Fermilab

FERMILAB-THESIS-2000-10

CDF/THESIS/JET/PUBLIC/5555
Version 1.0
12th February 2001

Calorimetric Measurements in CDF: A New Algorithm to Improve The Energy Resolution of Hadronic Jets

A Thesis Submitted
to
the University of Cassino
by

Giuseppe Latino

in Candidacy for the Degree
of
Doctor of Philosophy

Cassino, ITALY
February 2001

Abstract

Many interesting physics signatures in a collider experiment like CDF are characterized by final state quarks and gluons which fragment into *hadronic jets*. An improvement in the jet energy resolution is then very important for many Run II physics analyses and any effort in pursuing such result can be rightly considered as part of the complex CDF II upgrade program.

The jet energy resolution comes from many sources, which can be grouped into two categories: detector and physics effects. In the work exposed in this thesis, detector effects were extensively treated developing a new algorithm which allows to optimize the jet energy resolution using calorimetry as well as additional detector information. For the first time in a hadron collider the full granularity of the detector was adopted to perform corrections at "tower level" rather than at the usual "jet level", this allowing the new algorithm to be exploited regardless of the clustering algorithm chosen to reconstruct the jet.

In the present work the new algorithm has been optimized and implemented in an offline code which in principle allows it to be applied on each data sample. A preliminary study on potential algorithm improvements was also performed. The algorithm test, as performed on two different Run Ib data samples, γ-jet and di-jet, showed very promising results up to $\sim 25\%$ of improvement with respect to the standard CDF jet corrections.

UNIVERSITY of CASSINO

Ph.D Thesis

Calorimetric Measurements in CDF: A New Algorithm to Improve The Energy Resolution of Hadronic Jets

Giuseppe Latino

ADVISORS:

Prof. **Giovanni Maria Piacentino**
(University of Cassino)

Prof. **Stefano Lami**
(Rockefeller University, New York)

To Raffaella, my North Wind Light.

Contents

Introduction

The *Collider Detector at Fermilab* (**CDF**) is a particle detector to study the high mass states and large transverse momentum phenomena produced by the collision of proton and antiproton beams at the Tevatron Collider. The data collected in the period 1992-1996 (Run I) allowed the discovery of the top quark.

Run II at the Tevatron Collider will start in March 2001, providing collisions with an instantaneous luminosity up to $2 \times 10^{32}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ (about one order of magnitude greater than in Run I) and with a center-of-mass energy increased from 1.8 TeV to 2.0 TeV. The upgrade of the accelerator will confirm the Fermilab Tevatron Collider as the high energy frontier of particle physics and put the CDF experiment in the exciting position to improve Run I measurements and search for new physics for many years to come.

The goal of this new run is the accumulation of an integrated luminosity of at least 2 fb^{-1} at $\sqrt{s} = 2.0$ TeV in the first two operational years. After that, the Tevatron Collider program will be hopefully extended for other years, with a higher luminosity, so that ~ 20 fb^{-1} of data could be collected before the new Large Hadron Collider at CERN will start running.

The increased luminosity and energy required some modifications to the experimental apparatus. Based on ten years of experience with CDF and Tevatron physics, the detector, as described in Chapter 1 , has been upgraded with many new powerful features. Best attention has been also devoted to the performance of both data acquisition and offline analysis.

Many interesting physics signatures in a collider experiment are characterized by a quark or a gluon in the final state. But quarks and gluons cannot be seen as free particles because they are subject to a fragmentation process making them to experimentally appear as *hadronic jets*. From Run I data analysis, we have experienced that a limited jet energy resolution is the main source of error in many processes. Therefore, an improvement in the jet energy resolution would have a big impact on some of future Run II physics results, like for instance the top quark mass resolution (where the reconstruction of jet energies is involved) and the search for the Higgs boson

(expected to mainly decay into two jets). Any effort in pursuing it can rightly be seen as part of the CDF II upgrade program.

Dedicated studies have shown that the jet energy resolution comes from many sources, which can be grouped into two categories: detector and physics effects. In the work exposed in this thesis, detector effects are treated developing a new algorithm which allows to optimize the jet energy resolution using calorimetry as well as additional detector informations. *For the first time in a hadron collider*, the full granularity of the detector is adopted to perform corrections at "tower level" rather than at the usual "jet level". In such a way the algorithm improvement can in principle be exploited regardless of the clustering algorithm chosen to reconstruct the jet.

Chapter 2 describes the standard jet reconstruction and energy correction in CDF and the impact on Run II physics of an improved jet energy resolution. Jet theory and physics are addressed in Appendix A.

A new algorithm to improve the jet energy resolution using calorimetry, tracking and Shower Max information is motivated and described in Chapter 3. In particular the original contribution of the candidate in optimizing and implementing the algorithm into an offline analysis code is described. Results of a preliminary study on potential algorithm improvements are also shown.

The application of the new jet algorithm on CDF data, as performed by the candidate, is described in the last two chapters. Chapter 4 reports the results of testing the new algorithm code on two different Run Ib data samples, γ-jet and di-jet, showing very promising results up to $\sim 25\%$ of improvement relative to the standard CDF jet corrections. Further jet studies and cross-checks of the new algorithm are shown in the final Chapter 5.

Chapter 1

CDF at Fermilab

After a brief introduction to the Tevatron Collider, a description of the CDF detector is given in this chapter, with a particular attention to the detector components which are mainly involved in the work exposed in this thesis.

1.1 Hadron colliders

Particle accelerators can be regarded as giant microscopes, the most sophisticate tools to peer into the innermost recesses of matter. Their continuous evolution allows High Energy Physics to make astonishing advances in discovering and understanding the fundamental particles and forces of Nature, from which the universe is built. The latest of particle accelerator projects are "colliders", based on the principle of colliding particle beams, which is the most economical way to achieve the highest energies.

High energies are needed to probe small distances in the scattering of one particle from another as well as to produce a new particle whose mass can be materialized only if there is enough spare energy available in the center-of-mass frame. The discovery of several physical processes, like the production of heavy quarks and W and Z vector bosons, has been achieved thanks to very energetic interactions between leptons (typically electrons and positrons) or partons (quarks and gluons confined inside the hadrons), with the highest available energy in the center-of-mass of the collision. These collisions can be obtained by sending a particle beam on a fixed target or making two particle beams to collide head-on (*collider*). In the former the high density of the target gives a higher interaction probability, but the center-of-mass energy $\sqrt{s}$ available to create a new particle scales with the beam energy as $\sqrt{s} \propto \sqrt{E}$, while in the latter $\sqrt{s} \propto E$. So we can understand the great interest in designing and building bigger and bigger colliders even if the low

density of the beams gives a lower interaction probability.

Currently different kinds of collider are operating in the world: in lepton colliders (like LEP at CERN) the collision occurs between electron and positron beams[1]; hadron colliders use beams of protons and antiprotons (like Tevatron at Fermilab) or protons and electrons (like HERA in Hamburg). The hadron collider LHC under construction at CERN will provide proton-proton interactions beginning in 2005. The feasibility of a muon-antimuon collider is presently under study.

While in lepton colliders the interaction directly involves the beam particles, in hadron collisions the interaction involves the partons (quarks and gluons) confined inside the hadrons which typically carry only a low fraction of hadron energy[2]. On the other hand the negligible incidence of bremsstrahlung losses allows to obtain higher energies in hadron colliders with respect to lepton ones.
The most important parameters characterizing a collider performance are:

- $\sqrt{\mathbf{s}}$: the center-of-mass energy of the interacting beam particles
- $\mathcal{L}$: the *Instantaneous Luminosity*

The former is a critical parameter which sets the upper limit to the particle mass which can be created in the collision.
The latter is typically defined as (in the ideal case of relativistic beams completely overlapping) [1]:

$$\mathcal{L} = B f_0 \frac{N_p N_{\bar{p}}}{\Sigma} \tag{1.1}$$

Where B is the number of bunches circulating in the ring for each beam, $N_p(N_{\bar{p}})$ is the number of protons (antiprotons) in each bunch, f_0 is the single bunch revolution frequency and Σ is the bunch cross section. The luminosity is related to the number dn/dt of events of a given process (characterized by a cross section σ) which are produced in the time unit:

$$dn/dt = \mathcal{L}(t)\sigma \quad \Longrightarrow \quad n(T) = \sigma \int_0^T \mathcal{L} dt \tag{1.2}$$

The quantity $\mathcal{L}_{int} = \int \mathcal{L} dt$ (called *Integrated Luminosity*) is a very important parameter in a collider experiment giving the *total* number of events $n(T)$

[1] The beam particles are actually grouped into bunches so that a collision can be obtained when two opposite bunches overlap.

[2] At high energies, like at Tevatron, this parton confinement inside the hadrons tends to disappear (*Asymptotic Freedom*) so that the proton and antiproton beams are actually "parton beams" with a wide momentum band.

(for the given process) collected during the time T of data taking.
From the above: $[\mathcal{L}] = [cm^{-2}][sec^{-1}]$ while typically $[\mathcal{L}_{int}] = [b^{-1}]$ where the barn(b) is the standard unit for nuclear cross section (1b = 10^{-24} cm^2).

1.2 The Tevatron

The *Tevatron Collider* at Fermi National Accelerator Laboratory (Fermilab), located 50 Km west of Chicago, is a proton-antiproton collider which allows a center-of-mass interaction energy $\sqrt{s}$ = 2 TeV [3]. This energy will be the highest available in the world until around 2005 when LHC (at CERN) will be operative at $\sqrt{s}$ = 14 TeV.

The Tevatron is a double acceleration ring with a radius of 1 Km where proton and antiproton bunches circulate in opposite directions bent by superconducing magnets. The bunches can collide in two experimental areas, B0 and D0, where the experiments CDF II and D0 are respectively placed. This accelerator is also designed for fixed target experiments: 0.8 TeV proton bunches are extracted and sent to other experimental areas [2].

During the last four years this accelerator has been upgraded in order to increase both $\sqrt{s}$ and $\mathcal{L}$. In particular, the construction of the *Main Injector* and of the *Recycler Ring* was finalized to increase $\mathcal{L}$ (mainly limited by the particle density in the antiproton bunches) by increasing the number of bunches circulating in each beam. The new operational period (Run II) is scheduled to begin on March 2001.

The Tevatron is actually the last stage of the accelerator system which is schematically shown in fig. 1.1. The different parts of this complex are used in successive steps to produce and store the particle bunches sent to the Tevatron. The beam production can be schematically described as follows.

Proton Beam: H^- ions are produced by ionization of gaseous hydrogen and then accelerated up to an energy of 750 KeV by a Cockroft-Walton electrostatic accelerator. Then they are injected in a 150 m long linear accelerator (*Linac*) which increases their energy to 400 MeV. After being focused, the H^- ions are made to collide on a thin carbon target and in this interaction they lost the two electrons becoming protons. The protons are then injected in the *Booster*, a 75 m radius synchrotron where they reach the energy of 8 GeV and are grouped into bunches each containing up to $\sim 5 \cdot 10^{12}$ particles. These bunches are then sent to the *Main Injector*, a synchrotron accelerating them to 150 GeV. Finally the protons are injected into the Tevatron, a 1 Km radius synchrotron using superconducting magnets producing a field of 5.7

[3] In the period 1992-1996 (Run I) this collider was operated at $\sqrt{s}$ = 1.8 TeV allowing the discovery of the **top** quark.

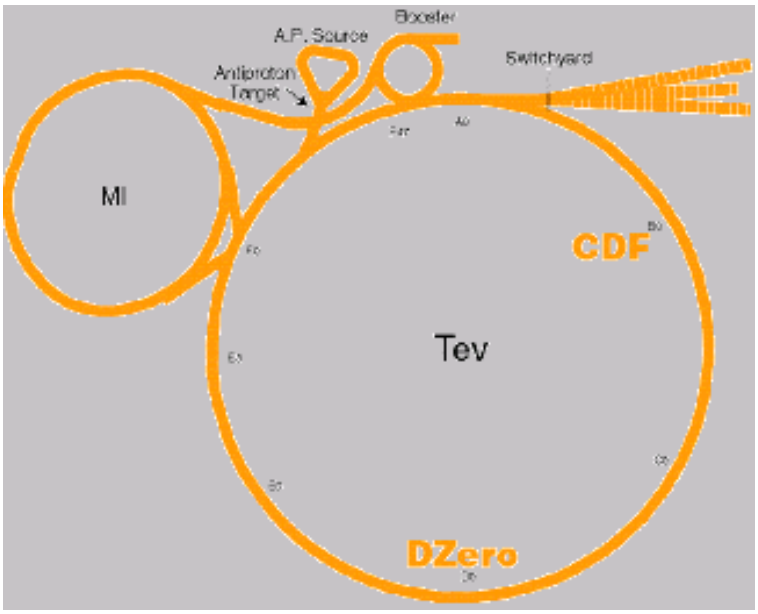

Figure 1.1: *Fermilab accelerator complex.*

T, where they reach the final energy of 1 TeV.
The Main Injector is the most important achievement in the Fermilab upgrade program for Run II [2]. It has been designed to solve most of the problems met with Run I when the *Main Ring*, a conventional synchrotron placed over the Tevatron in the same ring, served as injector. The Main Ring could not accept all protons from the Booster and also it created inefficiencies and backgrounds during the data acquisition from the Tevatron being the two accelerators in the same tunnel. Conversely the Main Injector allows an optimal interconnection between the Booster and the Tevatron: we expect an increase of 50% in protons for each production pulse and of 60% in proton pulse frequency.

Antiproton Beam: a proton bunch of 120 GeV, extracted from the Main Injector, is sent to a rotating nichel target after being focused. In this collision several nuclear products are generated among which antiprotons with a wide angular and momentum spread. A suitable magnetic field and a litium magnetic lens allow to select and focus the antiprotons in bunches having 8 Gev in average energy. The antiprotons are then sent to the *Debuncher*, a storage ring which narrows theirs momentum distribution using the *Stocastic Cooling* technology. At the same time the particle spatial spread is increased so to reduce the bunches to a continuous beam which is sent to the *Accumulator*, another storage ring placed inside the Debuncher. Here the

antiprotons are continuously stored during the various Debuncher cycles and subject to further stocastic cooling. This accumulation process continues until a current sufficient to give 36 (121)[4] bunches with acceptably high density is stored. The antiprotons are then injected into the Main Injector which accelerates them to 150 GeV and finally in the Tevatron where, circulating in opposite direction with respect to the proton bunches previously injected by the Main Injector, they reach the energy of 1 TeV.
After each accumulation-acceleration-interaction cycle when the beams are degraded, the antiprotons still circulating inside the Tevatron are decelerated first in the Tevatron and then in the Main Injector up to 8 GeV and then sent to the *Recycler*. This is a Stocastic Cooling Ring placed in the same tunnel of the Main Injector where the antiprotons coming both from the Accumulator and from the Tevatron (through the Main Injector) are stored till a new boosting injection in the Main Injector and then in the Tevatron. In such a way the Recycler, which is able to store up to $5 \cdot 10^{12}$ antiprotons, allows an "antiprotons recycling" resulting in an increased luminosity.

In Run IIa (Run IIb), 36 (140) proton bunches and 36 (121) antiproton bunches circulate in opposite directions inside the Tevatron, colliding every 396 (132) ns in the center of the CDF II and D0 detectors. At each bunch crossing the interaction region is characterized by a gaussian spatial distribution with $\sigma_z \sim 16$ cm in the beam direction and $\sigma_t \sim 16$ μm in the transversal one.
The increase in the number of beam bunches with respect to Run I (6 bunches per beam) allows to increase the luminosity reducing at the same time the bunch particle density which results in a lower average interaction number in each bunch crossing and so in "cleaner" physical events. The Fermilab Accelerator Complex upgrade is expected to allow the Tevatron to deliver an instantaneous luminosity of $2 \cdot 10^{32} cm^{-2} sec^{-1}$ which will provide $2 fb^{-1}$ of integrated luminosity at $\sqrt{s} = 2$ TeV after 2 years of data acquisition [2].
On the other hand the reduced time between two bunch crossing ($\sim$3500 ns in Run I) and the increase in the collision energy have influenced most of the CDF Detector upgrade program: from the choice of the active part of new subsystems to the design of the new readout electronics.
Table 1.1 reports the most important beam parameters for the different Tevatron operating configurations [2].

[4]These two numbers are related to the two different operative configurations planned for the Tevatron: 36 antiproton bunches for a beams interaction every 396 ns (Run IIa), 121 for a beams interaction every 132 ns (Run IIb).

Run	Run Ib	Run IIa	Run IIb
# of bunches ($p \times \bar{p}$)	6×6	36×36	140×121
p per bunch	$2.3 \cdot 10^{11}$	$2.7 \cdot 10^{11}$	$2.3 \cdot 10^{11}$
$\bar{p}$ per bunch	$6 \cdot 10^{10}$	$7 \cdot 10^{10}$	$7 \cdot 10^{10}$
p rms bunch length (cm)	63	37	37
$\bar{p}$ rms bunch length (cm)	59	37	37
Beam energy (GeV)	900	1000	1000
Bunch spacing (ns)	≈3500	396	132
Luminosity ($cm^{-2}sec^{-1}$)	$1.6 \cdot 10^{31}$	$> 5 \cdot 10^{31}$	$2 \cdot 10^{32}$
Interactions/collision average #	2.5	2.3	1.3

Table 1.1: *Comparison among the most important beam parameters for the different Tevatron operating configurations. Run Ib was the last period of data acquisition of the previous run (1994-96). Run IIa is scheduled to start March 2001 while Run IIb after the first two operative years of the new Tevatron.*

1.3 The CDF II Detector

The *Collider Detector at Fermilab* (**CDF**) [3, 4, 5] is a collection of detection devices designed to study a wide range of physical processes produced in proton-antiproton collisions and characterized by final state particles with high transverse momentum. This equipment is finalized to detect photons, electrons, muons, hadronic jets and (indirectly) neutrinos, measuring their position, energy and momentum. In order to achieve these goals and considering the energy equivalence of the two beams, it is characterized by an almost complete coverage in solid angle and by an azimuthal and "back-forward" symmetry with respect to the nominal interaction point.

CDF is a general purpose detector. With the accelerator providing up to about 7.6 million collisions per second, it can pick up online just the most interesting events according to several different physics programs. The first version of the CDF detector collected data starting 1987. During Run I (1992-96) the CDF collaboration achieved very important physics results: the discovery of the ***top*** quark [6], precision measurement of the W mass [7], the preliminary measurement of $sin2\beta$ [8] and the discovery of the B_c meson [9].

In the last four years the CDF detector underwent a major upgrade with the addition of almost a million new detector elements that will allow CDF to make even more precise measurements and search for new phenomena for many years to come. The **CDF II** detector will be able to cope with the new Tevatron operative conditions (increase in collision energy and frequency) and to overcome some detector limitation seen during Run I (track

reconstruction, muon detection, forward calorimetry).

CDF adopts two coordinate systems according to the following convention: *cartesian coordinates* (x, y, z) with origin in the nominal interaction point, z axis along the beam direction according the proton motion, x axis going outside the Tevatron ring and y axis up perpendicular to the ring plane; *polar coordinates* (r, ϕ, θ) oriented according to the usual convention with respect to the cartesian ones. Like other collider experiments the following variables are very often used: the *pseudorapidity* η which is related to the polar angle θ by

$$\eta = -log[tan(\frac{\theta}{2})] \tag{1.3}$$

and the projection on the transverse plane of the particle energy and momentum, $E_T = E \cdot sin\theta$ and $P_T = P \cdot sin\theta$.

The whole detector system is composed of three regions: the *central* one covering the η range $|\eta| \leq 1$ and the two *plug* ones which symmetrically cover the "backward" and "forward" regions with $1 \leq |\eta| \leq 3.6$. Starting from the nominal interaction point the central detector can be divided into three main substructures:

- *Tracking System*
- *Calorimetric System*
- *Muon System*

The two plug detectors are essentially calorimetric and muon systems being an extension up to $|\eta| \leq 3.6$ of the corresponding central systems.

Finally the *Trigger System* allow to select the rare physical events of interest to be written on tape with a drastically reduced rate respect to the bunch crossing one.

As the work exposed in this thesis is part of the CDF upgrade program, the CDF II upgraded system will be described in this section. Nevertheless, as both simulation and data sample were used relatively to Run I, a brief description of CDF I will be also given in next section.

Given the complexity of the CDF apparatus, only a general description is provided, with a particular emphasis to the detector components mostly involved in the work of the present thesis. The main detector structures are sketched in fig. 1.2.

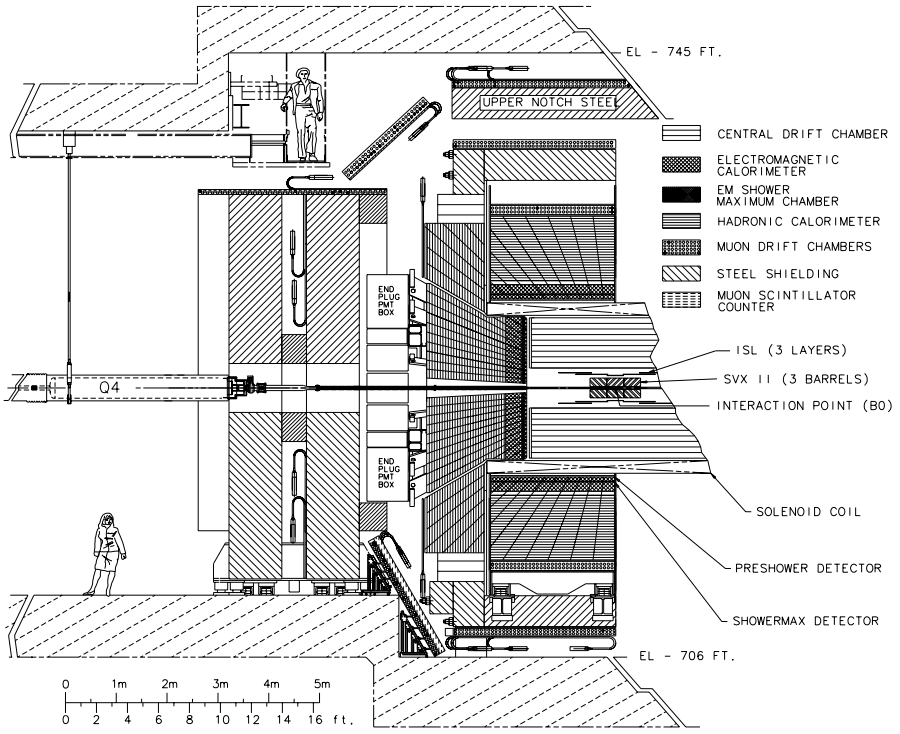

Figure 1.2: *Side view of the CDF II detector.*

1.3.1 Tracking System

This system is placed inside a 4.8 m long and 3 m wide superconducting solenoid, which provides an uniform magnetic field parallel to the beam axis of about 1.4 Tesla. It allows a 3-D track reconstruction for charged particles traversing the field region and so, by track deflection, to measure their charge and transverse momentum. It consists of a microstrip silicon section (L00+SVXII+ISL), which gives a precise tracking around the beam, surrounded by a cylindrical drift chamber (COT).

Layer00

The Layer00 (**L00**) is a vertex silicon detector placed just outside the berillium beam pipe at 1.6 cm from the beam axis [10]. It is made by a layer of single sided microstrip silicon detectors parallel to the beam giving track information in the r-ϕ plane. Fig. 1.3 shows one of the two parts constituting this device. A new technology allows to polarize these sensors to high voltage ($\sim$500 V) which gives a good signal-to-noise ratio (S/N) even after high radiation doses ($\sim$5 MRad). The L00 sensors, covering longitudinally

Figure 1.3: *Drawing of one half of L00. Four sensor groups (blue) starting at $z = 0$ and the readout hybrids (magenta) starting at $z \approx 45$ cm are shown. The cooling system is shown in red.*

the beam pipe for about 80 cm, have a strip separation of 25 μm and a 50 μm readout pitch. This allows a 6 μm single point resolution in the transverse plane. To reduce the radiation damage effects the sensors need to be operated at about 0 ^{o}C, which requires a cooling system embedded in their carbon fiber support. The readout electronics is placed in a separate zone to avoid both silicon heating and radiation damage problems and is connected to the about 16,000 channels by special cables. Detector signals are processed by the SVX3D chips which are the same, like the combined data acquisition system, used for all the silicon tracking system.

SVXII

Outside the L00, between 2.4 and 10.6 cm from the beam axis, five double sided microstrip silicon layers are placed which constitute the Silicon VerteX Detector II (**SVXII**) [5]. Three of them allow to get the track position on the r-ϕ plane from the readout of one side (microstrips parallel to the beam axis) while the z coordinate is read from the other side (microstrips perpendicular to the beam axis). The other two layers give track information on r-ϕ from one side and r-ϕ' from the other, having the microstrips tilted

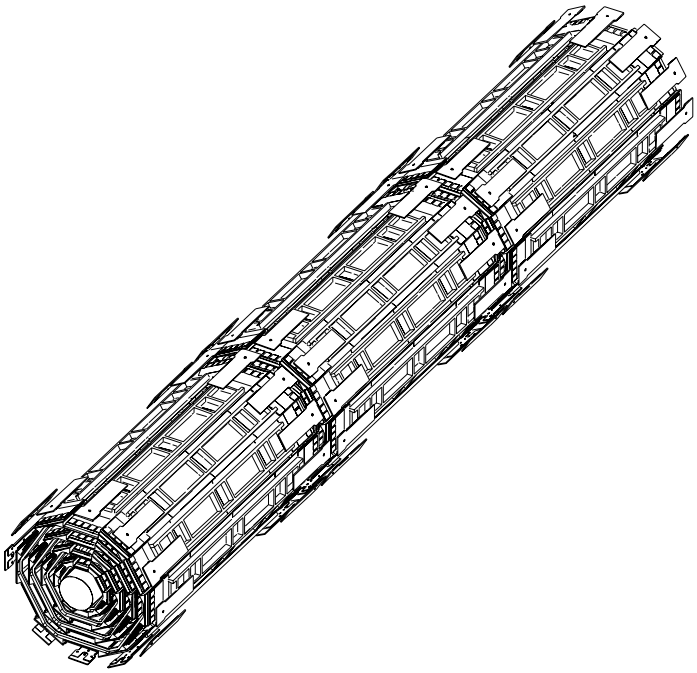

Figure 1.4: *Final SVXII assembly with the silicon detectors mounted on the three barrels.*

of 1.2^o in a stereo configuration respect to the z axis[5]. In such a way it is possible to have a 3-D track reconstruction with an approximatively uniform efficiency. The strip pitch is 60 μm on the r-ϕ readout side while is varying from 60 to 140 μm on the r-ϕ' and z one. This detector has been designed to give (with an improved coverage of the pseudorapidity region) a 3-D vertex reconstruction and to provide track information up to $|\eta| \leq 2$ with an impact parameter resolution $\sigma_\phi < 30$ μm and $\sigma_z < 60$ μm for central high momentum tracks. The sensors are mounted on three mechanical structures (*barrels*) covering a total length of 96 cm along the z coordinate and are segmented into 12 *wedges* in ϕ. The barrels also support the water cooling channels for the readout electronics. Fig. 1.4 gives a sketch of the complete barrels assembly. The almost 406,000 channels of SVXII are read by the readout chips SVX3D placed on its external sides along z. Both SVX3D chips and SVXII detectors are designed and tested to be "radiation-hard" considering the strong radiation field in which they have to operate. We expect that after about 3 fb^{-1} of integrated luminosity the two innermost layers (Layer0 and Layer1) will have a significant worsening, but the additional information from Layer00 will allow good inner tracking up to ~ 5 fb^{-1}.

[5]This sensors with one side 1.2^o tilted strips are obtained from new 6 inch silicon wafers allowing to obtain two sensors from an only wafer so lowering the production cost.

Intermediate Silicon Layers

The Intermediate Silicon Layers Detector (**ISL**) is a charged particle detector placed in the radial intermediate region between SVXII and the central tracking system (COT) [5]. Fig. 1.5 shows the position of the ISL silicon layers. They are in the radial range $20 < r < 30$ cm extending to $|z| = 65$ cm for the inner layer and $|z| = 87.5$ cm for the outer one and covering an overall pseudorapidity zone up to $|\eta| \leq 2$. For $|\eta| \leq 1$, where the COT tracking information is more complete, ISL is made by a single detector layer at $r = 22$ cm; for $1 \leq |\eta| \leq 2$ two layers are placed at $r = 22$ and 28 cm.

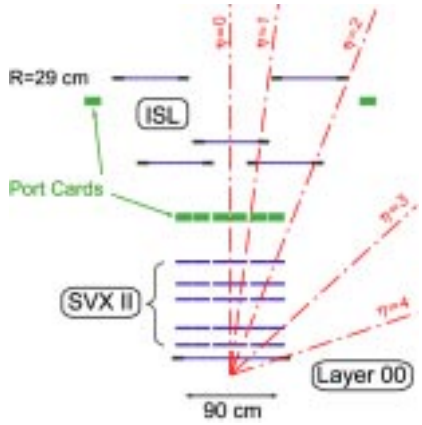

Figure 1.5: *Positions of all the silicon layers constituting the inner silicon tracker system. The pseudorapidity coverage regions are also shown. The z scale has been compressed.*

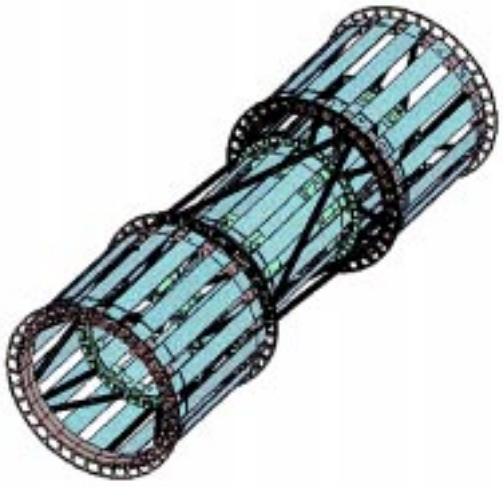

Figure 1.6: *Overall ISL assembly. The readout modules are installed on a carbon fiber structure also containing the SVXII detector.*

The total ISL extension in z is about 2 m for an overall silicon active surface of about 3.5 m^2. The azimuthal segmentation into "wedges" is such that a complete ϕ coverage is obtained without inefficiencies. The overall ISL assembly is shown in fig. 1.6.

The microstrip silicon crystals used in ISL (58mm×76mm) are the same as SVXII and are double-sided with axial strips on one side and small angle (1.2^o) stereo strips on the other. The microbounding of 3 crystals connected to their readout electronics (same SVX3D chips as SVXII) gives a *ladder* which is the basic electric readout unit . ISL has 296 of such ladders for a total of about 300,000 readout channels connected to the data acquisition.

In principle a charged particle, before entering the COT, leaves 13 measurement points in the inner silicon system (L00+SVXII+ISL). So this three detector can be considered an independent tracking system for $|\eta| \leq 2$ giving a 3-D track reconstruction. As consequence tracking, lepton identification and b-tagging capabilities can be extended over this full region. Moreover an improvement in the overall track fit is attended in the central region where the ISL will provide a further measurement point between the SVXII and the COT. From preliminary simulations [5]: the single hit resolution is expected to be $\sim 20\mu$m in the transverse plane; the impact parameter resolution is expected to be $\delta d \sim 15\mu$m while for the transverse momentum resolution $\frac{\delta P_T}{P_T} \sim 0.4\% \cdot P_T$. A 8% fake track incidence is finally expected.

Central Outer Tracker

For $|\eta| \leq 1$, tracking at large radii is performed by the Central Outer Tacker (**COT**), a large cylindrical open cell drift chamber [5], which substituted the old drift chamber (CTC) used in Run I. This device is 3 m long occupying the radial range $44 < r < 132$ cm (see fig. 1.7). The COT is designed to operate with a maximum drift time of 100 ns by reducing the maximum drift distance (obtained with a reduced cell dimension) and by using a gas mixture ($50 : 35 : 15\ Ar - Et - CF_4$) with a faster drift velocity ($\sim 100\mu m/ns$) [6]. Considering a bunch crossing every 132 ns as operative limit for Run II, this avoids dead times and also allows to use the COT information for the level 1 trigger.

The COT is made by 96 wire coaxial layers. They are grouped into 8 cell *superlayers*. Four, with wires parallel to the beam axis, giving the $r - \phi$ coordinate and four giving the z coordinate with a stereo ($\pm 3^o$) wire configuration. A 3-D track reconstruction is so possible. Each basic drift cell contains 12 sense wires alternating with shaper wires. The drift electric field

[6] The 706 ns CTC maximum drift time was the fundamental reason for it to be substituted with an improved device (COT).

is 2.5 kV/cm which, for the CDF II central magnetic field, gives a Lorentz deflection angle of 35^o [7].

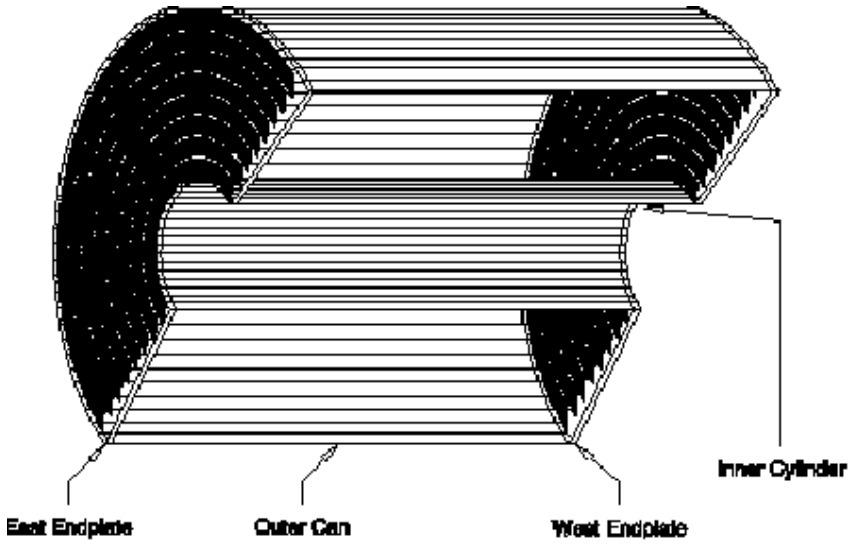

Figure 1.7: *Cross view of one half of the COT.*

The cell geometry is sketched in fig 1.8. Its electrostatic field configuration is given by a cathode (*"field panel"*), which is made by gold on a 0.25 mm thick Mylar sheet, and by wires attached on Mylar strips (*"shaper panel"*) which close both electrically and mechanically the ends of each cell. In such a way the loss effect due to a broken wire is minimized as contained within one cell. Between the field panels, sense wires alternate in a plane with potential wires. Both are 1.6 mm gold-plated tungsten wires. Signals from the about 30,000 COT readout channels are processed (discriminated, shaped and amplified) and sent to a TDC. Finally they are sent in parallel to the trigger and data acquisition systems.

The COT measurement precision is strongly related to the geometry regularity and to the electrostastic field uniformity. The single hit resolution is expected to be $\sim 180\mu$m in the transverse plane while the transverse momentum resolution is expected $\frac{\delta P_T}{P_T} \sim 0.3\% \cdot P_T$. These resolutions are quite similar to those obtained with the CTC in Run I.

[7] The Lorentz angle is the deflection angle of the drift particles with respect to the electric field direction which is due to the solenoid magnetic field. To get an uniform readout the cells are tilted by this angle respect to the radial direction.

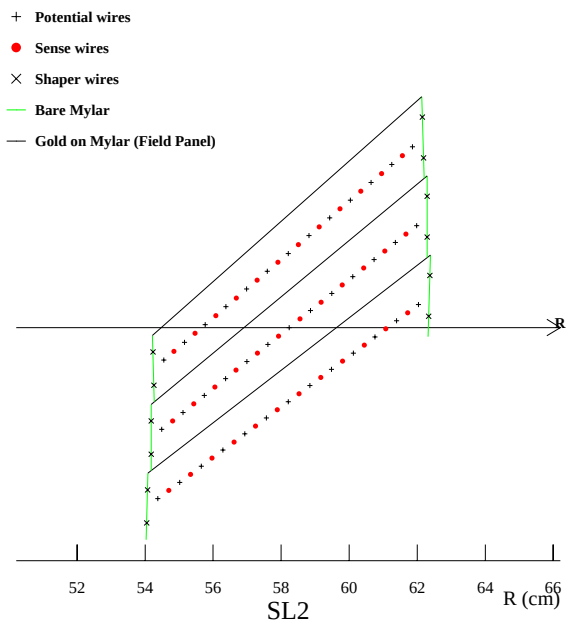

Figure 1.8: *$r-\phi$ view of a single COT cell.*

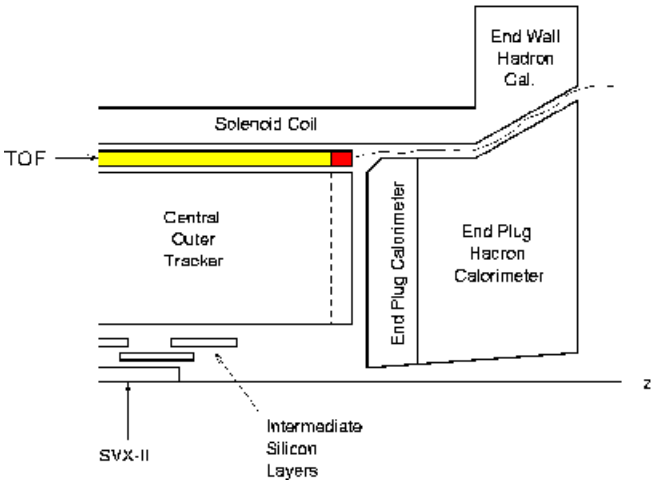

Figure 1.9: *TOF scintillator bars in CDF II.*

Time of Flight Detector

The Time of Flight (**TOF**) scintillating bars [10] are placed between the COT and the superconducting solenoid cryostat arranged in a cylinder structure. Each of the 216 bars is 2.8 m long and is placed at a radial distance of 138 cm from the beam axis parallel to it (see fig. 1.9). Two PMTs read the light from both the bar ends. The TOF measures the times between the bunch crossing

and the signals produced into its scintillators by the particles originated in the collision. The particle speed is evaluated by this time and combining this information with the momentum from the tracking system its mass is measured.

Fast precision electronics is used for the time of flight measurements which is expected to give a time resolution of $\sim$ 100 ps on single particle. A 1σ separation between π and K with $P < 2.2$ GeV/c is expected.

1.3.2 Calorimetric System

Outside the solenoid coil, sampling calorimeters cover the pseudorapidity region $|\eta| \leq 3.6$ and the entire 2π azimuthal angle [5]. Different size and thickness plastic scintillator and absorber layers are alternatively stacked forming the electromagnetic calorimeters (allowing the energy measurement for photons and electrons) and the hadronic calorimeters (measuring hadron particle energies). The primary particle produces a shower of secondary particles inside the absorber. The shower particles deposit a fraction of their energy in the sampling material producing a light signal read by photomultipliers (PMTs) through wavelength shifting (WLS) light guides or optical fibers. The original particle energy is then obtained with a calibration based on test beam data.

The calorimetric system is composed of two sub-systems: the *central* calorimeter and the *plug* calorimeter. Both of them are segmented in the η and ϕ coordinates in order to have a projective tower geometry pointing back to the nominal interaction point (see fig. 1.2 and fig. 1.10).

In each tower the electromagnetic compartment is backed by the hadronic one, both readout by different PMTs. A comparison of electromagnetic and hadronic energy can be made on a tower-by-tower basis allowing the distinction between photons/electrons and hadrons. The tower segmentation also gives a measure the angle at which the particle emerged from the interaction point.

The central calorimeter will be described in next section being the same as in Run I. Here we describe the new plug calorimeter replacing the old plug and forward gas calorimeters no more compatible with the crossing rates for Run II.

Plug Calorimeter

The calorimetric coverage in the backward/forward region ($1 \leq |\eta| \leq 3.6$) is provided by two identical **Plug** Calorimeters [5] (see fig. 1.10). Both the electromagnetic (EM) and hadronic (HAD) sections of this new calorimeter

are sampling devices having scintillator tiles as active elements read out by WLS optical fibers embedded in the scintillator. The WLS fibers are spliced to clear fibers carrying the light out to PMTs placed on the back plane of each endplug. Active and absorber elements are arranged in a tower geometry, common to both EM and HAD sections, whose $\Delta\eta\times\Delta\phi$ segmentation varies with θ from 0.1×7.5^o to 0.6×15^o.

The EM calorimeter is composed of 23 layers for a total thickness of about 21 X_0 (corresponding to $\sim$ 1 λ_0)[8] at normal incidence. Each layer is composed of 4.5 mm lead and 4 mm scintillator. The scintillator tiles of the first layer are made out of 10 mm thick scintillator and are read by multi-anode photomultipliers (MAPMTs). They will be used as a *Preshower Detector*. A *Shower Max Detector* is also located at the deep of maximum shower development (approximately 6 X_0). This position detector is made of scintillator strips read out by WLS fibers carrying the light to MAPMTs. This device will improve the $e^{\pm}$-γ/π^0 separation. From test beam data [11], the energy resolution of the EM section results $\sim 15.5\%/\sqrt{E}\oplus 1\%$ [9].

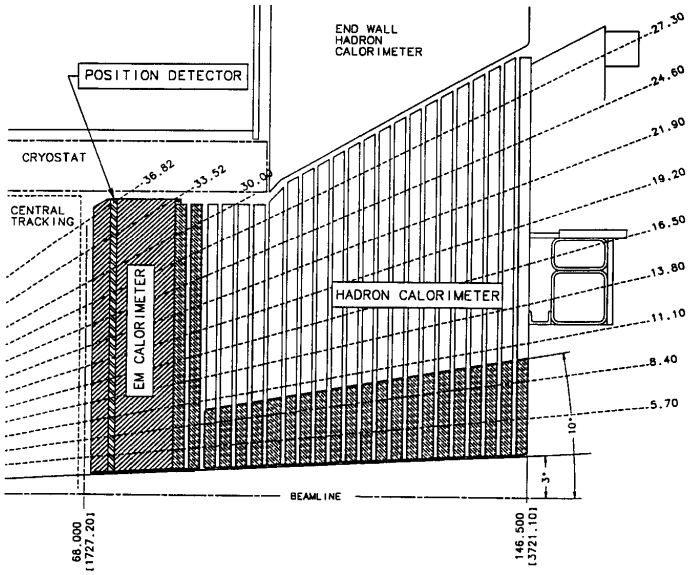

Figure 1.10: *r-z view of the upper part of the new CDF II plug calorimeter.*

[8]The *radiation length* X_0 and the *nuclear interaction length* λ_0 of a given material are respectively defined as the mean distance over which a high-energy electron loses about 1/e of its energy by bremsstrahlung and the mean free path for a hadron (usually a π^+) to undergo a nuclear inelastic interaction. They are the appropriate scale length to describe high-energy electromagnetic and hadronic showers, respectively.

[9]Here and in the following, the symbol $\oplus$ indicates a sum in quadrature.

The HAD calorimeter is a 23 layer sampling device with an unit layer made of 5 mm iron and 6 mm scintillator. The overall thickness for a normal incidence is $\sim 7\ \lambda_0$ and the energy resolution for pions is estimated to be $\sim 78\%/\sqrt{E} \oplus 5\%$ from test beam data [11].

1.3.3 Muon System

The muon detection will be provided, in the pseudorapidity range $|\eta| \leq 2.0$, by four systems of scintillators and proportional chambers integrating and extending the old muon system [5]. The muon candidate z and ϕ coordinates are provided by the chambers while the scintillator detectors are used for triggering and spurious signals rejection. The calorimeter steel, the magnet return yoke, additional steel walls and the RunI forward muon toroids act as hadron absorber for this system, drastically reducing the "*punch-through*" incidence.

In the central region ($|\eta| \leq 0.6$) the **CMU** chambers are placed at $r \sim 3.5$ m just outside the calorimeter while at $r \sim 5$ m, behind a 60 cm iron wall, there are the **CMP** chambers with an external scintillator layer. At the same radial distance the $0.6 \leq |\eta| \leq 1.0$ range is covered by the **CMX** chambers, provided with a double scintillator coverage. Finally the **IMU** chambers, also integrated with scintillators, give muon detection in the range $1.0 \leq |\eta| \leq 2.0$.

1.3.4 Trigger System

In Run II operative conditions the Tevatron will provide a bunch crossing rate up to ~ 7.6 MHz while physical events can be written to magnetic tape with a $\lesssim 50$ Hz rate [5]. So a tiered "deadtimeless" trigger system (implemented in a complex hardware-software architecture) is used to properly select the rare physical processes of interest. Each event is considered sequentially at three levels of approximation with each level providing a rate reduction sufficient for the next level to have minimal deadtime.

Like the data acquisition system (DAQ), the trigger is fully pipelined with the Level-1 and 2 using a custom hardware on a limited subset of the data while the Level-3 using a processor farm running on the full event readout. This system is very flexible allowing to accommodate over 100 separate trigger selections. With a 40 kHz Level-1 accept rate and a 300 Hz rate out of Level-2 a deatime $< 10\%$ is expected at full luminosity.

A detailed description of the main trigger selection criteria is not possible at the moment as they are not yet completely defined. Here we briefly describe the main improvements with respect to the corresponding CDF I system. The track finding (previously available only at Level-2) is added to

Level-1 allowing an improved electron an muon identification. At Level-2 the Silicon Vertex Tracker (SVT) provides the trigger on tracks with large impact parameters, very important to select events with b quarks in the final state which characterize a large number of interesting processes. At the same time an improved Level-2 track momentum resolution and electron/photon and muon identification will be achieved.

1.4 The CDF I Detector

As shown in the next chapters, the work exposed in this thesis was performed using both simulation and data samples relative to Run I. A brief description of CDF I is then needed regarding the detector components directly involved in this study: tracking, central calorimetry and trigger system.

1.4.1 Tracking

The CDF I tracking system consisted of three separate tracking devices placed inside the solenoid magnetic field [6].

Outside the 1.9 cm radius beryllium beampipe was a four layer *Silicon Microstrip Vertex Detector* (**SVX**) consisting of two identical cylindrical modules for a total length of 51 cm and a complete azimuthal coverage up to $|\eta| \leq 1.9$. Each module was made of 12 *wedges* consisting of four silicon detector layers placed at a radial distance ranging from 3 cm to 7.9 cm. This device provided track informations on the transverse plane with a single-hit resolution, measured on data, $\sigma \sim 13$ μm and an impact parameter resolution measured to be δd ~ 17 μm. Its performance resulted to be critical for the quark top discovery.

Surrounding the SVX, a *Vertex Drift Chamber* (**VTX**) provided tracking information up to $r = 22$ cm and $|\eta| \leq 3.25$ also measuring the $p\bar{p}$ interaction vertex along the z axis with a resolution of 1 mm. This device was composed of 8 layers of drift chambers arranged in an octagonal structure.

Both the SVX and VTX were placed inside the *Central Tracking Chamber* (**CTC**), a 3.2 m long drift chamber with an outer radius of 132 cm providing very precise charged particle transverse momentum measurements up to $|\eta| \leq 1.1$. The CTC contained 84 concentric cylindrical layers of anodic sense wires grouped into 9 superlayers. Five of them had 12 (axial) wires parallel to the beam axis and provided tracking in the r-ϕ plane with a single-hit resolution ~ 200 μm. The other four superlayers had 6 (stereo) wires tilted at $\pm 3^o$ with respect to the beam direction which, combined with the axial wires, provided tracking in the r-z plane with a resolution on $z \sim 4$ mm. A 3-D

track reconstruction was so obtained with this device. In each superlayer the wires were grouped into cells tilted of 45^o respect to the radial direction to account for the Lorentz angle. The resolution on isolated track transverse momentum was $\frac{\sigma_{P_T}}{P_T} \sim 0.002 \times P_T$. If track reconstruction was made also using the SVX information, this resolution was $\frac{\sigma_{P_T}}{P_T} \sim (0.0009 \times P_T) \oplus (0.0066)$ [6].

1.4.2 Calorimetry

Surrounding the solenoid were sampling calorimeters which covered 2π in azimuth and the pseudorapidity range $|\eta| \leq 4.2$ [4]. The whole system was composed of three separate η regions: central, end plug and forward (see fig. 1.11). Each region had an electromagnetic calorimeter (respectively CEM, PEM and FEM) with lead as absorber and behind it a hadronic calorimeter (respectively CHA/WHA, PHA and FHA) whose absorber was iron. The active part was plastic scintillator in the central region while in the plug and forward ones was a mixture of 50% argon and 50% ethane (with a small percentage of alcohol to prevent glow discharge). The calorimeters were segmented in η and ϕ to form a projective tower geometry. The $\Delta\eta \times \Delta\phi$ segmentation was 0.1×15^o for the central towers and 0.1×5^o for the others. The CDF I calorimeters coverage, resolution and thickness are summarized in table 1.2.

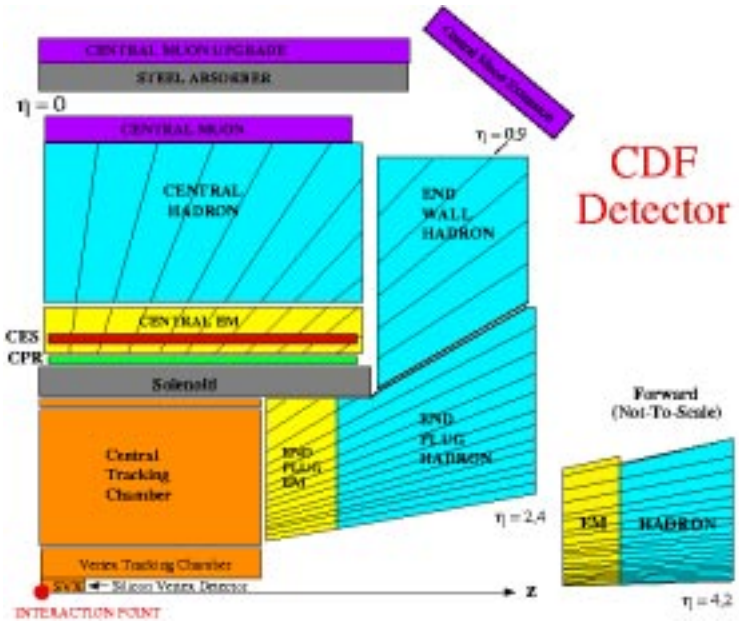

Figure 1.11: *The CDF I experimental setup. Central calorimetry is the same as for CDF II.*

System	η range	Resolution	Thickness
CEM	$\mid \eta \mid < 1.1$	$13,7\%/\sqrt{E_T} \oplus 2\%$	18 X_0
PEM	$1.1 < \mid \eta \mid < 2.4$	$22\%/\sqrt{E_T} \oplus 2\%$	18-21 X_0
FEM	$2.2 < \mid \eta \mid < 4.2$	$26\%/\sqrt{E_T} \oplus 2\%$	25 X_0
CHA	$\mid \eta \mid < 0.9$	$50\%/\sqrt{E_T} \oplus 3\%$	4,5 λ_0
WHA	$0.7 < \mid \eta \mid < 1.3$	$75\%/\sqrt{E_T} \oplus 4\%$	4,5 λ_0
PHA	$1.3 < \mid \eta \mid < 2.4$	$106\%/\sqrt{E_T} \oplus 6\%$	5,7 λ_0
FHA	$2.4 < \mid \eta \mid < 4.2$	$137\%/\sqrt{E_T} \oplus 3\%$	7,7 λ_0

Table 1.2: *Angular coverage, resolution and thickness of the CDF I calorimeters. E_T is in GeV.*

Central Calorimeter

As already mentioned, this part of the CDF detector was not subject to any substantial change in the CDF II upgrade program.

The *Central Calorimeter* is azimuthally arranged in 48 physically separated 15^o wide modules (*wedges*) each segmented in η into ten towers [12, 13]. Neighboring towers belonging to different wedges are physically separated by not instrumented cracks (*"ϕ-cracks"*), whereas tower separation inside the same wedge is obtained collecting the light from different cells into different PMTs. The boundary between the two halves of the central calorimeter and between the wedges and endwall modules constitutes another not instrumented region (*"η-cracks"*).

The *Central Electromagnetic Calorimeter* (**CEM**) is overlapped by a hadronic section split into two parts, *Central Hadronic* (**CHA**) and *Wall Hadronic* (**WHA**). In each tower both electromagnetic and hadronic sections are read by two PMTs from the opposite ϕ faces.

Proportional chambers are located between the solenoid and the CEM forming the *Central Preradiator Detector* (**CPR**) which provides r-ϕ information on electromagnetic showers initiating in the material of the solenoid coil.

Located six radiation lengths deep in the CEM calorimeters (approximately at shower maximum) is the *Central Electromagnetic Strip Detector* (**CES**) [14]. These gas multiwire proportional chambers are divided in each wedge into two halves in z providing pulse height readout in two orthogonal directions: the anode wire channels give the shower pulse height distribution as a function of the azimuthal ϕ coordinate with the only readout for five towers, while the cathode strips provide the shower profile along the z axis. The CES measures the electromagnetic shower centroid position with

a resolution of about 2 cm along both directions.

In front of each ϕ-crack dead region is a 12 X_0 tungsten bar with a gas proportional chamber behind it called *Central Crack Chamber* (**CCR**) [15]. Each of the 48 crack modules has 10 read-out pads in η corresponding to the 10 CEM towers. This device allows to detect and measure the energy of electrons and photons falling in such not instrumented zone.

1.4.3 Trigger

The CDF I *Trigger System* was a three level system [6].

Level 1 used fast outputs coming from the muon chambers for muon triggers and from the calorimeters for electrons/photons and jets triggers. Both electromagnetic and hadronic calorimeter towers were summed into trigger towers in a window $\Delta\eta\times\Delta\phi = 0.2\times15^o$. The trigger signals from the detector were sent to the trigger electronics and separately stored until a level 1 decision was made. No deadtime was introduced at this level as a reset signal was sent in time for the next beam crossing if a level 1 accept was not satisfied at a given crossing. Level 1 calorimeter triggers required the sum of E_T for all calorimeter towers (individually above a given threshold, typically set to 1 GeV) to be greater than a fixed threshold ($30 \div 40$ GeV). At a typical luminosity of $5 \cdot 10^{30}$ cm^{-2}s^{-1} the rate of level 1 triggers was about 1 kHz.

The level 2 trigger started on a level 1 accept. A hardware clustering processor searched for energetic clusters only considering towers above a programmable threshold. Electromagnetic and hadronic energies were separately summed up for all towers identified as belonging to the same cluster and the E_T, η and ϕ mean values of the clusters were formed and sorted in a list. CTC tracks, provided by a fast (10 μs) hardware tracking processor (**CFT**), were matched to the electromagnetic clusters and muon system segments to make candidate electrons and muons. The final trigger was a selection on muons, electrons, photons, jets and $\not\!\!\vec{E}_T$. The level 2 trigger output rate was about 12 Hz.

The third triggering level was constituted by commercial processors reading the events selected by the level 2 trigger and submitting them to the same software reconstruction algorithms used in the "off-line" analysis. Most of the execution time was used for the three-dimensional track reconstruction in the CTC. The events accepted by this filter algorithm were finally stored on magnetic tape, with about 5 Hz output rate, for off-line processing.

Chapter 2

Jets at CDF

Jets are among the most interesting "objects" produced by a high energy collider, as they characterize the experimental signature of many known physical events as well as of new physics. Jet physics and phenomenology are described in detail in Appendix A. This chapter describes how jets are reconstructed and corrected in CDF to reproduce well the energy and direction of the partons originating them. The importance of improving the jet energy resolution and its impact on some of the Run II main physics goals is also shown.

2.1 Jet Reconstruction

In a collider experiment a jet appears typically as an energy deposit shared among several calorimeter towers. A reconstruction algorithm is then needed to recognize and reconstruct a jet starting from the energy information on each calorimeter tower.

Fig. 2.1 give us an idea of the jet "development" in the CDF detector. It is the event display of a typical di-jet event where two jets, balancing each other in the transverse plane, are produced.

The CDF jet reconstruction process can be divided into two parts: first a list of towers is assigned to each jet by a clustering algorithm (the *cone algorithm*), then the energetic and geometrical information of each tower is combined to reconstruct the jet energy and direction (jet four-momentum). In this section both of them will be described as they were performed in Run I. The cone algorithm implementation for Run II will be almost the same with some difference due to the changed geometry of the plug calorimeter towers. However complementary algorithms are at present under study [16].

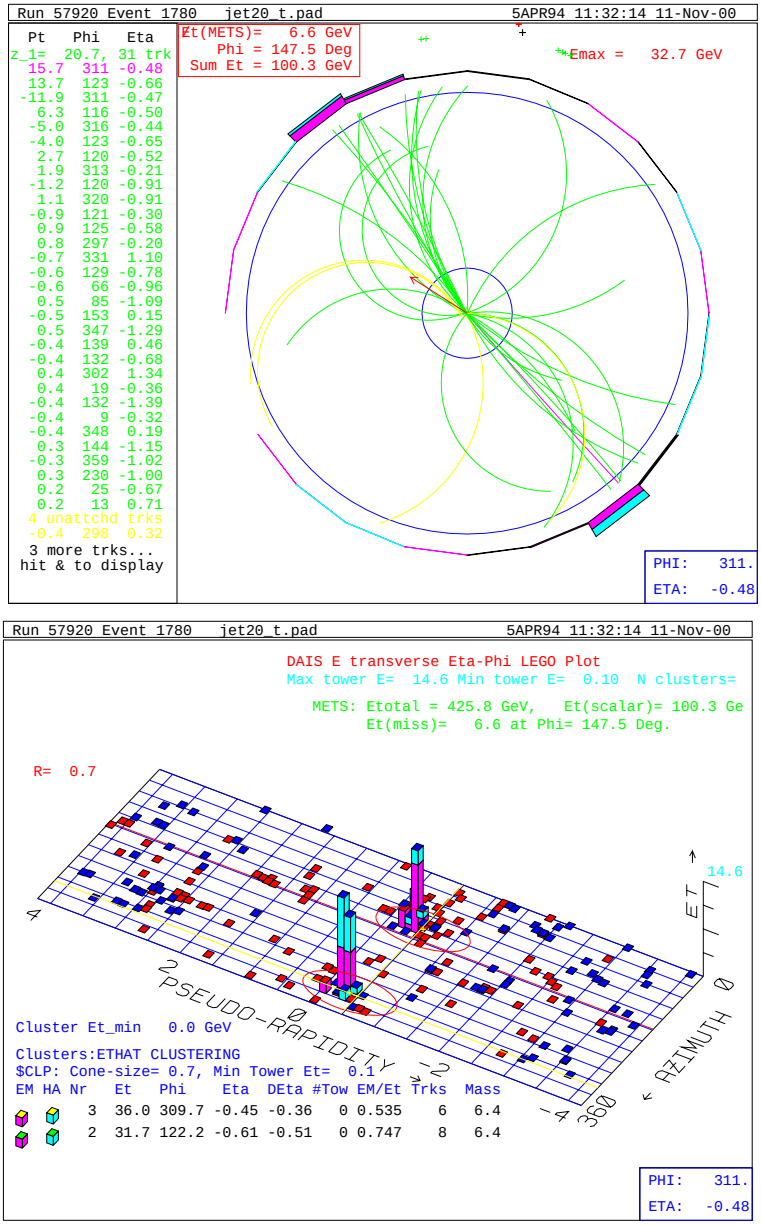

Figure 2.1: *A typical CDF event where two jets are produced balancing each other in the transverse plane. The CTC-plot (upper) allows to see the charged particles (mainly $\pi^{\pm}$) associated to the jet while the lego-plot (lower) shows the jet energy sharing among the calorimeter towers. The electromagnetic deposit (magenta) is mainly due to $\pi^0 \to \gamma\gamma$ while the hadronic one (blue) is due to both charged and neutral hadrons. Only towers with energy above the clustering threshold (0.1 GeV) are shown. Also shown is the reconstruction cone radius set to 0.7.*

2.1.1 CDF Jet Clustering Algorithm

The CDF jet clustering algorithm uses a cone of a fixed radius to define a jet, the CDF calorimeter towers being its basic units. Energetic towers are assigned to jet clusters by an off-line routine (JETCLU) [17, 18] implementing the jet-finding algorithm in three steps:

- **Preclustering**: a list of "seed towers" above a fixed E_T threshold (set to 1.0 GeV) is created and sorted in order of decreasing E_T. In the plug and forward calorimeters, towers are grouped together in sets of three in ϕ to have a $\Delta\phi = 15^o$ segmentation corresponding to the central one. Seed towers are grouped into *preclusters* consisting of an unbroken chain of continuous towers with decreasing energy. If a tower is outside a 7×7 window around the largest E_T seed, it is used to form another precluster. The precluster list is the starting point for the next step.

- **Cone algorithm**: using the true tower segmentation, the E_T weighted centroid of each precluster is found and a cone in η-ϕ space of radius R (typically R = 0.4, 0.7 and 1.0) is formed around it [1]. A loop is then performed over all towers with E_T above 0.1 GeV, including them in a cluster if their centroid is inside the cone around it. A new cluster centroid is then recalculated from the new set of towers belonging to it and a new cone is drawn around this position. The loop over towers is repeated using the new centroid. This process is iterated until the tower list for each cluster remains unchanged in two consecutive passes.

- **Merging and/or solving overlaps**: the cone algorithm is such that in multijet events two clusters can overlap with some towers being assigned to more than one cluster. In this step clusters are merged into one or left alone so to have each tower uniquely assigned to a cluster. When one cluster is completely contained in another, the smaller of the two is dropped. In the overlapping situation, an overlap fraction is defined as the E_T sum of the common towers divided by the E_T of the smaller cluster. When this fraction is above a cutoff (usually 0.75) the two clusters are combined into one, otherwise they are kept intact assigning each tower in the overlap region to the closest cluster in η-ϕ space. Finally, after the cluster separation or merging, the new centroid of each cluster is recalculated using the new tower list.

[1] R is usually referred as "cone radius" and is chosen to best fit the jet topology of each particular analysis.

2.1.2 Jet Energy and Momentum Reconstruction

From the list of towers associated with the cluster, JETCLU then calculates the jet four-momentum components according the following definitions:

$$E_J = \sum_{i=1}^{N} E_i \tag{2.1}$$

$$P_{x,J} = \sum_{i=1}^{N} E_i \sin\theta_i \cos\phi_i \tag{2.2}$$

$$P_{y,J} = \sum_{i=1}^{N} E_i \sin\theta_i \sin\phi_i \tag{2.3}$$

$$P_{z,J} = \sum_{i=1}^{N} E_i \cos\theta_i \tag{2.4}$$

where i is the tower index and N the number of towers in the cluster. The angles ϕ_i are evaluated according to the CDF coordinate system (see Chapter 1) while the angles θ_i are calculated respect to the event vertex along the beam axis.

Note that, according to these definitions, jets are not massless.

Using the above quantities, the jet transverse energy and momentum $E_{T,J}$ and $P_{T,J}$ are defined as:

$$P_J = \sqrt{P_{x,J}^2 + P_{y,J}^2 + P_{z,J}^2} \tag{2.5}$$

$$P_{T,J} = \sqrt{P_{x,J}^2 + P_{y,J}^2} \tag{2.6}$$

$$E_{T,J} = E_J \frac{P_{T,J}}{P_J} \tag{2.7}$$

2.2 Jet Energy Corrections

As will be shown in next chapter, the jet energy is affected by mismeasurement due both to physics and to detector effects. So proper corrections need to be applied in order to reconstruct, from the *measured* jet P_T, the true momentum of the parton generating it [18]. The pure reconstructed jet E_T and P_T are usually referred as the **Raw** jet E_T and P_T.

2.2.1 JTC96 Jet Corrections

In this section we describe the standard CDF jet corrections used in Run I and which we will consider as a reference in the following of the present thesis. Analogous corrections will be developed for Run II using new data and new Monte Carlo samples accounting for the different run and detector conditions.

These corrections are usually referenced as "**JTC96 Jet Corrections**" (being implemented by the off-line routine JTC96 [19, 20]) and are performed in four separate steps:

- **Relative** corrections: accounting for the effects of gaps and edges inside the calorimeters ("η-cracks") and for non-uniform response of different calorimeters.
- **Absolute** corrections: correcting for non-linear response of the calorimeters to low momentum particles and for effects due to not detected fragmentation products. The *true* parton P_T is estimated from the observed raw jet P_T.
- **Underlying Event (UE)** corrections: the contribution to the jet P_T coming from energy not associated with the hard scattering process is estimated and subtracted.
- **Out-Of-Cone (OOC)** corrections: a fragmentation model is used to estimate the amount of jet energy carried by particles going outside the clustering cone. Such quantity is then used to correct the jet P_T.

A brief description of the method used to extract each correction will be now given.

Relative Corrections

η-cracks effects and possible differences in the detector response among the central, plug and forward calorimeters are accounted for by these corrections providing an uniform jet response as function of η.

The corrections are parametrized as a function of the jet η and raw P_T and are derived by equating in P_T all jets to an equivalent *central* jet [19, 20]. The central region (usually in the $0.2 \leq |\eta| \leq 0.7$ range so to be far from cracks) is considered because here the calorimeter response is flat and non-linearities are well understood from extensive test-beam measurements and checked during data acquisition using tracking informations.

To perform such corrections di-jet events, with at least one central jet, are selected and a relative response function is extracted requiring the P_T

back-to-back balancing of the central jet (usually referred as "*trigger* jet") with the other one (usually referred as "*probe* jet") falling in each calorimetric region. A dependence of such function on the probe jet η is considered.

Di-jet events are selected from the "jet triggers" data samples according the following cuts:

- At least one ("trigger") jet at $0.2 \leq |\eta| \leq 0.7$.
- One additional ("probe") jet with $P_T > 15$ GeV/c.
- No other jets with $P_T > 15$ GeV/c.
- z coordinate of the event vertex $|z| \leq 60$ cm.
- One vertex in the event [2].
- $\Delta\phi_{J1-J2} > 2.7$ radians.

Bias effects from the on-line trigger cut, are avoided requiring ΣP_T(Raw) for the two leading jets to exceed twice the single jet trigger threshold value.

In a perfect detector, from momentum conservation in the transverse plane, the *missing* E_T ($\vec{\not\!\!E}_T$) variable (defined in CDF as the vectorial sum over all energetic towers above a fixed threshold - usually 0.1 GeV - with $|\eta| < 3.6$) is expected to have a very low value being randomly directed in each direction. The occurrence of a high value of such variable (usually above a given trigger threshold) is used to indirectly indicate the presence of an undetected neutrino.

A neutrino is not present in a typical di-jet event, so the $\vec{\not\!\!E}_T$ is attributed to jet mismeasurements and is expected to be correlated to their directions (see also section 4.2.3).

We call $\vec{\not\!\!E}_T$ *projection fraction* (MPF) the ratio between the $\vec{\not\!\!E}_T$ vector projection along the $\vec{P}_T^{Probe}$ vector direction ($\widehat{P}_T^{Probe}$) and the mean value of P_T^{Probe} and $P_T^{Trigger}$:

$$MPF = \frac{2(\vec{\not\!\!E}_T \cdot \widehat{P}_T^{Probe})}{P_T^{Probe} + P_T^{Trigger}}$$

In the $\not\!\!E_T \ll P_T^{Probe}$ hypothesis, $\vec{\not\!\!E}_T \cdot \widehat{P}_T^{Probe} \simeq P_T^{Trigger} - P_T^{Probe}$.
So we obtain:

$$MPF = \frac{2(P_T^{Trigger} - P_T^{Probe})}{P_T^{Trigger} + P_T^{Probe}}$$

[2] For Run Ib this cut was loosened to two in order to improve the statistics, as a large fraction of the events had more than one vertex due to high luminosity [20].

Finally, defining the relative jet scale correction factor β as

$$\beta = \frac{P_T^{Trigger}}{P_T^{Probe}}$$

we get

$$\beta = \frac{2 + MPF}{2 - MPF}$$

which results to be a function of P_T and η.

The relative corrections are derived by fitting the β distribution with a continuous curve as a function of the probe jet η in different ΣP_T(Raw) ranges [20].

Figure 2.2 reports the Run Ib relative corrections obtained with this method.

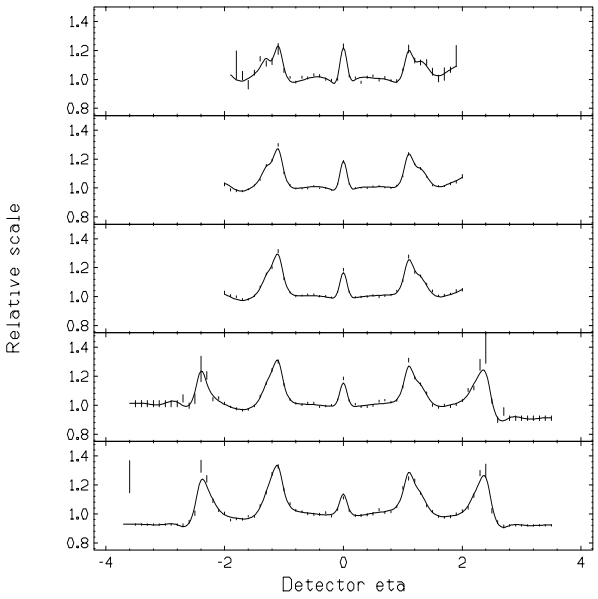

Figure 2.2: *Run Ib relative corrections for a cone radius R = 0.4 [20]. Each ΣP_T bin (mean values from top to bottom: 363, 241, 172, 125 and 64 GeV/c) corresponds to each jet trigger sample: JET20, JET50, JET70 and JET100. The effect of "η-cracks" is in particular evident.*

Absolute Corrections

These corrections are introduced to get the best energy estimate of the original parton generating the jet. The parton energy is usually underestimated mainly because of nuclear absorption, particle leakage and nonlinear calorimeter response.

Monte Carlo simulations are used to obtain these corrections according the following definitions:

- A given parton is associated to a jet if their directions are close in the η-ϕ space below a fixed distance (usually $\Delta R < 0.4$).
- The jet P_T is evaluated summing the P_T of all particles falling into the clustering cone.
- A P_T dependent correction factor is defined as the ratio between the parton P_T and the jet raw P_T:

$$\alpha(P_T) = < \frac{P_T^{Parton}}{P_T^{Jet}(Raw)} >$$

Because of the nonlinearity in calorimeter response, the observed jet energy is a function not only of the incident parton energy but also of the momentum spectrum of the particles produced in the jet fragmentation. Particular attention was so made in order to properly reproduce in the simulation the *observed* jet fragmentation. The parameters of the Monte Carlo event generator accounting for parton fragmentation (SETPRT), implemented in the QFL simulation package of CDF, were tuned in order to reproduce the longitudinal and transverse (charged) fragmentation properties observed using the CTC [18].

Underlying Event Corrections

The underlying event corrections (**UE**) take into account all the contributions to the jet energy not coming from the original parton. Hadron collisions are in fact characterized by an "ambient energy" (usually referred as *underlying event*) which is produced by soft interactions of spectator partons (the "beam-beam remnants") and by initial state gluon radiation (ISR) effects [18, 21] (see fig. 2.3).

Both data and Monte Carlo samples are used to study the UE effects. The (ambient) energy density, to be used in the UE corrections, is estimated from "minimum bias" events[3] considering the $\sum E_T$ over the calorimeter towers for $|\eta| \leq 1.0$, normalized to the overall coverage. The absolute correction factor is applied to the tower energy to get the correction at parton level.

It is clear that for a bigger jet reconstruction cone, a bigger UE incidence is observed so these corrections need to be derived for different jet cones.

[3]Minimum-bias events are selected by a trigger just demanding the occurrence of a collision.

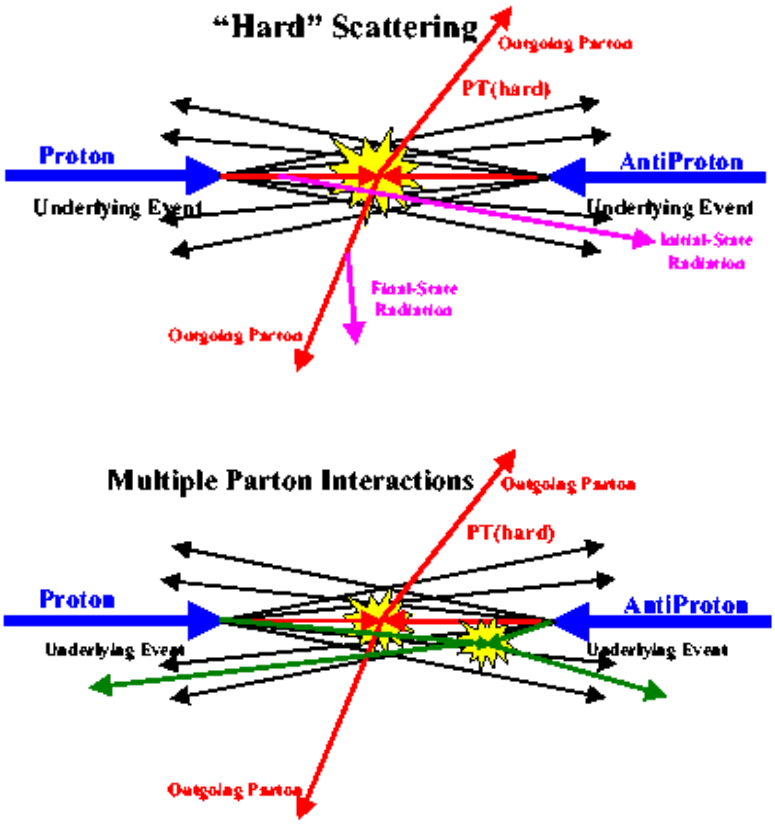

Figure 2.3: *Illustration of two typical $p\bar{p}$ collisions with a "hard" 2-to-2 parton scattering originating a di-jet event. Top: only a primary interaction is present. The final state contains particles coming from the two outgoing partons (*"true" *jet particles), including the final state radiation (FSR), and particles coming from the UE. The UE consists of the beam-beam remnants plus ISR. Bottom: a multiple parton interaction has occurred. In addition to the "hard" 2-to-2 parton scattering also a "semi-hard" one is present contributing particles to the UE.*

Moreover, as the UE effects are also due to particles coming from other interactions in the same bunch crossing (multiple vertices), the UE corrections are also parametrized in function of the the number of vertices in the event (N_v) found by the SVX. Table 2.1 shows the different parametrization of these corrections for the three jet reconstruction cones, with a dependence on N_v. No P_T^{Parton} dependence is observed for these corrections.

Clustering Cone	0.4	0.7	1.0
UE for MC	0.370	1.133	2.312
UE for data (a)	0.297·N	0.910·N	1.858·N
UE for data (b)	0.65	1.98	4.05
OOC for data and MC	1.95+0.156·P_T	1.29+0.052·P_T	0.54+0.022·P_T

Table 2.1: *UE and OOC corrections used in JTC96. N = N_v - 1, N_v being the number of vertices found in the same $p\bar{p}$ bunch crossing. Line (a) must be subtracted before absolute corrections (to have a better jet raw P_T estimate), line (b) after them. The OOC corrections are added as last step. All values are given in GeV.*

Out-of-Cone Corrections

The out-of-cone (**OOC**) corrections recover the jet energy falling out of the clustering cone because of fragmentation effects and gluon radiation (see fig. 2.3 top).

The study of cone losses is performed with the same Monte Carlo sample used to obtain the absolute corrections. A dependence on the reconstruction cone size is obviously expected. The correction is defined as:

$$\Delta P^{OOC}(P_T, R) = P_T^{Parton} - P_T^{Jet}$$

where P_T^{Parton} is given by the sum of the P_T of all particles coming from the parton while P_T^{Jet} is given by the sum of the P_T of all particles coming from the parton and falling inside the reconstruction cone. ΔP^{OOC} can be linearly parametrized as a P_T function, as shown in table 2.1.

2.2.2 Uncertainties in Energy Scale

The absolute jet energy scale is subject to a systematic error mainly due to the uncertainties in the single-particle response of the calorimeters and in the jet fragmentation function. Dedicated studies have shown an uncertainty ranging from $\sim$ 10% for low E_T jets ($E_T \sim 25$ GeV) to $\sim$ 4% for very energetic central jets ($0.1 \leq |\eta| \leq 0.7$) [18]. Usually a $\sim$ 5% uncertainty is quoted in most of analyses where high-E_T central jets are present.

2.2.3 Specific Corrections

The standard JTC96 jet corrections are not suitable in reconstructing the right energy for jets generated by b-quark fragmentation (b-jets). The ana-

lysis of HERWIG $t\bar{t}$ Monte Carlo samples has in fact shown significant disagreements between reconstructed and primary parton energies [22]. The origin of this disagreement can be traced on the fact that b-jets are characterized by a secondary decay vertex inside them due to B hadron decay. So a relevant part of the jet energy can be carried away by additional particles going outside the clustering cone or by a not detected neutrino or a muon (not interacting in the calorimeters) which are generated in the semileptonic B hadron decay.

On the other hand the b-quarks jets occurrence is typical of top-quark production so, for the top mass analysis, an additional set of jet corrections (usually referred a "*AA Corrections*") has been developed to be applied after the JTC96 ones so to bring the jet energy to agree with the HERWIG parton energy [22] [4].

The AA corrections are applied using a parametrization which depends on the B hadron decay mode (i.e. distinguishing if semileptonic or not and, in the former occurrence, if an electron or a muon is present). The parameters are obtained with a fitting method using the HERWIG Monte Carlo.

2.3 Jets in Physics Events

As already mentioned, many interesting physics signatures in an experiment like CDF have quarks or gluons in the final state which, because of the fragmentation process, experimentally appear as hadronic jets. The importance of improving the jet energy resolution is then evident. In the following, while briefly describing the CDF II physics search program, we will give two important examples where an improved jet energy resolution will play a leading role: top quark mass measurement and light Higgs search.

2.3.1 The CDF II Physics Program

Several fields of interest in particle physics will be investigated in Run II data analysis [5]. The CDF II physics program can be summarized into five main goals:

- Top quark properties studies.
- Higgs boson and new phenomena beyond the SM direct search.

[4] These corrections are indeed fundamental in every analysis where b-jet characterize the final state. An example is given by the Higgs boson search in the associate production channel WH $\rightarrow$ W$b\bar{b}$.

- Precision electroweak physics measurements.
- B meson physics.
- Test of perturbative QCD at NLO and large Q^2.

The two measurements more sensitive to the jet energy resolution will be now detailed.

Top Quark Mass

The most important physics result obtained by the CDF collaboration in Run I was the discovery of the top quark and the first direct measurement of its mass and cross section [6, 23] (see fig. 2.4).

Combined with the W mass, m_{top} gives information about the mass of the Higgs boson, the missing particle foreseen by the Standard Model. Fig. 2.5 shows the predicted limits on the Higgs mass given by the present resolutions on top and W masses. It is so evident that the top quark mass will be one of the most important electroweak measurements performed at the Tevatron in Run II.

Currently, the statistical and systematic uncertainties on CDF top mass measurement are both about 5 GeV. The statistical uncertainty should scale as $1/\sqrt{N}$. The CDF collaboration is confident to reduce it to ~ 1 GeV/c^2 in the optimized lepton + $\geq$ 4-jet sample with at least one b-tagged jet [5].

In Run II, systematics will dominate the uncertainty on m_{top}. With the new integrated tracking, the acceptance for double-tagged lepton + $\geq$ 4-jet events can increase by about a factor of 2.5. In these events, the probability of misassociation among jet and parton is lower. By reducing this kind of systematic uncertainty, the top mass resolution will improve by $\sim 20\%$. Moreover the systematics due to the b-tagging bias may be better understood for this class of events.

Almost all the remaining systematics in the measurement of m_{top} are coupled to the reliability of the Monte Carlo models to get the spectrum of fit masses in signal and background. Assuming an accurate theory model, most of the uncertainty is related to detector resolution effects. Instrumental contribution include calorimeter nonlinearity, losses in cracks, dead zones and absolute energy scale. A larger and more difficult part of the energy resolution concerns the reliability of the extrapolation to the original parton energy [5].

An improved jet energy resolution is so very important in reducing the experimental errors on the m_{top} measurement as part or all (depending from the decay channel involved) of the top mass is carried away by hadronic jets.

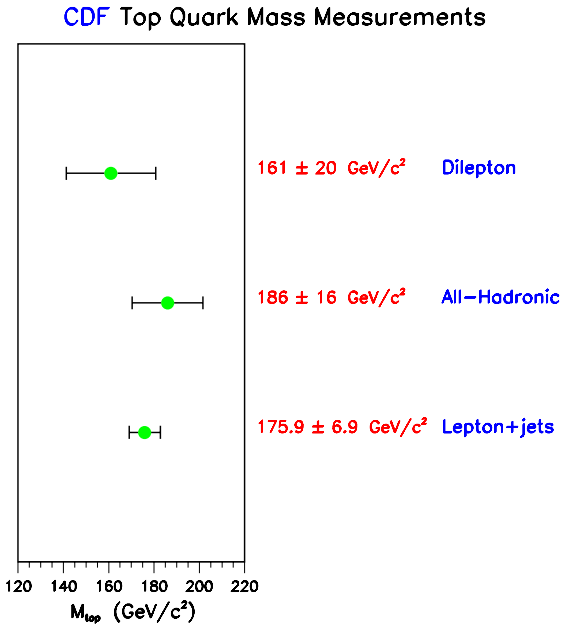

Figure 2.4: *The CDF top mass values achieved in Run I in each decay channel. Statistical and systematic uncertainties are summed in quadrature.*

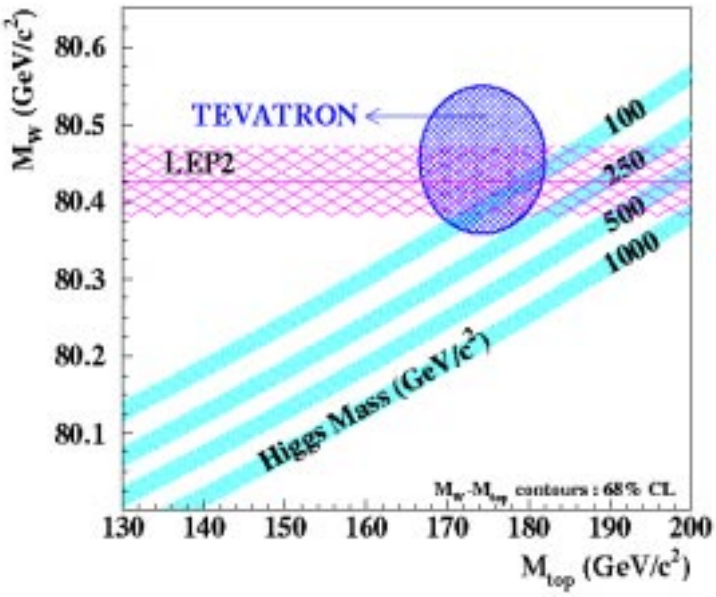

Figure 2.5: *Higgs mass constraints as obtained from the latest m_W and m_{top} measurements performed at LEP2 and Tevatron (CDF and D0 average values).*

Light Higgs Physics

Despite the success of the Standard Model, the dynamics responsible for electroweak symmetry spontaneous breaking is still unknown. The simplest model for this mechanism is the standard Higgs model, based on a doublet of fundamental scalar fields [24]. This model predicts the existence of a new particle, the *Higgs boson*, of unknown mass but with fixed couplings to other particles.

The search for the Higgs boson is one of the primary goals of present and future high energy physics experiments.

Some indication of a potential Higgs signal around 114 GeV, has recently been observed at LEP II in the ALEPH experiment [25]. However the collected statistics does not allow to well reveal the origin of such anomaly. Anyway a lower limit on the Higgs mass up to 114 GeV, is at the moment quoted from LEP II [26, 27, 28].

A lower limit up to $\sim$ 120 GeV, can be reached at CDF with 2 fb^{-1} of data which will be collected in Run IIa (see fig. 2.6).

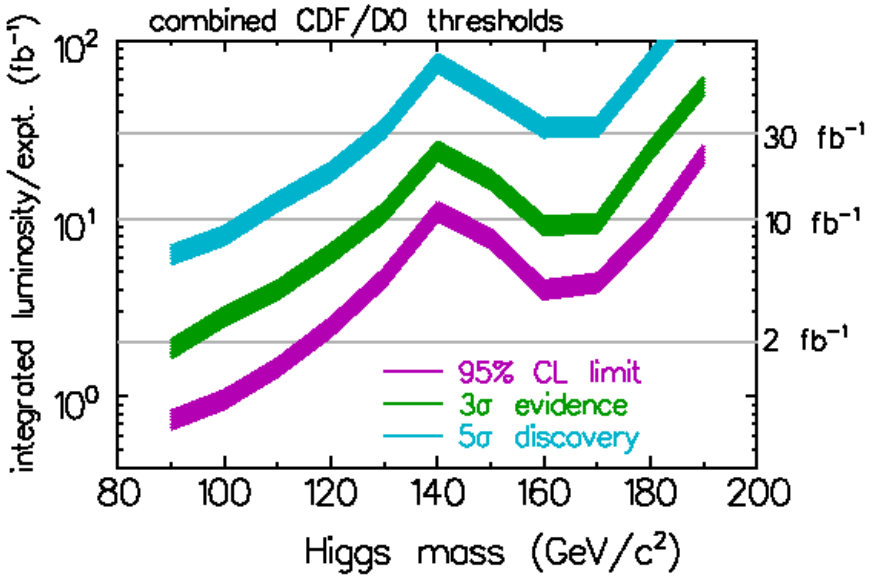

Figure 2.6: *The reach of the Standard Model Higgs search at the upgrade Tevatron. The integrated luminosities delivered per experiment, which are required to exclude the Higgs boson at 95% CL, observe it at the 3 sigma level or discover it at the 5 sigma level, are shown as a function of Higgs mass. The lower edge of the bands is the calculated threshold; the bands extend upward from these nominal thresholds by 30% as an indication of the uncertainties in b-tagging efficiency, background rate, mass resolution, and other effects [29].*

Higher masses will be explored by the CERN Large Hadron Collider (LHC) [30, 31]. However, the light intermediate-mass region $m_H \lesssim 130$ GeV, which is also the favored region for a light Higgs boson predicted in the minimal supersymmetric theory, is the most difficult to investigate at the LHC. The CMS detector intends to cover this region with the rare decay $H \to \gamma\gamma$ [30]. Also the ATLAS detector covers down to $m_H \approx 110$ GeV with this decay mode [31].

The dominant decay mode of the Higgs boson in this mass range is $H \to b\bar{b}$ (the two b quark as usual originating two jets), with a branching ratio of about 80%. The process $q\bar{q} \to WH$, followed by $H \to b\bar{b}$ and the leptonic decay of the W boson (the lepton triggering the event), is considered as the ideal one to discover the light intermediate-mass Higgs boson at the Tevatron [32, 33, 34]. This signal may be more difficult to detect at LHC due to very large top quark background.

Let's see in more detail the conditions to discover a light Higgs boson at CDF in Run II.

$q\bar{q} \to WH$, $H \to b\bar{b}$

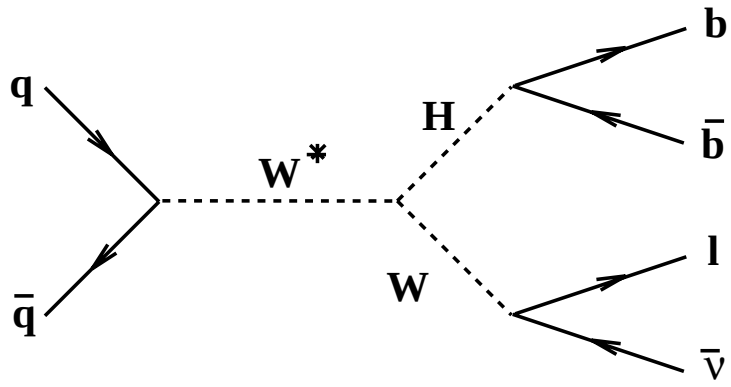

Figure 2.7: *Production of Higgs boson in the process WH with* $H \to b\bar{b}$.

The associate production of a Higgs boson and a W or Z boson, with the Higgs decaying to $b\bar{b}$ and the W or Z decaying leptonically (e or μ), is a possible way to detect the Higgs in the mass range up to 130 GeV (see fig. 2.7).

The main source of background, simulating the associated WH production process, will be W + 2-jet events [5].

From Run I data, we expect in Run II a W + 2-jet background about 1000 times larger than signal. The b-tagging techniques [6], previously developed in the top search, can considerably reduce this background. Moreover the new Silicon Vertex Detector (SVX II) is expected to improve both the fiducial acceptance to $|\eta| \lesssim 2.0$ and the efficiency of b-tagging algorithm.

A preliminary study also considered the impact on the Higgs search of a $\sim 30\%$ improved jet energy resolution [35]. Such a improvement is expected to be in principle achieved with dedicated studies like the work exposed in this thesis. Some results of this study are shown in fig. 2.8 and fig. 2.9. The former reports the expected signal+background distributions using 10 fb^{-1} of data for Higgs masses of 80 and 100 GeV. The present jet resolution is considered in these plots. Clearly this resolution does not allow to see any signal, specially at higher masses. The latter reports the same plots with the improved jet resolution.

A better jet energy resolution, which implies a more precise di-jet mass measurement, results so crucial in the search of the Higgs boson as well as of new physics producing a "bump" in the di-jet mass spectrum.

[5] Other backgrounds to this process will be:

1. $Wb\bar{b}$ production
2. WZ with $Z \to b\bar{b}$
3. $W^* \to tb$ (single top production)
4. W-gluon fusion ($t + q + b$ final state)
5. $Wc\bar{c}$

[6] These methods, mainly relying on track informations from the Silicon Vertex Detector, allow to tag a jet as being originated from a b-quark.

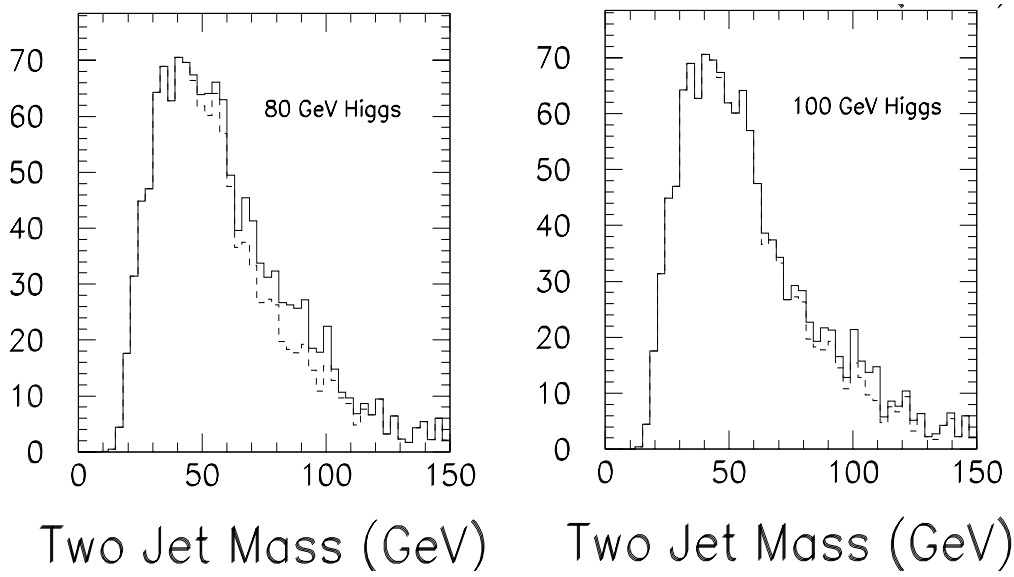

Figure 2.8: *Expected signal+background mass distribution for the WH process with 10* fb^{-1} *of data at 2 TeV. The solid line is signal+background, the dashed one is the sum of all backgrounds. The nominal jet resolution is assumed.*

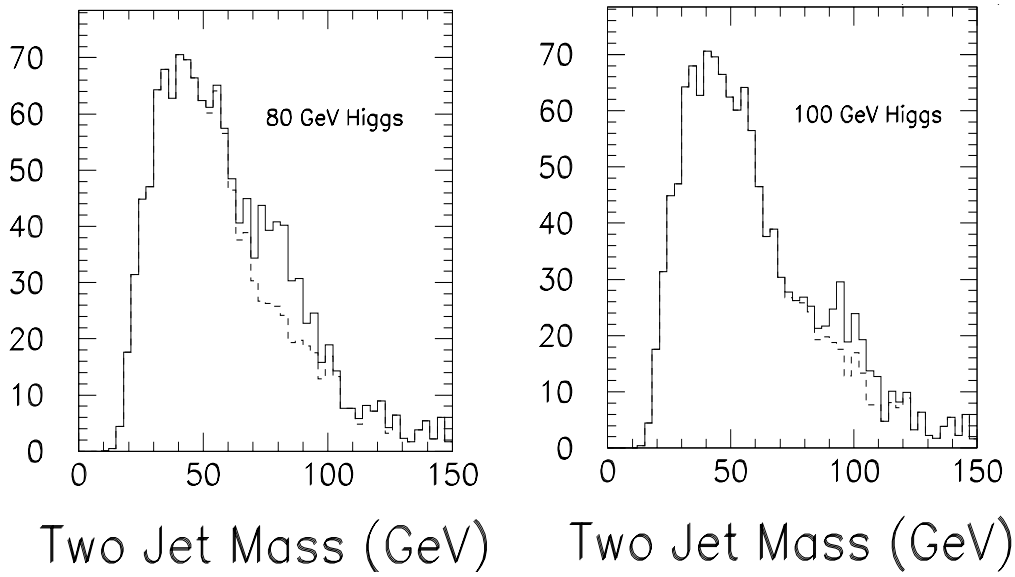

Figure 2.9: *Expected signal+background mass distribution for the WH process with 10* fb^{-1} *of data at 2 TeV. The solid line is signal+background, the dashed one is the sum of all backgrounds. A* 30% *better jet energy resolution than the nominal one is assumed.*

Chapter 3

A New Jet Correction Algorithm

In this chapter a new method to correct the jet energy for detector effects due to calorimeter non-linearities and fluctuations in the single charged particle response is introduced. For the first time the full calorimeter granularity is used to perform energy corrections at "tower level" rather than the usual "jet level". Using the information on charged particle momentum from the Central Tracking Chamber (CTC) and on neutral electromagnetic cluster energy from the Central Shower Max (CES), towers are classified into four groups and then corrected in energy according to the kind of particles hitting them. The algorithm is optimized using both data and CDF detector simulation and implemented in a offline code allowing it to work on each sample. Further potential algorithm developments are also discussed.

3.1 Introduction

The jet energy resolution originates from many sources which can be grouped into two categories: physics effects, such as fluctuations in the energy outside the clustering cone, and detector effects, such as calorimeter resolution.

Physics Effects

Dedicated studies have pointed out the main physics effects involved in jet energy resolution with a particle level study of Monte Carlo $W \rightarrow q\bar{q}$ events [36]. As the main goal of improving the jet energy resolution is to get a better di-jet mass measurement, these studies directly looked at the impact on di-jet invariant mass resolution. Jet reconstruction uncertainties coming from physics (which would be also present if the energy of each particle in the

jet would be exactly known) were separately studied and compared to pure detector effects.

Fig. 3.1 shows the different impact on the di-jet mass distribution of simulated $W \rightarrow q\bar{q}$ events, due to these two contributions. The physics effects tend to dominate the distribution tails while the detector ones give their main contribution to the core of the distribution. In these studies,

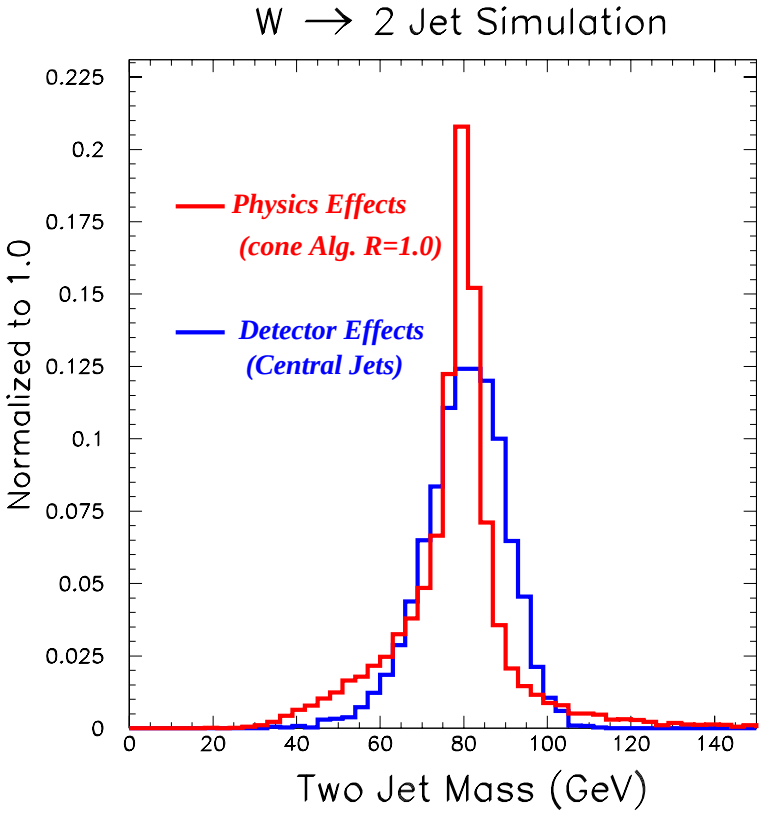

Figure 3.1: *The $W \rightarrow q\bar{q}$ mass distribution with pure algorithm effects (red), and pure detector effects (blue).*

four main different physics effects were considered: the natural width of the decaying object, underlying event fluctuations, final state gluon radiation (FSR) and misidentified jets from initial state radiation (ISR). The dominant effect was found to be the FSR causing a significant tail at low masses. No significant difference was found considering different clustering algorithms. An exhaustive study made to improve this tail showed a merging radius of 1.0 and a cut at 8 GeV on extra jet activity to be optimal.

For these reasons in the following we will always consider a reconstruction cone radius of 1.0 and a cut at 8 GeV on extra jets.

Detector Effects

A first study made to well understand these effects was performed both on the CDF detector simulation and on a suitable Run I data sample [36, 37, 38].

Detector effects mainly include calorimeter non-linearity, magnetic field effects, and uninstrumented detector areas. In central calorimeters the response non-linearity is the largest contribution to the mismeasurement, due both to the non-compensating calorimeter ($e/h > 1.0$) and to the calibration procedure.

The absolute calibration of the calorimeter modules was originally established in a test-beam where the response to high energy ($E > 10$ GeV) pions and electrons was measured. The EM compartment was calibrated using electrons, while the HAD response was measured using pions leaving only minimum-ionizing signals in the EM calorimeter. A standard procedure made by three separate steps (using both radioactive sources and light from laser and LED devices) allows calorimeter calibration during data acquisition so to bring the detector response to the original test-beam calibration.

Using this calibration prescription, and taking the tower energy as the sum of the energy of the EM and HAD compartments, the response to low energy ($E < 10$ GeV) isolated tracks (mainly pions) has been studied in minimum bias events [39]. The average calorimeter response to low energy particles was found non-linear with energy and is reasonably well reproduced by QFL and CDFSIM, the CDF detector simulation programs. The origin of such non-linearity (even present in a perfect detector with no cracks and infinite thickness) is inherent in any calorimeter based on particle showering and having different responses to hadronic and electromagnetic showers.

Jet particles have a low P_T spectrum (typically below 10 GeV) so the non-linearity effect is relevant for them. Furthermore, because of the fragmentation process, there are large fluctuations both in the neutral/charged mixture and in the energy sharing between hadrons inside the jet. Thus the non-linear calorimeter response causes a degradation of the jet energy resolution and a bias, since the jet energy with high particle multiplicity is underestimated.

To overcome the above problems, one idea can be that of considering the information on charged particle momentum given by the Central Tracking System to reduce the non-linearity effect in the measurement performed by the calorimeters. This procedure can also improve the calorimeter response to single charged particle as the particle momentum resolution from the tracker is considerably better than the calorimeter resolution on the same particle energy (see Chapter 1).

3.2 Preliminary Study With the CDF Simulation

Because of the lack of complete CDF II simulation codes and of data sample obtained in this experimental configuration, the present study and the following algorithm development were performed considering the CDF I setup. In particular the track reconstruction is performed by the Central Tracking Chamber (CTC) covering only the central detector zone ($|\eta| < 1.1$). The new tracking chamber (COT) has the same geometrical and resolution characteristics of the old one (see Chapter 1) so all the results obtained with the CTC are expected to be well reproduced with the COT. The algorithm implementation for CDF II will be discussed in section 3.8.

The standard CDF jet-finding and reconstruction algorithm JETCLU (see Chapter 2) combines the energies of the electromagnetic and hadronic calorimeters of each tower, but makes no use of tracking information. On the other hand, almost 2/3 of the jet energy is carried by charged particles whose trajectories are efficiently reconstructed by the CTC [1].

For charged hadrons, the CTC momentum measurement is much more precise than the calorimetric energy measurement for most of the energy range of interest. Even without vertex constraints, the crossover for $\Delta P_T/P_T = 0.002 \times P_T$ (the resolution with which track momenta are measured by the CTC) and $\Delta E_T/E_T = 0.5/\sqrt{E_T}$ (the calorimeter energy resolution for isolated charged particles) is about 50 GeV. Even jets well above this energy can potentially be better measured using the CTC information as usually they fragment into several charged hadrons.

So it is important to understand whether at least part of calorimetric information on the jet energy can be replaced by the tracking information on the jet charged particles. The main problem which may restrict such an attempt is that the CTC does not give energetic information of *neutral* particles such as photons (mainly from π^0 decay), neutrons and K_Ls.

Previous works made in this direction by several authors of the CDF collaboration were not very encouraging (the best improvement achieved was about 7%) [40, 41, 42, 43].

Confusion between neutral and charged particles can seriously compromise the effort. Nevertheless for low energy jets the particles are sufficiently spread out so that their overlap may be not a great problem. In this situation tracking system and calorimetry may provide complementary information.

[1] The CTC has a track reconstruction efficiency of better than 80% in the core of jets (for jet E_T up to 100 GeV). For isolated tracks the efficiency is better than 98% [18].

In particular for these low energy jets, where the calorimetric measurement gives the largest error, tracking may provide a significant improvement. For high energy jets, where track-finding may fail because of the increased track density in the jet core and where track momentum measurement becomes less accurate, the improvement might be considerably reduced.

Preliminary studies and considerations showed the importance of distinguishing overlapping particles inside a tower and to take into account that usually hadron showers are NOT contained into an only tower while electromagnetic ones usually are [38].

To justify the development of a new algorithm further studies were performed with a γ-jet Monte Carlo sample.

Using both the known particle list and the CDF detector simulation a check was performed on the energy resolution potential improvement if track momenta measurement for charged particles and calorimeter measurement for neutral particles is used. The jet energy resolution has been studied moving gradually from a perfect detector condition (which exactly measures the energy of each particle) to the full CDF detector resolution.

Five situations were taken into account:

1. Perfect detector. The energy of all particles falling into the jet reconstruction cone is exactly measured. Only the energy information of the particles going outside is lost. The resolution comes from the fluctuation on the amount of such missing energy.

2. Magnetic field effect. Charged particles with low P_T ($\lesssim$ 350 MeV) cannot reach the calorimeters because they are bent by the magnetic field ("curl up effect"). They don't contribute to the jet energy and the fluctuation in their incidence give a worse resolution.

3. Nominal resolution for neutrals. In a real detector each particle energy is known with finite resolution. At this step we take into account the neutral particles contribution by applying the nominal calorimeter resolution. For photons we apply the CEM resolution ($13.5\%/\sqrt{E_T}$) while for neutral hadrons we use the CHA resolution ($50\%/\sqrt{E_T}$). We assume that all neutrals are detected (no cracks effect). The charged hadron energy is still exactly measured.

4. Nominal resolution for charged. Now we apply the nominal resolution for charged particles as measured by the central hadronic calorimeter ($50\%/\sqrt{E_T}$). No cracks effect is accounted for.

5. Full CDF detector. All detector effects are included applying the full CDF detector simulation. The "JTC96 corrections" are also applied.

Table 3.1 summarizes the jet energy resolution obtained in the five cases. We used γ-jet Monte Carlo sample events with two different thresholds for the photon P_T and a cone radius $R = 1.0$ for the jet clustering. The jet energy resolution was taken as the width of the balancing distribution that is [2]:

$$\frac{P_T^{jet} - P_T^{\gamma}}{P_T^{\gamma}}$$

	$P_T^{\gamma} > 15$ GeV	$P_T^{\gamma} > 50$ GeV
Perfect detector	4.8%	1.9%
Perfect but 350 MeV cut	5.1%	2.3%
Nominal for neutrals	7.1%	3.2%
Nominal for tracks	16%	8.8%
Detector resolution	19%	11%

Table 3.1: *Jet energy resolution going from a perfect detector condition to the full CDF detector simulation. PYTHIA Monte Carlo γ-jet events, with ISR turned off, were used.*

Going from the third to the fourth step we can note how the jet energy resolution is dominated by the one of the hadronic calorimeter. An improvement can in principle be achieved substituting as much hadron calorimeter information as possible with track momentum. The highest potential improvement is considerable (around 60%) but requires a *perfect* particle identification.

On the other hand achieving such a perfect particle identification inside the jet is quite difficult. So the limit to the true improvement is mainly set by our ability in separating the different particle types hitting each calorimeter tower.

3.3 The Classification Method

In order to achieve the best particle identification possible inside the jet, for the first time the full granularity of the CDF central calorimeters was considered. Additional detector informations were then used to determine which

[2] In γ-jet events $P_T^{parton} \cong P_T^{\gamma}$ and so the Photon-Jet balancing can be seen as a direct Parton-Jet comparison.

kind of particle (charged or neutral) had released some or all of its energy in each tower. In the last step proper energy corrections were applied at tower level using the most appropriate detector information (CTC, calorimeter or both) to get the best estimate of the tower energy.

It is clear that the improvement achieved introducing the track measurement is reduced by the fluctuations in the energy relased by charged hadrons in the electromagnetic calorimeter. Furthermore, using the track momenta and calorimeter response to estimate the particle energy, we need to be very careful to avoid double counting of contributions.

The simplest situation is when we can use either tracking or EM calorimeter information only. For every jet it is possible to identify a set of such "golden" towers specially in low energy jets were the particles are enough spread out. Here we expect a higher improvement as for low P_T jets the "JTC96 corrections" work worse.

In higher P_T jets, specially in the jet core, we expect a greater incidence of the more complicated situation were more than a particle is hitting a single tower. Anyway in principle also for these towers the track momentum can be used to improve the energy measurement but a suitable combined tracking and calorimetry information needs to be adopted.

An operative solution was to "classify" each tower depending on the particle kind hitting it, the energy correction automatically following. Hence from now on the method will be referred as the "**Classification Method**".

Using tracking and CES cluster informations, central calorimeter towers were subdivided into four main tower classes.

3.3.1 Track Tower

Each track was associated to a tower by extrapolating it to the central calorimeter at the strip chambers radius ($R_{ces} = 184$ cm) where the impact point coordinates z and ϕ were evaluated [3].

Each tower hit by a track was then called "target tower" and labeled as "*track tower*".

As hadronic showers are not usually completely confined in the target tower but extend also to the neighbor towers, the charged particle shower leakage needs to be accounted for.

The isolated charged particle response is usually studied summing up the calorimeter energies in a 3×3 window centered around the target tower and comparing this sum to the track P_T measured by the CTC (see for instance

[3] It is a helical extrapolation to the center of the solenoid coil followed by a straight line extrapolation outside the solenoid to the strip chamber radius.

ref. [44]). However, while this choice proved to be the best one on minimum bias events where particles are very well isolated, it is too conservative to classify towers in jet events.

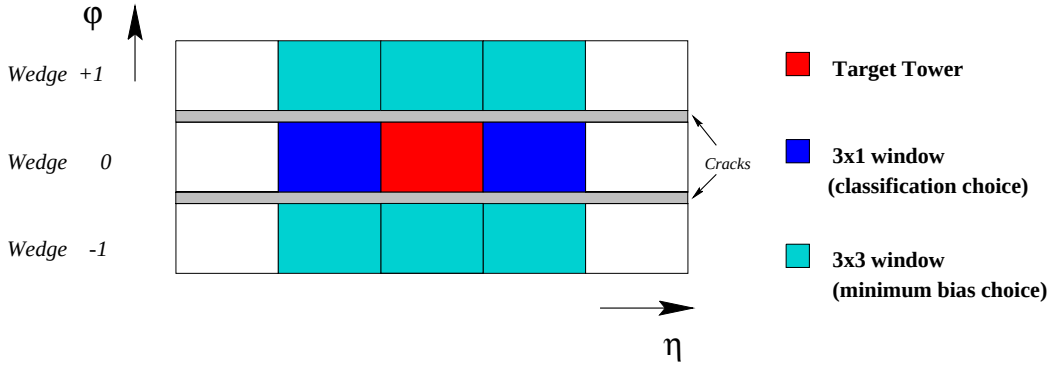

Figure 3.2: *Schematic view of the* 3×3 *window centered around the "target tower" (red). This nine towers matrix is considered in minimum bias studies. The "*3×1 *window" (blue towers) was found optimal for the Classification Algorithm. The figure is not to scale, the "*ϕ*" tower extent being about twice the "*η*" one.*

The energy released by the track shower in the η and ϕ-neighboring towers was studied using a Monte Carlo sample of 10 GeV isolated tracks. The calorimeter response (obtained summing the electromagnetic and hadronic energies) of the towers around the target tower were compared (fig. 3.2). Fig 3.3 shows the energy reconstructed in five (target + four nearby towers), three (target + two nearby towers in the same wedge) and one tower (the target tower only). We concluded that the 3×1 window was the best compromise containing most of the shower energy for a reasonable number of "track towers". This choice is also physically motivated from the fact that the "ϕ extent" of a tower is about twice the η one and, at the same time, two ϕ-neighboring towers are separated by not instrumented cracks which results in a very reduced ϕ-side leakage. Moreover (as we can see in fig. 3.4) most of the tracks in a jet have low energy (well below 10 GeV) which corresponds to a main shower leakage into the "target wedges".

In conclusion, as illustrated in fig. 3.2, we classify a tower like "*track tower*" if a reconstructed track (or more) is pointing to it or to one of the two η-neighboring towers (3×1 window). At the same time we require no CES cluster inside the same tower (see next subsection).

This kind of tower is attributed to a charged hadron (typically $\pi^{\pm}$).

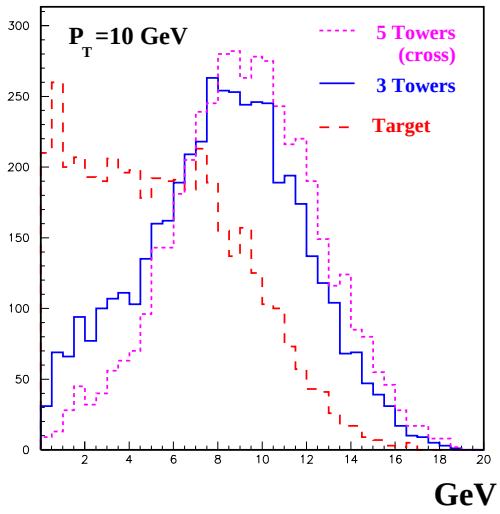

Figure 3.3: *MC 10 GeV isolated tracks energy reconstructed in five (target + four nearby towers), three (target + two η-nearby towers) and one tower (the target tower only).*

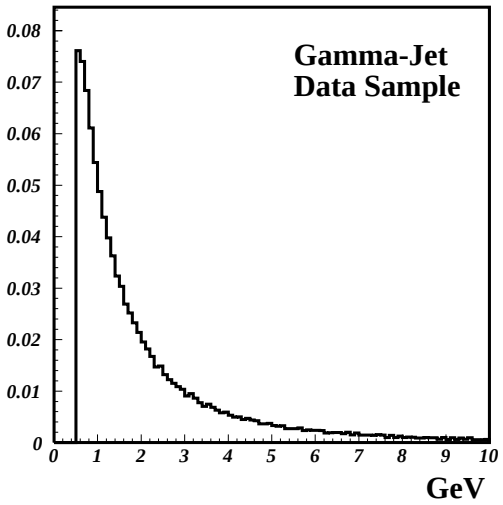

Figure 3.4: *Track momentum distribution for the Run Ib γ-jet data sample.*

3.3.2 Gamma Tower

The presence of photons (mainly coming from π^0 decays) is provided by electromagnetic clusters in the Strip Chamber Detector (CES) (see Chapter 1). These clusters are detected on both z (strip channels) and r-ϕ (wire channels) coordinates with a position resolution of about 2 cm, being reconstructed by

a suitable clustering algorithm [4].

Because of the CES clustering seed threshold, the photon detection is fully efficient only for $E^{\gamma} \gtrsim 1$ GeV decreasing to zero for $E^{\gamma} \lesssim 0.4$ GeV. This not complete photon identification performed by the CES, limits our classification potential, but releasing the threshold cut would introduce a not negligible incidence of fake CES clusters [5].

As for the charged hadron showers, the electromagnetic showers leakage into neighbor towers needs also to be accounted for. Test-beam results showed that electromagnetic showers are well confined in only one tower if the particle hits the tower center while no leakage in the ϕ-boundary tower is present even if the photon (or electron) is falling close to the ϕ-crack [46]. This is because of the two 5 mm thick steel sheets on wedge boundaries. For this reason also for electromagnetic showers our study was restricted to the two closest towers in η (3×1 window).

Using the full detector simulation (QFL) we have studied the CEM response to 4 GeV photons as a function of hit position. Fig. 3.5 shows the calorimeter response (defined as the CEM energy collected in the target tower divided by true photon energy) to a single isolated MC 4 GeV photon for different z positions (corresponding to different photon η) inside the same wedge. As expected, the response is about one if the photon hits the tower center with a not negligible shower leakage in the nearest tower when the photon falls near the tower η-boundaries. Fig. 3.6 (left) reports the same variable as function of the relative η position in one tower. It is evident that, if a photon falls near the tower boundary ($\lesssim 20\%$ of the tower z size) some leakage is present in the neighboring tower [6].

[4] This algorithm can be summarized as follows [14, 45]:

- A 11 wires (or strips) window is considered around a seed wire (or strip).
- All wires (or strips) channels above a fixed threshold (250 MeV for Run Ib) are considered "seed" candidates and ordered in energy.
- Clustering starts from the highest energy seed candidate and continues through all candidates, with the elimination of wires (or strips) used in previously found clusters.
- The shower fit is then performed over the 11 wires and strips for each cluster using a fitting procedure optimized with test-beam electrons in the 10-100 GeV energy range.

[5] A dedicated study, performed by the candidate to simulate the impact of an enhanced noise on the CES readout electronics, showed that the classification algorithm is very sensitive to the seed threshold cut. About 5% (10%) in jet energy resolution is degraded increasing the seed threshold to 500 (750) MeV.

[6] The lower face of the CEM tower is considered, the z size being ~ 20 cm for tower 0,

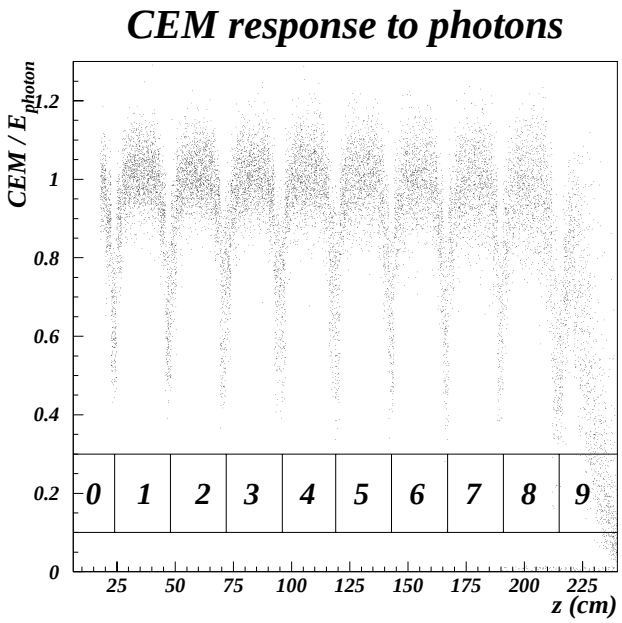

Figure 3.5: *CEM calorimeter response to a MC 4 GeV photon as a function of the z position in the wedge. Tower boundaries and standard numeration are shown. Good response at the tower center, η-adjacent tower leakage and the different behavior of tower 0 and 9 (due to the η-cracks) are also evident.*

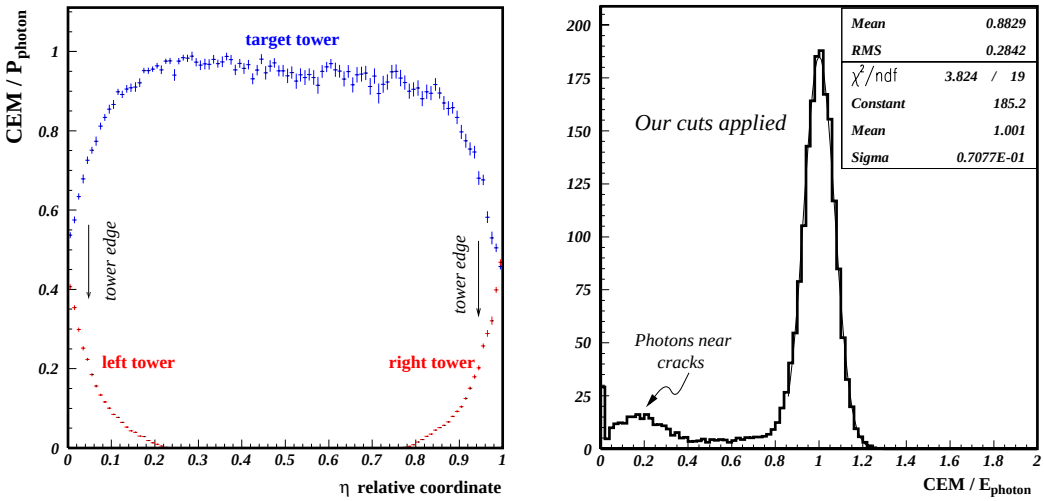

Figure 3.6: *Left: CEM tower response as a function of hit point. Right: total CEM tower response for "gamma towers" related to the true photon energy. A Monte Carlo sample of isolated 4 GeV photons was used.*

∼ 24 cm for towers 1-8 and ∼ 25 cm for tower 9.

From the above considerations we flag a tower like "*gamma tower*" when a CES cluster is present either in that tower or in a η-neighboring tower if the z distance between the CES cluster and the tower edge is less than 20% of the tower z length. The absence of a track pointing to the tower and to the two η-neighboring towers is also required.

This kind of tower is substantially attributed to a π^0 decaying to photons.

A check was performed using the Monte Carlo 4 GeV isolated photons studying the total CEM response of all towers of the event labeled as "gamma". The result is shown in fig. 3.6 (right). The energy resolution of the peak is in agreement with the nominal CEM resolution measured in the test-beam ($13.5\%/\sqrt{E_T}$). The low detector efficiency for photons near the ϕ cracks is also evident.

3.3.3 Not Assigned Tower

A tower which is energetic but with no track pointing to it (or to the two η-neighboring towers) *and* with no CES cluster present is labeled as "*not assigned tower*". Monte Carlo studies have shown that such towers are mainly due either to low energy (E $<$ 1 GeV) photons, not producing a CES cluster, or to a neutral hadron (n,$\bar{n}$,K_L). The less likely occurrence of hadronic shower leakage outside the wedge limits is addressed by flagging a "not assigned tower" as "track tower" if a track is falling in a 3×3 window around it. In such a way, double counting in the jet energy reconstruction is avoided.

3.3.4 Mixed Tower

We label as "*mixed tower*" each tower satisfying *both* prescriptions required for "track" *and* "gamma" towers. This is the most complicated case where two or more particles (with different interaction characteristics inside the calorimeters) overlap in the same tower. Since the energy collected in a mixed tower mainly comes from a track ($\pi^{\pm}$) and from a photon (π^0), we have to estimate how much energy is released by the track inside the CEM to subtract its contribution, so avoiding double energy counting. The neutral hadrons contribution to the CHA deposit needs also to be accounted for.

Previous works performed in the past to include the CTC information in the jet energy reconstruction, met a similar problem. However the overlap problem is now limited to a subsample of towers. In the remaining subsample, made of "golden" towers, we can unambiguously use tracking or EM informations to improve the jet energy resolution. The final results will depend on the fraction of "golden" towers respect to the "mixed" towers.

3.3.5 CES Fake Clusters

In order to achieve the best classification possible, which will be reflected in a better tower energy correction, we need to reduce as much as possible any origin of classification error.

Our classification method relies on track and CES informations. Track reconstruction is affected by a very low rate of fake incidence but signals attributed to photons are more sensitive to fake CES clusters manly due to track losses and to readout electronics noise.

Non MIP Tracks Losses

The Δz distance between the extrapolated track and the CES cluster for "mixed towers" having both charged particle and photon falling inside them is shown in fig. 3.7. The $\Delta z = z_{trk} - z_{CES}$ distribution is expected to be triangular as the charged and neutral particles directions are essentially not correlated at production. However, from fig. 3.7 we can see an excess of events around $\Delta z = 0$. It can be explained in terms of a charged particle starting its shower before the Strip Chamber radius, so producing a CES cluster. In this case the CES cluster is not due to a photon being a fake that we have to remove to avoid the wrong classification of a "track tower" like a "mixed tower".

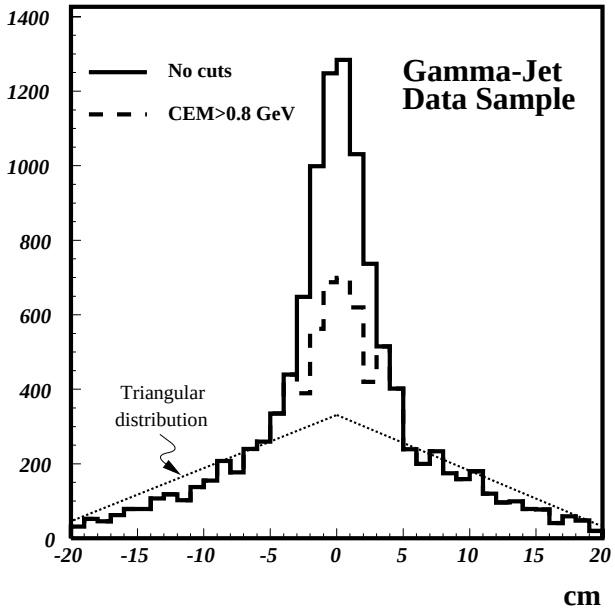

Figure 3.7: *The $\Delta z = z_{trk} - z_{CES}$ distribution obtained applying (dashed) or not (full) a cut on CEM energy. The triangular distribution is also shown. The "mixed towers" in the Run Ib γ-jet data sample were used.*

It is clear that such "fake" CES clusters are very close to the track, but at the same time this zone is also populated by true clusters as $\Delta z = 0$ is the most probable value in a triangular distribution. Anyway a low CEM energy is expected for "fake" clusters as coming from the charged particle shower. From MC studies we decided to consider as "fake" the CES clusters with $|\Delta z| < 3$ cm *and* with a CEM energy less than 0.8 GeV. As shown in the dashed plot of fig. 3.7, applying this cut we obtained a distribution more similar to the triangular one.

Noise Effect CES Clusters

Another origin of fake CES cluster can be found in noise effects. To investigate the incidence of this problem we used a γ-jet data sample applying our "gamma tower" or "mixed tower" classification criteria to the jet towers. In about 0.8% of the events we found at least one of such towers with NO electromagnetic deposition. Considering that the tower energy threshold is 0.1 GeV while the CES cluster seed threshold is 0.250 GeV, we attributed this CES cluster to noise effects.

To overcome this problem we assigned a CES cluster to a tower ONLY if its electromagnetic energy was not zero.

Fig. 3.8 finally gives a graphic summary of tower classification.

3.4 Monte Carlo Study of the Classification Method

A preliminary check of the tower classification was made on a γ-jet Monte Carlo sample [38]. In such a way we can test the classification algorithm as we exactly know from the Monte Carlo particle list what particle is hitting each tower.

In particular with this simulation was found that only few towers are hit by two kind of particles. However we expect a different real situation with a higher fraction of "mixed tower" because in the particle level simulation particles do not shower (this is equivalent to consider a simple 1×1 window for each particle).

Fig. 3.9 is the lego plot of an event were only the central calorimeter is shown. The tower heights are proportional to the collected energy. In fig. 3.9 (left) the tower energies from the Monte Carlo particle list are reported, the

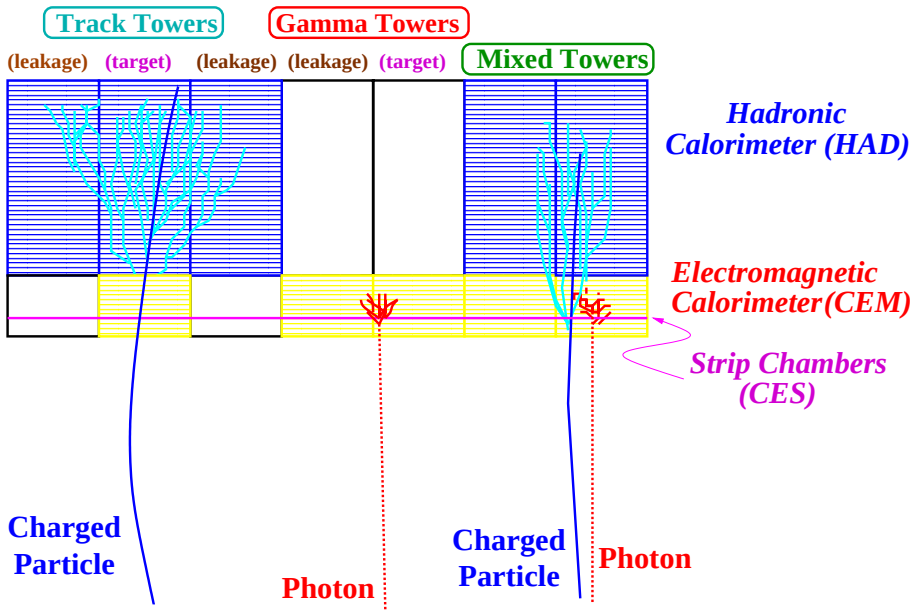

Figure 3.8: *Graphic summary of tower classification. Both target and leakage track, gamma and mixed towers are shown.*

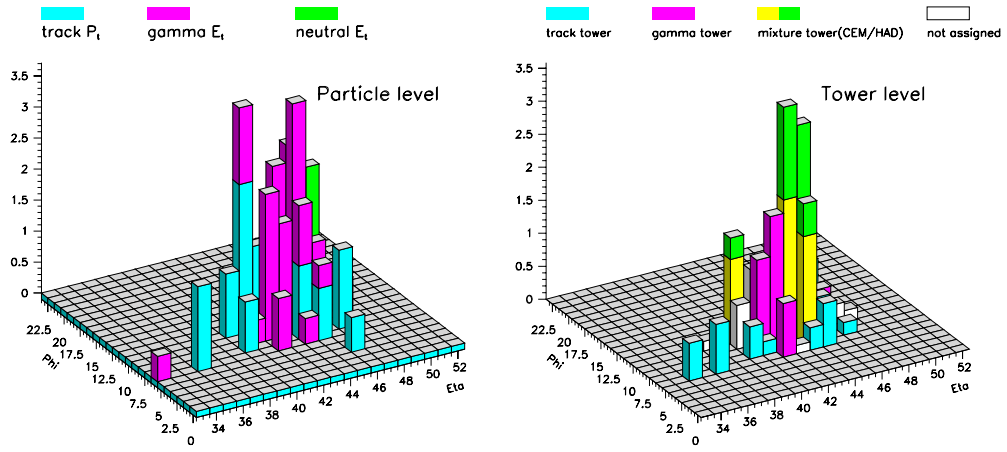

Figure 3.9: *Left: lego plot view of the central calorimeter for a Monte Carlo event. From the Monte Carlo particles list we know the tower energy due to each single particle. Right: the same event is shown after the classification method is applied.*

photon being removed. A red tower means that a gamma fell in that tower, a blue one means that a charged particle hit the tower while a green tower means that a neutral hadronic particle is present. If a tower has two colors, then two different kinds of particles are hitting it.

In fig. 3.9 (right) we can see the same event after applying our classification scheme. The blue towers are classified as "track towers", the red ones are "gamma towers", the white ones are "not assigned towers" and the yellow/green (yellow for CEM energy and green for HAD energy) ones are "mixed towers". We can see how for particles spread outside the jet core, the classification is well working. In the jet core, being the particles very close to each other, several mixed towers are present.

At this point it is clear that the potential improvement we can get with the classification approach is strongly depending on the "mixed towers" incidence in the event. The lower this incidence (and consequently the higher that of "golden towers") the higher the improvement in the jet energy measurement.

3.5 Incidence of Different Classified Towers

A statistic study of the incidence of towers labeled according our classification method was performed using the whole Run Ib γ-jet data sample. Table 3.2 summarizes the results.

Tower type	*Tower fraction*	*CEM fraction*	*HAD fraction*	*Total*
Track	58.6%	22.6%	43.1%	31.0%
Gamma	8.8%	11.6%	2.5%	7.9%
Mixed	17.8%	62.6%	50.8%	57.8%
Not Assigned	14.8%	3.2%	3.6%	3.4%

Table 3.2: *Tower classification for the γ-Jet Run Ib data sample. A cut on the photon energy ($P_T^\gamma \geq 40$ GeV) was applied.*

From this table we can notice as, in average, for each event about 60% of the towers are "track towers" carrying only $\sim 30\%$ of the jet energy. This difference can be attributed to shower leakage outside the target tower being more than one of such towers associated with each track.

The difference between total CEM ($\sim$ 12%) and total HAD ($\sim$ 2.5%) energy in the "gamma towers" gives an indication that the classification is well working for this kind of towers.

For "mixed towers" it is very important to note the difference between the tower fraction ($\sim 18\%$) and the jet energy fraction ($\sim 58\%$). This means that

overlap problems are present in few towers but the corresponding amount of energy is very relevant. These towers are mainly due to the overlap of energetic particle in the core of the jet with a higher contribution of neutral hadrons these particles not being deflected by the solenoid magnetic field.

Finally, we can see that few towers are classified like "not assigned" ($\sim$ 15%) carrying only a small energy amount ($\sim$ 3.5% of the total). Even in this case the results are in agreement with expectations so confirming the validity of the classification approach.

3.6 The New Definition of Tower Energy

The classification process labels each tower according to the particle kind hitting it. The tower energy is then re-assigned as follows:

Track towers

The calorimeter energy is replaced by the sum of the momenta of the tracks hitting the tower.

Gamma towers

We only consider the CEM energy rescaled by a correction factor $\alpha^{em} = 1.049$ accounting for absolute CEM energy calibration in Run Ib data (which we consider in the present analysis)[7].

Mixed towers

A more complex method to combine CTC, CEM and HAD informations is adopted. In fact, some double counting problems can arise as charged hadrons can release part of their energy in the CEM compartment. So we have to subtract from the CEM energy (mainly due to photons) the expected contribution coming from these charged particles.

The CEM energy released by charged particles in a given tower $\langle CEM \rangle_{trk}$ is parametrized as:

$$\langle CEM \rangle_{trk} = \sum_{Target} \alpha^0 \cdot P_T^{trk} + \sum_{Left} \alpha^+ \cdot P_T^{trk} + \sum_{Right} \alpha^- \cdot P_T^{trk} \qquad (3.1)$$

where the contribution from all the tracks impinging in the 3×1 window around the tower is taken into account.

[7] This CEM scale correction is taken into account in the standard JTC96 jet corrections rescaling the jet electromagnetic energy fraction.

The total energy carried by electromagnetic particles hitting a single mixed tower is then expressed as:

$$E_{Pho} = \alpha^{em} \cdot CEM_{mix} - \langle CEM \rangle_{trk} \tag{3.2}$$

As seen, "mixed towers" are mostly in the jet core so we also expect some neutral hadronic particles contribution, as the neutrals are not deflected by the magnetic field. A further correction term is so introduced:

$$E_{neu} = \alpha^{neu} \cdot HAD_{mix} \tag{3.3}$$

In such a way we "recover" part of the hadronic compartment energy as due to neutral hadrons.

In conclusion the energy of a "mix tower" is defined as

$$E_{MIX} = E_{trk} + E_{pho} + E_{neu}.$$

In detail:

$$E_{MIX} = \sum_{Target} P_T^{trk} + (\alpha^{em} \cdot CEM_{mix} - \langle CEM \rangle_{trk}) + \alpha^{neu} \cdot HAD_{mix} \tag{3.4}$$

Not assigned towers

The CEM and HAD energy sum is considered because we expect that the energy released in these towers is coming from neutral hadronic particles and/or photons. This tower energy needs also to be rescaled by an absolute energy scale correction accounting for calorimeter non-linearities.

Now a new energy is assigned to each kind of tower using both tracking and calorimeter informations which will give an improved jet energy resolution as we will show in the next chapter.

We can see how this method needs a suitable choice of the unknown five parameters used in the above expressions. We have also to note how most of such parameters occur just in mixed towers which (see table 3.2) in average carry more than half of the jet energy. Their determination is crucial to accomplish a reasonable well working algorithm.

Table 3.3 summarizes the new tower energy definition.

Tower Type	*Energy tower definition*
Tracks	$\sum_{Target} P_T^{trk}$
Gamma	$\alpha^{em} \cdot CEM$
Mixed	$\sum_{Target} P_T^{trk} + (\alpha^{em} \cdot CEM_{mix} - \langle CEM \rangle_{trk}) + \alpha^{neu} \cdot HAD_{mix}$ $\langle CEM \rangle_{trk} = \sum_{Target} \alpha^0 \cdot P_T^{trk} + \sum_{Left} \alpha^+ \cdot P_T^{trk} + \sum_{Right} \alpha^- \cdot P_T^{trk}$
Not Assigned	$\alpha^{na} \cdot (CEM + HAD)$

Table 3.3: *The new tower energy definition.*

3.6.1 Parameter Selection

To be as close to the detector response as possible, we decide to use a suitable data sample to get the unknown parameters. The best candidate is the γ-jet sample where the photon P_T, well reconstructed by the electromagnetic calorimeter (with resolution $\sim 13.5\%/\sqrt{E_T}$), gives a good estimate of the transverse momentum of the parton (originating the jet) balancing it in the transverse plane. The selection criteria for this sample will be detailed in next chapter. From $P_T^{parton} \sim P_T^{\gamma}$, the photon-jet balancing

$$f_b = \frac{P_T^{jet} - P_T^{\gamma}}{P_T^{\gamma}} \tag{3.5}$$

gives a direct parton-jet comparison.

The first direct parameter determination was the α^{na} value. We supposed this number to be equal to the overall multiplicative factor needed to rescale the raw calorimeter jet energy to the photon energy scale. This factor was so derived from fig. 3.10 such that the mean value of the gaussian fit of the f_b distribution was centered to zero. To only account for calorimeter non-linearities, central jets far from η-cracks ($0.1 \leq |\eta^J| \leq 0.7$) and no secondary jet activity were required. The result was:

$$\alpha^{na} = 1.28$$

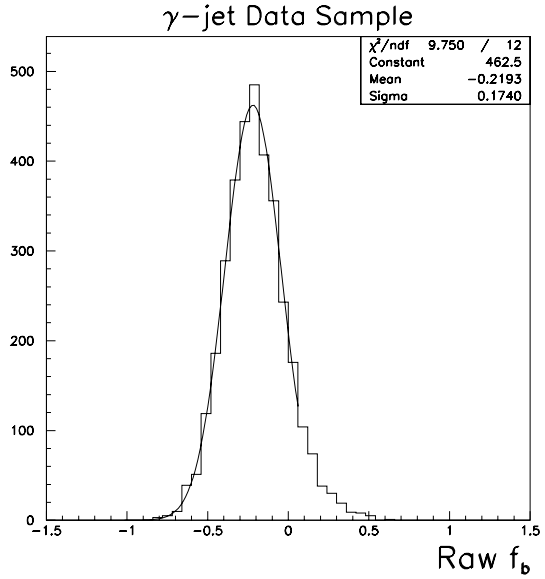

Figure 3.10: *Photon-Jet balancing in real data using the raw jet P_T. The calorimeter response non-linearity is the main source of the imbalance as both photon and jet were chosen in central calorimeters outside crack zones and no secondary jet activity was also required. From the mean value of the gaussian fit for the peak is possible to extract an estimate of the scale correction factor.*

Naturally, simply using this scale factor for all kind of tower would result in no substantial jet resolution improvement. As shown in section 3.2, we have in fact to use as much tracking information as possible.

To get the other four parameters we decided to reconstruct the jet P_T on a jet-averaged basis relying on the towers list belonging to the jet cluster previously found by JETCLU. The jet tower energy corrections (summarized in table 3.3) were made leaving the parameters free. For symmetry $\alpha^+ = \alpha^-$ was imposed while a track P_T dependent α^0 parameter was searched in the form:

$$\alpha^0(P_T^{trk}) = a + b \cdot P_T^{trk} \tag{3.6}$$

It is in fact well known (see for instance ref. [44]) that the response of the CEM detector to the single charged hadron is dependent on its P_T.

Then a minimization code (MINUIT) was used to obtain the best parameters by minimizing the photon-jet balancing (eq. 3.5).

A preliminary consistence check of the three most important parameters (a,b,$\alpha^{\pm}$) was performed for the track deposition estimate inside the CEM (eq. 3.1). The expected quantity was compared with the observed one in

"track towers". In fact in such towers, by construction, there is no photon, so the difference $\langle CEM \rangle_{trk} - CEM$ is expected to be zero within experimental errors. As shown in fig. 3.11 our CEM deposition estimate (which from the above is very critical for our algorithm) results well valued.

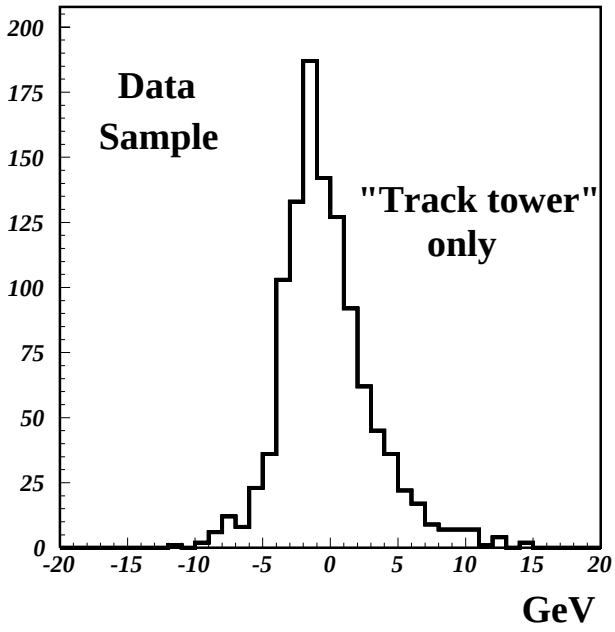

Figure 3.11: *$\langle CEM \rangle_{trk} - CEM$ distribution for "track tower" in the Run Ib γ-Jet data sample. This cross-check shows that the energy released by charged hadrons in the electromagnetic calorimeter can be successfully parametrized using the sum of the momenta of all the tracks impinging in the 3×1 window around a tower.*

Previous analyses got the best results ($\sim$ 30% of jet energy resolution improvement) using a set of parameters which is dependent on the photon P_T [37, 38, 47]. Five different sets were obtained minimizing f_b for each of the P_T^γ bins in which the γ-jet sample was divided. The origin of this dependence on the jet energy scale can be attributed to the dependence of the P_T spectrum of the tracks on the jet P_T.

The problem of this approach is connected to the impossibility to make the algorithm working on another data sample.

Consequently the first aim of the present work was to find a different way to get the best set of parameters and to develop a stand-alone tool working in principle on each sample. Some attempts were first made with *fixed* parameters NOT depending on the photon P_T which gave a simpler and more general implementation of the algorithm. But the best results

were obtained with P_T^{jet}(raw) dependent parameters [8]. The γ-jet sample was divided into three intervals of the P_T^{jet}(raw) spectrum (<25 GeV, 25-35 GeV, >35 GeV) and three set of parameters (12 in total) were obtained with the usual balancing method. Table 3.4 summarizes the parameter values.

Raw P_T^{jet} *(GeV)*	a	b	$\alpha^{\pm}$	α^{neu}	α^{em}	α^{na}	*Scale factor*
< 25	0.15	0.017	0.15	0.61	1.049	1.28	1.04
25-35	0.45	0.003	0.16	0.74	1.049	1.28	1.05
> 35	0.57	0.002	0.16	0.71	1.049	1.28	1.12

Table 3.4: *Summary of parameter values.*

Results on jet energy resolution obtained using the above parameters are reported in next chapter. The algorithm implementation in a stand-alone offline analysis code and some further potential developments are described in the following.

3.7 A New Jet Correction Code: JCOR2K

In its conceptual design the classification algorithm (usually referred as the **JCL98** Algorithm from the name of its first preliminary implementation in an offline code) classifies and corrects the energy of each calorimeter tower. The energy information of the tower bank (TOWE) is then redefined and finally the standard CDF cone clustering algorithm JETCLU (or, in principle, some other else) is applied to reconstruct the jets.

However, as it will be shown in the following, this approach is not worthy to be applied on Run I data.

The "Tracks 00" Problem

Sometimes a track is associated to a calorimeter tower with no energy deposition in both electromagnetic and hadronic compartments (zero CEM energy and zero HAD energy, hence "Track 00"). This anomaly is mainly due to the presence of cracks, to stoppage into the coil and to PMTs gain fluctuation and is known since early analyses (see for example ref. [44]). This particular kind of tower is classified as "track" by our algorithm but it is not inside the TOWE bank (which only include energetic towers above 0.1 GeV). So,

[8] Notice that also the standard JTC96 corrections are P_T^{jet}(raw) dependent.

even if inside the jet cone, these towers are not present in the jet tower list previously provided by JETCLU. We will refer to such towers as *"00 towers"*.

Fig. 3.12 shows the P_T distribution of these tracks for a γ-jet data sample. In fig. 3.13 (left) the sum of these track P_T's into a cone 1.0 jet is plotted, and in fig. 3.13 (right) the same quantity is divided by the jet raw P_T. Events in the $25 < P_T^{\gamma} < 30$ GeV range have been selected. The fraction of energy recovered is $\sim 5\%$ of the jet raw energy. This fraction is shown in fig. 3.14 as a function of the jet P_T. Our results are consistent with previous CDF studies [44].

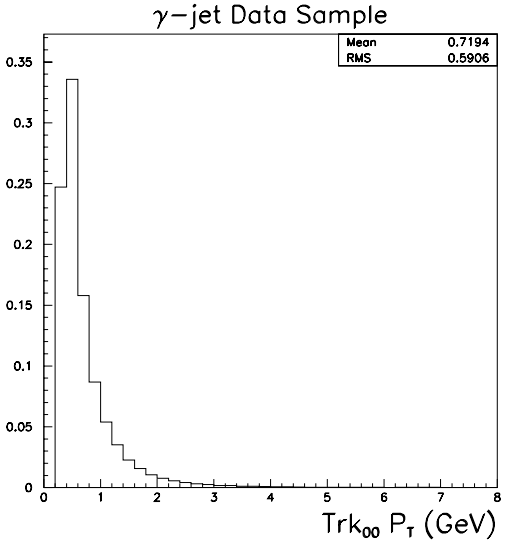

Figure 3.12: *P_T distribution of tracks pointing to "00 towers" (zero in both electromagnetic and hadronic compartments). The distribution is normalized to 1.*

When reassigning the tower energy it is not possible to include the "00 towers" to the tower list used by JETCLU in reconstructing the jets. This *track* information is lost as it is not technically possible, in the Run I offline FORTRAN architecture, to easily change the TOWE bank length adding new towers [9]. So using the JCL98+JETCLU approach we lost in average about 5% of the jet raw P_T (see figures 3.13 and 3.14). To recover the jet energy scale for this "tracks 00" effect we need a scale factor > 1.05 which automatically results in loosing more than 5% in jet energy resolution improvement [10].

[9] The new C++ environment for Run II will easily allow such a job.

[10] If a gaussian distribution, like the jet P_T distribution from a parton of given P_T, is rescaled by a generic factor, also its σ results multiplied by such factor.

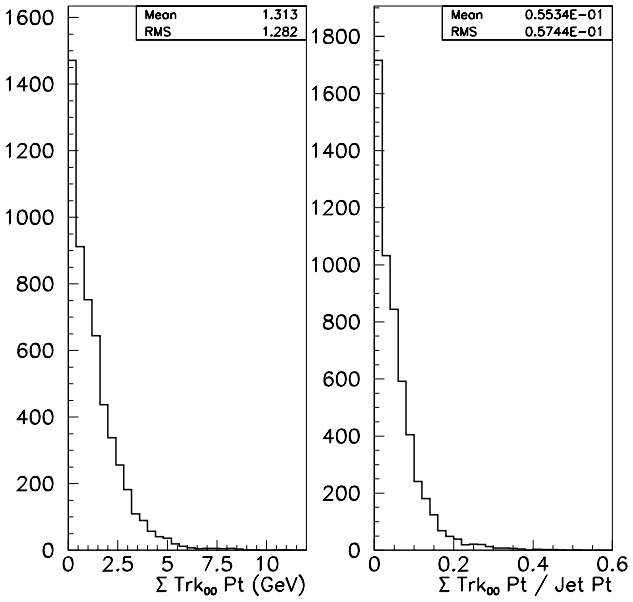

Figure 3.13: *ΣP_T of tracks in a cone 1.0 around the jet axis pointing to "00 towers" (left), divided by the jet raw P_T (right).*

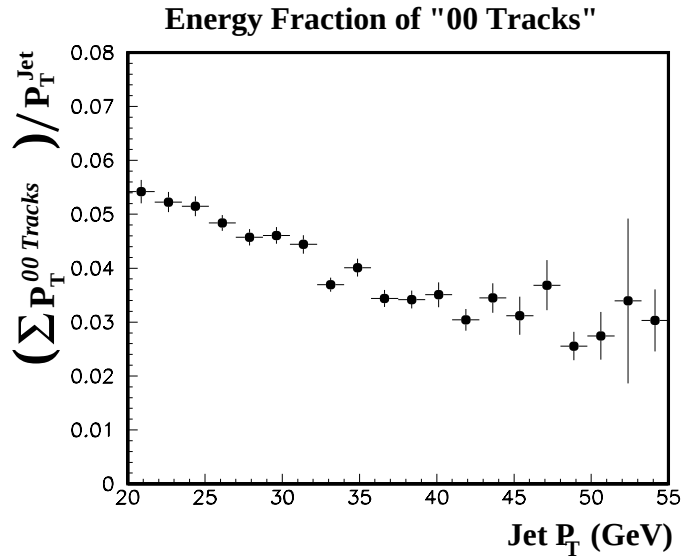

Figure 3.14: *The fraction of the jet P_T recovered thanks to the addition of the P_T of the tracks pointing to zero-energy towers is shown as a function of the jet P_T.*

On the other hand, waiting for the new Run II data, we want to implement the algorithm in a Run I offline code which can easily be converted into a C++ Run II offline code and which allows to test it on different data samples.

The best idea to overcome the above problem is to consider an alternative approach respect to JETCLU to reconstruct the jet energy.

The code used in the previous analyses referred to the JCL98 module for what concerns the tower classification but presented some differences with respect to JETCLU regarding the jet energy reconstruction and correction. Here this correction was made on a jet-averaged basis relying on the tower list belonging to the jet cluster previously found by JETCLU (this is the same jet energy reconstruction used to get the best classification parameters). So the jet centroid was not recalculated, the jet energy correction resulting only in a correction factor.

3.7.1 JCOR2K

The "hybrid" approach used in the past has the problem of reconstructing a jet P_T (making a *scalar* sum over the tower energies) which actually is more close to the JETCLU jet E_T (see Chapter 2). This fact did not allow a perfect "apple-with-apple" comparison between the old and new correction algorithms having also a not negligible effect when comparing the absolute energy scale [48].

In the present study we pay particular attention to the different definition of P_T and E_T of a jet, so to use the same variable for the JETCLU+JTC96 versus the new algorithm comparison. In the past, as only the improvement on the jet energy resolution was studied, the difference between the use of the jet E_T instead of P_T was negligible. But it makes a difference in the study of the energy scale for instance in terms of photon-jet balancing.

To overcome all the above problems a new offline code (**JCOR2K**) has been developed with the following characteristics:

- Tower classification: identical to JCL98.
- Tower correction: as in table 3.3 with jet *raw* P_T dependent parameters (three jet raw P_T bins were considered, see table 3.4).
- Towers 00: included in the jet tower list if inside the R = 1.0 cone around the jet axis.
- Jet 4-momentum: reconstructed with a JETCLU simulation.
- OOC correction: same as JTC96 (R =1.0 cone).

- Absolute energy scale: jet P_T rescaling factor for each of the three jet raw P_T bins.

The rescaling factors (see table 3.4) were derived from the γ-jet sample by centering to zero the photon-jet P_T balancing. The out-of-cone corrections, due only to physics effects, are the same as JTC96.

Some caveats:

1. The jet tower list is the same of that previously found by JETCLU as its simulation is only used to reconstruct the jet 4-momentum. "00 towers" are added to this list.

2. The JETCLU simulation resulted in a reconstructed jet raw $P_T \sim 1\%$ lower than using the real JETCLU.

This code, even if optimized for the γ-jet sample, can run on a generic data sample reproducing the same results of the Classification+JETCLU ideal corrections when the "00 towers" are included.

3.8 Future Algorithm Developments

Results on data will be presented in next chapter, showing a clear improvement obtained by the new algorithm for two different data samples. Anyway these results are far to be the best achievable.

The tower classification and correction can in principle be improved using more detector informations such as Crack Chamber (CCR), Central Preradiator (CPR) and tower PMTs balancing (see Chapter 1) and further Monte Carlo studies. Some hints are shown in the following.

As shown in a preliminary study, the CCR information can be used to distinguish photon falling into ϕ-cracks also recovering their energy lost in such not instrumented zone [49].

The CPR, originally designed to improve the π^0/γ separation, actually gives a sampling in the $r-\phi$ view of the early development of electromagnetic showers in the material of the solenoid coil. So, it can in principle indicate the presence of a photon not detected by the CES giving a better tower classification [11].

The combined η and ϕ informations on charged tracks and on CEM and CHA clusters position can be used to decide if lateral energy leakage from

[11] The CES clustering algorithm is fully efficient only for photons with $E^\gamma \gtrsim 1$ GeV.

a tower to the η or ϕ-neighboring one is consistent or not [12]. In principle this can give a better separation among overlapping particles inside the same tower and also a better neutrals (K_L,n..) contribution estimate to the tower energy.

Test-beam studies have shown that a measure of the ϕ position of both CEM and CHA showers can be extracted from the logarithm of the signal ratio for the two photomultipliers ("Left" and "Right") reading the opposite ϕ sides of both electromagnetic and hadronic compartments of a tower. The collected light was in fact found to drop exponentially with the distance between the shower and the photomultipliers [13, 46]. Fig. 3.15 shows log(L/R) for the hadronic section of a central tower (CHA) as function of x, where x is the distance (along the ϕ direction) of a 50 GeV pion beam from the tower center considered at the Shower Maximum [13]. Similar results (with a different linear fitting function) were found studying the CEM response maps to a 50 GeV electron beam [46].

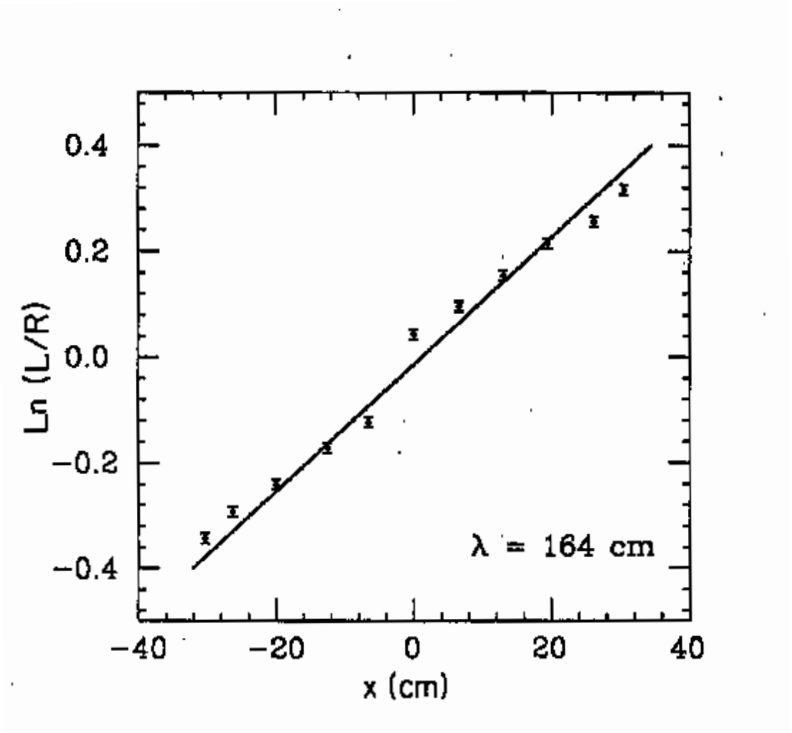

Figure 3.15: *The logarithm of the ratio of the left (L) and right (R) PM pulse heights in a CHA tower, as a function of the distance of the pion beam from the tower center. An equivalent attenuation length (λ) of 164 cm was found.*

[12]The CES wires ϕ information was not considered till now because of the ambiguity arising from the fact that the ϕ coordinate CES readout is related to five towers inside the same wedge [14].

A More Refined Approach

From several studies (see for example [39, 44, 50]) it is known that the charged hadron CEM energy deposition is a function of its P_T and impact point on the tower surface. So in principle we can get the α^0 and $\alpha^{\pm}$ algorithm parameters (that is the most important ones) as a function of P_T and the relative η and ϕ tower coordinates, directly from data combining tracking and calorimetry information. The minimum bias data sample is the best candidate as it contains well isolated charged tracks. The advantage of this approach is that the parameters obtained in such a way are not sensitive to the particular data sample used [13] and are always related, and so automatically calibrated, to the calorimeter response.

A first preliminary dedicated study (still in progress) has been performed on isolated charged tracks on the γ-jet sample [51].

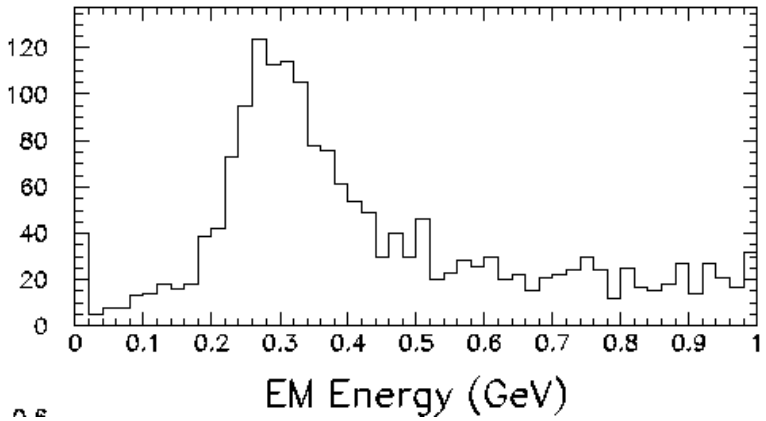

Figure 3.16: *CEM energy deposition for central isolated charged tracks with $P > 2$ GeV/c in the γ-jet sample [51]. The MIP peak around 0.3 GeV is evident.*

With this approach a better resolution can also be in principle achieved using the CPR and CES information to distinguish between a MIP charged hadron and a non-MIP one inside the solenoid coil and the CEM respectively. Reference [44], for instance, reports a $\sim 20\%$ and $\sim 50\%$ probability for an inelastic nuclear interaction respectively inside the superconducting coil and the CEM for low-energy pions. This occurrence can be so evidenced by the CPR and/or by a suitable CES clustering algorithm in a window around the extrapolated direction of the track.

[13] As it will be clear from the results of the next chapter, the present parameter selection is biased by the energy spectrum of the data sample used.

The importance of distinguishing between MIP and non-MIP charged hadrons is shown in fig. 3.16, were the CEM MIP peak (around 0.3 GeV) for isolated tracks is well separated by the remaining flat energy deposition due to tracks starting their shower before the CHA. It is important to remember that the α^0 parameter is presently related with the mean value of such a distribution for the whole track P_T range.

Plug Region Extension

The algorithm has a strict geometrical limitation as it can only correct central jets because tracking information in Run I was only available for central ($|\eta| \lesssim 1$) charged particles. Incidentally we have to remember that the jet energy resolution improvement obtained with the new algorithm is expected to be reproduced in the Run II setup as the new COT tracking chamber has the same geometrical and resolution characteristics of the old CTC.

However, in Run II this limitation can be overcome as the combined information from the L00+SVXII+ISL silicon system will provide an independent tracking up to $|\eta| \leq 2$ (see fig. 3.17) with an expected track momentum resolution $\frac{\delta P_T}{P_T} \sim 0.4\% \cdot P_T$. Moreover the new plug electromagnetic calorimeter is provided with a Shower Max detector (see Chapter 1). So the new algorithm can be easily extended to this wider pseudorapidity range.

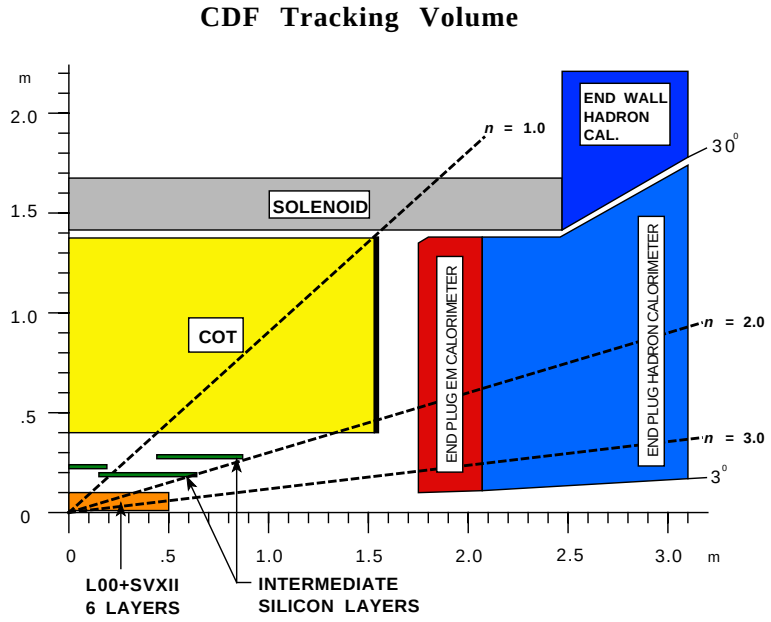

Figure 3.17: *The CDF II tracking volume. The extension to the* $1 \leq |\eta| \leq 2$ *region is shown.*

Chapter 4

Testing the New Algorithm

In order to check the new jet correction algorithm, the Run Ib γ-jet data sample was reanalized with respect to a previous analysis with optimized selection criteria which allow to reduce the background while retaining enough statistics. A better agreement between data and Monte Carlo simulation has also been achieved. The JCOR2K stand-alone offline code (which implements the algorithm using a set of parameters tuned on the same γ-jet sample) was used giving an improvement on the jet energy resolution of $\sim$ 25% compared to the standard CDF jet corrections.

A further test was performed on a different Run Ib data sample, the di-jet sample, using the same code. Results show that the new algorithm provides a clear improvement in both jet energy resolution ($\sim$ 15%) and $\vec{\not\!\!E}_T$ resolution. These results represent the first check that the algorithm can in principle work on any data sample.

4.1 The γ-jet Sample

Direct photon data is the ideal sample to compare different jet energy reconstructions in the photon-jet P_T balancing, as the photon (well measured by the electromagnetic calorimeter) provides a good energy estimate of the parton originating the jet. Fig. 4.1 shows the CDF event display of a typical γ-jet event where a well isolated photon is balancing the jet in the transverse plane. The leading order diagrams for photon production in $p\bar{p}$ collisions are shown in fig. 4.2. The two upper (Compton) diagrams dominate at CDF [14].

A set of appropriate cuts is needed to reduce the incidence of other physical processes which mimic the direct photon production. This background is mainly coming from di-jet events with one jet mostly consisting of an isolated neutral meson decaying into photons (π^0, η,...). As it will be shown in next

chapter, after most of the di-jet background is rejected by a strict photon isolation requirement, also cosmics background can play some role.

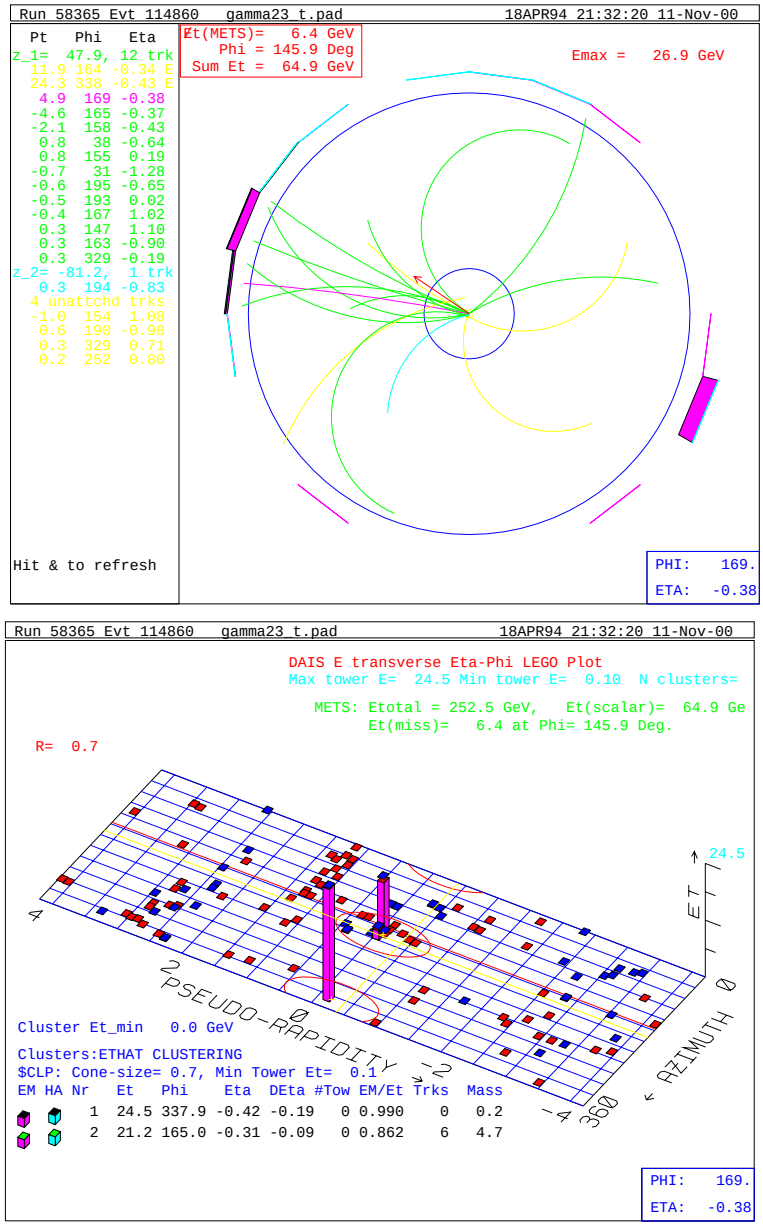

Figure 4.1: *Typical γ-jet event. A well isolated photon is balancing the jet in the transverse plane.*

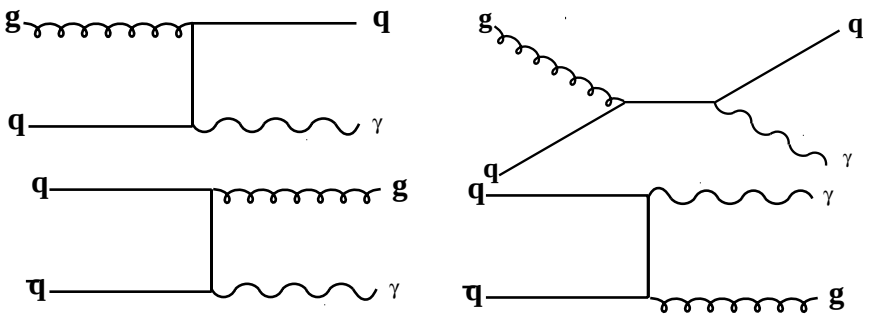

Figure 4.2: *Leading order diagrams for photon production in $p\bar{p}$ collisions: Compton QCD diagrams (upper); annihilation diagrams (lower).*

4.1.1 Previous Studies

A previous analysis performed a first algorithm check on the γ-jet sample [37, 38]. To compare the new jet energy reconstruction method with the standard CDF one, JETCLU + JTC96 jet corrections, a technique developed by the UA2 collaboration was adopted [52, 53]. In order to extract the detector effect contributions to the jet energy resolution, the imbalance vector $\vec{P}_T = \vec{P}_T^{jet} + \vec{P}_T^{\gamma}$ was projected onto the azimuthal angular bisectors of the photon-jet system (see fig. 4.3). The two components $P_{T\xi}$ and $P_{T\eta}$ are sensitive to different effects. The calorimeter energy resolution is the main source of the $P_{T\xi}$ component, while gluon radiation (physics) effects are common to both components. Both of them are also subject to other small effects such as angular measurement errors, underlying event and out-of-cone fluctuations.

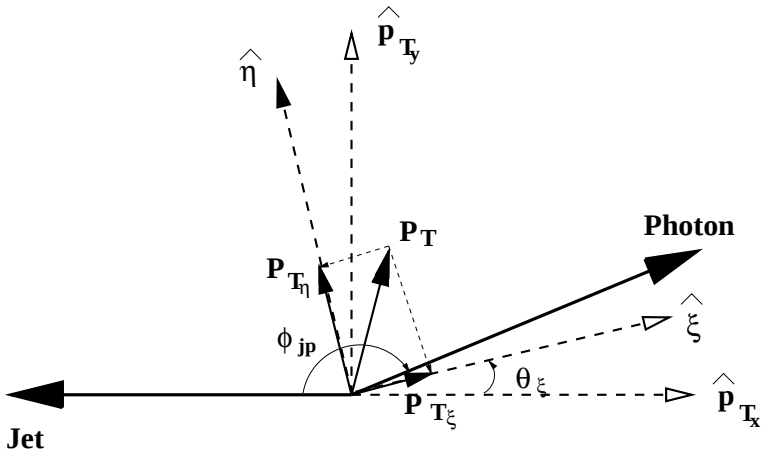

Figure 4.3: *The resulting transverse momentum $\vec{P}_T$ of the γ-jet system has been decomposed into the two components $P_{T\eta}$ and $P_{T\xi}$, which are sensitive to different effects responsible for generating the $\vec{P}_T$ vector.*

After hard gluon emission is reduced by applying a cut on the second jet energy, the soft contribution can be removed by subtracting in quadrature σ_η from σ_ξ (that is the σ of the gaussian fit for the $P_{T\eta}$ and $P_{T\xi}$ distributions). In this way we subtract the contribution due to finite angle resolution (σ_η) from the jet energy resolution σ_ξ defining an *effective* jet calorimeter resolution as:

$$\sigma_D = \sqrt{\sigma_\xi^2 - \sigma_\eta^2} \tag{4.1}$$

Fig. 4.4 shows the central detector resolution σ_D for both methods as a function of P_T^γ as obtained by the previous analysis with the selection criteria of ref. [37]. An improvement is noticed using the new algorithm for all the photon energies. The central detector resolution was evaluated as $\sigma/P_T \sim 109\% \ / \ \sqrt{P_T}$ using the JTC96 corrections and $\sigma/P_T \sim 78\% \ / \ \sqrt{P_T}$ with the classification method.

These results were obtained using a P_T^γ dependent set of parameters and reconstructing the jet P_T on a "jet averaged" basis. Only the jet resolutions were compared, without paying too much attention to the absolute energy scale.

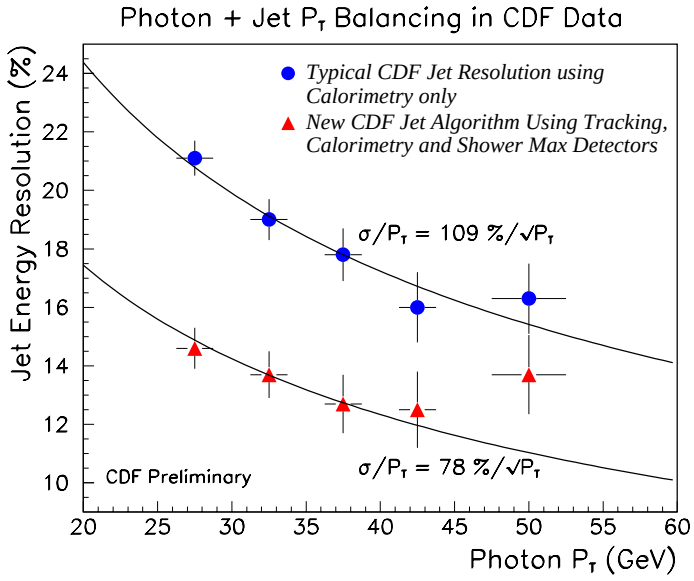

Figure 4.4: *The central detector resolution σ_D is plotted as a function of P_T^γ for the two methods. This plot was obtained by the previous analysis described in ref. [37].*

4.1.2 Data-Monte Carlo Comparison

Since the above analysis we undertook a data-simulation [1] study in order to optimize the set of parameters necessary in the new algorithm. The not-so-good agreement between the real and the simulated data called for a better reduction of the background in the photon data sample. So a new analysis of the Run Ib photon sample was performed with new selection criteria to reduce the background (mainly the photon isolation cut), while releasing other cuts in order to retain enough statistics. The new stand-alone offline code JCOR2K (see section 3.7) was used to apply the classification algorithm. We will show that the photon-jet balancing of this new sample provides a better jet energy resolution in the central calorimeter relative to the previous analysis, for both standard JETCLU+JTC96 corrections and the new algorithm. Furthermore, the new jet algorithm shows an improvement on the jet energy resolution of $\sim 25\%$ compared to JTC96 corrections.

4.1.3 The Photon Sample Selection

We used the standard photon selection on the Run Ib photon sample, with events passing the Level 3 inclusive photon trigger with $P_T > 23$ GeV and an isolation cut on the energy within a 0.4 cone less than 4 GeV. Photon selection criteria are in PHO94.CDF. Events with more than one class-12 vertex have been rejected. A 1.0 jet cone and a cut on extra jet activity with $P_T > 8$ GeV were adopted as found in previous studies to be optimal for the reconstruction of isolated light jets [36]. So, in order also to select a good jet-photon balancing, we required one and only one jet with raw $P_T > 8$ GeV to avoid hard gluon radiation in the event. As is ref. [37] we considered only central jets with $|\eta| < 0.7$. No special corrections have been applied to the towers near the $\eta = 0$ crack. For this reason, cuts on $|\eta_\gamma| > 0.1$ and $|\eta_{jet}| > 0.1$ have been applied.

At this point there are 18,609 events in the $25 < P_T^\gamma < 55$ GeV range. We investigated several different options to further reduce the background in the photon sample, improve the data-QFL agreement and still retain enough statistics to allow a study of the energy resolution as a function of the photon P_T. The final selection criteria which differ from ref. [37] are:

1. Tighten the photon isolation cut in a 0.4 cone around the photon from 4 GeV to 1 GeV, as already used for the Run Ib prompt photon cross section [54].

[1] γ-jet Pythia Monte Carlo sample + a full CDF detector simulation (**QFL**) .

2. Release the cut on J1-clnc, the energy clusterized in the leading jet which was deposited outside the central calorimeter, from 0.5 GeV to 3 GeV.

3. Apply the cut on the charged fraction Chf $= \Sigma P_T^{tr}/P_T^J(raw) > 0.2$ suggested by ref. [55] (also described in next chapter) to suppress residual cosmics background and events with low energy jets whose underestimate of the transverse energy is due to out-of-cone leakage.

The last cut reduces the sample by only ~2.3%. The distribution of the energy inside a 0.4 cone around the photon and of the out-of-CC energy for the leading jet for the remaining 18,185 events are shown in figg. 4.5 and 4.6, respectively.

The second cut was applied in ref. [37] in order to select super-central jets. However, with a cluster cone with radius equal to 1.0, about half of the events were rejected. We don't expect that this cut discriminate between signal and background events, as it can be seen in fig. 4.7 (left) where we compare the distribution of the out-of-CC energy from a signal sample (photon isolation cut in 0.4 cone < 1 GeV) with a control sample with a photon isolation cut > 3 GeV, for events in the $25 < P_T^\gamma < 35$ GeV bin. We decide to set the J1-clnc cut to < 3 GeV, rather than completely eliminate it, to cut the tail of events with too much deposit outside the central calorimeter.

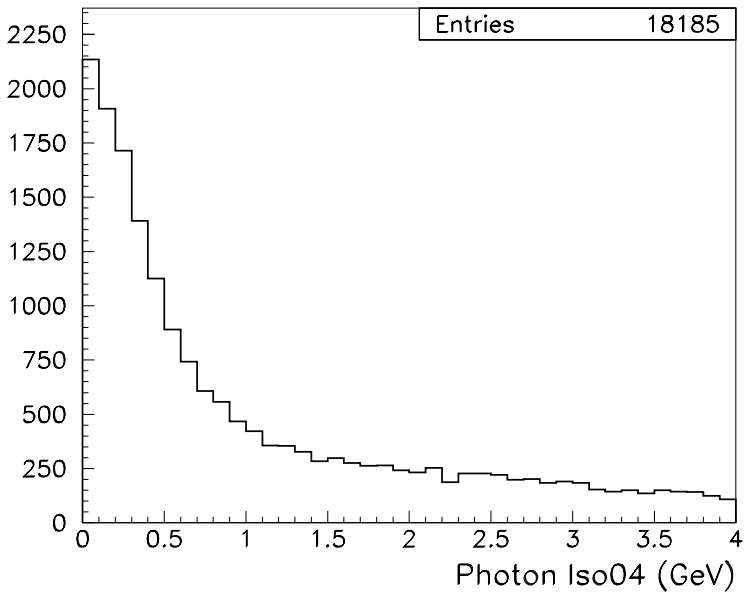

Figure 4.5: *Distribution of the energy inside a 0.4 cone around the photon.*

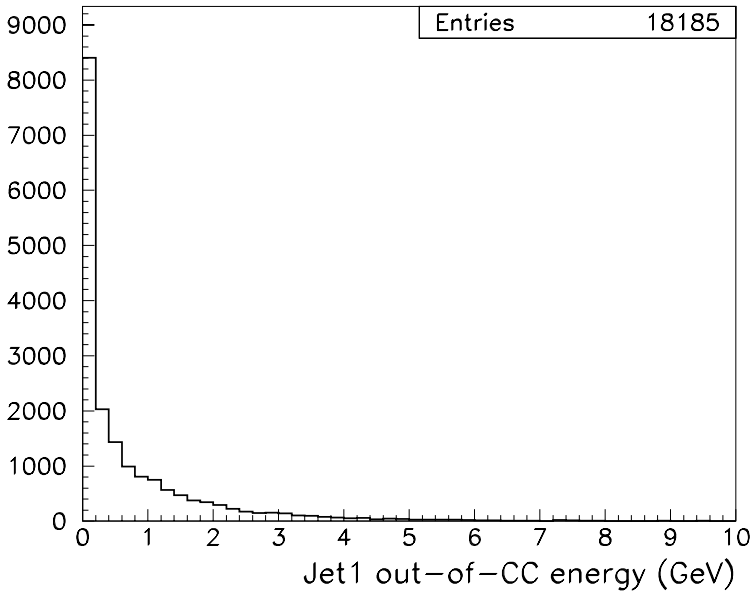

Figure 4.6: *Distribution of the out-of-Central Calorimeter energy for the leading jet.*

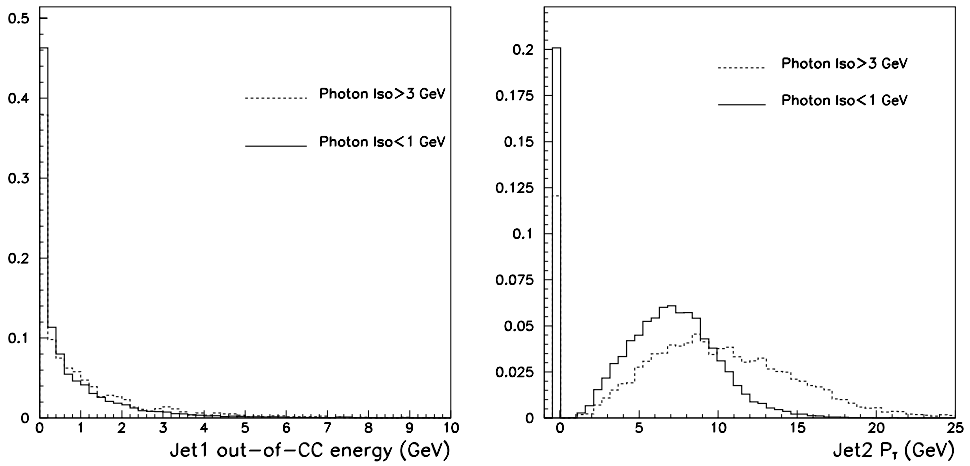

Figure 4.7: *Energy of the leading jet deposited outside the central calorimeter (left) and transverse energy of the second jet (right) for a sample with photon isolation < 1 GeV (solid) and > 3 GeV (dashed). Events are selected in the $25 < P_T^\gamma < 35$ GeV bin and the distributions are normalized to 1.*

Tab. 4.1 summarizes the number of events and the detector resolution σ_D for both JTC96 and classification corrections, for different combinations of the three cuts. The table refers to events in the $25 < P_T^\gamma < 35$ GeV bin. The tightening of the photon isolation cut is the one that noticeably improves the resolution. Fig. 4.7 (right) shows also the difference in extra jet activity (in this case the second jet P_T) for the two samples with photon isolation < 1 GeV and > 3 GeV, respectively.

Cuts	*N.events*	$\sigma_D^{Class}(GeV)$	$\sigma_D^{JTC96}(GeV)$
PhoIso04<4 GeV	14,575	4.5	5.6
PhoIso04<4 GeV, J1-clnc<0.5 GeV	9,056	4.5	5.8
PhoIso04<1 GeV	9,053	3.4	4.6
PhoIso04<1 GeV, J1-clnc<0.5 GeV	5,792	3.4	4.7
PhoIso04<1 GeV, J1-clnc<3 GeV	8,682	3.3	4.6
PhoIso04<1 GeV, J1-clnc<3 GeV, Chf>0.2	8,482	3.2	4.5

Table 4.1: *Number of events and detector resolution σ_D for both JTC96 and classification corrections, for different combinations of the three cuts: PhoIso04 is the energy in a 0.4 cone around the photon, J1-clnc is the energy of the leading jet deposited outside the CC, and Chf is the charged fraction of the jet energy as measured by the CTC track momentum. Events in the $25 < P_T^\gamma < 35$ GeV bin only are selected. The error on each energy resolution is about ± 0.2 GeV.*

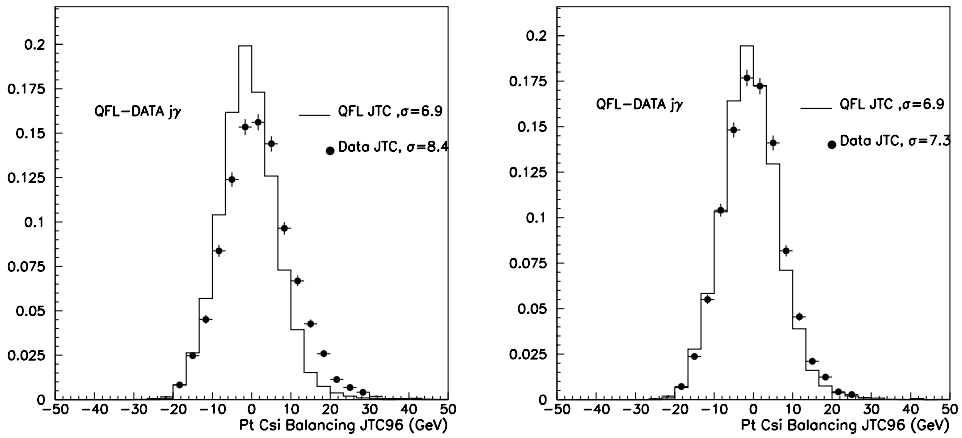

Figure 4.8: *Distribution of the γ-jet balancing $P_{T\xi}$ component for the data (dots) and a Pythia+QFL sample (histo). JTC96 corrections are applied. Events are selected in the $25 < P_T^\gamma < 35$ GeV bin and the distributions are normalized to 1. The new photon selection criteria (right) provide a better agreement between data and MC than the previous study of ref. [37] (left).*

With the new selection criteria we are left with 10,996 events in the $25 < P_T^\gamma < 55$ GeV range, a statistics similar to the previous analysis. However, the better quality of this photon sample can be seen in the comparison between the data and a Pythia+QFL direct photon sample. The bad agreement data-QFL with the previous selection does not depend on the jet algorithm used. Therefore in fig. 4.8 we compare QFL predictions to the data with standard JTC96 corrections, for the $P_{T\xi}$ balancing component and for events in the $25 < P_T^\gamma < 35$ GeV bin. The new selection (right) shows a much better agreement than the previous selection (left).

4.1.4 Jet Energy Resolution

As in ref. [37], the γ-jet sample was divided into 5 GeV bins for the $25 < P_T^\gamma < 55$ GeV range. The last two bins have been merged due to lack of statistics.

The jet energy resolution obtained with the new corrections was compared to the standard JTC96, by projecting the resulting P_T along the azimuthal angular bisectors of the γ-jet system.

The gaussian fit width of the $P_{T\xi}$ component, σ_ξ, is supposed to increase with P_T^γ if it is sensitive to the detector resolution. This is the case, as shown in fig. 4.9 (left), and we notice an improvement by using the new corrections rather than the JTC96.

The σ_η width is supposed to have a flat dependence on P_T^γ, and we don't expect that the classification method improve this component, as it cannot recover the angular resolution due to physics effects, as shown in fig. 4.9 (right).

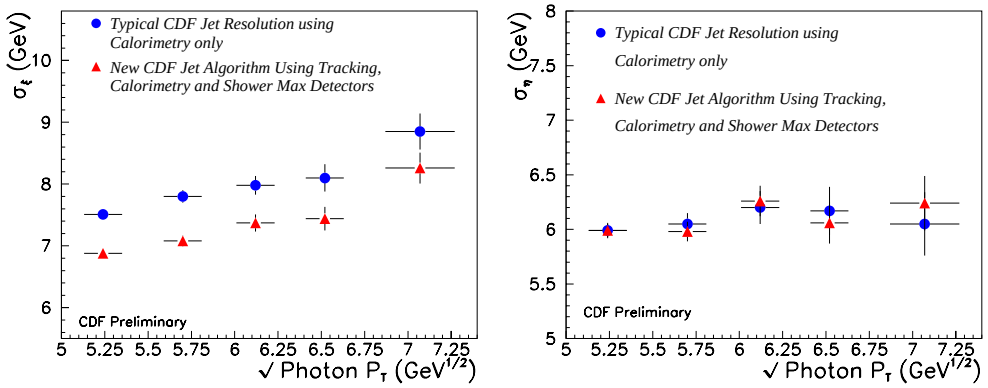

Figure 4.9: *The value of σ_ξ (left) and of σ_η (right) are plotted as a function of P_T^γ for the two methods.*

We defined an effective jet calorimeter resolution as $\sigma_D = \sqrt{\sigma_\xi^2 - \sigma_\eta^2}$, and we plotted it in fig. 4.10 (upper) divided by the central value of each P_T bin. First of all we notice an improvement for both the energy reconstruction methods relative to the previous analysis of ref. [37], thanks to the new photon selection criteria. Furthermore, an improvement of $\sim 23\%$ is noticed using the JCOR2K corrections relative to the JTC96 ones, for all the photon energies. The curves in fig. 4.10 are a best fit to the data, and we evaluated the central detector resolution as $\sigma/P_T \sim 83\% \ / \ \sqrt{P_T}$ using JTC96 corrections, while with the new classification method we reached $\sigma/P_T \sim 64\% \ / \ \sqrt{P_T}$.

The energy scale has been studied by plotting, for both the correction methods, the $(P_T^{jet} - P_T^{\gamma})/P_T^{\gamma}$ balancing offset as a function of P_T^{γ} in the bottom plot of fig 4.10. JCOR2K is well inside the $\sim 5\%$ systematic error quoted for JTC96 [18].

Similar results for the jet energy resolution have been obtained when using the E_T of the jet, as shown in fig. 4.11 (top). However, as expected when a jet mass is generated, the balancing offsets are few percents higher (fig. 4.11, bottom). In this case, we evaluated the central detector resolution as $\sigma/E_T \sim 87\% \ / \ \sqrt{E_T}$ using JTC96 corrections, while with the new classification method we reached $\sigma/E_T \sim 65\% \ / \ \sqrt{E_T}$. An improvement of $\sim 25\%$ is so observed.

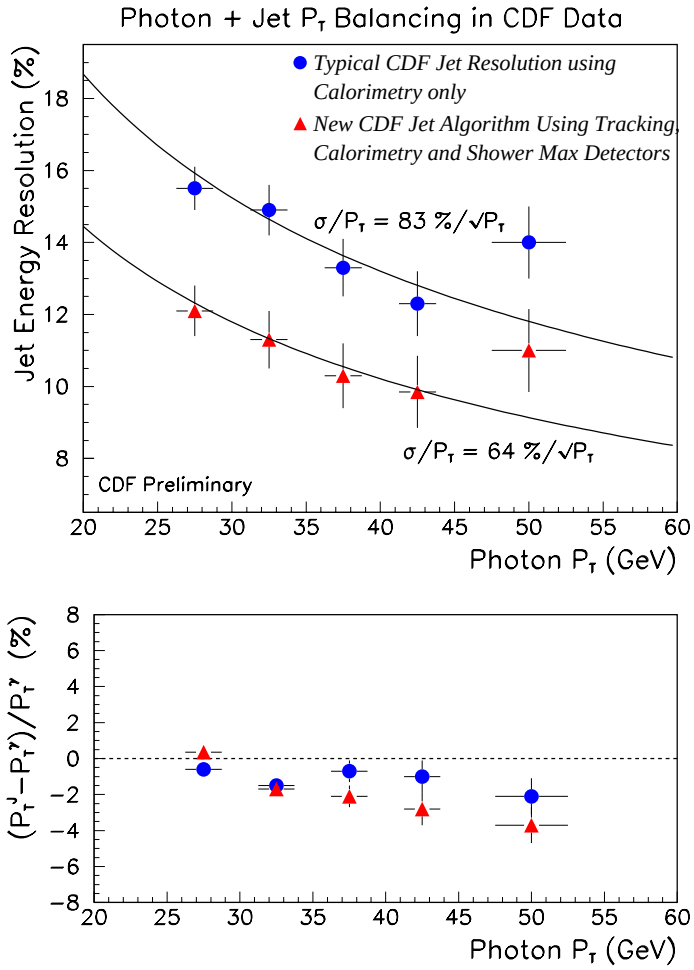

Figure 4.10: *Top: The central detector resolution σ_D is plotted as a function of P_T^{γ} for the two methods. Bottom: The $(P_T^{jet} - P_T^{\gamma})/P_T^{\gamma}$ balancing offset is plotted as a function of P_T^{γ} for the two methods. The P_T of the jets has been used for the photon-jet balancing.*

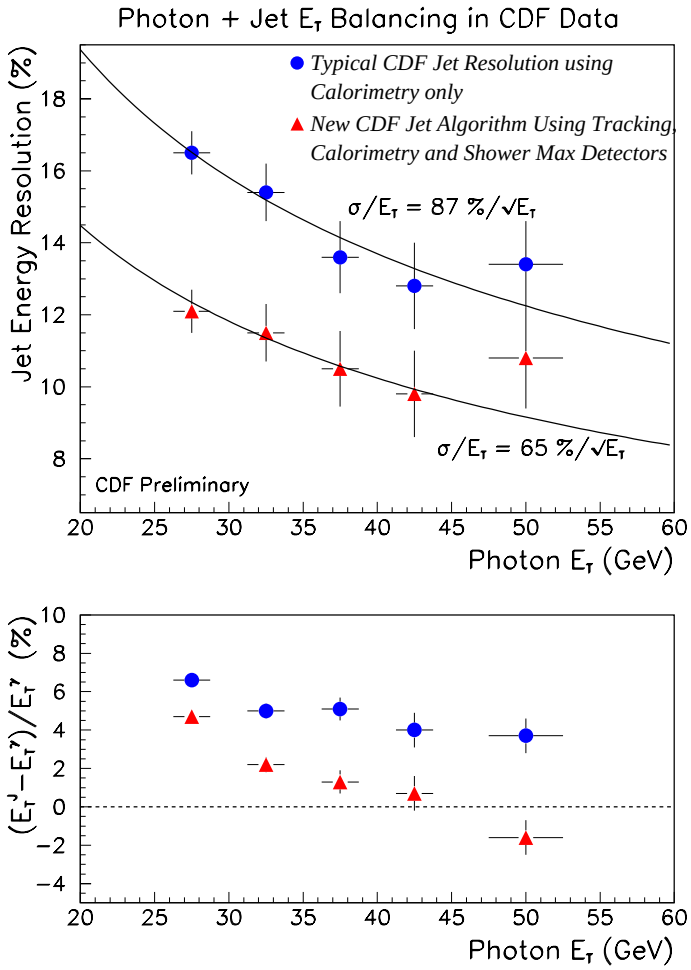

Figure 4.11: *Top: The central detector resolution σ_D is plotted as a function of P_T^γ for the two methods. Bottom: The $(E_T^{jet} - P_T^\gamma)/P_T^\gamma$ balancing offset is plotted as a function of P_T^γ for the two methods. The E_T of the jets has been used for the photon-jet balancing.*

4.1.5 Further Studies

No Vertex Cut

In order to select a clean sample to study the photon-jet P_T balancing, events with more than one class-12 vertex have been rejected in our analysis. Here we want to show what happens when a cut on the number of vertices is completely removed. Due to the other selection cuts however, the statistics increases only of about 30%. Fig. 4.12 shows the resulting resolutions for

the two different corrections, JTC96 and JCOR2K. The energy resolution worsens of some percent for both methods, a bit more for the new one which is more sensitive to the increased particle multiplicity in the calorimeter.

It is worthy to note that also this result will be reproduced in Run II where a lower average vertex number per bunch crossing is expected (see Chapter 1), then resulting in a better resolution improvement.

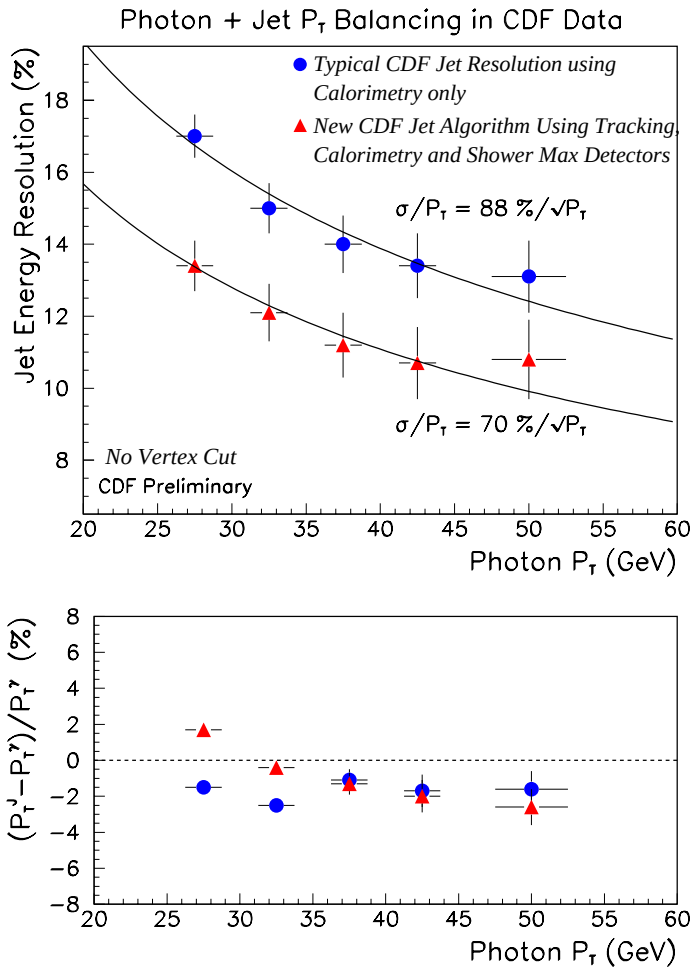

Figure 4.12: *Top: The central detector resolution σ_D is plotted as a function of P_T^γ for the two methods. Bottom: The $(P_T^{jet} - P_T^\gamma)/P_T^\gamma$ balancing offset is plotted as a function of P_T^γ for the two methods. The P_T of the jets has been used for the photon-jet balancing. No vertex cut has been applied.*

Dependence on the Fitting Procedure

The definition of the calorimeter resolution as $\sigma_D = \sqrt{\sigma_\xi^2 - \sigma_\eta^2}$ implies a procedure to extract the values of σ_ξ and σ_η from the photon-jet P_T balancing projected along the bisectors ξ and η.

In fig. 4.13 the two components $P_{T\eta}$ and $P_{T\xi}$ are shown for JTC96 (top) and for JCOR2K (bottom), for events in the $25 < P_T^\gamma < 30$ GeV range. In the present analysis, as well as in the previous one, we assumed a gaussian curve to be a good approximation of the distributions and we took the sigmas of the fitted curves as our estimate of σ_ξ and σ_η. The $P_{T\eta}$ and $P_{T\xi}$ distributions have been fitted over the whole range of the shown plots.

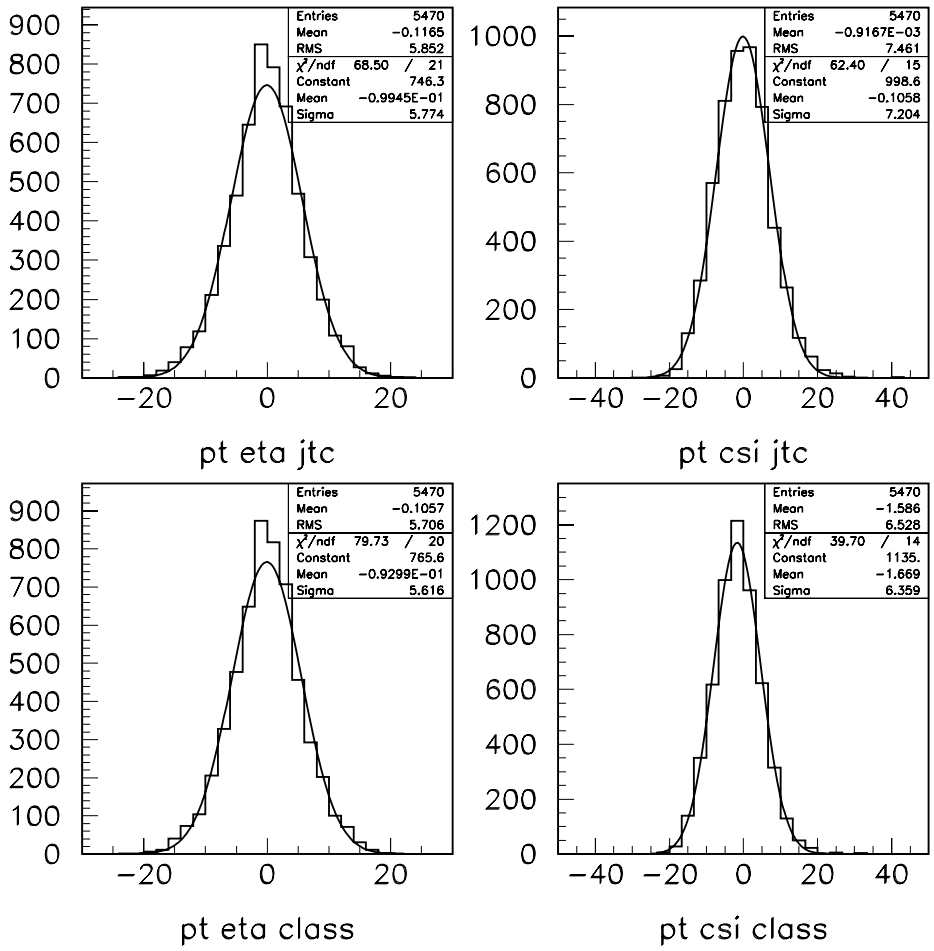

Figure 4.13: *Photon-jet P_T balancing projected along the bisectors η (left) and ξ (right) for standard JTC96 corrections (top) and the classification method (bottom); events in the $25 < P_T^\gamma < 30$ GeV range only are shown in the plots.*

We studied the dependence of the energy resolution on the histogram binning and on different fitting methods, like for instance a gaussian fit to the core of the distributions, without the tails. We found a difference of few percent only, much smaller than the statistical error on each point, and of the same amount and in the same direction for both JTC96 and JCOR2K.

However, in the extreme case where we consider the RMS of the distributions as our sigma values, bigger point-to-point fluctuations are observed (fig. 4.14) resulting in worse overall energy resolutions. For the time being no systematic error coming from the fitting procedure has been added to the energy resolution plots.

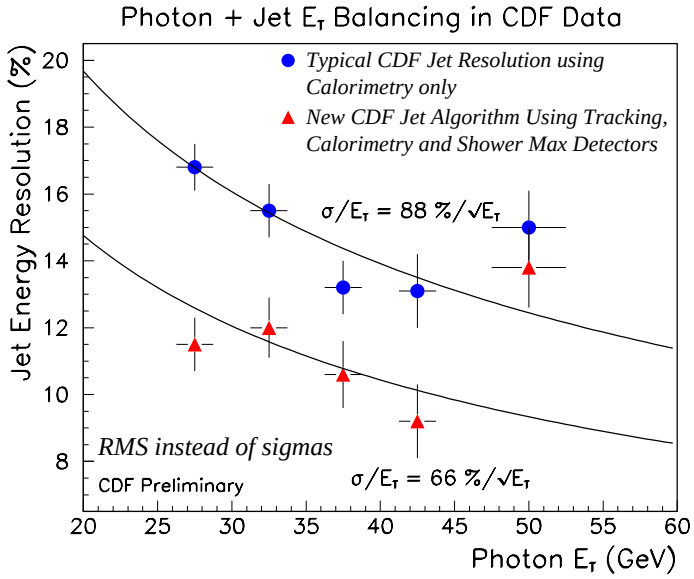

Figure 4.14: *The central detector resolution σ_D is plotted as a function of P_T^γ for the two methods. In this case we considered the RMS of the distributions as our sigma values.*

4.2 The Di-jet Sample

Till now the classification algorithm has been optimized and checked only on the γ-jet data sample from which the parameters were derived. It is very important to demonstrate that our approach could work in a "sample-independent" manner, by testing it on a different data sample using a set of parameters previously derived from the γ-jet one. Furthermore, we understand the need of different data samples with more statistics at higher P_T's to get a better set of parameters which could work well on high P_T jets. The best candidates for this purpose are the di-jet samples (from jet20, jet50, jet70 and jet100 triggers) being the standard ones used to make the JTC96 corrections over the whole P_T range. So further studies are needed on these new data samples to understand how the algorithm is working on them.

We present a first step made in this direction.

Fig. 2.1 shows the CDF event display of a typical di-jet event were two jets are produced balancing each other in the transverse plane.

4.2.1 The Di-jet Sample Selection

Di-jet samples from the jet20, jet50, jet70 and jet100 trigger datasets have been used in the past to perform di-jet balancing studies to get the relative jet corrections used in the JTC96 module [19, 20] and to make studies on relative jet energy resolutions [56].

The present analysis is intended to test the algorithm on a di-jet sample using a set of parameters obtained from the γ-jet sample. Given the relatively low P_T spectrum of the events in this sample, the best check we can do is to consider the jet20 trigger dataset as a starting point.

The Run Ib Jet20 trigger Stream B PADs dataset (QJ2B-5P), stored on tape, was processed with version 7.12 of the offline package producing ntuples similar to those used in the γ-jet analysis (now containing informations on the first three jets of the event). The JCOR2K code was used to apply the classification algorithm.

The following standard cuts were applied in order to reduce the background:

- $|Z_v| < 60$ cm
- $|\vec{\not{E}}_T|/\sqrt{\Sigma E_T} < 6$
- $\Sigma E_T \leq 2$ TeV

Di-jet events were selected with the following additional cuts:

- $N_{jets} \geq 2$
- $N_\gamma = 0$ ($P_T^\gamma > 10$ GeV)
- $0.1 < |\eta^{J1,J2}| < 0.7$
- $d\phi^{J1-J2} > 140^o$
- $P_T^{J1}(raw) > 15$ GeV
- $P_T^{J3}(raw) < 8$ GeV
- $N_{class12}^V = 1$
- $E_{nc}^{J1,J2} < 3$ GeV
- $Chf^{J1,J2} > 0.2$
- 25 GeV $< P_T^{J1}$(JTC96) < 55 GeV

Most of these cuts are the same as in the γ-jet analysis in order to reproduce the same kinematical limits [47]. Similarly we used the cone radius R = $\sqrt{\Delta\eta^2 + \Delta\phi^2} = 1.0$ to reconstruct jets with JETCLU.

The cut on N_γ is intended to avoid di-jet events actually being γ-jet ones.

As we want to apply our algorithm on both jets, we require both of them to be central. The η=0 region is avoided because no special corrections were applied for the towers near it.

No third jet with raw $P_T > 8$ GeV was required to avoid hard gluon radiation.

$E_{nc}^{J1,J2}$ is the energy clusterized in jet1 or jet2 which is released outside the central calorimeter. A cut on this variable is required to limit the amount of jet charged energy not detected by the CTC.

$Chf^{J1,J2}$ is the fraction of the jet1 or jet2 raw P_T carried by charged particles. In next chapter we will show (see also ref. [55]) that this cut is very efficient in removing residual cosmics background and events with lower energy jets due to a greater incidence of out-of-cone activity. This cut retains about 98% of "good" jets.

The JTC96 corrected jet1 P_T was chosen in the range 25-55 GeV to be as close as possible to the same energy range as the γ-jet sample [2].

All the above cuts reduced our sample to ∼5600 events from the original ∼610K events on tape.

[2] In the next subsection will be shown as this variable is the best to fix the energy scale of the event.

We also preferred to avoid the following standard cut usually applied to the jet20 trigger dataset in the di-jet balancing studies:

$$\sum_{i=1}^{2} P_T^{Ji}(Raw) > 50GeV \tag{4.2}$$

This cut is intended to reduce the incidence of biased events coming from the on-line trigger. If we consider events with $\Sigma P_T(Raw) \leq 2\times$(single jet threshold), we select a biased sample because of the following reasons:

1. a "good" central jet triggering the event and a secondary jet which can be underestimated because of cracks [19]

2. a "*feed up* effect" when the triggering jet, actually originating from a parton with energy below the threshold, results having higher energy because of fluctuations coming from the energy resolution [56].

The cut 4.2 hardens the P_T spectrum of the jets (only $\sim 27\%$ of the events is in the 30-35 GeV P_T range and no event below 30 GeV). This does not allow the JCOR2K corrections to work well using the set of parameters optimized on the γ-jet sample, which is characterized by a lower P_T spectrum. Furthermore, considering that our algorithm makes use of tracking information, we aspect better results in events were the jet energy is underestimated because of cracks. For the same reason we did not apply a cut on the second jet P_T.

This aspect will be better investigated in section 4.2.4.

4.2.2 Event Energy Scale

In the JTC96 module the jet relative energy correction functions are parametrized as a function of the jet η in five different $\Sigma P_t(Raw)$ ranges [20] (see Chapter 2). In other studies on the jet energy resolution from di-jet balancing, it was considered the energy dependence as a function of the *corrected* ΣP_t (ΣP_t(JTC96)), as the final goal was to determine the resolution for a jet of a given corrected P_T [56].

In the present study we would like to reproduce (at least qualitatively) the results of the γ-jet analysis shown in fig. 4.10. The problem is that now we don't have a "good probe" of the energy scale like the photon. So the question is: what is the best variable to represent the energy scale of the event in order to study the jet energy resolution as a function of it?

It is clear that using the average value of the raw P_T of the two leading jets ($\Sigma P_t(Raw)/2$) is not the best choice because we know that the raw jet energy is underestimated by about 20%.

Considering that in average the JTC96 corrections reproduce the right energy scale of a jet, a better choice seems to be $\Sigma P_t(\mathrm{JTC96})/2$. But in the present analysis, for the reasons explained in the previous subsection, no cut is given to the P_T of the second jet and to $\Sigma P_t(Raw)$. Therefore, in principle it is possible to have events with a big difference in the corrected energies of the two leading jets. Consequently $\Sigma P_t(\mathrm{JTC96})/2$ can fake the right energy scale.

The last choice we can think is to consider the $P_T^{J1}(\mathrm{JTC96})$. In fact in general the most energetic jet, being the one which mainly triggered the event, will be more central so having less probability to fall near cracks and will be the best measured as the jet energy resolution improves with the increase of its energy.

In order to make a quantitative comparison between these two choices, the projected $\vec{\not{E}}_T$ (PME) balancing technique, described in section 4.2.3, was applied to the JTC96 corrected jets to get an estimate of the single jet energy resolution. Fig. 4.15 shows how the typical function $K/\sqrt{P_T}$ fits well the jet energy resolution distribution obtained binning in $P_T^{J1}(\mathrm{JTC96})$ (circles) while it does not using the different binning according to $\Sigma P_t(\mathrm{JTC96})/2$ (triangles). The full function $(103\%/\sqrt{P_T})$ is the best fit to the circles with reduced χ^2 $\sim$ 1.1 while the dashed one (corresponding to $98\%/\sqrt{P_T}$) is the best fit for the triangles with reduced $\chi^2 \sim 17$.

We conclude that P_T^{J1}(JTC96) is the best variable to fix the energy scale of the event and we will use it in the following.

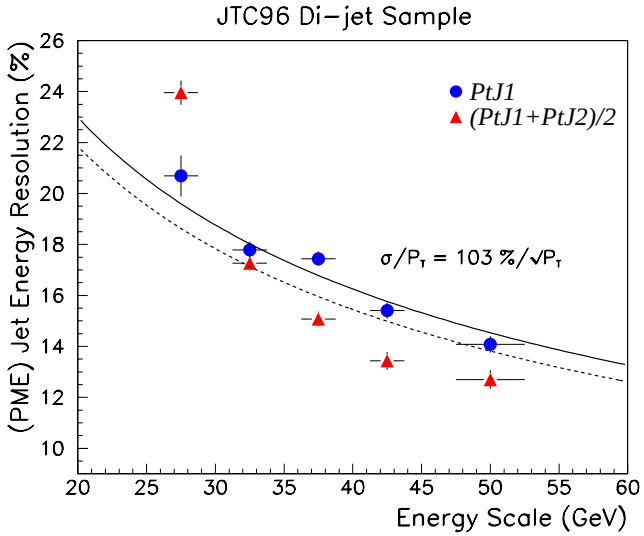

Figure 4.15: *Central detector resolution for JTC96 corrected jets as function of the energy scale of the event obtained with two different criteria relying on the JTC96 corrected jet P_T's.*

4.2.3 Jet Resolution Measurements

In order to check the algorithm we want to compare its impact on the jet energy resolution with respect to the standard JTC96 corrections. The di-jet balancing techniques can directly give an estimate of the jet resolution while an indirect study can be done looking at the $\vec{\not\!\!E}_T$ resolution.

A measurement of the single jet energy resolution was obtained using two independent balancing standard methods: the K_T and the projected $\vec{\not\!\!E}_T$ (PME). In order to have symmetric distributions around zero, in both studies the jet labeling was made in a random way relying on their azimuthal angle: label 1 (*"probe jet"*) was attributed to the lower ϕ jet regardless of its energy.

K_T Balancing Studies

In the γ-jet analysis the central detector jet resolution was obtained using an approach à la UA2 previously applied in CDF to jet resolution studies made on di-jet samples (see for instance ref. [53]). In the present analysis we started using a similar technique.

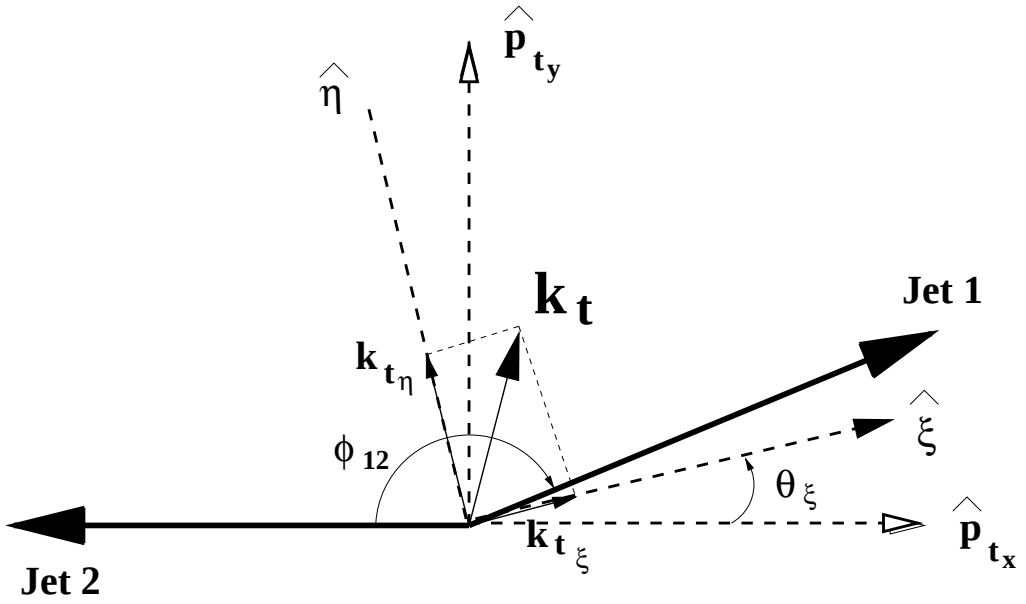

Figure 4.16: *The total transverse momentum $\vec{K}_T$ of the di-jet system is decomposed into two components K_{T_η} and K_{T_ξ} sensitive to different effects generating it. To have symmetric distributions the label 1 is randomly attributed to the lower ϕ jet.*

The two-jet imbalance vector $\vec{K_T} = \vec{P}_T^{jet1} + \vec{P}_T^{jet2}$ was decomposed in two components K_{T_ξ} and K_{T_η} along the azimuthal angular bisectors of the two-jets system (see fig. 4.16). The situation is quite similar to the γ-jet study. These components are sensitive to different effects generating the $\vec{K_T}$ vector: the main source of K_{T_ξ} is the calorimeter energy resolution while angular measurements errors due to gluon radiation effects, being the main origin of K_{T_η}, are common to both components. Hard gluon emission effects are reduced with the cut on the third jet P_T (8 GeV). Soft gluon effects are removed by subtracting in quadrature the sigmas of the K_{T_ξ} and K_{T_η} distributions. In this way, subtracting the finite angle effects (σ_η) from the global jet energy resolution (σ_ξ), we can get an estimate of the *effective* central calorimeter jet energy resolution which is therefore defined as:

$$\sigma_D = \sqrt{\sigma_\xi^2 - \sigma_\eta^2} \tag{4.3}$$

In a di-jet sample the resolution obtained with this method is actually coming from the convolution of the single jet resolutions. So, in order to obtain the single jet energy resolution, we have to scale σ_D by $1/\sqrt{2}$.

Fig. 4.17 (top) reports the jet energy resolutions we got with this method using σ_ξ and σ_η as obtained from a gaussian fit of the K_{T_ξ} and K_{T_η} distributions. The standard JTC96 corrections and JCOR2K are separately applied on both jets. The drawn functions are the best fits to data. A clear improvement in jet energy resolution ($\sim 15\%$) is observed.

Incidentally, we also note how the jet energy resolutions we obtain in this data sample are higher than in the γ-jet one for both methods (see fig. 4.10 (top)). This aspect can be attributed to the convoluted contribution of both jets to σ_ξ and σ_η which we simply accounted for with a $1/\sqrt{2}$ factor. So having an indication of how difficult is an absolute measurement of the jet energy resolution.

As in the γ-jet sample, we also considered the absolute jet energy scale. To check that our algorithm is correctly reproducing it also in the di-jet sample, we compare the energy scale of the two correction methods using the variable:

$$f_s = (P_T^{JCOR2K} - P_T^{JTC96})/P_T^{JTC96} \tag{4.4}$$

for the two leading jets of the event. The results are shown in fig. 4.17 (bottom). JCOR2K is reproducing the JTC96 jet energy scale within its systematic uncertainty ($\sim 5\%$). These results are similar to those seen in fig. 4.10 (bottom) for the γ-jet sample.

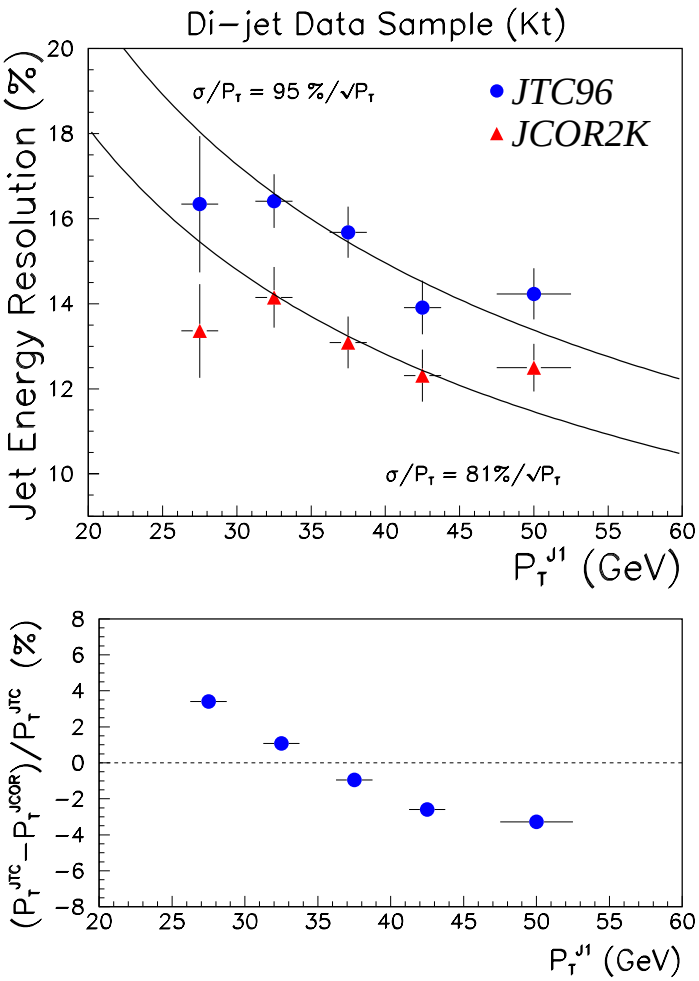

Figure 4.17: *Top: The single jet central detector resolution $\sigma_D/\sqrt{2}$ as a function of the energy scale of the event parametrized by the JTC96 corrected P_T of the leading jet. Both jets are JTC96 or JCOR2K corrected. Bottom: energy scale comparison $(P_T^{JCOR2K} - P_T^{JTC96})/P_T^{JTC96}$ for the two leading jets of the event as function of the energy scale.*

Projected $\vec{\not{E}}_T$ Studies

An alternative approach to the previous study can be made considering the projection along the probe jet axis of the $\vec{\not{E}}_T$ vector, defined by:

$$PME = \vec{\not{E}}_T \cdot \hat{P}_T^J(probe) \tag{4.5}$$

This variable is usually preferred in di-jet balancing studies because, being the $\vec{\not{E}}_T$ derived from all towers above a given threshold with $|\eta| < 3.6$, it is less dependent on third jet activity and on whether it is clusterized or

not. Projecting the $\vec{\not\!\!E}_T$ along the probe jet axis (which actually is almost coincident with the axis of the other jet) minimizes the effect of the *"K_T kick"* due to the third jet which tends to be perpendicular to this axis. So the corresponding di-jet balancing resolution is better as the PME, accounting event per event for the effects of additional jets, will be mainly due ONLY to calorimetry effects [57].

Fig. 4.18 reports the azimuthal angular separation between the leading (full) and secondary (dashed) jet and the $\vec{\not\!\!E}_T$ vector. As expected, in a di-jet event the $\vec{\not\!\!E}_T$ is mainly due to jet mismeasurement so its direction is mainly along the di-jet system axis oriented according to the secondary jet.

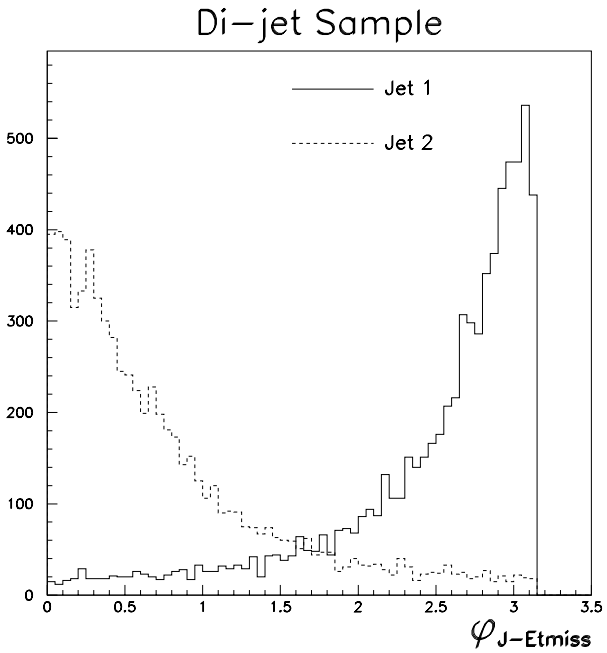

Figure 4.18: *Azimuthal angular separation between the leading (full) and secondary (dashed) jet and the $\vec{\not\!\!E}_T$ vector in the di-jet events. As the $\vec{\not\!\!E}_T$ vector is mainly originated by jet mismeasurement, its direction is strongly correlated with the di-jet system axis being oriented along the secondary jet.*

Now we do not need to subtract the soft third jet contribution from the jet resolution and the sigma of the PME distribution (σ_{PME}) can be directly taken as a measurement of the jets energy resolution. As both jets contribute to it, the single jet resolution is derived scaling by $1/\sqrt{2}$.

To compare the JTC96 and JCOR2K corrections with this method, we have to correct the $\vec{\not\!\!E}_T$ (derived from the tower raw energy) for the difference in P_T between corrected and uncorrected jets.

The *corrected* $\vec{\not{E}}_T$ is defined according the following equation:

$$\vec{\not{E}}_T^{\,corr} = \vec{\not{E}}_T^{\,raw} - \sum_i (\vec{P}_T^{J_i}(corr) - \vec{P}_T^{J_i}(raw)) \tag{4.6}$$

Were "corr" stands for JTC96 or JCOR2K and the sum is made on the three leading jets of the event. The third jet is JCOR2K corrected only if central ($|\eta^{J3}| < 0.7$) otherwise is JTC96 corrected.

The results on jet energy resolutions obtained applying the PME method are shown in fig. 4.19. Data points come from a gaussian fit of the PME distributions while the drawn functions are the best fit to data. A net improvement in jet energy resolution is obtained with the new algorithm ($\sim$ 15% like the result obtained with the K_T method). We can see also how we got a more regular trend with this method. The origin of that can be attributed both to a better gaussian fit of the PME distributions and to the fact that, unlike the K_T balancing method, now we are not combining in quadrature the sigmas of two gaussian fits which results in more fluctuations.

The jet energy resolution measurements are a bit greater than those obtained with the other method (see fig. 4.17). Anyway in the present analysis we are not interested in understanding what is the best estimate of such quantities, as we rather want to make a comparison between the two jet energy correction approaches.

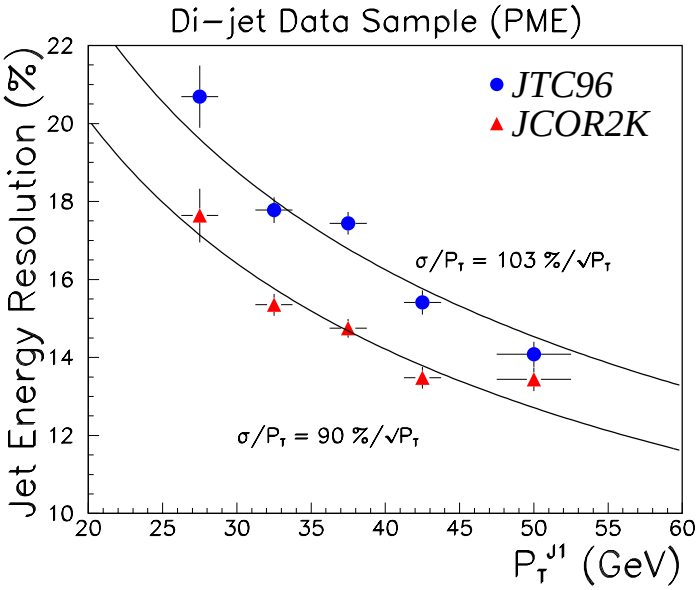

Figure 4.19: *The single jet central detector resolution $\sigma_{PME}/\sqrt{2}$ as a function of the energy scale of the event parametrized by the JTC96 corrected P_T of the leading jet. The distributions are obtained after correcting the $\vec{\not{E}}_T$ for JTC96 or JCOR2K jet corrections.*

4.2.4 $\vec{\not\!\!E}_T$ Resolution Studies

As an alternative check of the new algorithm we looked at its impact on the $\vec{\not\!\!E}_T$ resolution. As in the previous subsection, the comparison between JTC96 and JCOR2K was made correcting this variable separately for the two methods according to eq. 4.6.

For the $\vec{\not\!\!E}_T$ resolution studies we used the standard definition given by:

$$\sigma_{\not\!\!E_T} = \sqrt{< \not\!\!E_T^2 >} \tag{4.7}$$

In fact, if the $\vec{\not\!\!E}_T$ x and y components are expected to be gaussian with the same σ, the $\not\!\!E_T^2$ distribution will be exponential with $< \not\!\!E_T^2 >= 2\sigma^2$.

The $\vec{\not\!\!E}_T$ resolution is usually parametrized in function of the total scalar E_T of the event: $\sigma_{\not\!\!E_T} = K \cdot \sqrt{E_T^{tot}}$ were $K \sim 0.7$ in minimum-bias events [58]. From fig. 4.20, were $\sigma_{\not\!\!E_T}$ is plotted versus the total scalar raw E_T of the event, we can see the improvement in $\vec{\not\!\!E}_T$ resolution obtained with the new method. Data points were derived from an exponential fit of the $\not\!\!E_T^2$ distributions.

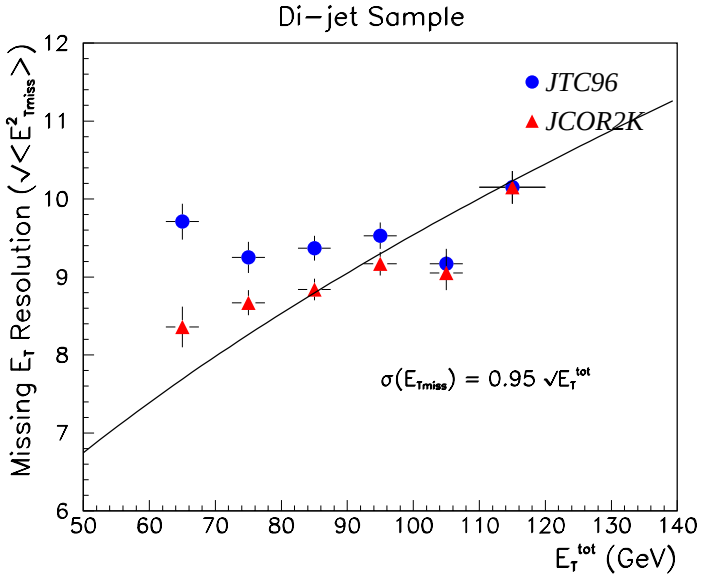

Figure 4.20: *$\sigma_{\not\!\!E_T}$ as a function of the total scalar E_T of the event. The plots are obtained from an exponential fit of the $\not\!\!E_T^2$ distributions after correcting the $\vec{\not\!\!E}_T$ for the JTC96 or JCOR2K jet corrections.*

It is important to notice that the standard trend $K \cdot \sqrt{E_T^{tot}}$ is not satisfied mainly applying the JTC96 corrections and in particular at low E_T^{tot} values where a worse resolution is observed. The function plotted in fig. 4.20 is the best fit to JCOR2K data. We attribute that to the fact that our sample is

biased by events with a secondary jet which can be much underestimated. The differences in the plots of fig. 4.20 are so attributed to the different approach of our algorithm in recovering part of the jet energy using track information which results in a better resolution just at lower E_T^{tot} in events were part of the jet energy is lost in cracks.

To check this hypothesis we reselected the di-jet sample applying the following cuts which are standard for the jet20 trigger dataset:

- $\sum_{i=1}^{2} P_T^{Ji}(Raw) > 50$ GeV
- $P_T^{J2}(raw) > 15$ GeV
- $P_T^{J3}(raw) < 15$ GeV

Where the cut on the third jet was released to 15 GeV in order to have more statistics.

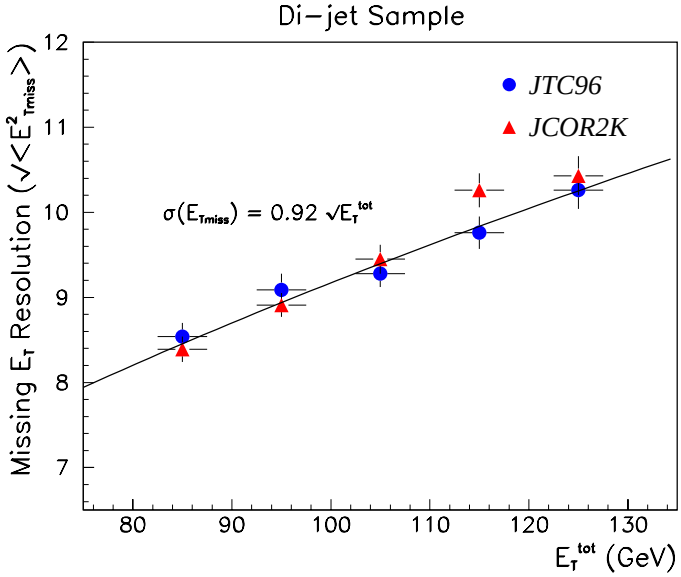

Figure 4.21: *$\sigma_{\not\!\!E_T}$ as a function of the total scalar E_T of the event. The distributions are obtained after correcting the $\vec{\not\!\!E}_T$ for the JTC96 or JCOR2K jet corrections. The trigger bias is removed applying standard cuts.*

The results on $\sigma_{\not\!\!E_T}$ obtained with the new cuts are reported in fig. 4.21. We can see how the typical $K \cdot \sqrt{E_T^{tot}}$ trend is well satisfied for both distributions. As expected, the improvement we got with the new method is greatly reduced because the cuts selected a sample with a harder jet P_T spectrum.

Anyway indications are found that the new algorithm can also improve the $\vec{\not\!\!E}_T$ resolution at least in events where part of the energy is clusterized in jets.

Chapter 5

Jet Studies With the New Algorithm

In this chapter we show how it is possible to extract (requiring the absence of "mixed-target towers") a not negligible γ-jet sub-sample where the jets are not well measured. This effect is not dependent on the new tower energy corrections as it is also present with the standard JTC96 corrections. Two concomitant effects contribute: a stronger incidence of cosmics events faking γ-jet production (which is connected with the specific data sample used to test the algorithm but which actually is the less relevant effect) and a greater incidence of out-of-cone activity (which we presume to be sample independent as we also observed it in the di-jet sample).

5.1 A Golden Subsample

In Chapter 3 we introduced a new method to form the energy of a calorimeter tower relying on the classification derived from the kind of particles hitting it. The new jet energy is then obtained combining the new energy information of the towers belonging to it.

Because of fluctuations in the fragmentation process, the fraction of jet towers which we classified in each group is different from jet to jet, strongly depending on the topology of the particles inside a jet. In particular we expect a lower fraction of mixed towers in broader jets, where particles are well separated. The opposite effect is expected in the case of collimated jets. Sometimes we find a jet where no mixed tower is present. Still considering the γ-jet sample, the fraction of events without mixed tower as a function of the photon P_T is shown in fig. 5.1. As expected, the lower the photon energy the higher the fraction of events without mixed towers, as a less energetic

jet (less collimated) is boosted by the photon. Consequently such events are characterized by a lower jet P_T spectrum.

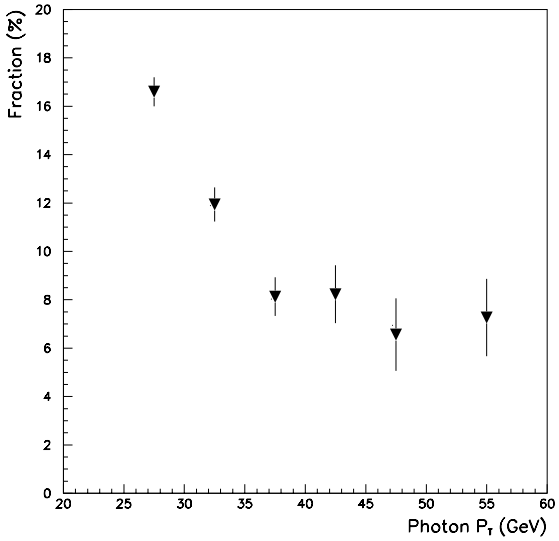

Figure 5.1: *Fraction of events without Mix-Towers as a function of the photon P_T.*

We expect such cases to be "golden" jets, in which the resolution is the best achievable, i.e. an upper limit to our efforts in improving jet energy resolution. In fact, most of the parameters we need are related to mixed towers. However, in this chapter we present results which show how such a "no mix-towers" sub-sample is actually characterized by a high cosmic rays background and low energy jets with an underestimated transverse energy due to out-of-cone leakage which actually do not allow us to use this sub-sample to obtain this upper limit on jet energy resolution improvement.

5.2 Something Strange

As seen in Chapter 4, as first monitor of the resolution we can consider the distribution of:

$$f_b = (P_T^{jet} - P_T^{\gamma})/P_T^{\gamma} \tag{5.1}$$

The width of this distribution is dominated by the jet resolution, thus the distribution gets narrower as the jet resolution improves. Incidentally we

have to remember that the classification parameters were obtained from the γ-jet sample as the ones which minimize the f_b distribution.

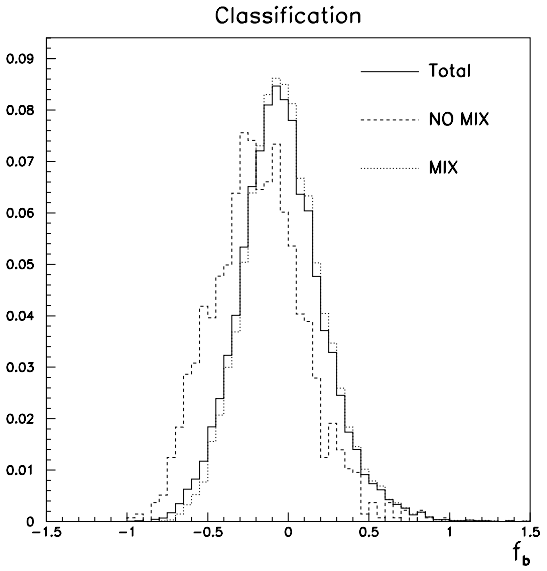

Figure 5.2: *f_b distributions for total (full), NO MIX (dashed) and MIX (dotted) samples. The jet energy is evaluated using the classification algorithm. Distributions are normalized to 1.*

As already mentioned, we expect events without mixed towers to be characterized by the best resolution achievable with our method. The f_b distribution for our classification algorithm is reported in fig. 5.2. In these plots the full histogram is relative to total sample, the dashed one to events WITHOUT mixed towers ("NO MIX" sub-sample), while the dotted one is for events WITH mixed towers ("MIX" sub-sample). We see how, contrary to our expectations, the f_b distribution for jets without mixed towers is not symmetric, is broader and shifted to negative values. This means that in this sub-sample jets are not well measured and in average underestimated.

It is important to notice that this effect is present both using our algorithm and the JTC96 correction (see fig. 5.3). So we can claim that it is not a bias due to our method of correcting for the towers energies but it is only connected with the selection made by the tower classification. From this figure, we can also see that in the NO MIX sub-sample our method does not give a good improvement. *From now on we will use the JTC96 method to perform our studies on this sub-sample.*

To understand where those events are coming from, we looked at some of

them with the CDF event display. With this tool we found some indication that the NO MIX sub-sample is rich in events due to cosmic rays passing the standard cuts. One peculiar event without mixed towers is reported in fig. 5.4. The next step was to find an objective way to recognize and remove these background events.

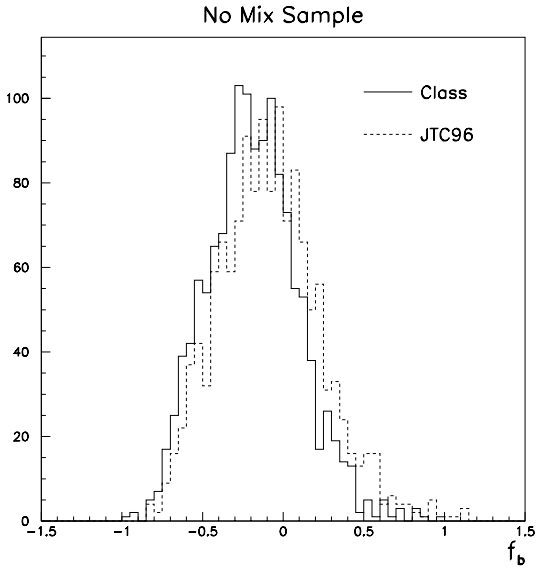

Figure 5.3: *f_b distribution in the "NO MIX" sub-sample for the classification method (full) and JTC96 (dashed).*

5.3 Cosmics Background Studies

In CDF note 4256 [59] it was shown that the timing information from hadronic calorimeter TDCs can be used to separate good photon candidates from cosmics background. The energy threshold to fire the TDCs is about 300 MeV so a not negligible fraction of central photons above 25 GeV can have timing information being the energy leakage from an EM cluster in the hadronic calorimeter about 1%. Using a new algorithm to form the time for an ELES object and a sample of 70 GeV photon triggers (very rich in cosmic rays), it was shown that cosmic candidates have a characteristic FLAT time spectrum while the photon candidates have a sharp peak around 7 ns which is about the time separation between the interaction point and the central hadronic calorimeter for prompt particles. We used the same algorithm on our γ-jet sample finding about 40% of events with timing information on the photon. We continued our study on this sub-sample were we could use this further information as a tool to reject cosmics background.

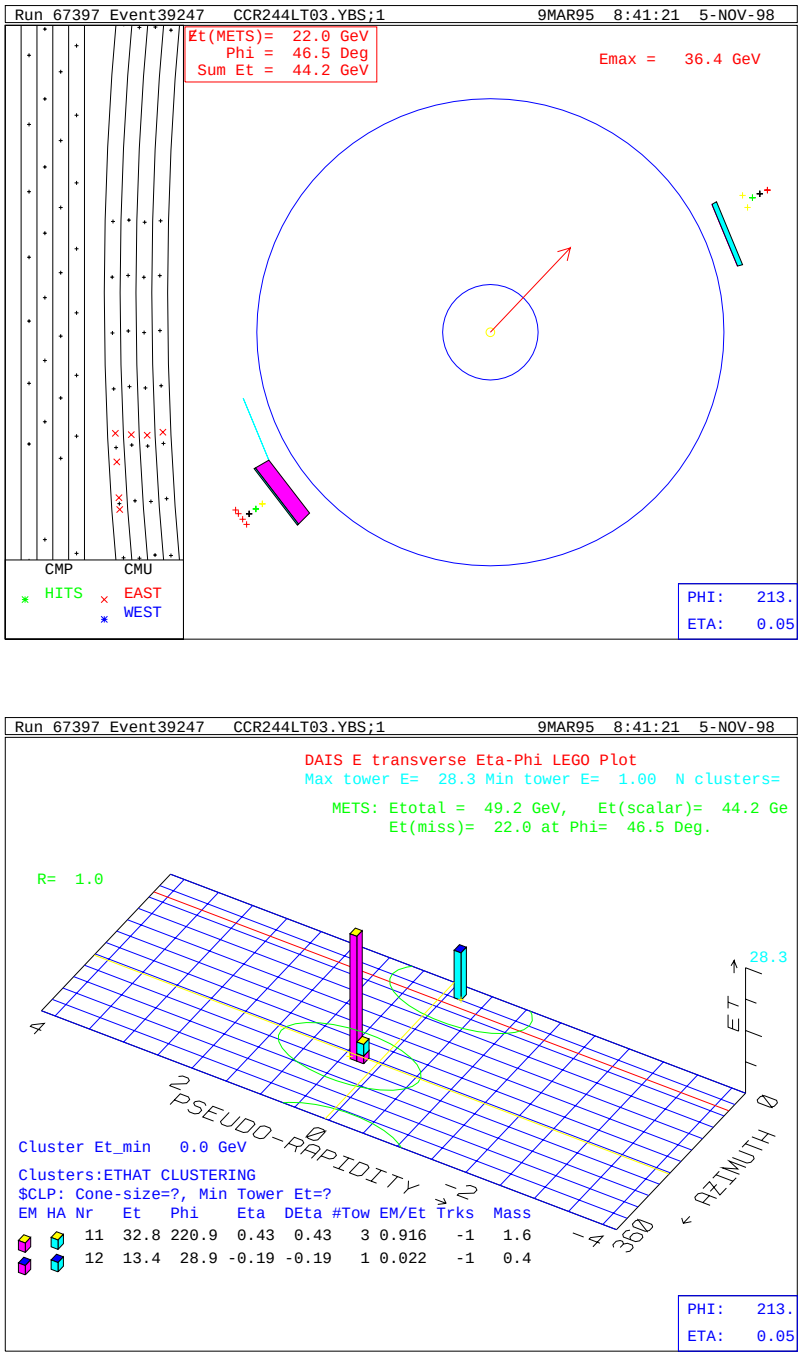

Figure 5.4: *Peculiar event in the sub-sample without mixed towers attributed to cosmics background.*

Fig. 5.5 shows a flat distribution (cosmics) superimposed to a peak (photons)[1] for both NO MIX and MIX sub-samples. This confirms the fact that in our sample there is a residual contamination of cosmic rays passing the standard cuts. In particular we can see how the NO MIX sub-sample is richer in cosmics than the MIX one.

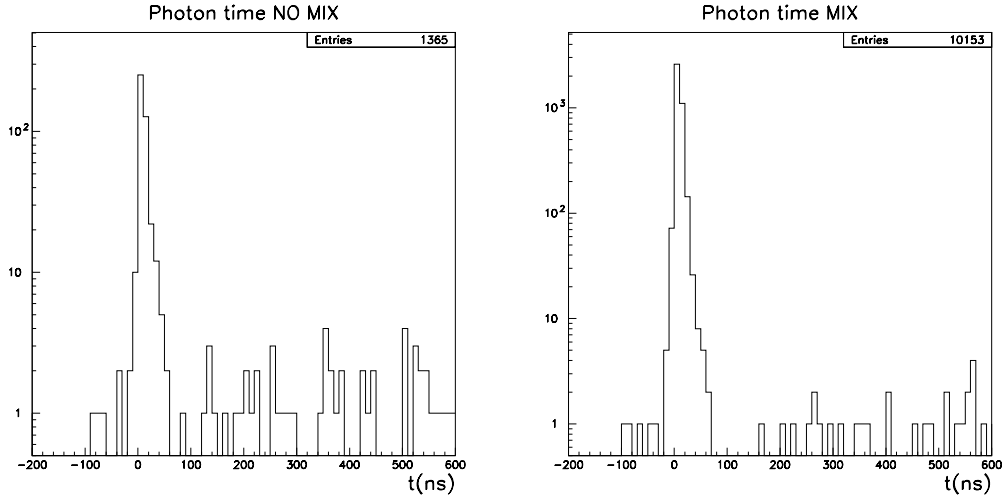

Figure 5.5: *Photon timing distribution in NO MIX and MIX sub-samples. The former (left) is richer in events in the flat zone attributed to cosmics.*

We decided to consider as *"Good Photon"* a photon candidate with:

$$\text{-10 ns} \leq t_\gamma \leq 45 \text{ ns}$$

and *"Cosmic"* otherwise. This "time window" was chosen so to flatten the distributions in fig. 5.5 outside it.

With this time cut we noticed that the NO MIX sub-sample is richer in cosmics than the MIX one for about a factor of 10. An estimate of the residual cosmics background in the "Good Photon" sample was obtained from the number of events between 100 ns and 600 ns scaled to the width of the time window. The residual background for this sub-sample resulted reduced of about a factor of 10 respect the original sample. Table 5.1 summarizes these results.

From fig. 5.6 we can see that in the sub-sample with photon timing information we substantially got the same f_b distributions as in the original

[1] A blow up of the peak gives us a photon timing curve consistent with that found for good photons in reference [59].

Sub-sample	Events	Events with t_γ	Cosmics	Residual Cosmics
NO MIX	1365	489 ($\sim 36\%$)	62 ($\sim 12.7\%$)	5.6 ($\sim 1.3\%$)
MIX	10153	3990 ($\sim 39\%$)	48 ($\sim 1.2\%$)	3.1 ($\sim 0.08\%$)
Total	11518	4479 ($\sim 39\%$)	110 ($\sim 2.5\%$)	8.7 ($\sim 0.2\%$)

Table 5.1: *Number of events in NO MIX and MIX sub-samples. Column 4 shows how the NO MIX one has a cosmics background about a factor 10 bigger compared to MIX one. From column 5 we can see how the timing cut reduces cosmics by about a factor of 10.*

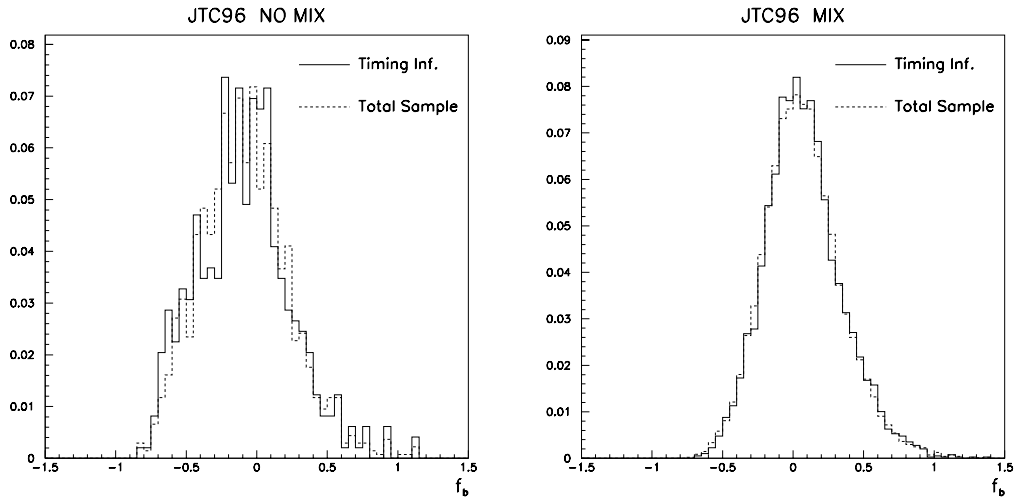

Figure 5.6: *f_b distributions for NO MIX (left) and MIX (right) sub-samples with (full) and without (dashed) photon timing information. Distributions have been normalized to 1.*

sample. So no bias was introduced by the time information requirement. In particular the "shoulder" at lower values near -0.5 still remains.

Using time information to distinguish between photon and cosmics, we got the results shown in fig. 5.7. We can see how the cosmics background gives its main contribution just in the "shoulder" and how its rejection gives a narrower f_b distribution. Nevertheless this new distribution continues to be not well symmetric and with a residual shoulder. To check if this effect can be attributed to the residual cosmics in the "Good Photon" sample, we tried to make an estimate of the number of events contributing to this shoulder.

This was obtained from fig. 5.8 evaluating the counts in the NO MIX distribution corresponding to the area (yellow zone) in the shoulder above

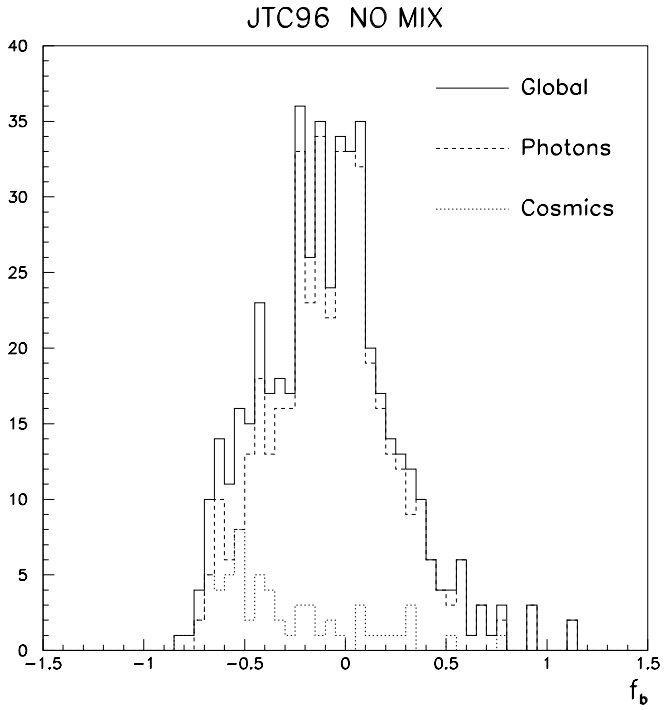

Figure 5.7: *f_b distributions for global (full), photon (dashed) and cosmics (dotted) events in the NO MIX sub-sample. Cosmics give their main contribution just in the "shoulder" zone. All distributions are normalized to 1.*

the MIX distribution which was supposed to have the "right" trend. To do that the NO MIX distribution was shifted by 0.1 so to match its peak with that of the MIX one while the common normalization was the number of events in the NO MIX sample.

In this way we got about 32 events corresponding to the yellow area while from table 5.1 we know that the cosmics contamination of the "Good Photons" sample is less than 6 events.

As a check we considered the variable Chf (charged fraction) defined as:

$$Chf = \frac{\Sigma P_T^{tr}}{P_T^J(raw)} \tag{5.2}$$

Where the sum is on the tracks inside the jet cone.

We know that [60] in γ-jet events due to cosmics, the jet is very likely to have few tracks because it comes from a coincidence between a cosmic ray bremsstrahlung photon and a minimum bias collision or is due to particles

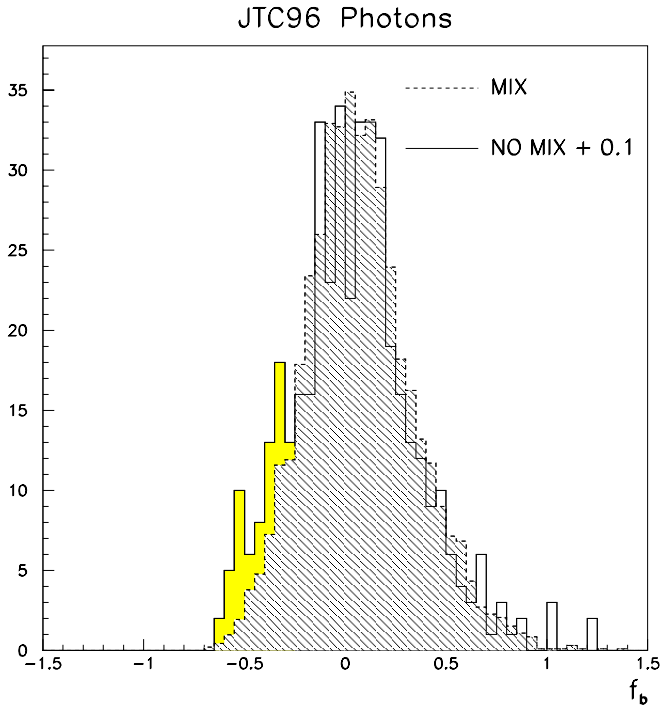

Figure 5.8: *f_b distributions for photon events in NO MIX (full) and MIX (dashed) sub-samples. The NO MIX one is shifted by 0.1 to match the two peaks. The MIX one is normalized to the number of events in the NO MIX sample.*

generated by cosmics interaction inside calorimeters with tracks not well reconstructed by the CTC being not originated in the interaction zone. So we aspect that the Chf variable will have low values. Fig. 5.9 (left) confirms that: "bad photon" events are mainly clusterized at lower values while "good photon" ones have a quite different distribution with a maximum near 0.85. From the same figure (right) we can see how the NO MIX photon sub-sample is characterized by a flat spectrum which confirms us that our previous cosmics residual background estimate is good. The difference in Chf distribution between total and NO MIX photon samples is attributed to the requirement of no mixed-target tower in the latter which selects events with less tracks inside the jet and so less energy carried by charged particles. This aspect will be further investigated in next section.

Incidentally we note how the Chf variable can be used to get rid of cosmics background in events were photon timing information is not available. A cut at 0.2, for instance, reduces cosmics of about 50% retaining about

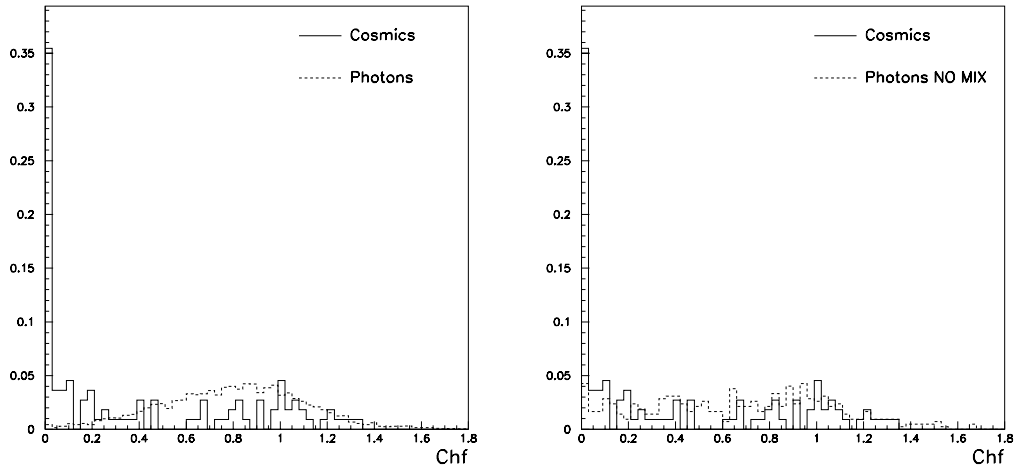

Figure 5.9: *Chf distributions for photons and cosmics (left) and for photon NO MIX and cosmics (right). Values greater than 1 derive from the SCALAR sum on tracks P_T. No indication is seen of a significative cosmics background in the photon NO MIX sub-sample. All distributions are normalized to 1.*

97.5% of photons. Anyway in the following we continue to distinguish good photons only relying on timing information so to have a very low residual cosmics incidence.

We concluded that the "residual shoulder" in the f_b distribution for the NO MIX sample cannot be substantially attributed to cosmics background and the origin of this anomaly needs to be further investigated together with that of the shift towards negatives values.

5.4 Jet Studies

In the previous section was shown that the NO MIX sub-sample is characterized by a f_b distribution with a strange behavior even after the cosmics background removal: an asymmetric shape and a mean value underestimated of about 10%

As expected, low f_b values correspond to events with a jet of low P_T ($\leq$ 20 GeV) (see fig. 5.10) which results boosted by a photon in the 25-35 GeV range. The NO MIX sub-sample has a percentage of such events four times ($\sim$ 21% respect to $\sim$ 5%) higher than the MIX one as a consequence of the lower P_T^J spectrum (see fig. 5.11 (left)). This fact led us to attribute the

shoulder at low f_b values to an effect of the photon P_T cut at 25 GeV. From fig. 5.11 (right) we noticed some indication of a greater extra-jet activity, which is consistent with the lower P_T spectrum, for the NO MIX events but no significative improvement was obtained with the tighter cut $P_T^{J2} \leq 8$ GeV.

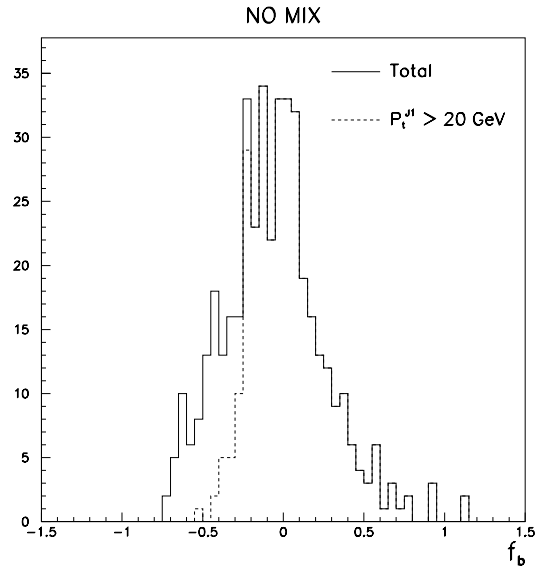

Figure 5.10: *f_b distributions for total NO MIX sample (full) and with $P_T^{J1} > 20$ GeV (dashed).*

From an empirical point of view we can say that a jet without mixed towers is more likely to be not well measured than a jet with and this fact is not connected with extra-jet activity. Nevertheless we need to understand why this happens in order to be sure that our algorithm is correctly working.

The first step of this study (*made in the "Good Photon" sub-sample*) was to look at some informations which could help in understanding if there is some "pathology" in such jets. In fact it is natural to think that the absence of mixed-target towers is connected either with the loss of some tracks, for example in ϕ-cracks or in out-of-cone effects, or with the loss of information from CES detector in events were the detector is not perfectly working.

From fig. 5.12 we can see how, as expected, in the NO MIX sub-sample jets have in average less tracks then in the MIX one. This fact together with no indication from the study of the badrun-flag information rejected the above second hypothesis. Fig. 5.13 shows how the difference in ntracks cannot be attributed to different efficiency of the tracking system for different η_J being the two distribution similar for both sub-samples. Further studies

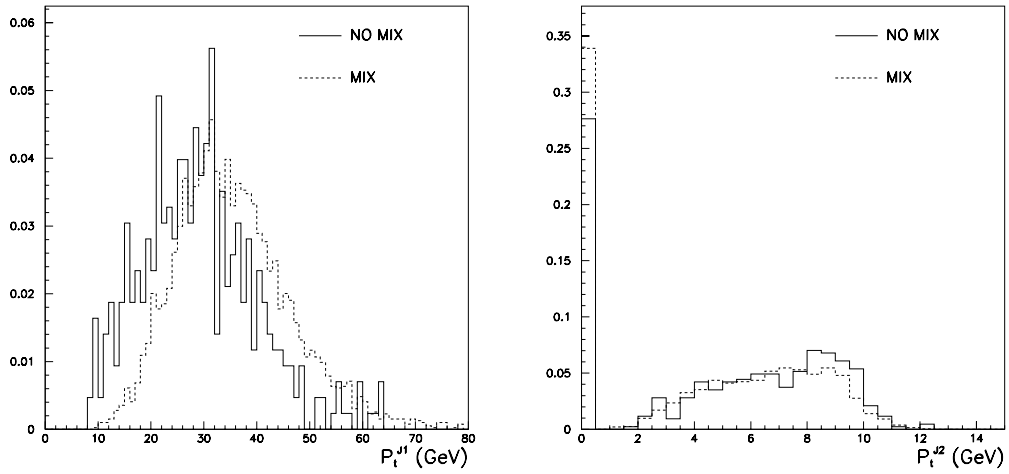

Figure 5.11: *JTC96 corrected Jet P_T distributions in NO MIX (full) and MIX (dashed) sub-samples for leading (left) and secondary (right) jet. Distributions are normalized to 1.*

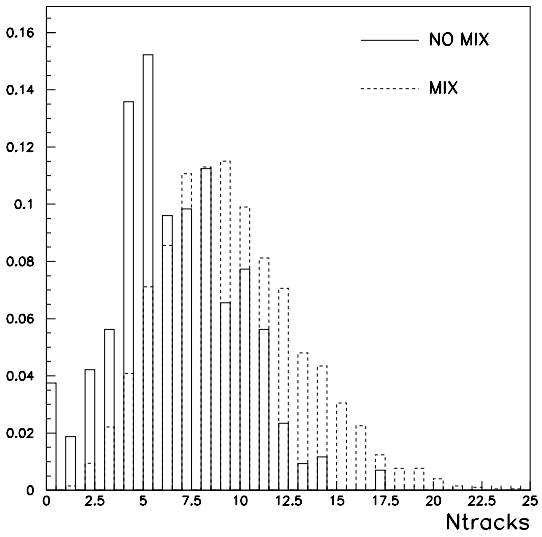

Figure 5.12: *Distributions of the number of tracks inside the jet for NO MIX and MIX sub-samples. In the former (full line) jets have in average less tracks than in the latter (dashed line). Both distributions are normalized to 1.*

on the energy in crack chambers (CCR) inside the jets and on the ϕ of jet centroids showed no substantial difference between the two sub-samples. So no indication was found that the NO MIX sub-sample is more rich in events with tracks inside jets falling in ϕ-cracks.

Another interesting feature was also found in the fact that the shift to low balancing values in events without mixed towers cannot be attributed to the low average number of tracks. How we can see in fig. 5.14 this effect still remains in the NO MIX sub-sample also for events with an high number of tracks inside the jet ($\geq$5) while is not present in the MIX one in events with a low number of tracks ($<$5).

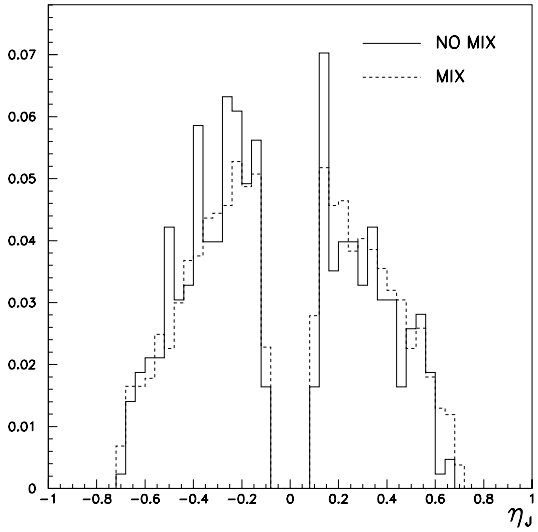

Figure 5.13: *η_J distributions for NO MIX (full) and MIX (dashed) sub-samples. Distributions are normalized to 1.*

The last hypothesis to check was the incidence of out-of-cone effects. Fig. 5.15 indicates a lower charged energy inside jets (Chf variable defined in eq. 5.2) for the NO MIX events which could be connected with a greater number of tracks going out the clustering cone.

To make a quantitative study we naively considered a cone of radius 1.5 around the jet axis (which roughly corresponds to the hemisphere opposite to the photon direction) and then defined the variable :

$$Chf_{out} = \frac{\Sigma P_T^{tr}(cone1.0 - 1.5)}{\Sigma P_T^{tr}(cone1.5)} \tag{5.3}$$

that is the fraction of charged energy in cone 1.5 *outside* the jet cone compared to *total* charged energy in cone 1.5. Chf_{out} should be able to quantify out-of-cone activity.

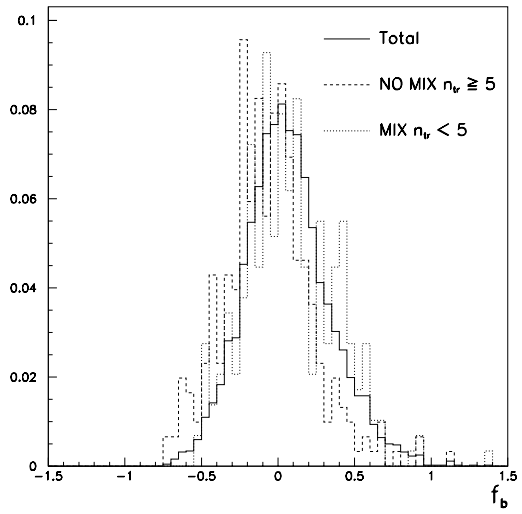

Figure 5.14: *f_b distributions for NO MIX and MIX sub-samples requiring a number of tracks inside the jet respectively ≥ 5 (dashed) and < 5 (dotted). No significative change is observed in both sub-samples comparing these distribution with that of total photon sample (full). Common normalization is 1.*

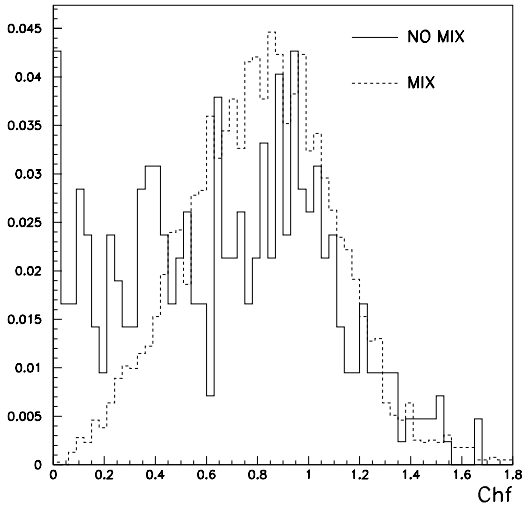

Figure 5.15: *Chf distributions for NO MIX (full) and MIX (dashed) sub-samples. Distributions are normalized to 1.*

Fig. 5.16 indicates some correlation between the Chf and Chf_{out} variables: as expected a lower charged energy inside the jet (which we indicatively selected using the same cut on Chf at 0.2, introduced in the previous section) is in average connected with a higher out-of-cone activity. So a cut on Chf, besides reducing the cosmics background, can also help in rejecting events with a jet energy mismeasurement due to a high out-of-cone leakage.

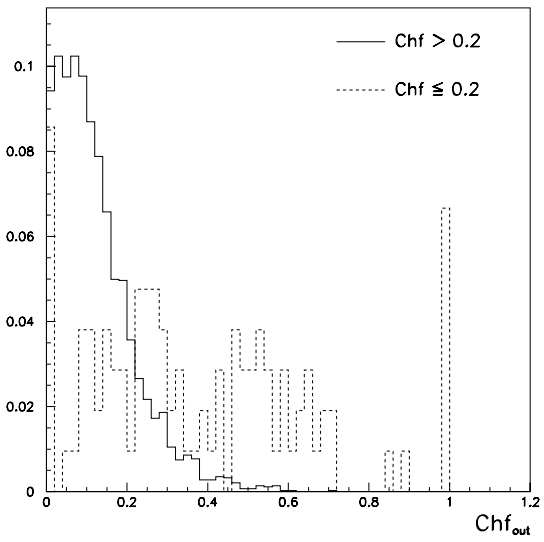

Figure 5.16: *Chf_{out} distributions in the "good photon" sub-sample requiring Chf $>$ 0.2 (full) and Chf $\leq$ 0.2 (dashed). Distributions are normalized to 1.*

From fig. 5.17 we can see how actually, connected to the lower charged energy inside jets, out-of-cone effects are more relevant in the NO MIX subsample. The spike at 1 corresponds to events where no track is inside the jet cone which we know to come mostly from cosmics. The observed number of such events is 7 which is consistent with our estimate of residual cosmic background made in the previous section (about 6 events).

To study the effect of out-of-cone incidence, we applied a cut on Chf_{out} at different values. Fig. 5.18 plots f_b distributions in the NO MIX sample for 3 different cuts (0.25, 0.15, 0.05). We see that as the cut becomes tighter the peak at positive values becomes more relevant while the shoulder at lower values tends to disappear.

Fig. 5.19 shows how increasing the cut on Chf_{out} (that is requiring a lower and lower out-of-cone activity) the f_b distribution for NO MIX events tends to be symmetric and superimposed to the MIX one. Table 5.2 quantifies

the effects of these cuts on the mean values of the f_b distributions. The low values tail still remaining even with $Chf_{out} < 0.05$ (see lower right of fig. 5.19) resulted due to extra-jet activity (it was removed requiring $P_T^{J2} < 6$ GeV) while the spike was removed with a tighter cut on Chf_{out}.

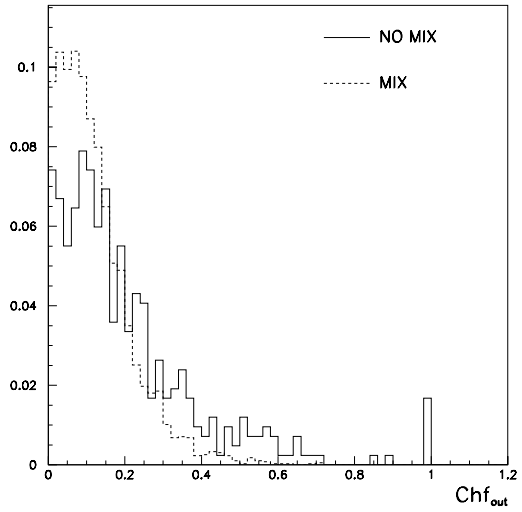

Figure 5.17: *Chf_{out} distributions for NO MIX (full) and MIX (dashed) sub-samples. Distributions are normalized to 1.*

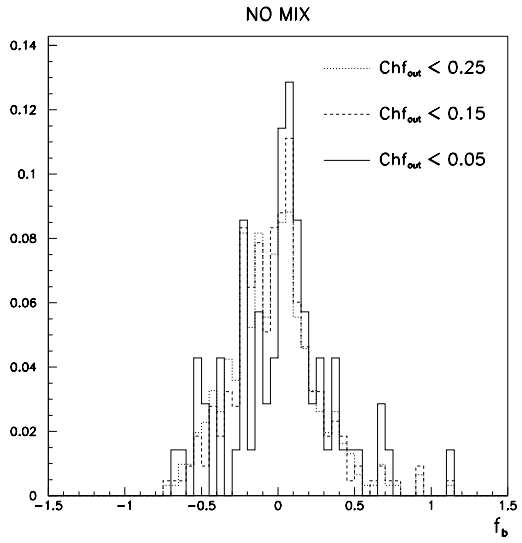

Figure 5.18: *f_b distributions for NO MIX sub-sample requiring different cuts on Chf_{out}: 0.25 (dotted), 0.15 (dashed), 0.05 (full). Common normalization is 1.*

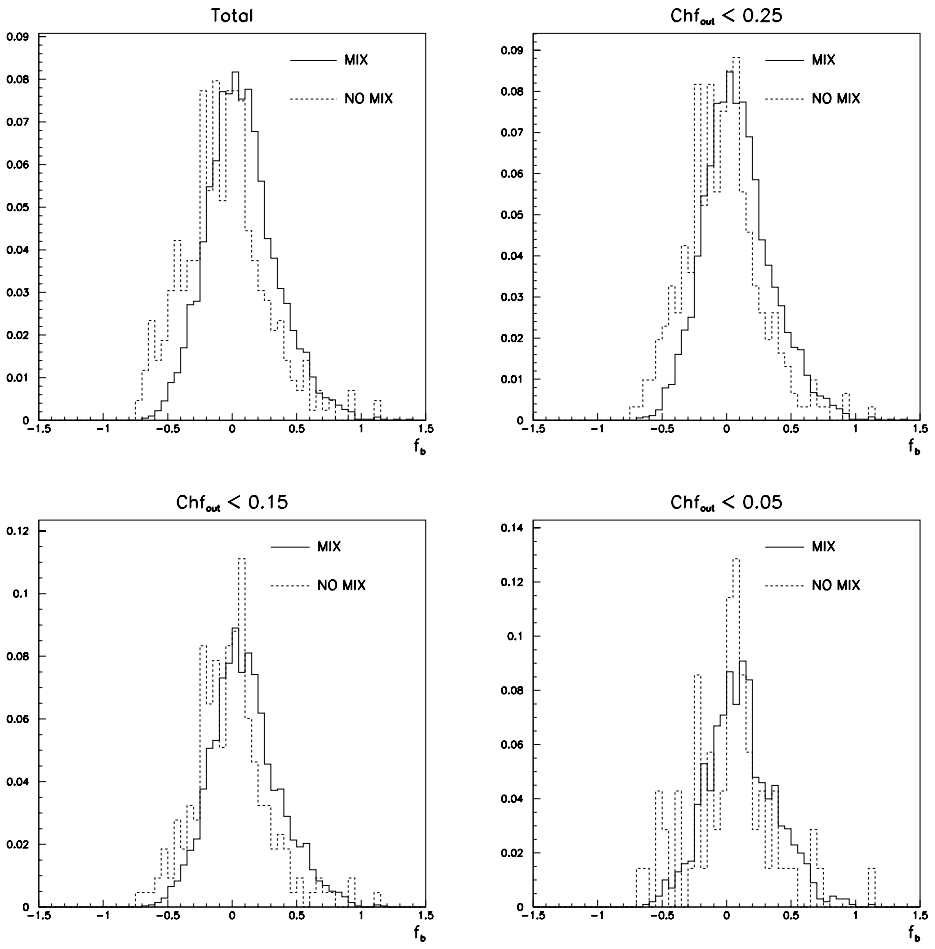

Figure 5.19: *f_b distributions for MIX (full) and NO MIX (dashed) sub-samples without Chf_{out} cut (upper left) and with Chf_{out} <0.25, 0.15, and 0.05. Requiring a lower out-of-cone incidence leads the NO MIX distribution to be symmetric and superimposed to the MIX one. Distributions are normalized to 1.*

Chf_{out} cut	μ_{f_b} NO MIX	μ_{f_b} MIX
None	-0.082 ± 0.014	0.036 ± 0.004
0.25	-0.052 ± 0.015	0.039 ± 0.004
0.15	-0.022 ± 0.015	0.053 ± 0.005
0.05	0.076 ± 0.014	0.062 ± 0.008

Table 5.2: *Mean values of the gaussian fit of the f_b distributions plotted in figure 5.19. Errors are defined as $\sigma/\sqrt{N}$.*

Finally we investigated if a similar effect is also found in the di-jet sample. Fig. 5.20 reports the Chf_{out} distributions for leading and secondary jet requiring (full) or not (dashed) the absence of mixed target towers inside the jet. A similar behavior as in the γ-jet sample (fig. 5.17) is observed indicating this effect is not connected to the particular data sample we considered to test the algorithm.

Some indication of residual cosmics background is also evident in the di-jet sample.

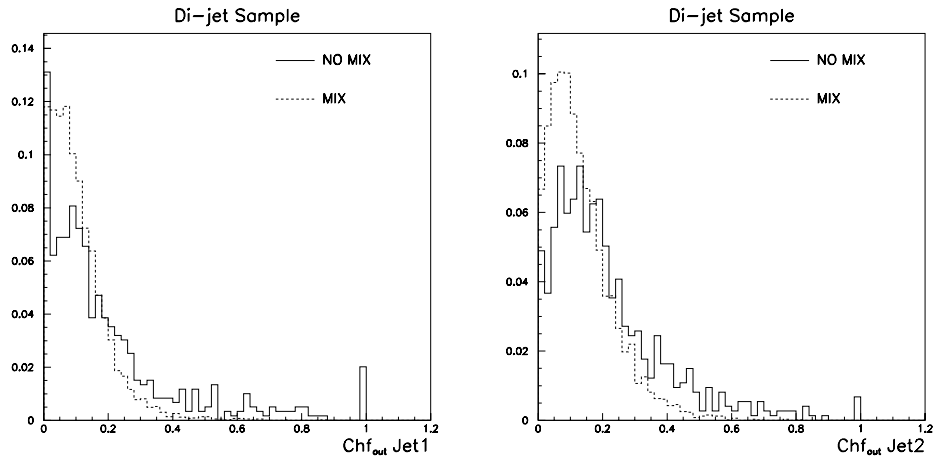

Figure 5.20: *Chf_{out} distributions requiring (full) or not (dashed) the absence of mixed target towers inside leading (left) and secondary (right) jet in the di-jet sample. Distributions are normalized to 1.*

In conclusion, we found that our algorithm can select jets where out-of-cone effects could be relevant. This aspect needs to be further tested when checking the algorithm on other data samples and considered in future when correcting for an absolute jet energy scale.

Appendix A

Jets in QCD

Chapter 2 dealt with the reconstruction of hadronic jets in the CDF experiment. This appendix provides a brief theoretical and phenomenological description of the jet production process in $p\bar{p}$ collisions.

A.1 The Standard Model

In the last decades a new theory was developed that describes all of the known forces among elementary particles [1] : the *Standard Model of Particle Interactions* [61, 62]. According to this theory, all of the fundamental interactions derive from a single general principle, the requirement of *local gauge invariance* (under suitable transformations) of the Lagrangian functions describing them.

Many years of high energy physics scientific research have confirmed the validity of the Standard Model in a wide range of experimental tests. The main goal of present and future high energy experiments, is to further check its predictions (like the Higgs boson existence) also looking at some not expected signature indicating new physics beyond this theory.

A.1.1 Fundamental Forces

The Standard Model actually is a collection of related *local gauge theories* incorporating the three known fundamental elementary particles interactions:

- *Quantum Electrodinamics* (**QED**), describing the *electromagnetic* interactions like the force binding electrons to nuclei.

[1] Exception is made for *Gravity* which, as far as we know, is too weak to play any significant role in ordinary nuclear and sub-nuclear processes.

- *Electroweak Theory*, which describes the *weak* forces responsible of such processes like the beta-decay of nuclei and which unifies them to the electromagnetic ones as being different aspects of a single (*"Electroweak"*) interaction.

- *Quantum Cromodinamics* (**QCD**), the theory of *strong* interactions which, acting at very short distances, bind quarks together to make nucleons (protons and neutrons) and nucleons to make nuclei.

These forces are transmitted by specific carriers, spin-1 particles usually referred as *gauge bosons* [2]. The electromagnetic interaction is described in terms of *photon* exchanges between charged particles, weak forces are transmitted by the $W^{\pm}$ (*charged current* weak interactions) and Z^0 (*neutral current* weak interactions) bosons while the strong interactions are mediated by 8 different *gluons* coupling only to the quark "*color charge*". The γ and gluons are massless while the $W^{\pm}$ and Z^0 bosons (observed for the first time at CERN S$p\bar{p}$S hadron collider in 1983 [63, 64]) are very massive ($\sim$ 80 and 91 GeV respectively).

A.1.2 Leptons and Quarks

The subnuclear physical world is characterized by the existence of many particles. Most of them are unstable and decay into the more familiar particles which constitute ordinary matter (photons, electrons, protons and neutrons). Anyway only few can be considered as the building block of the physical world. According to our present understanding, they are "point-like" elementary particles which can be grouped into two main "families ": *leptons* and *quarks*.

Electrons (e), muons (μ) and taus (τ) (which can be seen as heavier replicas of the electrons) and their associated neutrinos (respectively ν_e, ν_μ and ν_τ), constitute the lepton family. Leptons are spin-$\frac{1}{2}$ fermions not having strong interactions and are coupled into three *weak isospin generations*. Each generation consist of a charged lepton and the associated neutrino and cannot couple to another by weak interactions [3].

Apart from the leptons and the γ, $W^{\pm}$ and Z^0 gauge bosons, all other observed particles have strong interactions and are called *hadrons*. They include a variety of particles among which protons and neutrons. There are many strong (even if indirect) experimental evidences that hadrons are

[2] According to the relativistic quantum theory, these carriers can equivalently be described in terms of fields or particles.

[3] This corresponds to the leptonic quantic number invariance in weak interactions.

Quarks	$\begin{pmatrix} u \\ d \end{pmatrix}$	$\begin{pmatrix} c \\ s \end{pmatrix}$	$\begin{pmatrix} t \\ b \end{pmatrix}$
Leptons	$\begin{pmatrix} \nu_e \\ e \end{pmatrix}$	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}$	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}$
Gauge Bosons	$W^{\pm}$ Z^0	Photon	8 Gluons
Mediated Interaction	Weak	Electromagnetic	Strong

Table A.1: *Standard Model elementary particles and force carriers.*

made out of elementary subnuclear constituents called *quarks* held together by strong forces, but also experiencing electromagnetic and weak interactions. Six different kinds of quarks (also referred as "*flavours*") account for all the several tens of known hadrons. The u, c and t ("up", "charm" and "top" [4]) quark flavours have electric charge $\frac{2}{3}$ while d, s and b ("down", "strange" and "bottom") quarks have charge -$\frac{1}{3}$. Like the leptons, all of them are spin-$\frac{1}{2}$ fermions and are grouped into three generations of weak isospin. However, contrary to the leptons, weak interactions can also couple quarks belonging to different generations, even if with a suppressed probability (parametrized by the 3×3 Cabibbo-Kobayashi-Maskawa matrix elements).

Leptons, quark and the gauge bosons are the basis of our present understanding of the physical world [5]. Table A.1 lists these particles also showing the three generations in which they are grouped into.

Each elementary particle has an associated antiparticle, with same mass and spin but opposite charge. The photon, Z^0 and gluons are identical to their antiparticles.

[4] The top quark was the last to be discovered. Its experimental evidence was found in 1995 at the Tevatron by the CDF and D0 collaborations [6, 65] .

[5] This list is actually missing of the not yet detected *Higgs particle* (H^O) [24], the spin-0 boson foreseen by the Standard Model to explain the mechanism for spontaneously breaking of the $SU(2)_L \times U(1)_Y$ electroweak symmetry giving mass to the $W^{\pm}$ and Z^0 bosons and to quarks and leptons. Its discovery can rightly be considered a fundamental benchmark in further confirming the Standard Model validity.

A.2 Hadron Structure and Confinement

Two classes of hadrons are found: *baryons* and *mesons*. Baryons are fermions (half-integral spins) consisting of three quarks [6] while meson are bosons (integral spins) made of one quark and one antiquark. The quark configurations of hadrons establish their charge and other quantum numbers. These constituent quarks are usually referred as "*valence*" quark. From the laws of quantum mechanics, a fluctuating cloud (or "*sea*") of virtual gluons and neutral $q\bar{q}$ pairs is also expected to be present in each hadron [7]. Quarks and gluons inside a hadron are also generally referred as *partons*.

Quarks and gluons alone experience and transmit strong forces. According to the corresponding theory (QCD), these interactions are described in terms of *"color charge"*. Each quark is supposed to have one of three possible color: "red", "blue" and "green". Antiquarks carry the corresponding "anticolor", while gluons carry two labels: one color and one anticolor. Only colored particles can emit or absorb gluons so to conserve color at each particle-particle-gluon vertex. Leptons and the other gauge boson, being colorless, are not subject to strong interactions.

Because of color, the strong forces differ significantly from the electromagnetic ones even if both are transmitted by massless gauge bosons. For instance, differently from photons, gluons can couple directly to other gluons. The most remarkable consequence is the "*color confinement*": only colorless (color-neutral) states are allowed as physical hadrons. Neither quarks nor gluons can appear as isolated particles, they can only exist within (colorless) composite hadrons [8].

The strong coupling constant α_s also has a qualitatively different behaviour with Q^2 (where Q is the parton interaction four-momentum transfer giving the energy scale of the process) than the QED coupling constant α_{QED}.

[6] The nucleus constituents, protons and neutrons, are the lightest baryons.

[7] Such "sea" quarks do not affect the hadron quantum numbers but can play a role in high energy hadron collisions.

[8] The simplest colorless quark configurations are:

- three quarks with different colors (one red, one blue and one green), forming a baryon;
- quark + antiquark with the same color and anticolor (symmetrical superimposition of red + antired, blue + antiblue and green + antigreen), forming mesons.

These are exactly the valence quark combinations that describe the known baryons and mesons.

To a first approximation in $Q^2/{\Lambda^2}_{QCD}$ [66]:

$$\alpha_s(Q^2) = \frac{12\pi}{(33 - 2N_f)\ln\dfrac{Q^2}{{\Lambda^2}_{QCD}}} \tag{A.1}$$

N_f is the number of quark flavors involved which depends on Q^2 [9]. The parameter Λ_{QCD}, experimentally determinated to be about 0.2 GeV, gives the scale magnitude at which $\alpha_s(Q^2)$ becomes strong.

The theory is applicable only for $Q^2 \gg {\Lambda^2}_{QCD}$ for which $\alpha_s(Q^2)$ is small (*perturbative QCD*). When the momentum transfer becomes comparable with the masses of the light hadron (*i.e.* $Q^2 \simeq 1$ GeV), the constant $\alpha_s(Q^2)$ becomes large and perturbative theory is no more valid. This could give an indication that the confinement of quarks and gluons inside hadrons is actually a consequence of the growth of the coupling at the low scales. This large value of the coupling constant is the source of most of the mathematical complexities and uncertainties that still surround QCD calculations at low Q^2. On the other hand it is of great importance that this "running" coupling goes to zero in the infinite Q^2 limit. This fact, called *asymptotic freedom*, allows perturbation theory to be used in theoretical calculations to produce experimentally verifiable predictions for hard scattering processes (where a good reference value for α_s results to be $\alpha_s({m^2}_Z) = 1.118 \pm 0.003$ [67]).

The above considerations lead to guess that the effective interaction potential between a generic $q\bar{q}$ pair, could be approximated by a combination of a short-range asymptotic freedom contribution (due to single gluon exchange) and a long-range confining potential increasing with r, such as:

$$V(r) = -\frac{4}{3}\frac{\alpha_s}{r} + \lambda r \tag{A.2}$$

Such a potential allows, for instance, to give a very good account of the spectrum of mesons made of heavy quarks such as $\psi(c\bar{c})$ [68].

The effect of injecting energy into a hadron is not to separate the quarks, but to stretch the color lines of force among them which break with the creation of a new quark-antiquark pair and hence new hadrons [10]. So, when a quark or gluon recoils energetically from a hard collision, the broken lines of force behind it lead to a "*jet*" of hadrons according to a process usually referred as *fragmentation* or *hadronization*.

[9] In general, a quark i with mass m_i is expected to contribute only when $Q^2 \geq 4{m^2}_i$.

[10] This effect is somewhat similar to break a magnetic bar: as isolated N or S monopoles do not exist, a new pair of poles will be created.

A.2.1 Fragmentation

Colored quark and gluons inside hadrons can be regarded as free particles during a hard collision, but subsequently color forces will organize them into colorless hadrons with a fragmentation process which typically involves the creation of many additional quark-antiquark pairs.

Fragmentation is governed by soft non-perturbative processes that cannot be calculated from scratch. A semi-empirical description is so adopted which describes this complex process with the introduction of some hadronization models (*independent fragmentation model*, *string model*, *cluster model* and so on) implemented in explicit Monte Carlo constructions to be tuned on data.

From the empirical point of view, the parton fragmentation is usually parametrized by a "k-to-h *fragmentation function*" $D^h{}_k$. This function is defined so that $D^h{}_k(z)dz$ is the probability that the hadron h, produced in the fragmentation of the original fast hadron k, is carrying a fraction between z and $z + dz$ of the original hadron energy. For instance, in the independent fragmentation hypothesis, where it is supposed that partons fragment independently of each other, $D(z)$ is often parametrized in the form [66]:

$$D(z) = \frac{f(1-z)^n}{z} \tag{A.3}$$

where f is a constant and n is derived so to parametrize the observed behaviour at large z.

The integral of $D^h{}_k(z)$ over the full physical range of z ($z_{min} \to 1$, where $z_{min} = m_h/E_k$) is the average number of hadrons h (also referred as the *mean multiplicity of* h) in the complete jet originating from parton k.

A.3 A brief Jet History

Jets were first observed at e^+e^- colliders in 1975 when the c.m. energy reached 6-8 GeV at SPEAR [66]. When PEP and PETRA reached energies of 30-40 GeV, jets were found to be the dominant feature of hadron production.

Jets in hadron collisions were observed for the first time in early 80s at the CERN ISR colliding proton beam with a c.m. energy of 63 GeV [69]. The higher energy necessary to see jets in hadron collisions, can be explained considering that in such interactions the hard scattering occurs between constituents partons carrying only a fraction of the hadron momentum. Only with the 540-630 GeV c.m. energies of the CERN $Sp\bar{p}S$, jet production became the most striking feature of events with large transverse energy [70].

A.4 Jet Phenomenology

Since they cannot be isolated, quarks and gluons can be studied only indirectly. In particular, the study of a jet gives informations about the initiating parton as it consists of a group of energetic particles which are emitted spatially collimated along the original parton direction.

From the experimental point of view, in a typical collider experiment, jets appear as showers of electromagnetic and hadronic matter. They are observed as clusters of energy located in adjacent detector towers. Typically, a jet contains tens of neutral and charged pions, a lesser extent of kaons and very few light baryons (such protons and neutrons) each of which showers into multiple cells [11]. Each jet is so characterized by a *charged* fraction (mainly $\pi^{\pm}$), a *neutral electromagnetic* fraction (mainly photons from $\pi^0 \to \gamma\gamma$ decays) and a *neutral hadronic* one (mainly K_Ls and neutrons) (see fig. 2.1).

A short description of jet production in $p\bar{p}$ interactions will now be given.

A.4.1 Jets in $p\bar{p}$ Collisions

In $p\bar{p}$ collisions, jet production can be understood as a point-like collision of a quark or gluon from the proton and a quark or gluon from the antiproton. After colliding, because of fragmentation, these scattered partons manifest themselves as "sprays" of particles or "jets".

In general, any *inelastic* scattering between a proton and an antiproton can be described as an *elastic* collision between a single proton and antiproton constituent. The non-colliding constituents of the incoming proton and antiproton are called "beam fragments" or "spectators" and contribute to the "underlying event" (see Chapter 2).

In this parton model picture, the proton and antiproton can be seen as "broad band" beams of partons carrying a fraction x of the momentum of their parent hadrons. Transverse momenta of parton are neglected. *Parton distribution functions* (**PDF**) $f_i(x, \mu^2)$ are so introduced giving the probability for the i-th kind parton to have fractional momentum between x and $x + dx$ (μ is a factorization scale). Figure A.1 reports an example of such parton distribution functions [12].

[11] In the CDF experiment, for instance, a single jet "illuminates" roughly 20 calorimetric towers on average.

[12] Even if in principle they can be extracted from the QCD lagrangian, the small Q^2 scale, due to the bound states of partons inside hadrons, imply non-perturbative calculations. So, the only way to obtain these functions is to extract them from experimental data in Deep Inelastic Scattering (DIS) of leptons on nucleons. As data cover only a finite range

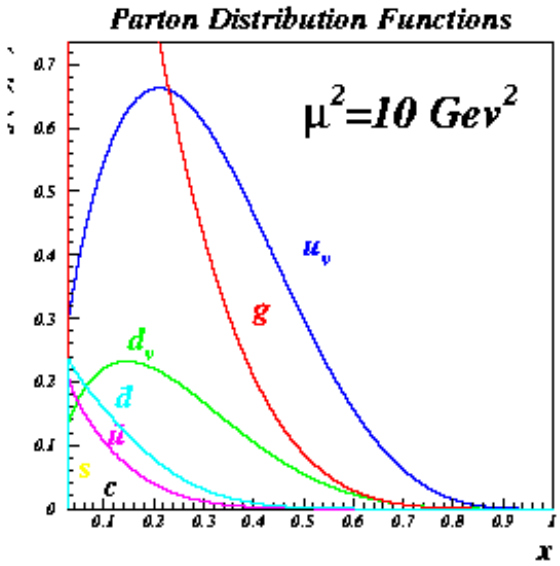

Figure A.1: *Parton distribution functions as obtained from a global next-to-leading-order QCD fit to DIS and other data at scale* $\mu^2 = 10 GeV^2$ *[71].*

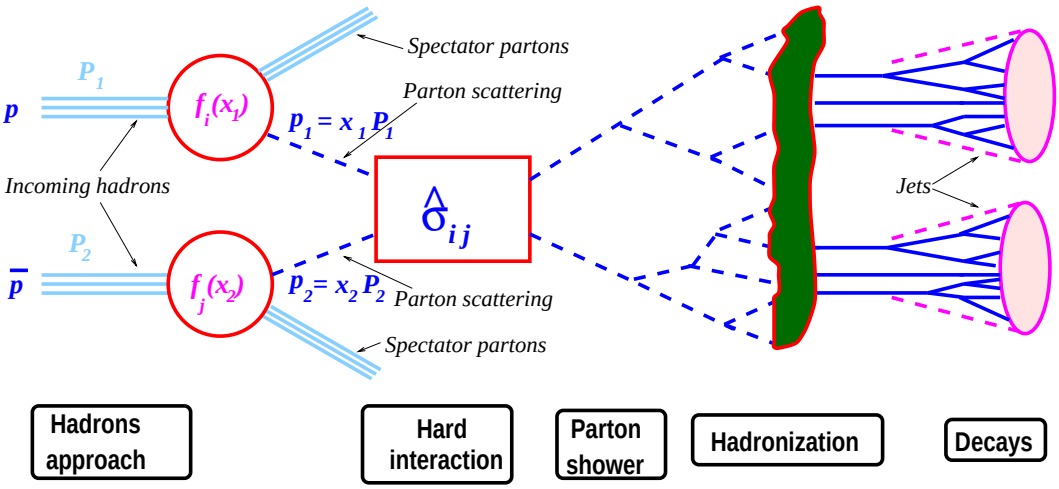

Figure A.2: *Hard "two-body" parton interaction producing a di-jet event in a proton-antiproton collision.*

of Q^2, the evolution of such functions with Q^2 is computed, using perturbation theory, with the *Altarelli-Parisi equation.*

Predictions for jet production are given by folding the parton distribution functions with perturbatively calculated "two-body" scattering cross sections $\hat{\sigma}_{i,j}$. Any cross section of interest is so calculated as:

$$\sigma = \sum_{i,j} \int f_i(x_1, \mu^2) f_j(x_2, \mu^2) \hat{\sigma}_{i,j}(x_1 P, x_2 P, \alpha_s({\mu_R}^2), Q^2/\mu^2, Q^2/{\mu_R}^2) dx_1 dx_2$$

Fig. A.2 shows the representation of an elementary "two-body" interaction between two partons in a $p\bar{p}$ collision producing a di-jet event.

The *factorization scale* μ discriminates whether a parton, inside the incoming hadron, takes part or not in the hard scattering: if the momentum of a parton is greater than the scale μ, it contributes to the short-distance cross section (as the partons i and j in fig. A.2); if its momentum is less than the scale μ, it is considered part of the hadron structure not involved in the hard interaction (spectator parton).

Next-to-leading order (NLO) calculations include one additional parton emission. For instance, a final state quark can radiate an additional gluon *before* hadronization (final state radiation) the entire scattering process being now proportional to ${\alpha^3}_s$. Depending on the proximity of the other partons, a third jet can be experimentally reconstructed.

Conclusions

During the last four years the CDF experiment at Fermilab underwent an extensive upgrade program both to allow its data acquisition to cope with the Run II Tevatron operative conditions and to improve the detector performance. The work presented in this thesis is part of this complex upgrade program.

Particle jets characterize the final state signature of many physics processes of interest for the CDF Run II physics program. Therefore, the resolution with which we succeed in reconstructing the energy of partons originating them is of crucial importance.

The many sources contributing to the jet energy resolution can be grouped into two categories, detector and physics effects. The incidence of detector effects has been extensively investigated and a new method to correct for low energy non-linearities and to improve the single charged particle response has been developed. Using track momenta from the central tracking system and Shower Max clusters, the calorimeter towers have been divided into different classes relying on the kind of particles hitting them and the tower energy has been reassigned according to this classification.

The new algorithm has been optimized and implemented in an offline analysis code allowing it to work in principle on every data sample. A test on different data samples, like γ-jet and di-jet data, has been performed by the candidate. The results are very encouraging, showing a jet energy resolution improvement up to $\sim 25\%$ relative to the standard jet corrections.

Acknowledgements

In the last three years several people helped me in developing the work exposed in this thesis. It is then a pleasure for me to thank all of them.

First, I would like to thank my advisor Prof. Giovanni Maria Piacentino who, trusting in my physicist skills (was he right ?), introduced me to the young Physics Group of the Faculty of Engineering at the University of Cassino. His friendly supervision and continuative encouragement were fundamental for me to work at best.

I am especially grateful to Prof. Stefano Lami, my "Di-jet Mass Group" advisor, with whom I worked in close contact almost daily ...in spite of the remarkable distance between Italy and New York. Many years of friendship unite us and my adventure in the "dirty" topic of the jet reconstruction study has further confirmed the intellectual and human pleasure of making research with him. His patient guide and useful advice were very precious in the work and preparation of this thesis.

Special thanks are also due to the other Di-jet Mass Group guys: Andrea Bocci, Davide Costanzo and Riccardo Paoletti, who initially introduced me into the "deep mysteries" of the new jet correction algorithm, and to Dr. Steve Kuhlmann, our "deus ex machina", whose help and hints were fundamental in fixing and solving a lot of problems connected with my analysis.

Finally, I would like to thank the University of Cassino for giving me the opportunity of getting the PhD degree in Industrial Engineering, and the INFN CDF-Pisa Group who hosted me for more than two years of activity in the Di-jet Mass Group and allowed me to perform part of it at the Fermilab laboratory, where I spent several months of pleasant work with my colleagues.

Bibliography

[1] D. H. Perkins, *"Introduction to High Energy Physics"*, Addison-Wesley (1987).

[2] Fermilab Beam Division, *"Run II Handbook"*, http://www-bd.fnal.gov/lug/runII_handbook/RunII_index.html.

[3] F. Abe *et al.*, CDF Collaboration, Nucl. Instr. Meth. A 271, 387 (1988).

[4] F. Abe *et al.*, CDF Collaboration, The Collider Detector at Fermilab *"A compilation of articles reprinted from Nuclear Instruments and Methods in Physical Research-A"* NORTH-HOLLAND - Amsterdam.

[5] CDF Collaboration, *"The CDFII Detector Technical Design Report"*, FERMILAB-Pub-96/390-E CDF (1996).

[6] F. Abe *et al.*, CDF Collaboration, Phys. Rev. D**50**, 2966 (1994).

[7] F. Abe *et al.*, CDF Collaboration, Phys. Rev. Lett **75**, 11 (1995).

[8] F. Abe *et al.*, CDF Collaboration, Phys. Rev. Lett. **81**, 5513 (1998).

[9] F. Abe *et al.*, CDF Collaboration, Phys. Rev. Lett. **81**, 2432 (1998).

[10] CDF Collaboration *"Proposal for Enhancement of the CDF II Detector. An Inner Silicon Layer and Time of Flight Detector"*, FERMILAB-Proposal-909 (1998).

[11] G. Apollinari, CDF Collaboration *"Test Beam Results from the CDF Plug Upgrade Calorimeter"*, FERMILAB-CONF-97/243-E (1997).

[12] L.Balka *et al.*, *"The CDF Central Electromagnetic Calorimeter"*, Nucl. Instrumen. Meth. **A262** 272-279 (1988).

[13] S.Bertolucci *et al.*, *"The CDF Central and Endwall Hadron Calorimeter"*,
Nucl. Instrumen. Meth. **A267** 301-314 (1988).

[14] F.Abe *et al.*, CDF Collaboration, Phys. Rev. D**48**, 2998 (1993).

[15] M. Dell'Orso, P. Giannetti, G. Punzi, *"The Central Calorimeter ϕ-Cracks and the Jet Energy Measurement"*,
CDF/MEMO/CDF/CDFR/896 (1989).

[16] G.C. Blazey *et al.*, *"RunII Jet Physics"*,
CDF/PUB/JET/PUBLIC/5293 (2000).

[17] B. Flaugher, J. Mueller, *"A Guide to JETCLU: the CDF Jet Cluster Algorithm"*,
CDF/DOC/JET/CDFR/1814 (1992).

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

[19] B.Flaugher, L.Keeble, *"New Jet Correction Function QDJSCO 2.0"*,
CDF/ANAL/JET/CDFR/1513 (1991).

[20] N. Eddy, *"New Relative Jet Corrections for Run Ia and Ib Data"*,
CDF/ANAL/JET/CDFR/3534 (1996).

[21] R. Field, D. Stuart *"The Underlying Event in Hard Scattering Processes"*,
CDF/PHYS/MIN-BIAS/PUBLIC/5244 (2000).

[22] A.Caner, L.Galtieri, J.Lys, W.Yao, A.Yagil, *"Jet Corrections for Top Mass Fitting"*,
CDF/ANAL/TOP/CDFR/2469 (1994).

[23] F.Abe *et al.*, CDF Collaboration, Phys. Rev. Lett. **73**, 225 (1994).

[24] P. W. Higgs, Phys. Rev. Lett. 12 132 (1964).

[25] Aleph Collaboration, CERN-EP/2000-138 (2000).
(Submitted to Phys. Lett. B).

[26] L3 Collaboration, CERN-EP/2000-146 (2000).
(Submitted to Phys. Lett. B).

[27] Opal Collaboration, CERN-EP/2000-156 (2000).
(Accepted by Phys. Lett. B).

[28] Delphi Collaboration, CERN-EP/FINAL (2000).
(Accepted by Phys. Lett. B).

[29] M. Carena *et al.*, *"Report of the Tevatron Higgs Working Group"*, FERMILAB-Conf-00/279-T; hep-ph/0010338 (2000).

[30] CMS Collaboration, *"CMS Technical Proposal"*, CERN/LHCC/94-38, LHCC/P1 (1994).

[31] ATLAS Collaboration, *"ATLAS Technical Proposal"*, CERN/LHCC/94-43, LHCC/P2 (1994).

[32] A. Stange, W. Marciano, and W. Willenbrock, Phys. Rev. D**49**, 1354 (1994); D**50**, 4491 (1994)

[33] S. Mrenna and G. Kane, CALT-68-1938 (1994).

[34] A. Balyaev, E. Boos and L. Dudko, Mod. Phys. Lett. **A10**, 25 (1995).

[35] S. Kuhlmann, *"Will We Find the Higgs in Run II ?"*, CDF/PHYS/ELECTROWEAK/PUBLIC/3342 (1995).

[36] S. Kuhlmann D. Costanzo R. Paoletti S. Lami, *"Particle Level Study of $W \to q\bar{q}$ and Jet Algorithms"*, CDF/ANAL/JET/CDFR/4259 (1997).

[37] A.Bocci, D. Costanzo, S. Kuhlmann, S. Lami and R. Paoletti, *"Jet Energy Resolution Using Calorimeter, Tracking and Shower Max Information"*, CDF/ANAL/JET/CDFR/4681 (1998).

[38] A. Bocci, *"A Study of Jet Energy Measurement at CDF"*, Laurea Thesis, CDF/THESIS/JET/PUBLIC/4668 (1998).

[39] S.Behrends, M.D.Shapiro, E.Pare, *"Single Pion Response in the Central Calorimeter"*, CDF/ANAL/CAL/CDFR/1066 (1989).

[40] B. Cousins, *"On Using CTC Track Momenta to Improve Jet Energy Resolution"*, CDF/MEMO/JET/PUBLIC/1567 (1991).

[41] S. Behrends *et al.*, *"Charged Track Correction to Jets and Effect on Jet Energy Resolution"*, CDF/ANAL/JET/CDFR/1015 (1989).

[42] S.dell'Agnello, *"Effect of the Charged Track Correction on Jet Energy Resolution in 88/89 Data"*,
CDF/ANAL/CAL/CDFR/1063 (1989).

[43] G.Punzi, *"Invariant Mass Spectrum of Jet-Jet Events in Proton-Antiproton Collision at* $\sqrt{s} = 1.8$ *TeV"*,
Ph.D. Thesis, Scuola normale Superiore, PISA (1990).

[44] S. Behrends, S. Kuhlmann, D. Amidei, F. Rimondi, *"Response of the Central Calorimeter to Low Energy Charged Particles"*,
CDF/ANAL/CAL/CDFR/0583 (1987).

[45] N. Solomey, A. Wicklund *"Study of Electromagnetic Showers in the CDF Central Electromagnetic Calorimeter Strip Chamber"*,
CDF/DOC/CDF/PUBLIC/0247 (1987).

[46] S.Bertolucci *et al.*, *"Response Maps of the CDF Central Electromagnetic Calorimeter with Electrons"*,
Nucl. Instrumen. Meth. **A267** 315-329 (1988).

[47] S. Lami, G. Latino, *"An Update of the Jet Energy Resolution Using Calorimeter, Tracking and Shower Max Information"*,
CDF/ANAL/JET/CDFR/5192 (1999).

[48] A. Bocci, S. Kuhlmann, S. Lami, G. Latino, *"Latest Results on the Jet Energy Resolution Using Calorimeter, Tracking and Shower Max Information"*,
CDF/ANAL/JET/CDFR/5416 (2000).

[49] S. Kuhlmann, *"Z→ee Peak Using Central Crack Chambers"*,
CDF/ANAL/JET/CDFR/4798 (1998).

[50] A.A.Bhatti, *"Isolated Charged Particle Response in 1992-93 Minimum Bias Data"*,
CDF/ANAL/JET/CDFR/2838 (1994).

[51] S. Kuhlmann, *"Charged Particle Energy Depositions in the Central Electromagnetic Calorimeter"*,
CDF/ANAL/JET/CDFR/5238 (2000).

[52] P. Bagnaia *et al.*, UA2 Collaboration, Phys. Lett. 144**B**, 283 (1984).

[53] Rick St. Denis, *"Dijet Angular Distributions in* $p\bar{p}$ *Collisions at The Fermilab Tevatron"*, CDF/ANAL/JET/GROUP/0805 (1988).

[54] D. Partos, S. Kuhlmann, J. Lamoureux, *"Run Ib Prompt Photon Cross Section"*,
CDF/ANAL/JET/CDFR/4910 (1999).

[55] A. Bocci, D. Costanzo, S. Kuhlmann, S. Lami, G. Latino, R. Paoletti, *"Jet Studies in γ+Jet Sample Using New Energy Reconstruction Algorithm"*,
CDF/ANAL/JET/CDFR/5179 (1999).

[56] E. Kovács, *"Jet Energy Resolutions Froms Dijet Balancing"*,
CDF/ANAL/JET/CDFR/2608 (1994).

[57] D. Brown, *"A Study of Dijet P_T Balancing Using 1987 Data"*,
CDF/ANAL/JET/CDFR/0835 (1988).

[58] A. Baden, M. Curatolo, B. Esposito, S. Miscetti and P. Sestini, *"The Missing E_T Resolution for '87 Run Data"*,
CDF/ANAL/MISSING-ET/CDFR/0715 (1988).

[59] S. Kuhlmann, *"Timing Information For Central Photons and the ee$\gamma\gamma$ MET Event"*,
CDF/ANAL/EXOTICS/CDFR/4256 (1997).

[60] S. Kuhlmann, *"Cosmic Ray Backgrounds to Direct Photons (and Jets)"*,
CDF/ANAL/JET/CDFR/980 (1989).

[61] L. Glashow, Nucl. Phys. **22**, 579 (1961).

[62] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967).

[63] G. Arnison *et al.*, UA1 Collaboration, Phys. Lett. 122**B**, 103 (1983).

[64] M. Banner *et al.*, UA2 Collaboration, Phys. Lett. 122**B**, 476 (1983).

[65] S. Abachi *et al.*, D0 Collaboration, Phys. Rev. Lett. **74**, 2632 (1995).

[66] V. Barger, R. Phillips, *"Collider Physics"*, Addison-Wesley (1987).

[67] Particle Data Group, Phys. Rev. **D 54** (1996).

[68] P. Collins, A. Martin, *"Hadron Interactions"*, Adam Hilger LTD (1984).

[69] V. Cavasinni, Riv. Nuovo Cimento **13**, 1 (1990).

[70] P. Bagnaia *et al.*, Z. Phys. **C20**, 117 (1983).

[71] A. Martin *et al.*, Phys. Rev. D. **50** 6734 (1994).