At the end of this process, there may be many track candidates associated with the original COT track. The candidate that has hits in the largest number of silicon layers is chosen as the *real* track; if more than one candidate has the same number of hits, the χ^2 of the fit in the silicon is used to choose the best track.

4.1.3 Vertex reconstruction

The position of the interaction point of the $p\bar{p}$ collision (*primary vertex*) is of fundamental importance for event reconstruction. At *CDF* two algorithms can be used for primary vertex reconstruction.

[3] A road is a band containing the track.

One is called *PrimVtx* [3] and is used, as an example, in b quark identification. *PrimVtx* starts from the beamline z-position (*seed vertex*) measured during collisions. Then the following cuts (with respect to the seed vertex position) are applied to the tracks:

- $|z_{trk} - z_{vertex}| < 1.0$ cm;
- $|d_0| < 1.0$ cm, where d_0 is track impact parameter;
- $\frac{d_0}{\sigma} < 3.0$, where σ is error on d_0.

Tracks surviving the cuts are ordered in decreasing p_T and used in a fit to a common vertex. Tracks with χ^2 relative to the vertex greater than 10 are removed and the remaining ones are fit again to a common point. This procedure is iterated until no tracks have $\chi^2 > 10$ relative to the vertex.

The second vertex finding algorithm developed in CDF is *ZVertexColl* [4]. This algorithm starts from pre-tracking vertices, i.e. vertices obtained from tracks passing minimal quality requirements. Among these, a lot of fake vertices are present: *ZVertexColl* cleans up these vertices requiring a certain number of tracks with $p_T > 300$ MeV be associated to them. A track is associated to a vertex if it is within 1 cm from silicon standalone[4] vertex (or 5 cm from *COT* standalone vertex). Vertex position z is calculated from tracks positions z_0 weighted by ther error δ:

$$z = \frac{\sum_i \frac{z_{0_i}}{\delta_i^2}}{\sum_i \frac{1}{\delta_i^2}} \tag{4.1}$$

Vertices found by *ZVertexColl* are classified by quality flags according to the number of tracks with silicon/*COT* tracks associated to the vertex. Associated *COT* tracks have shown to reduce the fake rate of vertices thus higher quality is given to vertices with *COT* tracks associated:

- Quality 0: all vertices
- Quality 4: ≥ 1 track with *COT* hits

[4] A *standalone* vertex is a vertex reconstructed completely inside a single detector - silicon detector or *COT* - without any input from other detectors.

- Quality 7: ≥ 6 tracks with silicon hits, ≥ 1 track with COT hits
- Quality 12: ≥ 2 tracks with COT hits
- Quality 28: ≥ 4 tracks with COT hits
- Quality 60: ≥ 6 tracks with COT hits

4.2 Jets

4.2.1 Reconstruction algorithm

Jets are reconstructed applying a clustering algorithm to calorimeter data. This algorithm determines the number of jets in an event, their energies and directions.

- Each calorimetric tower is assigned a vector in the $r\eta\phi$ space: it originates in the interaction point and point towards the tower energy barycenter[5](see figure 4.3). Its module is equal to the total transverse energy of the tower.
- Towers with $E_T > 1$ GeV are ordered according to their decreasing energy and adjacent towers are grouped in pre-clusters.
- A fixed radius cone is drawn around each precluster in the $\eta\phi$ plane ($\Delta r = \sqrt{\Delta\eta^2 + \Delta\phi^2}$); radius size is chosen depending on the physics process under study: for example, for high multiplicity events a smaller value (typically $\Delta r = 0.4$) is preferred, while a greater radius ($\Delta r = 0.7$) is chosen in other cases. The cone axis is the vector with maximum module.
- All vectors falling inside a cone are summed and the axis is re-estimated. This step is repeated until all vectors are assigned to a cone.
- Remaining vectors with $E_T > 0.1$ GeV are associated to the cone containing them and the axis is re-estimated until no new vector is found inside the cone.

[5]The tower barycenter is located at 6 radiation lengths X_0 for electromagnetic calorimeters and 1.5 interaction lengths λ for hadronic ones, i.e. it is assumed that all energy has been released at the average depth of calorimeters.

If two cones overlap, two solutions are possible, depending on how much is the energy they have in common: if the less energetic one has more than 75% of its energy in common with the other, the two cones are replaced by a single one. Otherwise, they are kept distinguished and common vectors are assigned to the closest cone in the $\eta - \phi$ plane.

Finally, summing all vectors in a cone, jet four-momentum is computed. Jet four-momentum *(E,**p**)* is computed assuming that each vector corresponds to a massless particle that deposited all its energy in the tower barycenter. So jet four-momentum is computed summing the four-momenta of all vectors that make up the jet:

$$E = \sum_i (E_i^{had} + E_i^{em}) \tag{4.2}$$

$$p_x = \sum_i (E_i^{had} \sin\theta_i^{had} + E_i^{em} \sin\theta_i^{em}) \cos\phi_i \tag{4.3}$$

$$p_y = \sum_i (E_i^{had} \sin\theta_i^{had} + E_i^{em} \sin\theta_i^{em}) \sin\phi_i \tag{4.4}$$

$$p_z = \sum_i (E_i^{had} \cos\theta_i^{had} + E_i^{em} \cos\theta_i^{em}) \tag{4.5}$$

Starting from the above quantities, jet transverse energy, transverse momentum and rapidity are calculated:

$$P_T = \sqrt{p_x^2 + p_y^2} \tag{4.6}$$

$$E_T = P_T \frac{E}{P} \tag{4.7}$$

$$\eta = 0.5 \ln \frac{E + p_z}{E - p_z} \tag{4.8}$$

Jet quadrimomentum explained so far is computed starting from raw calorimetric energies. Because of calorimeters intrinsic limits, raw energies differ from real deposited energies, thus jet four-momenta need to be corrected, as we will discuss in the next section.

4.2.2 Jet corrections

Jet energies measured in calorimeters suffer from intrinsic limits of both calorimeters and jet reconstruction algorithm. Different particles produce different responses

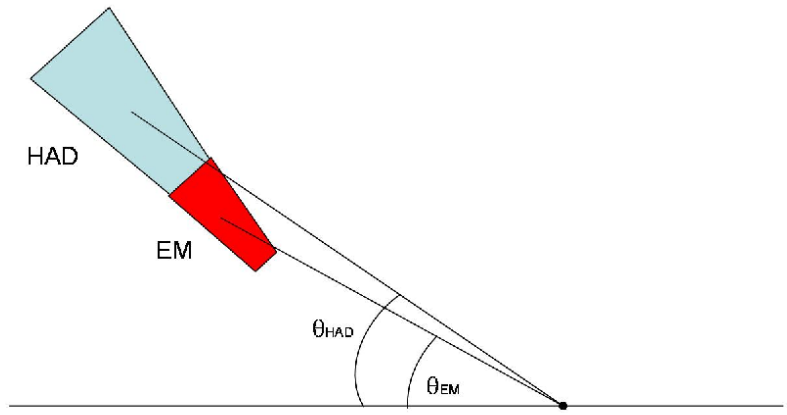

Figure 4.3: Calorimetric tower for jet reconstruction: each tower is assigned a vector originating in the interaction point and pointing towards the tower energy barycenter.

in calorimeters and some of them can fall in uninstrumented regions of the detector. Moreover, calorimeter response to particle energies is non-linear. The jet clustering algorithm, on the other hand, doesn't take into account multiple interactions and energy that can be radiated outside the fixed radius cone. For all these reasons, a set of corrections has been developed in order to scale measured jet energy back to the energy of the particle originating the jet[5].

- **Relative energy scale correction**

  Relative (or η-dependent) jet energy corrections [6] [7] are applied to raw jet energies to correct for non-uniformities in calorimeter response along η. Calorimeter response in each η bin is normalized to the response in the region with $0.2 \leq |\eta| \leq 0.6$, because this region is far away from detector cracks and it is expected to have a stable response[6]. The correction factor is obtained using the dijet balancing method applied to dijet events

  This method starts selecting events with one out of two jets in the region $0.2 \leq |\eta| \leq 0.6$. This jet is defined as *trigger jet*. The other jet is defined as *probe jet*. If both jets are in the region of $0.2 \leq |\eta| \leq 0.6$, trigger and

[6] Central calorimeters are better calibrated and understood.

probe jet are assigned randomly. The transverse momentum of two jets in a 2→2 process should be equal and this property is used to calculate first a p_T balancing fraction $\Delta p_T f$:

$$\Delta p_T f = \frac{\Delta p_T}{p_T^{ave}} = \frac{p_T^{probe} - p_T^{trigger}}{(p_T^{probe} + p_T^{trigger})/2} \tag{4.9}$$

and then a correction factor to make, on average, the probe jet scale equal to trigger:

$$\beta = \frac{p_T^{probe}}{p_T^{trigger}} = \frac{2 + \langle \Delta p_T f \rangle}{2 - \langle \Delta p_T f \rangle} \tag{4.10}$$

In figure 4.4 we show the correction factor as a function of η.

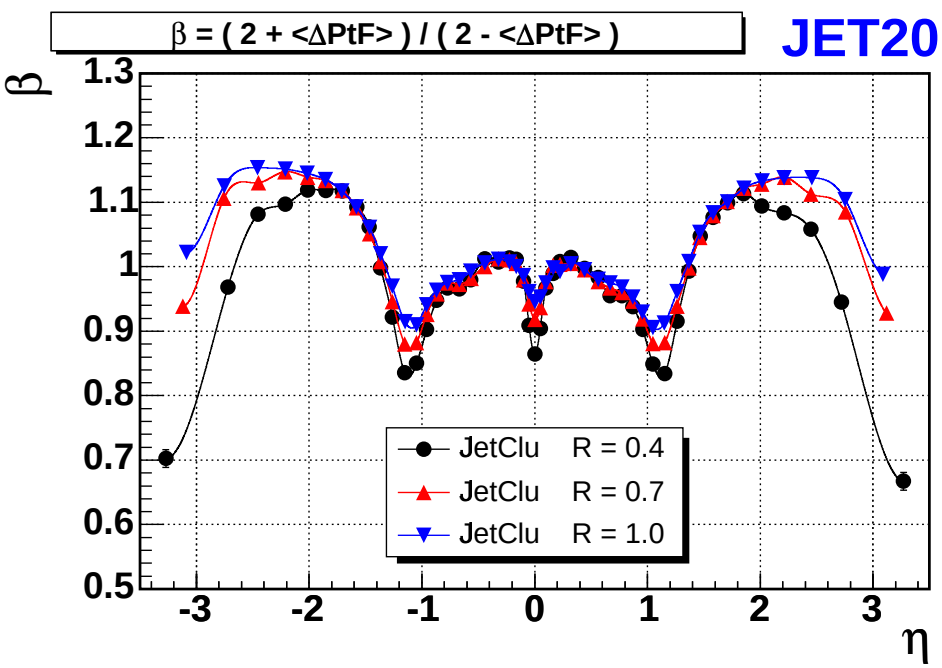

Figure 4.4: Relative energy scale correction factor as a function on η for three different values of cone radii. *Jet20* is the name of the sample on which correction was calculated: it is a sample of events collected with a trigger requiring at *Level 1* one calorimetric tower with energy above 20 GeV.

- **Multiple interactions correction**

  At current instantaneous luminosity[7] and with 36 bunches, we expect on average one hard interaction per beam crossing. However in a fraction of events more than one $p\bar{p}$ interaction can occur. Energy from these non overlapping minimum bias events may fall into the jet clustering cone of the hard interaction causing thus a mismeasurement of jet energy. A correction for this effect is extracted using a sample of minimum bias events [8] [9]: for each event, transverse energy E_T inside cones of different radii (0.4, 0.7 and 1.0) is measured in a region far away from cracks ($0.1 \leq |\eta| \leq 0.7$); then, the distribution of average E_t as a function of the number of quality 12 vertices[8] is fitted with a straight line and the slope of the fitting lines are taken as corrections factors (see fig. 4.5).

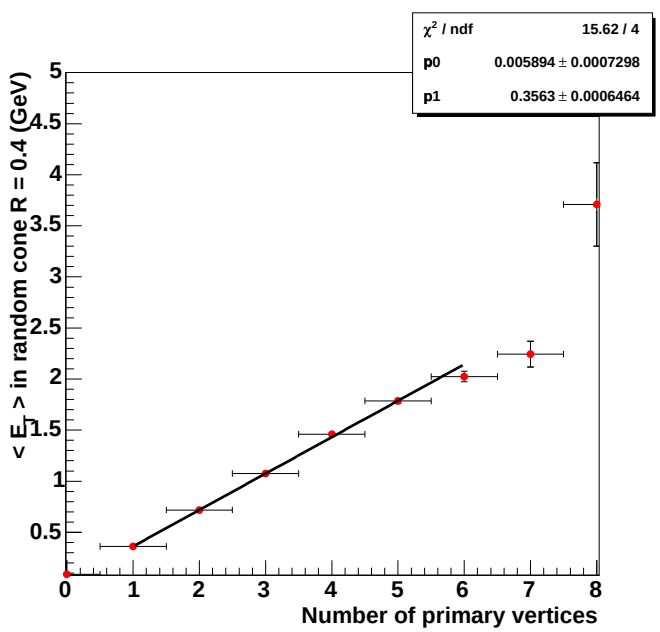

Figure 4.5: Average transverse energy as a function of the number of primary vertices in the event: a correction factor for multiple interaction is extracted from the slope of the fitting line.

[7] $\mathcal{L} \sim 10^{32}$ cm^{-2}s^{-1}

[8] See section 4.1.3 for further details on vertex classification.

- **Absolute energy scale correction**

  A jet contains different types of particles with wide momentum spectra. As calorimeter response to a particle depends on its momentum, position, incident angle and type of particle, jet momentum at hadron level is in general different from its momentum measured at calorimeter level. Absolute energy scale correction converts the calorimeter cluster transverse momentum p_T to the sum of transverse momenta of the particles in the jet cone: calorimeter energy is converted to particle energy [10]. The procedure to extract a calorimeter-to-hadron correction factor is based on the following steps :

  - generate a large sample of MC events with full *CDF* simulation to cover the p_T range from 0 to 600 GeV;
  - create clusters of calorimeter towers and of HEPG particles using the same *CDF* standard cluster algorithm;
  - associate calorimeter-level jets with hadron-level jets;
  - parametrize the mapping between calorimeter and hadron-level jets as a function of hadron-level jets;
  - as a correction factor, extract the probabilities of measuring a jet with p_T^{cal} given a jet with fixed value of p_T^{had}.

  In fig. 4.6 absolute corrections as a function of calorimeter-level jet momentum are showed for different cone sizes.

- **Underlying event correction**

  In a hadron-hadron collision, in addition to the hard interaction that produces the jets in the final state, there is also an *underlying* event, originating mostly from soft spectator interactions. In some of the events, the spectator interaction may be hard enough to produce soft jets. Energy from the underlying event can fall in the jet cones of the hard scattering process thus biasing jet energy measurements. A correction factor for such effect has been calculated using a sample of minimum bias events as for multiple interaction correction, but selecting only those events with one vertex [9]. For each event, transverse

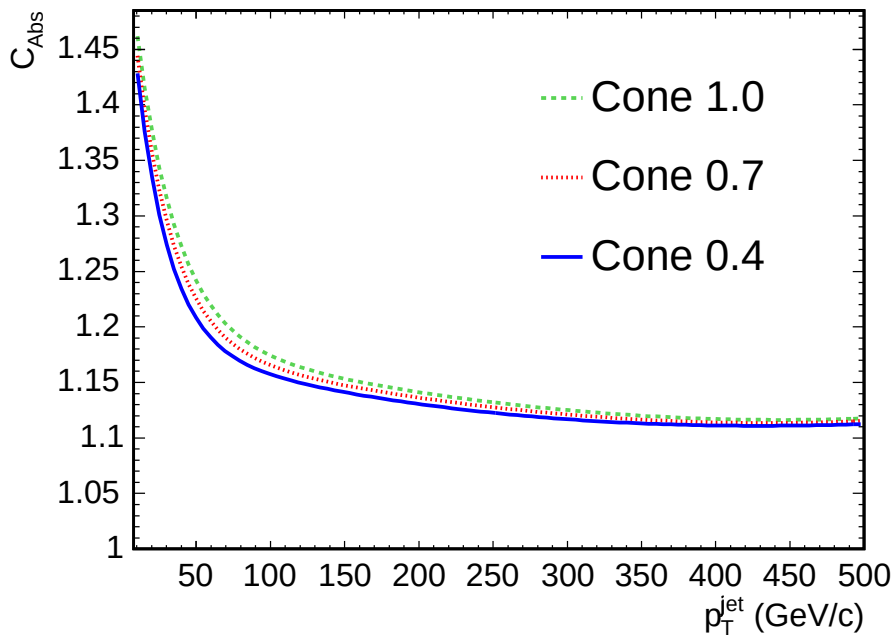

Figure 4.6: Absolute jet energy scale corrections for jet with different cone sizes as a function of jet momentum.

energy E_T inside cones of different radii (0.4, 0.7 and 1.0) is measured in a region far away from cracks ($0.1 \leq |\eta| \leq 0.7$). The correction factor is extracted from the mean values of E_T distribution (fig. 4.7).

- **out-of-cone correction**
  The jet clustering may not include all the energy from the initiating partons. Some of the partons generated during fragmentation may fall outside the cone chosen for the clustering algorithm. This energy must be added to the jet to get the parton level energy. A correction factor is obtained using MC events [11]: hadron-level jets are matched to partons if their distance in the $\eta - \phi$ plane is less than 0.1. Then the difference in energy between hadron and parton jet is parametrized using the same method as for absolute corrections (see fig. 4.8).

We have seen different corrections that account for different sources of jet energy mismeasurement. Depending on the physics analysis, all of them or just a subset

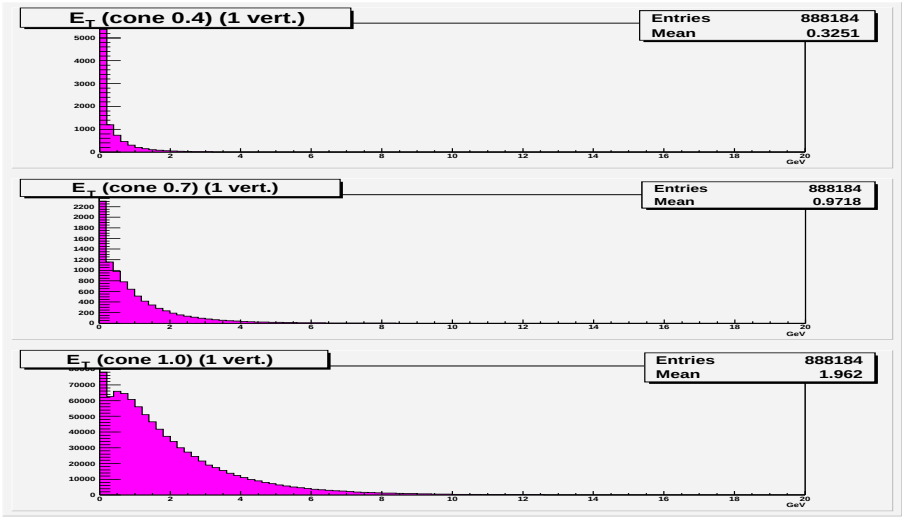

Figure 4.7: E_T in cones of different radii in minimum bias events with just one quality 12 vertex.

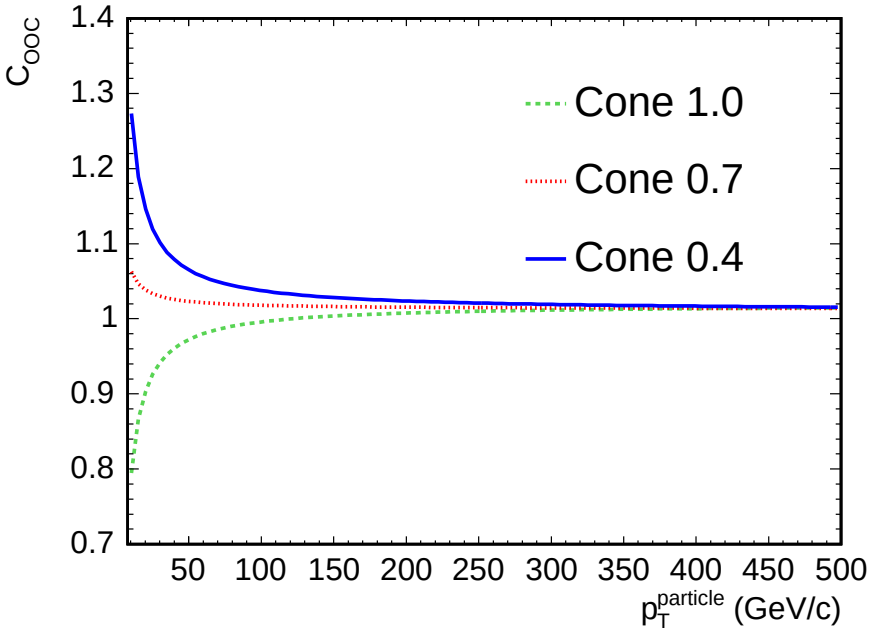

Figure 4.8: Out-of-cone correction factor as a function of jet momentum for different cone sizes.

can be applied.

Corrections are applied to the raw measured jet momentum according to:

$$P_T(R, P_T, \eta) = (P_T^{raw}(R) \cdot f_\eta(R, P_T, \eta) - M_{p\bar{p}I}(R)) \cdot f_{abs}(R, P_T) - UE(R) + OOC(R, P_T) \tag{4.11}$$

where R is the clustering cone radius, P_T is the raw energy measured in the cone and η the pseudorapidity of the jet; f_η, $M_{p\bar{p}I}$, f_{abs}, UE and OOC are respectively relative, multiple interactions, absolute, underlying event and out-of-cone correction factors.

4.3 Leptons

4.3.1 electron

Being a charged particle, an electron traversing the detector first leaves a track in the tracking system and then looses its energy in the electromagnetic calorimeter. So a good electron candidate is made of a cluster in an electromagnetic calorimeter (central or plug) and one or more associated tracks; if available, shower max cluster and preshower clusters can help electron identification. The shower has to be narrow and well defined in shape, both longitudinally and transversely. The ratio between hadronic and electromagnetic energies has to be small and track momentum has to match electromagnetic cluster energy [12] [13].

4.3.2 muon

Muons can leave a track in the tracking system and then loose all their energy in muon detectors. Muons are reconstructed using the informations coming from muon chambers (*CMU, CMP, CMX, BMU*) and muon scintillators (*CSP, CSX, BSU, TSU*). The first provide measurements of drift time, which is then converted to a drift distance, i.e. a distance from the wire to a location that the muon has occupied in its flight, in the plane perpendicular to the chamber sense wire. Scintillators, on the other hand, only produce timing information. The output of chambers and scintillators produce muon hits. Fitting these hits a muon track segment, a *stub*, is obtained. Finally *COT* tracks are extrapolated to the muon chambers and matched to muon stubs in the $r - \phi$ plane [14] [15].

4.3.3 tau

Tau lepton can decay leptonically into electron or muon (and the corresponding neutrinos) or semileptonically into charged and neutral pions; the first case is not distinguishable from a leptonic decay from W, while the second has a precise signature: tau decays preferably into 1 or 3 charged pions (One/Three prong event) and in most cases neutral pions are present. So a well isolated jet with low track

multiplicity and neutral pions is a good tau candidate.

τ reconstruction procedure exploits information from calorimeter and tracking systems. One looks for an isolated narrow cluster above a certain energy threshold and then matches it to *COT* tracks. For more details on τ reconstruction, see section 6.1.

4.3.4 neutrino

Neutrinos produced in a $p\bar{p}$ collision escape the detector without leaving any sign, but their presence can be inferred from momentum conservation.

Colliding proton and antiproton have nearly opposite momenta so total vector sum of momenta of final state particles must be zero. As many particles escape down the beam pipe, we can't apply *total* momentum conservation; but we are allowed to apply *transverse* momentum conservation , i.e. momentum conservation in the plane perpendicular to the beam, if we consider that particles escaping down the beam pipe have negligible transverse momentum. If the sum of transverse momenta of detected particles is significantly different from zero, we can attribute this discrepancy to the presence of one or more undetected neutrinos whose momentum is therefore the opposite of the total detected transverse momentum. As neutrino mass is negligible at a center-of-mass energy of 1.96 TeV, undetected neutrinos total momentum, called *missing* E_t ($\not{E}_T$), is calculated as the opposite of the vector sum of all calorimetric towers, both electromagnetic and hadronic:

$$\vec{\not{E}}_T = -\sum_{towers} (E_i \sin\theta_i)\vec{n}_i$$

where E_i is the energy of the *ith* tower, θ_i is the polar angle of the line pointing from the interaction point to the *ith* tower and $\vec{n}_i$ is the transverse unit vector pointing from the interaction point to the center of the tower [16].

4.4 Photons

A photon traversing CDF detector interacts only with electromagnetic calorimeter and shower max detector. Thus photon identification starts looking for clusters

of energy around a seed tower with energy greater than 3 GeV. Total energy of the hadronic towers located behind the photon cluster has to be negligible with respect to the photon cluster energy. Photon cluster isolation is required: the difference between photon energy and the energy in a 0.4 cone around the seed tower has to be less than 15% of photon energy. Moreover the sum of transverse momenta of all tracks pointing to the 0.4 cone is required to be less than 2 GeV/c. Electromagnetic shower shape shall be transverse and no matching tracks have to be present. The line connecting the primary event vertex to the CES shower position determines photon direction [17].

4.5 b quark tagging

As top quark decays almost exclusively in a W boson and a b quark and considering that the process $t\bar{t} \to b\bar{b}W^+W^-$ has a cross section $\sim 10^{-3}$ times smaller than the total inelastic cross section, an efficient b quark tagging algorithm is essential to reduce signal to background ratio. b tagging relies on two properties of b quark: it has a long lifetime, so it travels a non-negligible distance before it decays[9] and it decays semi-leptonically through the decay mode $b \to l\nu c$. From these properties it is possible to construct algorithms that will tag a jet if it comes from a b quark (*b-jets*). At *CDF* three b-tagging algorithms have been implemented so far:

- *SecVtx* algorithm;
- *JetProbability (JP)* algorithm
- *Soft Lepton Tag (SLT)* algorithm

4.5.1 SecVtx algorithm

SecVtx algorithm [18] [19] can tag a b-jet exploiting the fact that its vertex is displaced from the primary vertex, due to long b quark lifetime.

[9] $\tau \sim 1.5$ ps, so for a typical $p_T \sim 40$ GeV, a b quark can travel for a distance of some millimeters before it decays.

The algorithm starts from COT and silicon tracks inside a cone and as a first step, using as discriminating variable their impact parameter, it removes tracks identified as K_S, Λ or γ daughters, or consistent with primary vertex or too far from it. Then a three dimensional common vertex constrained fit is performed using two tracks; if the $\chi^2 < 50$, the two tracks are used as seed to find other tracks that point towards the same secondary vertex. If at least three tracks are found to be compatible with a secondary vertex, they are considered *potential tag*. The *potential tag* becomes a b-tag if it survives the following cuts:

1. $|L_{xy}| < 2.5$ cm, where L_{xy} is the decay length of the secondary vertex (see fig 4.9); this cut helps rejecting conversions from the first layer of $SVXII$;
2. $\frac{L_{xy}}{\sigma_{L_{xy}}} > 3$;
3. if M_{trk} is the invariant mass of the tracks, $|m_{Ks} - M_{trk}| > 0.01$ GeV and $|m_\Lambda - M_{trk}| > 0.006$ GeV
4. $|L_{xy} \cdot (M_{trk}/p_{trk})| < 1$ cm.

The tags are classified depending on where the secondary vertex is located with respect to the jet cone axis. Secondary vertices on the same side of the interaction point as the jet cone axis are *positive* tags, otherwise they are classified as *negative* tags. Negative tags can arise from tracks mismeasurements (see fig. 4.10).

4.5.2 JetProbability algorithm

This algorithm exploits the information of the tracks associated to a jet to determine the probability that the jet comes from the primary vertex [20]. The probability distribution is uniformly distributed for a jet arising from the primary vertex , while it shows a peak on zero for a long-lived jet (fig. 4.11). The probability is based on track impact parameters and on their uncertainties.

All tracks associated to a primary jet should originate from the primary vertex; due to finite tracking resolution, these tracks are reconstructed with a non-zero impact parameters that have equal probability to be either positive or negative signed. The width of the impact parameter distribution from these tracks is solely due to

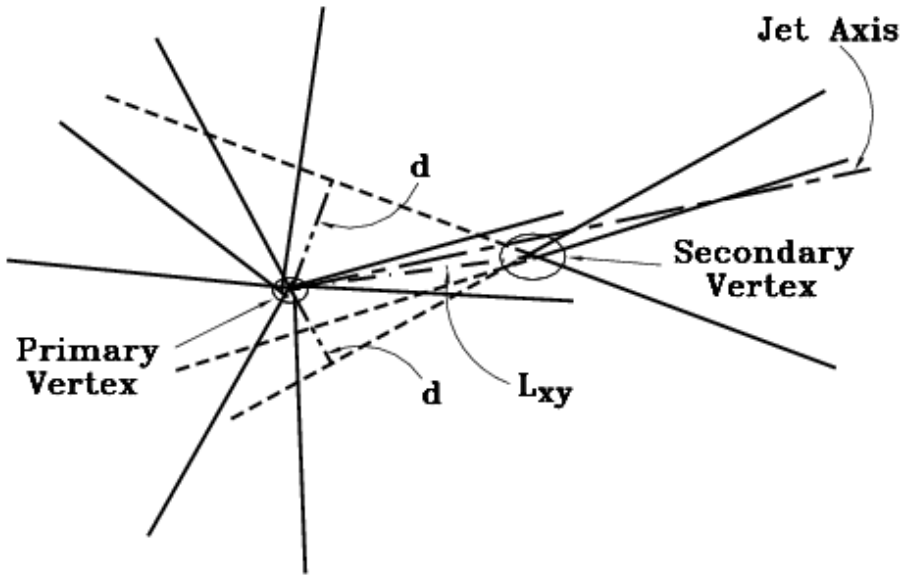

Figure 4.9: Schematic 2D view of an event containing a jet with a secondary vertex.

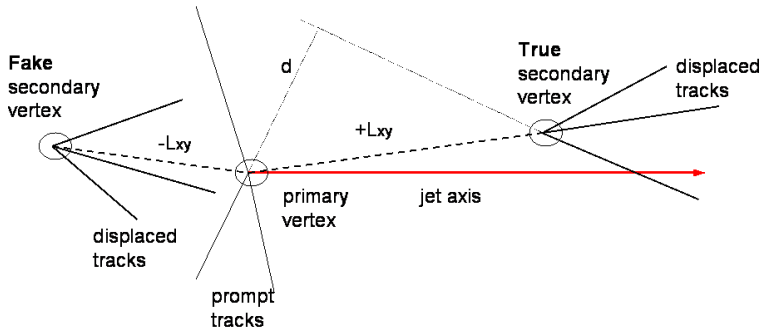

Figure 4.10: Real and fake secondary vertices as seen in the transverse plane.

the tracking detector resolution and multiple scatterings. A long-lived particle, on the other hand, will travel some distance along the jet direction before decaying and its decay products will preferentially have positive signed impact parameter (fig. 4.12). To minimize the contribution of mismeasured tracks, the final probabi-

lity is computed using the signed impact parameter significance (ratio of the impact parameter to its measured error) instead of the parameter itself. Given a track with impact parameter significance S_{d_0}, the probability that a track from a light quark has a larger value of S_{d_0} is calculated. Combining probabilities for all tracks in a jet, one obtains a per jet probability. By construction, this probability is flat for jets originating from light quarks, peaked towards zero otherwise.

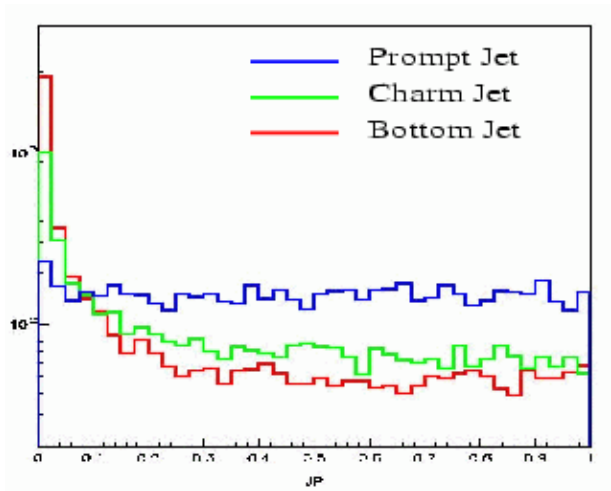

Figure 4.11: Jet probability distribution from prompt, charm and bottom jets.

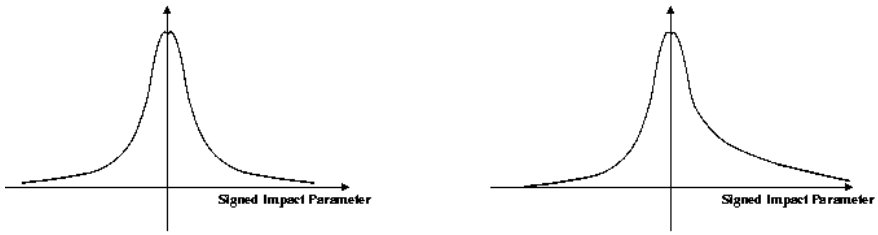

Figure 4.12: Signed impact parameter distribution for tracks from primary vertex (left) and from secondary vertex (right).

4.5.3 Soft Lepton Tag algorithm

This algorithm is based on the fact that ~ 20 % of b quarks decay to muons. In general, muon identification relies on the presence of a stub in the muon chambers, associated with a track and minimum ionization energy deposition in the electromagnetic and hadronic calorimeters. Muons coming from b quarks are not isolated so information from calorimeters can't be used. Moreover, multiple scattering of muons in the material of *CDF* detector has to be taken into account: it causes a deflection in the muon path that ranges from about a few millimeters for a momentum of 2 GeV/c to about half a meter for a 50 GeV/c muon. *SLT* algorithm procedure can be divided in two steps [21]:

1. The first step consists in finding *taggable* tracks, i.e. tracks that could have been left from muons. To take into account the fact that the muon may non have had enough momentum to reach the muon chambers, tracks whose momentum is lower than 2.8 GeV are rejected. Moreover, it has to point to the fiducial volume requested by the user: one can choose to define a track taggable if it points to a volume limited by the physical edges of the muon chambers, or a distance 3 σ_{MS} inside/outside the physical edges, where σ_{MS} is the standard deviation of the maximum deflection expected from multiple scattering through the detector's material.

2. If a track is taggable and has a stub associated to it, the algorithm performs a likelihood comparing all the available information about the muon candidate with the expected values, obtained from a sample of muons derived from data. Besides variables from muon detectors, for the likelihood one can use also some track quality information, like the number of *COT* hits, the beamline-corrected impact parameter and the track z_0 position.

Bibliography

[1] Paolo Gatti, **Performance of the new tracking system at *CDF II***, *CDF/THESIS/TRACKING/PUBLIC/5561*, 2000

[2] W.M.Yao, K.Bloom, **"Outside-In" silicon tracking at *CDF***, *CDF/TRACKING/DOC/CDFR/CDF5991*, 2002

[3] H.Stadie, W.Wagner, T.Müller, **Vxprim in RunII**, *CDF/DOC/TRACKING/CDFR/6047*, 2002

[4] J.-F.Arguin, B.Heinemann, A.Yagil, **The z-Vertex Algorithm in Run II**, *CDF/ANAL/TRACKING/CDFR6238*, 2002

[5] CDF collaboration, Jet Energy Group, **Jet Energy Corrections at CDF**, *CDF/PUB/JET/PUBLIC/7543*, 2005

[6] A.A.Bhatti, K.Hatakeyama, **Relative jet energy corrections using missing E_T projection fraction and dijet balancing**, *CDF/ANAl/JET/CDFR/6854*, 2004

[7] A.A.Bhatti, K.Hatakeyama, **Update on the relative jet energy corrections**, *CDF/ANAL/JET/CDFR/7354*, 2005

[8] B.Cooper, M.D'Onofrio, G.Flanagan, **Multiple interaction corrections**, *CDF/PHYS/JET/CDFR/7365*, 2004

[9] Jean-Francois Arguin, Beate Heinemann, **Underlying event corrections for Run II**, *CDF/ANAL/JET/CDFR/6293*, 2003

[10] Anwar Bhatti, Florencia Canelli, **Absolute corrections and their systematic uncertaities**, *CDF/DOC/OFFLINE/CDFR/5456*, 2003

[11] A.Bhatti, F.Canelli, L.Galtieri, B.Heinemann **Out-of-Cone Corrections and their Systematics Uncertainties**, *CDF/ANAL/TOP/CDFR/7449*, 2005

[12] Robert G.Wagner, **Electron identification for Run II: algorithms**, *CDF/DOC/OFFLINE/CDFR/5456*, 2003

[13] C.Issever, C.Mills, A.Robson, B.Wagner, **Plug Electron Baseline Cuts Defined in Summer 2003 and Cut Efficiencies**, *CDF/DOC/ELECTRON/CDFR/6789*, 2004

[14] James N.Bellinger et al., **A guide to muon reconstruction and software for Run 2**, *CDF/DOC/COMP_UPG/CDFR/5870*, 2002

[15] V.Martin, **High-p_T Muon ID Cuts and Efficiencies for use with 5.3.1 Data and 5.3.3 MC**, *CDF/DOC/MUON/CDFR/7367*, 2005

[16] **Home page of the MET Working Group**, `http://www-cdf.fnal.gov/internal/physics/exotics/susy/r2a/met`

[17] **Baseline Analysis Cuts for High Pt Photons V2.3**, `http://www-cdf.fnal.gov/internal/physics/photon/docs/cuts.html`

[18] D.Glenzinski et al., **A detailed study of the SECVTX algorithm**, *CDF/ANAL/SEC_VTX/CDFR/2925*, 1995

[19] C.Neu et al., **SecVtx Optimization Studies for 5.3.3 Analysis**, *CDF/ANAL/SEC_VTX/CDFR/7585*, 2005

[20] D.Acosta et al., **Introduction to Run II jet probability heavy flavor tagging**, *CDF/DOC/CDF/CDFR/6315*

[21] L.Cerrito, A.Taffard, **A Soft Muon Tagger for Run II**, *CDF/DOC/TOP/CDFR/6305*, 2003

Chapter 5

Neural Networks

The mail goal of our analysis is extracting $t\bar{t} \rightarrow \tau + jets$ events among the huge background due to QCD events and to do this we use neural networks having as inputs the variables that most discriminate between signal and background. Before going in detail on our analysis, in this chapter we will first give an introduction on the main concepts about neural networks and then describe the learning strategy we use, the *Reactive Tabu Search* (*RTS*).

5.1 Brief introduction to the main concepts

An Artificial Neural Network may be defined a computing system of Von Neuman type aimed at approximating efficiently a given mapping from a subset $\mathcal{D}$ of $\mathbb{R}^n$ into $\mathbb{R}^m$ on the basis of known examples. The examples form a so called *training set*. The mapping is defined in terms of a number of parameters, called *weights* and organised in a hierarchical structure, named *architecture*. The process of getting the approximated mapping (the *learning* process) is usually a *trial and error* one, in fact searching for a minimum of a positive definite *error function* confronting expected values of the mapping against those actually given by a specific configuration of the neural network. Such error function (also called *cost function*) is a parametric one, depending on the weights for a fixed architecture and evaluated over all of the given examples.

The weights compose a varying vector in some $\mathbb{R}^k$ space, where the error function

defines a hyper-surface that we call *surface error* : we may depict the learning process as building trajectories on the error surface, searching there for a (optimal) minimum.

According to [1], as a model we may refer to the classical case of approximating a one-dimensional function using polynomials or truncated Taylor series

$$F(x) = w_0 + w_1(x - x_0) + w_2(x - x_0)^2 + \cdots + w_2(x - x_0)^k \,, \tag{5.1}$$

pictorially represented in fig.5.1: $z \to 1$, $z \to z^1$, $z \to z^2$, ... , $z \to z^k$ in the circles (or *nodes*) are in a sense *cellular automata* or *primitive functions* of a computing network, whose *weights* are the coefficient a_0, a_1, a_2, ... a_k. The output node adds all incoming information and produces the final polynomial value.

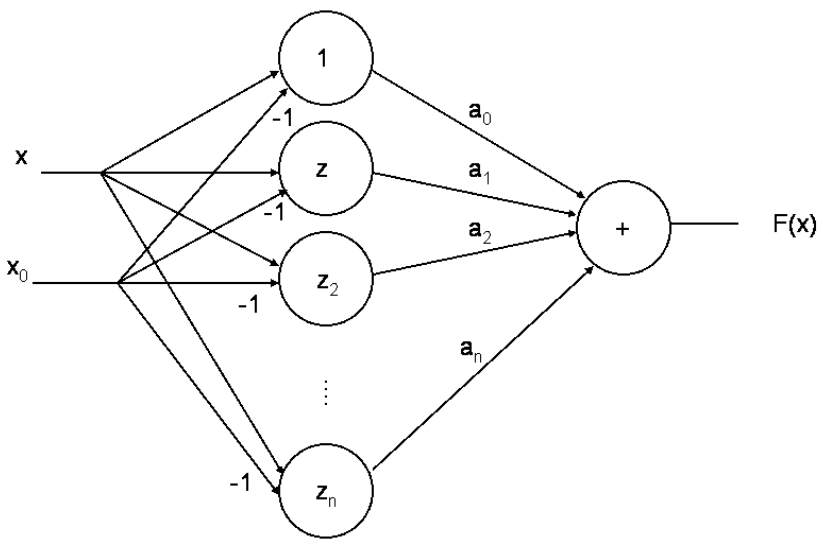

Figure 5.1: A polynomial network.

Given the analytical expression of a function, indeed, the weights of (5.1) can be calculated as the coefficients of the first $k + 1$ terms of a Taylor expansion, as well known; but such analytical expression is equivalent to the exact knowledge of the function over its complete definition set.

Artificial Neural Network's are intended to deal with the more general case when a mapping is incompletely and noisily known, namely through its value in a finite

number of example points, constituting however a statistically significant sample that could allow ANN to reach satisfactory generalization ability, i.e. ability to approximate the mapping everywhere. In any of their nodes, called *perceptrons*, the primitive functions are general threshold ones, whose inputs are properly weighted: the same as for neurons in brain and their input information [2] (see also [1], [3], [4] for introduction and review).
We represent input and weights as vectors like $\vec{x}$ and $\vec{w}$ respectively and we call θ a primitive threshold function, whose output depends on a test as

$$\vec{w} \cdot \vec{x} \geq x_0 \tag{5.2}$$

or in the *extended form* as

$$\vec{w} \cdot \vec{x} \geq 0 \tag{5.3}$$

using a w_0 component equal to -1 for $\vec{w}$ (fig.5.2); x_0 is usually called *bias*. Some

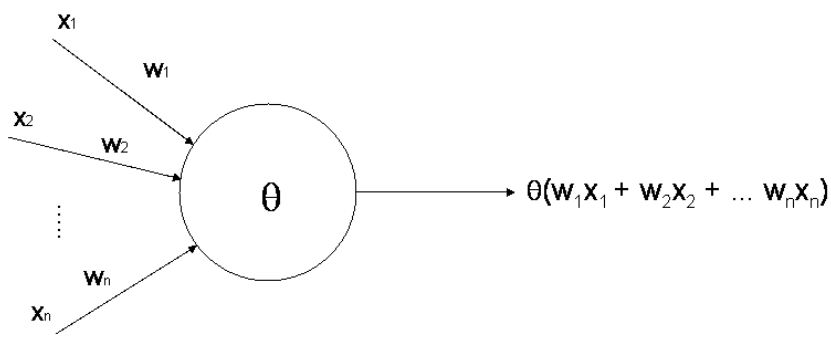

Figure 5.2: Abstract neuron primitive function.

mapping may divide its definition set into two linearly separated subsets, i.e. subsets separated by some hyper-plane: it is easily proved that in this case a single perceptron is sufficient to reproduce the mapping, because the equation $\vec{w} \cdot \vec{x} = 0$ defines in fact an hyper-plane, while condition (5.2) determines one (and consequently the other) of the two half-spaces on its sides. The most simple examples are the bi-dimensional logic functions AND and OR as can be seen in fig.5.3. More often, however, a mapping does divide its definition set into two or more subsets

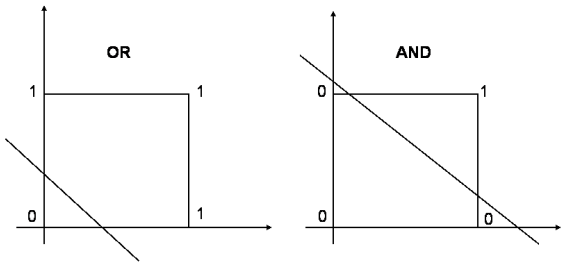

Figure 5.3: Linearly separated sets.

that are not separable by a flat surface. The most simple example is given by the two-dimensional XOR or (exclusive OR) logic function, for which a continuous line of at least two straight pieces is needed to separate the two subsets into which its definition set is divided (fig.5.4). It is understandable that in this case some network of a number of perceptrons could cooperate to build some suitable piece-wise flat hyper-surface able to approximate the needed, generally curved hyper-surface of separation.

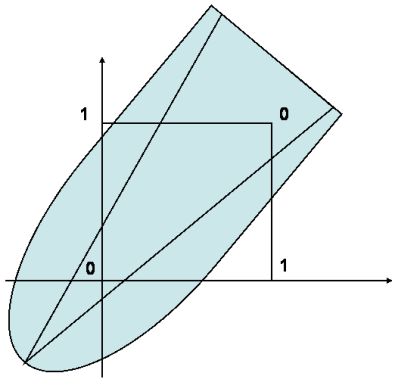

Figure 5.4: Non linearly separated sets.

Mappings that separate their definition set into subsets are those occurring in classification problems through pattern recognition and are the typical ones in (high-energy) physics analysis. What should make Artificial Neural Networks more pow-

erful then the usual sequential-cuts attack to classification problems may be viewed in the two-dimensional hypothetical case shown in fig.5.5, because there quite hardly any cut on the projection on any axis could give any good classification result; instead the blue piecewise straight line is what one could expect from an Artificial Neural Network.

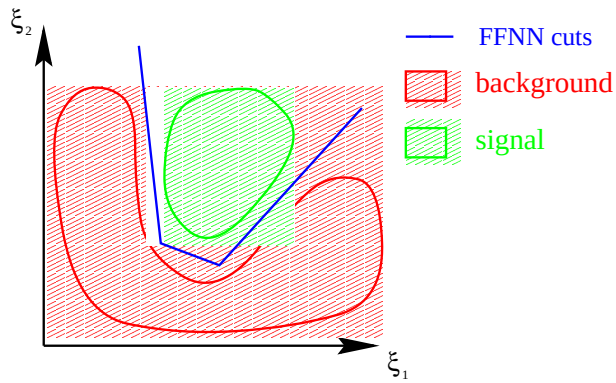

Figure 5.5: Hypothetical two-dimensional classification case. The blue piecewise straight line could be a Neural Network cut.

ANN architectures most usually are identified by a number of *layers*, each composed of a number of (abstract) *neurons*, the computing units sketched in fig.5.2, whose primitive function is called *activation function* and typically consists of a smooth variant of the *step* function, like the sigmoid:

$$S(x) = \frac{1 - e^{-x}}{1 + e^{-x}} \tag{5.4}$$

These networks are generally called **multi-layer perceptrons**; the set of input nodes is called the *input layer*, the set of output nodes is called the *output layer*; all the layers not directly connected to input or output are called *hidden layers* (fig.5.6). We are interested in so called **feed-forward network**s, such that information proceeds from input to output along successive layers. In fig.5.7 it is shown a **two layer** perceptron, thus a network having two layers of weights and using sigmoidal hidden units, that has proven to be an important class for practical applications. In fact it is possible to see that such networks can approximate arbitrarily well any functional

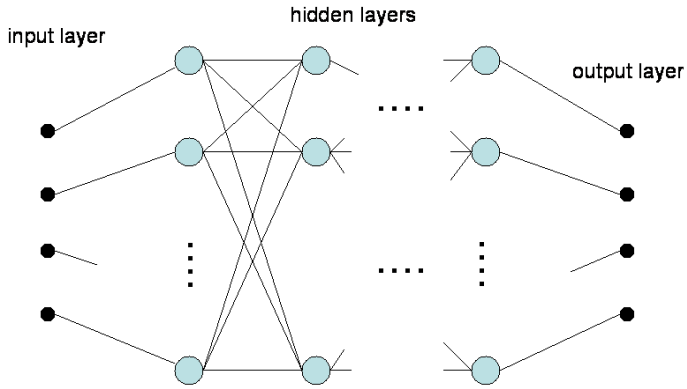

Figure 5.6: Multi-layer network.

continuous mapping from a finite-dimensional space to a finite-dimensional space, provided a sufficient number of hidden units is given. Discussion of this important point is found for instance in [1], [3], [4], [5] and literature cited therein. What is of most interest to us is the corollary that in the context of classifications problems networks with sigmoidal nonlinearities and two layers of weights can approximate any decision boundary to arbitrary accuracy; in other words such networks provide universal non-linear discriminant functions and thus they *model* **posterior probabilities of class membership**.

5.2 Reactive Tabu Search

Several approaches to learning with feed-forward multi-layer perceptrons have been realized (see [1], [3], [4], [5] reviews); they mostly use derivative based algorithms. We use instead an non-derivative based strategy devised by R. Battiti and GP. Tecchiolli, formulated as a *combinatorial optimization* problem and solved by means of a heuristic operational method called *Reactive Tabu Search* (**RTS**) [6, 7]. RTS is based on the construction of *search trajectories* in the space of the binary strings of length $L = N * B$, into which N *weights* are suitably coded using B bits per weight. The search is intended to locate an optimal minimum on the *cost surface*

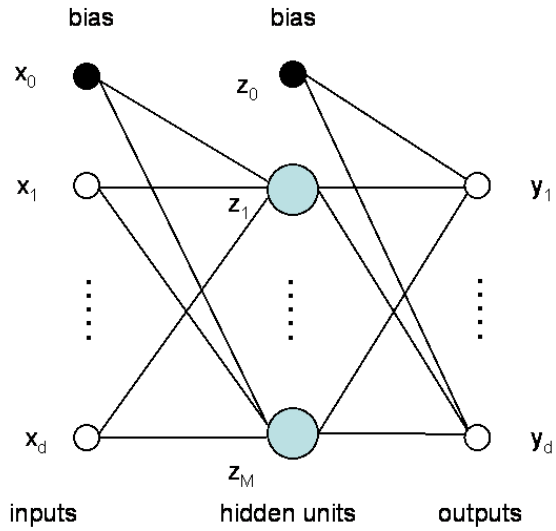

Figure 5.7: Two-layer network.

by means of a sequence of elementary moves, each consisting of a single bit-flip in the string of weights.

According to the known Glover's *Taboo Search* strategy ([8]), when a move is done, its inverse is forbidden for a *prohibition period* of T successive steps, thus forcing some amount of *diversification* in the search process: as a matter of fact, if no such action is undertaken, the resulting "greedy" moves toward low values of the *cost function* would easily cause *trapping* in local minima or *cycling's* around them. RTS remarkably enhances such diversification, dynamically adjusting the T parameters through a simple mechanism that evaluates and *reacts to* the local shape of the cost surface: this way it escapes rapidly from local minima and cycling's. Evaluation of the local shape of the cost surface is done recalling history of the recent moves, so that an important feature of RTS is that it is intensively **history-based**.

Each effective elementary move is decided on the basis of the best in a *sample* of admissible (i.e. not taboo) random exploratory steps in a neighbourhood of the current position. In order to have such exploratory moves effectively done in a *vicinity*, the well known *gray* code is used to represent the weights, because then a single bit-flip on the string of the weights most probably produces a small change

of just one of the weights; alternatively the usual binary coded is implemented, but then the bit to be flipped (*toggle-bit*) on a randomly picked weight is chosen according to a linear distribution, decreasing with the bit significance. Fig.5.8 shows the basics of RTS.

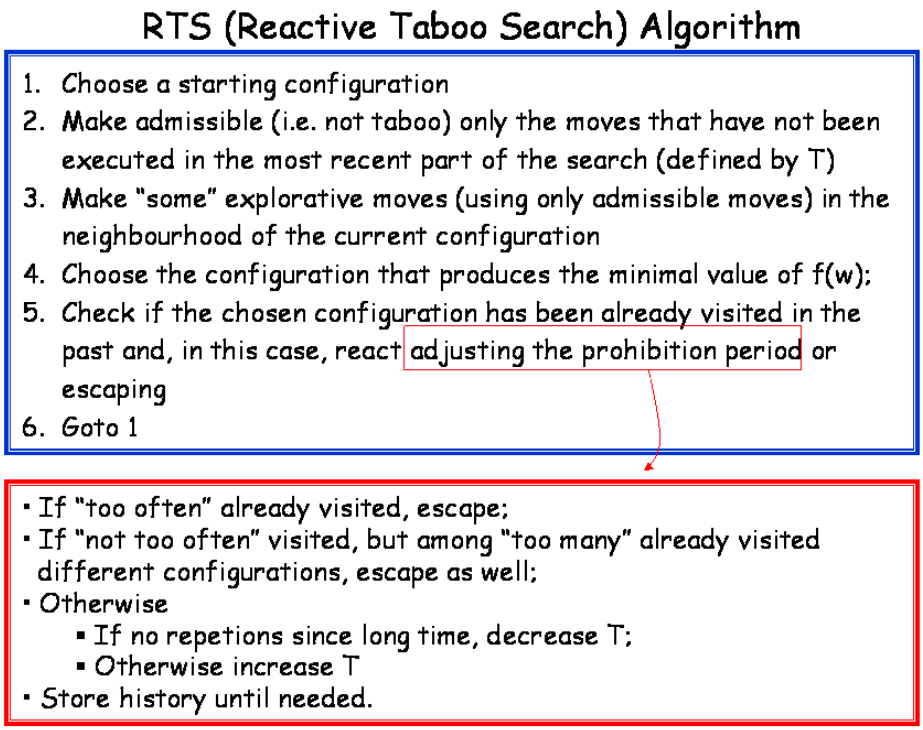

Figure 5.8: Basics of RTS operations.

Bibliography

[1] R. Rojas, **Neural Networks, a systematic introduction**, *Springer-Verlag (1996)*

[2] F. Rosenblatt, **The perceptron: a probabilistic model for information storage and organization in the brain**, *Psycological Review, Vol.65,pp.386-408*

[3] C.M. Bishop, **Neural Networks for Pattern Recognition**, *Oxford University Press, (1995)*;

[4] J. Hertz, A. Krogh and R.G. Palmer, **Introduction to the Thory of Neural Computation**, *Addison-Wesley, (1991)*

[5] B.D. Ripley, **Pattern Recognition and Neural Networks**, *Cambridge University Press (1996)*

[6] R. Battiti, G. Tecchiolli, *ORSA J. Comp.,* **6***, 126 (1994).*

[7] R. Battiti and G. Tecchiolli, **Training neural nets with the reactive tabu search**, *IEEE Transactions on Neural Networks 6(5),1185-1200 (1995)*

[8] F. Glover., *ORSA J. Comp.,* **1***, 190 (1989)*

Chapter 6

$t\bar{t} \rightarrow \tau + jets$ selection

In this chapter, after a brief description of τ reconstruction method at *CDF*, we will describe the variables we use to discriminate $top \rightarrow \tau + jets$ events from background. Such variables are used as inputs for neural networks: we explore two techniques, the first based on a single net, the other on two networks in cascade.

6.1 τ candidate reconstruction

At *CDF* τ is reconstructed exploiting information from tracking system, calorimeters and shower max detectors. Reconstruction procedure is based on the following steps [1] [2]:

1. a calorimeter tower with transverse energy $E_T^{seed_twr}$ greater than 6 GeV is found (*seed tower*);

2. a number N^{twr} of adjacent towers (*shoulder towers*) with transverse energy greater than 1 GeV is added to the seed tower to build a calorimetric cluster; the sum of transverse energies of the towers is used as the transverse energy of the τ candidate;

3. a *COT* track pointing to the cluster and with momentum $p_T^{seed_trk}$ greater than 4.5 GeV is associated to the cluster (*seed track*); if several such tracks

are found, the one with the highest momentum is chosen; the direction of the seed track becomes a *reference direction* for all following steps;

4. other *COT* tracks close to the seed one are associated to the tau candidate (*shoulder tracks*); shoulder tracks must have momentum greater than 1 GeV and small separation $\Delta z^{sh\text{-}trk}$ (5 cm) from the z-intercept of the seed track; moreover, they have to be within a 3D angle θ_{iso} with respect to the reference direction. Tracks within angle θ_{sig} with respect to the seed track are considered τ decay products while ones with $\theta_{sig} < \theta < \theta_{iso}$ are called *isolation tracks* and are used to veto τ candidates. θ_{sig} and θ_{iso} define *signal* and *isolation* cones respectively (see fig. 6.1). θ_{sig} depends on τ energy and is defined as

$$\theta_{sig} = min(0.17, max(\frac{5GeV}{E_{calo}}, 0.05))$$

while θ_{iso} is 0.525 rad (30°).

5. the last step is π^0 reconstruction and matching to τ candidates. π^0s are reconstructed as clusters in the *CES*. Raw wire and strip *CES* clusters are obtained using a clustering algorithm that starts with a seed wire or strip and combines up to five wires or strips into a cluster. Raw strips and wire clusters are then matched to each other to form 2D *CES* clusters. If matching is not unique, the measured energy is used to identify which wire and strip clusters are coming from the same cluster. Each matched cluster is classified as a π^0 candidate if no *COT* tracks with $p_T > 1$ GeV/c is found nearby. If only one π^0 is found in correspondence of a particular calorimeter tower, the energy assigned to the π^0 is the full EM energy of this tower minus the expected energy deposited from all tracks traversing the tower. If more than one π^0 candidate interests the same tower, the available EM energy is divided between them proportionally to their respective *CES* cluster energies. As in the tracks case, for π^0s a cone around the seed track is defined:

$$\alpha_{\pi^0} = min(0.17, max(\frac{5GeV}{E_{calo}}, 0.10))$$

All π^0s candidates inside the cone, having $E_T > 1$ GeV and having matched in both x and z views of *CES* are associated to the τ candidate.

$E_T^{seed_twr}$	$>$	6.0 GeV
$E_T^{sh_twr}$	$>$	1.0 GeV
N^{twr}	$\leq$	6
θ_{sig}	$=$	$min\left(0.17, \frac{5.0 rad/GeV}{E^{\tau_cl}}\right)$
θ_{iso}	$=$	0.525 rad
$p_T^{seed_trk}$	$>$	6.0 GeV
$p_T^{sh_trk}$	$>$	1.0 GeV
Δz^{sh_trk}	$<$	5.0 cm

Table 6.1: τ reconstruction parameters.

In table 6.1 τ candidate reconstruction parameters are summarized.

6.2 τ identification variables

The variables we use to separate $t\bar{t} \rightarrow \tau + jets$ events from background are divided in two groups: *global* variables are those which describe the events in its entirety, used to discriminate from background, in our multijet trigger sample [1], $t\bar{t}$ events additionally satisfying necessary conditions to the $\tau + jets$ channel, like significant total missing energy; other variables (τ mass, number of tracks in the τ cone, ...), describe specifically the τ candidate: we use them to distinguish between *real* taus and jets faking taus.

6.2.1 Input variables: global

- **njets**

  In a clean $t\bar{t} \rightarrow \tau + jets$ event we expect five jets: two of them come from one of the two Ws, two from b quarks and one from τ. Number of jets (*njets*) distribution is therefore peaked on 5, while background events are peaked on 4 (Figure 6.2). *njets* refer to *tight* jets, i.e. jets with $E_T > 15.0$ GeV and

[1] Multijet trigger requirements will be described in section 6.3

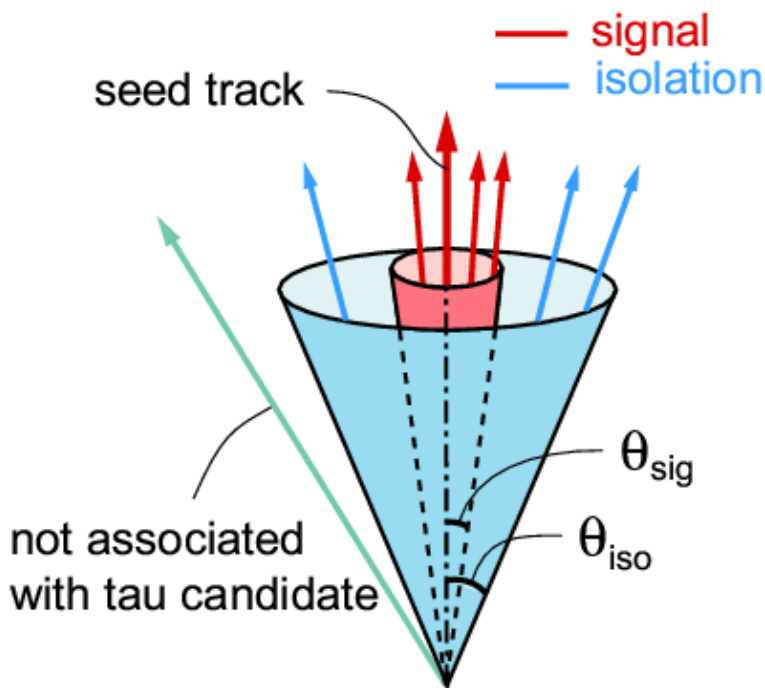

Figure 6.1: τ candidate reconstruction: seed track, signal and isolation cones.

$|\eta| < 2.0$.

- **aplanarity** and **sphericity**

  Both are calculated starting from the normalized momentum tensor M_{ab}

$$M_{ab} = \frac{\Sigma_i p_{ia} p_{ib}}{\Sigma_i p_i^2} \tag{6.1}$$

  where a and b indexes run over the three space directions and p_i is the momentum of the i_{th} jet [3]. We are interested in finding the axis $\vec{n}$ such that the normalized sum of the square components of the jet momenta along it has a maximum

$$max \quad \frac{\sum_i (\vec{p}_i \vec{n})^2}{\sum_i \vec{p}_i^2} \tag{6.2}$$

  This quantity easily characterizes space direction distribution of such momenta, namely the case they are mostly along $\vec{n}$ against more isotropical distributions.

The ratio in eq. 6.2 can be written as

$$\frac{\sum_i (\vec{p}_i \vec{n})^2}{\sum_i \vec{p}_i^2} = \sum_{a,b=1}^{3} n_a \frac{\sum_i (p_{ia} p_{ib})}{\sum_i \vec{p}_i^2} n_b = \sum_{a,b=1}^{3} n_a M_{ab} n_b \quad (6.3)$$

M matrix is symmetric and definite positive. Its unit eigenvectors $\vec{n}_{(1)}$, $\vec{n}_{(2)}$ and $\vec{n}_{(3)}$ have eigenvalues Q_j that satisfy the condition $Q_1 + Q_2 + Q_3 = 1$ because the trace of M is null.

If we order the eigenvalues such as $0 \leq Q_1 \leq Q_2 \leq Q_3$, the axis $\vec{n}$ we are looking for is the normalized eigenvector $\vec{n}_{(3)}$ of M with the highest eigenvalue $Q_{(3)}$.

Sphericity S is thus defined by

$$\begin{aligned} S &= \frac{3}{2}(Q_1 + Q_2) = \frac{3}{2}(1 - Q_3) = \\ &= \frac{3}{2}\left(1 - \frac{\sum_i (\vec{p}_i \cdot \vec{n}_{(3)})^2}{\sum_i \vec{p}_i^2}\right) = \frac{3}{2}\left(\frac{\sum_i \vec{p}_{iT(3)}^2}{\sum_i \vec{p}_i^2}\right) \end{aligned} \quad (6.4)$$

where subscript T denotes momentum component transverse with respect to $\vec{n}_{(3)}$ axis; S is null in the limiting case that the momenta are all exactly along $\vec{n}_{(3)}$ and approaches 1 for a perfectly isotropic jet momenta distribution, when $Q_1 = Q_2 = Q_3 = \frac{1}{3}$.

Aplanarity is defined as

$$A = \frac{3}{2} Q_1 \quad (6.5)$$

and it is null when the sum of jet momenta has null component on $\vec{n}_{(1)}$ axis; this is the case of coplanar or collinear events. On the contrary, when jet momenta have isotropic distribution, $Q_1 = Q_2 = Q_3 = \frac{1}{3}$ and A reaches its maximum, $\frac{1}{2}$.

Jets emerging from a $t\bar{t}$ pair are expected to be uniformly distributed and, as a consequence, they will tend not to lie in the same plane. Thus, for $t\bar{t}$ events we expect higher aplanarity and sphericity values (Figures 6.3 and 6.4).

- **centrality**

  Jets produced by $t\bar{t}$ pairs decaying hadronically are emitted preferably in the transverse plane ($r - \phi$ plane); the centrality C of the energy flux is calculated as

  $$C = \frac{\Sigma_i E_{Ti}}{\sqrt{\hat{s}}} \tag{6.6}$$

  Signal events are expected to have centrality values nearer to 1 (greater amount of energy emitted in the transverse plane) with respect to background ones (Fig. 6.5).

- $\boldsymbol{\Sigma E_{T3}}$

  In QCD events, the two most energetic jets are produced by $q\bar{q}$ processes while the least energetic ones are produced by gluon *bremsstrahlung*. On the contrary, in a $t\bar{t}$ event up to 6 jets can be produced by hard processes. As a consequence, ΣE_{T3}, the scalar sum of jet transverse energies, excluded the two leading ones, can discriminate between signal and background (Fig. 6.6).

- **Missing Et significance**

  In a $t\bar{t} \to \tau + jets$ event two neutrinos are produced (from W and τ decays): therefore we expect a significant amount of *Missing Energy* ($\not{E}_T$) in signal events.

  $\not{E}_T$ significance is calculated as

  $$MetSigf = \frac{\not{E}_T}{\sqrt{\Sigma E_T}} \tag{6.7}$$

  Signal events have larger missing E_T significance with respect to background ones (Fig. 6.7).

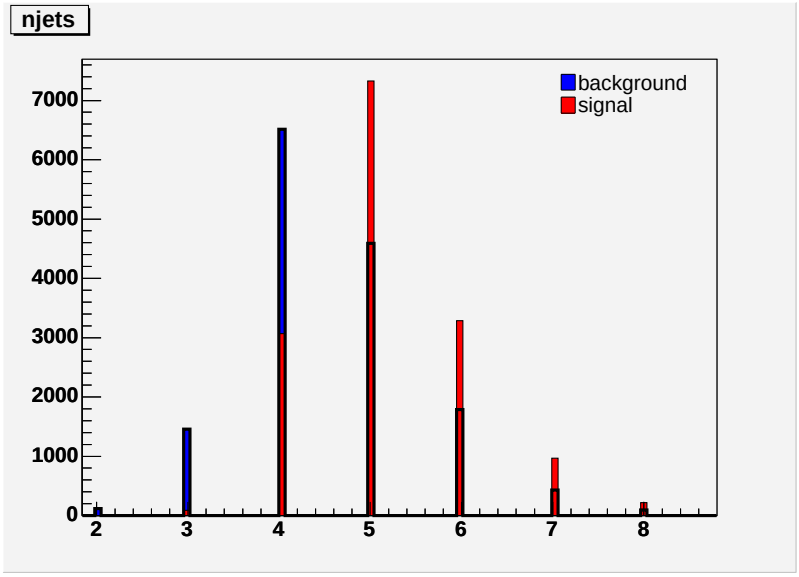

Figure 6.2: Distribution of tight jets (jets with $E_T > 15.0$ GeV and $|\eta| < 2.0$) for $t\bar{t} \rightarrow \tau + jets$ and background events.

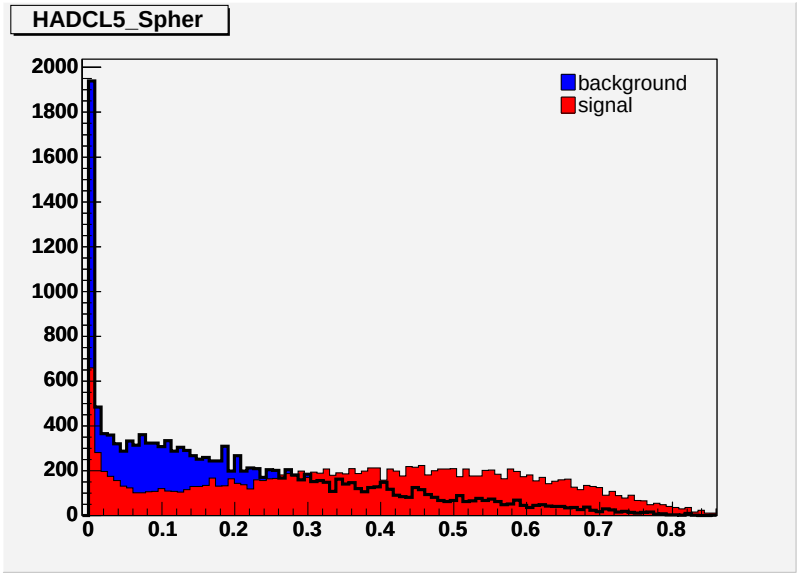

Figure 6.3: Sphericity distribution: jets emerging from $t\bar{t}$ pairs are uniformly distributed, thus they tend to have sphericity values higher than background ones.

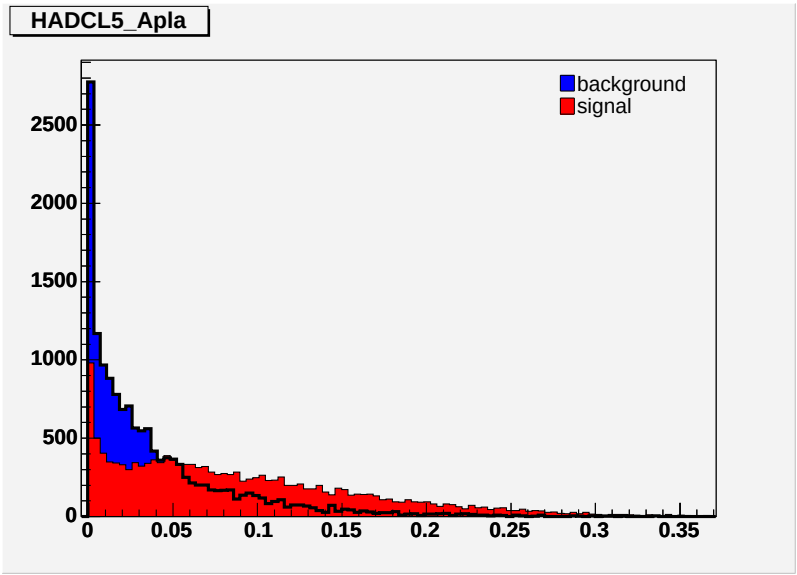

Figure 6.4: Aplanarity distribution: as jets emerging from $t\bar{t}$ pairs are uniformly distributed, they tend not to lie in the same plane having thus aplanarity values higher than background events.

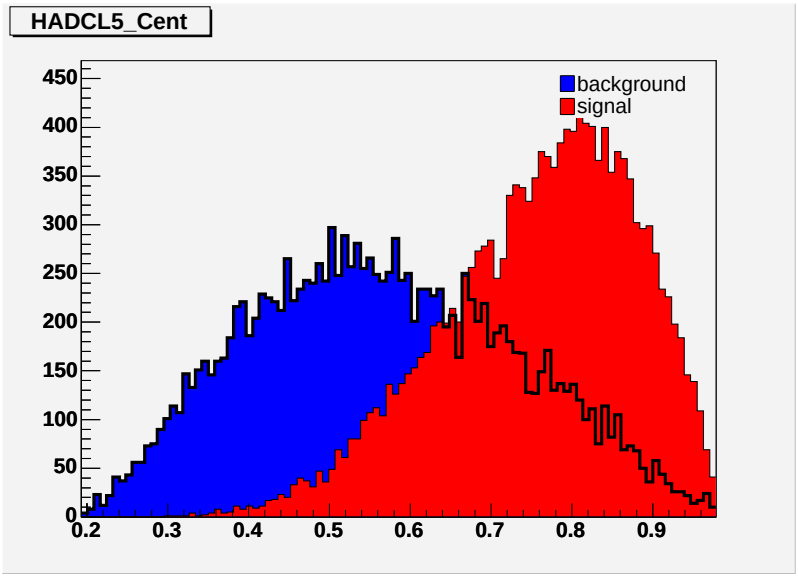

Figure 6.5: Centrality distribution: jets produced by $t\bar{t}$ pairs are emitted preferably in the transverse plane, thus they have centrality values nearer to 1 (greater amount of energy in the transverse plane) with respect to background.

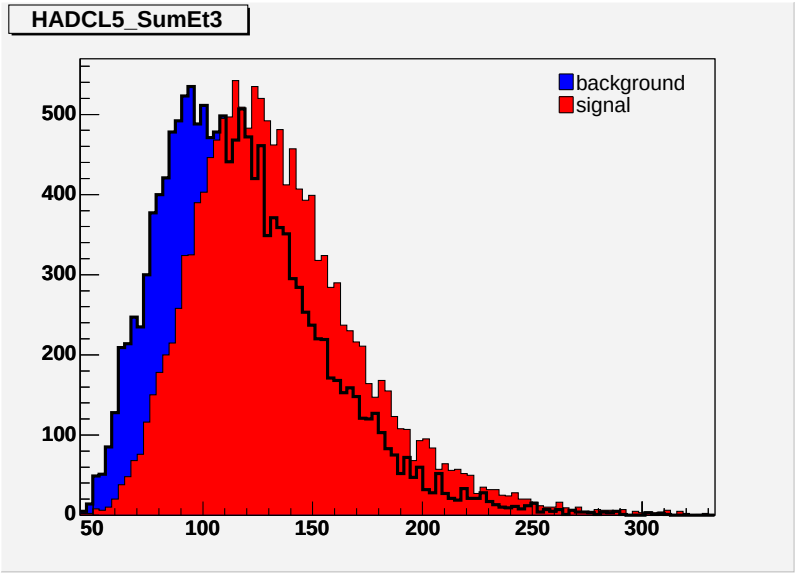

Figure 6.6: ΣE_{T3} distribution: the scalar sum of jet transverse energies, excluded the two leading ones, is greater in $t\bar{t}$ events with respect to background, because in the first case up to six jets can be produced by hard processed, against two in a QCD process.

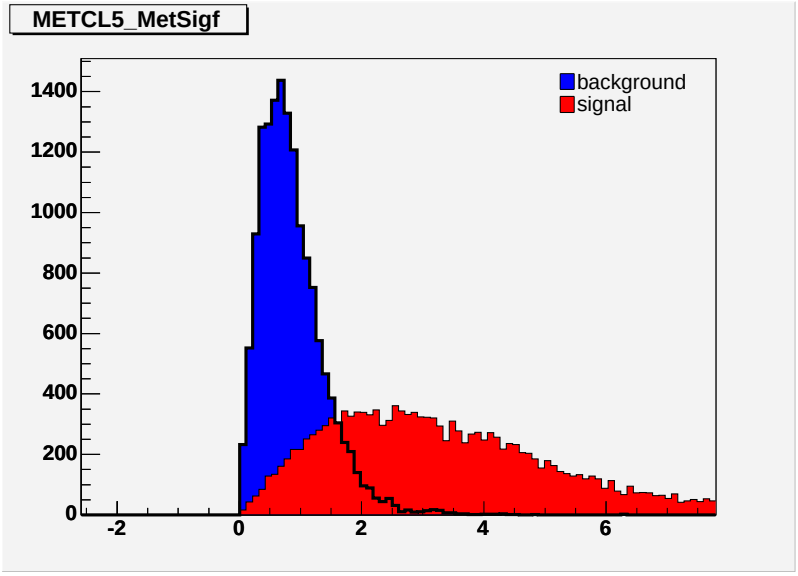

Figure 6.7: Missing transverse energy significance distribution: in a $t\bar{t} \rightarrow \tau + jets$ events, due to the production of two neutrinos, missing energy has significance greater than background events.

6.2.2 Input variables: tau specific

- **ntrks**

  A τ decaying semileptonically will produce in most cases one or three charged pions; thus *ntrks*, the number of tracks in the τ signal cone, has a distribution peaked on these two values (fig. 6.8). On the contrary, quark or gluon jets arising from QCD processes have a broader distribution peaking at different track multiplicities while photons, electrons and muons have none or just one track.

- **nmuhits**

  A muon could fake a τ decaying into just one charged pion and without neutral pions; thus the presence of muon hits in correspondence of a τ candidate can be used as a discriminating variable (Fig. 6.9);

- **etateta** and **phiphi**

  τ cluster Eta-Eta and Phi-Phi moments:

  $$\sqrt{<(\eta_i - <\eta>)^2>}, \quad \sqrt{<(\phi_i - <\phi>)^2>}$$

  where η_i and ϕ_i are pseudorapidity and ϕ coordinate of the ith tower belonging to the τ cluster. Jets from real taus are narrower than those from fake taus (figs. 6.10 and 6.11).

- **caloetiso, trackiso, nisopi0s**

  As already stated, a jet coming from a τ is narrow and well isolated; we can quantify this isolation in different ways, exploiting the definition of *isolation* cone (section 6.1): *trackiso* is the transverse momentum of all tracks in the isolation annulus (Fig. 6.13), while *caloetiso* is the total calorimetric energy in the same region (Fig. 6.12). A real τ should have null these quantities. For the same reason should be zero the number of neutral pions in the isolation cone, *nisopi0s* (Fig. 6.14).

- **visible mass**

  The so-called *visible mass* (fig. 6.15) is τ candidate invariant mass calculated

using momenta and energies of both tracks and neutral pions. τ candidates from real taus have reconstructed masses lower than fake taus

- **τ boost**
  The size of a jet can result from different factors, one of them being jet's boost: bigger the boost, smaller the size; thus a jet with low track multiplicity can fake a jet from τ if its boost is great enough. The Lorentz boost parameter

$$\tau_{boost} = \gamma\beta = \frac{|\vec{p}|}{m_\tau} \tag{6.8}$$

  calculated using τ candidate momentum and *visible mass* (Fig. 6.16) can help in discriminating real taus from fakes.

- **angleseedtocluster**
  The angle between the seed track and the cluster barycentre is plotted in figure 6.17: true taus tend to have smaller values with respect to fake taus.

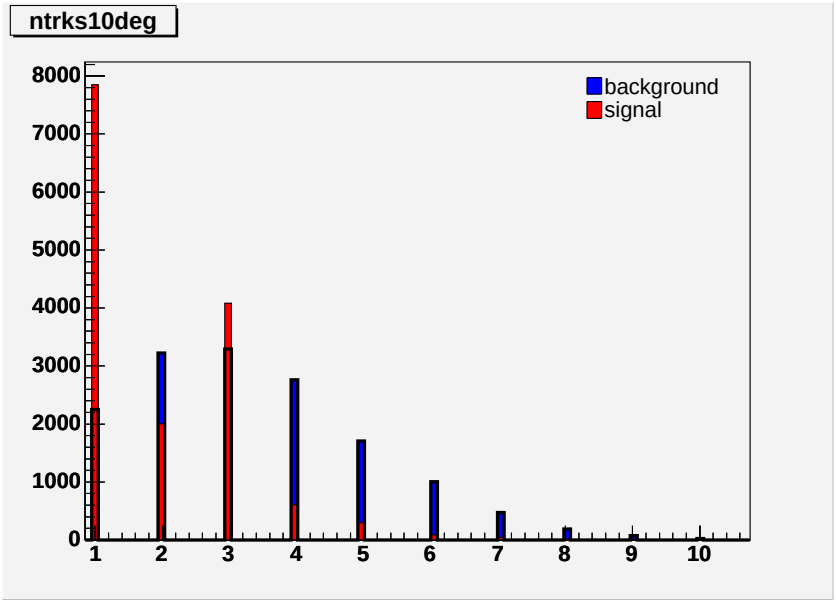

Figure 6.8: Distribution of the number of tracks in the τ signal cone: a jet from a true τ has one or three charged tracks.

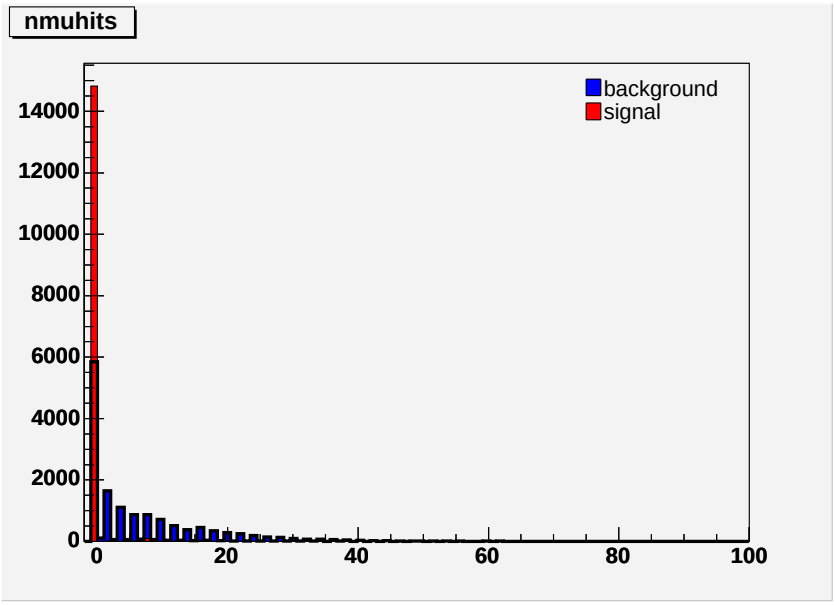

Figure 6.9: Distribution of muon hits in correspondence of τ candidates; a muon can fake a τ decaying into just one charged piona real τ this variable can discriminate between true taus and muons faking taus.

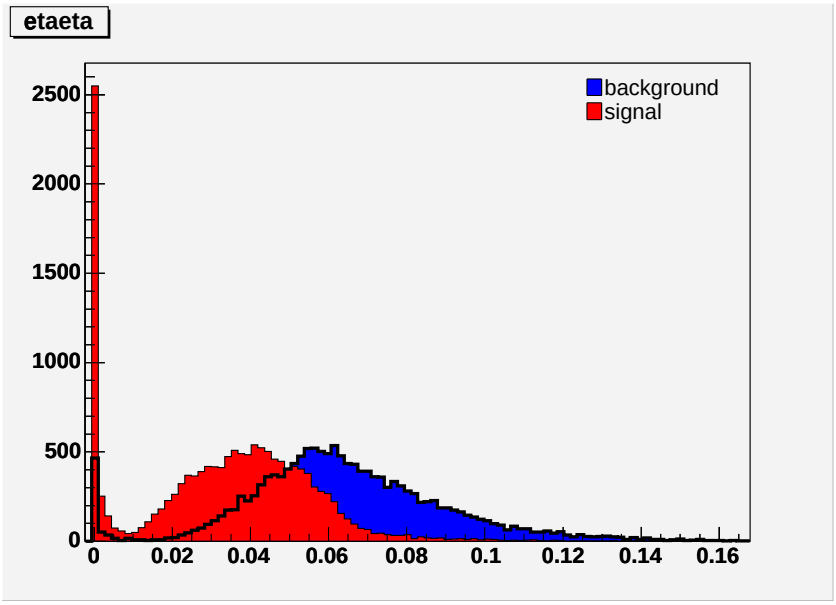

Figure 6.10: Distribution of τ cluster Eta-Eta moment: true taus produce a narrower jet with respect to fake taus, thus cluster Eta-Eta moment is lower.

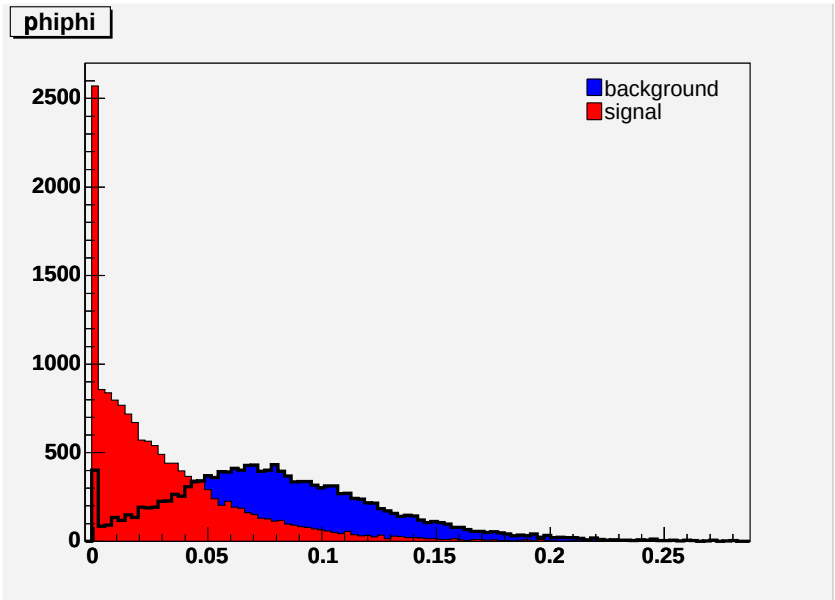

Figure 6.11: Distribution of τ cluster Phi-Phi moment: true taus produce a narrower jet with respect to fake taus, thus cluster Phi-Phi moment is lower.

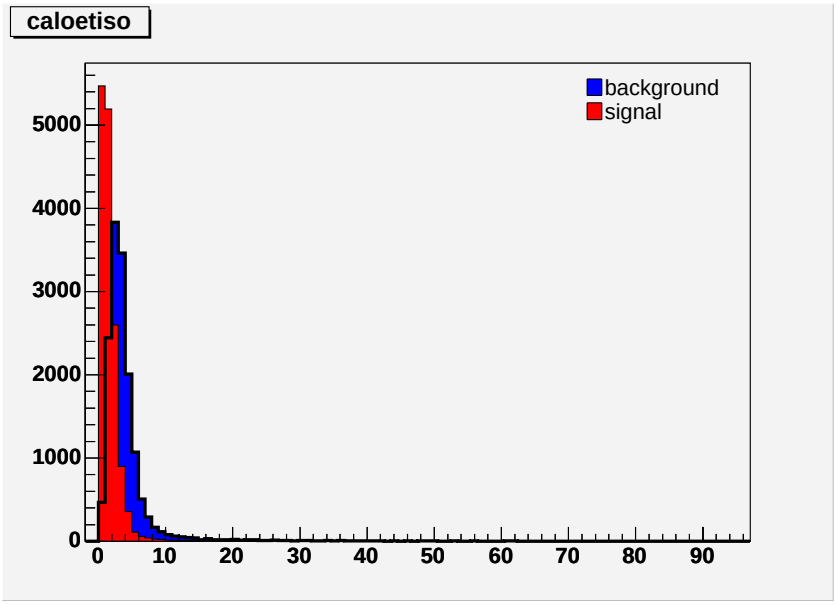

Figure 6.12: Distribution of the total calorimetric energy in τ jet isolation annulus. It is peaked towards zero for real taus.

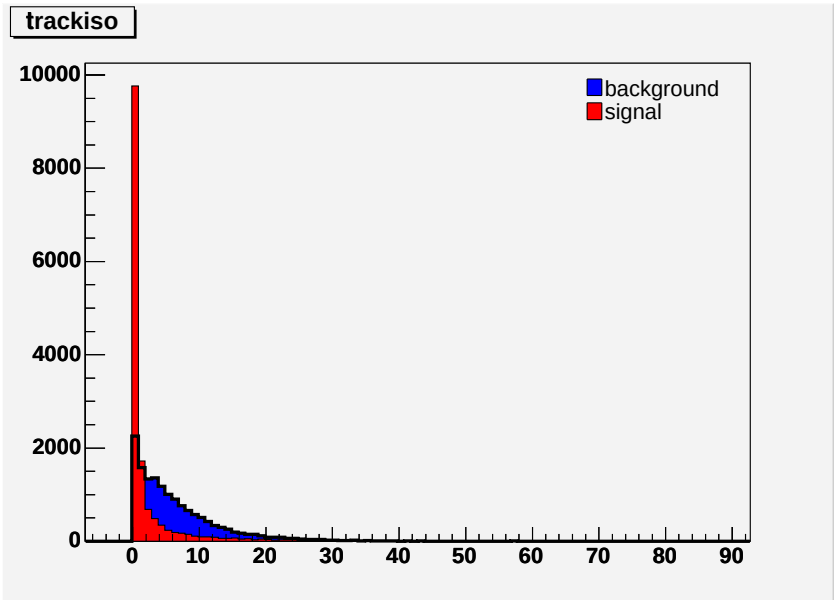

Figure 6.13: Distribution of transverse momentum of all tracks in τ jet isolation annulus, peaked towards zero for real taus.

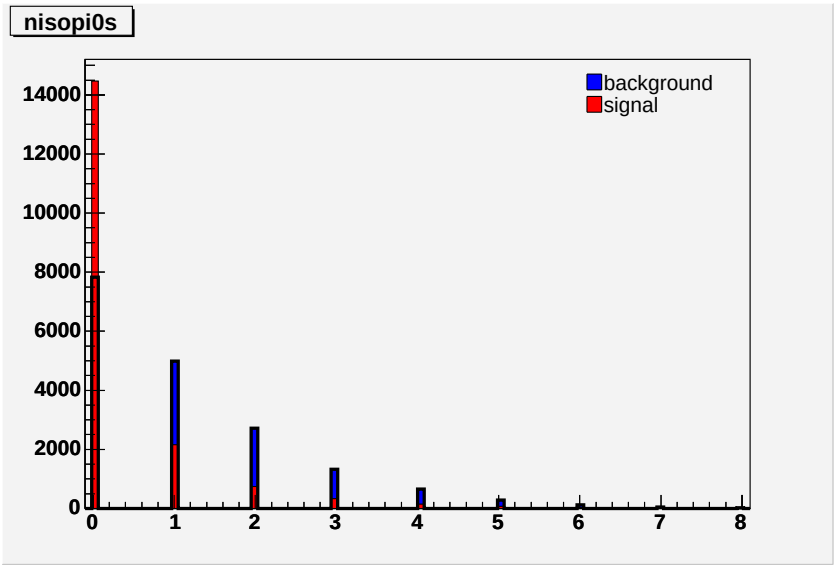

Figure 6.14: Number of neutral pions in τ jet isolation annulus. Real taus, in most cases, have no neutral pions in the isolation region.

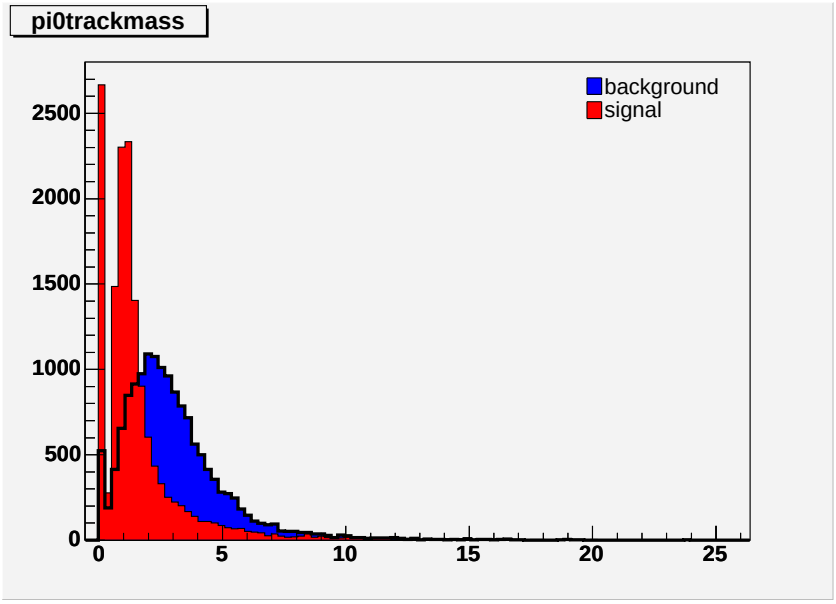

Figure 6.15: Visible mass of τ candidates, obtained summing over all tracks and neutral pions in signal cone. τ candidates from real taus have reconstructed masses lower than fake taus (and approaching more the true value 1.777 GeV/c^2.)

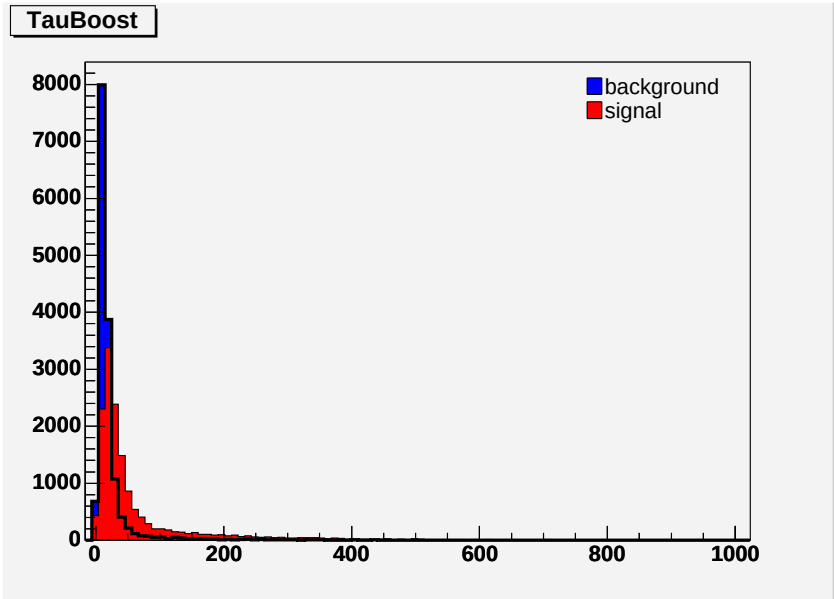

Figure 6.16: Distribution of Lorentz boost parameter for real and fake taus: real taus show greater boosts compared to fakes.

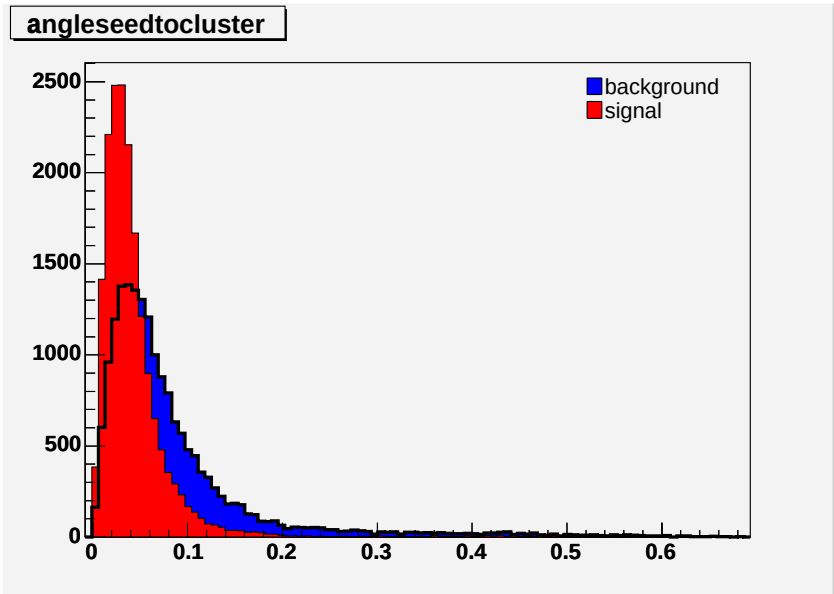

Figure 6.17: Distribution of the angle between τ candidate seed track and cluster barycentre. It is peaked towards zero for real taus.

6.3 $t\bar{t} \to \tau + jets$ selection using NNs

We want to select $t\bar{t} \to \tau + jets$ events in 311 pb^{-1} of data collected up to August 2004 (*gset0d* dataset) This dataset has been obtained through **TOP_MULTIJET** trigger that requires necessary conditions for high transverse momenta hadronic final states:

- at **Level 1** at least 1 calorimetric tower with $E_T \geq 10$ GeV,
- at **Level 2** at least 4 calorimetric clusters with $E_T^{cluster} \geq 15$ GeV and total transverse energy $\Sigma E_T \geq 125$ GeV,
- at **Level 3** at least 4 jets (with fixed cone radius 0.4) with $|\eta| \leq 2.0$ and transverse energy $E_T^{jet} \geq 10$ GeV.

This trigger is thus optimized for *all-hadronic* channel, but it fits also to our analysis having $t\bar{t} \to \tau + jets$ channel a signature similar to *all-hadronic* one: the main differences are a substantial amount of missing energy, due to two neutrinos and five jets instead of six. In order to clean our dataset as much as possible before analysis, we apply some preliminary cuts (**prerequisites**); these are very general cuts, not affecting acceptance on the channel under study [4]:

- we apply Good Run List version 7.0 as from the Data Quality Web page and remove by hand runs taken with old trigger i.e. those whose run number is smaller than 143938;
- we ask that at least one highest-SumPt class-12 vertex is present[2];
- in order to restrict our analysis to events with a well centered beam crossing, we ask that $z_{vert} < 60$ cm, where z_{vert} is z-position of the highest-SumPt class-12 vertex reconstructed with *ZVertexColl* algorithm; moreover, we ask that

[2] A class-12 vertex is a vertex reconstructed by *ZVertexColl* algorithm using at least two tracks having *COT* hits; vertices found by *ZVertexColl* are used by jet clustering algorithm (see section 4.1.3).

the distance between this vertex and the one from *PrimVtx* algorithm[3] is less than 5 cm ($|z_{vert} - z_{PrimVtx}| < 5$ cm), to ensure that the vertex used for tagging is indeed close to the one used for jet clustering;

- to avoid overlap with other leptonic analysis we remove events with a good, high P_T electron or muon.

After prerequisites application, we are left with nearly 4 millions events in 311 pb^{-1} of data among which, from Standard Model, we expect ~ 121 $t\bar{t} \to \tau + jets$ events[4]. The $\frac{S}{N}$ ratio before any analysis is thus $\sim \frac{1}{32000}$.

In a $t\bar{t} \to \tau + jets$ event two jets are from b quarks, so the first step to reduce background is considering only those events that have at least one *SecVtx*[5] b-tagged jet[6]. As one can see from tables 6.2 and 6.3, were effects of cuts applied to *gset0d* (data) and *ttopvh* ($t\bar{t}$ MC) datasets are shown, this request throws away a certain amount of signal (35%) but it makes a great rejection of background (92%). Moreover, having a b-tagged jet in an event makes event reconstruction easier. For the same reason we also reject those events that do not have at least one reconstructed τ candidate: we loose some signal (40%) but make event reconstruction easier.

After prerequisites, b-tagging and τ candidate request, we are left with a sample with $\frac{S}{N} \sim \frac{1}{6000}$. This is our starting point. How can we further increase $\frac{S}{N}$? Our strategy is based on neural networks. In the following we will describe the two techniques we explored: the first is based on a single neural network, the second on two of them in cascade.

[3] *PrimVtx* is another algorithm used to reconstruct primary vertex; vertices found by *PrimVtx* are used for b-tagging (see section 4.1.3).

[4] The number of expected signal events is estimated from *ttopvh* dataset (MC events generated with Herwig with top mass = 175 GeV) according to

$$N_{MC} = N_{data} \frac{\mathcal{L}_{data}}{\mathcal{L}_{MC}}$$

where $\mathcal{L}_{MC}$ and $\mathcal{L}_{data}$ are MC and data luminosities.

[5] b-tagging algorithms are described in section 4.5.

[6] We do not require two b-tagged jets because this request would lower signal acceptance too much: we would be left with only 20% of signal events, against an efficiency of 65% requiring just one b-tagged.

	number of events	*efficiency (%)*
Total Data events	5900266	100.0
Good Run	4249634	72.0
z vertex	3901254	91.8
No lepton	3897615	99.9
Events with at least one b-tagged jet	308577	7.9
Events with at least one τ cand	209804	68.0
Total τ candidates	**278343**	

Table 6.2: Effects of preliminary cuts on Data events (*gset0d* dataset, 311 pb^{-1})

	number of events	*efficiency (%)*
Tot events	219479	100.0
Good Run	193855	88.3
trigger	125489	64.7
z vertex	120454	96.0
No lepton	107050	88.9
τ + jets	11286 (121 in 311 pb^{-1})	10.5
Events with at least one b-tag	7340	65.0
Events with at least one τ cand	4331 (47 in 311 pb^{-1})	59.0
Total τ candidates	**4331**	

Table 6.3: Effects of preliminary cuts on MC events (*ttopvh* dataset, $M_{top} = 175$ GeV).

6.3.1 Neural Network analysis

In this section, we briefly recall how a neural network based analysis is structured. We use feed-forward neural network with one hidden layer, whose architecture is thus mainly characterized by n_i-n_h-n_o triplet, where n_i is the number of input features, n_h is the number of hidden neurons and n_o is the number of output neurons. A neural network analysis starts with *training*: we drive the net to recognize patterns given in a set of known examples (*training set*). In our case, we ask the net to assign an *output* value as near as possible to 1000 (0) to signal (background) event, searching for the best issue iteratively. We have to avoid *over-training*, which happens when the neural network learns "too well" on the training set and is not effective on a different sample of events (it is inefficient in *generalizing*). To avoid over-training, each n iterations performing a *validation* test: neural network trained so far is applied to a set of events different from training ones, but still known (the *validation* set); we eventually save the best neural network configuration on the basis of the best issue on the validation set.

In our present analysis we found that a neural network reaches its best performance after 20000 - 30000 iterations.

neural network performance is described by two parameters, *purity* and *efficiency*, which are functions of a *cut* on neural network output. Signal events are classified as *True Positives* if their neural network output is greater than a fixed bound, the *cut*, *False Negatives* otherwise. In a similar way, background events are *False Positives* if their neural network output is greater than cut, i.e. they have been incorrectly classified as signal, *True Negatives* otherwise. Using these definitions, *Purity* is

$$Pur(cut) = \frac{TruePositives(cut)}{TruePositives(cut) + FalsePositives(cut)} \tag{6.9}$$

and *Efficiency*

$$Eff(cut) = \frac{TruePositives(cut)}{TotalSignalEvents} \tag{6.10}$$

Purity measures how well the neural network can discriminate between signal and background events, while *Efficiency* is a measure of the neural network capability in recognizing signal events. An ideal neural network can recognize signal and dis-

criminate it from background with infinite precision, so both *Purity* and *Efficiency* are 100% and the *Purity* vs *Efficiency* plot is a step function; the more the training plot approaches the the ideal one, the better the training.

6.3.2 Training samples

Signal events are taken from $t\bar{t}$ MonteCarlo datasets generated with Herwig. Signal training sample is made of a mixture of events from MC generated with different top masses, in order to make training independent from top mass. In table 6.4 the datasets used for training are listed. We consider top masses from 167.5 GeV to 180 GeV: the current top mass estimate, combination of *CDF* and *DØ* results, is 172.7 ± 2.9 GeV/c^2. Before any neural network analysis, we apply to MC events *TOP_MULTIJET* trigger simulation and prerequisites. Then we select $t\bar{t} \rightarrow \tau + jets$ events at *hepg* level. Not all *hepg* $\tau + jets$ events have exactly one reconstructed τ candidate: some events haven't, because τ reconstruction algorithm isn't 100% efficient[7], and some events can have more than one, because the algorithm requests are rather loose - not to reject signal - thus a narrow jet with low track multiplicity can be classified as τ *candidate* even if it doesn't come from a τ. So, we make signal selection requiring that each *hepg* $\tau + jets$ event has at least one τ candidate matching the *hepg* τ. Matching is performed in $\eta - \phi$ plane: we decide that a τ candidate matches an *hepg* τ if their distance in $\eta - \phi$ is less than 0.2. In fig. 6.18 such distances are plotted. In table 6.5 the effects of all preliminary cuts on MC events are shown. After all requirements, our signal events are τ candidates matching *hepg* taus in $t\bar{t} \rightarrow \tau + jets$ events.

As far as **background** is concerned, we randomly select τ candidates from *gset0d* dataset after prerequisites application and b-tagging requirement. We can consider those events as background because $\frac{S}{N}$ after all preliminary cuts is $\sim \frac{1}{6000}$; during training session, as we will see later, we won't use more than a few thousands of events so signal contamination in such a sample is small enough that we can neglect

[7] taus with $\eta > 1.0$ can't be reconstructed because the algorithm uses only tracks from *COT*, whose η coverage is $|\eta \leq 1.0|$

top mass (GeV)	*dataset name*
167.55	ttopbh
170	ttopch
172.5	ttopdh
175	ttopvh
177.5	ttopfh
180	ttopgh

Table 6.4: MC datasets used to build training samples.

	ttopbh	*ttopch*	*ttopdh*	*ttopvh*	*ttopfh*	*ttopgh*
Tot events before Good run list	209946	233233	233448	219479	225501	229394
goodRunCount	185450	207609	207941	193855	199877	203770
goodTrigCount	116101	132025	133545	125489	130470	133746
goodZVertCount	111556	126877	128234	120454	125396	128404
noLeptonCount	99406	112933	114020	107050	111152	113815
$\tau + jets$ after prereq	10336	11602	11963	11286	11915	12168
$\tau + jets$ after with b-tag	6794	7621	7911	7340	7782	7906
$\tau + jets$ with matched τ cand	**4118**	**4481**	**4696**	**4331**	**4710**	**4740**

Table 6.5: Effects of preliminary cuts on MC datasets used to build training samples.

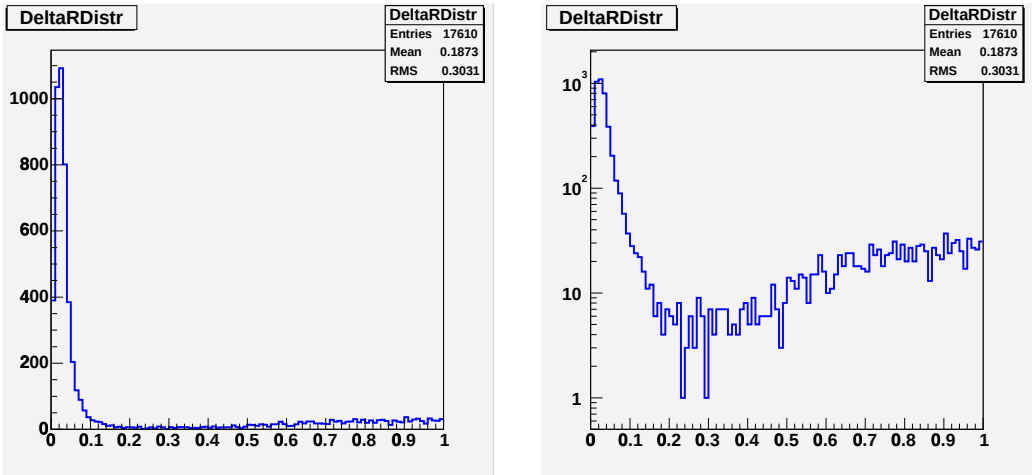

Figure 6.18: Distribution of distances in $\eta - \phi$ plane between hepg taus and τ candidates: linear (left) and logarithmic (right) scale; tau candidates whose distance from *hepg* taus is greater than 0.2 are considered fake taus.

it.

We have to teach a network to distinguish between signal and background: **how many training events** should we use? We have at our disposal as many background events as we need ($\sim$ 278000, from table 6.2) but only $\sim$ 4000 events for each mass as signal, for a total of $\sim$ 24000 signal events (table 6.5). Common sense (and previous studies [5]) suggests that more examples we submit to the net during training session, better results we obtain. But the greater the number of events, the slower the net. So we have to find a good compromise between a statistically significant number of events and net speed.

To find the best number of training events we fix a net architecture and do training using samples with different number of events, from 12 (just a provocation!) to 36000. Then we compare signal and background distributions and efficiencies on signal on a fixed, independent validation set made up of 18000 signal and 18000 background events.

In figures 6.19, 6.20 and 6.21 results are shown. Figure 6.19 represents the distribution of signal events processed by the neural network. We asked the net to assign to every event a value as much near as possible to 1000 if it recognizes

it as signal, near to 0 otherwise. Each bin between $NNout$ and $NNout + \delta_{NNout}$ contains the number of events whose neural network output lies in that interval. As signal event distribution is peaked towards the output value 1000, in figure 6.19 we consider only neural network outputs greater than 900. Figure 6.20 has to be read in the same way, with the difference that it shows background events distribution and thus only neural network output smaller than 100 are considered. Figure 6.21 shows the efficiency on signal events as a function of the cut on neural network output. Looking at all these plots we can conclude that 6 events are obviously too few to describe signal/background, thus the neural network can't distinguish between them at all; but if we move to 120, 1200 and more events we can't recognize a specific trend: all trainings seem to give the same good results. Indeed, using 36000 events the net seems to recognize signal and background with more difficulty: too much information can result in a more difficult search of the absolute minimum for a fixed number of iterations. So we decide to use 6000 training events for our analysis.

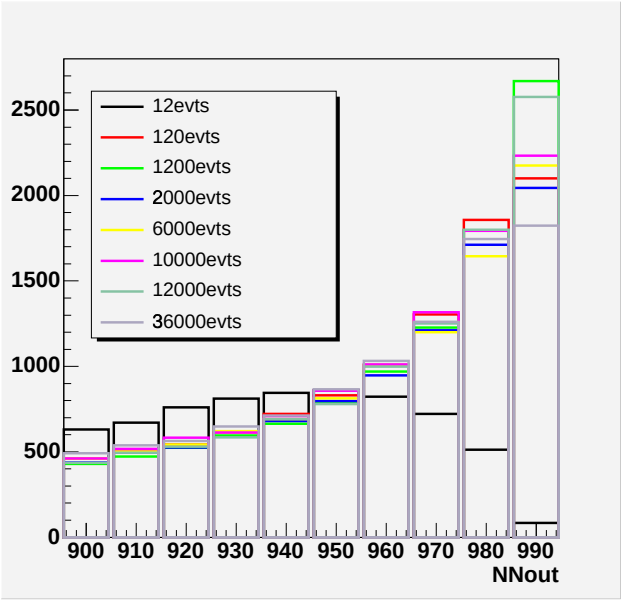

Figure 6.19: Distribution of signal events processed by neural networks trained with different numbers of training events. Zoom on the *signal* region ($900 \leq$ NNout < 1000).

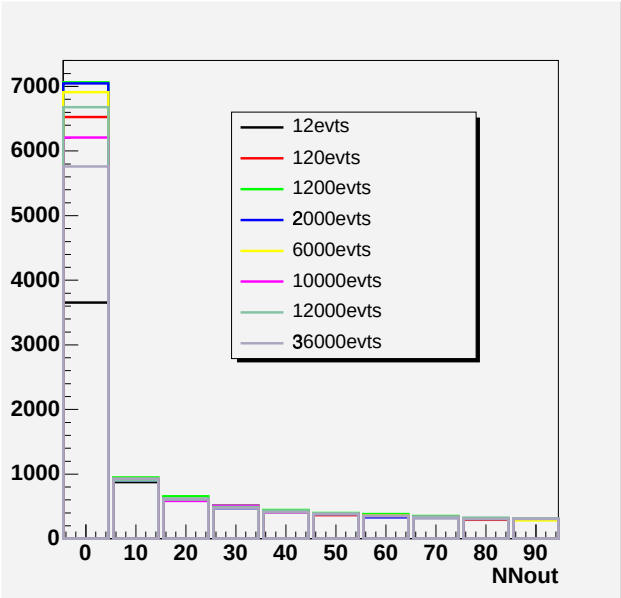

Figure 6.20: Distribution of background events processed by neural networks trained with different numbers of training events. Zoom on the *background* region ($0 \leq$ NNout < 100).

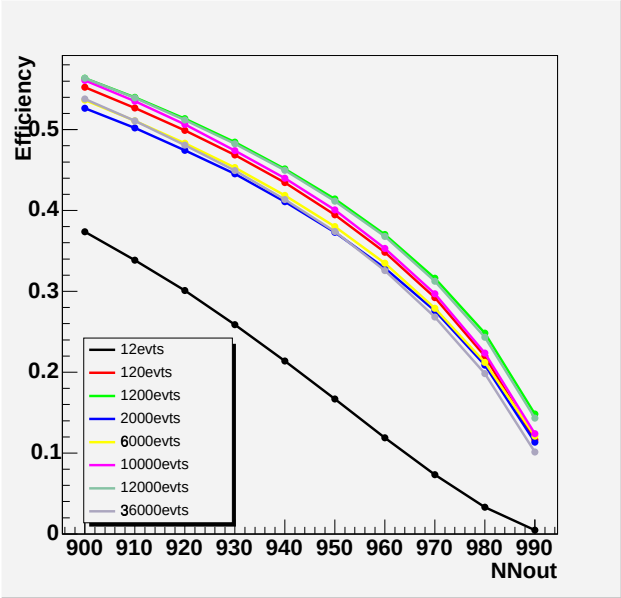

Figure 6.21: Efficiency as a function of NN output for neural networks trained with different numbers of training events. Zoom on the *signal* region ($900 \leq$ NNout < 1000).

6.3.3 Single NN

The first technique we explored is based on a single neural network, trained with both global and τ specific variables [6]. We made 30000 iterations with a validation every 100 iterations. The architecture used is 16-20-1. In figg. 6.22 and 6.23 training results are shown: signal and background are very well separated and efficiency on signal is very high for all cuts on NN output; if we consider NNout = 950, we have 77.3% efficiency on signal. Purity as a function of NN output has a very good trend, 99.7% at NNout = 950. Thus purity versus efficiency plot approaches the ideal one - a point at 100%. But such good training results do not mean that we will obtain the same results on real data. To check goodness of training we have to apply it to a sample different from training one. As signal we use 1100 events not overlapping with training ones and taken from *ttopvh* dataset (M_{top} = 175 GeV). These 1100 events correspond to 43 signal events in 311 pb^{-1}. Background events are *all gset0d* events passing preliminary cuts. Validation results are shown in figg. 6.24 and 6.25. Signal-background separation is good and so is efficiency on signal. As far as purity is concerned, the net can't reach the high values obtained during training, where signal and background were present in the same proportions. Such low purity values are reflected in the $\frac{S}{N}$ ratios as a function of NN output reported in table 6.6, together with efficiency and number of signal events expected in 311 pb^{-1}.

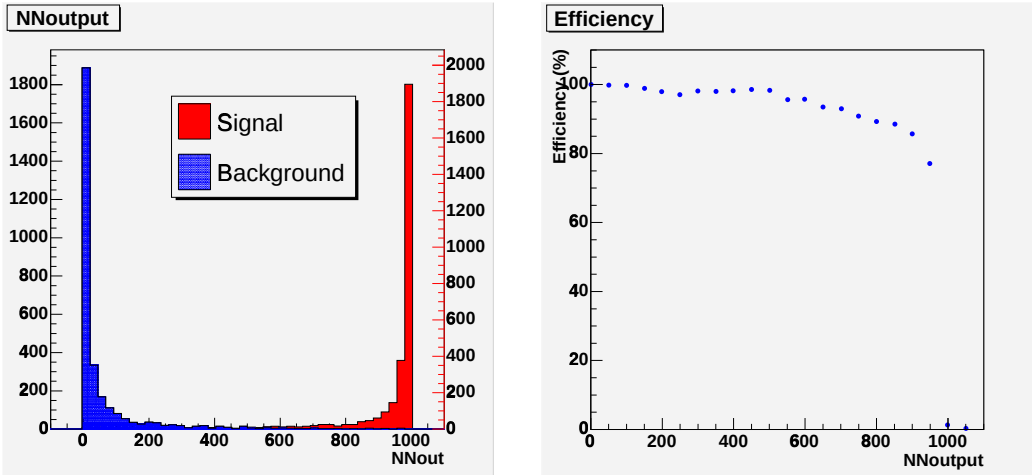

Figure 6.22: Single NN training results: output distribution (left) and efficiency on signal as a function of cut on NN output (right)

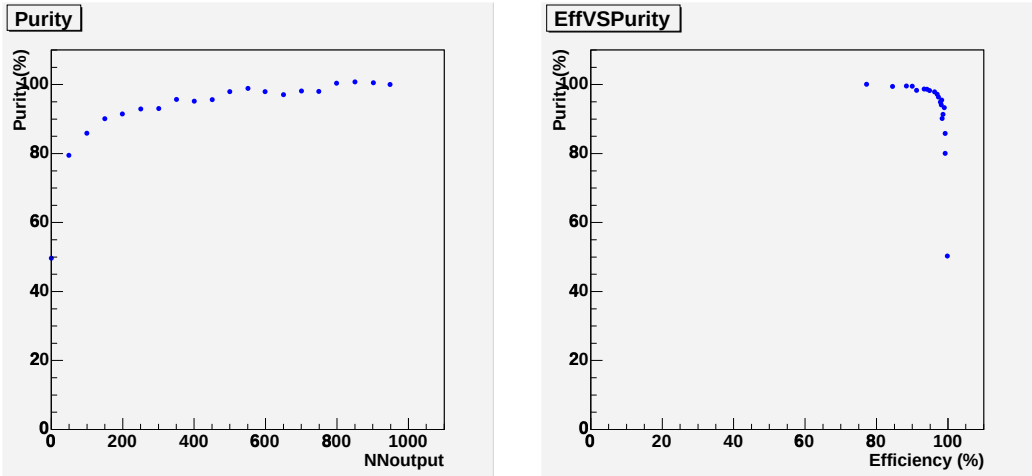

Figure 6.23: Single NN training results: purity as a function of cut on NN output (left) and efficiency vs purity (right)

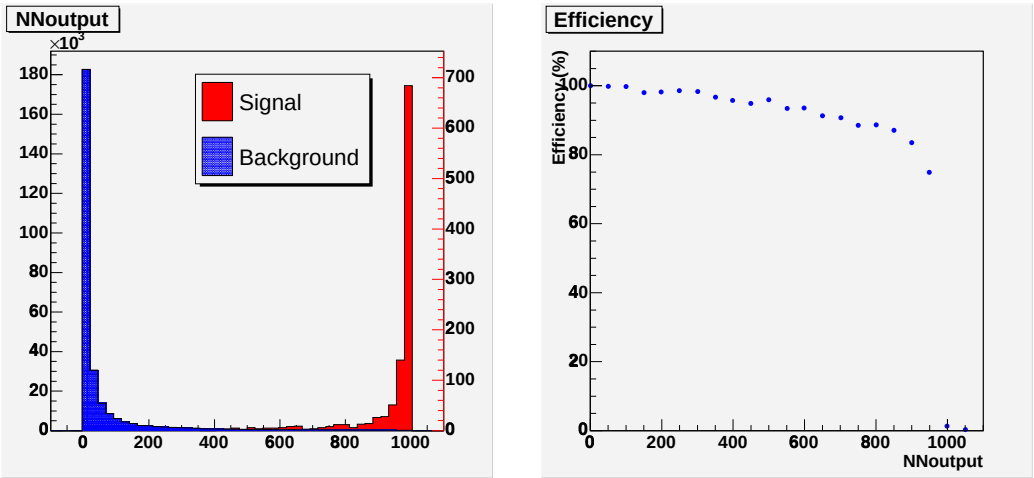

Figure 6.24: Single NN validation results: output distribution (left) and efficiency on signal as a function of cut on NN output (right)

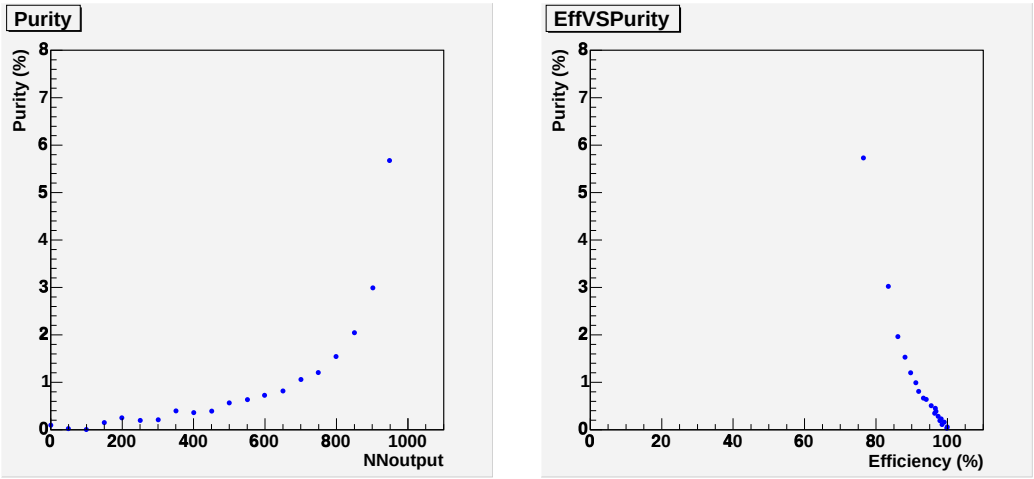

Figure 6.25: Single NN validation results: purity as a function of cut on NN output (left) and efficiency vs purity (right)

Cut	True Pos.	False Neg.	False Pos.	Eff(%)	Pur(%)	S/N	In 311pb^{-1}
0	1100	0	278337	100	0.0	1/5978	46.6
50	1098	2	60813	99.8	0.1	1/1308	46.4
100	1092	8	40230	99.3	0.1	1/870	46.2
150	1087	13	30318	98.8	0.2	1/658	46.0
200	1079	21	24351	98.1	0.2	1/532	45.7
250	1075	25	20022	97.7	0.2	1/439	45.5
300	1067	33	16638	97.0	0.3	1/367	45.2
350	1063	37	13773	96.6	0.3	1/305	45.0
400	1062	38	11799	96.5	0.4	1/261	44.9
450	1057	43	10065	96.1	0.4	1/224	44.7
500	1050	50	8577	95.5	0.5	1/192	44.4
550	1040	60	7347	94.5	0.6	1/166	44.0
600	1028	72	6234	93.5	0.7	1/142	43.5
650	1016	84	5190	92.4	0.8	1/120	43.0
700	1002	98	4326	91.1	1.0	1/101	42.4
750	989	111	3435	89.9	1.2	1/81	41.9
800	971	129	2652	88.3	1.5	1/64	41.1
850	948	152	2004	86.2	2.0	1/49	40.1
900	913	187	1248	83.0	3.0	1/31	38.6
910	902	198	1095	82.0	3.4	1/28	38.2
920	888	212	963	80.7	3.8	1/25	37.6
930	877	223	813	79.7	4.4	1/21	37.1
940	864	236	729	78.5	4.8	1/19	36.6
950	844	256	585	76.7	5.8	1/15	35.7
960	809	291	468	73.5	6.8	1/13	34.2
970	760	340	369	69.1	8.0	1/10	32.2
980	684	416	255	62.2	10.2	1/8	28.9
990	568	532	120	51.6	16.7	1/4	24.0
995	426	674	75	38.7	19.4	1/3	18.0

Table 6.6: Single NN validation results: *True Positives* are signal events passing the cut, *False Negatives* are signal events not passing the cut, *False Positives* are background events (*gset0d* events) passing the cut. $\frac{S}{N}$ ratio is calculated as $\frac{S}{B-S}$ where B are *False Positives* and S is the number of signal events expected in 311 pb^{-1} and passing the cut.

NN1	NN2
nmuhits	ntrks10deg
aplanarity	nisopi0s
centrality	etaeta
sphericity	phiphi
ΣE_{T3}	trackiso
Missing Et significance	caloetiso
	visible mass
	angleseedtocluster
	τ boost

Table 6.7: Input variables for *NN1* and *NN2*.

6.3.4 Two NN cascade

The second technique we're going to illustrate is based on two NNs in cascade. The idea is to apply the second net only to those events whose *NN1* output is greater than a certain value, *BestCut*. *NN2* thus works on events that tend all to be recognized as signal by *NN1*, even background ones, and makes a further cleaning in the sample. In table 6.7 the input variables for the two nets are listed: with the first net we want to concentrate the selection on generical $t\bar{t}$ multijet events, so we use in input mostly *global* variables.The second net is trained by τ specific variables only, to identify specifically $t\bar{t} \to \tau + jets$ events.

NN1 training and validation

NN1 training procedure is the same as for the Single NN: same training sample and same number of iterations; the architecture however changes and the one which gives the best results appears to be 7-14-1. In figg. 6.26 - 6.29 training and validation results are shown. We are using just a subset of the variables used for the single NN technique, so we expect that the net has more difficulty in discriminating signal from background at this stage. The second net will help increasing $\frac{S}{N}$ assuming that we optimally determine the cut on the first net.

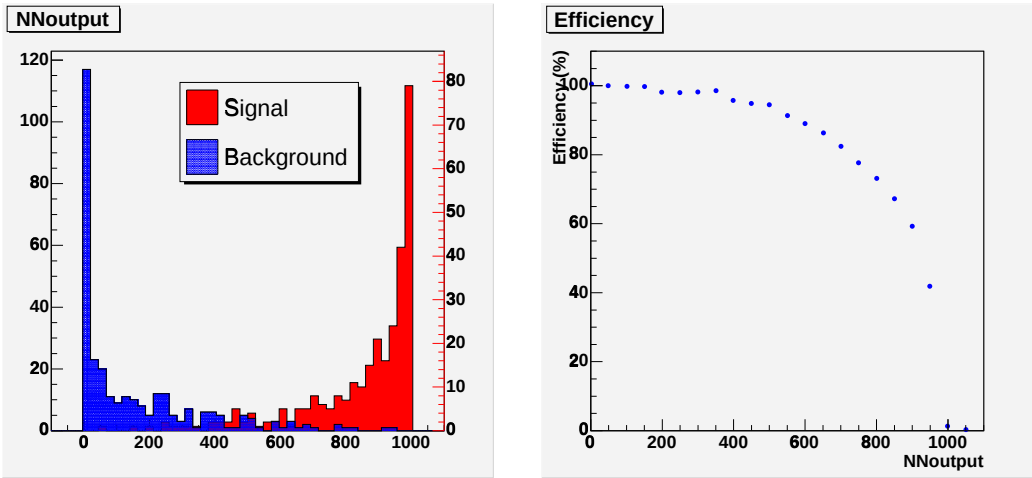

Figure 6.26: NN1 training results: output distribution (left) and efficiency on signal as a function of cut on NN output (right)

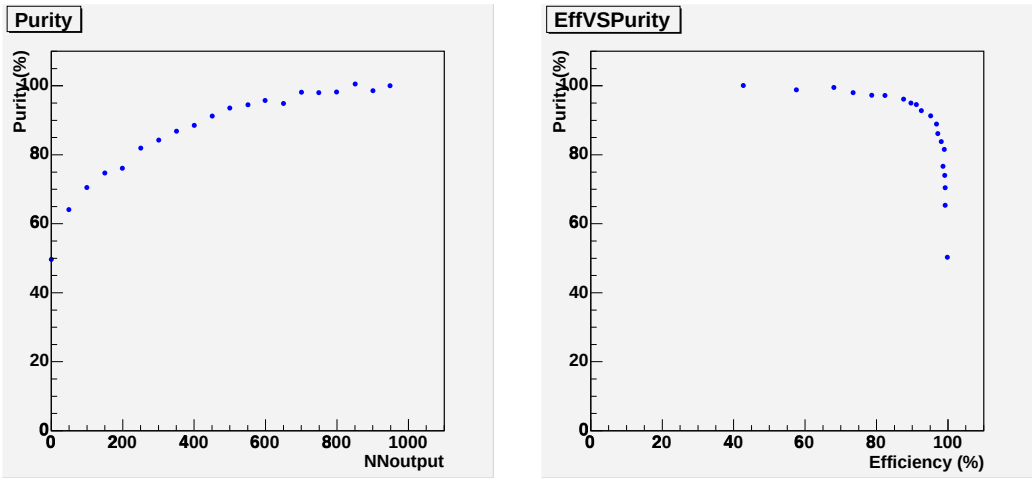

Figure 6.27: NN1 training results: purity as a function of cut on NN output (left) and efficiency vs purity (right)

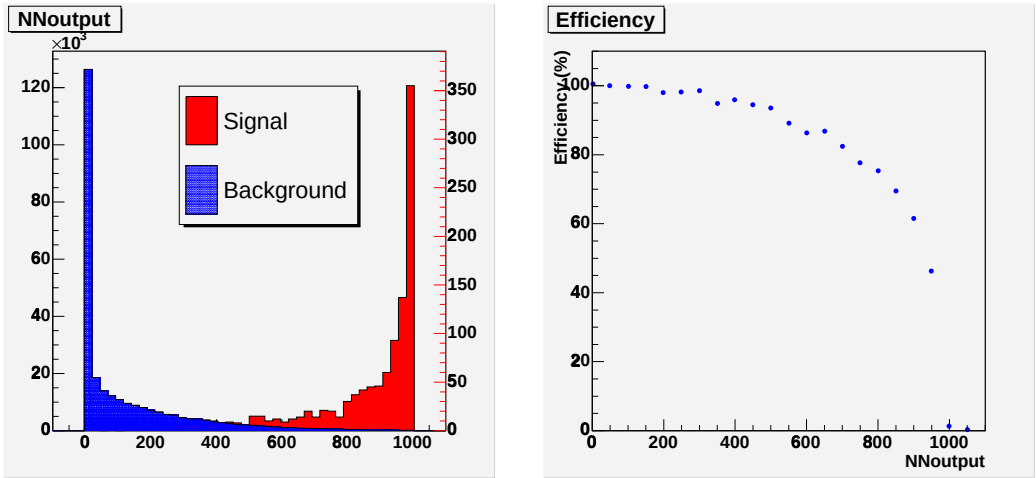

Figure 6.28: NN1 validation results: output distribution (left) and efficiency on signal as a function of cut on NN output (right)

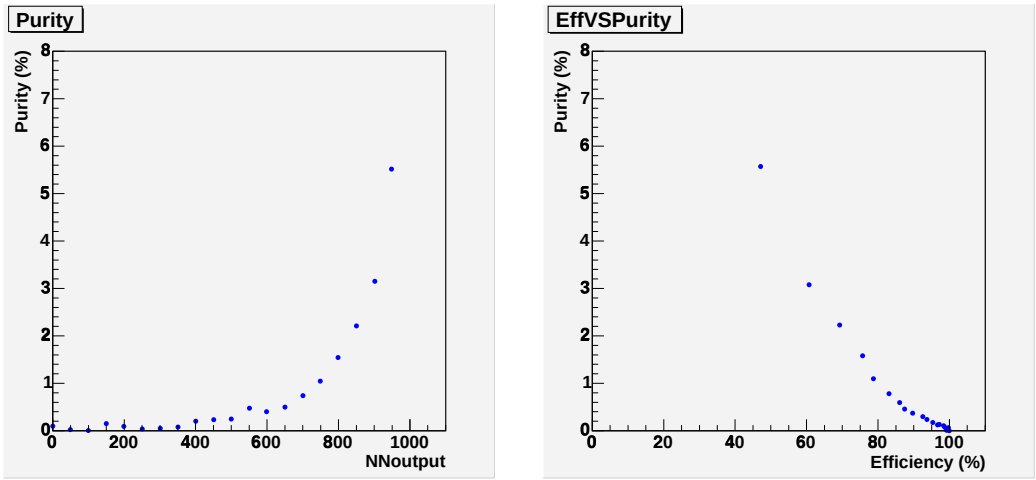

Figure 6.29: NN1 validation results: purity as a function of cut on NN output (left) and efficiency vs purity (right)

BestCut determination

To determine *BestCut*, for each MC mass[8] we make 100 samples with realistic proportions of signal and background (initial $\frac{S}{N} \sim \frac{1}{6000}$). Signal events are randomly picked from MC datasets *ttopbh,...ttopgh*, while as background we consider all *gset0d* events having missing transverse energy significance less than 3 GeV$^{1/2}$. This cut assures that we are considering only background events, where signal contamination is negligible at all (recall $\not{E}_T$ distribution in figure 6.7). After *NN1* application, we calculate signal significance $\frac{S}{\sqrt{S+B}}$ for different values of *NN1* output cut. In figure 6.30 we plot signal significance as a function of *NN1* output for all masses. each point is the mean value over the 100 measures. The value that maximizes $\frac{S}{\sqrt{S+B}}$ is 960, independentently from top mass.

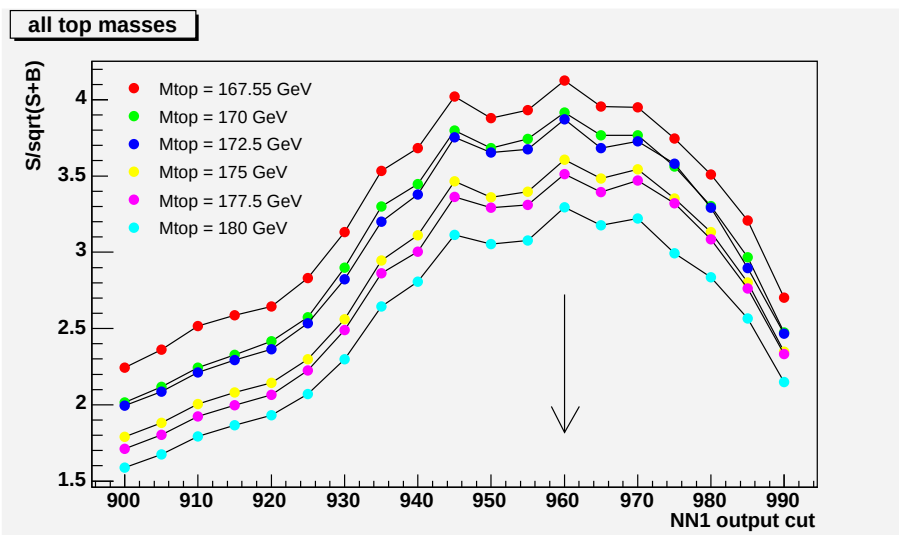

Figure 6.30: Signal significance as a function of *NN1* output cut; the arrow indicates the value chosen as *BestCut* (960)

NN2 training and validation

Of course *NN2* should be trained with events whose *NN1* output is greater than *BestCut*. Only 291 *gset0d* events survive this cut, however, and among them there is

[8] To check if *BestCut* depends on top mass

a significant amount of signal, thus we can't consider all those events as background. As, at the moment, we do not have at our disposal MC events simulating our background, we decide to train *NN2* with the same sample used for Single NN and *NN1*. In figg. 6.31 and 6.32 *NN2* training results are shown. Validation is made on those MC and *gset0d* events passing *NN1 BestCut* and results are plotted in figg. 6.33 and 6.34. Final $\frac{S}{N}$ ratios as a function of cut on *NN2* output are listed in table 6.8.

6.3.5 Comparing the two techniques

To decide which technique gives the best results we compare the $\frac{S}{N}$ we obtain in the two cases. With a single neural network the best value we obtain is $\frac{1}{3.2}$ if we cut at NNout = 995: we are left with 75 events among which we expect **18** signal events. With two neural networks in cascade, if we cut at *NN2* output = 450, we have $\frac{S}{N} = \frac{1}{1.3}$ and the same acceptance on signal (42 events surviving the cut, **18** signal events expected). If we accept a lower efficiency on signal, we can obtain even better signal-to-noise ratios, as seen in table 6.8. Thus for our future analysis, we prefer the double NN technique.

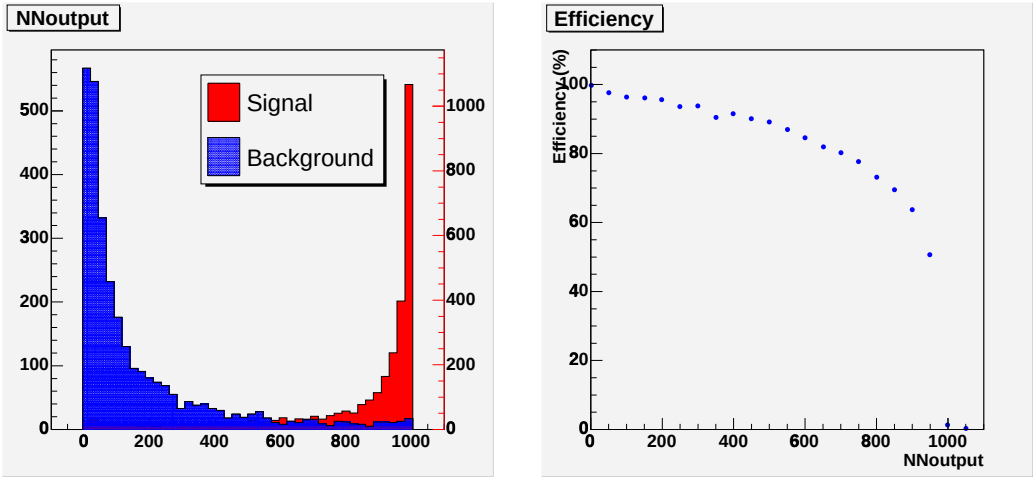

Figure 6.31: NN2 training results: output distribution (left) and efficiency on signal as a function of cut on NN output (right)

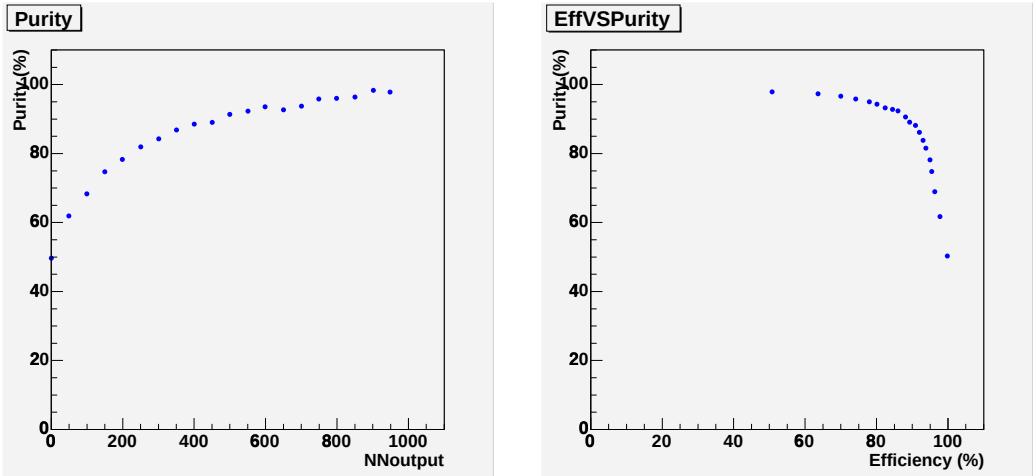

Figure 6.32: NN2 training results: purity as a function of cut on NN output (left) and efficiency vs purity (right)

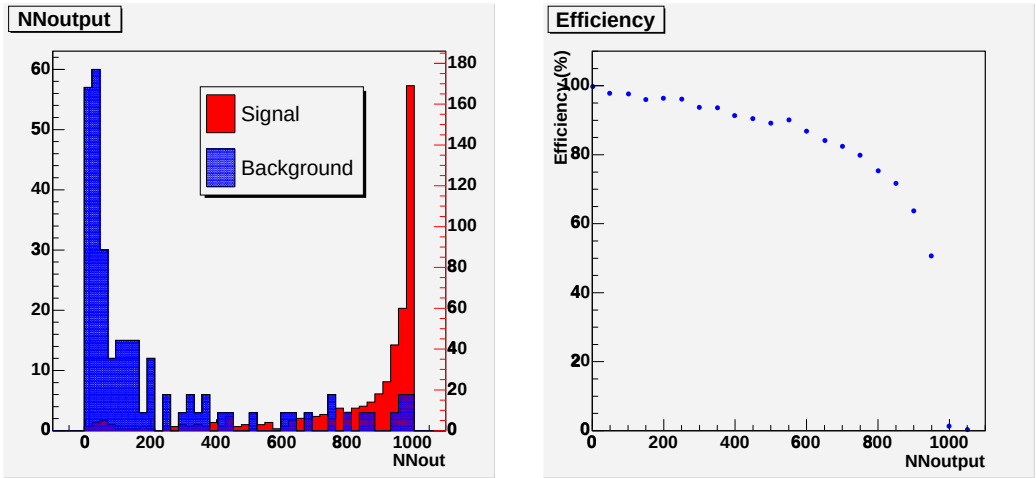

Figure 6.33: NN2 validation results: output distribution (left) and efficiency on signal as a function of cut on NN output (right)

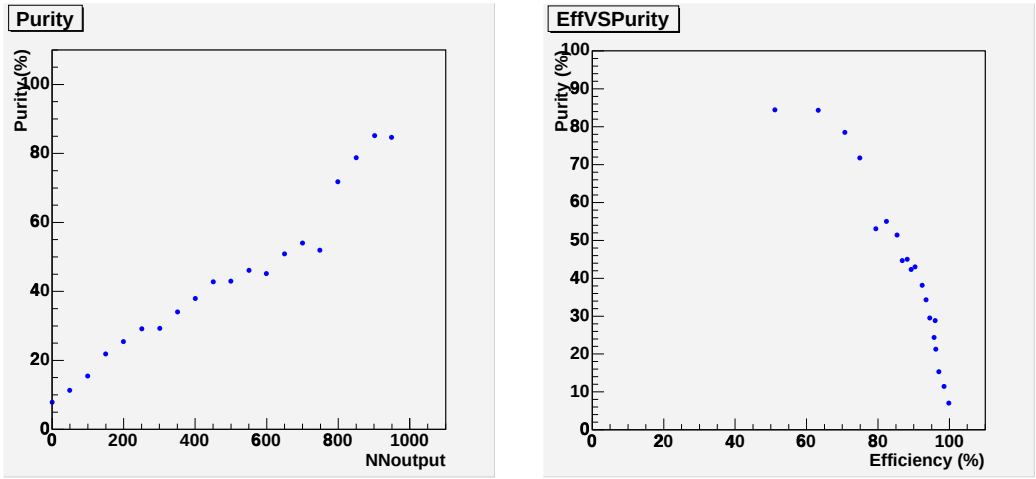

Figure 6.34: NN2 training results: purity as a function of cut on NN output (left) and efficiency vs purity (right)

Cut	True Pos.	False Neg.	False Pos.	Eff(%)	Pur(%)	S/N	In 311pb^{-1}
0	471	0	291	100	6.8	1/13.6	19.9
50	463	8	165	98.3	11.9	1/7.4	19.6
100	456	15	123	96.8	15.7	1/5.4	19.3
150	455	16	93	96.6	20.7	1/3.9	19.3
200	452	19	78	96.0	24.5	1/3.1	19.1
250	450	21	66	95.5	28.9	1/2.5	19.0
300	445	26	63	94.5	29.9	1/2.3	18.8
350	440	31	54	93.4	34.5	1/1.9	18.6
400	433	38	48	91.9	38.2	1/1.6	18.3
450	425	46	42	90.2	42.8	1/1.3	18.0
500	419	52	42	89.0	42.2	1/1.4	17.7
550	415	56	39	88.1	45.0	1/1.2	17.6
600	409	62	39	86.8	44.4	1/1.3	17.3
650	402	69	33	85.4	51.6	1/1.0	17.0
700	388	83	30	82.4	54.7	1/0.9	16.4
750	374	97	30	79.4	52.8	1/0.9	15.8
800	354	117	21	75.2	71.3	1/0.4	15.0
850	332	139	18	70.5	78.1	1/0.3	14.1
900	298	173	15	63.3	84.1	1/0.23	12.6
950	239	232	12	50.7	84.3	1/0.2	10.1

Table 6.8: *NN2* validation results: *True Positives* are signal events passing the cut, *False Negatives* are signal events not passing the cut, *False Positives* are background events (*gset0d* events) passing the cut. $\frac{S}{N}$ ratio is calculated as $\frac{S}{B-S}$ where B are *False Positives* and S is the number of signal events expected in 311 pb^{-1} and passing the cut.

Bibliography

[1] **Tau working group web page**, `http://www-cdf.fnal.gov/internal/physics/exotic/tau/`

[2] A.Anastassov et al., **Tau Reconstruction Efficiency and QCD Fake Rate for Run 2**, *CDF/ANAL/EXOTIC/PUBLIC/6308*, 2003

[3] Vernon D. Barger, Roger J.N.Phillips, **Collider physics**, *Addison-Wesley*, 1987

[4] **All-hadronic group web page**, `http://www-cdf.fnal.gov/internal/physics/top/RunIIAllHadronic/webpages/index.html`

[5] A.Sidoti et al., **Investigation of $t\bar{t}$ in the full hadronic final state at CDF with a neural network approach**, *Nucl.Instr.Meth. A466:538-550,2001*

[6] S.Amerio, A.Gresele, I.Lazzizzera, **Top $\rightarrow$ tau + jets event selection using neural networks**, *CDF/ANAL/TOP/CDFR/7745*, 2005

Chapter 7

$t\bar{t}$ cross section measurement

In the previous chapter we described two neural network based techniques we explored in order to select $t\bar{t} \rightarrow \tau + jets$ events among events collected with *TOP_MULTI_JET* trigger and we concluded that using two nets in cascade, for a fixed acceptance on signal, we can obtain a greater $\frac{S}{N}$ ratio with respect to a single neural network technique. Now, using this technique, we will turn to $t\bar{t}$ cross section measurement.

Measurement on data will be the last step; first we will introduce the method we use to get cross section and the measurements of the efficiencies we need for the cross section estimate. Then we will check the method applying it to samples where signal and background are known. After an estimate of the systematic uncertainties associated to our measure, we will finally move to data and measure the cross section.

7.1 The method

To select $t\bar{t} \rightarrow \tau + jets$ events we use two NNs in cascade. The first NN, *NN1*, is trained by variables describing globally the event mainly as a $t\bar{t}$ production; then *NN2* is trained by τ specific variables. *NN2* is applied to events whose *NN1* output is greater than a certain value, *BestCut*, optimally selecting $t\bar{t}$ events in data. *BestCut* is in fact the value that maximizes the signal significance $\frac{S}{\sqrt{S+B}}$. *NN2* thus acts on an optimal $t\bar{t}$ sample classified by $NN1$, and makes on it further selection of $t\bar{t} \rightarrow \tau + jets$ events. Then we perform a binned likelihood fit on *NN2* output

in order to estimate the number of signal (n_s) and background (n_b) events in the sample. We then use n_s to calculate cross section according to

$$\sigma_{t\bar{t}} = \frac{n_s}{\epsilon_{sel}\epsilon_{NN1}L} \tag{7.1}$$

where ϵ_{sel} is the efficiency of the cuts we apply before NN analysis, ϵ_{NN1} is *NN1 BestCut* efficiency and L is the integrated luminosity.

The likelihood we use is based on the expected probabilities $f_s(i)$ and $f_b(i)$ of finding signal and background in the *ith* of the bins of *NN2* output distribution. These probabilities are given by the fraction of signal and background events we find in the bins when applying *NN2* to the training known sample. Thus the expected number of signal events in the *ith* bin is $f_s(i)$ of total n_s while

$$\frac{f_s(i)n_s}{n_s + n_b} \tag{7.2}$$

is the chance of finding signal in that bin; analogously

$$\frac{f_b(i)n_b}{n_s + n_b} \tag{7.3}$$

is the chance of finding background in the *ith* bin. Consequently

$$\left(\frac{n_b f_b(i) + n_s f_s(i)}{n_s + n_b}\right)^{k_i} \tag{7.4}$$

is the chance of finding k_i signal and background events in the *ith* bin, while

$$\prod_{i=1}^{\#bin}\left(\frac{n_b f_b(i) + n_s f_s(i)}{n_s + n_b}\right)^{k_i} \tag{7.5}$$

is the chance of the final overall distribution of signal and background in the bins. Our likelihood

$$L(n_s, n_b) = \frac{(n_s + n_b)^N e^{-(n_s+n_b)}}{N!}\prod_{i=1}^{\#bin}\left(\frac{n_b f_b(i) + n_s f_s(i)}{n_s + n_b}\right)^{k_i} \tag{7.6}$$

is given by subjecting 7.5 to the condition that n_s and n_b are the expectation values of a priori poisson probability that out of the N events k of them are found as signal and $N - k$ as background, summing them over all the possbile k:

$$\sum_{k=0}^{N} \frac{n_s^k e^{-n_s}}{k!} \frac{n_b^{N-k} e^{-n_b}}{(N-k)!} = \frac{(n_s+n_b)^N e^{-(n_s+n_b)}}{N!} \tag{7.7}$$

Minimizing $-\ln L$ we obtain n_s and n_b.

7.2 Efficiencies

To measure cross section we need to estimate the efficiencies on signal of prerequisites, b-tagging, τ candidate requests and *NN1* output cut.

Prerequisite efficiency

In fig. 7.1 and table 7.1 prerequisite efficiencies for different masses are shown. Efficiency increases with increasing masses: trigger requirements[1] are based on calorimetric energies and an increasing top mass produces jets with increasing energies; thus a greater number of events survives trigger cuts.

[1]At *Level 1* 1 calorimetric tower with $E_T \geq$ 10GeV, at Level 2 4 calorimetric clusters with $E_T^{cluster} \geq 15GeV$ and $\Sigma E_T \geq$ 125 GeV, at *Level 3* 4 jets with $|\eta| \leq 2.0$ and $E_T^{jet} \geq$ 10 GeV

Dataset	M_{top}(GeV)	Good run	prereq	Eff.	$\tau + jets$	Eff.
ttopbh	167.55	185450	99406	**0.536 ± 0.003**	10336	**0.0557 ± 0.0006**
ttopch	170	207609	112933	**0.544 ± 0.003**	11602	**0.0559 ± 0.0005**
ttopdh	172.5	207941	114020	**0.548 ± 0.003**	11963	**0.0575 ± 0.0005**
ttopvh	175	193855	107050	**0.552 ± 0.003**	11286	**0.0582 ± 0.0006**
ttopfh	177.5	199877	111152	**0.556 ± 0.003**	11915	**0.0596 ± 0.0006**
ttopgh	180	203770	113815	**0.558 ± 0.003**	12168	**0.0597 ± 0.0006**

Table 7.1: Effects of prerequisites on events used to build training samples.

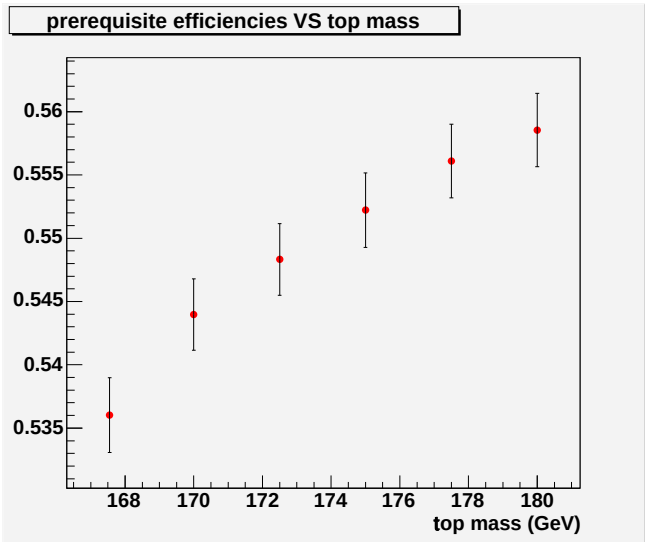

Figure 7.1: Prerequisite efficiency as a function of top mass

b-tagging efficiency

b-tagging efficiency is calculated as the ratio between events having at least on b-tagged jet and all events. It is shown in fig 7.2 and table 7.2 and within statistical errors it is independent from top mass.

When applying the method to data, we shall multiply b-tagging efficiency estimated on MC events by a *scale factor*[2] [1], that accounts for imperfections in MC simulations.

Dataset	top mass (GeV)	Good run evts	At least 1 b tag	Efficiency
ttopbh	167.55	10336	6794	**0.657 ± 0.014**
ttopch	170	11602	7621	**0.657 ± 0.014**
ttopdh	172.5	11963	7911	**0.661 ± 0.013**
ttopvh	175	11286	7340	**0.650 ± 0.014**
ttopfh	177.5	11915	7782	**0.653 ± 0.013**
ttopgh	180	12168	7906	**0.650 ± 0.013**

Table 7.2: b-tagging efficiency on $t\bar{t} \rightarrow \tau + jets$ events

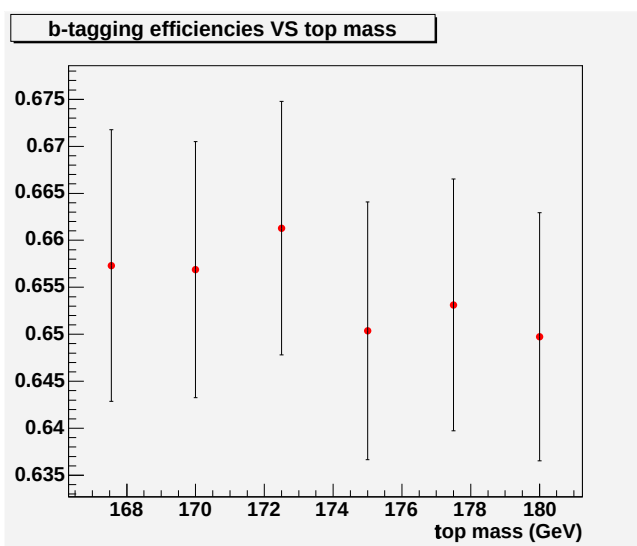

Figure 7.2: b-tagging efficiency on $t\bar{t} \rightarrow \tau + jets$ events as a function of top mass

[2] 0.909±0.015(stat)±0.058(syst)

τ finder algorithm efficiency

τ finder algorithm efficiency on our channel is estimated counting all $t\bar{t} \rightarrow \tau + jets$ events with a τ candidate matching an hepg τ. In table 7.3 and fig. 7.3 efficiency as a function of top mass is shown. Within statistical errors it is independent from top mass.

Dataset	top mass (GeV)	Good run evts	Evts with matched τ	Efficiency
ttopbh	167.55	10336	6193	**0.599 ± 0.014**
ttopch	170	11602	6807	**0.587 ± 0.013**
ttopdh	172.5	11963	7138	**0.597 ± 0.013**
ttopvh	175	11286	6632	**0.588 ± 0.013**
ttopfh	177.5	11915	7160	**0.601 ± 0.013**
ttopgh	180	12168	7295	**0.600 ± 0.012**

Table 7.3: tau finder efficiency on $t\bar{t} \rightarrow \tau + jets$ events

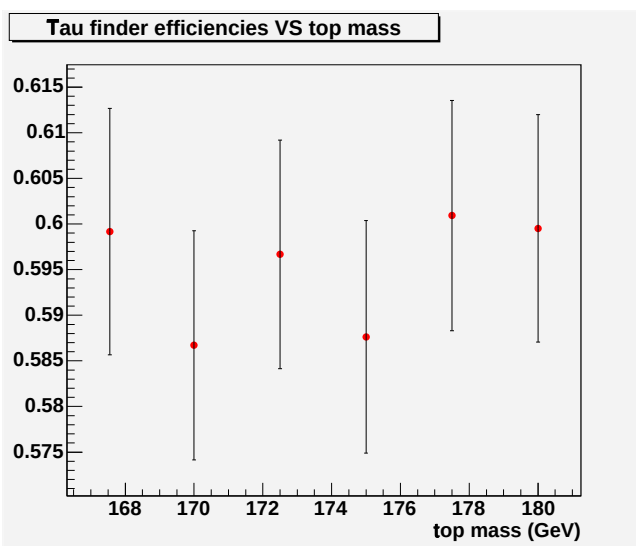

Figure 7.3: τ finder efficiency on $t\bar{t} \rightarrow \tau + jets$ events as a function of top mass

NN1 output cut efficiency

NN1 output cut efficiency is calculated submitting to *NN1* all $t\bar{t} \rightarrow \tau + jets$ events passing prerequisites, b-tagging and τ matching cuts and counting those whose NN output is greater than $BestCut = 960$. In fig. 7.4 and table 7.4 *NN1* output cut efficiency as a function of top mass are reported. No sensible dependence on top mass is evident, in particular for masses from 170 to 180 GeV/c^2.

Dataset	top mass (GeV)	Tot events	NNout > 960	Efficiency
ttopbh	167.55	4118	1709	**0.415 ± 0.017**
ttopch	170	4481	1935	**0.432 ± 0.016**
ttopdh	172.5	4696	2054	**0.437 ± 0.016**
ttopvh	175	4331	1890	**0.436 ± 0.017**
ttopfh	177.5	4710	2036	**0.432 ± 0.016**
ttopgh	180	4740	2082	**0.439 ± 0.016**

Table 7.4: *NN1* output cut efficiency on $t\bar{t} \rightarrow \tau + jets$ events

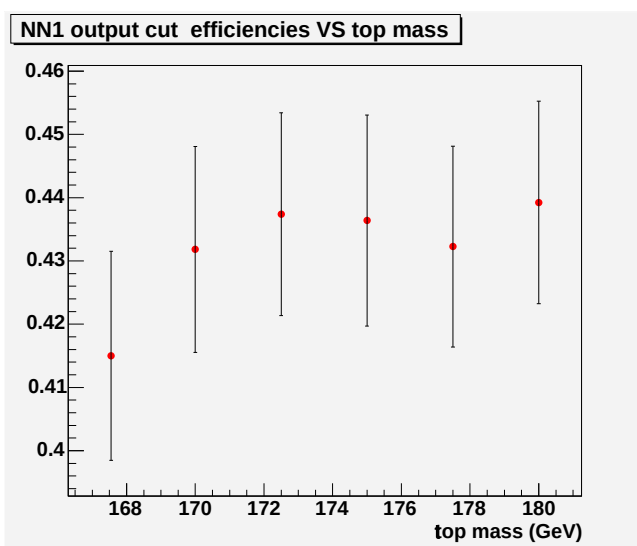

Figure 7.4: *NN1* output cut efficiency on $t\bar{t} \rightarrow \tau + jets$ events as a function of top mass

7.3 Validation on MC samples

To validate our method, we use those 100 samples for each mass built to determine *BestCut* (section 6.3.4). For each sample we perform the following steps:

1. apply *NN1*;
2. select all events whose *NN1* output is greater than *BestCut*;
3. apply *NN2*;
4. perform a binned likelihood fit on *NN2* output distribution to extract the number of signal and background events in the sample;
5. calculate cross section.

In section 6.3.4, looking at signal significance as a function of cut on *NN1* output, we determined that *BestCut* is 960. As a first check of out method, we want to study cross section as a function of *BestCut*, to assure that 960 gives the best result. For each of 100 samples, we repeat steps 1-5 varying *BestCut* from 700 to 990.

Starting with loose *NN1* cut, we expect to have poor final purity and an overestimation of the cross section, such effect reducing more and more with tighter cuts: the hope is that a cross section plateau in correspondence of the tighter cuts comes out, corresponding to the correct value of cross section. Indeed this is what one can see in figures 7.5, 7.6 and 7.7. The red line is the theoretical value of cross section calculated from [2] and [3]. Each point is the mean over the 100 measures (one for each sample) and the error bar is the standard deviation. Increasing *NN1* output cut, standard deviation increases because efficiency on signal lowers. For each mass the plateau begins at *NN1* cut 910 and for almost all masses the cross section value approaching more the theoretical one is found for *BestCut* 960.

In fig. 7.8 cross section as a function of top mass is plotted for *BestCut* 960. The theoretical values are superimposed and, as in the previous plots, each point is the mean over the 100 values obtained for each mass. Within statistical errors, cross section measurements are compatible with theoretical values.

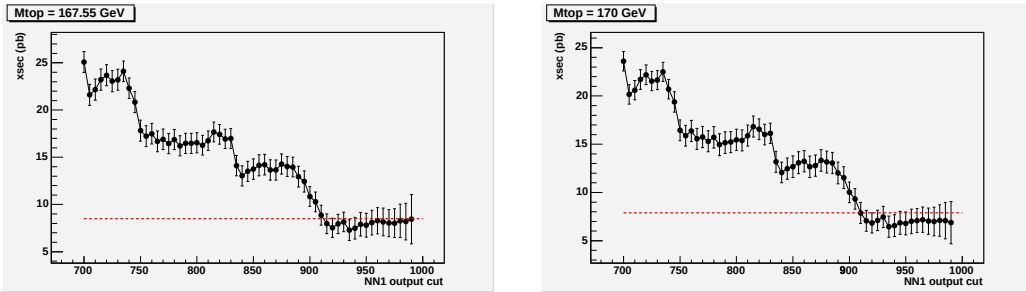

Figure 7.5: Cross section as a function of *NN1* output cut for samples with known signal and background; red line is the theoretical value. **left**: $\sigma_{th} = 8.5$ pb; **right**: $\sigma_{th} = 7.9$ pb. Each point is the mean over 100 measures and error bar is standard deviation.

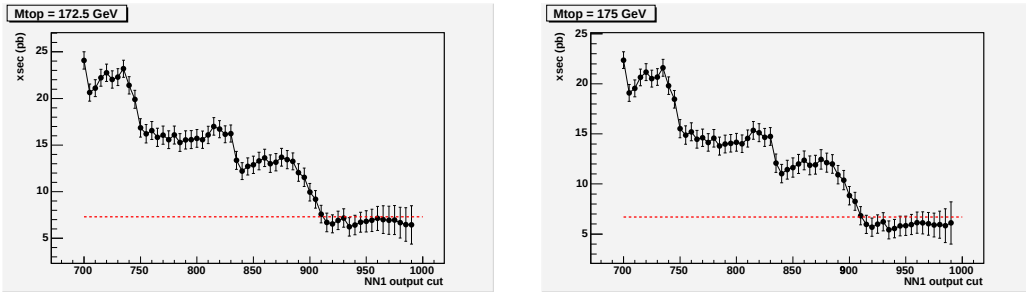

Figure 7.6: Cross section as a function of *NN1* output cut for samples with known signal and background; red line is the theoretical value. **left**: $\sigma_{th} = 7.3$ pb; **right**: $\sigma_{th} = 6.7$ pb.

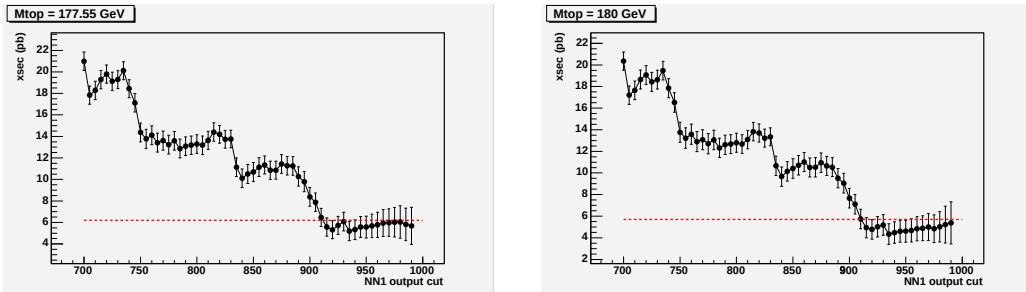

Figure 7.7: Cross section as a function of *NN1* output cut for samples with known signal and background ; red line is the theoretical value. **left**: $\sigma_{th} = 6.2$ pb; **right**: $\sigma_{th} = 5.7$ pb.

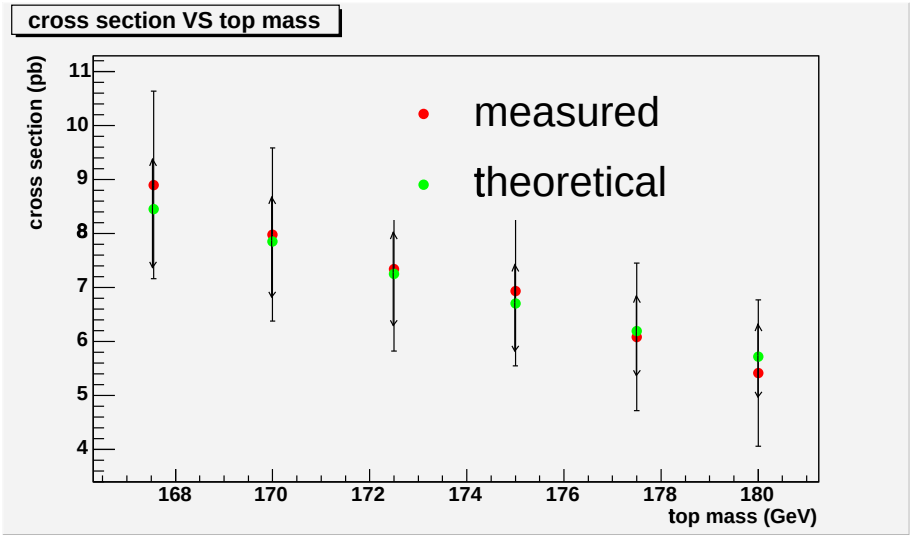

Figure 7.8: Cross section measured on validation samples as a function of top mass; each point is the mean over 100 measurements made on samples with known signal and background events. The error bar with the arrows is the theoretical error, the other error bar is the standard deviation over 100 measurements.

7.4 Systematics

Before moving to measurement on data, we want to estimate the systematic uncertainties which will be associated to our cross section estimate.

7.4.1 Systematic error introduced by NNs

Using neural networks we introduce a systematic error due to the finite number of events used for training and to the training procedure itself. The greater the number of events we use for training, the better the separation between signal and background; as we do not have an infinite number of events at our disposal, we obviously introduce a systematic error in our measure. Moreover, during training procedure, the trajectory followed to locate an optimal minimum on the space of weights associated to NN nodes, can differ from training to training: we recall that a move on such space consists of a bit-flip in the strings of weights and the bit to flip is chosen randomly.

To quantify the systematic error introduced by the use of NNs, we apply the following procedure:

- we build training samples with different cardinalities[3] and for each cardinality we build four non overlapping samples[4];
- for each sample we make 5 trainings with the same architectures for both *NN1* and *NN2*;
- for each of the 5 trainings and for each of the 4 non-overlapping samples with the same cardinality, we make 100 applications[5] (pseudo-experiments) on samples with n_s signal events and n_b background events. The first are selected randomly from *ttopvh* MC dataset and are the number of expected signal events[6] in 311 pb^{-1}. n_b are all *gset0d* events whose *missing* E_T significance is

[3] 2000, 6000, 10000, 12000 events.

[4] This small number of non overlapping samples has been forced by the limited number of signal events at our disposal ($\sim$ 24000 events)

[5] $NN1 \rightarrow BestCut \rightarrow NN2 \rightarrow fit$.

[6] 47 events if we consider top mass = 175 GeV.

smaller than 3.0 GeV $^{\frac{1}{2}}$. This cut assures that in background events there is no signal contamination.

- having done 100 pseudo-experiments for each of the 5 trainings made with each of the 4 non overlapping samples, we calculate the overall mean of the number of events estimated at the end of each pseudo-experiment:

$$(\bar{n}_s)^{overall} = \frac{1}{N_{non-overl} N_{tr} N_{pe}} \sum_{i=1}^{N_{non-overl}} \sum_{j=1}^{N_{tr}} \sum_{k=1}^{N_{pe}} n_{s_{ijk}} \tag{7.8}$$

  where $N_{non-overl}$ is the number of non overlapping samples, N_{tr} is the number of trainings we made for each non-overlapping sample, N_{pe} is the number of pseudo-experiments and $n_{s_{ijk}}$ is the number of events estimated by the kth pseudoexperiment of the jth training made with the ith non overlapping sample;

- to quantify the systematic error introduced by the use of different non-overlapping samples and trainings, for each training j made with a specific non-overlapping sample i we calculate a mean just over the pseudo-experiments:

$$(\bar{n}_s)^{pe}_{ij} = \frac{1}{N_{pe}} \sum_{k=1}^{N_{pe}} n_{s_{ijk}} \tag{7.9}$$

- we assign as systematic error to n_s the standard deviation of $(\bar{n}_s)^{pe}$ with respect to $(\bar{n}_s)^{overall}$:

$$\delta^2_{n_s} = \frac{1}{N_{non-overl} N_{tr} - 1} \sum_{j=1}^{N_{non-overl}} \sum_{i=1}^{N_{tr}} |(\bar{n}_s)^{pe}_{ij} - (\bar{n}_s)^{overall}|^2 \tag{7.10}$$

  In the ideal case, the use of a certain training sample instead of another with the same number of events, should not introduce any error and the standard deviation should be null. But in the real case, using a training sample with a *finite* number of events introduces a systematic effect due to the choice of the training events. The same for different trainings made with the same sample: if we could perform infinite iterations we would always reach the same result but as we do a finite number of iterations we introduce a systematic effect on our measurement.

# events	$\bar{n}_s$	δ_{n_s}	δ_{n_s}/n_s (%)
2000	18.20	1.37	**7.5**
6000	17.85	1.75	**9.8**
10000	18.35	1.84	**10.0**
12000	19.62	1.92	**9.8**

Table 7.5: Systematic error introduced by neural networks computed for training samples with different cardinalities.

In table 7.5 systematic errors for different cardinalities are reported. The systematic error seems to be independent from the number of events used for training, mainly using 6000, 1000 and 12000 events. With 2000 events we obtain a smaller systematic uncertainty: this effect will be investigated in the future repeating the whole procedure with other cardinalities around 2000. For our measure, we used 6000 training events, thus the systematic error introduced by the net is $\frac{\delta_\sigma}{\sigma} = 9.8$ %.

7.4.2 Other sources of systematic errors

MC generator

For NN training we used MC events generated with Herwig. Different MC generators can differ in the fragmentation and generation models they use, introducing thus a systematic effect in the measure. To estimate the systematic effect introduced by the choice of a particular MC generator, we apply to event generated with Phytia the same cuts applied to Herwig events and compare the efficiencies obtained (table 7.6).

Efficiency calculated on Herwig events is $\epsilon_H = 0.0223 \pm 0.0004$ while on Phytia events we obtain $\epsilon_P = 0.0217 \pm 0.0004$. We derive a relative systematic uncertainty

$$\left(\frac{\delta\epsilon}{\epsilon}\right)_{MC} = \frac{1}{2}\frac{|\epsilon_H - \epsilon_P|}{\epsilon_H} = 1.3\% \tag{7.11}$$

	Pythia (ttop0z)	*Herwig (ttopvh)*
Total MC events	190831	193855
Good Run	170987	125489
Trigger	108086	120545
z vertex	103660	107050
No lepton	92669	11286
tau+jets & at least one b tag	6181	6632
matched tau	3711	4331

Table 7.6: Effects of cuts on events generated with different MC generators, Pythia (*ttop0z*) and Herwig (*ttopvh*), both generated with the same top mass (M_{top} = 175 GeV))

ISR/FSR

To calculate the systematic uncertainty that arises from incomplete knowledge of initial and final state radiation (*ISR*/*FSR*), we apply our cuts to events generated with more/less *ISR*/*FSR*, i.e. with or without simulation of initial and final state radiation. We use the following samples, generated with Pythia:

- *ttopbr* (less *ISR*)
- *ttopdr* (more *ISR*)
- *ttopfr* (less *FSR*)
- *ttopkr* (more *FSR*)

Effects of cuts on these sample are listed in tables 7.7 and 7.8. We obtain the following efficiencies:

- $\epsilon_{ISR}^{less} = 0.0203 \pm 0.0004$
- $\epsilon_{ISR}^{more} = 0.0210 \pm 0.0004$
- $\epsilon_{FSR}^{less} = 0.0206 \pm 0.0004$
- $\epsilon_{FSR}^{more} = 0.0212 \pm 0.0004$

	less ISR	*more ISR*
Total MC events	178448	174436
Good Run	108286	107784
Trigger	104846	104324
z vertex	93984	93519
No lepton	9269	9276
tau+jets & at least one b tag	6034	6057
matched tau	3620	3656

Table 7.7: Effects of cuts on events generated with/without initial state radiation.

We compare the previous efficiencies with the value obtained on the sample generated with Pythia at the same top mass (*ttopel* dataset, table 7.9, $\epsilon_P = 0.0214 \pm 0.0001$). As $|\epsilon_{ISR}^{less} - \epsilon_P| > |\epsilon_{ISR}^{more} - \epsilon_P|$ and $|\epsilon_{FSR}^{less} - \epsilon_P| > |\epsilon_{FSR}^{more} - \epsilon_P|$ we obtain

$$\left(\frac{\delta\epsilon}{\epsilon}\right)_{ISR} = \frac{1}{2}\frac{|\epsilon_{ISR}^{less} - \epsilon_P|}{\epsilon_P} = 3.2\% \tag{7.12}$$

and

$$\left(\frac{\delta\epsilon}{\epsilon}\right)_{FSR} = \frac{1}{2}\frac{|\epsilon_{FSR}^{less} - \epsilon_P|}{\epsilon_P} = 2.5\% \tag{7.13}$$

Adding in quadrature the two contributes, we obtain an overall systematic uncertainty $\left(\frac{\delta\epsilon}{\epsilon}\right)_{ISR/FSR} = 4.1\%$.

Jet energy scale

To compute systematics due to jet energy corrections, we vary all the different sources of jet energy corrections by $\pm 1\sigma$, recalculate all variables related to jet energies and finally calculate the variation on signal efficiency. We recall that our method consists in the following steps:

1. application of prerequisites;
2. selection of events with at least one b-tagged jet;
3. selection of events with at least one τ candidate;

	less FSR	*more FSR*
Total MC events	172562	184271
Good Run	105672	114542
Trigger	102388	110751
z vertex	91932	99369
No lepton	9084	9924
tau+jets & at least one b tag	5892	6506
matched tau	3558	3898

Table 7.8: Effects of cuts on events generated with/without final state radiation.

	ttopel ($M_{top} = 178\ GeV$)
Total MC events	1150043
Good Run	1021924
Trigger	647365
z vertex	622086
No lepton	556572
tau+jets & at least one b tag	36856
matched tau	21889

Table 7.9: Effects of cuts on *ttopel* dataset.

4. application of the first neural network *NN1*;
5. selection of those events whose *NN1* output is greater than a certain value *BestCut*;
6. application of a second neural network *NN2*.

The effects of jet energy corrections enter our method in steps 2 and 4: b-tagged jets are searched among those with corrected transverse energy greater than 15 GeV and *NN1* is trained with variables strongly related to jet corrections, such as ΣE_{T3}, sphericity, aplanarity and centrality. *NN2*, on the contrary, is trained only with τ specific variables and τ energy and momentum aren't subjected to jet corrections. As far as step 2 is concerned, the selection of jets with $E_{T_{corr}} > 15$ GeV is affected very slightly by $\pm 1\sigma$ variation of jet corrections: the vast majority of the jets have energies greater than 15 GeV, and the $\pm 1\sigma$ variation increases/decreases the number of surviving events of a few units. So we decide to estimate the systematic uncertainty due to jet corrections comparing *NN1 BestCut* efficiencies.
In table 7.10 we show the effects of *BestCut* on *NN1* outputs obtained using samples with default and $\pm 1\sigma$ correction factors.

We obtain the following efficiencies:

- $\epsilon_{def} = 0.98 \pm 0.02\%$
- $\epsilon_{+1\sigma} = 1.00 \pm 0.02\%$
- $\epsilon_{-1\sigma} = 0.95 \pm 0.02\%$

The systematic uncertainty due to jet correction is thus

$$\left(\frac{\delta\epsilon}{\epsilon}\right)_{JES} = \frac{1}{2}\frac{|\epsilon_{+1\sigma} - \epsilon_{-1\sigma}|}{\epsilon_{def}} = 2.6\% \tag{7.14}$$

PDFs

To evaluate the systematic uncertainty due to pdf dependence, we follow the official top group recipe [4] to reweight the $t\bar{t}$ sample using the pdf libraries. We obtain the following systematic uncertainties:

	default corr.fact.	$+1\sigma$	-1σ
Total MC events	219479	219479	219479
Good Run	193855	193855	193855
Trigger	125489	125489	125489
z vertex	120454	120454	120454
No lepton	107050	107050	107050
tau+jets & at least one b tag	7340	7346	7336
matched tau	4331	4334	4329
NN1 out > *BestCut*	1890	1939	1852

Table 7.10: Effects of preliminary and *BestCut* cuts on events corrected with default correction factor and with factors obtained varying all sources of jet corrections by $\pm 1\sigma$ (*ttopvh* dataset).

- 2.0 %, difference between MRST72 and CTEQ5L cross sections;
- 1.6 %, difference between MRST cross sections for two different values of α_s (MRST75 vs MRST72);
- 0.6 %, quadrature sum of the maximum $\pm$ differences for the 20 CTEQ eigenvectors.

Summing in quadrature all these uncertainties, we obtain a total systematic uncertainty of 2.6 %.

7.5 Measurement on data

Let's apply the overall method to all *gset0d* events and extract cross section measurement. We assume that top mass is 175 GeV. In section 7.3, we studied cross section as a function of cut on *NN1* output; a plateau corresponding to the correct cross section value was evidenced for a range of *NN1* output cuts containing the *BestCut* value we had estimated using signal significance (section 6.3.4). In figure 7.9 cross section as a function of cut on *NN1* output is shown for *gset0d* data.

The error bar represents the statistical error which is dominated by the error coming from the fit. The plateau is not as evident as that we obtained during validation on MC samples: in that case it started at *NN1* output cut 910 while here it begins at 960. We explain such effect considering that background events in validation samples were selected requiring missing transverse energy significance smaller than 3 $\mathrm{GeV}^{\frac{1}{2}}$. This cut assures that in background events there is no signal contamination but makes *NN1* and *NN2* tasks easier: the first can discriminate better between signal and background because we do not sumbit it *difficult* background, i.e. background events with large missing energy that can be easily taken for signal; as a consequence, the second net is submitted much cleaner samples even for *NN1* output cuts lower than *BestCut* and the plateu starts well before *BestCut*.

In table 7.11 we report efficiencies and luminosity values used for the measure and the number of signal events extracted by the fit for *BestCut* = 960. The main source of statitiscal error is in the number of signal events: the fit gives an error of 32%. In table 7.12 we summarize systematic errors.

The cross section value for *BestCut* 960 is

$$\sigma(M_{top} = 175GeV) = 7.1 \pm 2.3(stat) \pm 0.9(syst)pb \tag{7.15}$$

in agreement with the theoretical valueof 6.7 pb.

Systematic error is dominated by the systematic uncertainty introduced by our method (9.8%).

The statistical error, as already stated, is dominated by the error on the estimate of n_s. We are quite confident that it can be reduced with higher statistics (1 fb^{-1} of data will soon be available to be processed). Moreover, tau finder algorithm is subject to continuous improvements from which our analysis will surely gain efficiency on signal. Statistical error is also related to the purity level of the sample submitted to the second net: greater the purity smaller the error associated to n_s by the fit. Thus better neural network performances, that can be achieved exploring new architectures and input variables, can improve measurement precision.

n_s	**19.5 ± 6.3**
ϵ_{prereq}	**0.058 ± 0.001**
$\epsilon_{b_{tag}}$	**0.591 ± 0.016**
$\epsilon_{\tau_{find}}$	**0.588 ± 0.013**
ϵ_{NN1}	**0.436 ± 0.017**
Luminosity	**311 pb^{-1}**

Table 7.11: Ingredients for cross section estimate: number of signal events, efficiencies and luminosity values. Errors are statistical only.

NN	**9.8 %**
MC generator	**1.3 %**
ISR/FSR	**4.1 %**
JES	**2.6 %**
PDFs	**2.6 %**
Luminosity	**5.8 %**
total	**12.7 %**

Table 7.12: Systematic uncertainties affecting our measure: the main contribute is from the neural network method we use.

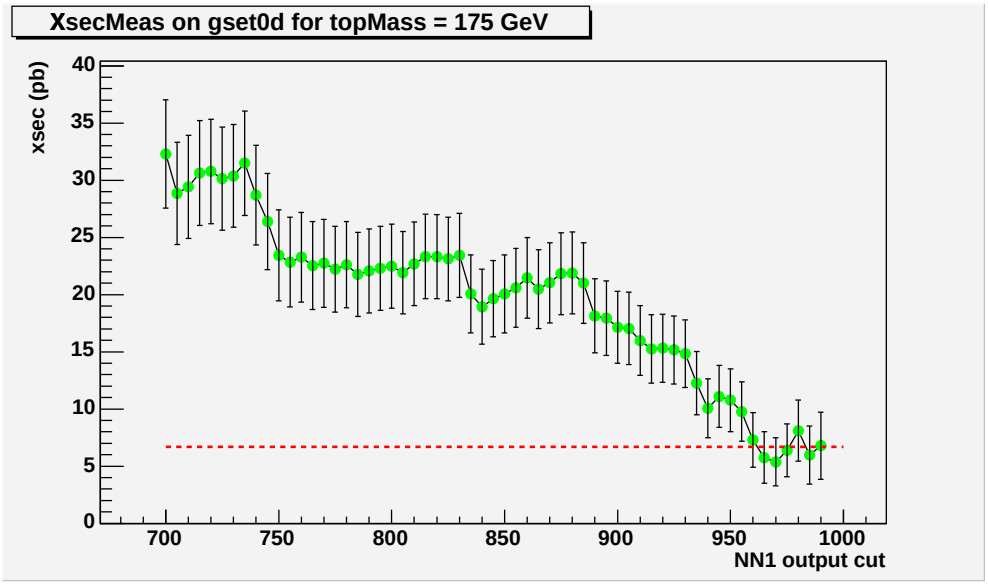

Figure 7.9: Cross section as a function of cut on *NN1* output; red line is the theoretical value (6.7 pb) for M_{top}=175 GeV. Each point is obtained applying the overall method ($NN1 \rightarrow BestCut \rightarrow NN2 \rightarrow fit$) varying the cut on $NN1$ output in order to evidence on data the same plateau we obtained on samples with known signal and background. The error bar is the statistical error calculated propagating errors on efficiencies and on n_s, the number of signal events obtained by the likelihood minimization. n_s is the main source of statistical error while errors on efficiencies are negligible. The value obtained for *NN1* output cut 960 ($BestCut$) is our final cross section estimate (7.1±2.3(stat)±0.9(syst)pb)

Bibliography

[1] H.Bachacou et al., **Combining the SecVtx B-Tagging Scale Factors for 5.3.3 Analysis**, *CDF/DOC/SEC_VTX/CDFR/7480, (2005)*

[2] M.Cacciari et al., **The $t\bar{t}$ cross-section at 1.8 and 1.96 TeV: a study of the systematics due to parton densities and scale dependence**, *CERN-TH/2003-054, hep-ph/0303085*, 2003

[3] S.Catani et al., **The top cross section in hadronic collisions**, *CERN-TH/96-21, hep-ph/9602208*, 1996

[4] Stephen Miller's instruction., `http://cdfrh0.grid.umich.edu/miller/pdf/pdf_acceptance.html`

Conclusions

In this thesis we study a selection of $t\bar{t}$ pairs decaying in a final state with five jets, two from b quarks, two from W bosons and one from a τ lepton decaying semileptonically into charged and neutral pions. We search for this signal in 311 pb^{-1} of data selected with *TOP_MULTIJET* trigger up to August 2004. $\tau + jets$ channel suffers from a large background from QCD and before any analysis $\frac{S}{N}$ is $\sim \frac{1}{32000}$. A first cleaning of the sample can be done exploiting b-tagging and τ finder algorithms: $\frac{S}{N}$ after requiring at least one b-tagged jet and a τ candidate in each event, is $\frac{1}{6000}$. Using two neural networks in cascade, the first trained mostly by global variables, the second by tau specific ones, we can reach a $\frac{S}{N} = \frac{1}{1.3}$ with 18 expected signal events. This technique, compared to a standard one with just one network, allows to obtain a $\frac{S}{N}$ three times better.

Making a binned likelihood fit on the output of the second net, we measure $t\bar{t}$ production cross section. For a top mass = 175 GeV/c^2 we obtain

$$\sigma(M_{top} = 175GeV) = 7.1 \pm 2.3(stat) \pm 0.9(syst)pb$$

The measured value is in agreement with the theoretical value of 6.7 pb.

We are working on the reduction of statistical error: the analysis will be repeated in the next future on a greater sample of data (1 fb^{-1}) with improved tau finder algorithm. New net architectures and input variables will be tested, to increase the discriminating power of the first net, which is strongly related to statistical error on the number of signal events estimated by minimization procedure.

Acknowledgements

First I would like to thank my supervisor, Ignazio Lazzizzera, for the fundamental support he gave me during the last three years and for his beautiful (and ad personam) lectures on particle physics and quantum field theory.

I'm very grateful to Ambra, dear friend, qualified colleague and ... demanding personal trainer: she always tries to transmit me her sporting spirit, indeed with small success, sorry! This last year would have been very difficult without you!

Thank you Gabriele, for your support and your meticulous corrections!

Many thanks to all CDF Padova group, especially Giovanni Busetto, Patrizia Azzi, Tommaso Dorigo and Donatella Lucchesi, from whom I tried to learn the job of experimental particle physicist.

Special thanks to Julien, for being a sincere friend and a good drinking buddy...

Thanks to Melisa, who shared with me sad and happy days in Fermilab and Udine... and she should thank me for the coffees!

Alessandra, Anna, Erica, Lisa and Gigi, Stefania and Gianluca thank you for being so close to me even if our lives are taking different directions. From you, during these years, I learned that real friendship exists, if we want to ...

Finally I want to thank Lorenza, Mauro, Francesco and my parents Gloria and Graziano: I would be nothing without their love.