A Measurement of the Top Quark Mass in the Lepton + Jets Channel using the Ideogram Technique at DØ

by

Vivek S. Parihar

M.Sc., IIT Kanpur; India, 2002

A dissertation submitted in partial fulfillment of the requirements for
the Degree of Doctor of Philosophy
in The Department Of Physics at Brown University

PROVIDENCE, RHODE ISLAND

May 2013

This dissertation by Vivek S. Parihar is accepted in its present form by
The Department Of Physics as satisfying the
dissertation requirement for the degree of
Doctor of Philosophy

Date_____ _____

Professor Ulrich Heintz, Advisor

Recommended to the Graduate Council

Date_____ _____

Professor David Cutts, Reader

Date_____ _____

Professor Meenakshi Narain, Reader

Approved by the Graduate Council

Date_____ _____

Dean Peter M. Weber
Dean of the Graduate School

Acknowledgments

The work in this dissertation has been made possible due to the invaluable guidance of Prof. Ulrich Heintz. I 'am extremely thankful for his constant and untiring supervision. I would also like to thank the Brown University Physics Department for providing me support at various stages. I 'am thankful to my reading committee.

Much of the work in this thesis was carried out at Fermilab and I would particularly like to thank several of my DØ colleagues. At the seminal stages of the project, I had received great help from Dr. Michele Weber and Dr. Pieter Houben. Throughout the course of my stay at Fermilab, fellow students, postdocs and officemates have extended their helping hands to me. I would like to thank a few (not in any particular order): Dr. DooKee Cho, Dr. Daniel Boline, Dr. Amitabha Das, Dr. Lidija Zivkovic, Dr. Shabnam Jabeen and Dr. Zhenyu Ye. The facilities provided at DØ were conducive for research and I 'am thankful to the management for that.

It would not be justified, if I didn't thank my teachers and mentors who have taught me in various capacities and a special mention in this regards would go to Prof. Allan Widom and Prof. Yogendra Srivastava. In the same spirit, I would also like to thank the Boston University Physics Department for accomodating me as a graduate student from Jan 2007-Aug 2009.

Last but not the least, I 'am grateful to my family members for their love, patience and support. I wish to acknowledge my late father, who passed away last year, for his life has been an inspiration to me.

Abstract of " A Measurement of the Top Quark Mass in the Lepton + Jets Channel using the Ideogram Technique at DØ," by Vivek S. Parihar, Ph.D., Brown University, May 2013

This thesis presents a measurement of the mass of the top quark. The top quarks are produced through proton-antiproton collisions with $\sqrt{s}$=1.96 TeV at the Fermilab Tevatron collider using the DØ detector. The method employed for the measurement is called the Ideogram Technique. The events used in the measurement are selected such that the top quark decay signatures (as seen in the detector) are at least one lepton (electron or muon), four or more jets with at least one of them tagged as originating from a b-quark and missing transverse momentum. These events are then fitted using a kinematic fitter. The event-by-event likelihood is calculated using templates that depend on the kinematically reconstructed top quark mass. The events are weighted by their probability to be signal or background, using topological information. The measurement of the top quark mass is $175.7 \pm 1.98\ GeV/c^2$.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER **One**

Top Quark: Theoretical Perspective

1.1 Standard Model of Particle Physics

The Standard Model (SM) entails our current understanding of particle physics. It has been laid on foundation of experimental discoveries and theoretical advancement and relates to the observed particles and their interactions. It describes two classes of matter: spin $\frac{1}{2}$ fermions and spin 1 bosons (Fig. 1.1) that carry three fundamental forces: the electromagnetic force, the weak force and the strong force with the first two forces unified into electroweak force. The SM represents gauge group $SU(3) \times SU(2) \times U(1)$. A gauge theory is one that possesses invariance under a set of local transformations i.e. transformations whose parameters are space-time dependent. The Standard Model has been phenomenologically successful but it does not describe a complete theory of fundamental interactions e.g. gravity is not included, neither is the viable dark matter that possesses the properties deduced from observational cosmology etc.

Fermions are identified as quarks and leptons as shown in Fig. 1.1. The quarks carry color charge and thus interact via the strong force (as well as electromagnetic and weak force). The six quarks (and corresponding anti quarks) are classified into three genera-

tions with increasing mass for each generation. Every generation contains one quark with charge $(+2/3)\,e$ and one quark with charge $(-1/3)\,e$, which are ofter referred to as the up type and down type respectively. Leptons are also classified into three generations with each generation having a charged lepton with charge (-e) and a corresponding uncharged neutrino.

The masses of the particles (determined experimentally) are also shown in the Fig. 1.1. The top quark is the heaviest of all the particles. The mass of the quarks is dependent on the renormalization scheme. This is discussed at the end of this dissertation.

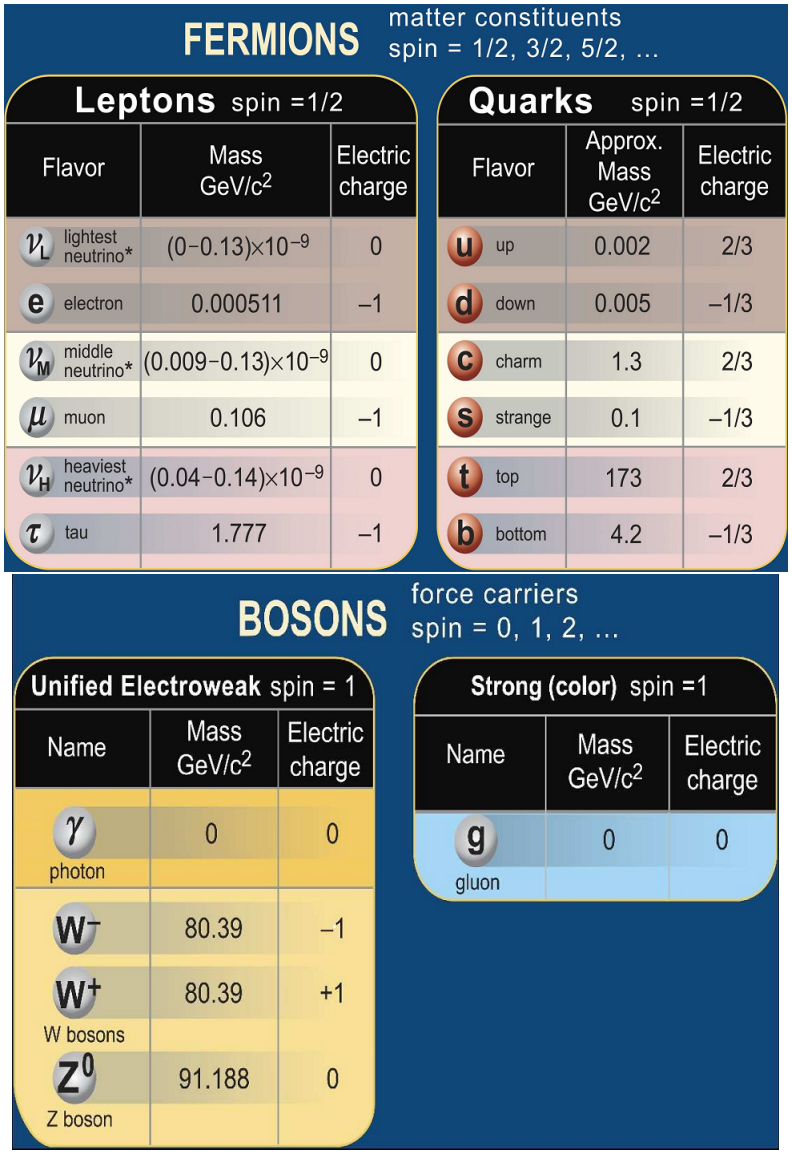


Figure 1.1: Constituents of matter and the force carriers in the Standard Model.

1.1.1 Electroweak Theory

The electroweak force is represented by $SU(2) \times U(1)_Y$ gauge group. This is spontaneously broken and results in separate electromagnetic and weak forces. The process of spontaneously broken symmetry is discussed in Sec. 1.1.3. The carrier of electromagnetic force is the massless spin-1 particle γ (photon). This feature of the photon makes electromagnetic

force, a long range force. The carriers of weak force are the $W^{\pm}$ and Z bosons, which are massive with masses 80.385± 0.015 GeV/c^2 and 91.1876±0.0021 GeV/c^2 respectively. This feature of the carriers makes the weak force short range. Weak force can change the flavor of quarks. The amplitude of mixing between quark flavors mediated by the W boson is described by the CKM matrix (Cabibbo-Kobayashi-Maskawa). The terms in the CKM matrix are complex and the matrix is required to be unitary in the Standard Model. The matrix is parametrized by three mixing angles and a CP-violating phase.

1.1.2 Quantum Chromodynamics

Quantum Chromodynamics (QCD) is the theory of strong interactions. It is a non-abelian theory with the gauge group representation of SU(3). The quarks are described by a field ψ_i where $i = 1,2,3$. The quantum number associated with the label i is called color. The eight gauge bosons which are represented by the generators of the gauge group are called gluons. These are taken to be carriers of the strong force. The coupling for strong interactions is the QCD gauge coupling, g_s. It is usually redefined in terms of α_s as :

$$\alpha_s = \frac{g_s^2}{4\pi} \tag{1.1}$$

This coupling is energy dependent and due to the non-abelian nature of QCD, α_s decreases as energy increases. Thus, when higher order perturbative calculations are performed in the Feynman formalism, the loop diagrams have the effect of "dressing" the couplings. This is shown in Fig. 1.2 for one-loop corrections to the gluon propagator. These diagrams contain ultraviolet divergences and need to be renormalized e.g. by subtracting at some renormalization scale μ. If the squared momenta of all the particles coming into the vertex is Q^2, then

$$\alpha_s(Q^2) = \alpha(\mu^2) - \alpha_s(\mu^2)^2\beta_0 \ln(Q^2/\mu^2) + ... \tag{1.2}$$

The coefficient β_0 is calculated to be :

$$\beta_o = \frac{11N_c - 2n_f}{12\pi} \tag{1.3}$$

where, N_c is the number of colors (=3), n_f is the number of active flavors i.e. the number of flavors whose mass threshold is below the momentum scale Q^2. A more precise analysis shows that the effective coupling obeys the differential equation

$$\frac{\partial \alpha_s(Q^2)}{\partial ln(Q^2)} = \beta(\alpha_s(Q^2)) \tag{1.4}$$

The solution of which depends on the boundary value and is often quoted at the Z boson mass

$$\alpha_s(M_Z^2) = 0.1161 \pm 0.004$$

This is one of the free parameters of the SM. The running of α_s is shown in Fig. 1.3.

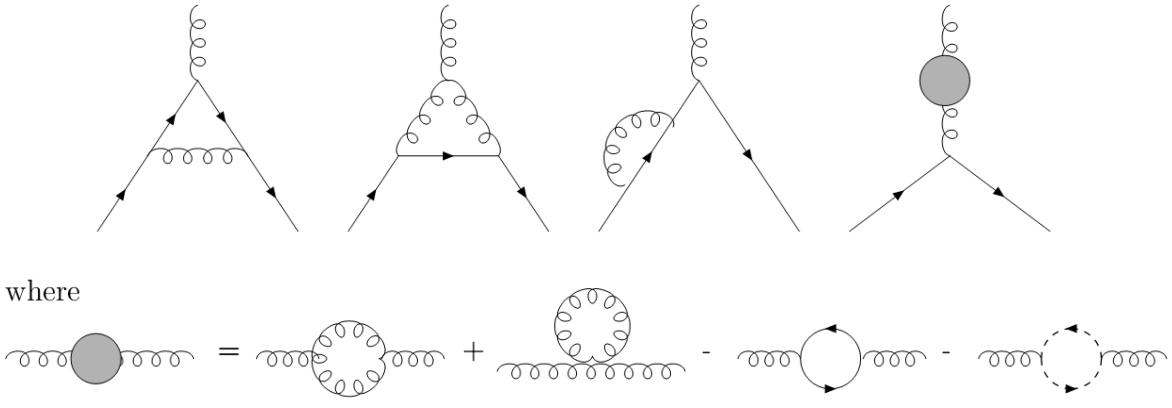


Figure 1.2: One loop corrections to gluon propagator.

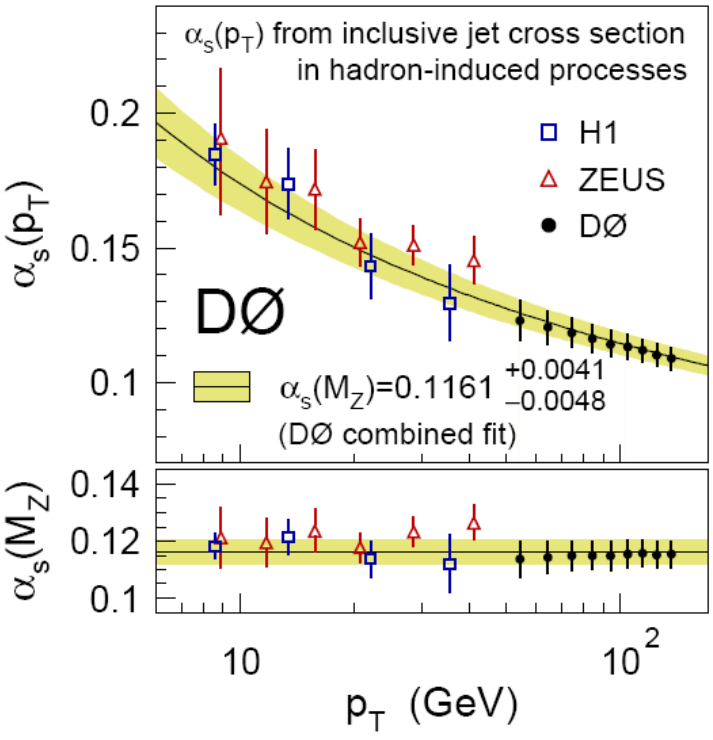


Figure 1.3: Running of strong coupling constant.(Ref. [2])

Gauge invariance requires that the gauge coupling for the interaction between gluons must be exactly the same as gauge coupling for interaction between quarks and gluons.

Converse to the situation described above, the coupling grows when we go to larger distances and therefore the complicated system of gluon exchanges, which leads to the binding of quarks (anti quarks) inside hadrons leads to stronger and stronger binding as the quarks are pulled apart. This process is called confinement. Thus the only free particles that can be observed at macroscopic distances from each other are color singlets. This process is not completely understood, nevertheless Monte Carlo programs have been developed that can simulate the hadronization in such a way that the results of short distance perturbative calculations at the level of quarks and gluons can be confronted with experiments.

1.1.3 Mass Generation mechanism

In 1950's Yang and Mills extended the idea of gauge invariance to local non-abelian transformations such as SU(2). In this case, one needs a set of massless gauge fields. In order that such a gauge theory could be applied to weak interactions, particles that transform into one another under weak interactions were looked at e.g. u and d quark. The three gauge bosons were interpreted as the $W^{\pm}$ and Z bosons that mediate the weak interactions. The weak interactions were known to be short range and had to be mediated by massive vector bosons, whereas Yang-Mills fields are required to be massless in order to preserve the gauge invariance. This paradox was resolved by the electroweak symmetry breaking mechanism (Ref. [8, 9, 10]) and is often referred to as the mechanism of spontaneous symmetry breaking.In this scenario, one starts with a theory that possesses the required gauge invariance but the ground state of the theory is not invariant under gauge transformations. The breaking of the invariance arises in the quantization of the theory whereas the Lagrangian contains terms that are gauge invariant. One of the consequences of this is that the gauge bosons acquire a mass. It also lead to a renormalizable theory whereby infinities that arise out of the higher order calculations can be reabsorbed into the parameters of the Lagrangian. Another consequence of this mechanism is the existence of a scalar (spin 0) particle, the Higgs boson.

Thus the union of QCD and the electroweak gauge theory is known as the Standard Model. It has nineteen fundamental parameters, most of which are associated with the masses of the gauge bosons, the quarks, leptons and Higgs. These are not all independent and because of the renormalizable nature of the theory, perturbative calculations can be performed at higher order that can predict the cross sections and decay rates for both strongly and weakly interacting particles. The predictions and data are becoming more and more precise, making the tests of the Standard Model increasingly stringent.

1.2 Top Quark's role in precision electroweak analyses

The gauge, matter and Higgs sectors of the Standard Model depend on five parameters: the three gauge couplings, g_S, g, g' and the Higgs-field vacuum expectation value and self interaction, v and λ. At tree level, all electroweak quantities depend on three of the parameters g, g' and v. The three best measured electroweak quantities to determine these parameters at tree level are :

$$\alpha = \frac{1}{4\pi}\frac{g^2 g'^2}{g^2+g'^2} = \frac{1}{137.03599976}$$

$$G_F = \frac{1}{\sqrt{2}v^2} = 1.16637 \times 10^{-5} GeV^{-2}$$

$$M_Z = \frac{1}{2}\sqrt{g^2+g'^2}v = 91.186\,GeV$$

The value of α is extracted from low energy experiments, G_F is extracted from the muon lifetime and M_Z is measured from LEP. At tree level, the W boson mass can be written as :

$$M_W^2 = \frac{1}{4}g^2v^2 = \frac{1}{2}M_Z^2\left(1+\sqrt{1-\frac{4\pi\alpha}{\sqrt{2}G_F M_Z^2}}\right) \tag{1.5}$$

By defining $s_W^2 \equiv 1 - \frac{M_W^2}{M_Z^2}$, which is also referred to as the on-shell definition of $sin^2\theta_W$, one can write M_W at tree level as :

$$M_W^2 = \frac{\frac{\pi\alpha}{\sqrt{2}G_F}}{s_W^2} \tag{1.6}$$

At one loop level, this is modified to :

$$M_W^2 = \frac{\frac{\pi\alpha}{\sqrt{2}G_F}}{s_W^2\,(1-\triangle r)} \tag{1.7}$$

where $\triangle r$ contains the one-loop corrections. The top quark contributes to $\triangle r$ via the one loop diagrams shown in Fig. 1.4 which contribute to the W and Z masses:

$$(\triangle r)_{top} \approx -\frac{3G_F m_t^2}{8\sqrt{2}\pi^2}\frac{1}{t_W^2} \tag{1.8}$$

where $t_W^2 \equiv tan^2\theta_W$. This one-loop correction depends quadratically on the top quark mass.

The Higgs boson also contributes to $\triangle r$ via the loop diagrams shown in Fig. 1.5

$$(\triangle r)_{Higgs} \approx \frac{11G_F M_Z^2 c_W^2}{24\sqrt{2}\pi^2}\ln\frac{m_h^2}{M_Z^2} \tag{1.9}$$

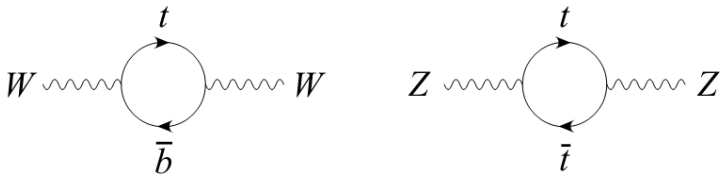


Figure 1.4: Virtual top quark loops contributing to W and Z bosons.

Figure 1.5: Virtual Higgs-boson loops contributing to the W and Z bosons.

where $c_W^2 \equiv cos^2\theta_W$. This correction depends logarithmically on the Higgs-boson mass. Hence to predict M_W at one loop level, one needs not just α,G_F, M_Z but also $m_t$and m_h.

In other words, prediction of Higgs boson mass requires α, G_F, M_Z and m_t, M_W. Fig. 1.6 shows the plot of M_W vs. m_t indicating lines of constant Higgs mass. The green ellipse indicates the 68 % confidence level measurements of M_W and m_t with $m_W = 80385 \pm 15\,MeV$ and $m_t = 173.2 \pm 0.9\,GeV/c^2$. Mass of Higgs boson is predicted to be less than $152\,GeV/c^2$ at 95% confidence level.

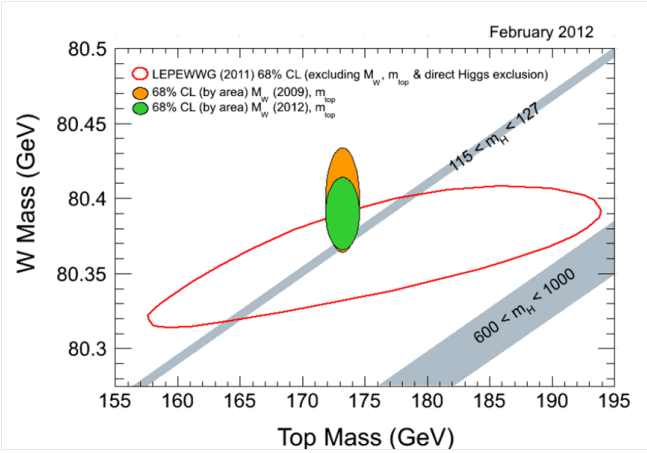


Figure 1.6: Lines of constant Higgs mass on a plot of M_W vs m_t The green ellipse shows 68 % CL direct measurement of the most recent M_W and m_t. (Ref. [3])

1.3 Top quark production and decay

At $\sqrt{s} = 1.96$ TeV center of mass energy of the Tevatron, the top quark is primarily produced in $t\bar{t}$ pairs which proceeds through either $q\bar{q}$ annihilation (85%) or gg fusion (15 %).

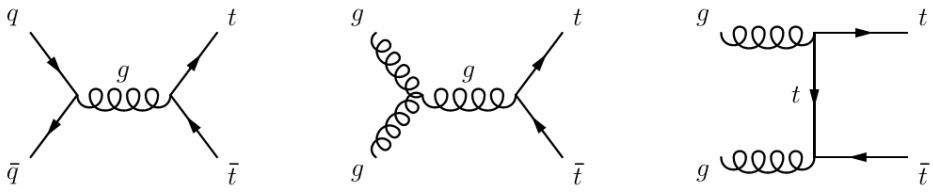


Figure 1.7: Feynman diagrams for $t\bar{t}$ production via $q\bar{q}$ annihilation and gg fusion.

The production process for the top pair production proceeds via the strong interaction and at high energies can be described by using perturbation techniques. The structure of the colliding proton and antiproton can be resolved into two parts. The hard scattering process that takes place between the constituents of the colliding hadrons (quarks/antiquarks or gluons) and the internal structure of the colliding proton and antiproton are responsible for the top quark production. In practice, a factorization scale μ_F^2 is introduced to separate the hard scattering partonic cross section from the modeling of the constituents of proton and antiproton. The modeling of the constituents is independent of the hard-scattering process and parton distribution functions (PDFs) $f_{PDF}^a(x,\mu_F^2)$ are introduced that describe the probability density to find a parton a with a longitudinal momentum fraction x inside a colliding proton or antiproton. The PDFs are determined from the fits to the experimental data. In this analysis, CTEQ6M (Ref. [11]) with a factorization scale of $\mu_F^2 = (175\,GeV)^2$ was used. The physical observables do not depend on the factorization scale, the overall dependence on it remains if the calculations are not done to an infinite order in the perturbation theory via the PDFs.

The leading order Feynman diagrams for the hard-scattering process of $t\bar{t}$ are shown in Fig. 1.7. If the contributions from higher order diagrams are taken into account, the renormalization of divergent quantities becomes necessary. This leads to the introduction of renormalization scale μ_R^2. In this analysis, the factorization and renormalization scales were chosen to be the same.

The $t\bar{t}$ production cross section in hadron collisions can be written down by integrating over all possible initial state parton momenta and then summing over all contributing initial state parton species.

$$\sigma(P_1,P_2) = \sum_{a_1,a_2} \int dx_1 dx_2 \, f^{a_1}_{PDF_1}(x_1,\mu_F^2)\, f^{a_2}_{PDF_2}(x_2,\mu_F^2)\, \hat{\sigma}\left(x_1P_1, x_2P_2, \alpha_s(\mu_F^2), \frac{Q^2}{\mu_R^2}\right) \tag{1.10}$$

In the expression (1.10), $\hat{\sigma}$represents the partonic cross section, $P_1$and P_2 are the momenta of the incoming hadrons and a_1, a_2 represent the parton species that initiate the hard nteraction. The $t\bar{t}$ cross section measured at DØ is found to be $7.78^{+0.77}_{-0.64}\,pb$ (Ref. [12]) and is in good agreement with the theoretical prediction of approximate next-to-next-to leading order (NNLO) value of $7.48^{+0.56}_{-0.72}\,pb$ (Ref. [13, 14, 15, 16]).

The top quark decays weakly into a W boson and a down type quark. The CKM matrix element $|V_{tb}|$ is nearly equal to one. Since, $(t \to Wb)/(t \to Wq) = |V_{tb}|^2/(|V_{tb}|^2 + |V_{ts}|^2 + |V_{td}|^2)$, it can be seen that the top quark produced will decay to a b quark nearly 100 % of the time. When measuring the top mass, we assume that the $t\bar{t}$ will always decay into $W^+bW^-\bar{b}$. The top quark decay width (Ref. [17]) is much larger than Λ_{QCD} and so the decay process happens before the top quarks can hadronize. Thus, the decay mode is determined by the decays of the two W bosons. Thhe three modes that are commonly studied are dilepton, lepton+jets (l+jets) and all-jets topologies. These do not include events with one or more tauonic W boson decays which are not trivial to reconstruct and provide less information about the mass. Thus the word lepton represents electron aor muon when referring to the decay topologies of the top quarks.

The branching fractions of various topologies are shown in Fig. 1.8.

- Dilepton mode: About 5% of the $t\bar{t}$ events have W bosons that decay into an electron or a muon plus the corresponding neutrino. These are called dilepton events and are characterized by two oppositely charged isolated energetic leptons, two b quark jets and missing transverse energy due to the presence of the neutrinos in the W decays.

- Lepton+jets topology: The lepton+jets events are thos 29% of the $t\bar{t}$ events that have one $W \to e\nu$ or $W \to \mu\nu$ and one hadronically decaying W boson. Thses events are characterized by one isolated energetic lepton, atleast four energetic jets (two of which are b jets) and missing transverse energy. The main physics background comes from the events where a leptonically W boson is produced in association withfour jets. Multijet background where one jet mimicks an isolated electron, also plays a role. In lepton+jets events, the transverse momentum components of the neutrino can be obtained from the missing transverse momentum and the vent kinematics is over-constrained if one assumes equal masses of the top and antitop quarks and invariant $l\nu$ and $q\bar{q}'$ masses equal to the W boson mass.

- All jets events: In about 46% of the $t\bar{t}$events, both W bosons decay hadronically and hence the event topology in this case is characterized by 6 energetic jets, no charged leptons and no significant missing transverse energy.

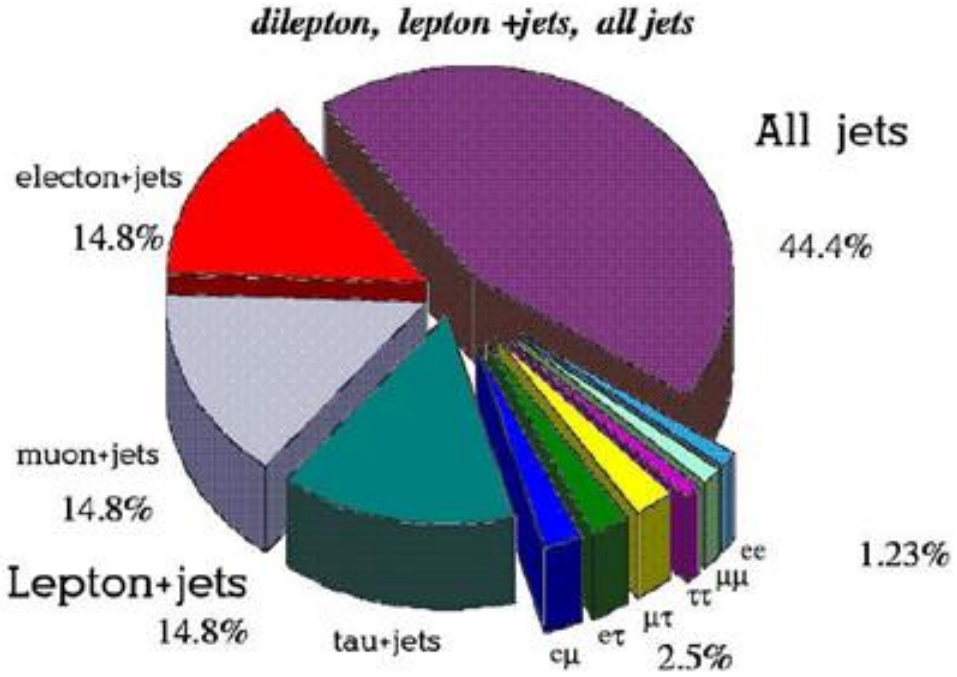


Figure 1.8: $t\bar{t}$ decay topologies based on the branching fractions of the W boson.

CHAPTER **Two**

The Experiment

2.1 Fermilab Tevatron Collider

The Tevatron collider (Ref.[18]) is located at the Fermi National Accelerator Laboratory in Batavia (IL, USA), where the collisions of the protons and anti protons at a center of mass energy of 1.96 TeV, have taken place. It is a ring of 6.3 km in circumference which is a part of a sophisticated accelerator complex. The layout of the accelerator complex is shown in the Fig.2.1 . The particle acceleration goes through various stages :

- 750 keV Cockroft-Walton pre-acceleration: The hydrogen gas is ionized by knocking of the peripheral electron and the H^- ions thus produced are accelerated, before being passed through a 150 meter long linear accelerator (LINAC). The ions in this system are accelerated to 400 MeV using oscillating electric fields. After the removal of the negatively charged electrons, the protons, thus produced, are passed on to the *Booster*.

- Booster: This is the first in the series of synchrotons involved in the step-by-step acceleration. The protons are accelerated to 8 GeV by going around in circular orbits.

- Main Injector: This can perform multiple tasks.
  - Accelerates protons to 150 GeV.
  - Accelerates protons to 120 GeV and send them to *Antiproton source* where the proton beam is targeted on a nickel target that produces, amongst many particles, the anti protons, which are then stored in an *accumulator ring* and sent back to Main Injector for acceleration.
  - It injects the protons and anti protons into the Tevatron for the final acceleration.
- Tevatron: It accelerates the particles from the Main Injector to 980 GeV. The proton and anti proton beams are accelerated in opposite directions and made to collide at two crossing points, the DØ (Ref.[19]) and the CDF (Ref.[20]) detectors. The net resulting center of mass energy is 1.96 TeV. The particle beams are kept on their trajectory and in focus through a series of dipole and quadrupole magnets. The proton and anti proton beams are counter circulated in 36 bunches each with a time pulse of 396 ns. Each bunch has about 10^{11} protons and 10^{10} anti protons resulting in the instantaneous luminosity of about $4x10^{32}$ $cm^{-2}s^{-1}$.

The last collisions at the Tevatron took place on 30 September 2011 which is when the operations were ceased, due to the fact that the machine has been superseded in the center of mass energy by the *Large Hadron Collider*, which uses protons on protons for collisions.

2.2 The DØ Detector

The DØ detector is a general purpose detector that has been used to identify and record interesting collision events. The design of the detector is based on accomplishing the physics goals of the experiment. These include, but are not limited to, study of the top quark, the W and the Z bosons and search for the elusive Higgs boson.

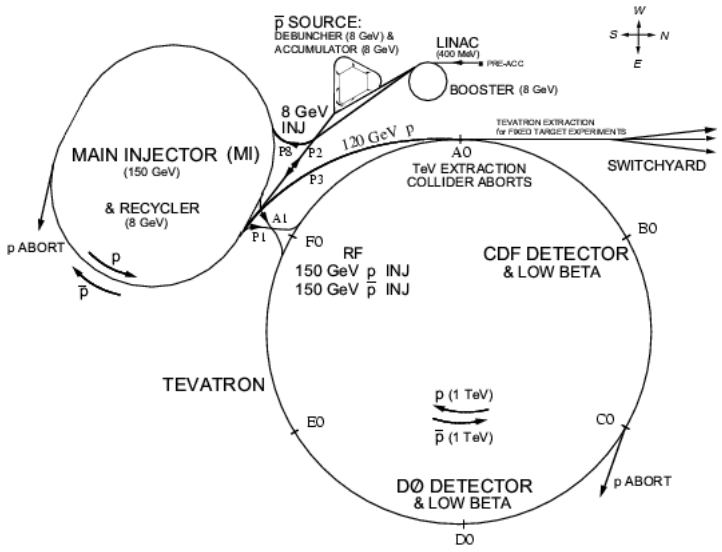


Figure 2.1: Layout of the accelerator complex at Fermilab

The schematic design of the full detector is shown in Fig. 2.2. Surrounding the beam pipe, the detector has the following main components.

- Central Tracker enclosed inside a 2 Tesla magnetic field : This system acts as a magnetic spectrometer. Interactions with the material are minimal here but just enough to identify the track of the charged particles. The magnetic field enables measurement of the energy and sign of the charged particle, which in turn facilitates, electron identification and calorimeter calibration.

- Calorimeter : This system consists of three sampling calorimeters which are primarily made of uranium/liquid-argon and an intercryostat detector. Maximum interactions with the material happen here and most particles produced in collisions are "stopped" in this region except for muons and neutrinos.

- The Muon system : A dedicated detector to detect and measure the momentum of the muons.

A brief discussion of these components will follow in the next sections.

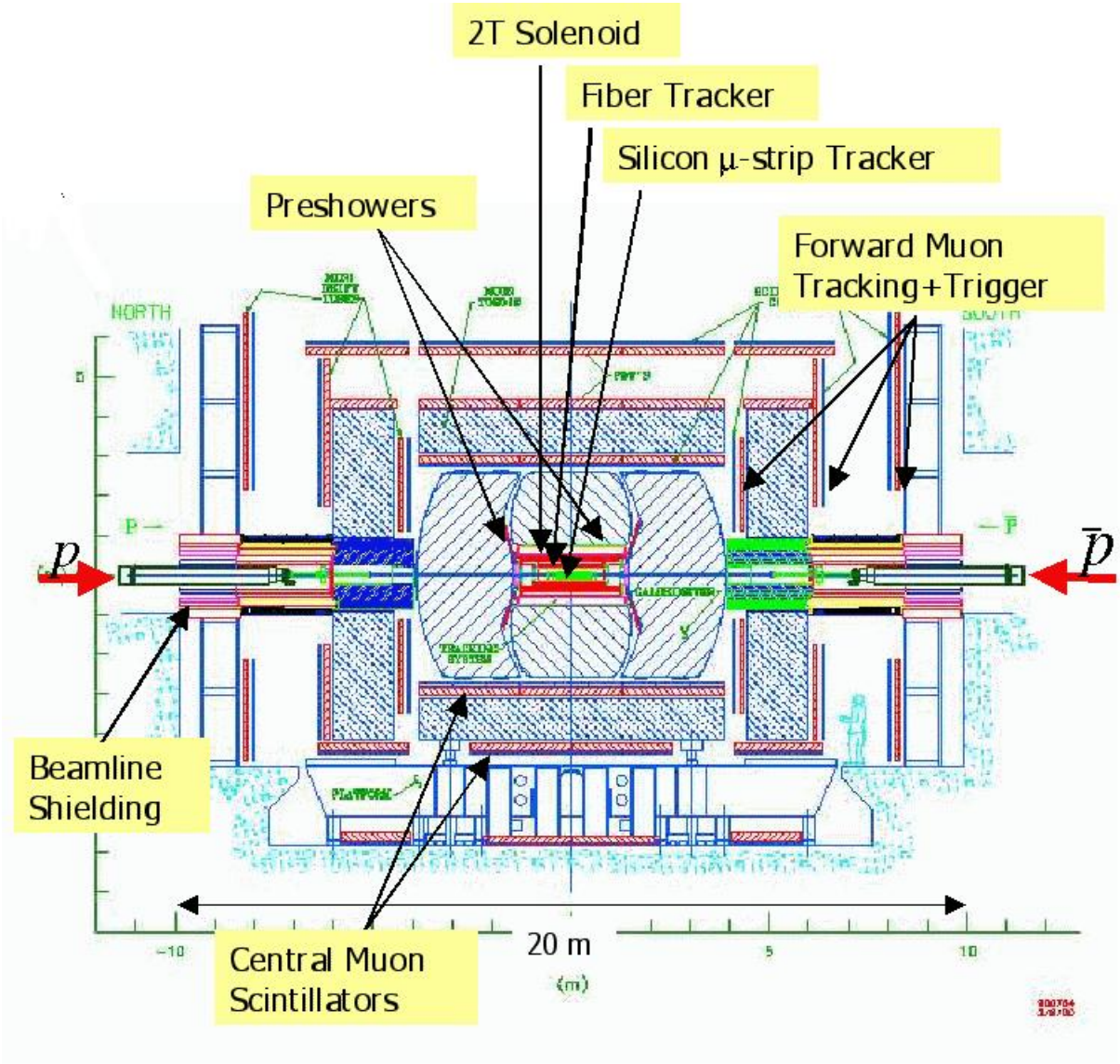


Figure 2.2: Schematic layout of the DØ detector

DØ uses a right handed Cartesian coordinate system in which the z-axis (also known as the longitudinal direction) points in the direction of the proton beam, the x-axis is in the plane of accelerator whereby the increasing direction of x is pointing away from the center of the Tevatron ring. The y-axis points vertically upwards. The xy plane is also referred to as the $r\phi$ plane in the cylindrical coordinate system. Thus, the xy plane or the $r\phi$ plane are both perpendicular to the beam direction. If, referring to the spherical coordinate system, the angle θ is taken as polar angle such that $\theta = 0$, points along the proton beam. Since, most particles of interest are produced in an ultra-relativistic regime, it customary to use the pseudo-rapidity (η) which is related to the polar angle θ as follows :

$$\eta = -\ln\left[tan\left(\frac{\theta}{2}\right)\right] \tag{2.1}$$

The pseudo-rapidity is a high energy approximation of the rapidity (y) which relates to the energy of a particle (E) and its longitudinal momentum (p_z) as follows :

$$y = \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right) \tag{2.2}$$

η is often a useful approximation, when the mass and momentum of the particles resulting from collision are unknown. The sum of longitudinal momentum of particles, produced in the collision process, is often difficult to measure for each event due to the collision remnants draining down the beam pipe. In such a scenario it is often wiser to refer to the transverse momentum $(p_T = p\,sin\theta)$, since the total momentum in the transverse plane $(xy\,or\,r\phi)$ is known to be zero, balancing the initial state of longitudinally colliding protons and anti-protons.

When a collision event is read/interpreted, the point of collision is not usually at the center of the detector. Essentially, the hard scattering process of the partons, which carry some fraction of the momentum of the colliding proton (anti-proton) is usually accessible.

The bunches of colliding protons are approximately 30 cm long and thus the collision may occur within $\sim \sqrt{2} \times 30\,cm$ in the $z-$direction. The actual position, where the collision occurs is the origin of physics eta (η_{phys}). While describing the particle trajectories, it is more convenient to do so in a coordinate system, where the origin of collision coincides with the center of the detector. A correction is usually applied to the z-position in such cases and the pseudo-rapidity here is referred to as η_{det}. The difference in the two definitions is shown in Fig. 2.3.

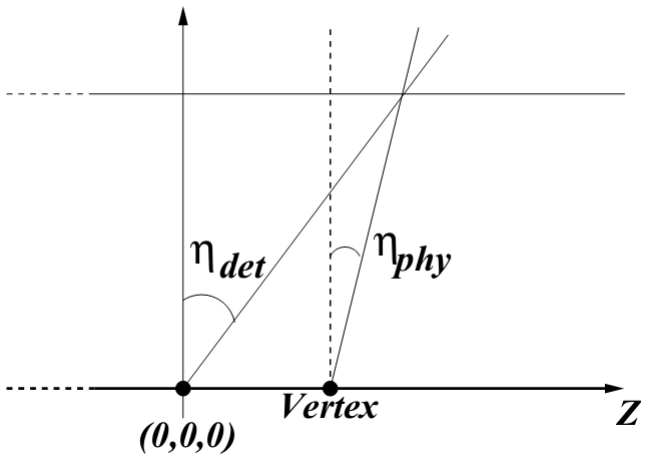


Figure 2.3: Difference in definition of physics and detector η

2.2.1 Tracking Detector

The tracking proceeds through layers of sub-detectors specialized to track various charged particles passing through it. The central tracking system is enclosed in a solenoidal magnet that produces a uniform magnetic field of 2.0 Tesla. The purpose of this magnet is to bend the trajectory of charged particles so that their momentum could be measured precisely.

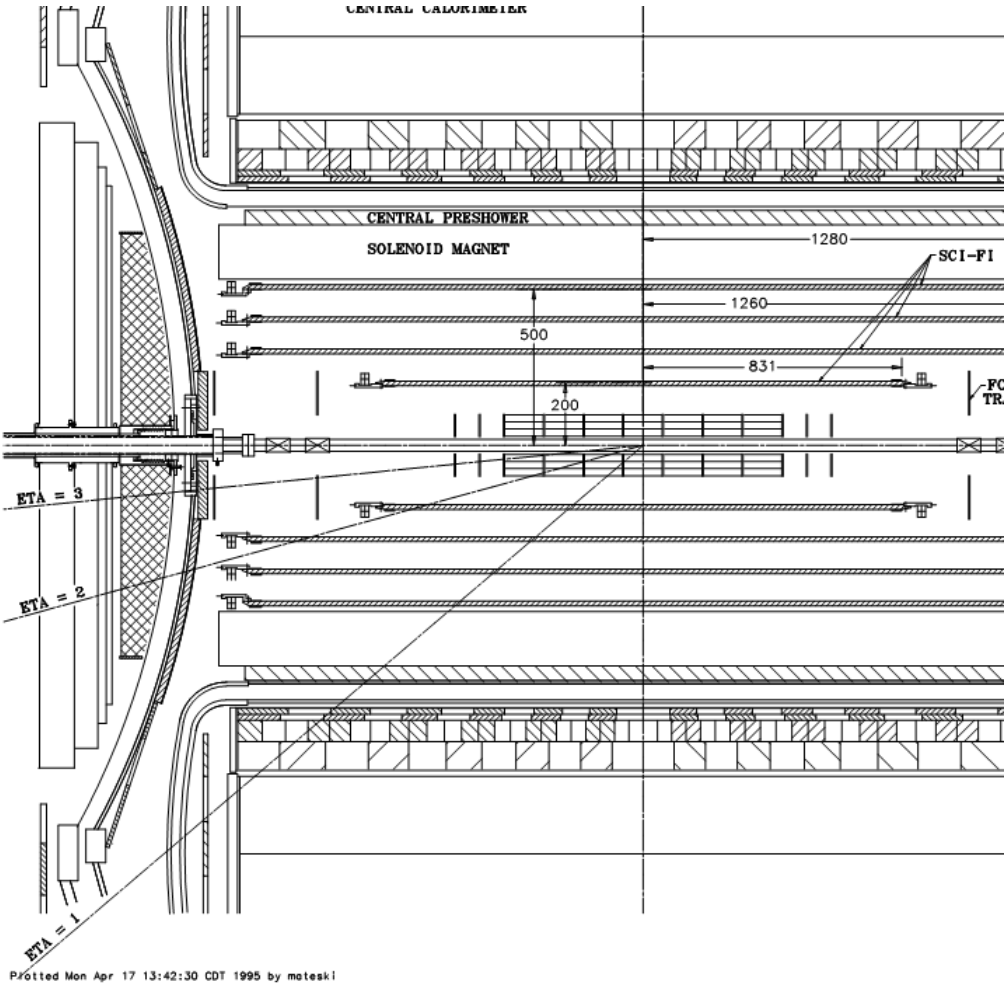


Figure 2.4: Zoomed in view of the DØ tracking detector

Silicon Microstrip Tracker (SMT)

The silicon microstrip tracker surrounds the beam pipe and is the detector element closest to the interaction point. Hence, it plays a crucial role in tracking and vertex determination of the charged particles released as the product of collisions. The length of the interaction region is about 25 cm, which dictates the length of this device. Since the tracks emanating from the interaction point are almost perpendicular to the detector surfaces for most of the η^{det} range, it is imperative to have the design of SMT in the form of barrel modules, interspersed with disks in the central region and assemblies of disks in the forward

region. This architecture allows for a three dimensional reconstruction of tracks and vertices, with the barrels covering low values of η and measuring the $r-\phi$ coordinates, while the disks provide $r-z$ coordinates and are useful for high values of η . Fig. 2.5 shows a representative view of the SMT detector. Each of the six barrels in the central region has four silicon readout layers and are referred to as *ladders* . Layers 1 and 2 have 12 ladders while layers 3 and 4 have 24 ladders each. Each barrel is covered at the end with a disk of 12 double-sided wedge detectors, referred to as "F-disks". In the far forward regions, not straddling the barrels, are the so called "H-disks" that provide tracking for high values of η . The 4 layers in the barrel detectors are implanted as two double sided layers and two single sided layers with the silicon sensor in each layer, segmented into a series of parallel strips of roughly 50 μm pitch. The two sided layers have the strips parallel to the beam pipe on one side while a stereo geometry of either a 2° or 90° on the other side. The whole detector sums up to a total of 912 readout modules with close to 800 K channels. The pitch of the strips is sufficient to provide a position resolution of 10 μm. The detectors are made out of n-type silicon wafers that are 300 μm thick, with a p^+ implants, along the length of the detector. A dielectric medium between the strips and the aluminum coating provides a capacitive coupling between the detector and the readout electronics. A sensor is a poly silicon resistor that is designed to act as a reverse biased diode. A charged particle creates electron-hole pairs while traversing through the sensor with the electrons produced accelerating towards the p^+ implants. The image charge formed on the aluminum coating is stored in an analog pipeline in the readout chips. The signal is digitized, when a Level 1 trigger accept is received and read out by a chip that is bonded to the sensor. More details can be found in (Ref.[21]).

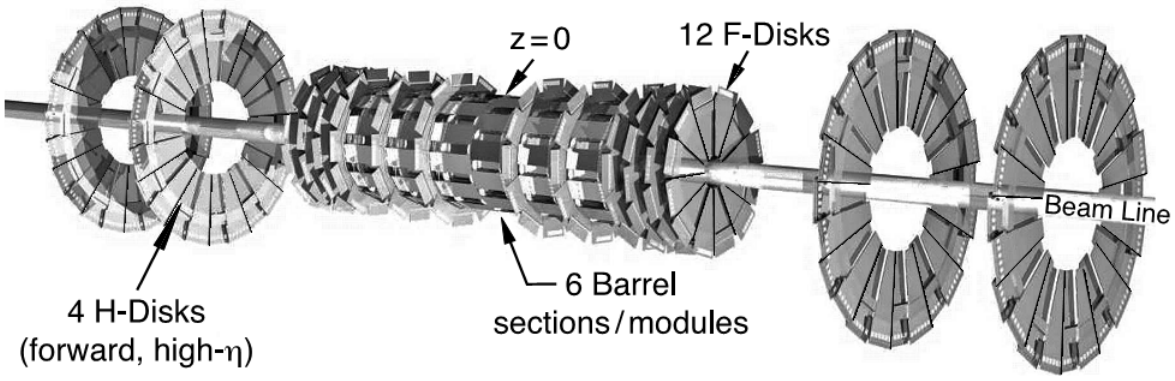
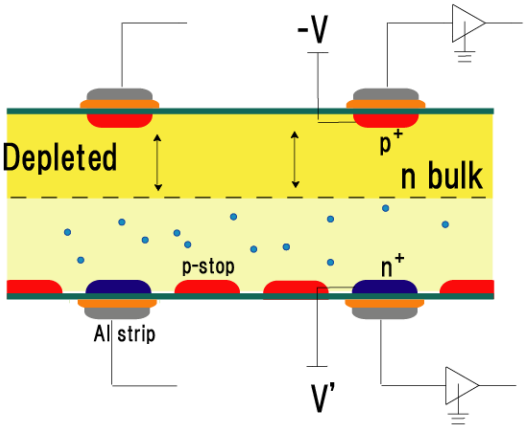


Figure 2.5: A representation of Silicon Microstrip Tracker (SMT)



A double-sided Silicon sensor.

Central Fiber Tracker (CFT)

The central fiber tracker works through three components :

- Scintillating Fibers, which convert the charged particle hits into scintillating light signal.
- Clear Fiber waveguides, which convert the scintillation light signal into visible light signal.

- Visible light photon counter cassettes (VLPC), which convert the visible light signal into an electrical signal for readout purposes.

Eight concentric cylinders with radii of 20-52 cm from the beam pipe, support the scintillating fibers. Each cylinder has one doublet of axial fibers (orientation along the beam pipe) and a second doublet of stereo fibers (orientation of $\pm 3^{\circ}$ with respect to beam pipe) mounted on it. These fibers are made of polystyrene which is doped with a fluorescent dye called paraterphenyl. Excitations in the polystyrene are transferred to the dye via dipole-dipole interactions. The paraterphenyl has a short fluorescence decay of few nanoseconds and a short emission wavelength. But the mean free path of the emitted light is also short ($\sim$ few hundred microns) and so a secondary wavelength shifter made out of 3-hydroxyflavone is added as a dopant in low concentrations. This allows for re-emittance of a 340 nm wavelength light to a 530 nm wavelength light, which then is freely transmitted in the polystyrene. Small diameter of the fiber (835 micrometer) allows CFT a resolution of around 100 micrometers. More details of the scintillating fibers can be found in Ref.[22].

The 1.66-2.52 meters of scintillating fibers, mounted on the supporting cylinders, are then connected to clear fiber waveguides. The other end of the clear fibers are coated with reflective aluminium. They are very much akin to the scintillating fibers, in their composition but without the fluorescent dye. The scintillation photons travel upto 12 m inside the clear fibers. On the other end, these fibers are connected to the VLPC cassettes. The $r\phi$ view of the fiber layout is shown in Fig. 2.6.

The VLPCs are impurity-band silicon wafers that act as photo detectors, capable of detecting single photons. The variations in gain, quantum efficiency and thermal noise requires the VLPC wafers to operate at an optimal signal-to-noise ratio between 6 and 8 Volts. Also housed in the cassettes are the analog front end boards (AFE) which act as preamplifier to the digitized signal received from the VLPCs and also provide discrimina-

tor signals, temperature control and bias-voltage control electronics. These optimizations allow CFT to work with quick response, high gain and great quantum efficiency for it to function even in high background environment.

The signal from axial fibers is used to make L1 triggering decisions (described later in the chapter) at a p_T thresholds of 1.5 GeV/c and higher. These signals are then passed on to a second level of triggering system. A third level of trigger system (L3, described later) uses the full CFT information provided through the readout electronics in VLPCs. There are about 77000 readout channels and the photelectron yield per fiber is determined to be around 8.5 and about 99.5% of the thermal noise is found to be below the the threshold of 1 photoelectron.

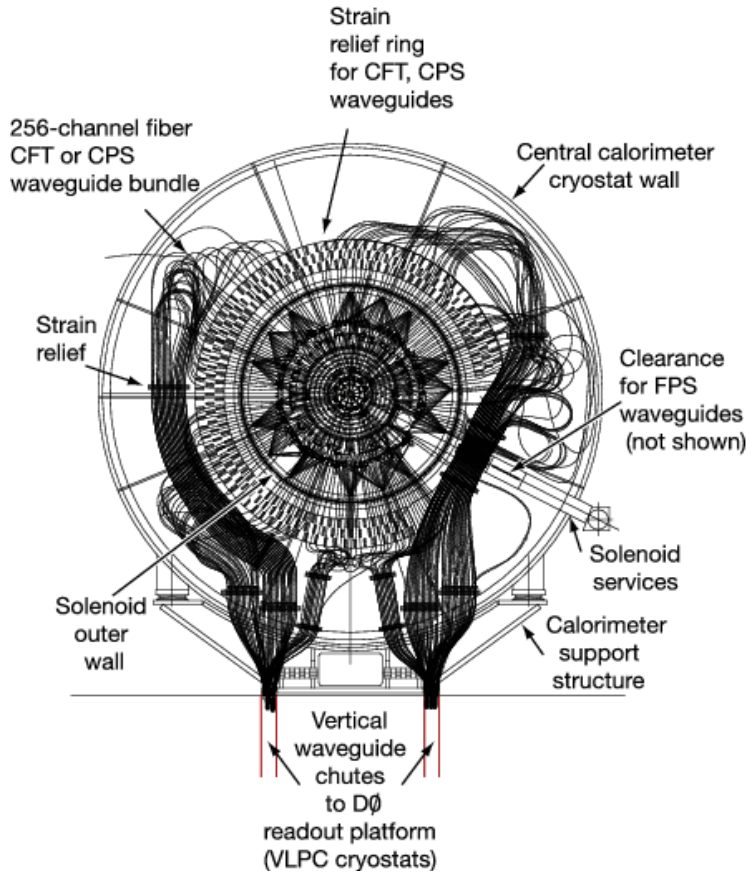


Figure 2.6: $r\phi$ view of CFT clear fiber waveguides

Solenoid Magnet

A superconducting solenoid magnet is wrapped around the central tracking detector system, providing a uniform central field of 2 Tesla. This provides a momentum measurement of the charged tracks by bending the charged particle trajectory. With about 2.73 m in length and 1.42 m in diameter, the solenoid is wound by two layers of multi-filamentary Cu:NbTi wires, stabilized with aluminum. Good field uniformity is achieved by having larger current density at the end of the coils. The thickness of the magnet system is about 1 radiation length (A radiation length X_0 is defined as the mean distance over which an electron loses about $\frac{1}{e}$ of its energy). The polarity of the solenoid is reversible, which is a boon for many measurements as it helps to remove biases in track momentum determination.

Forward and Central Preshower Detectors

The central preshower (CPS) is placed just outside the solenoid magnet to supplement the energy and momentum measurement of charged particles. It helps mainly in triggering on an electron and offline electron identification. The electrons lose about $1X_0$ while traversing through the solenoid magnet material, while the CPS provides up to 2 X_0 radiation lengths for the electrons. It consists of 5.5 mm lead and three concentric cylindrical layers of triangular shaped scintillators arranged in a *xuv* geometry ($x = axial$, $u, v = \pm 22^\circ$ aligned to beam pipe, in a stereo geometry). Using Monte Carlo simulations, the position resolution of a 10 GeV electron is estimated to be $\backsimeq$ 1.4 mm for the CPS detector (Ref.[23]).

The forward preshower (FPS) covers $1.5 < |\eta^{det}| < 2.5$ and is mounted outside the solenoid magnet and in front of the end cap calorimeter. This is used for electron identification in the forward region. The material used is similar to that of the CPS but here only the *uv* stereo geometry is in place with a layer of lead ($2X_0$) sandwiched between doublets of stereo fibers. The layout of the preshower detectors is shown in Fig. 2.7.

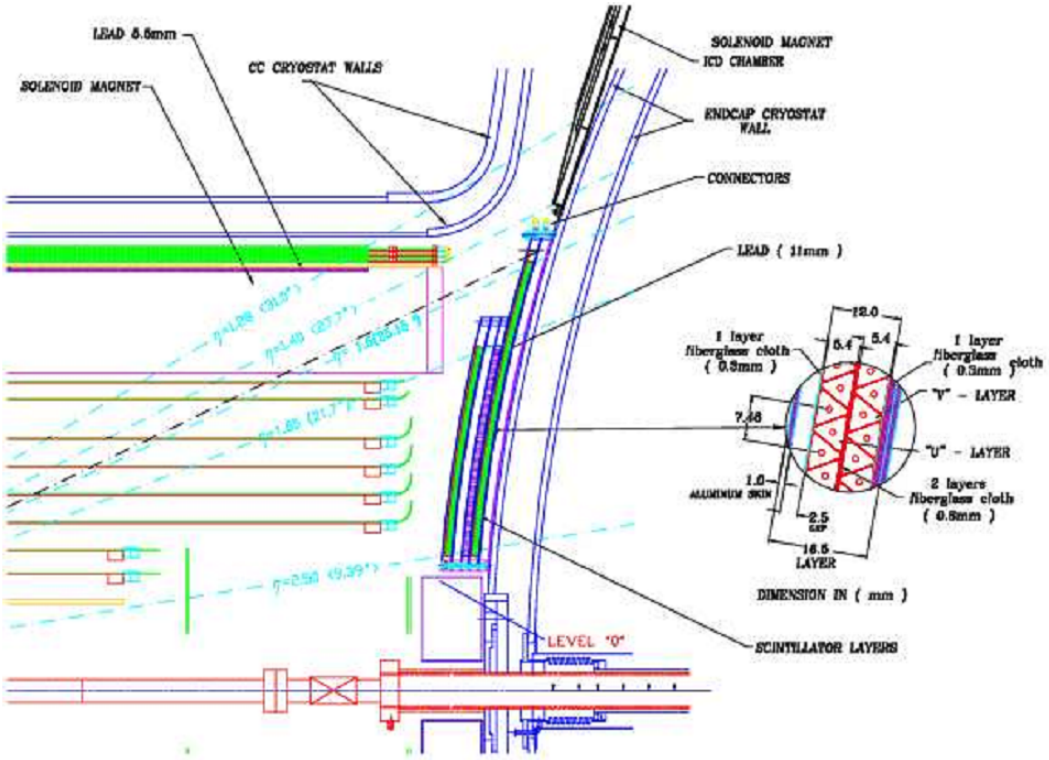


Figure 2.7: Forward and Central Preshower detectors

2.2.2 Calorimeter

The calorimeter measures the energy of particles. For highly energetic particles, produced in the collisions, it is necessary to "stop" the particles to measure the incident kinetic energy. For this purpose, usually a material with high atomic number (Z) is used. When the incident particle that interacts electromagnetically, say an electron, is suddenly made to interact with such a medium, it undergoes a process of braking radiation (*Bremsstrahlung*) due to charge particle interactions with the Coulomb field of the nuclei of the material. The result is the emission of photon. The photons thus produced are usually energetic enough to go through pair creation ($\gamma \rightarrow e^+e^-$), which in turn could go through another *Bremsstrahlung* process. Hence incident electrons or photons go through a chain of the above described process until the energy losses are sufficient and further loss of energy

occurs only through ionization. Such a process is called *electromagnetic showering*. The loss of energy in this process is gauged by the quantity *radiation length* (X_0).

Incident hadronic particles also go through the "stopping" process but the energy loss here is mainly due to the strong interactions of the hadrons with the nucleus at the core of the material. Thus a material with high atomic weight (A) is desirable. A chain of such inelastic collisions slows down the incident particles until the stopping process culminates into that of the ionization mechanism. The shower thus produced is referred to as *hadronic shower*. The measure of the loss of energy in this process is gauged by the quantity *nuclear interaction length* (λ_I).

DØ uses depleted Uranium (with copper and stainless steel) as the material which has the desired characteristic (high Z and A) as described above. Uranium has a X_0 of about 3.2 mm and a λ_I of about 10.5 cm. The ionizing medium used is Liquid Argon (LAr). A system of interleaved absorber (Uranium) and ionizing medium (LAr) is used to measure energy. As most of the energy is absorbed in the highly dense, inert absorber, only a fraction of incident energy can be measured and hence it is called a *sampling calorimeter*. This process is statistical in nature and hence the resolution achievable on the energy measurement of incident particles is limited. The ratio of the signal measured to the energy of the incident particle is defined as the *response* of the calorimeter. It is intuitive to expect that the response of the electromagnetic and hadronic showers would be different due to the difference in nature of interactions for the two processes, for example, the neutrinos (ν) produced in the pions (π) and Kaons (K) decays will escape undetected. It is also noteworthy that the hadronic showers have an electromagnetic component owing to photons (γ) produced in the decay of uncharged pions (π^0) and η particles. For the resolution of energy to be not affected by convolution of electromagnetic component of hadronic showering and the electromagnetic showering itself, it is desired that the ratio of the response of electrons to that of the pions (e/π) be close to 1. At DØ , this ratio ranges from 1.04 to

1.11, depending on the energy of the incident particle.

The DØ calorimeter is divided into large number of modules, with each module built of interleaved absorber plates (depleted Uranium), the active medium (LAr) and signal boards (to measure the signal). A transverse cross section of one calorimeter cell is shown in Fig. 2.8. The distance between the absorber plates and the signal boards is 2.3 mm. The signal boards are made of copper pads surrounded on each side, by a 0.5 mm thick sheets of G-10 to provide a dielectric medium. The surfaces of the sheets facing the active medium are coated with a resistive coating. While the calorimeter is in operation, the copper pad is grounded while the resistive coats are at 2-2.5 kV positive potential. When the showering occurs, the charged particles that ionize the Argon, induce a signal on the copper pad through capacitive coupling. The drift time of the electrons is roughly 450 ns. The signal from adjacent signal boards in a module are summed to form *readout* cells.

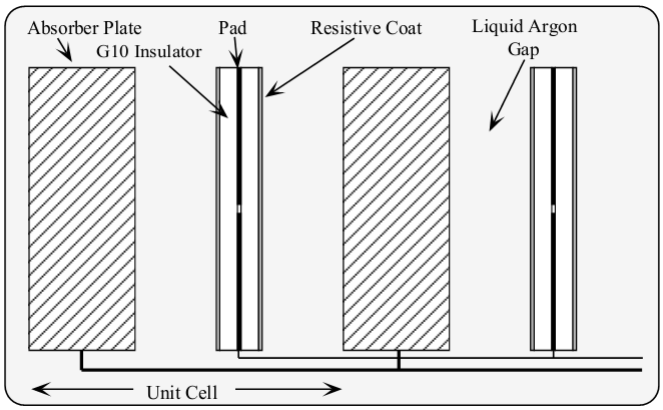
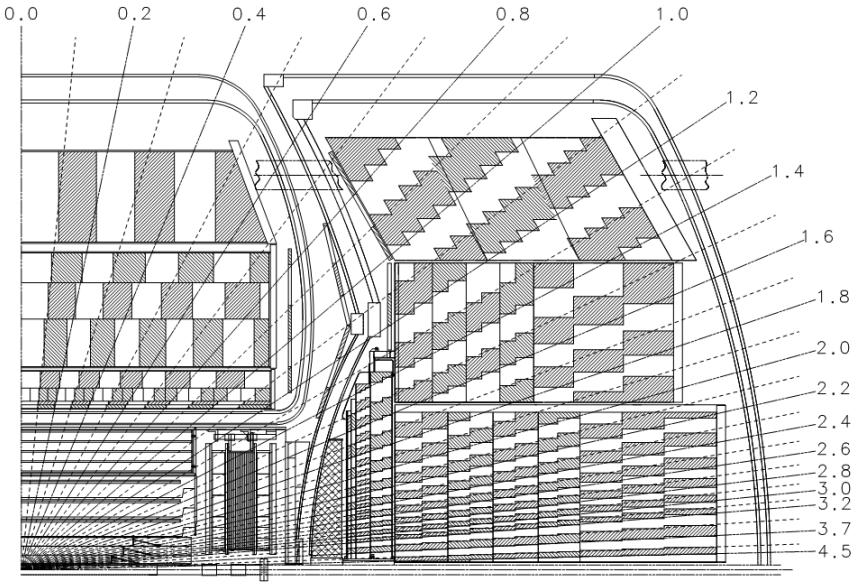


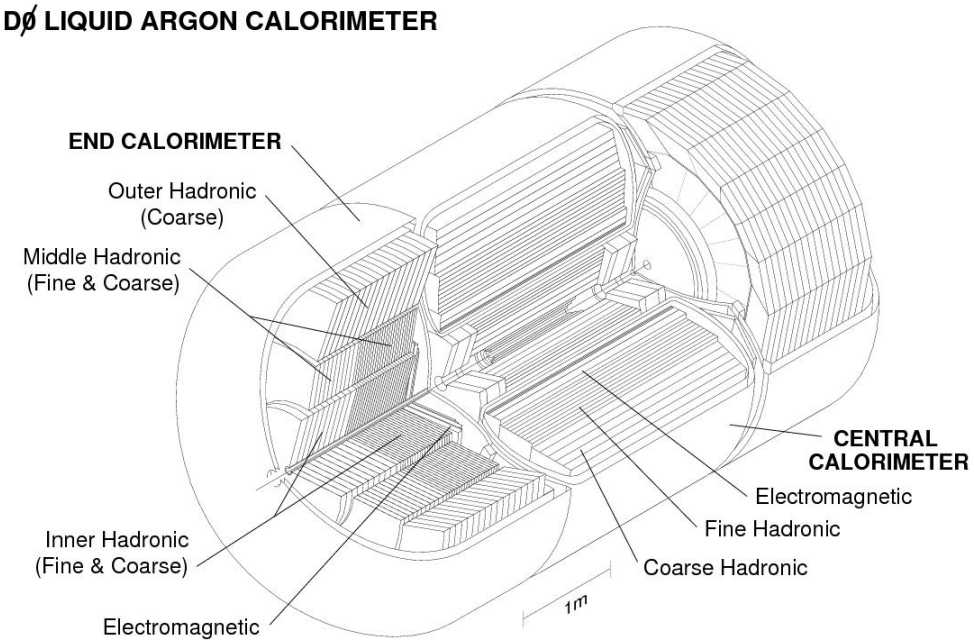
Figure 2.8: Cross sectional view of a calorimeter cell

The calorimeter cells are arranged in three assemblies, one central calorimeter (CC) and two end calorimeters (EC). Fig.2.9 shows the cross sectional view and the granularity of these assemblies. CC has a coverage of $|\eta^{det}| < 1.2$ while the EC covers the region $1.1 < |\eta^{det}| < 4.5$. The central calorimeter has a toroidal geometry with several layers of the calorimeter modules described above. These are classified as 4 electromag-

netic layers (EM) with 32 ϕ modules, 4 fine-hadronic (FH) layers with 16 ϕ modules and 1 coarse-hadronic (CH) layer with 16 ϕ modules. The FH and CH measure mostly the hadronic showering and these are placed beyond the EM layers, due to the larger spatial extent of the nuclear interaction length (λ_I). The end cap is similar to CC in architecture with the geometry differing in the fact that the electromagnetic module is disc shaped with cylindrical shaped fine and coarse-hadronic modules beyond it. In CC, the cells span $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$, with the exception of the third electromagnetic layer, where maximum electromagnetic showering occurs for particles with $p_T > 20$ GeV. In this layer, the granularity is finer for precise determination of the position of showers with $\Delta\eta \times \Delta\phi = 0.05 \times 0.05$. The position resolution is around 0.8-1.2 mm, varying coarsely as $1/\sqrt{E}$ for the electrons (Ref.[24, 25, 26, 27, 28]). More details of the calorimeter can be found in Ref.[29].



(a) η coverage of CC and EC.



(b) granularity of CC and EC

Figure 2.9: View of the central and end cap assemblies of the DØ calorimeter.

2.2.3 Muon System

The muon system measures the momentum and trajectory of muons produced in the collisions. Owing to their large mass ($\approx$ 200 times mass of electron), muons traverse most of the central detector without interacting. The dominant process of energy loss is ionization. A dedicated system beyond the the coarse-hadronic (CH) puchthrough region is in place to bend and measure the trajectory of muons. The energy loss of a relativistic heavy particle, such as the muons produced at Tevatron energies, is governed by the Bethe-Bloch equation (Ref.[30]):

$$-\frac{dE}{dx} = Kz^2\frac{Z}{A}\frac{1}{2\beta^2}\left[\ln\frac{2m_ec^2\beta^2\gamma^2T_{max}}{I^2} - \beta^2 - \frac{\delta}{2}\right] \tag{2.3}$$

wherein, Z,A characterize the material used for absorption and represent the atomic number and atomic weight respectively. $\beta = \frac{v}{c}$ and represents the relativistic regime of the charged particle that passes through the medium. T_{max} is the maximum kinetic energy that can be imparted to the charged particle and I represents the mean excitation energy of the atoms in the absorber. At DØ , the muons lie in the minimum ionizing regime of $\beta\gamma$.

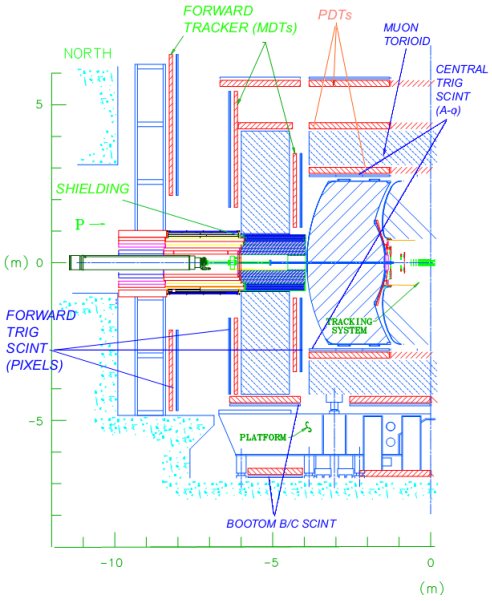
The system consists of toroidal magnets, just outside the calorimeter, the drift chambers for position of the hit determination and also for triggering purposes, the scintillation counters for fast muon identification and triggering. A map of the magnetic field of the toroidal magnets can be seen in Fig. 2.10. The central toroidal magnets operate at 1.8 Tesla while the end toroids operate at nearly 2 Tesla. The central part of the muon system provides tracking up to $|\eta| < 1$, while the forward part covers $1 < |\eta^{det}| < 2$.

The central muon system has three layers of proportional drift tubes (PDTs) with layer A inside the toroidal magnet and layers B and C outside the magnet. Fig. 2.10 shows the placements of the drift tubes in the detector. The drift chambers have large rectangular cross sections and are formed out of extruded aluminum tubes. The layers have 72-96 cells,

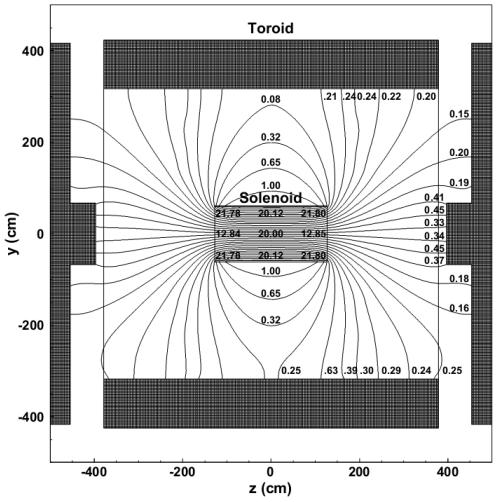
which are 10.1 cm in length. Each cell is centered with an anode wire made of gold plated tungsten in the center and diamond shaped cathode pads (copper-clad G10), sandwiching the wire, to provide information on hit position along the wire. The drift tubes contain a gas that is a mixture of 84% argon, 8% methane and 8% tetra-fluoro carbon. This provides an ionizing medium for the charged particles. When in operation, the sense wire at the center is held at a positive voltage of 4.7 kV, while the cathode pads are held at 2.3 kV. A charged particle, passing through the drift tubes ionizes the gas and the electrons accelerate towards the positively charged sense wire. This causes further ionization and the the result is an avalanche of electrons, providing an amplified signal. The drift velocity is approximately 10 $^{cm}/_{\mu s}$ for a maximum drift time of 500 nanoseconds. Each layer has a certain number of decks of the cells, for which the sense wires are ganged together, which are then readout by the electronics. Along with the electron drift time, the time difference between the arrival of the signal pulse from the cell that got the hit and that of its readout partner also recorded. This, along with the charge deposition on the cathode pad, are used to determine the hit position, along the wire. The drift distance resolution is $\sim$ 1 mm, while the resolution on the time interval varies between 10 cm to 50 cm. Besides the drift chambers, the central muon system is equipped with *Cosmic caps and bottom counters.* These are installed on the top, sides and bottom of the outer layer of the PDTs to provide a fast timing signal to associate a muon hit in the PDT with the appropriate bunch crossing and also help discriminate against the cosmic ray background. Another component is the $A\phi$ scintillation counter that covers the layer A of PDTs. This facilitates fast triggering on muons with high p_T, by matching the tracks to those triggered in the CFT, and identifying the low p_T muons that do not make it to the outside of the toroid.

The forward muon system utilizes mini drift tubes (MDTs) which have the characteristics of low electron drift time ($\sim$132 ns), good coordinate resolution and high segmentation. These are also arranged in 3 layers with layer A inside the forward toroidal magnet. The

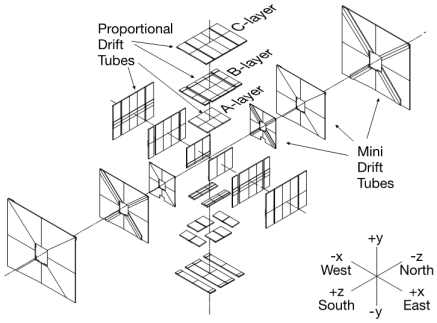
ionizing medium is a gaseous mixture of $CF_4 - CH_4$ (90%-10%). This is the main difference from the central muon PDTs. The scintillators covering layer A, B and C have a trapezoidal shape. Photo multiplier tubes mounted on the detector collect the signal in the form of light and send it out as an electrical pulse for further readout.



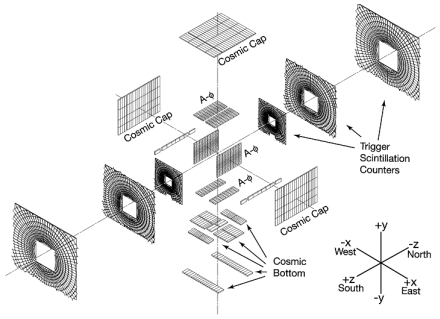
(a) $D\emptyset$ muon system



(b) Muon toroidal magnet field map



(c) Muon drift chambers



(d) Muon scintillator skeletal view

Figure 2.10: Various components of the central and forward muon system at DØ

2.2.4 Luminosity Monitoring System (LM)

The luminosity monitor determines the Tevatron luminosity at the DØ interaction region and also provides a fast measurement of the z coordinate of the interaction vertex. The basic idea is to measure the number of inelastic $p\bar{p}$ collisions in each beam crossing. The luminosity is defined as :

$$L = \frac{f\bar{N}_{LM}}{\sigma_{LM}} \tag{2.4}$$

wherein, f is the beam crossing frequency, σ_{LM} represents the effective cross section for the inelastic $p\bar{p}$ collisions, with the acceptance and efficiency of the luminosity monitor taken into account (Ref.[31]), while $\bar{N}_{LM}$ represents the average number of inelastic collisions. The products of these collisions are detected per beam crossing by two discs of plastic scintillators, segmented in 24 sectors along the ϕ direction, are placed at the far forward ends of the detector. At $z = \pm 140$ cm, these detectors span a radial area of beam pipe to the forward preshower detectors while providing an η coverage of [2.7,4.4]. The scintillation light produced in the Bicron BC-408 scintillator is detected by photo multiplier tubes (PMT), located at the center of each sector (see Fig. 2.11). The whole system is integrated with the Level 1 trigger (see next section). More details can be found in Ref.[32].

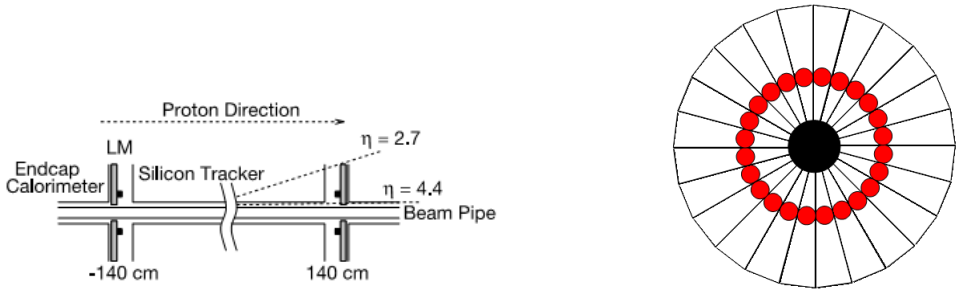


Figure 2.11: DØ luminosity monitoring system showing the location ($left$) and the geometry of LM counters with PMTs (red dots)

2.3 Trigger System

To select the interesting physics events out of the plethora of data available for an operational Tevatron with interaction rates close to 1.7 MHz, a three tier system is in place with each sieving through fewer events than its preceding level but with more detail and complexity. A cartoon of such a process, known as *triggering* , is shown in Fig. 2.12, with the three levels being termed Level 1 (L1), Level 2 (L2) and Level 3 (L3). At the first stage, L1 comprises of triggers based on sub-detectors for tracking, calorimeter and muon. L1 passes a digitized signal of accept/reject upon each beam crossing to the second stage (L2) at a rate of ∼2KHz. Here the hardware engines and embedded microprocessors associated with specific sub-detectors take a decision on whether to accept or reject an event based on the individual objects and object correlations. The accept rate for L2 is ∼ 1KHz. The decision is then passed onto a farm of computers in L3,where the sophisticated algorithms analyze the L2 accepted events and send a select few events to tape for offline reconstruction at a rate of ∼100 Hz. A more detailed overview of the three stages of triggering and elements involved in each stage is shown in Fig. 2.13.

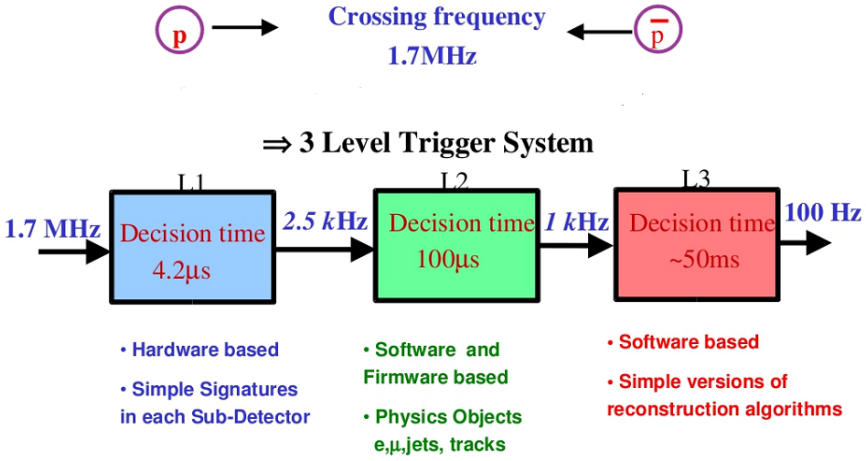


Figure 2.12: Basic trigger road map (Ref.[4])

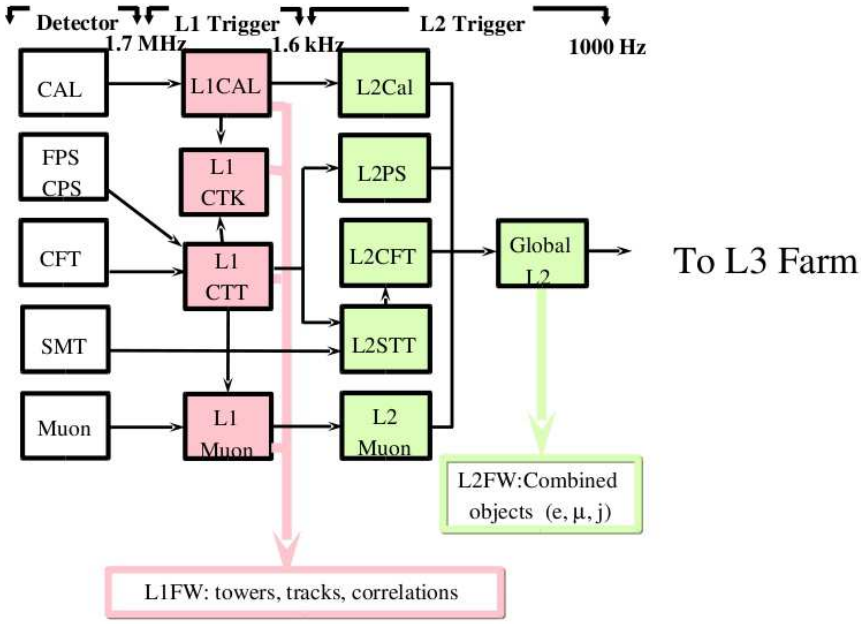


Figure 2.13: Detailed structure of trigger framework

2.3.1 Level 1

In this section, the Level 1 trigger elements are described. L1Cal identifies the patterns in transverse energy deposits in the calorimeter. L1CTT and L1Muon pass on the list of tracks with their characteristics to L1 Global trigger, which take decisions based on predefined thresholds. The rate of L1 trigger accepts is limited by the maximum readout rates of the participating sub detector system and by the motivation to minimize the *deadtime* associated with the readout.

L1CTT

The Level 1 central tracking trigger uses the information from axial fiber hits in the CFT and identifies the hit pattern with the pre-programmed look up tables (LUTs). These LUTs are digitized information encoded for know possible hit patterns in the CFT. The axial fiber information is divided in eighty ϕ segments of 4.5° each. The signals from these sectors

are processed by field programmable gate arrays (FPGA) that search for the patterns in LUTs. If a hit pattern is found in the data, a track candidate is generated. Each of these track candidates are characterized by a relative ϕ within a trigger sector, track momentum and track curvature. These tracks are then passed onto L1Muon and a second stage silicon track trigger (L2STT, described later) for triggering decisions.

L1Muon

The inputs to Level 1 Muon triggers come from the scintillation counters, the muon wire chambers and the L1CTT tracks. The scintillation counter is segmented in the same way as L1CTT and the tracks obtained via scintillation counters are processed by the FPGAs that performs a combinatorial logic to about 60,000 muon channels. The tracks thus obtained in L1CTT and scintillation counters are matched and sent to Muon trigger manager (MTM). Up to 480 tracks for every bunch crossing can be processed. The trigger terms from scintillation counters and muon wire chambers are summed up by the MTM and sent to the global trigger framework (TFW) for trigger decisions. L1Muon can handle up to 256 trigger terms and filters 32 of these to be sent to the TFW. Those tracks that can be classified as high p_T tracks are also made to go through the cosmic ray veto scintillation counters. These cosmic ray muons usually penetrate the DØ detector, but do not pass through its center. The timing information of the tracks left by these muons relative to that of the the bunch crossings can be used to reject cosmic ray background.

L1Cal

The Level 1 calorimeter trigger is responsible for making trigger decisions based on the information obtained from calorimeter. In this, the the electromagnetic modules and the fine-hadronic modules take part by summing the energies measured in the cells to form energy *towers* that span 0.2×0.2 in $\eta\phi$ space. Various sums are calculated viz. the total

electromagnetic energy (E^{em}), the total hadronic energy (E^{had}), the scalar sum of electromagnetic transverse energy (E_T^{em}), scalar sum of the hadronic transverse energy (E_T^{had}) , the total scalar transverse energy (E_T) and the missing transverse energy ($\not\!\!E_T$). These sums are compared to a set of programmable thresholds that yields a trigger term which is sent to the TFW for decisions. In all there are 16 such thresholds at level 1, including the fine-hadronic modules. Additional trigger terms can be constructed according to the desired calorimetry activity covering $|\eta| < 4$.

2.3.2 Level 2

The Level 2 triggering system consists of two components. The preprocessor elements are customized to run specialized software algorithms while the L2 global processor uses the preprocessor results to assert a decision on whether to accept or reject an event. Level 2 calorimeter (L2Cal) is a calorimeter preprocessor that uses clustering algorithms to construct basic jets or electron candidates using the information from L1Cal trigger towers. Similarly, L2CTT sorts the tracks in order of their p_T , using the information from L1CTT. The Level 2 silicon track trigger (L2STT) is the only preprocessor that combines the information from different sub-detector systems to run sophisticated algorithms of track finding. It is of great utility in identifying the b-quarks due to its precise reconstruction of impact parameter of tracks. This system is described in details below.

Silicon Track Trigger (L2STT)

The Level 2 silicon track trigger is the the preprocessor that uses the information from L1CTT and SMT to define and construct charged particle tracks during run time. It uses the axial fiber information for the six overlapping sectors of the CTT ϕ regions. About 46 tracks from each sector and up to 276 tracks per event, sent by L1CTT can be handled

by the STT. It constructs a circular trajectory from the information of the hit position in the outermost CFT layer (H-layer), the hit position in the innermost CFT layer (A-layer) and the average beam position. Thus the track position and momentum in the $r\phi$ plane is determined from the L1CTT information. It then combines this information with the silicon detectors (SMT), that are arranged into 12 sectors in ϕ. It then fits the tracks using the hits in the CFT and SMT layers, removing the constraint that the tracks originated from the point of interaction. This is particularly important for B-hadrons produced in the collisions as they have a typical lifetime of 1.5 ps and travel up to 500 μm before decaying. In such cases, the impact parameter or the distance of closest approach to the point of interaction can be determined to good accuracy. The data is processed via three components of the STT: the Fiber road card (FRC), the silicon track card (STC) and the track fitting card (TFC). Each of these plug into a common motherboard for use in a standardized crate. The communication between the components happens through mezzanine cards that use Low voltage differential signal (LVDS) cables. Each board also communicates with a common daughter board where the buffering and transfer of data to the data acquisition system (DAQ) occurs through PCI buses. Each STT crate processes data for two 30° sectors in ϕ. A brief description of each of these components follows.

Fiber Road Card

The information from L1CTT track candidates is used to define projective roads inside the silicon (SMT). Only the axial clusters in SMT that are found within the 2 mm wide road of the axial CFT hits are used as track candidates. Fig. 2.14 outlines the above principle. Fiber road card is designed receive the data from L1CTT and reformat it for the other daughter boards. After the data has been processed (fitted for tracks), it also broadcasts the results to the data acquisition on global level 2 accept for an event.

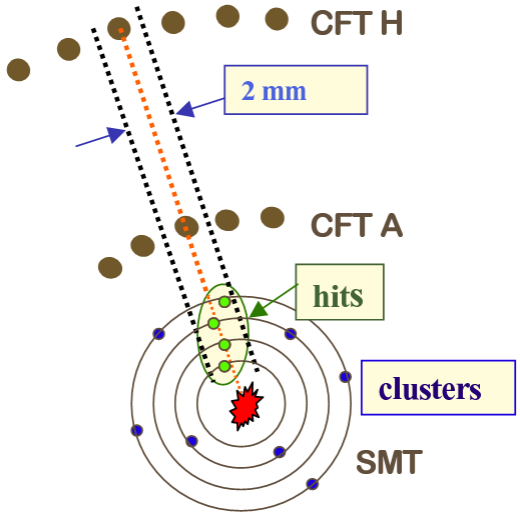


Figure 2.14: STT pattern recognition algorithm

Silicon Trigger Card (STC)

The Silicon track card receives digitized data from the 10 axial strips (including Layer 0) of the SMT, in each ϕ sector. It then masks out the noisy and dead silicon strip data while performing a strip by strip gain and offset correction.The data, thus filtered is allowed to go through a fast clustering algorithm. This involves successive passing of signal through thresholds. At first, each strip signal is tested by a strip threshold and is passed on to form a cluster only if the signal lies above it. Then the maximum of these adjacent strip signals is compared to a cluster threshold. If above the threshold, the cluster of adjacent strips passes on to next stage. The strip that records the maximum signal, along with its two neighboring strips inside a cluster are used to determine a weighted centroid for the respective cluster. Finally, the centroid is associated with the L1CTT tracks and for a centroid lying withing 2 mm road of the L1CTT track, the track is considered as selected. The list of centroids is then passed on to the track fitting card. Fig. 2.15 demonstrates the clustering algorithm implemented by the STC.

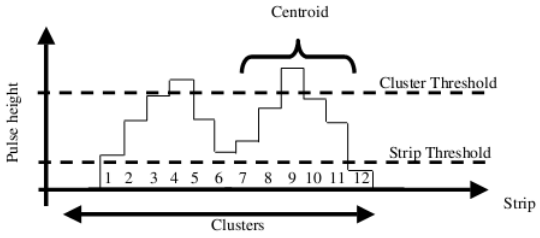


Figure 2.15: SMT clustering algorithm, implemented by STC

Track Fitting Card (TFC)

A circular trajectory is fitted using the SMT clusters and CTT tracks by the track fitting card. Clusters that are passed on by STC, are further required to go through another set of pattern recognition. Due to electronic noise, δ ray production and overlapping of signals from clusters of nearby particles, there is a possibility of finding multiple clusters per layer of SMT in the 2mm road defined earlier. To fit a track, it is desired that only one cluster per SMT layer is determined and hence the road is redefined to 1mm wide about the CTT tracks to filter the clusters further. Also, it is expected that the tracks be linear in the rz plane. The clusters in SMT are made to comply with this. Hence the algorithm looks at the silicon hits associated with the CTT track and in each layer, selects the hit that is closest to the track, under the assumption that the track originated from the interaction point. The track is fitted with a function in the $r\phi$ plane :

$$\phi\,(r) = \frac{b}{r} + \kappa r + \phi_0 \tag{2.5}$$

wherein, b represents the impact parameter of the track with respect to the center of the detector, κ represents the track curvature, while ϕ_0 is the azimuthal angle of the tangent to the track at the point of closest approach. The fitting is done using a goodness-of-fit measure defined as :

$$\chi^2 = \sum_{hits} \left[\frac{\phi_i - \phi(r)}{\sigma_i} \right]^2 \tag{2.6}$$

wherein, ϕ_i is the azimuthal position of the hit in the respective SMT layer, while σ_i is the resolution on the azimuthal position of the hit. The beam position is often fed back to the system, to keep the fitting up-to-date. Finally the output of impact parameter significance (b/σ_b) , b, ϕ_0, p_T derived from $1/\kappa$, the χ^2 , number of SMT layers used in the fit and the dE/dx derived from the pulse height values of the SMT hits are transmitted to the global Level 2 processor and the Level 3 processors. The resolution σ_b is found to be close to $\sqrt{40^2 + \left(\frac{54\,GeV/c}{p_T}\right)^2}\,\mu m$, including the beam spot resolution. More information of the STT system and hardware can be found in Ref.[33].

2.3.3 Level 3 and DAQ

The Level 3 data acquisition system (L3DAQ) communicates the fully digitized data from all detector subsystems to the level 3 trigger system for final filtering processes that run on a L3 trigger computer farm. When the global Level 2 trigger system issues an accept for an event, a total of up to 63 VME crates are read out. This a amounts to 250 kB of data per event, under normal conditions. A single board computer (SBC) in each crate sends out the accepted event fragment over an Ethernet network to a L3 trigger farm node. Here all the information of the event is combined and a modified version of the full reconstruction takes place in just about 100 ms, before the event is written to tape. The final trigger decisions on high level physics objects such as electrons, muons and jets also take place. Hence the two components, event builder and event filter work in tandem to provide an output rate of $\sim$ 100 Hz. A data flow structure of the level 3 triggering system is shown in Fig. 2.16. More details of the hardware and sophisticated computing languages involved can be found in Ref.[34] and the references therein.

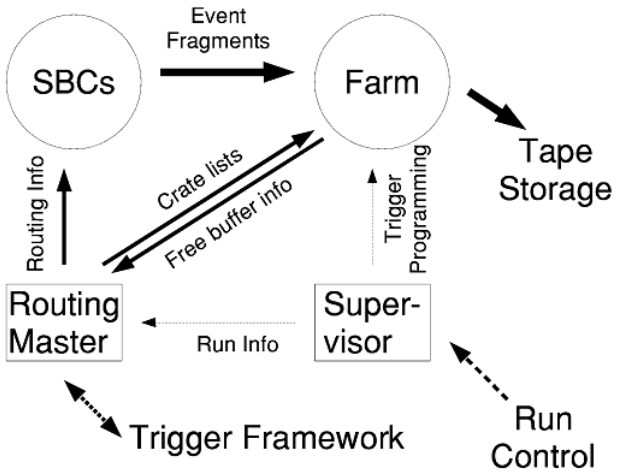


Figure 2.16: Data flow in Level 3 data acquisition system (L3DAQ)

CHAPTER **Three**

Particle Reconstruction: Algorithms and Identification

The data that is stored on tape is in digitized form, that is collected from various detector elements. This data, also referred to as the *raw* data is a collection of analogue to digital convertor (ADC) counts. To extract the interesting information of this raw data, various pattern recognition algorithms need to be employed so that the all important kinematic parameters of physics objects such as electron, muons, jets, $\not\!\!E_T$ can be determined, which would lead to identifying the particles produced in the $p\bar{p}$ collisions. This process is termed as *reconstruction*. The three main phases of reconstruction can be classified as finding of hits in the detector elements, forming clusters and eventually tracks of charged particles and identification of particles including discrimination from backgrounds. These process are discussed in the following sections.

3.1 Reconstruction in Central tracker

The two sub systems that aid in tracking are SMT and CFT (as described in the previous chapter) .

The clustering in SMT proceeds through a series of checks made by comparing the

signal recorded with thresholds and then constructing a centroid of the cluster that satisfies various requirements. SMT, being the closest tracking system to the point of interaction, the cluster building starts with its innermost layer. SMT strip that records a signal above a certain threshold ($\approx$ 8 ADC counts), is allowed to start a cluster, then the strip that is geometrically adjacent to the previous strip is compared to the threshold. All contiguous strips that are above the threshold form a cluster. With a gap in the above process, a new cluster is born. The position of the cluster is determined by the pulse height weighted average of the strips in a cluster. The axial and stereo hits are then combined to yield three dimensional global coordinates for the hits.

CFT clustering also proceeds in the same way. The signal pulse is replaced by the light yield per fiber which is eventually transformed to number of ADC counts. The thresholds in this case varies from sector to sector and between axial and stereo layers, depending upon the distance traveled by the light inside the fibers. The centroid position is just the mean of initial and final points of a cluster, in this case. Similarly, for the muon system, it matters, whether the hit was near the center of the active volume of a PDT or near the edge, which governs the drift time of electrons and hence the ADC thresholds.

The clusters formed above,do not yet give an estimate of the trajectory of the charged particle. The clusters could represent a superposition of adjacent tracks of charged particles and hence it is important to combine the clustering information in different layers of the tracking system. Dedicated algorithms are designed to construct a charged particle track and filter out the noisy, spurious clusters, arising due to combinatorial ambiguities. One kind of such an algorithm is called histogramming track finder (HTF) (Ref.[35]), whereby a Hough transformation is employed with the parameter space defined by the track curvature (κ) and the azimuthal angle of helix shaped tracks (ϕ). The spectrum of tracks that originate from interaction point and form clusters in various layers of SMT and/or CFT, have a certain trajectory and azimuthal angle in the transverse plane. These are mapped into straight lines

by the Hough transformation and for several tracks, the intersection of the transformed lines in $\kappa\phi$ space is taken as the transverse coordinates of the track. This corresponds to finding narrow peaks in the histogram of $\kappa\phi$. The appropriate track corresponding to the above coordinates can be selected by applying various criteria such as requiring clusters formed by the tracks in at least 4 layers of SMT or CFT, requiring that the track doesn't have more than three *misses* (defined by the absence of clusters in the track trajectory, in layer(s) of SMT or CFT, going outward from the point of interaction), requiring that there be no more than 2 misses inside SMT, number of total clusters found be 5 times more than the total number of misses, etc.

Tracks not originating from the point of interaction such as those of long lived particles, might not give a single point of intersection in the $\kappa\phi$ space and hence the HTF method cannot find them. To come over this problem, an *alternative algorithm* (AA) (Ref.[36]), is in place. This method is similar to the goodness-of-fit method, defined in Sec. 2.3.2, whereby the hits in the outward layer of the tracking detector are associated with the track if the χ^2 of the track fitting yields a value less than 16 and if the impact parameter of the track lies within 2.5 cm of the beam spot. To account for loss of energy through ionization when the charged particle traverses the material in the tracking volume, the hits associated with the above two methods are used by a Kalman fitter (Ref.[37]), which acts as a global multidimensional χ^2 minimization by predicting through propagation of the track in different layers. This whole procedure, provides a refined way of accurately determining the particle trajectory and its kinematics in the tracking volume.

3.2 Reconstruction in Calorimeter

As discussed in Sec.2.2.2, the particles (hadrons, electrons, photons) manifest themselves as showers obtained through ionization of LAr in the calorimeter. A digitized signal,

recorded by the readout electronics is converted into ADC counts. The ADC counts are converted into units of energy, taking into account the differences in cell response and readout electronics. The signal is corrected for gain and pedestal based on the calibrations obtained throughthe test beam in which particles of known energy are targeted on different sections of the calorimeter (Ref.[38]) and also via reconstruction of invariant mass of particles, whose mass is known. These ADC counts are converted to $\eta\phi$ coordinates for each cell and the four-momentum for the cell is is measured using the cell energy in the direction defined using the primary vertex of interaction as the origin.

$$P_{cell} \equiv E_{cell}(1,\hat{n}_{cell})$$

$$\hat{n}_{cell} = (sin\theta cos\phi, sin\theta sin\phi, cos\theta)$$

In the projective direction of $\hat{n}$, a tower consisting of cells with signal above 2.5σ of pedestal mean (cell noise), is constructed, with its associated four-momentum being :

$$P_{tower} \equiv (E_{tower},\vec{p}_{tower}) \equiv \sum_{i=cells} P_i$$

These tower energies and direction are used in reconstructing the energy and direction of physics objects, such as jets.

3.3 Muon Reconstruction

Sec.2.2.3, describes the muon system in detail. It was noted that the system is made of three layers of drift tubes (PDTs: central and MDTs: forward) and scintillators to account for the time of arrival of the signal and the position of the hit in the muon system. The muon reconstruction proceeds through linking segments in various layers and track fitting. Since the inner A layer is inside the toroid magnet while the two layers B and C , are placed outside the toroidal magnet, this allows the "segments" in the outer layers to be collinear. Hence a propagation of the track from the B and C layers to the layer A in the form of a circular helix can be made. This gives track parameters of position and momentum in layer A. Following this, a matching to tracks in the central tracking system is done, taking into account the losses in the medium (solenoid magnet, toroid magnet and calorimeter) through the error propagation technique. This gives track fitting measure χ^2 and a distance of closest approach (DCA) to the beam spot.

3.4 Vertex Reconstruction : Primary

The Primary Vertex (PV) is the longitudinal position (z_{PV}) along the beam pipe, where the hard scattering process of $p\bar{p}$ collisions take place. This is important to identify to define transverse energy and pseudo-rapidity of reconstructed objects as well as efficient identification of long lived particles such as hadrons containing b-quarks that travel some distance before decaying and hence leaving their signatures of displaced vertices. There are several soft inelastic collisions that occur in one beam crossing, resulting in tracks that are not the result of the hard scattering process. Such collisions are termed as *minimum bias* events and serve as a large instrumental background. This can be tackled using a three step process that involves track selection, vertex finding and vertex selection. Since a hard

scattering event is most likely to yield high transverse momentum (p_T) tracks, an algorithm proceeds with selecting tracks that have $p_T > 0.5$ GeV, and have at least two SMT hits, for the construction of a PV. Tracks are also selected on the basis of their *impact parameter significance* defined as b/σ_b, where b represents the distance of closest approach (DCA) of the track with respect to the beam pipe while σ_b is the associated uncertainty. The vertex finding algorithm then proceeds through forming clusters of the selected tracks (Ref.[39]). In the decreasing order of p_T , the tracks whose DCAs are within 2 cm of the mean value of an already existing cluster are added to the cluster. The tracks in each non-overlapping cluster is then fitted for a vertex using a Kalman fitter. Tracks contributing more to the χ^2 of the fit, are removed from the cluster and the fitting procedure is redone. This ends up in a list of possible vertices for PV, where the hard scattering process, could have occurred. Then comes the vertex selection procedure to eliminate the vertices associated with minimum bias interactions (Ref.[40]). This is done by assigning a MB (minimum bias) probability to each track, based on expected p_T distribution of tracks matched to the MB vertices that is obtained from the Monte Carlo simulations. These probabilities are combined to yield a MB probability for each vertex ($P_{MB} \equiv \xi \sum_{k=1}^{Ntrk} \frac{(-\ln \xi)^k}{k!}$; $\xi \equiv \prod_{Ntrk} P_{MB}^k$). In the end, the vertex with lowest value of P_{MB} is identified as the primary vertex. For this analysis, it is required that, $|z_{PV}| \leq 60\,cm$, within the SMT fiducial region and at least 3 tracks are fitted to the PV.

3.5 Particle Identification

3.5.1 Electrons

Electrons are primarily identified via the showering process that takes place in the electromagnetic calorimeter as described in Sec.2.2.2. There are several sources that can emulate

the showering produced by electrons as π^0 decays that overlap with a track, γ participating in pair creation, components of hadronic shower overlapping with electromagnetic showers etc. The shape of electron showering can be characterized by a cone of energy towers with radius $R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.2$, around the tower of highest energy. Furthermore, parameters such as electromagnetic fraction, defined as the ratio of the energy deposited in the electromagnetic calorimeter to the total energy deposited in all layers of the calorimeter, characterize the electron showering. Another important parameter is the isolation fraction, defined as :

$$f_{iso} = \frac{E_{tot}(R < 0.4) - E_{EM}(R < 0.2)}{E_{tot}(R < 0.2)}$$

wherein, E_{tot} is again the total energy deposited in all layers of calorimeter, while the E_{EM} is the energy deposited in electromagnetic calorimeter. For this analysis, it is required that this fraction is less than 15%. Electrons also interact with the tracking system and hence track matching provides a genuine way to measure E/p for electrons. This is done by extrapolating each track within 0.5×0.5 of $\eta\phi$ space and matching the electromagnetic cluster in the central tracking system to that of the electromagnetic energy deposition in the calorimeter. A matching fit χ^2 is defined as $\left(\frac{\delta\phi}{\sigma_\phi}\right)^2 + \left(\frac{\delta z}{\sigma_z}\right)^2 + \left(\frac{(E_T/p_T)-1}{\sigma_{E_T/p_T}}\right)^2$. Another χ^2_{HM} , defining a goodness of fit to seven variables that define an electron showering, such as fractional energy in the four physical depths of calorimeter, shower width in the azimuthal plane of the third layer of EM calorimeter, $\ln E_{shower}$ and $\ln|z_{PV}|$ are input as vectors to define a H-matrix χ^2 (Ref.[41]) :

$$\chi^2_{HM} = (\vec{v}_{measured} - \vec{v}_{MC})^T \, \mathbb{H}^{-1} \, (\vec{v}_{measured} - \vec{v}_{MC})$$

$$\mathbb{H}_{ij} \quad = \quad \frac{1}{N}\sum_{n=1}^{7}\left[x_i^n - \bar{x}_i^n\right]\left[x_j^n - \bar{x}_j^n\right] \qquad N = 7$$

$\bar{x}$ represent mean values of the Monte Carlo (MC) showers. A χ^2_{HM} of less than 50 is required in this analysis. A still tighter electron definition is employed by constructing an *electron likelihood.* This involves constructing a likelihood in the form of $L(x) = \frac{P_{signal}(x)}{P_{signal}(x)+P_{background}(x)}$, such that $0 < L(x) < 1$ and x is a combination of parameters described above (Ref.[42]). In this analysis, the electron likelihood value must be greater than 0.85 . Also, only electrons in the central calorimeter with $|\eta| < 1.1$, are kept in this analysis.

3.5.2 Muons

Muons are identified in the three layers (A, B and C) of the muon system. Due to less multiple scattering than in the toroidal magnet, the muon momentum can be more precisely measured in the tracking system. Scintillator hits provide the timing information. which helps to tag the muon that results from hard scattering process. In this analysis, muons with $|\eta| < 2$ are accepted. The muons are also required to have at least two A layer hits, at least one A layer scintillator hit, at least two B or C layer (outside toroid) PDT hits and at least one BC scintillator hit. The time interval between the beam crossing and the B or C layer scintillator hit, is required to be less then 10 ns. This also provides cosmic background rejection. The track reconstructed in the muon system must match a track reconstructed in the central tracker with $\chi^2/ndf < 4$, where ndf represents number of degrees of freedom. It is also required that the distance of closest approach (DCA) in the xy plane, be less than 0.2 cm, for the tracks that do not have any SMT hits, while $|DCA(x,y)| < 0.04$cm, is required for tracks with SMT hits. A muon is counted as tagged tight, when the energy deposited in the calorimeter, within the annular region of $0.1 < R < 0.4$, contains less than 8% of the

muon p_T and if the momenta of all tracks in the region $R < 0.5$, except for the one matched, adds up to less than 6% of the reconstructed muon p_T. To distinguish a muon resulting from the hard scattering process from the soft ones that result from semi-leptonic decays of B-hadrons, inside the calorimeter, a separation from jets by a distance of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ > 0.5 is demanded.

3.5.3 Jets

Jets are manifestation of final state quarks and gluons that hadronize to form various particles. The complexities arising from the type of particles that shower in the calorimeter, fraction of energy carried by various flavor of particles (b-quarks or light quarks), and the knowledge of the jet energy, limit the precision with which the underlying hard-scattering process of the partons (inside the protons and anti-protons) can be accessed. Jet algorithms are implemented after running a T42 algorithm (Ref.[43]). First, a *zero-suppression* algorithm rejects any calorimeter cells that record energies less than 2.5σ of the cell specific electronic noise. Then T42 algorithm selects those cells that have energy greater than 4σ and which has atleast one neigboring cell that records energy greater than 4σ. This allows for eliminating those cells that have an isolated energy tower, which is most likely to be a fluctuation. Following this, jets are reconstructed using a RunII legacy cone algorithm (Ref.[44]). In the following discussion, y refers to the rapidity as defined in Eq.2.2 rather than the pseudo-rapidity. This is done, so that the mass of the particles is not disregarded (even in the ultra-relativistic regime), especially for low p_T objects. Also, all the directions are defined with respect to the primary vertex (PV) of interaction, as the origin. The algorithm proceeds through iterative procedures of clustering into proto jets, splitting and merging. These processes are described briefly in the following discussion.

At first, the energy of the cells are summed up in the $\eta\phi$ space to form pseudo-

projective towers. Then the towers with energy greater than 0.5 GeV are ordered in p_T to form what are called *seeds* for pre-clustering. With the seeds in the center, an annular region of $\Delta R = 0.3$ is constructed and the four-momenta inside this cone is added to the four-momenta of the seed. This defines a pre-cluster for each seed. If the p_T of the pre-cluster is found to be above 1 GeV, it is tagged as a cluster. If the direction of the initial seed and that of the cluster agree within a small tolerance, the cluster becomes a proto jet. If the above is found not to be true, the pre-cluster is taken as a new seed and the above process is repeated. An axis alignment of the seeds with the clusters within 10^{-6} in the $\eta\phi$ space is necessary to get a list of proto jets. The proto jets thus formed, could be an overlap of two or more particle jets and hence processing them through merging and splitting is essential. The idea is to achieve a stable cone axis. At first, all proto jets that are close to one another are reprocessed through clustering. Then, based on the fraction of shared energy between two or more of the close enough proto jets (within $\Delta R = 0.1$), merging or splitting takes place. If the fraction of energy share is more than 50% of the proto jet with smaller p_T, amongst the two proto jets, both are merged to form a new proto jet and the clustering is repeated until a stable cone axis is achieved. Similarly, for the energy share less than 50%, the two are split using the mid-point of separation and treated as new proto jets. The important aspect of jet reconstruction algorithm, from the physics point of view is to achieve *Infrared and collinear* safety. This implies that the algorithm requires to be least sensitive to the soft gluon emission in the initial state (akin to the *bremsstrahlung* for accelerating charged particles, albeit color-charged in this case) as well as the collinear gluon radiation in the final state. These aspects could cause an artificial splitting or merging of jets arising from a single parton. The mid-point selected for merging andd splitting is a p_T weighted average of the $\eta\phi$ coordinates of the two proto jets involved. Since a mid-point algorithm is efficient enough to resolve a pair of jets, the loss of sensitivity of this algorithm occurs when more than two jets lie in close proximity of each other. A p_T cut is also applied to the

final collection of jets, to aid in the process of elimination of jets, contaminated with noise.

Jet Identification and Selection

To ensure that the jets that are selected in the analysis arise from the fragmentation of quaks and gluons, sources of backgrounds, such as mis-identification of electrons, photons and calorimeter noise, need to be minimized. This is done by placing requirements on jet propertes. A jet is termed a *good jet*, if it satisfies the following criteria :

- Jets that have less than 5% of their total energy in the EM calorimeter are rejected. This is done since the hadronic layers, fine-hadronic (FH) and coarse-hadronic (CH), are expected to be noisier than EM. On the other hand, jets with more than 95% of total energy in EM layer, ought to be from electrons or photons and hence these are also rejected. In other words, some activity in both EM, FH and CH is required.

- The coarse-hadronic (CH) calorimeter is the noisiest. It is required that jets deposit, less than 40% of the their total energy here.

- To alleviate electronic noise, which manifests itself as a *hot tower*, during the readout i.e. a single isolated tower of energy, it is required that the ratio of the transverse energy of the most energetic cell to that of the next most energetic cell be less than 10.

- Jets that are formed by coherent noise in some particular region of the calorimeter are removed by requiring the jets that have more than 90% of their total energy in one calorimeter tower be eliminated.

- Furthermore, all the jets are required to have passed the L1 trigger, in order to ensure that the jets are the result of the collision data. This is done by requiring the ratio $\frac{p_T^{L1}}{p_T^{reco}(1-CHF)}$, be greater than 0.5 for the central region. Here, p_T^{reco} represents the

reconstructed jet transverse momentum, while p_T^{L1} is that of a matched L1 object and CHF represents the coarse-hadronic fraction.

- To mitigate the effect of *minimum bias* vertices on the reconstruction of jets, tracks in the tracking system are selected that point towards the jet. A requirement that at least two of such tracks originate from the primary vertex (PV) makes the jets *vertex-confirmed.*

3.6 Jet Energy scale

The aim of the jet energy scale correction is to correct the reconstructed jet energy to the particle level. For these corrections to remain as model independent as possible, the effects of non-perturbative QCD viz. hadronization and underlying event are not corrected for. This is because the "particle level" energy corresponds to the incident particles on the calorimeter, which include the effects that are mentioned above. The relation between jet energy and particle energy, can be written as :

$$E_{particle} = \frac{E_{measured} - \hat{E}_O}{R.S} \tag{3.1}$$

wherein, E_o is the offset energy, R is the calorimeter response to the particles and S represents the correction factor, accounting for the migration of particles, inside and outside of the cone, also termed as showering effect due to scattering with the material of the calorimeter, etc. Hence all of these correction factors correct for the instrumental backgrounds. The offset energy is, on average, the energy that does not originate from hard scattering or underlying event. In principle, it depends on the Luminosity (L) and the number of primary vertices (n_{PV}). The offset energy can be classified into two categories, the first being that whose source is calorimetry noise and pile up (deposition of energy from

previous collisions) . This category is termed as NP. The second one arises from multiple $p\bar{p}$ interactions within one beam crossing and is termed as MI. The average offset energy is estimated for each calorimeter ring in $i\eta$, by summing overall energy towers in $i\phi$ (here $i\eta$ refers to integer value of $10.\eta$). The offsets are estimated by triggering on special kinds of events and are measured directly from the data. The "zero-bias" data is collected with the only requirement that the trigger should fire at the same time as the beam crossing. The "minimum-bias" events (as discussed in Sec. 3.4) are collected when the luminosity monitor (LM) receives hits at the same time on both disc arrays. This ensures than an inelastic collision of $p\bar{p}$ has taken place. This data is collected at a constant rate of 0.5 Hz. The NP offset is estimated using the zero-bias data with the veto on luminosity counter, indicating that no interaction occurred. The MI offset is estimated from the minimum bias events and it depends on the number of primary vertices (as the soft inelastic multiple $p\bar{p}$ collisions are looked for). This dependence is found to be linear and hence the offset for N number of interactions is estimated from events with N+1 PVs and from this the average energy measured for events with one PV is subtracted. To sum up, the offset energy (as a function of L, n_{PV}), can be written as :

$$\hat{E}_O^{ring}(i\eta, n_{PV}, L) = \hat{E}_{NP}^{ring}(i\eta, L) + \hat{E}_{MI}^{ring}(i\eta, n_{PV}, L)$$

$$\hat{E}_O^{ring}(i\eta, n_{PV}, L) = \hat{E}_{ZB}^{ring}(i\eta, L) + \hat{E}_{MB}^{ring}(i\eta, n_{PV}+1, L) - \hat{E}_{MB}^{ring}(i\eta, n_{PV}=1, L) \qquad (3.2)$$

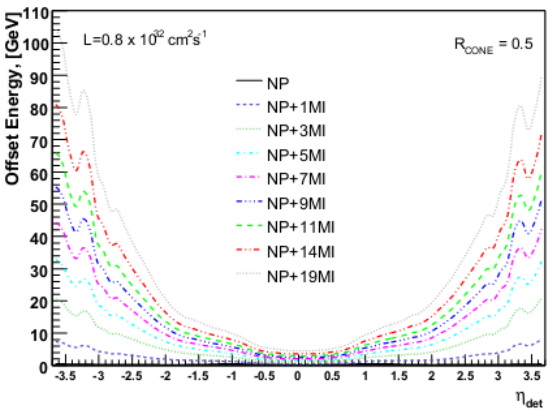


Figure 3.1: $\hat{E}_O$ as a function of $|\eta_{jet}^{detector}|$ for jets with cone size of $R = 0.5$. Different number of multiple interactions (MI) are shown

Jet Response

The next step is to determine R in Eq. 3.1 , which represents the detector response to the particles. This can be factorized as $R_{CC}(E) \times F_\eta(\eta, E)$, where R_{CC} represents the detector response in the central calorimeter, while F_η is the factor that normalizes the response of the calorimeter as a function of jet rapidity. This is derived from a sample of $\gamma + jets$ events (produced through $p\bar{p} \to \gamma q + X$, $p\bar{p} \to \gamma g + X$) with various thresholds on the transverse momentum of at least one electromagnetic cluster, found in the central region of the detector. The $\gamma + jet$ candidate thus selected is not free from contamination of dijet events, in which one of the partons fluctuates to a leading π^0 that decays into photons and thus the photon gets misidentified. This is not a sizeable background, it nevertheless affects low $p_{T_\gamma}^{measured}$. Hence a missing transverse energy projection fraction method (MPF) (Ref.[45]) is applied to measure the response from these selected events. This is based on the transverse plane momentum balance. A vector sum of all energy towers is projected, transverse to the beam. This equals the $\not{E}_T$ of the event. At the particle level, the photon is expected to balance the hadronic recoil, with $\vec{p}_{T,\gamma} + \vec{p}_{T,hadronic} = 0$. This method is also known as tag and probe method. Here, the photon (electromagnetic cluster) serves as the

tag, while the hadronic recoil in the calorimeter serves as the probe. In principle, this method could also be applied to $Z+jets$ or dijet events. To have a clear interpretation of $\not{E}_T$, the events in the $\gamma+jet$ sample are selected such that there is exactly one jet, back to back with the photon, by applying a selection cut on $\Delta\phi$ to be 3.0. Also, the response measurement is restricted to 1 or 2 PVs to have a lower luminosity environment for such events. Even in the ideal case scenario, the response to the particles is non-unity. One can thus write,

$$R_{em} \cdot \vec{p}_{T,\gamma} + R_{hadronic} \cdot \vec{p}_{T,hadronic} = -\vec{\not{E}}_T \tag{3.3}$$

wherein, $R_{em} \& R_{hadronic}$ are the electromagnetic and hadronic calorimeter responses, respectively. The electron energy scale, determined from using $Z \to e^+e^-$ events in the data, can be used to tune Monte Carlo simulations to reproduce the electron energy response in the data. This allows for the corrected photon energy, $R_{em} = 1$. With this into consideration, $R_{hadronic}$ can be estimated by projecting everything on the photon p_T unit vector $\hat{n}_\gamma$. Hence,

$$R_{hadronic} = 1 + \frac{\vec{\not{E}}_T \cdot \hat{n}_\gamma}{|\vec{p}_{T,\gamma}|} \tag{3.4}$$

becomes valid for the back to back photon and the hadronic activity. $R_{hadronic}$ can be considered synonymous with R_{jet}, for the selection of events, described above. This response is measured in different bins of jet p_T, but then the resolution on measured jet energy is found to be poor and hence a scaled variable $E' = p_{T,\gamma} \cdot cosh\,\eta_{jet}$ is used to measure the jet response. Owing to the limited statistics for $E' > 350\,GeV$, the measured response is extrapolated to up to 600 GeV using Monte Carlo models in which the response to single pions have been tuned to match the data. The response is then fitted with a quadratic function :

$$R^{\gamma+jet}_{MPF,CC}(E') = p_0 + p_1 \log(E'/E_0) + p_2 \log^2(E'/E_0) \tag{3.5}$$

with $E_0 = 100$ GeV. The fit parameters are shown in Fig. 3.2. Some of the uncertainties in this estimation arise from determination of photon energy scale, photon identification, fragmentation and parton distribution functions (PDFs) of which the first one is dominant.

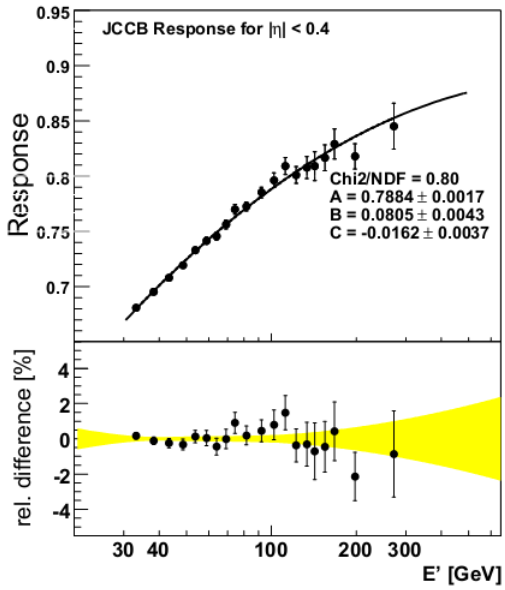


Figure 3.2: MPF response in CC, for jets with $R_{cone} = 0.5$, as a function of E'. Also, shown is the relative difference of fit with the data points.

With the energy offset correction, accounting for the irregularities arising due to ϕ dependence (hence summing over $i\phi$ in Eq. 3.2), the only directional dependence left is in η . This dependence arises, mainly due to varying radiation lengths (λ_I) traversed by the particles in the calorimeter. The jet response incorporates this by extending the central calorimeter (CC) response at forward pseudo-rapidity. This extrapolation is then fitted in the following way :

$$F_{\eta}^{\gamma+jet} \equiv \frac{R_{MPF,\eta}^{\gamma+jet}(E')}{R_{MPF,CC}^{\gamma+jet}(E')} = \frac{p_{0,\eta} + p_{1,\eta}\log(E'/E_0) + p_{2,\eta}\log^2(E'/E_0)}{R_{MPF,CC}^{\gamma+jet}(E')} \tag{3.6}$$

Fig. 3.3, shows the F_{η} for various values of E'.

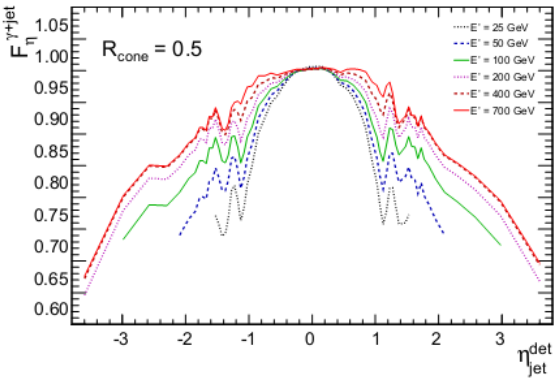


Figure 3.3: Relative MPF response, F_{η} for different values of E' in $\gamma+jet$ events.

Showering Corrections

The showering corrections include effects of energy profiles of particles that deposit energy in and out of the cone of $R = 0.5$. The profiles are created as a function of $\Delta R = \sqrt{(y_{particle} - y_{jet})^2 + (\phi_{particle} - \phi_{jet})^2}$, which essentially determines the the separation of incident particle from the jet cone axis. Then the correction factor can be defined as :

$$S = \frac{\sum_{\Delta R=0}^{Rcone} E_{in} + \sum_{\Delta R=0}^{Rcone} E_{out}}{\sum_{\Delta R=0}^{\infty} E_{in}} \tag{3.7}$$

wherein, E_{in} includes the sources of energy deposited from within, inside the cone, while E_{out} represents the sources of energy deposited from outside the cone. The energy profiles in the data are then compared to the Monte Carlo particles (without the offset applied) and the contributions from the NP and MI events, to get an absolute showering correction factor.

Closure

With all the correction factors of Eq. 3.1, in place, it is desired that the $< E_{jet}^{coorected} >$ matches with the $< E_{jet}^{particle} >$. This sort of ratio (known as internal closure) is expected to be one at least for the Monte Carlo simulations. This is shown in Fig. 3.4 for $|\eta_{jet}|$ spanned in intervals of 0.4. The dashed lines show the uncertainties expected and the major source is recognized as that of the η dependent response correction factor, shown in Fig. 3.5. A similar ratio of $\left(<E_{jet}^{corrected,DATA}> / <E_{jet}^{corrected,MC}> \right)$ is evaluated as a function of p_T in different $|\eta_{jet}^{detector}|$ regions. The data selected here is similar to that used for deriving MPF response correction, albeit a bit relaxed in the sense that the restriction on number of primary vertices is removed, for the sake of completeness. This is shown in Fig. 3.6. The uncertainties (dotted lines in the figure) are defined as the sum in quadrature of data and MC uncertainties.

In summary, the jet energy scale for a few values of jet transverse energy can be seen in Fig. 3.7. For most values of E_T, it is found to be less than 1.5. More details can be found in (Ref.[46, 47]).

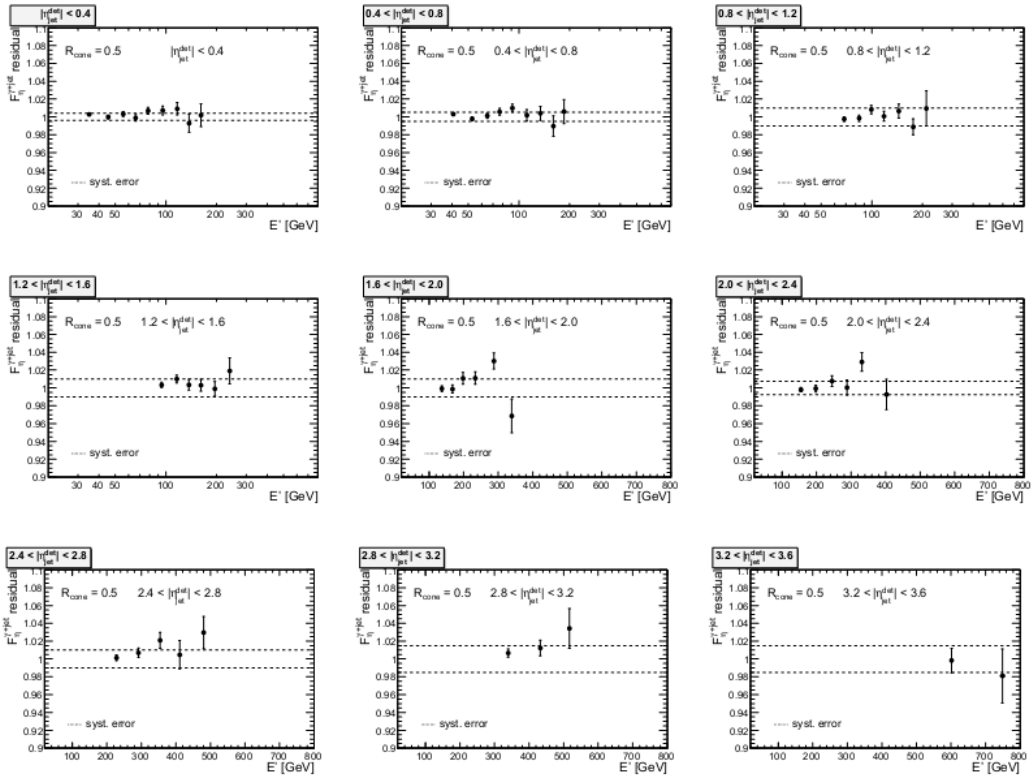


Figure 3.4: Internal closure of $\eta - dependent$ corrections for $R_{cone} = 0.5$ jets in $\gamma +$ *jet* sample.

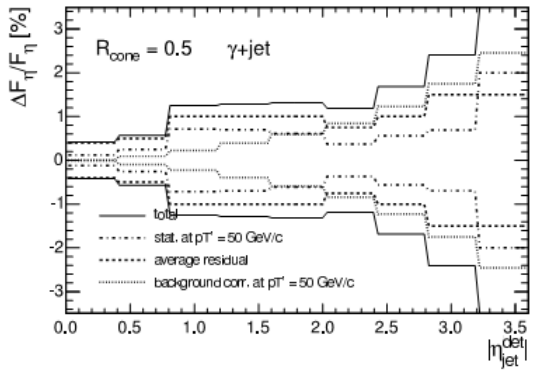


Figure 3.5: Relative uncertainty on F_{η}.

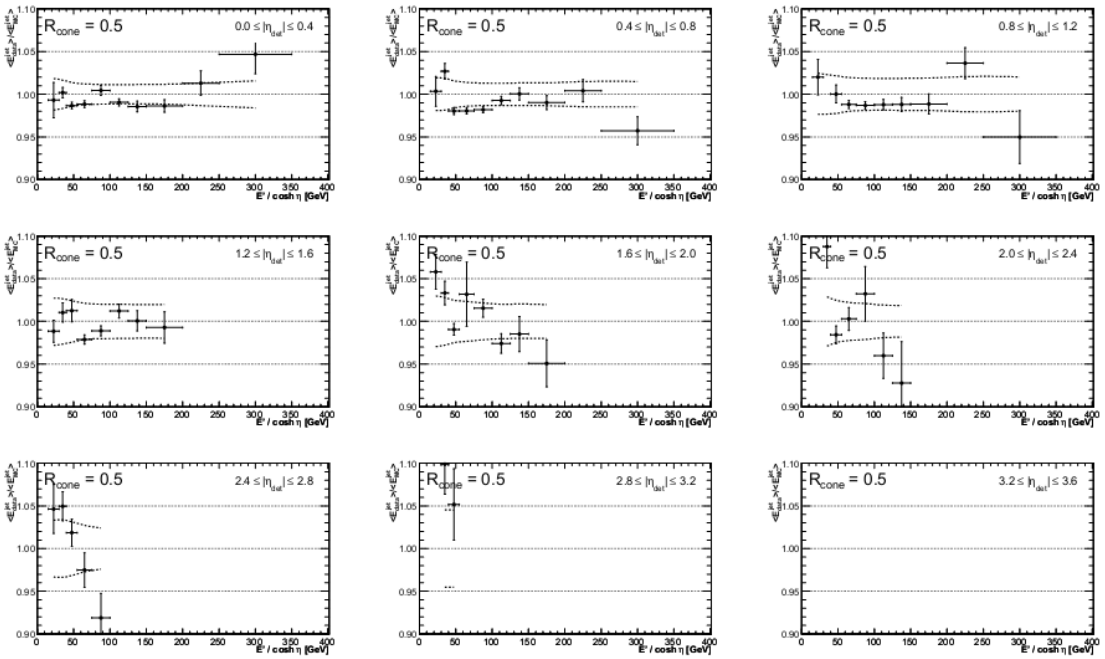


Figure 3.6: Data-to-MC closure as a function of p_T in different $|\eta_{jet}^{detector}|$ bins for $R_{cone} = 0.5$ jets in $\gamma + jet$ sample.

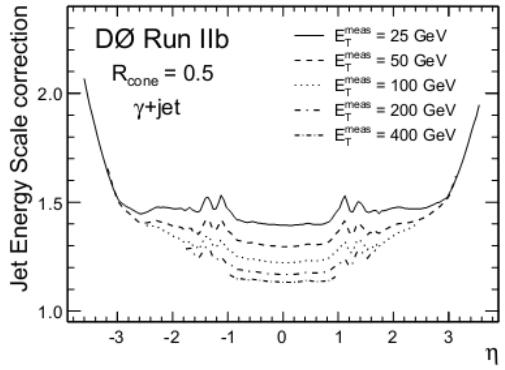


Figure 3.7: Jet Energy Scale correction factor as a function of pseudo-rapidity for a few measured values of E_T.

After applying the corrections, this analysis includes jets with $p_T > 15\,GeV$, within $|\eta_{jet}^{detector}| < 3.4$. Additionally, jets containing muons within $\Delta R(\mu, jetaxis) < 0.5$ are corrected for the momentum that is carried away by the muon and neutrino. Since, these muons could be have been generated, via semi-leptonic decay of B-hadrons (containing

b-quarks). The neutrino momentum is considered to be the same as that of the muon, for such cases.

3.6.1 Sample Dependent JES

In the previous section, the observed jet energy in data and simulations were corrected to the incident particle level. While, it was seen that such a correction involves a response that depends on the kinematics of the incident particle, no dependence on the flavor of the incident particle was considered. This section, explores such dependence, taking into account, the fact that the incident particle (hadron) could have been generated with parent parton as gluon, light quarks or heavy quarks, which in turn are products of the hard scattering process of the underlying event. The reason for considering this is that the particles of different flavor, decay in different topologies and hence the correction involving this aspect is called $flavor\ dependent$ jet energy scale or $sample\ dependent$ jet energy scale. The idea is to look for all spatially matched particles inside a cone of $R = 0.5$ and add the energies, weighted with their response for each particle. This quantity can then be compared between data and simulations. The measured jet energy taken here is without the offset corrections arising due to the NP and MI effects (as discussed in the previous section). The correction factor $F = \frac{\sum_i E_i \cdot R_i^{data}}{\sum_i E_i \cdot R_i^{MC}}$, is then multiplied to $(E_{jet}^{measured} - E_o)$, to which the corrections (briefed in the previous section), such as showering and η dependent response are applied. First, a closure for Monte Carlo simulations is obtained by comparing the measured energies and the truth values for the particles inside a jet. The response is measured for particles such as $\gamma, e^{\pm}, \mu^{\pm}, \pi^{\pm}, K^{\pm}, K_0^S, K_0^L, p^{\pm}, n \,\&\, \Lambda$, with each being parametrized as a function of p_T^{γ} in the $\gamma + jet$ sample as follows :

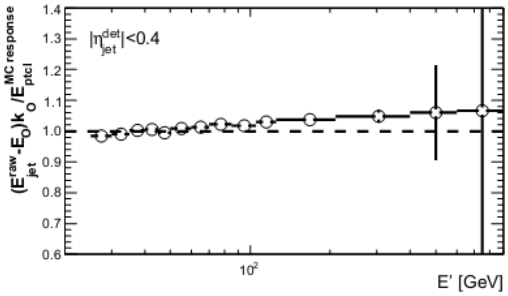
$$R_{\gamma}^{MC} = 0.25 \cdot p_{\gamma}^{0} \cdot \left[1 + Erf\left(\frac{E+p_{\gamma}^{1}}{\sqrt{2 \cdot p_{2\gamma}}}\right)\right]\left[1 + Erf\left(\frac{E+p_{\gamma}^{3}}{\sqrt{2 \cdot p_{\gamma}^{4}}}\right)\right]$$

$$R_{e^{\pm}}^{MC} = 0.25 \cdot \left[1 + Erf\left(\frac{E+p_{e^{\pm}}^{1}}{\sqrt{2 \cdot p_{e^{\pm}}^{2}}}\right)\right]\left[1 + Erf\left(\frac{E+p_{e^{\pm}}^{3}}{\sqrt{2 \cdot p_{e^{\pm}}^{4}}}\right)\right] \quad for \quad p_T > 0.3 \; else \; 0$$

$$R_{\mu^{\pm}}^{MC} = \left(p_{\mu^{\pm}}^{0} + p_{\mu^{\pm}}^{1} \cdot E\right).Landau(E, p_{\mu^{\pm}}^{2}, p_{\mu^{\pm}}^{3})$$

$$R_{h}^{MC} = p_{h}^{0} \cdot \left[1 - p_{h}^{1} \cdot (E/0.75)^{p_{h}^{2}-1}\right]$$

wherein, $R_{\gamma}, R_{e}, R_{\mu}, R_{h}$ represent the photon, electron, muon and hadronic response, respectively [p^0, p^1, p^2, p^3 being the parameters]. The ratio of the measured jet energy (without offset) and sum of the particle energies, with the above responses is shown in the figure below for $|\eta_{jet}^{detector}| < 0.4$, as a function of $E' = p_T^{\gamma} \cdot cosh\left(\eta^{jet}\right)$.



As expected, this ratio is close to 1. The responses of photon, $e^{\pm}$ and $\mu^{\pm}$ are similar in

data and hence to take this up to the next level, the hadron response R_h, needs to be tuned to the data. This is done by introducing 3 extra parameters (A, B, C) in the expression for R_h^{MC} above, such that :

$$R_h^{data} = C \cdot p_h^0 \left[1 - A \cdot p_h^1 \cdot (E/0.75)^{p_h^2 + B - 1}\right]$$

This is then tuned for data selected, with two different selections. First, a tight photon selection, in which a photon of high purity is demanded in $\gamma + jet$ events, and another one in which the track isolation requirement is reversed and thus selecting mainly dijet events. This tuning is shown in Fig. 3.8 for both kind of events.

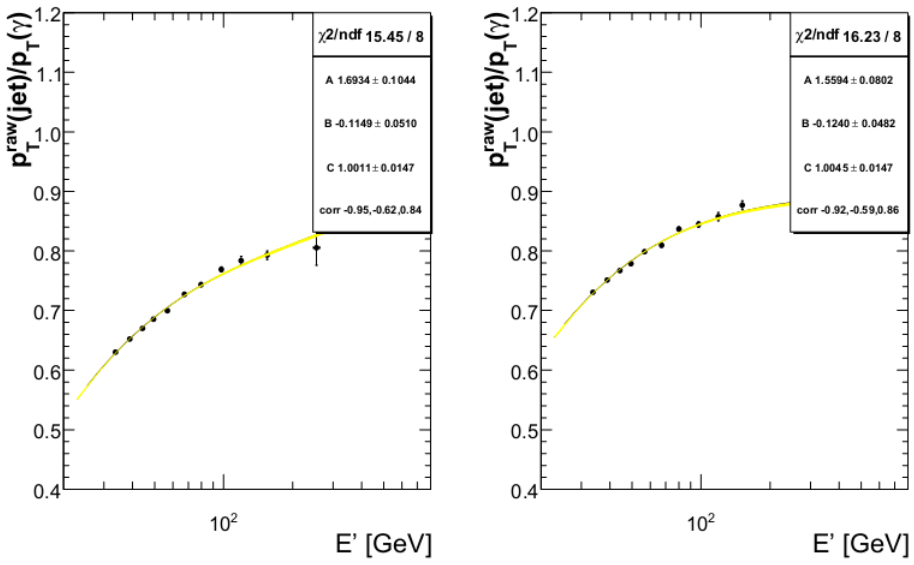
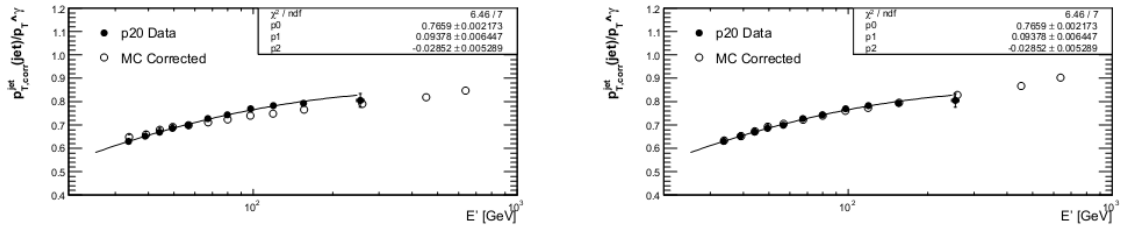


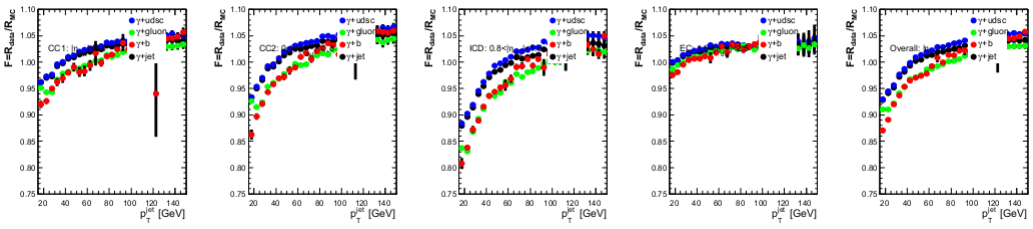
Figure 3.8: R_h tuning in MC for $|\eta_{jet}^{detector}|$ <0.4 for tight photon selection (left) and reversed track isolation (right)

The figure below, shows the difference between not tuned and tuned MC for a mixed sample of $\gamma + jet$ and dijet events with tight photon requirement. This is overlay-ed on the data (solid black dots). A much better description of data is achieved for the tuned MC for

$|\eta_{jet}^{detector}| < 0.4$, as shown below. This procedure is carried out up to an η of 2.5 and a similar agreement is seen in all the rapidity bins. (Ref.[48])



Including all the responses, the correction factor F , in various rapidity bins is shown in the figure to follow.Here, the blue dots represent gluon jets, the red ones are the b-quark jets , green represents light quark jets and black is the original $\gamma+$ *jet* sample. Including the systematic uncertainties (Ref.[49]), the ratio of the F factor of these 3 flavor of jets to that of the average $\gamma+$ *jet*, can be used to correct the jet energies in any sample. The deviation of such a ratio from 1, is shown in Fig. 3.9 , as a function of p_T^{jet} , in different rapidity bins.



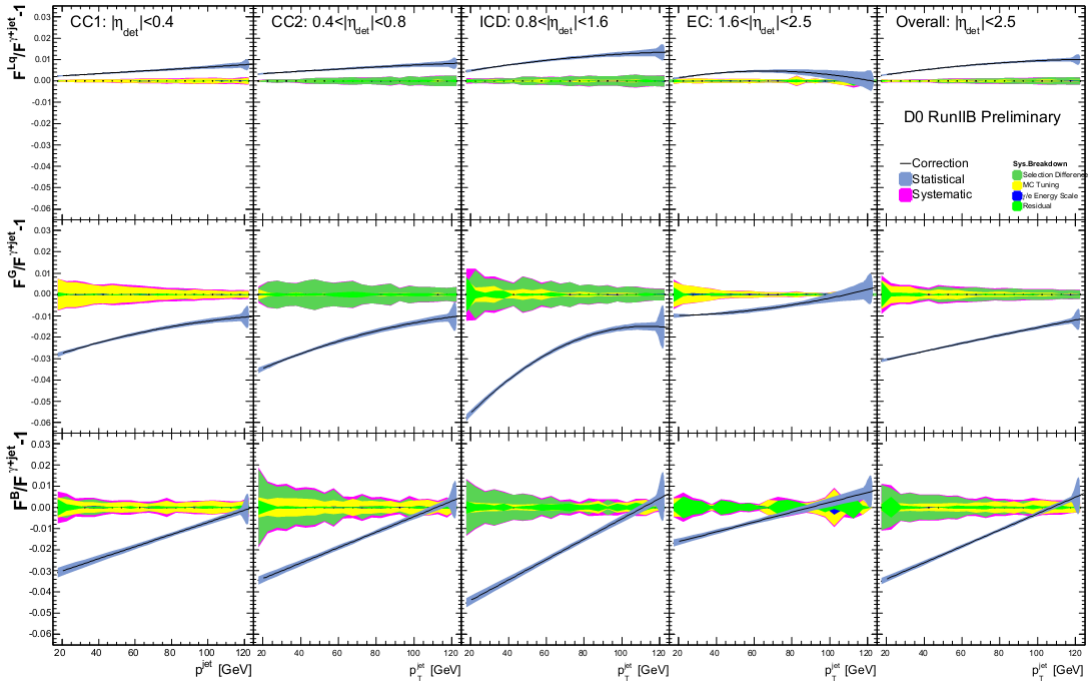


Figure 3.9: Correction factor for MC-data difference in jet response for different flavor of jets. [Light quarks(top), gluon(middle) and b-quarks(bottom)]

3.7 Missing Transverse Energy

Due to the initial state of unpolarized beams of p & $\bar{p}$ at the Tevatron, the momentum balance in the transverse plane is natural, to be asked for. Owing to the complex final states, it is not plausible to track down the energy carried away by the products of the collision, that are drained down the beam pipe. With this limitation, the only signature of particles that escape undetected i.e. neutrinos, exotics, is through the momentum balance in the transverse plane. This information is often contaminated with the calorimeter noise (especially in the coarse-hadronic section), pile up events, clusters in the calorimeter that do not end up in jets (unclustered energy), physics objects resolutions etc. Missing transverse energy (MET) is computed by taking the negative vectorial sum of all the energy recorded by calorimeter cells in the transverse plane (above certain threshold). The coarse-hadronic

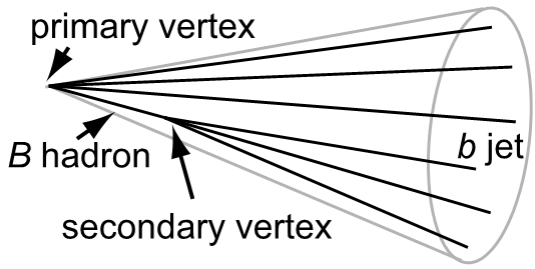
cells are added only if those were considered to form a jet. Hence one can write :

$$\vec{\not{E}}_T = - \sum_{E_T^{icell} > 0.1\,GeV} \vec{E}_T^{icell} \tag{3.8}$$

It can be seen MET ($\vec{\not{E}}_T$) is a vectorial quantity and can be projected into components in any direction in the transverse plane and hence it is also referred to as the missing transverse momentum. $\vec{\not{E}}_T$ is then corrected for the energy corrections, that are applied to the reconstructed objects such as (jet energy scale) and for the momentum of all the muons in an event, corrected for their energy loss in the calorimeter.

3.8 b-jet identification (tagging) and Secondary Vertices

b-quark identification is an important component of this analysis, since $t\bar{t}$ decay products contain two b-quark jets and hence serve as an important discriminator of the signal from rest of the background events. b-jet identification, not only requires good jet identification but also requires combining the information of the tracking system. To achieve this, clusters are formed using the tracking system information, to reconstruct, secondary vertices. Due to their relative long lifetime, B hadrons travel on the order of a mm, before decaying. Thus, the tracks originating from B-hadron decays, seem to come from a displaced vertex. The jet is termed 'taggable' if there are at least two tracks associated with the jet, with at least one hit in the SMT sub-detector system. One of the tracks has a $p_T > 1\,GeV$, while the other has $p_T > 0.5\,GeV$. The longitudinal distance between the secondary vertex and the distance of closest approach ($|z - d.c.a.|$) is required to be less than 0.4 cm.



As seen from the figure above, secondary vertex taggers (SVT) can be used to identify b-jets (Ref.[50]) and aid in the process of 'b-tagging'. There are other identifiable signatures that associate tracks to a corresponding b-jet. At $D\emptyset$, several such variables are combined in an artificial *neural network* (NN) to provide discrimination between b-jets and other light flavor jets. A rank is assigned to the following variables, that enter the NN :

- The impact parameter significance, defined in Sec. 2.3.2, which represents the significance of the distance between the primary vertex and the secondary vertex in the *xy* plane.
- A combined variable based on the number of tracks with specific values of the impact parameter significance
- a probability for all tracks to originate from a primary vertex
- a goodness-of-fit measure χ^2/n_{DF} for the secondary vertex with highest impact parameter significance
- number of tracks that originate from the secondary vertex with highest impact parameter significance
- the invariant mass of the above mentioned secondary vertex
- number of secondary vertices inside a jet.

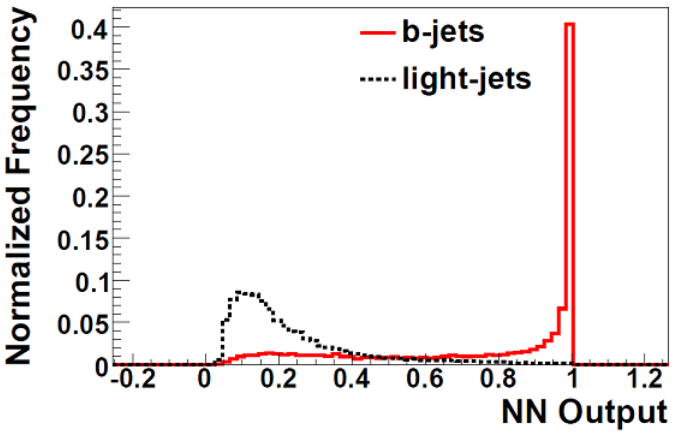
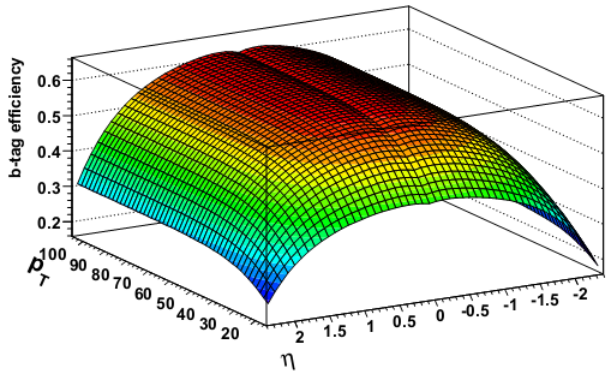


Figure 3.10: Neural network b-tagger output for b-jets and light flavor jets.

The NN , then combines the above variables to provide an output lying in $[0,1]$, per jet. Fig. 3.10 shows the NN output for b-quark jets and light flavor jets in a sample. As can be seen, the b-jets have the NN output peaking closer to 1, in contrast to that of the light flavor jets (u,d,s,g). In this analysis, a cut on the NN output value of 0.65 is applied. This is termed as MEDIUM operating point and has an efficiency 60% in most of the kinematic region. The light jets have an efficiency of less than 2% with this operating point. The $p_{T,}\eta$ based efficiency ,derived from the data is termed as tag-rate functions (TRF). This provides a probability for a particular jet to be "b-tagged", "c-tagged" or light flavor tagged. Such a parametrization for b-jets, is shown in the figure below.



Tag rate function (TRF) for b-jets.

CHAPTER **Four**

Event Selection

This chapter details the Monte Carlo simulation and data samples used for this analysis. It is important to compare the kinematic quantities between data and simulation. For this reason, a large number of events are simulated to describe a certain process and then filled in a histogram to be able to compare to the data. The data sample used in this analysis was collected between June 2006 and September 2009 and corresponds to a total integrated luminosity of about $4.3\,fb^{-1}$ of RunIIb running of $D\emptyset$. There are several detector and physics related steps that need to be taken into account, while simulating a process describing the initial and final states of a collision. This can be broken down into several steps as described below.

4.1 Monte Carlo event generation

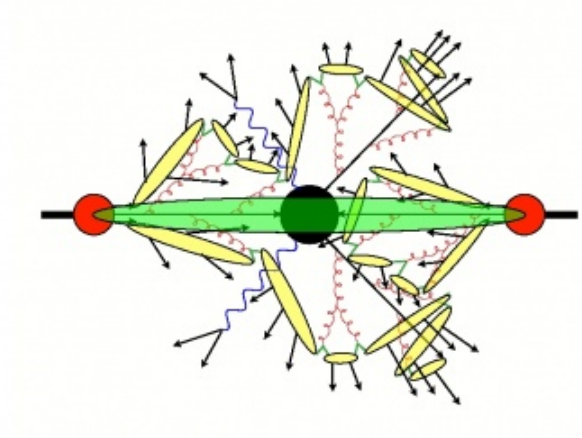


Figure 4.1: Stages of simulation, describing a hadron-hadron collisions by MC event generators [5].

The Monte Carlo method generates a large number of simulated collision events, entailing the final state particles and their momenta. It works via random sampling of a multidimensional phase space of the final state particles, for a given specified theoretical process such as top-anti top production. The idea is to assign certain probability for N points in the multidimensional phase space for a large N, such that the probability corresponds to the actual event produced in the real world. The method can be considered to follow various steps of evolution, starting from the *hard subprocess* , which represents the hard scattering of the partons (inside proton, anti proton) such as valence quarks and is shown as the black sphere in Fig. 4.1. The probability of finding the parton inside the proton or the anti proton, is given by *Parton density functions* and for this analysis, CTEQ6L1 parton distribution functions are used (Ref.[11]). These hard scattering partons carry a large momentum fraction of the incident hadrons and hence the outgoing fundamental particles (top, anti top or new hypothetical particle produced in the black region) represent high momentum scale (q^2). The outgoing color charged quarks or gluons radiate heavily in the process of acceleration

inside a color charged field and split into collinear partons, represented by the brown wavy lines in Fig. 4.1. The process is called *parton showering*. The final state radiation (FSR) is the parton showering of the outgoing partons in the hard sub process, while the initial state radiation (ISR) is the same for the incoming partons of the hard sub process. This happens, until the *virtuality scale* (q^2) reaches the *hadronization scale*, represented by yellow ellipses in the above figure. At this point, quarks get confined into hadrons and this process is modeled by Lund string model, in which the self interacting gluon field, acts as a "string" of color flux. The hadrons thus formed decay further and reach the detector. There are also some partons, that do not take part in the hard scattering process and constitute, what is known as *Underlying event*, represented by the green region in the above figure.

PYTHIA (Ref.[51]), is a leading order generator, which implies that it simulates only the lowest-order terms in the perturbative treatment of the hard sub process under consideration. The parton showering method used in PYTHIA, takes care of mostly soft partons, collinear with the original parton by summation of higher order logarithmic terms through a *Sudakov form factor*, that governs the parton shower evolution. Hence it allows for a good description of the jet substructure but a less accurate modeling of jet multiplicity. ALPGEN (Ref.[52]) is also a leading order generator but the hard sub process, ISR, FSR are modeled by employing a *matrix element* method. This method successfully describes the well separated partons with large transverse momenta, by including the quantum interferences between included diagrams through matrix elements of hard sub processes. Thus the number of radiated quarks or gluons, associated with the hard sub process is always mentioned, when simulating with ALPGEN. Thus a process such as top-anti top production is referred to as $t\bar{t}+nlp$, where $nlp = 0,1,2..etc.$, denotes the number of light partons associated with the hard scattering. ALPGEN models the final state partons pretty well but requires interfacing with another generator such as PYTHIA, to model the formation of jets through showering and hadronization. This process is referred to as ALPGEN+PYTHIA,

henceforth. In this process, PYTHIA is bound to generate additional jets through showering, which may have been already modeled by a higher value of nlp for a final state parton in ALPGEN. To avoid this double counting in the phase space, a MLM matching algorithm is used to determine whether the jets correspond to original final state partons from ALPGEN (Ref.[53])

In this analysis, ALPGEN+PYTHIA was used to generate $t\bar{t}+0lp, t\bar{t}+1lp, t\bar{t}+2lp$ with a top mass of 172.5 GeV. A dynamical factorization scale with $\mu_F^2 = m_t^2 + \sum \left(p_T^{jets}\right)^2$, was used. The cross section for $t\bar{t}$ production was scaled to its next-to-next to leading order (NNLO) value of $7.48^{+0.56}_{-0.72}\,pb$ (Ref.[13]). The main physics background of W+jets was also simulated using ALPGEN+PYTHIA. Three different sub samples were generated for this purpose : $Wb\bar{b}, Wc\bar{c} \& Wq\bar{q} + nlp$. Here, q stands for light quarks (u,d,s) and gluons and $nlp = 0,1,2,3$. Wc sub processes were included in the W+light parton sample with massless charm quarks. The number of $Wb\bar{b}\,and\,Wc\bar{c}$ events were increased, relative to W+light partons to match their respective NLO cross sections. The lesser significant backgrounds of $Z(\rightarrow e^+e^-, \mu^+\mu^-, \tau^+\tau^-) + jets$ were generated with ALPGEN+PYTHIA by breaking up into sub samples of $Zb\bar{b}$,$Zc\bar{c} \& Zq\bar{q}/Zgg$ plus light partons. The Z samples were normalized to their NNLO cross section value of $256\,pb$ (Ref.[54]). This is 1.3 times the value provided by leading order ALPGEN. In a similar fashion, the population of $Zb\bar{b}$ was scaled by a factor of 1.52 and that of $Zc\bar{c}$ by 1.67.

The other backgrounds include single top quark production, which was simulated using COMPHEP-SINGLETOP (Ref.[55]). The top quark mass was set to 172.5 GeV. The cross section for single top quark production was computed to NNLO and NNNLO threshold corrections in the s and $t-channels$ and was found to be $3.3\,pb$ (Ref.[56]). The diboson (WW, WZ, ZZ) were simulated using PYTHIA and scaled to their respective NLO cross sections of $12.3pb$, $3.7pb \& 1.4pb$. (Ref.[57]).

4.2 Detector Simulation

The stable long-lived particles that are produced as the product of the simulation process, described above, interact with the bulk material of the detector. These interactions and the interactions with the magnetic field, allow the particles to be identified. To simulate these interactions, a detailed model of detector composition and geometry , as well as a precise knowledge of solenoidal and toroidal magnetic fields is necessary. The evolution of particles through the detector is based on a software called GEANT3 (Ref.[58]). It uses Monte Carlo methods, random sampling of phase space that is associated with several distinct interactions such as ionization of bulk material, deflection of charged particle in a magnetic field. These are then translated to detector signals that is used as raw data to reconstruct physical objects such as jets, electrons, muons etc.

The Monte Carlo simulation of an event, represents a single collision of proton and anti-proton. In practice, there are multiple $p\bar{p}$ collisions that take place in each beam crossing. The detector signals, thus produced are to be accounted for using the minimum bias background events, which depend on the instantaneous luminosity. Such events collected (as described in Sec.3.4) are overlay-ed on to the MC events. The distribution, thus obtained of the instantaneous luminosity is called $luminosity\,profile$. This accounts for detector occupancy at high luminosity, for residual signals from preceding bunch crossings and the particles present in beam halo.

The events in the MC are not selected based on any triggers on the reconstructed objects, that would fire in case of actual data taking. To account for this, trigger probabilities are derived as a function of kinematically reconstructed objects and combined into a global trigger probability of an event. This parametrization of the trigger probabilities, provides only a coarse approximation of the actual trigger criteria. The temporal and luminosity dependence of such parametrization is taken into account through event weights.

Following the above procedure, some insignificant effects remain to be corrected to match the distribution of the simulated events to that of the data. In general, it is found that the resolutions of the reconstructed objects in the simulated events, are slightly better than those in the data. This can be attributed to the fact that the detector simulation does not take into account, detector aging and electronic noise arising from radiation effects etc. Some of these are accounted for, by re weighting the distribution of events in the simulation for example that of the z_{PV}, which is not perfectly Gaussian in nature for the data. The other inefficiencies in data, such as that of jet identification (JetID), vertex confirmation of jets are accounted for, by applying scale factors to simulated events, which are interpreted as probabilities that are less than 1. In such cases, the scale factors are applied by randomly rejecting the reconstructed objects, for which the inefficiencies in data were found, from the simulated events.

4.3 Event selection and background modeling

This analysis focuses on the semileptonic decay mode of $t\bar{t}$. The final state has the signature of an isolated lepton with high transverse momentum, several jets, large $\not{E}_T$. The following discussion describes the selection based on the reconstructed objects in two separate channels i.e. $e + jets$ and $\mu + jets$. Further, the selection criteria is designed to define a data sample enriched with $W + jets$ and $t\bar{t}$ events and hence events with 4 or more jets are considered. Selection, common to both the channels is as follows :

- Good jets (defined in Sec.3.5.3) emanating from a primary vertex with $|z_{PV}| < 60$ cm and at least 3 tracks attached to it, are required. These are also referred to as vertex confirmed jets.

- At least 4 such jets with $p_T > 20$ GeV and $|\eta| < 2.5$ are required, with an additional

requirement on the jet with highest p_T, also termed as *leading jet*, is required to have $p_{T.} > 40\,GeV$.

For $e + jets$ channel, additional selection requirement is as follows :

- An isolated electron with $p_T > 20\,GeV$ and $|\eta| < 1.1$ is required.
- The electron should have a $|\Delta z(e, PV)| < 1\,cm$.
- No second isolated lepton ($e^{\pm}, \mu^{\pm}$) with $p_T > 15\,GeV$ is allowed.
- Missing transverse momentum ($\not{p}_T$)is required to be greater than 20 GeV
- $\Delta\phi(e, \not{p}_T) > 2.2 - 0.045 \times \not{p}_T$ is required. This is referred to as *traingle cut*, henceforth.

The above selection allows 1002 data events in the $e + jets$ channel to pass through.

For $\mu + jets$ channel, additional selection requirement is as follows :

- An isolated muon with $p_T > 20\,GeV$ is required.
- Invariant mass of the selected muon and any other muon in the event must comply with $m_{\mu\mu} < 70\,GeV$ or $m_{\mu\mu} > 110\,GeV$ to reject any $Z(\rightarrow \mu\mu) + jets$ events.
- No second isolated lepton with $p_T > 15\,GeV$ is allowed.
- $|\Delta z(\mu, PV)| < 1\,cm$ is required as in $e + jets$ channel
- $\not{p}_T > 25\,GeV$ is required
- The triangle cut takes the form, $\Delta\phi(\mu, \not{p}_T) > 2.1 - 0.035 \times \not{p}_T$ in this channel.

A total of 807 events, survive the above selection cuts in the data.

The main two standard model processes that produce events with an isolated lepton, $\not{p}_T$ and at least four jets are $t\bar{t}$ production and $W + jets$ production. Single top, $Z + jets$

and diboson production can also give rise to such a final state but with much smaller cross sections. The next most important source of events are the *multijet* events in which the energy depositions from a jet is mis-identified as a lepton and the $\not{p}_T$ is mismeasured.

The multijet background model is estimated from the multijet events that enter the final data sample using measured selection efficiencies and a data sample, termed as '*loose*', that is a superset of the final data sample for this analysis. The loose sample is obtained by using less stringent cuts for the identification of the leptons. The number of events in the loose and final data samples are denoted by $N'$and N, respectively. N^{lj}denotes the combined number of events with genuine leptons in the loose sample, N^{jj} corresponds to the number of multijet events in the loose sample, ε_l is the efficiency for a lepton in the loose sample to also pass the final lepton selection, and ε_j is the efficiency for a misidentified jet in the loose sample to also pass the final lepton selection. With these definitions, one can write the following equations :

$$N' = N^{lj} + N^{jj} \quad and \quad N = \varepsilon_l N^{lj} + \varepsilon_j N^{jj}$$

Solving the above system of equations for N^{jl} , N^{jj} yields :

$$N^{lj} = \frac{N - \varepsilon_j N'}{\varepsilon_l - \varepsilon_j} \qquad and \qquad N^{jj} = \frac{\varepsilon_l N' - N}{\varepsilon_l - \varepsilon_j}$$

The efficiency ε_l for the true leptons is found to be $86.9 \pm 2.2\%$ for the $e + jets$ events and $93.9 \pm 2.2\%$ for $\mu + jets$ events. These numbers were obtained from the corresponding $W + jets$ and $t\bar{t}$ Monte Carlo samples. The efficiency ε_j for jets that get mis-identified as lepton, to pass the isolation selection was measured directly from the data. For this purpose, events with $\not{p}_T < 10 GeV$, which are dominated by misidentified leptons were used to calculate ε_j, as the ratio of number of events in the final and in the loose data samples. For $e + jets$ events, $\varepsilon_j = 13.0 \pm 3.0\%$ and for $\mu + jets$ events, $\varepsilon_j = 30.6 \pm 3.1\%$ was estimated.

The multijet background along with the other backgrounds (except $W + jets$), based on their respective cross section, is subtracted from the data. The resultant is then compared to the expected number of $W + jets$ events, to make up the resulting composition of the final data sample. The data sample compostion the the 2 channels is shown in Table 4.1

source	number of e+jets events	number of μ+jets events
$t\bar{t}$ production	573	401
single t production	10	7
W+jets	255	338
Z+jets	29	28
WW, WZ, ZZ+jets	20	16
multijets	115	16
data	1002	807

Table 4.1: Composition of the final data sample in $e + jets$ and $\mu + jets$ channel

4.4 Data and Monte Carlo comparison

Pseudo-data samples are constructed from the MC events to calibrate mass in the analysis. For this purpose, it is important that the distributions of topological and kinematic quantities agree between data and simulation. Such comparisons in various channels is shown below, after normalizing various components according to Table 5.1

- e+jets channel

Fig. 4.2, shows the Data-MC agreement for distribution of certain topological variables, which are defined in Chapter.6.

Fig. 4.3, shows the same for certain kinematic distributions. Fig. 4.4 depicts the distributions of other interesting quantities such as the W boson transverse mass and invariant mass of $t\bar{t}$ system in the data composed of simulated signals and various backgrounds (color code presented in the legend of the plots).

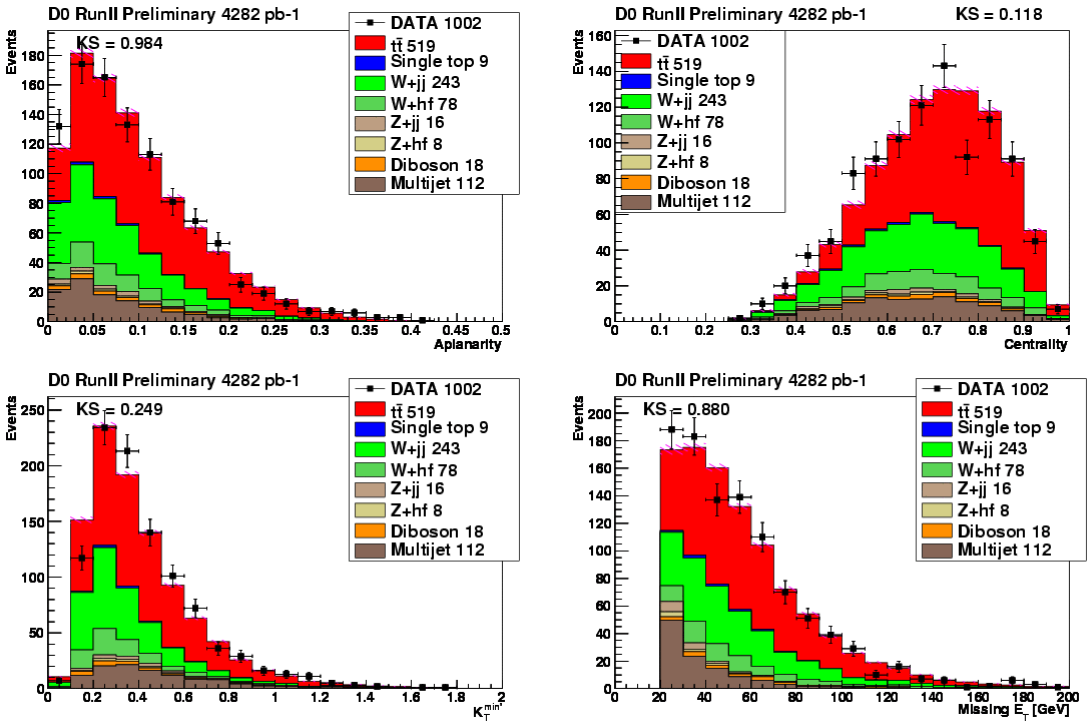


Figure 4.2: Data-MC comparison for topological variables : Aplanarity, Centrality, $K_T^{min'}$, $\not\!\!E_T$ in $e + jets$ channel

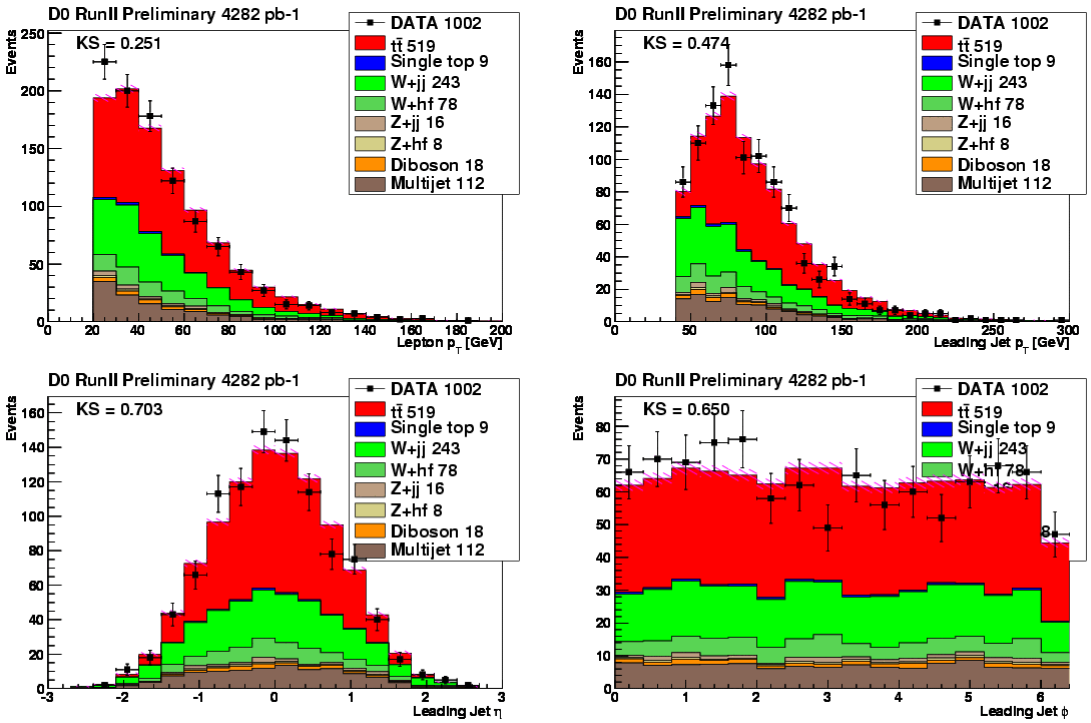


Figure 4.3: Data-MC comparison for lepton p_T, leading jet p_T, η, ϕ in $e + jets$ channel

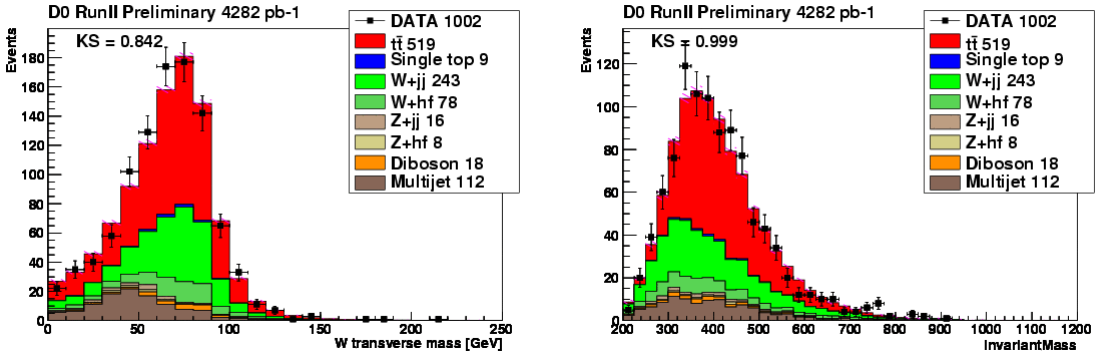


Figure 4.4: W transverse mass and invariant top mass distributions

- $\mu + jets$ channel

Based on the sample composition, described in Table 5.1, the control plots for the topological and kinematic quantities in $\mu + jets$ channel are shown in Fig.4.5-4.7.

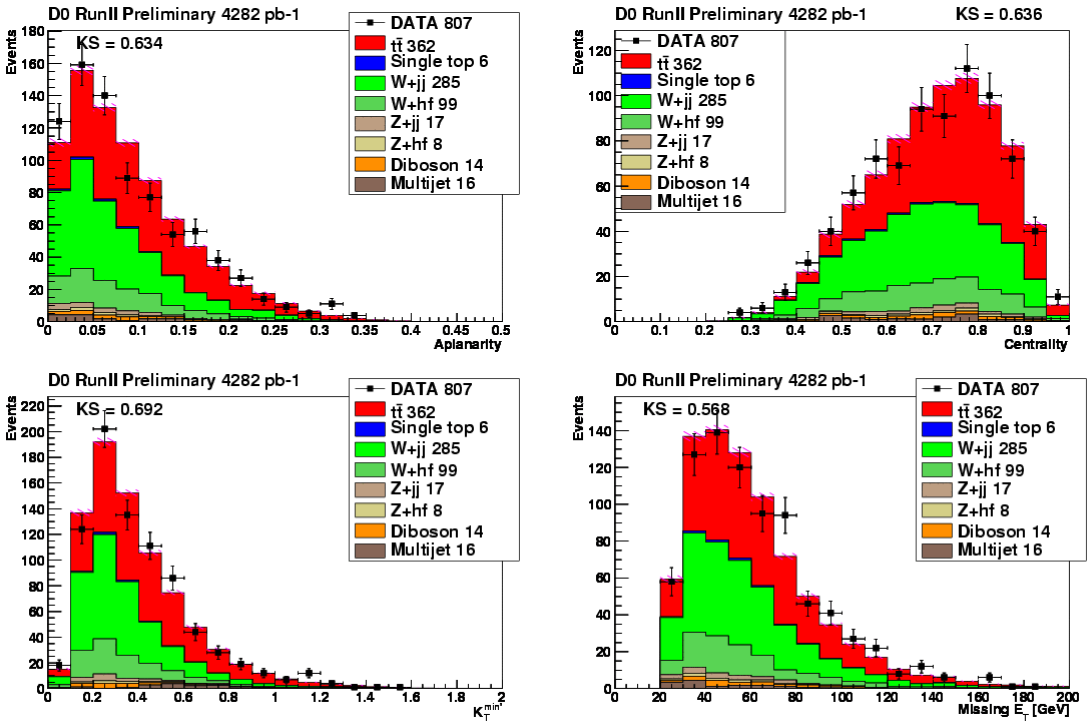


Figure 4.5: Data-MC comparison for topological variables : Aplanarity, Centrality, $K_T^{min'}$, $\not\!\!E_T$ in $\mu + jets$ channel

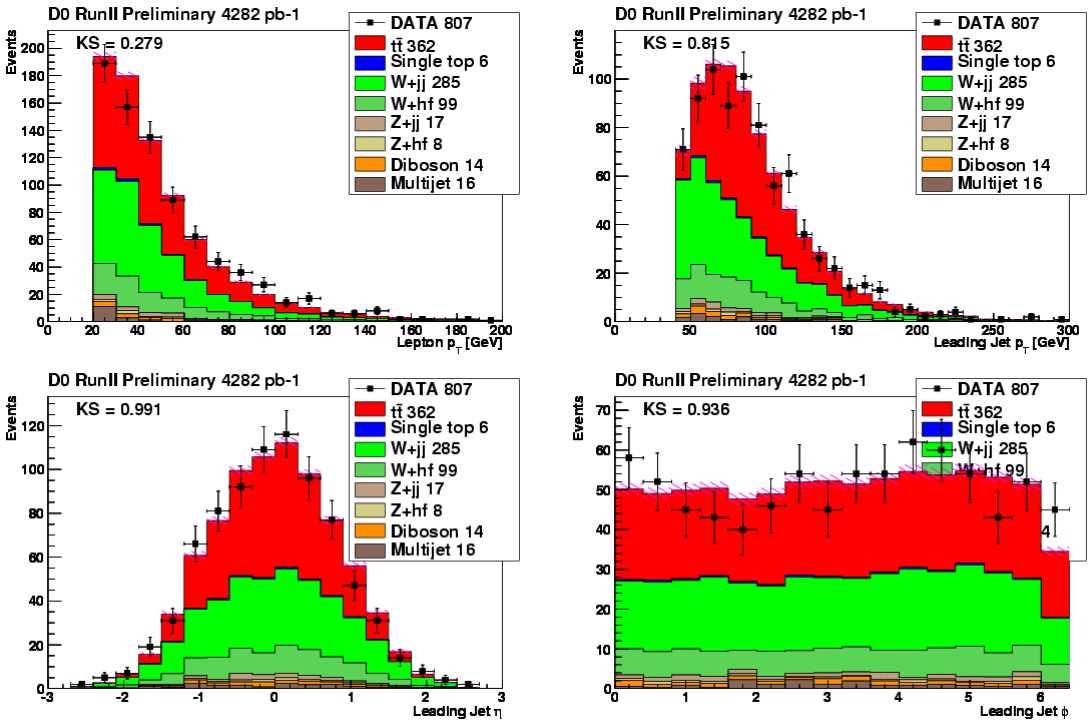


Figure 4.6: Data-MC comparison for lepton p_T, leading jet p_T, η, ϕ in $\mu + jets$ channel

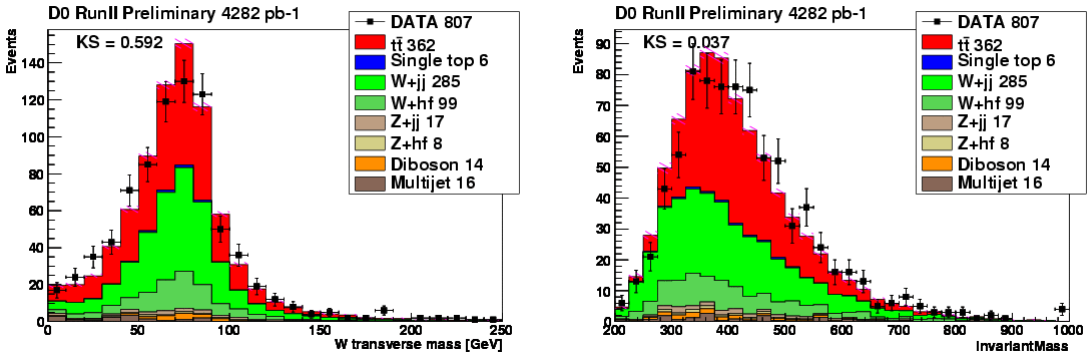


Figure 4.7: W transverse mass and invariant top mass distributions in $\mu + jets$ channel

4.4.1 Data-MC comparison based on b-tagging

Often times, as in this analysis, it is useful to split the samples in the order of the signal content. Since, $t\bar{t}$ decay products contain two b-quarks, the two channels under consideration ($e + jets$, $\mu + jets$) are further split according to the b-quark content. In this analysis,

the events are further sieved through in two different channels, using the NN b-tagging output (defined in Sec.3.8). For an NN output value greater than or equal to 0.65, the jet in the event is considered b-tagged. This allows to have different sample composition in two channels, characterized as follows :

- two or more b-tagged jets found in the event ($\geq 2tag$)
- exactly one b-tagged jet found in the event (1 tag)

The signal content in each of the 4 channels $[(1, \geq 2)tags \otimes (e, \mu) + jets]$ above, is good enough to perform calibrations of top mass in each of them separately. Following plots show the data-MC comparison of a few of the selected topological variables that are used in Chapter 7, to discriminate between signal and background.

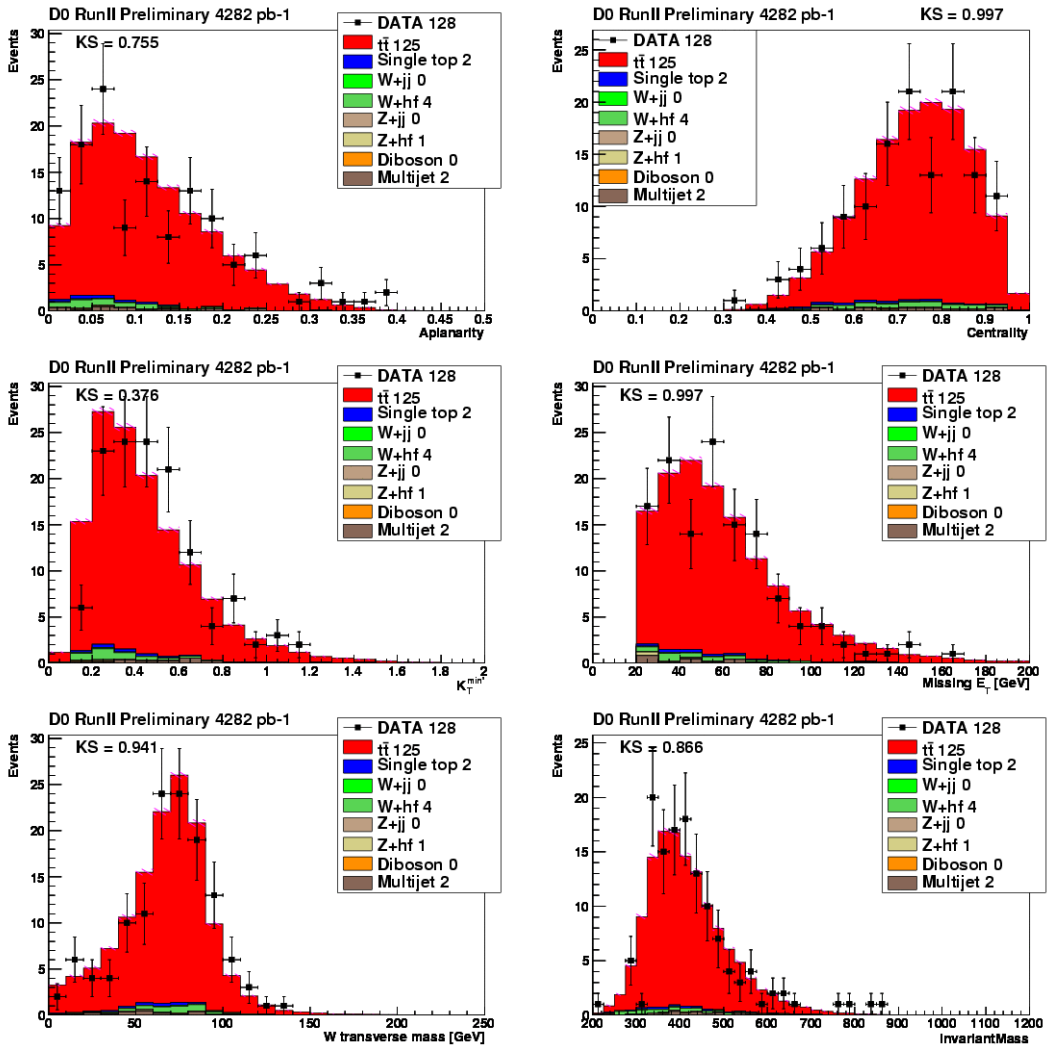


Figure 4.8: Data-MC comparison of Aplanarity, Centrality, $K_T^{min'}$,$\not\!\!E_T$, W transverse mass and $t\bar{t}$ invariant mass in $e+jets, \geq 2tags$ channel

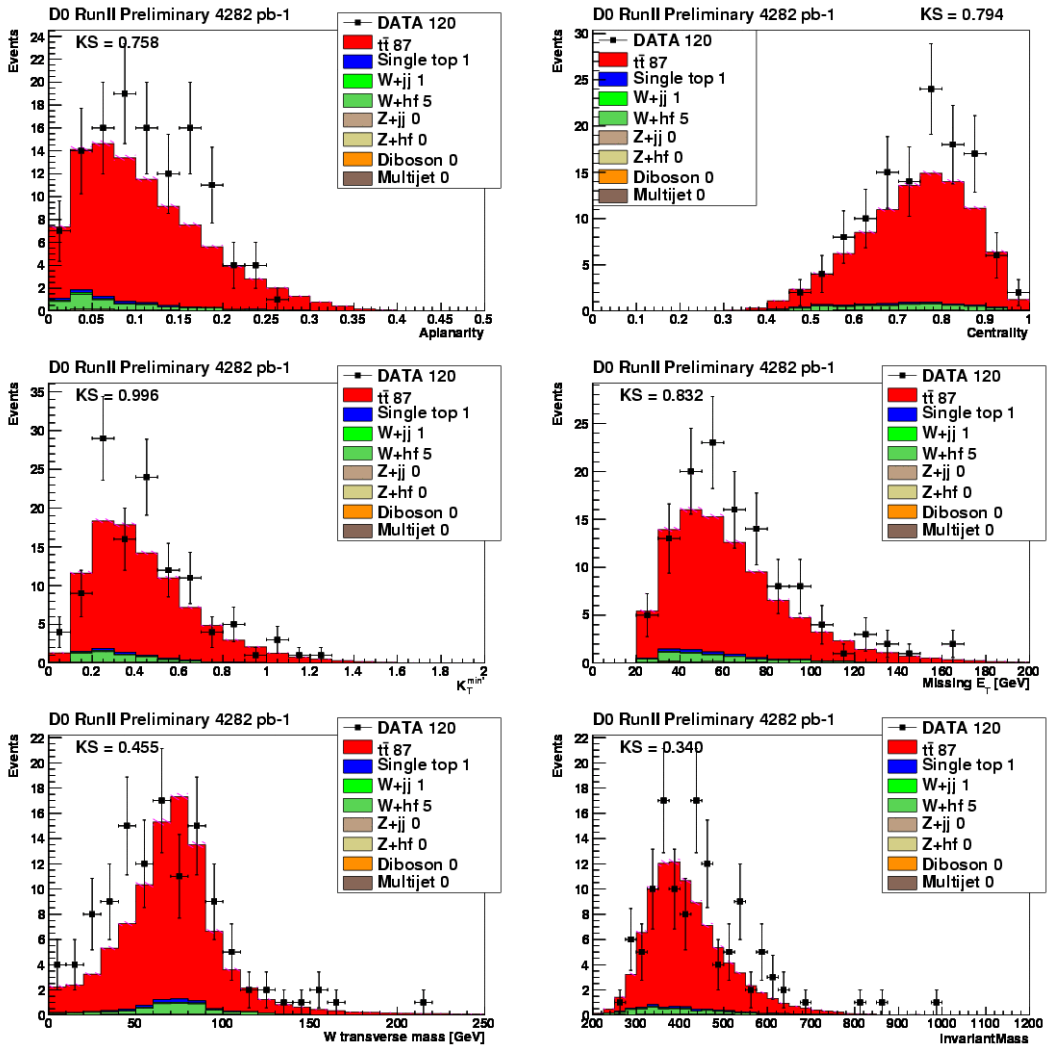


Figure 4.9: Data-MC comparison of Aplanarity, Centrality, $K_T^{min'}$,$\not\!\!E_T$, W transverse mass and $t\bar{t}$ invariant mass in $\mu + jets, \geq 2tags$ channel

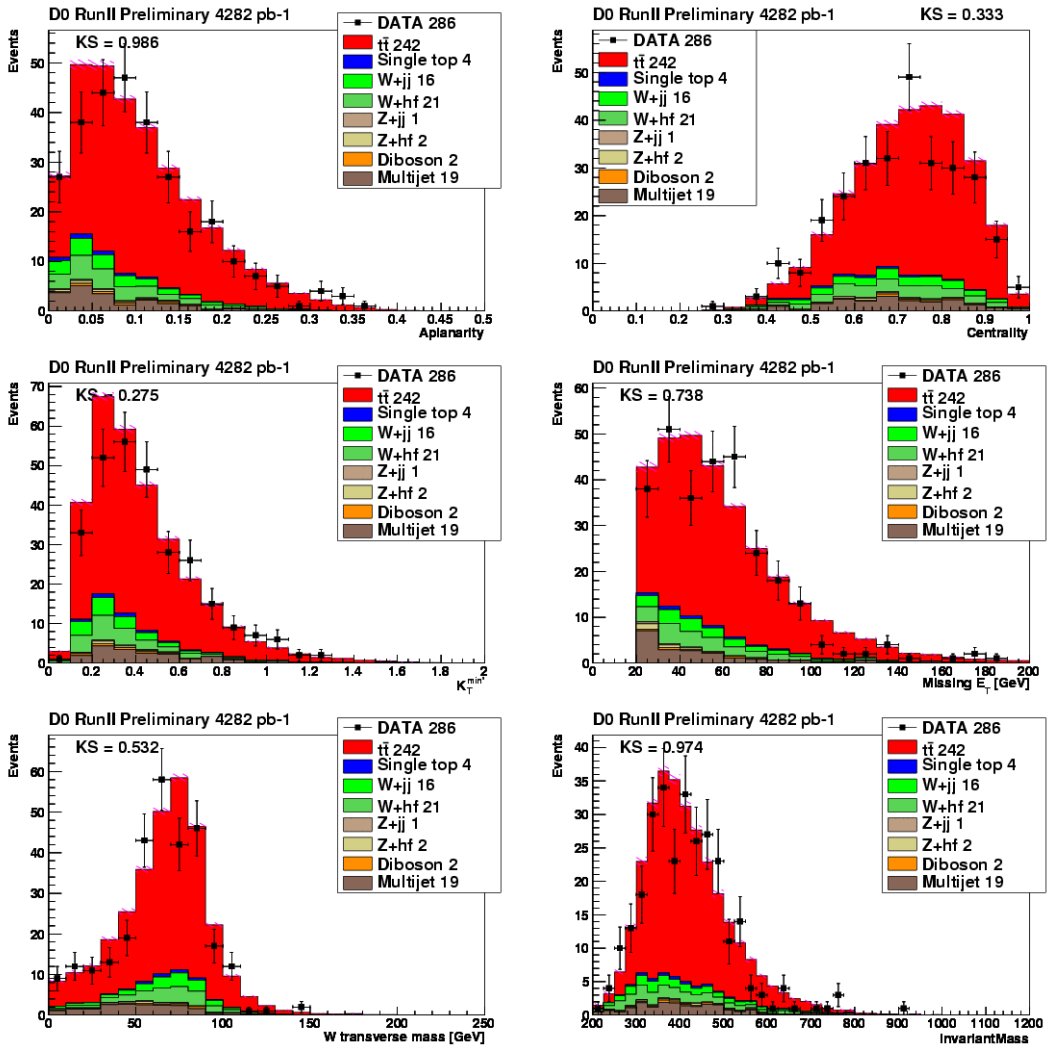


Figure 4.10: Data-MC comparison of Aplanarity, Centrality, $K_T^{min'}$,$\not\!\!E_T$, W transverse mass and $t\bar{t}$ invariant mass in $e+jets$, $1\,tag$ channel

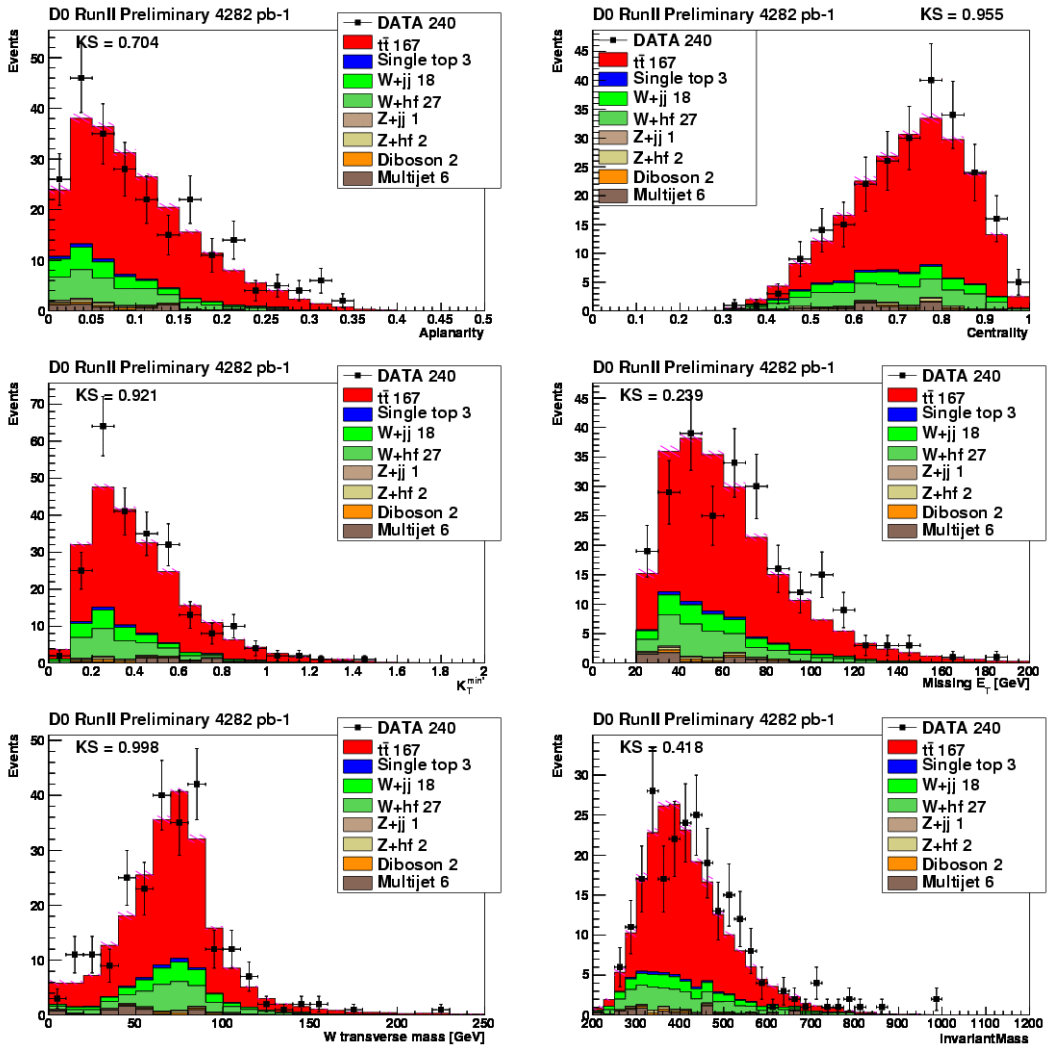


Figure 4.11: Data-MC comparison of Aplanarity, Centrality, $K_T^{min'}$,$\not\!\!E_T$, W transverse mass and $t\bar{t}$ invariant mass in $\mu + jets$, $1\,tag$ channel

CHAPTER **Five**

Kinematic Fitting for Top Mass

5.1 Introduction

The top quark event selection, as described in the previous chapter, selects events with four or more jets. In an ideal scenario, two jets arise from a b-quark and an anti b-quark and two from light quarks that are the result of a W boson decay. If the correspondence between jets and partons from the top decay were known, the top quark event could easily be reconstructed from the 4-vectors of the final state products. In practice this is not known and therefore all possible jet-parton assignments have to be considered as the input to a fitter that uses goodness-of-fit χ^2 to give a measure of how far the input quantities have to be moved to respect the constraints of a top quark event. Such a measure can be estimated for each jet-parton assignment. For four jets in an event, this amounts to 12 possible combinations per event [choice of two out of the 4 jets arising from W boson decay, one of the remaining two, originating from a b-quark while the other from anti b-quark $= \binom{4}{2} \times \binom{2}{1}$]. This ambiguity increases if more than 4 jets are used for assignments to the partons from top decay. Only the 4 jets with highest momenta are chosen as input to the fitter. This allows one to deal with manageable ambiguities. These 4 jets are often referred to as leading

jets, in the following discussion. For the reconstruction of the top quark event, one requires the 4-momenta of the partons from top decay rather than the 4-vector of the jets observed. This mapping is defined in the following sections.

5.2 Parton Level Corrections

The term Parton Level Corrections (PLC), Ref.[59, 60], is used to describe the relation between initial state partons and the final state partons. The final state partons undergo hadronization, fragmentation and hard gluon radiation. These final state partons manifests themselves as jets of particles whose response can be measured in the DØ calorimeter. The PLC's, correct the energy of the jets to the energy of the partons, from which they originated.

As described in Sec.3.6.1, the heavy-quark hadronization differs from the light-quark hadronization. The PLC are thus derived separately for jets originating from light quarks (*u,d,s,c*) and b-quarks. Also, the response varies with detector η and hence four η-ranges, crudely classifying central, inter-cryostat and forward end-cap regions are used to derive these corrections. $|\eta| \in [0, 0.5)$ is denoted as Region 1, $|\eta| \in [0.5, 1)$ is denoted as Region 2, $|\eta| \in [1, 1.5)$ as Region 3, and $|\eta| > 1.5$ as Region 4. The corrections are derived using Monte Carlo simulation of the process $t\bar{t} \rightarrow l + jets$ (l=e, μ) which are generated for top quark masses 150, 160, 165, 170, 172.5, 180, 185 and 190 GeV. The hard scattering process is modeled by the event generator ALPGEN+PYTHIA, as described in Sec.4.1. These generated events are then processed through the whole DØ reconstruction algorithm.

Using the Monte Carlo information, the primary partons from $t\bar{t}$ decay are matched to the jets using a jet cone algorithm with cone size $R = 0.5$. The jets are selected to be isolated with respect to all other jet objects by at least $\triangle R = 0.5$, and also matched to the final state parton within $\triangle R = 0.5$. The event selection is similar to the one described in

Chapter 5, albeit a bit relaxed.

- At least four jets with $p_T > 15$ GeV each with $|\eta| < 2.5$
- exactly one isolated charged lepton (e or μ) with $p_T > 20$ GeV and $|\eta| < 1.1$ (for electrons), $|\eta| < 2.0$ (for muons)
- Missing Transverse Energy ($\not{E}_T$) > 20 GeV

The closure tests for the corrections derived using these events are described later in this chapter. The following generator level information is used for closure :

1. The W boson mass formed by the two light quarks is compared to the nominal W boson mass of 80.4 GeV.
2. The top quark mass, which is obtained from the kinematic fit is compared to the generator level mass of the various $t\bar{t}$ samples.

Since, eventually, a simultaneous fit to the jet energy scale factor (α_{JES}, a multiplicative factor by which the energy of all jets in an event are scaled) is done in-situ, it is important to parametrize the corrections with respect to the correct choice of variables. The most likely factor is extracted from the data/simulation by using the hadronically decaying W boson in the event. The output JES is equal to input JES if, for a reconstructed W boson mass:

$$M_W(\alpha_{JES} \times E_1^{parton}, \alpha_{JES} \times E_2^{parton}) \quad = \quad \alpha_{JES} \times M_W(E_1^{parton}, E_2^{parton}) \qquad (5.1)$$

wherein, $E_1^{parton}, E_2^{parton}$ are the parton level corrected jet energies of the two jets that are used to reconstruct the W boson. The left hand side of the above equation being the

reconstructed W boson mass at the α_{JES}, while the right hand side being the reconstructed W boson mass at JES=1. Since the PLC (as will be seen latter) are nonlinear, we derive them as a function of jet energies so that Eq.5.1 holds true at least approximately. This is also the reason for using a relaxed event selection for deriving these corrections.

5.2.1 Light Quark Corrections

Based on the framework, described above, the matched jet-parton pairs for light quarks in the $t\bar{t}$ events are used to plot parton energy distributions in various jet energy bins. The choice of these bin ranges is made separately in each detector η region according to the statistics available. Appendix A shows the parton energy distribution in different jet energy bins. Each of these distributions is fitted with a Gaussian near the peak and the mean of the Gaussian is taken as the most likely parton energy value (E_{parton}) for that jet energy bin. The values of E_{parton} are then fitted with a polynomial of up to 6th degree as a function of E_{jet}, as shown in Fig.5.1-5.4. E_{jet} is the mean jet energy of the bin. The fits define the mapping of E_{jet} to E_{parton} for light quark jets.

$\lvert\eta^{det}\rvert$	p0(GeV)	p1(GeV)	p2(GeV)	p3(GeV)	p4(GeV)	p5(GeV)	p6(GeV)
$[0,0.5)$	7.85	0.75	0.00478	-6.42×10^{-5}	4.55×10^{-7}	-1.56×10^{-9}	1.99×10^{-12}
$[0.5,1.0)$	6.80	0.86	0.000867	-6.66×10^{-6}	3.42×10^{-8}	-8.53×10^{-11}	0.0
$[1.0,1.5)$	8.13	0.85	0.000467	-1.01×10^{-7}	-5.79×10^{-9}	3.96×10^{-12}	0.0
$[1.5,2.5)$	15.13	0.75	0.001385	-3.09×10^{-6}	0.0	0.0	0.0

Table 5.1: Fit parameters for $E_{parton} = p0 + p1\times E_{jet} + p2\times E_{jet}^2 + p3\times E_{jet}^3 + p4\times E_{jet}^4 + p5\times E_{jet}^5 + p6\times E_{jet}^6$ for light quarks.

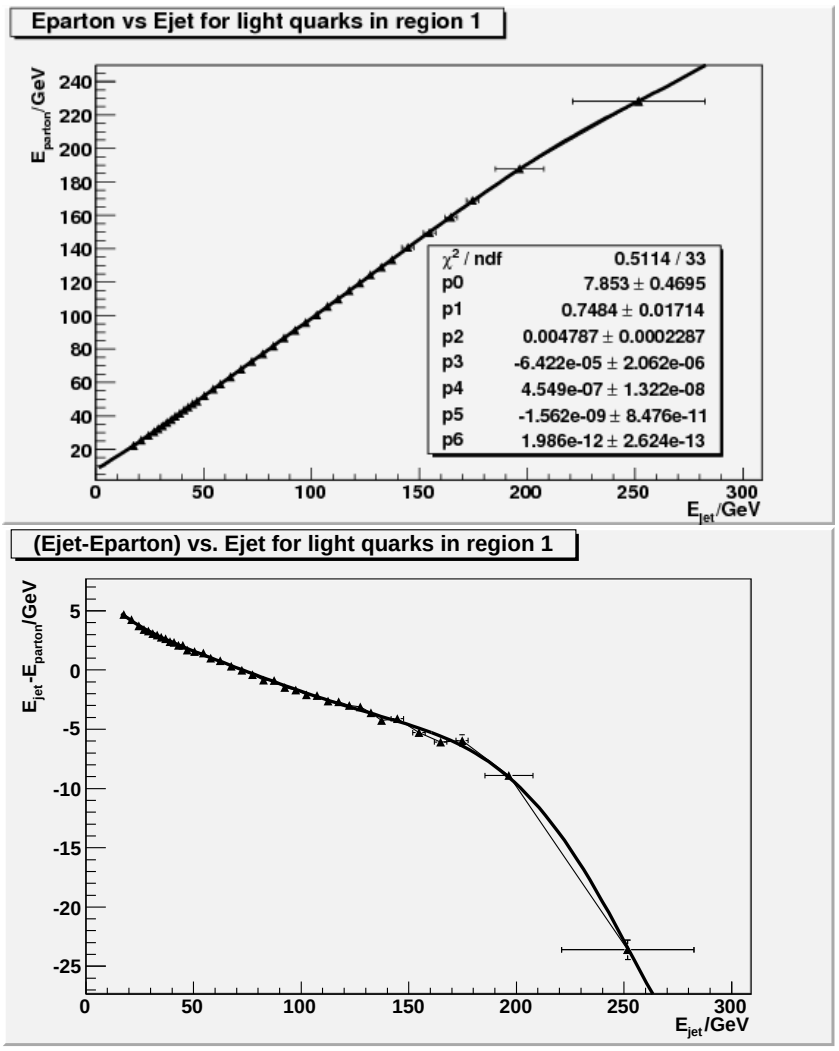


Figure 5.1: $E_{parton} vs. E_{jet}$ & $(E_{parton} - E_{jet})$ $vs.$ E_{jet} for light quark jets in Region 1, $|\eta| \in [0, 0.5)$.

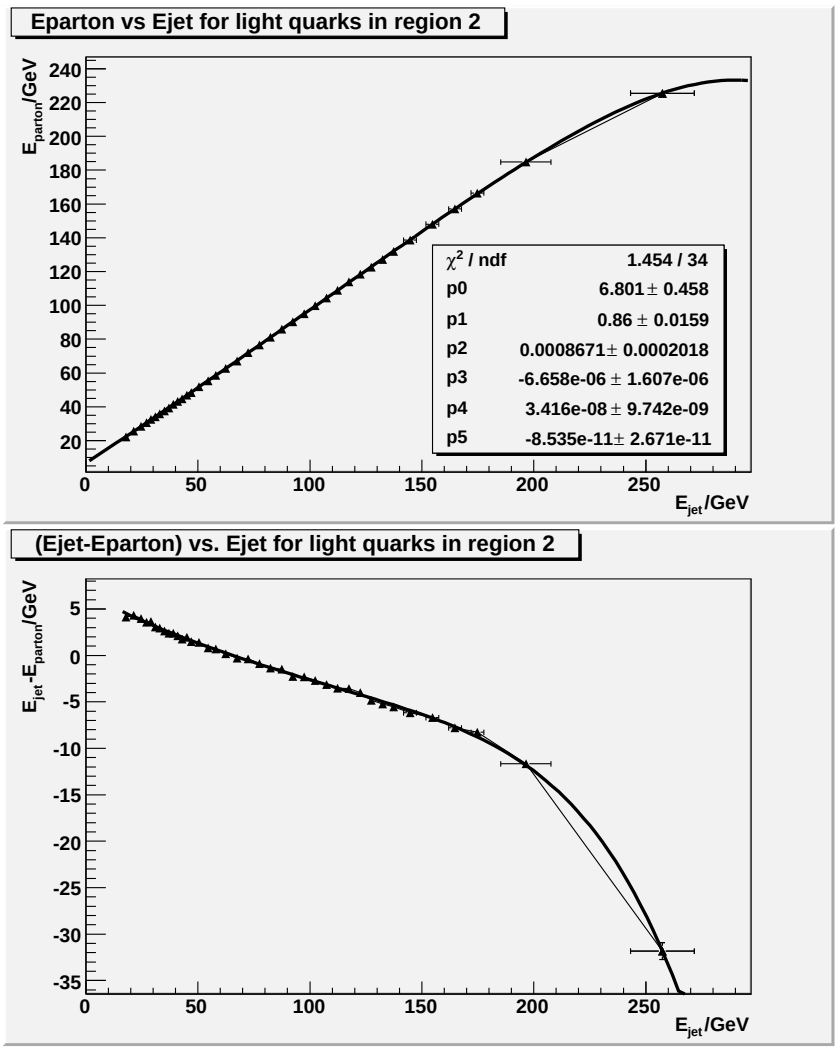


Figure 5.2: $E_{parton} vs. E_{jet}$ & $(E_{parton} - E_{jet})\, vs. E_{jet}$ for light quark jets in Region 2, $|\eta| \in [0.5, 1.0)$

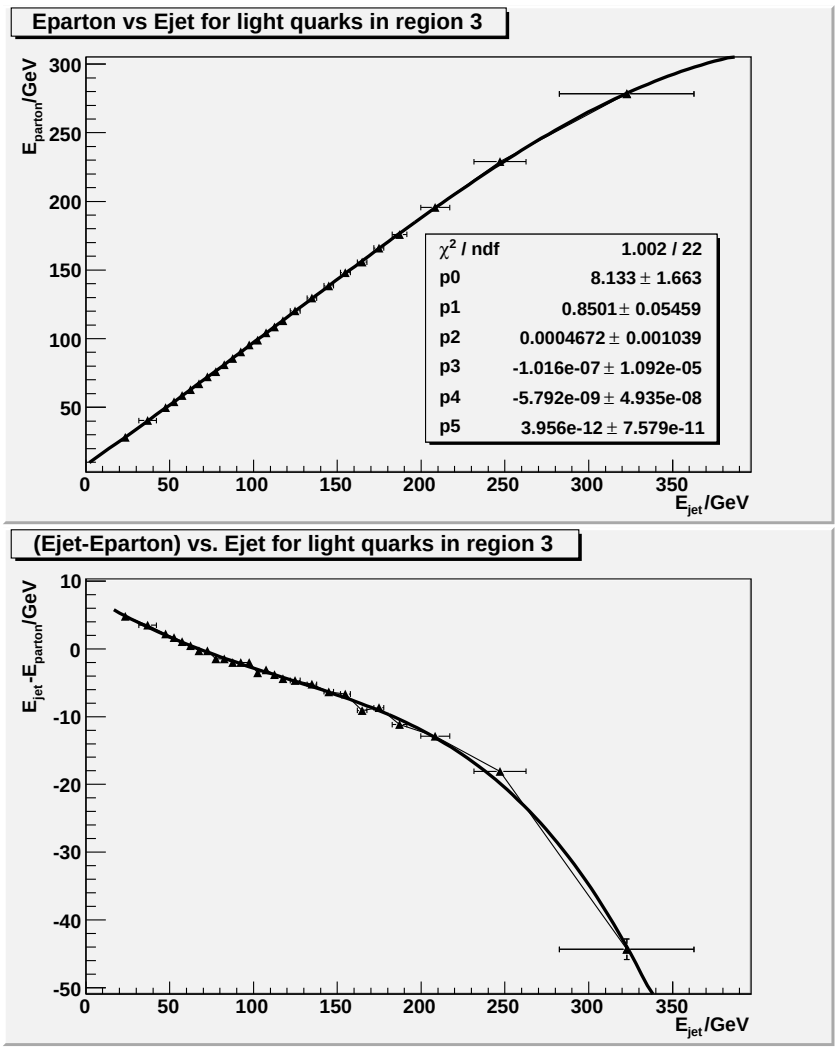


Figure 5.3: $E_{parton} vs. E_{jet}$ & $(E_{parton} - E_{jet})\, vs. E_{jet}$ for light quark jets in Region 3, $|\eta| \in [1.0, 1.5)$

5.2.2 b-quark Corrections

A similar procedure, as for light quark jets is used for the matched jet-parton pairs of b-quarks in $t\bar{t}$ events to derive the Parton Level Corrections. The Table 5.2 shows the fit

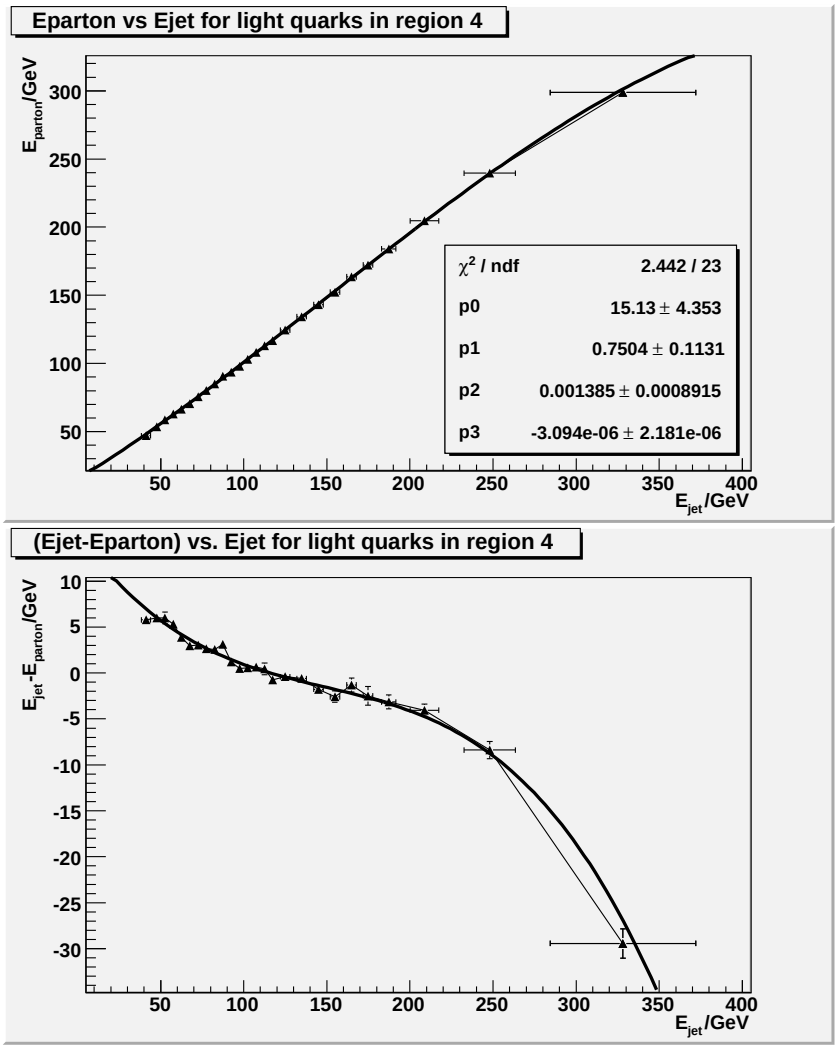


Figure 5.4: $E_{parton}\,vs.\,E_{jet}$ & $(E_{parton}-E_{jet})\,vs.\,E_{jet}$ for light quark jets in Region 4, $|\eta| \in [1.5, 2.5)$.

parameters. Appendix B shows the parton energy distributions in jet energy bins for the b-quarks. Fig.5.5-5.8, show the fits to E_{parton}vs. E_{jet} for b-quark jets.

$\lvert\eta^{det}\rvert$	p0(GeV)	p1(GeV)	p2(GeV)	p3(GeV)	p4(GeV)	p5(GeV)	p6(GeV)
$[0,0.5)$	16.85	0.674	0.00225	-6.33×10^{-6}	0.0	0.0	0.0
$[0.5,1.0)$	17.03	0.721	0.00167	-4.71×10^{-6}	0.0	0.0	0.0
$[1.0,1.5)$	12.73	1.047	-0.00286	2.52×10^{-5}	-9.52×10^{-8}	1.22×10^{-10}	0.0
$[1.5,2.5)$	10.73	1.409	-0.00876	6.22×10^{-5}	-1.92×10^{-7}	2.11×10^{-10}	0.0

Table 5.2: Fit parameters for $E_{parton} = p0 + p1\times E_{jet} + p2\times E_{jet}^2 + p3\times E_{jet}^3 + p4\times E_{jet}^4 + p5\times E_{jet}^5 + p6\times E_{jet}^6$ for b-quarks.

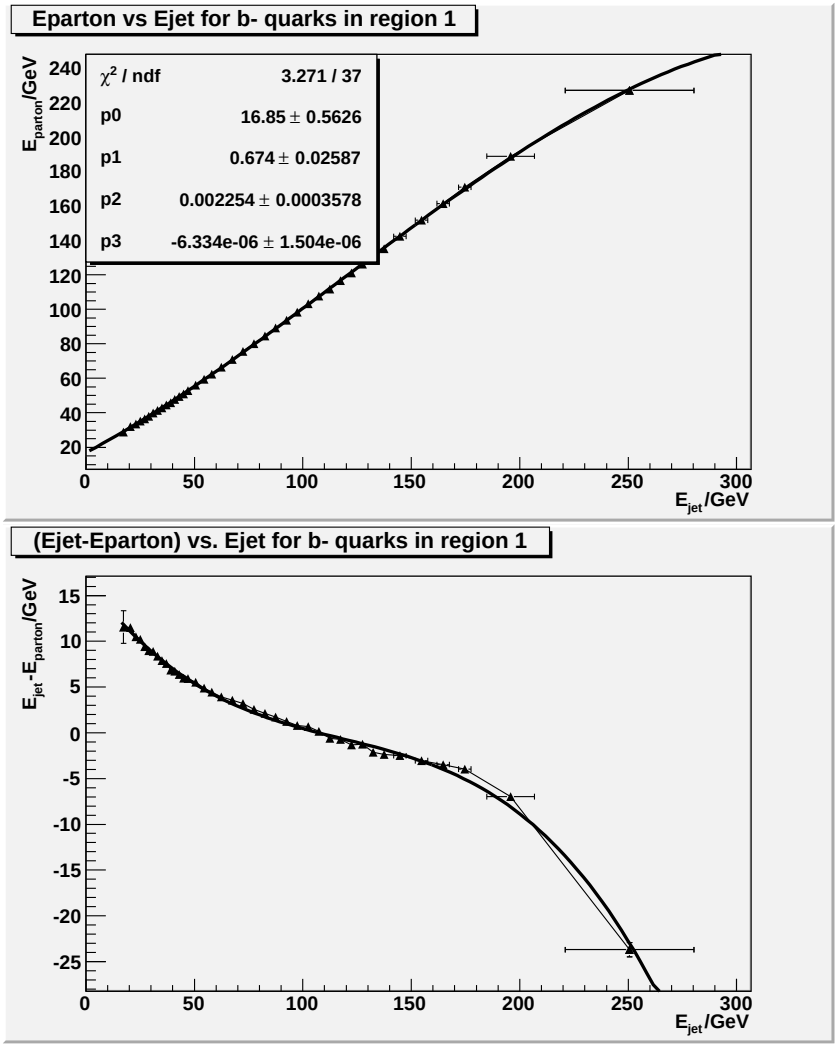


Figure 5.5: $E_{parton}\,vs.\,E_{jet}$ & $(E_{parton}-E_{jet})\,vs.\,E_{jet}$ for b-quark jets in Region 1, $|\eta| \in [0,0.5)$

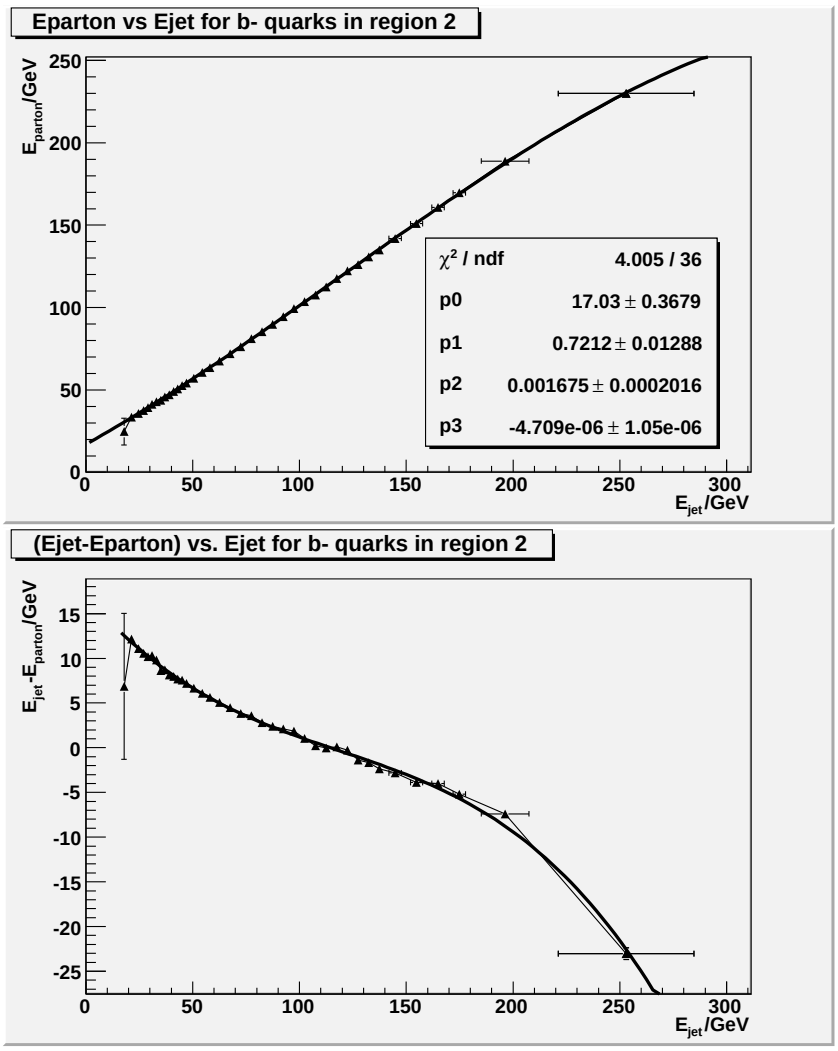


Figure 5.6: $E_{parton} vs. E_{jet}$ & $(E_{parton} - E_{jet})$ $vs.$ E_{jet} for b-quark jets in Region 2, $|\eta| \in [0.5, 1.0)$

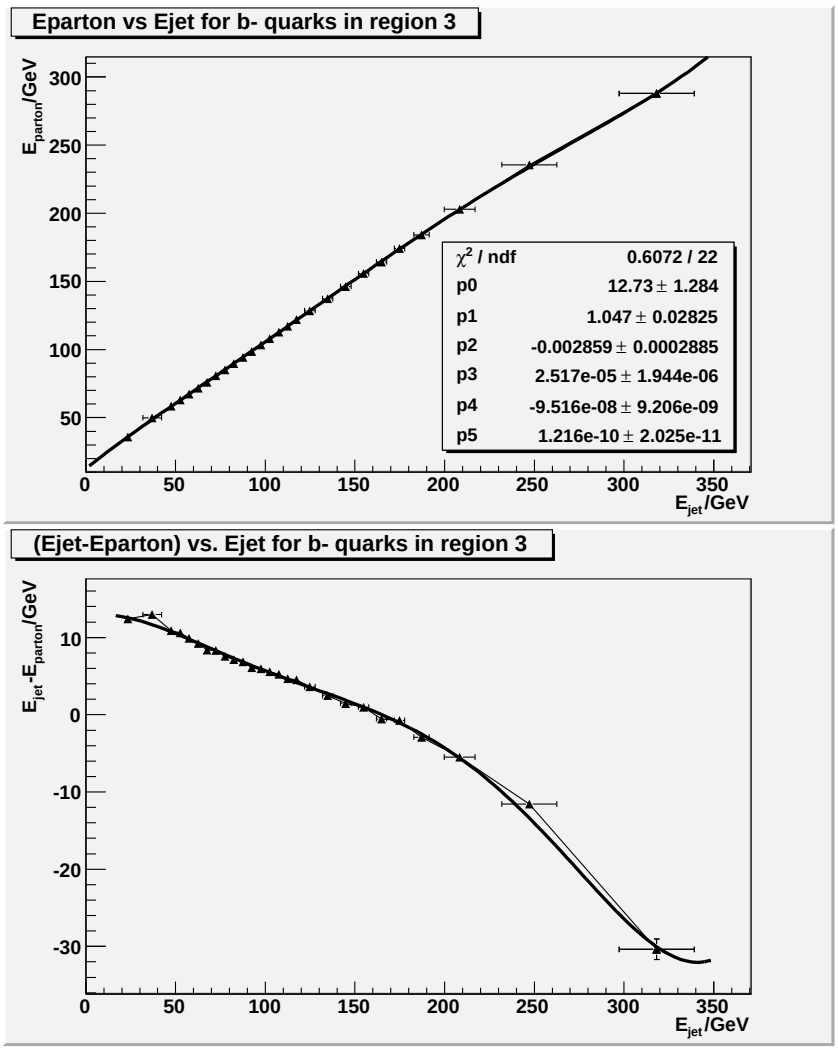


Figure 5.7: $E_{parton} vs. E_{jet}$ & $(E_{parton} - E_{jet}) vs.\ E_{jet}$ for b-quark jets in Region 3, $|\eta| \in [1.0, 1.5)$

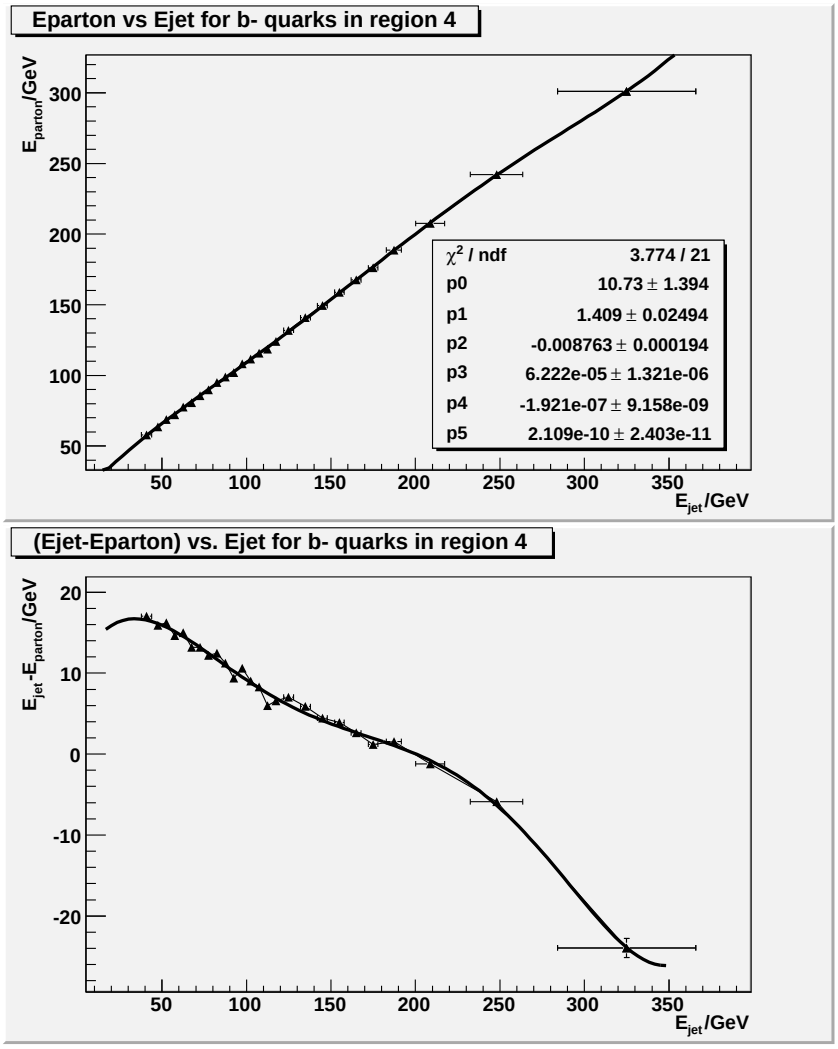


Figure 5.8: $E_{parton} vs. E_{jet}$ & $(E_{parton} - E_{jet}) vs.\ E_{jet}$ for b-quark jets in Region 4, $|\eta| \in [1.5, 2.5)$

5.3 Resolution Functions

The kinematic fitter described later in this chapter, uses resolution functions in the computation of χ^2. These are the uncertainties assigned to the kinematic variables of the reconstructed objects within which they are allowed to vary to reach an optimal solution for an event, respecting the given constraints. In the following sections, these resolution

functions are presented for parton level corrected jet energies and the angular jet variables pseudorapidity (η) and azimuthal (ϕ), for the light and b-quarks.

5.3.1 Light Quark Resolution

Since the jets are parton level corrected, the uncertainty on the parton energy for each jet energy bin is taken as the RMS of the Gaussian peak fitted to the parton energy distributions, shown in Appendix A (for the light quarks). This provides a parton energy resolution parametrized as a function of jet energy. Table 5.3 shows the parameters for the polynomial fit to $\sigma(E_{parton})\, vs.\, E_{jet}$, which is shown in Fig.5.9-5.10.

$\lvert\eta^{det}\rvert$	p0(GeV)	p1(GeV)	p2(GeV)	p3(GeV)	p4(GeV)	p5(GeV)
$[0, 0.5)$	4.55	0.05646	0.0001228	-4.55×10^{-6}	3.033×10^{-8}	-5.076×10^{-11}
$[0.5, 1.0)$	5.017	0.08429	-0.0004756	1.854×10^{-6}	0.0	0.0
$[1.0, 1.5)$	0.2849	0.5201	-0.007879	5.652×10^{-5}	-1.773×10^{-7}	2.033×10^{-11}
$[1.5, 2.5)$	-7.266	0.8105	-0.009131	4.847×10^{-5}	-1.123×10^{-7}	9.323×10^{-11}

Table 5.3: Parameters of polynomial fit to $\sigma(E_{parton})\, vs.\, E_{jet}$ for light quarks

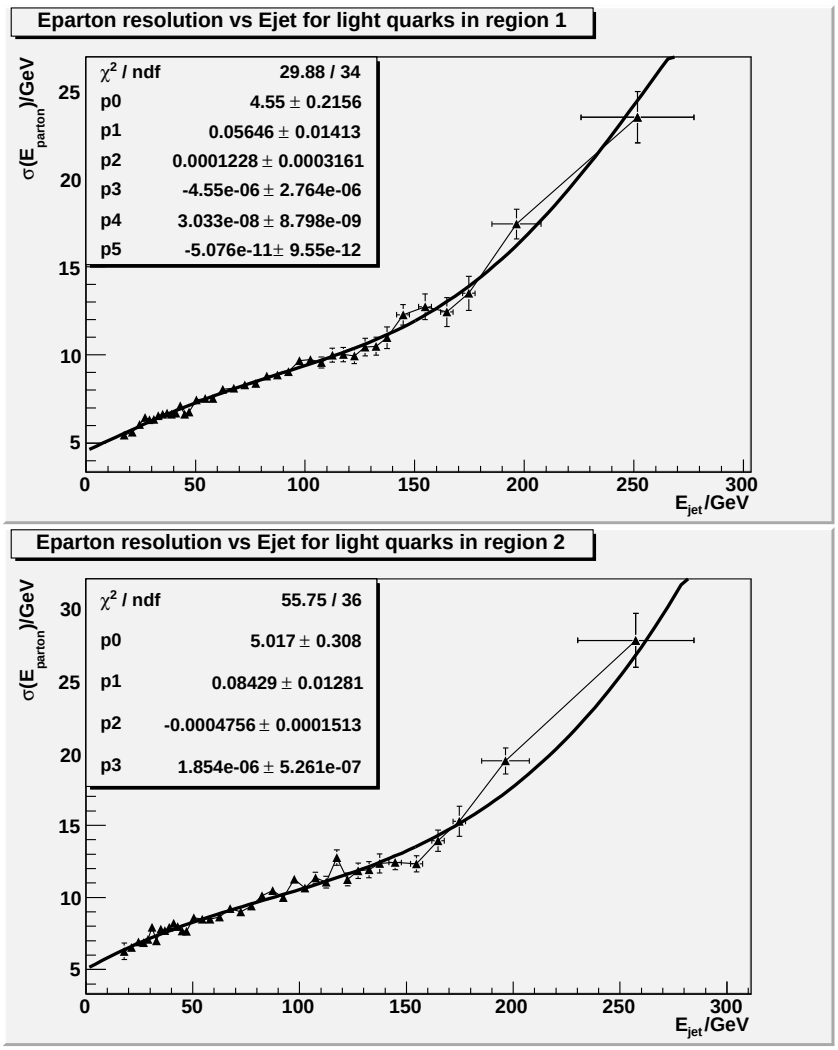


Figure 5.9: Energy resolution for light quarks in the detector η regions 1 and 2, plotted versus E_{jet}.

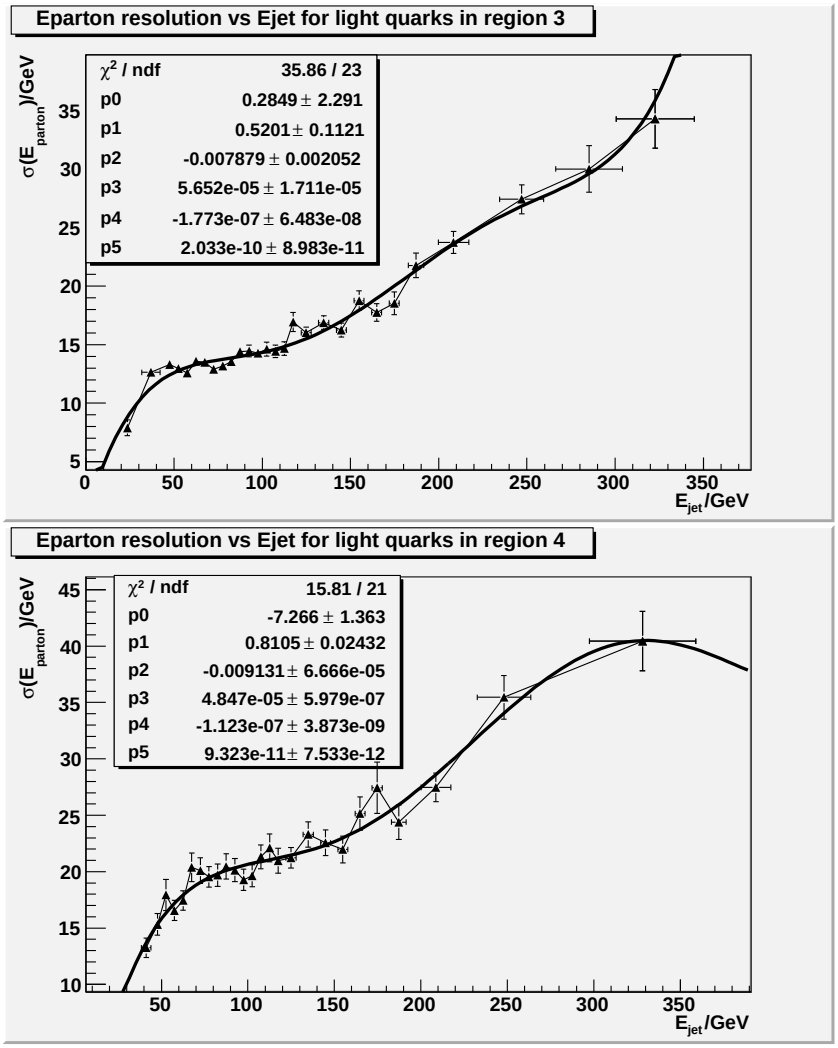
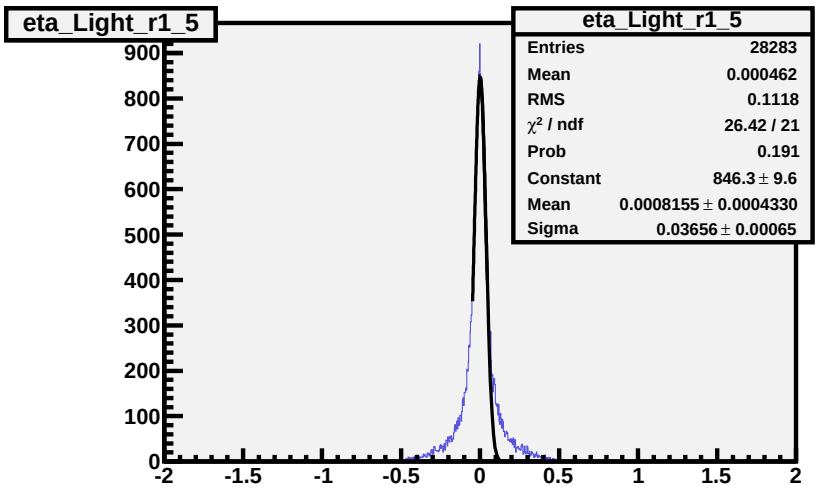
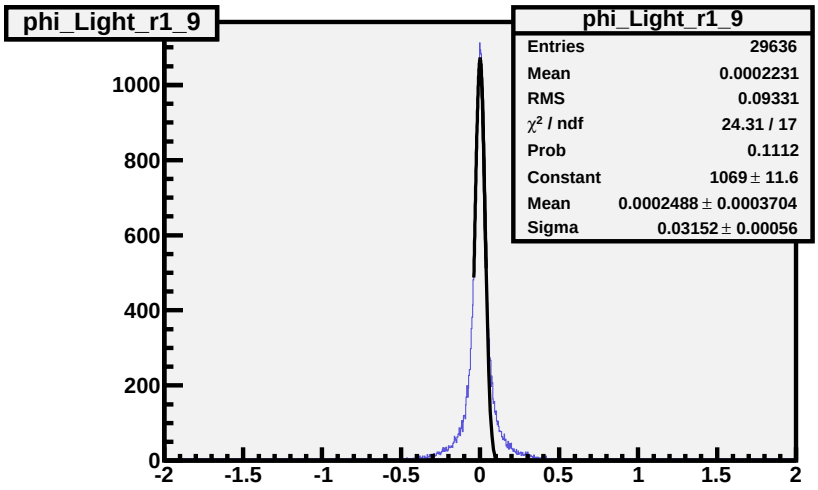


Figure 5.10: Energy resolution for light quarks in the detector η regions 3 and 4, plotted versus E_{jet}.

The vectorial quantities, pseudo-rapidity (physics η or y) and azimuthal angle (ϕ), for the parton matched jets also show some jet energy dependence. The pseudo-rapidity and azimuthal angles are not corrected to parton level. Fig.5.11 shows a typical distribution of $(\eta_{parton} - \eta_{jet})$ and $(\phi_{parton} - \phi_{jet})$ for light quarks in a particular jet energy bin. This distribution is narrowly centered around zero and hence does not need additional corrections.



(a)



(b)

Figure 5.11: Distribution of the difference of angular variables at the parton and jet levels for light quarks. (a) pseudo-rapidity (b) azimuth

The RMS of the distributions in Fig.5.11 are still parametrized as a function of jet energy in the four detector regions. Table 5.4 and Fig. 5.12-5.13 show the polynomial fits and parameter values for the pseudo-rapidity and Table 5.5 and Fig. 5.14-5.15 show the

same for azimuthal angle for light quarks.

$\lvert\eta^{det}\rvert$	$p0(GeV)$	$p1(GeV)$	$p2(GeV)$	$p3(GeV)$	$p4(GeV)$	$p5(GeV)$	$p6(GeV)$
$[0,0.5)$	0.12	-0.00425	8.25×10^{-5}	-8.65×10^{-7}	4.96×10^{-9}	-1.46×10^{-11}	1.73×10^{-14}
$[0.5,1.0)$	0.11	-0.00358	6.11×10^{-5}	-5.55×10^{-7}	2.71×10^{-9}	-6.66×10^{-12}	6.41×10^{-15}
$[1.0,1.5)$	0.066	0.000731	-3.89×10^{-5}	5.13×10^{-7}	-3.07×10^{-9}	8.68×10^{-12}	-9.43×10^{-15}
$[1.5,2.5)$	0.143	-0.00216	1.87×10^{-5}	-9.09×10^{-8}	2.47×10^{-10}	-3.47×10^{-13}	1.96×10^{-16}

Table 5.4: Parameters of polynomial fit to $\sigma(\eta_{parton}-\eta_{jet})\,vs.\,E_{jet}$ for light quarks.

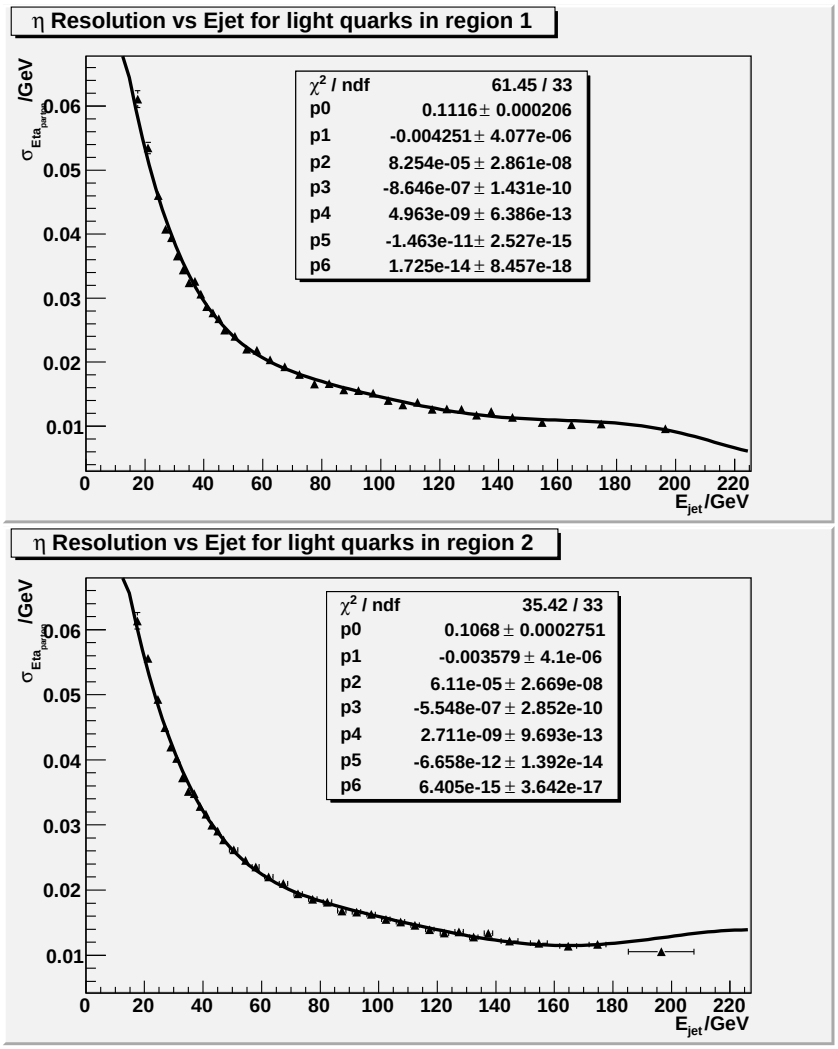


Figure 5.12: Pseudo-rapidity (η) resolution vs. E_{jet} for light quarks in detector $|\eta|$ regions 1 and 2.

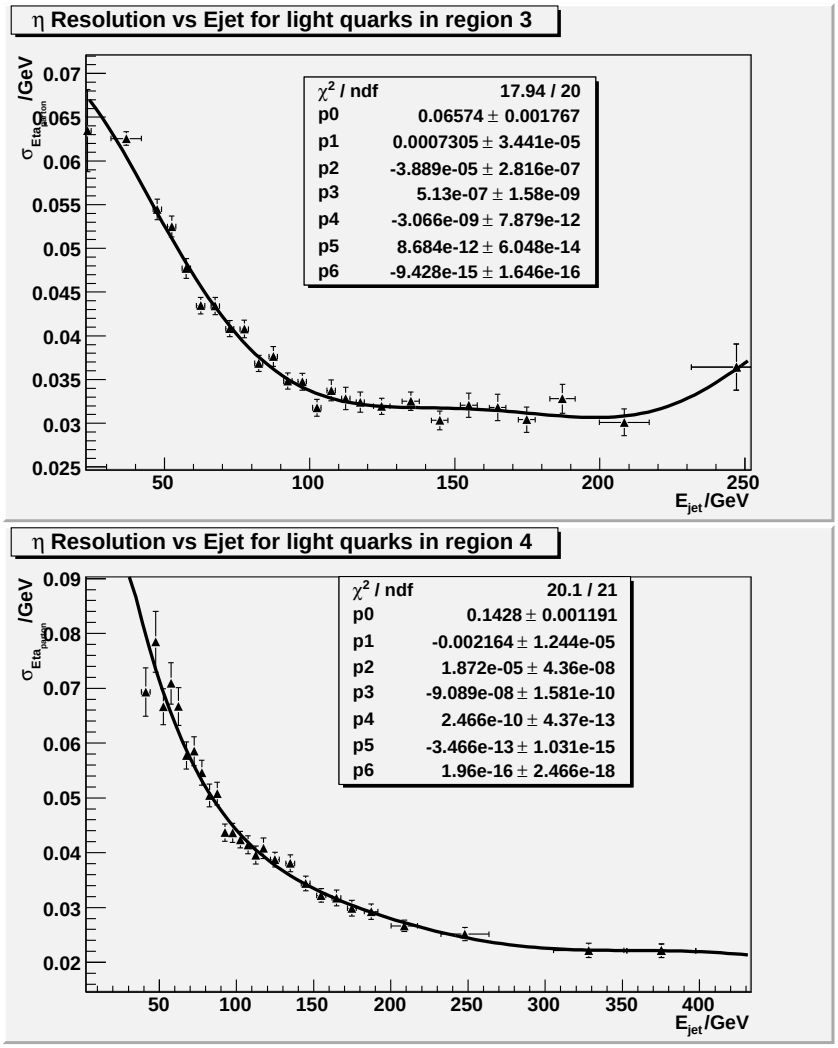


Figure 5.13: Pseudo-rapidity (η) resolution vs. E_{jet} for light quarks in four detector $|\eta|$ regions 3 and 4.

$\lvert\eta^{det}\rvert$	$p0(GeV)$	$p1(GeV)$	$p2(GeV)$	$p3(GeV)$	$p4(GeV)$	$p5(GeV)$	$p6(GeV)$
$[0,0.5)$	0.12	-0.0049	1.01×10^{-4}	-1.21×10^{-6}	6.76×10^{-9}	-2.07×10^{-11}	2.52×10^{-14}
$[0.5,1.0)$	0.11	-0.0034	5.85×10^{-5}	-5.47×10^{-7}	2.84×10^{-9}	-7.62×10^{-12}	8.23×10^{-15}
$[1.0,1.5)$	0.12	-0.0019	1.86×10^{-5}	-9.87×10^{-8}	3.10×10^{-10}	-5.53×10^{-13}	4.36×10^{-16}
$[1.5,2.5)$	0.09	-0.00037	-7.01×10^{-6}	9.55×10^{-8}	-4.80×10^{-10}	1.11×10^{-12}	-9.69×10^{-16}

Table 5.5: Parameters of polynomial fit to $\sigma(\phi_{parton}-\phi_{jet})\,vs.\,E_{jet}$ for light quarks.

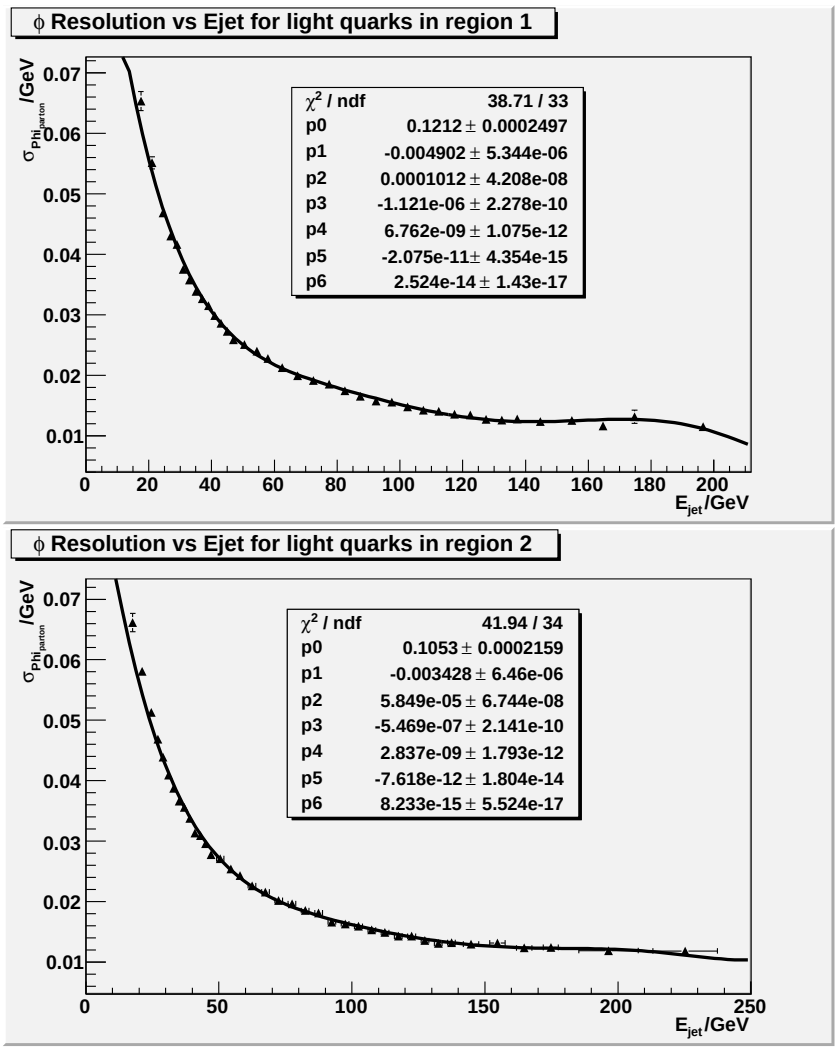


Figure 5.14: Azimuthal (ϕ) resolution vs. E_{jet} for light quarks in detector $|\eta|$ regions 1 and 2.

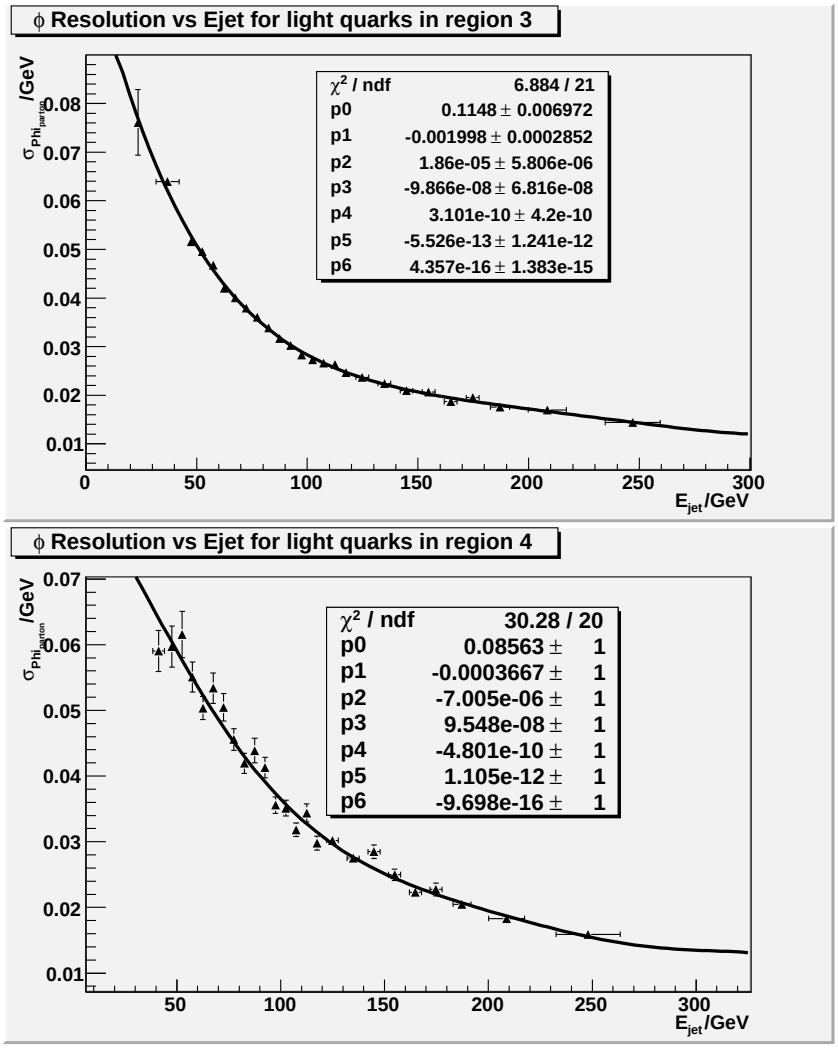


Figure 5.15: Azimuthal (ϕ) resolution vs. E_{jet} for light quarks in detector $|\eta|$ regions 3 and 4.

5.3.2 b-quark resolution

Of the four jets that enter the kinematic fitter, two are hypothesized to be b-jets for each jet-parton assignmen. As in the case of light quarks, the energy and angular resolutions are parametrized as a function of b-jet energy. Again, the parton energy resolution is derived from the parton energy distributions in various b-jet energy bins (shown in Appendix B). Table. 5.6 and Fig. 5.16-5.17 show the b-parton energy resolution parametrization as a

function of to b-jet energy in four detector regions.

$\lvert\eta^{det}\rvert$	$p0(GeV)$	$p1(GeV)$	$p2(GeV)$	$p3(GeV)$	$p4(GeV)$	$p5(GeV)$	$p6(GeV)$
$[0,0.5)$	12.5	-0.39	0.0106	1.23×10^{-4}	7.07×10^{-7}	-1.93×10^{-9}	2.01×10^{-12}
$[0.5,1.0)$	11.54	-0.18	0.0043	-3.92×10^{-5}	1.62×10^{-7}	-2.35×10^{-10}	0.0
$[1.0,1.5)$	-2.88	0.62	-0.0081	5.20×10^{-5}	-1.52×10^{-7}	1.67×10^{-10}	0.0
$[1.5,2.5)$	-14.79	1.19	-0.0153	9.23×10^{-5}	-2.56×10^{-7}	2.67×10^{-10}	0.0

Table 5.6: Parameters of polynomial fit to $\sigma(E_{parton})\,vs.\,E_{jet}$ for b-quarks.

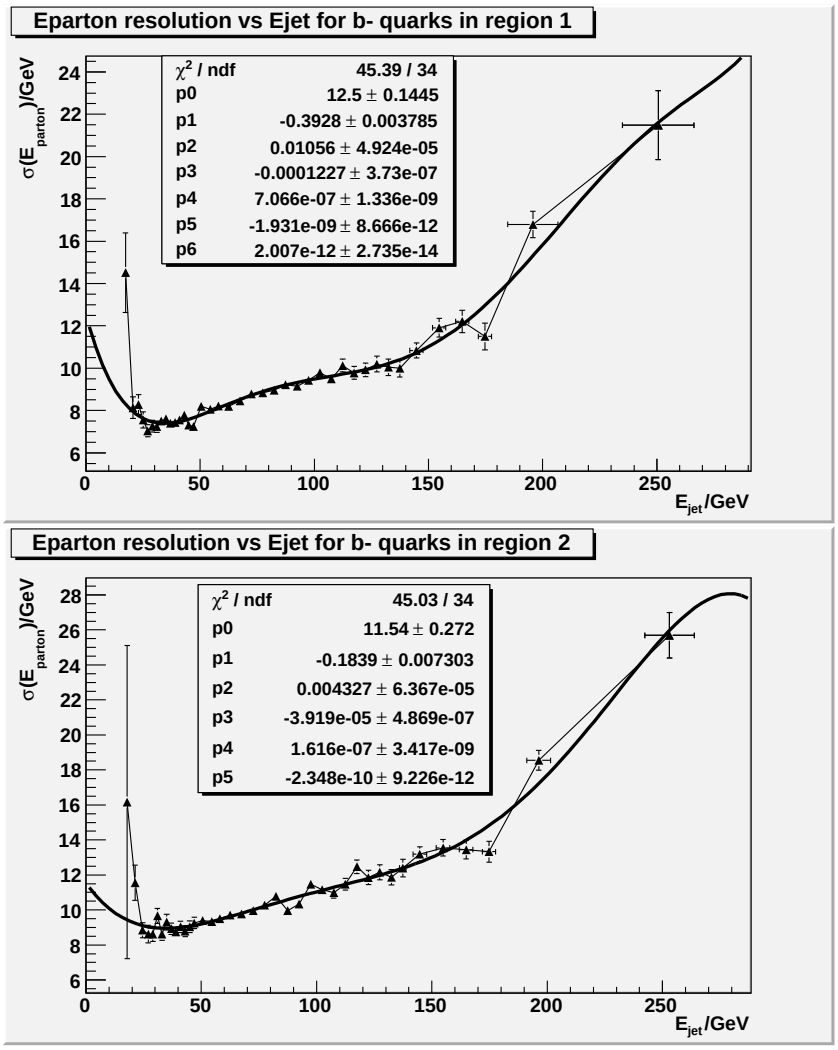


Figure 5.16: Energy resolution vs. E_{jet} for b-quarks in the detector η regions 1 and 2.

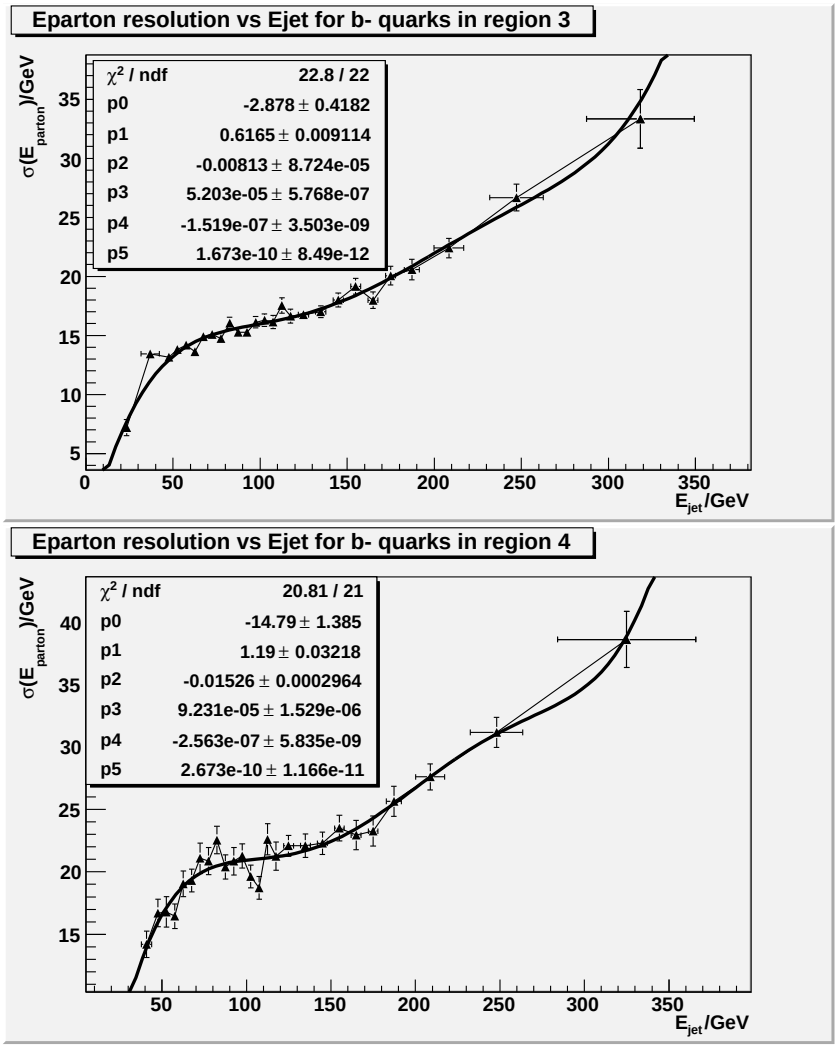
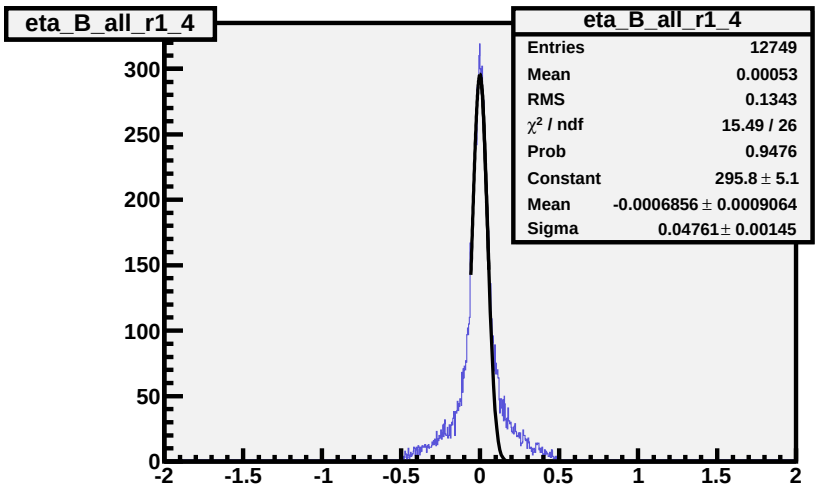
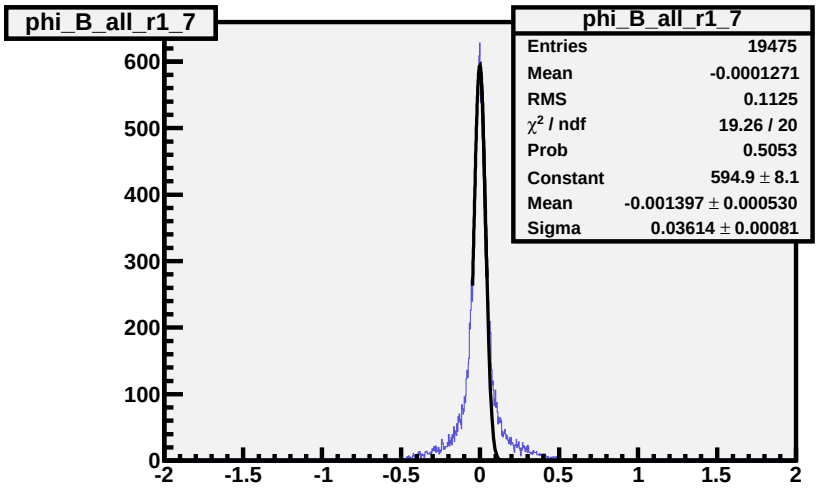


Figure 5.17: Energy resolution vs. E_{jet} for b-quarks in the detector η regions 3 and 4.

As in the case of light quarks, the angular variables for b-quarks are not corrected. Fig. 5.18, shows that the distribution of jet angular variables are close to the parton values and hence the RMS of the distribution is parametrized with respect to b-jet energy in different detector regions. This parametrization of the pseudo-rapidity of b-quarks is depicted in Table 5.7 and Fig. 5.19-5.20 while Table 5.8 and Fig. 5.21-5.22 shows the same for the azimuthal variable.



(a) $(\eta_{parton} - \eta_{jet})$ for matched b jet-parton pair.



(b) $(\phi_{parton} - \phi_{jet})$ for matched b jet-parton pair.

Figure 5.18: Distribution of the difference in angular variables for b-quark at parton and jet levels.

$\lvert\eta^{det}\rvert$	$p0(GeV)$	$p1(GeV)$	$p2(GeV)$	$p3(GeV)$	$p4(GeV)$	$p5(GeV)$	$p6(GeV)$	$p7(GeV)$
$[0,0.5)$	0.12	-0.0039	6.49×10^{-5}	-5.74×10^{-7}	2.77×10^{-9}	-6.84×10^{-12}	6.73×10^{-15}	0.0
$[0.5,1.0)$	0.12	-0.0038	5.76×10^{-5}	-4.78×10^{-7}	2.26×10^{-9}	-5.96×10^{-12}	7.94×10^{-15}	-3.99×10^{-18}
$[1.0,1.5)$	0.14	-0.0035	5.21×10^{-5}	-4.29×10^{-7}	1.99×10^{-9}	-4.79×10^{-12}	4.56×10^{-15}	0.0
$[1.5,2.5)$	0.17	-0.0028	2.44×10^{-5}	1.06×10^{-7}	2.14×10^{-10}	-1.14×10^{-13}	-1.07×10^{-16}	0.0

Table 5.7: Parameters of polynomial fit to $\sigma(\phi_{parton}-\phi_{jet})\,vs.\,E_{jet}$ for light quarks.

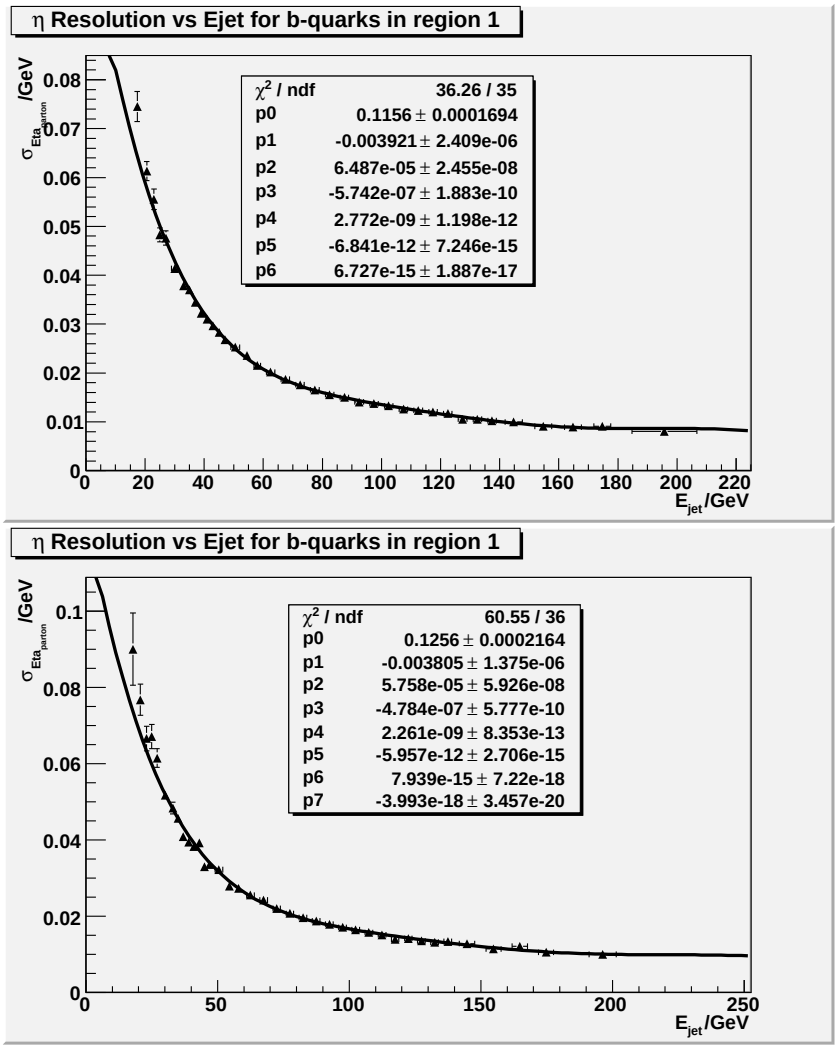


Figure 5.19: Pseudo-rapidity (η) resolution for b-quarks in detector $|\eta|$ regions 1 and 2, plotted versus E_{jet}.

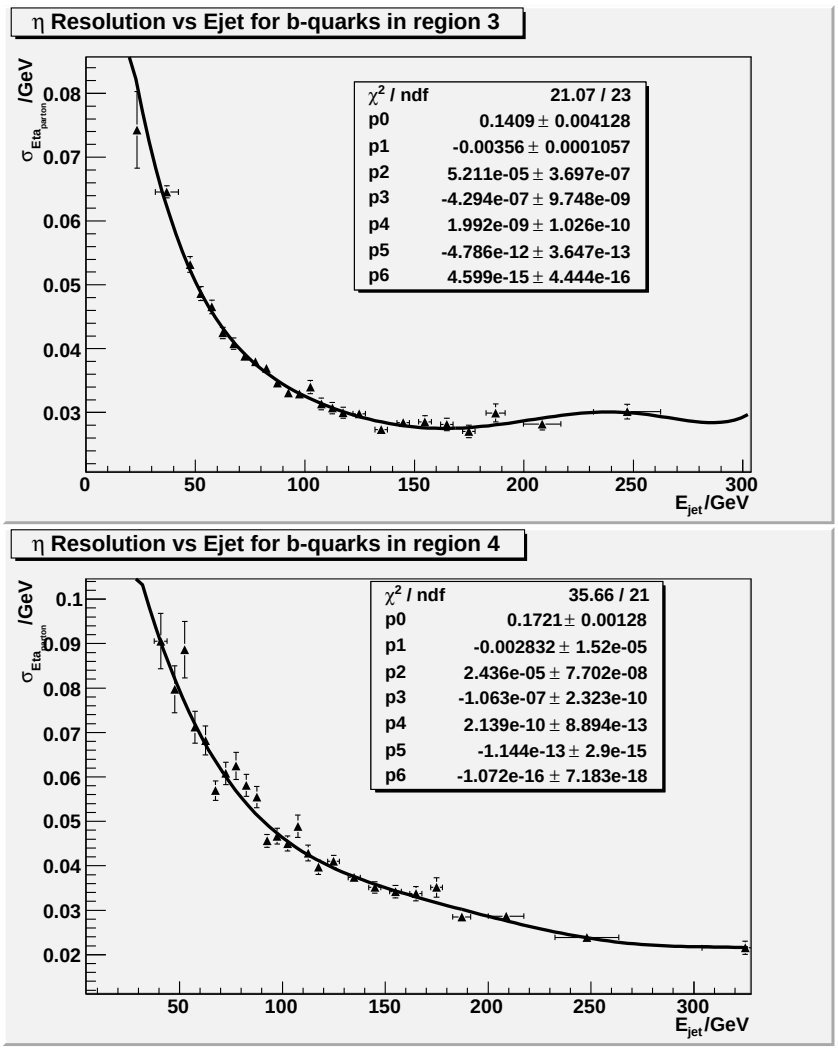


Figure 5.20: Pseudo-rapidity (η) resolution for b-quarks in detector $|\eta|$ regions 3 and 4, plotted versus E_{jet}.

$\lvert\eta^{det}\rvert$	$p0(GeV)$	$p1(GeV)$	$p2(GeV)$	$p3(GeV)$	$p4(GeV)$	$p5(GeV)$	$p6(GeV)$	$p7(GeV)$
$[0,0.5)$	0.11	-0.0033	4.27×10^{-5}	-1.97×10^{-7}	-6.43×10^{-10}	1.01×10^{-11}	-3.65×10^{-14}	4.43×10^{-17}
$[0.5,1.0)$	0.13	-0.0039	5.14×10^{-5}	-2.96×10^{-7}	1.27×01^{-10}	6.46×10^{-12}	-2.77×10^{-14}	3.64×10^{-17}
$[1.0,1.5)$	0.15	-0.0031	3.01×10^{-5}	-1.12×10^{-7}	-2.26×10^{-10}	3.14×10^{-12}	-9.06×10^{-15}	8.73×10^{-18}
$[1.5,2.5)$	0.11	-0.0012	5.91×10^{-6}	-1.26×10^{-8}	-2.17×10^{-11}	3.73×10^{-13}	-1.52×10^{-15}	2.05×10^{-18}

Table 5.8: Parameters of polynomial fit to $\sigma(\phi_{parton}-\phi_{jet})\,vs.\,E_{jet}$ for light quarks.

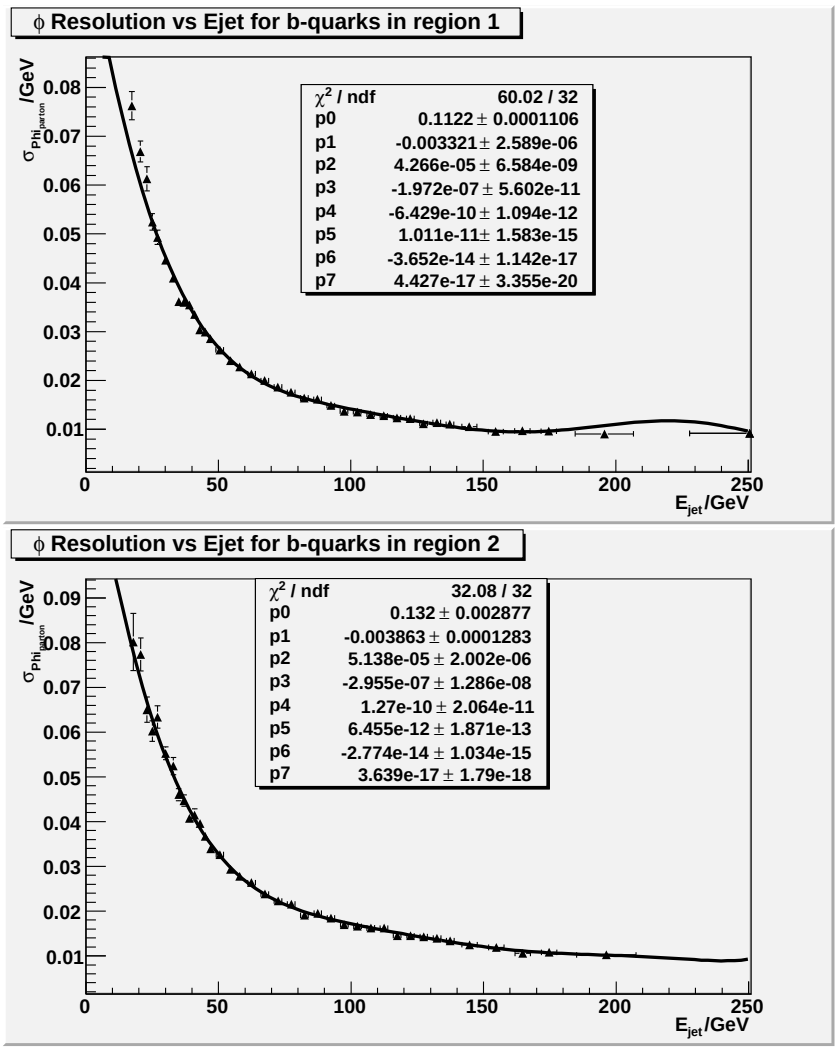


Figure 5.21: Azimuthal (ϕ) resolution for b-quarks in detector $|\eta|$ regions 1 and 2, plotted versus E_{jet}.

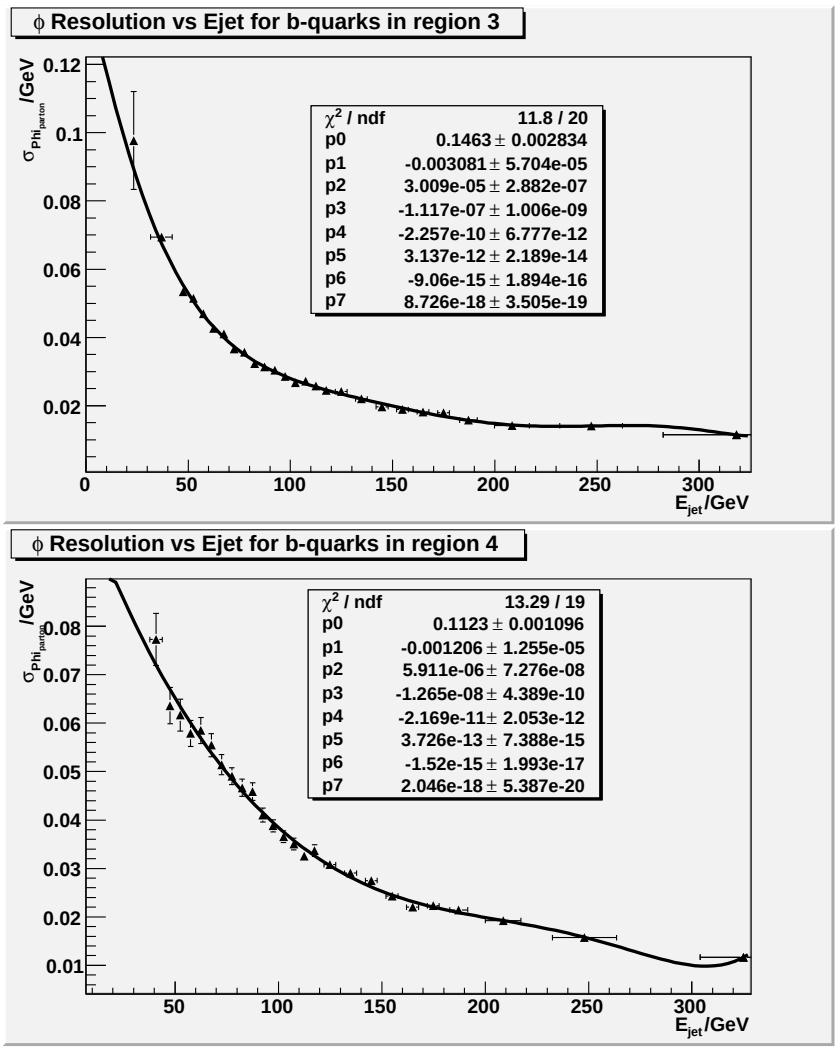
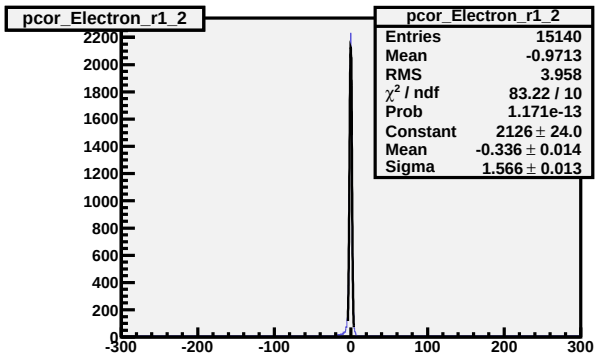


Figure 5.22: Azimuthal (ϕ) resolution vs. E$_{jet}$ for b-quarks in detector $|\eta|$ regions 3 and 4, plotted versus E_{jet}.

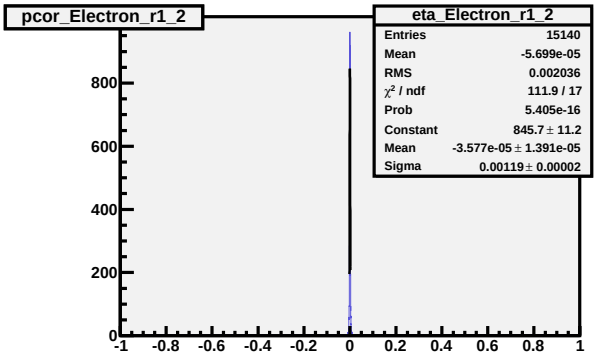
5.3.3 Lepton Resolutions

The energy (E), pseudo-rapidity (η) and azimuthal variables (ϕ) of charged lepton are compared between detector and parton level . This is done using the same set of events that were used to derive the parton level corrections and resolutions for light and b-quarks in the preceding sections. Fig. 5.23, shows the distributions in the differences of these attributes for electrons. Similar distributions exist for muons. Since the differences are small, there

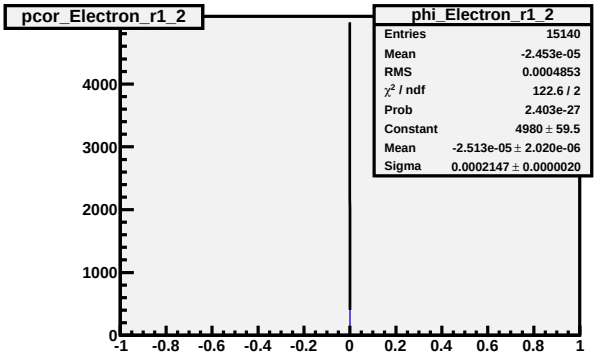
are no further corrections applied to the 4-vector of the charged leptons observed in the detector. The RMS of these distributions, are parametrized as a function of the lepton energy to yield the resolutions on the lepton E, η *and* ϕ. These are the inputs to the kinematic fitter, described in the next section.



(a) $(E_{parton} - E_{electron})$ distribution in an electron energy bin.



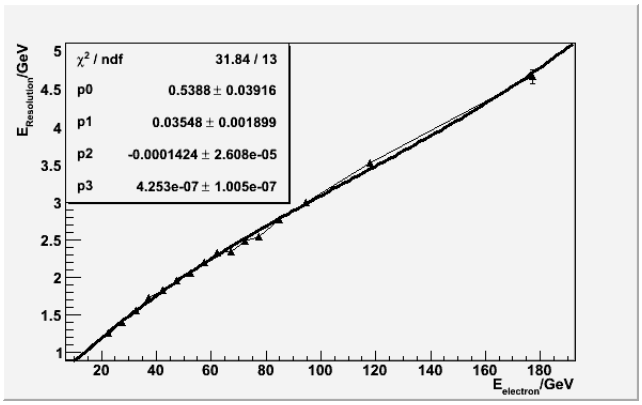
(b) $(\eta_{parton} - \eta_{electron})$ distribution in an electron energy bin.



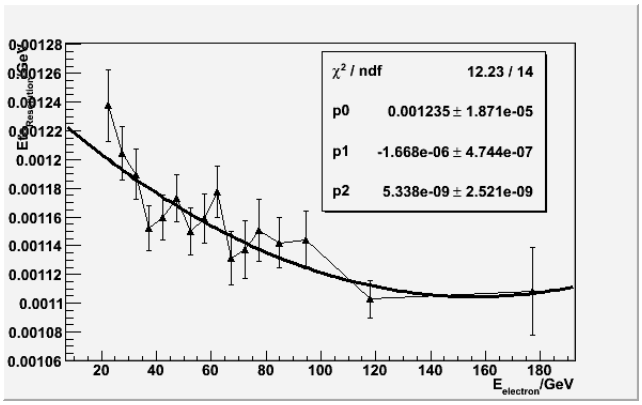
(c) $(\phi_{parton} - \phi_{electron})$ distribution in an electron energy bin

Figure 5.23: Distribution of differences in electron energy, pseudo-rapidity and azimuthal variables between detector and parton levels.

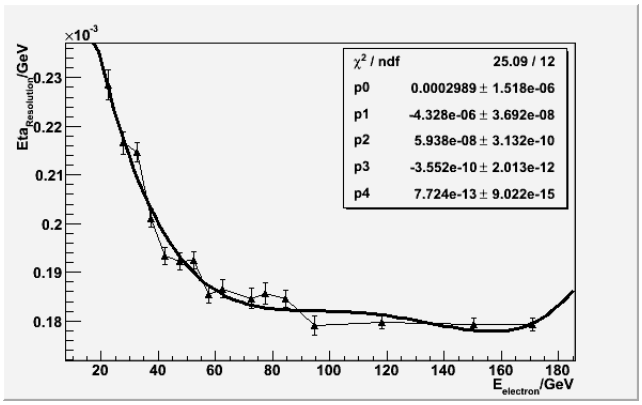
The electron resolution functions are parametrized in two detector η regions. Region 1 covers $|\eta^{det}| \in [0, 1.0)$ and Region 2 covers $|\eta^{det}| > 1.0$. Fig. 5.24-5.25, shows the polynomial fits to these parametrization for electrons.



(a) electron energy resolution vs. electron energy for detector $|\eta^{det}| \in [0, 1.0)$

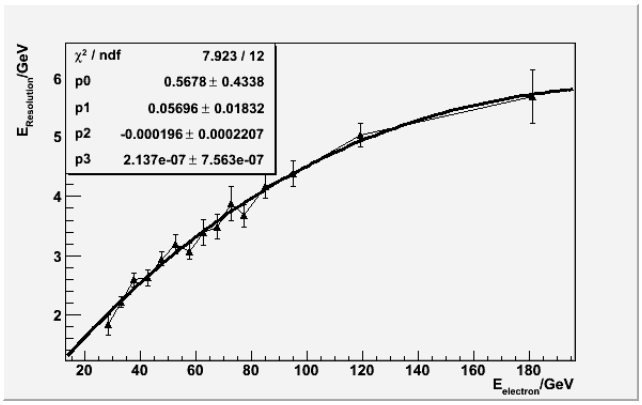


(b) electron pseudo-rapidity resolution vs. electron energy for detector $|\eta^{det}| \in [0, 1.0)$

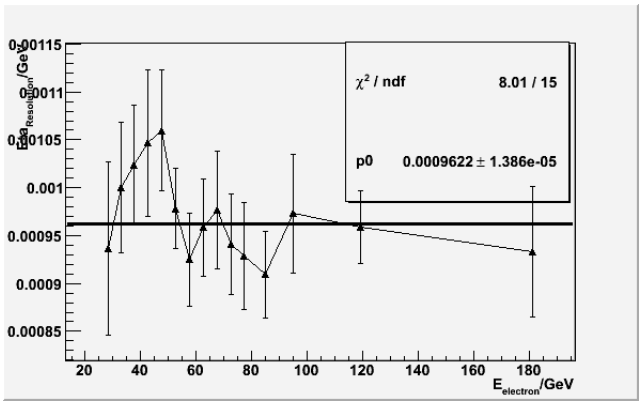


(c) electron azimuthal resolution vs. electron energy for detector $|\eta^{det}| \in [0, 1.0)$

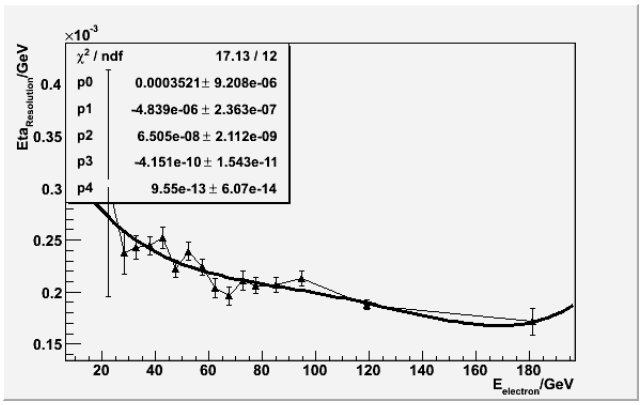
Figure 5.24: Resolutions of electron energy, pseudo-rapidity and azimuthal variables with respect to electron energy observed in the detector region 1 for $t\bar{t} \rightarrow e + jets$ events.



(a) electron energy resolution vs. electron energy for detector $|\eta^{det}| > 1.0$



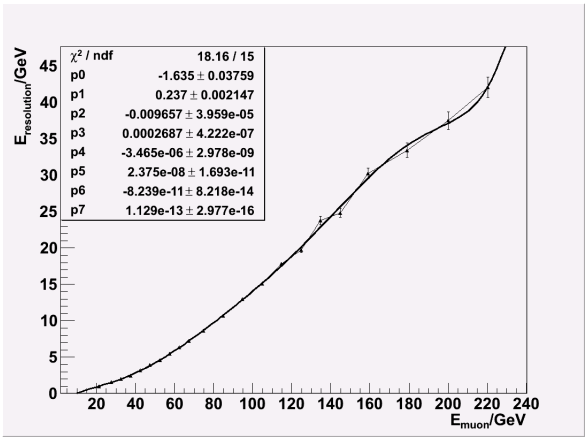
(b) electron pseudo-rapidity resolution vs. electron energy for detector $|\eta^{det}| > 1.0$



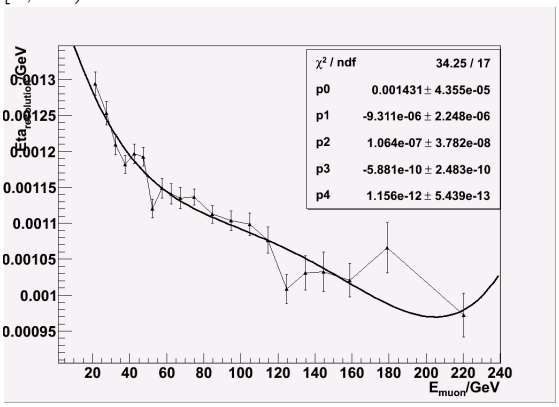
(c) electron azimuthal resolution vs. electron energy for detector $|\eta^{det}| > 1.0$

Figure 5.25: Resolutions of electron energy, pseudo-rapidity and azimuthal variables with respect to electron energy observed in the detector region 1 for $t\bar{t} \rightarrow e + jets$ events.

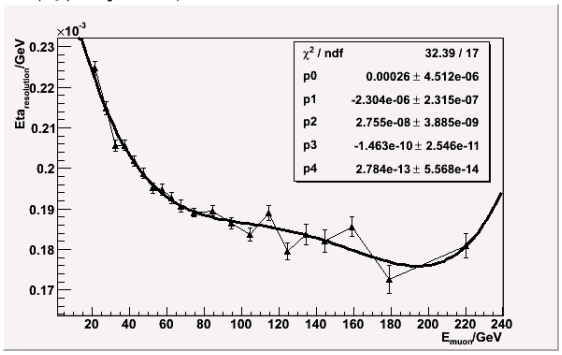
The muon resolution functions are also parametrized in two detector η regions. Region 1 covers $|\eta^{det}| \in [0, 1.6)$, and Region 2 covers $|\eta^{det}| > 1.6$. Fig. 5.26-5.27, shows the polynomial fits to these parametrization for muons.



(a) muon energy resolution vs. muon energy for detector $|\eta| \in [0, 1.0)$

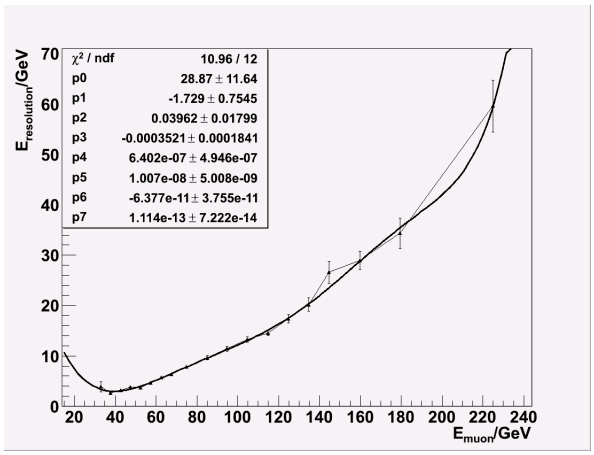


(b) muon pseudo-rapidity resolution vs. muon energy for detector $|\eta| \in [0, 1.0)$

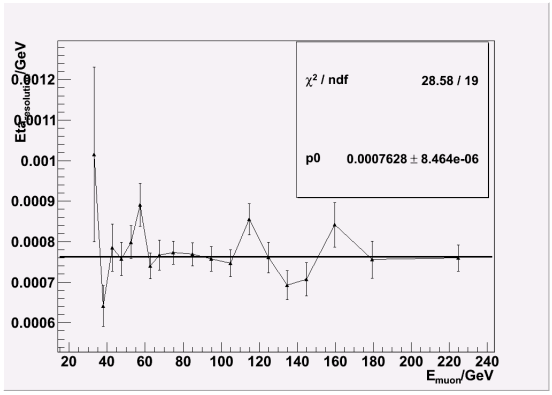


(c) muon azimuthal resolution vs. muon energy for detector $|\eta| \in [0, 1.0)$

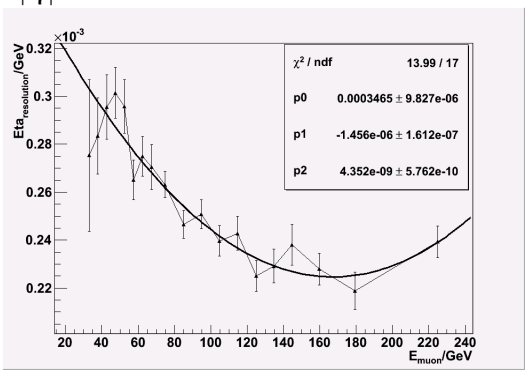
Figure 5.26: Resolutions of muon energy, pseudo-rapidity and azimuthal variables with respect to muon energy observed in the detector region 1 for $t\bar{t} \to \mu + jets$ events.



(a) muon energy resolution vs. muon energy for detector $|\eta| > 1.0$



(b) muon pseudo-rapidity resolution vs. muon energy for detector $|\eta| > 1.0$



(c) muon azimuthal resolution vs. muon energy for detector $|\eta| > 1.0$

Figure 5.27: Resolutions of muon energy, pseudo-rapidity and azimuthal variables with respect to muon energy observed in the detector region 1 for $t\bar{t} \rightarrow \mu + jets$ events.

5.4 HitFit: The kinematic fitter

The fitter takes as an input, the kinematic quantities of jets and leptons in the final state and varies them within their resolutions to obtain the best possible top quark mass (also referred to as the fitted top mass), under the following constraints :

1. $m_t^{had} = m_t^{lep}$
2. $m_W^{had} = m_W^{lep} = 80.4\,GeV$

The first condition states that the mass of the top and anti-top pair produced in the collision, are equal. This is a simple consequence of the CPT invariance, which is the basis of quantum field theory. Since, the semi-leptonic decay mode of top quarks is under consideration, both the top and anti-top decay to a $W^{\pm}$ boson, one of which decays hadronically (m_W^{had}), while the other decays leptonically to an electron or a muon and a neutrino (m_W^{lep}). The W boson mass is constrained to the known value. m_t^{had} refers to the invariant mass of the system of hadronically decaying W boson and a b-quark (anti b-quark), while m_t^{lep} refers to the invariant mass of the leptonically decaying W boson and b-quark (anti b-quark).

The kinematic fit, uses the measure of goodness of fit, that requires the measured observables (x_O), the fitted values (x_f) and the errors on the measured values ($\sigma(x_O)$), to minimize :

$$\chi^2 = \sum \frac{(x_f - x_O)^2}{\sigma^2(x_O)} \tag{5.2}$$

wherein, the summation runs over all observables. The list of the observables is as follows :

- The pseudo-rapidity (η), the azimuthal angle (ϕ) and the energy of the 4 leading jets.

- The pseudo-rapidity (η), the azimuthal angle (ϕ) and the energy of the charged lepton ($e^{\pm}$or $\mu^{\pm}$)

The constraints (1 and 2) are non-linear with respect to the above observables and hence the χ^2 is minimized iteratively, till the fitted values comply with the constraints within an infinitesimal tolerance. In this form of the fitting algorithm, the observables (measured quantities) are input as one-dimensional vectors ($\overrightarrow{x^m}$) while the fitted quantities are represented by $\overrightarrow{x^f}$. The χ^2 then takes the following form :

$$\chi^2 = (\overrightarrow{x^f} - \overrightarrow{x^m})^T \, G \, (\overrightarrow{x^f} - \overrightarrow{x^m}) \tag{5.3}$$

where, **G** is the inverse of the error matrix that contains the uncertainties of the measurements. At the initial step, the full event must be known to compute the above quantities. The missing transverse energy ($\not{E}_T$), only provides the initial estimate for the transverse momentum of the neutrino ($\overrightarrow{p_T^{\nu}}$). The longitudinal momentum of the neutrino (p_z^{ν}) is estimated using constraint 1. This yields a quadratic equation in p_z^{ν},

$$\left((p_z^{l^{\pm}b})^2 - (E^{l^{\pm}b})^2 \right) (p_z^{\nu}) + \alpha \, p_z^{l^{\pm}b} \, p_z^{\nu} - (E^{l^{\pm}b} \overrightarrow{p_T^{\nu}})^2 + \alpha^2/4 = 0 \tag{5.4}$$

where

$$\alpha = (m_t^{had})^2 - m_{l^{\pm}b} + 2 \, \overrightarrow{p_T^{\nu}} . \overrightarrow{p_T^{l^{\pm}b}} \tag{5.5}$$

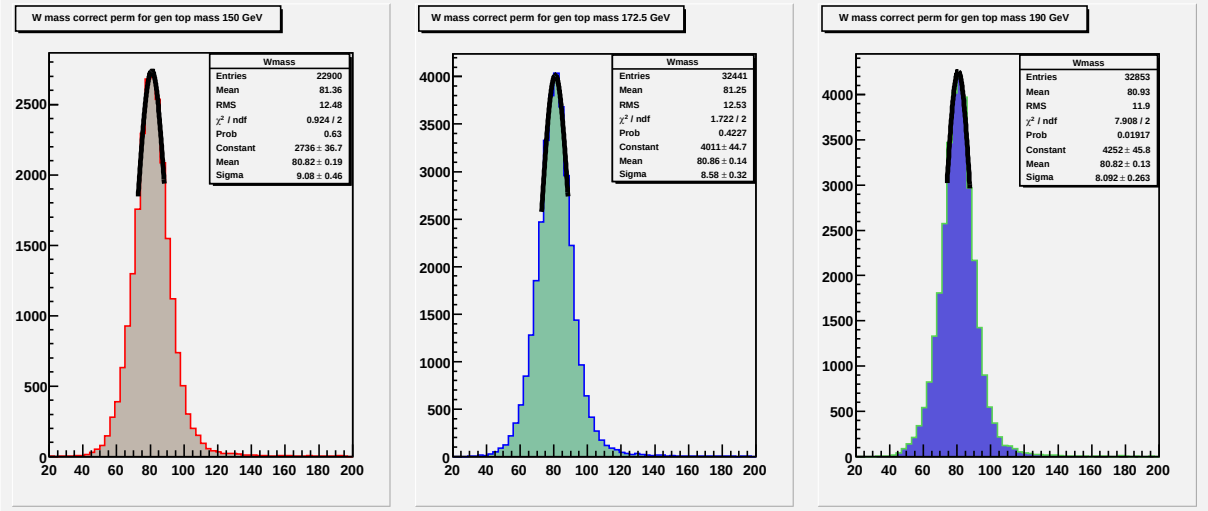
In the above equation, $l^{\pm}b$, refers to the system of the charged lepton and the b-quark involved in the leptonically decaying top. This procedure is carried out for each jet-parton assignment hypothesis for an event. As discussed before, there are 12 such possible assignments for the 4 leading jets scenario and with two solutions to Eq.5.4, there are 24 possible solutions and hence that many fitted top mass value per event. For each hypothesis, the

parton level corrections are applied to get the 4-momenta of the partons involved, depending on the flavor of the jets assumed for that hypothesis. For each step, the uncertainty and the 4-momenta of the neutrino are adjusted accordingly. The fitting procedure is based on a SQUAW algorithm. Details can be found in (Ref.[61]).

5.4.1 Performance of the fitter

The kinematic fitter, outputs a minimum χ^2 and a fitted top mass for each of the 24 solutions per event. In this chapter, we started out with semi-leptonic decay modes of $t\bar{t}$ samples simulated with ALPGEN+PYTHIA for various input masses of 150 to 190 GeV. The 4-vectors of the jets were corrected to the parton level using the matched jet-parton pairs in these samples. The credibility of these corrections depends on the closure tests. For these tests, events which have all four of the leading jets matched to their respective partons are used. This way, the contamination due to subsidiary arising from gluon radiation effects is minimal. We also know that for such events, there is at least one solution for the event in which the jet-parton assignment hypothesis was correct. There could be at most two fitted top mass values for such events owing to the possible two neutrino solutions that one gets by solving Eq. 5.4. We refer to these solutions as the "correct solutions". The closure can be classified into two categories :

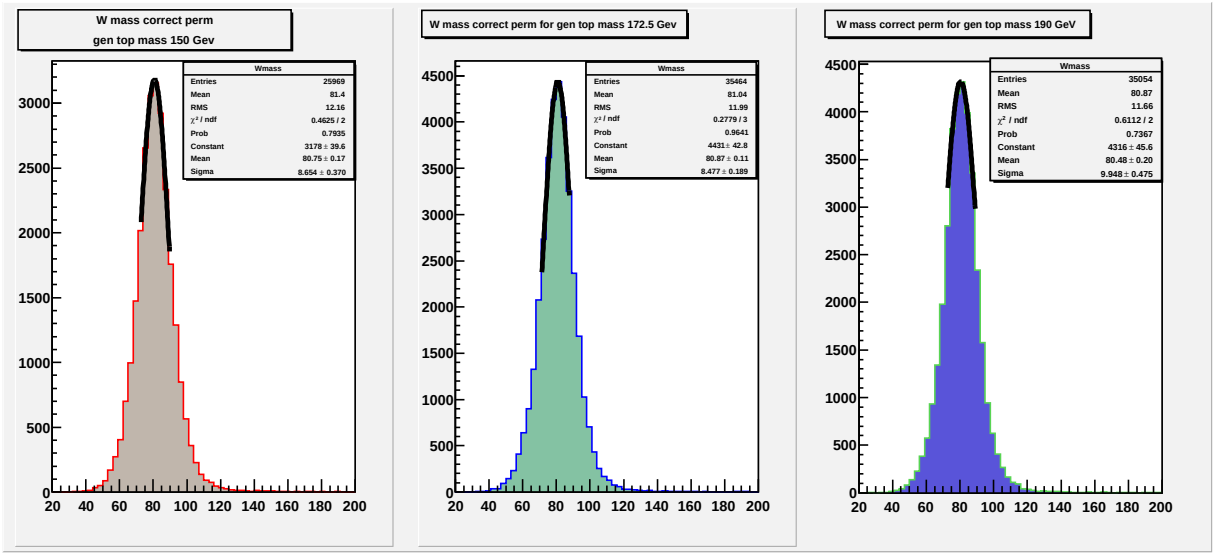
- The hadronic W mass (m_W^{had}) for the events with correct combinations is compared to the generator level value of 80.4 GeV. Fig. 5.28, shows that the peak values of such distributions are independent of the generated top mass and close to (within $\sim 0.3\%$) the true value. This provides a check for the light parton corrections and hence the correct mapping of the $W \to q\bar{q}$ process in the top quark events.



(a) m_W^{had} for correct combination in e+jets events generated with top mass = 150 GeV

(b) m_W^{had} for correct combination in e+jets events generated with top mass = 172.5 GeV

(c) m_W^{had} for correct combination in e+jets events generated with top mass = 190 GeV



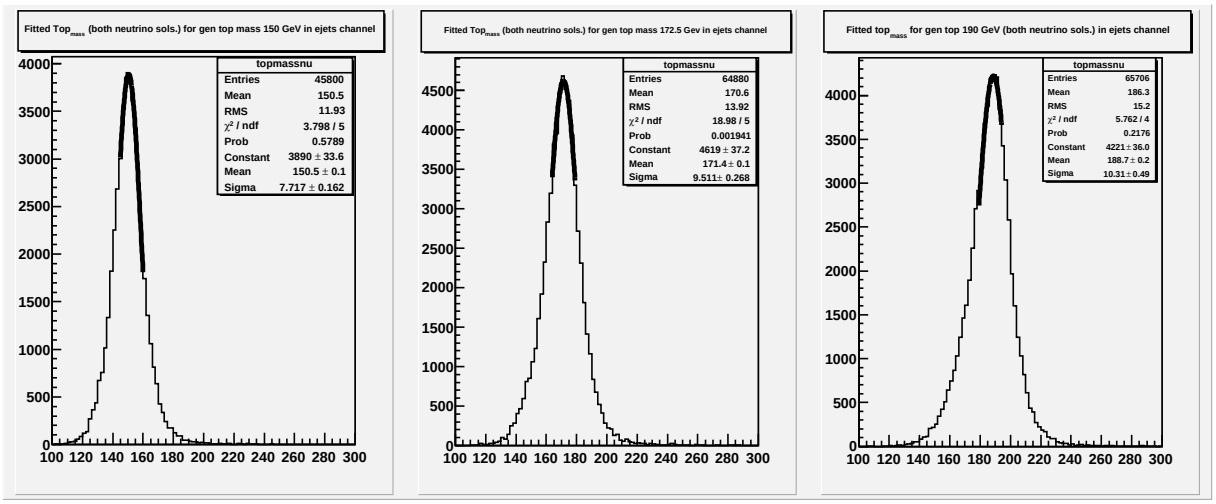
(d) m_W^{had} for correct combination in μ+jets events generated with top mass = 150 GeV

(e) m_W^{had} for correct combination in μ+jets events generated with top mass = 172.5 GeV

(f) m_W^{had} for correct combination in μ+jets events generated with top mass = 190 GeV

Figure 5.28: Parton level corrected hadronic W mass distributions for various top mass samples.

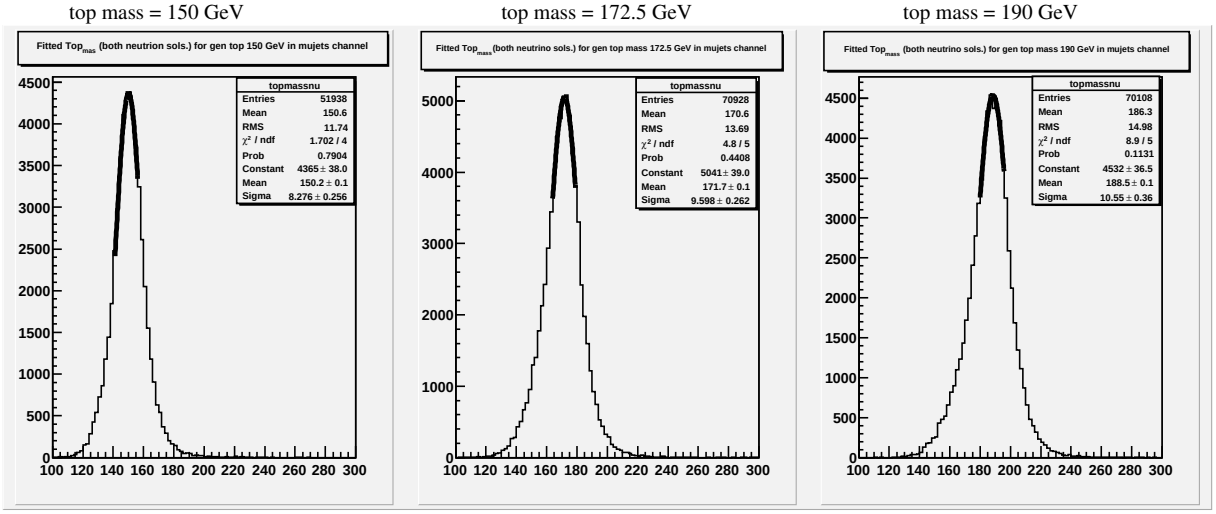
- The peak distributions of the fitted top quark mass at minimum χ^2 for the correct solutions are shown in Fig. 5.29. These distributions include both the neutrino solutions for the correct combination.



(a) fitted m_{top} for correct combination in e+jets events generated with top mass = 150 GeV

(b) fitted m_{top} for correct combination in e+jets events generated with top mass = 172.5 GeV

(c) fitted m_{top} for correct combination in e+jets events generated with top mass = 190 GeV



(d) fitted m_{top} for correct combination in μ+jets events generated with top mass = 150 GeV

(e) fitted m_{top} for correct combination in μ+jets events generated with top mass = 172.5 GeV

(f) fitted m_{top} for correct combination in μ+jets events generated with top mass = 190 GeV

Figure 5.29: Fitted top mass distributions for top mass samples generated with different input mass.

CHAPTER **Six**

Ideogram Method

6.1 Introduction

The performance of the kinematic fitter demonstrated in the previous chapter only delineates the accuracy of the resulting fit values given that the correct jet-parton assignments are known. This is not usually the case when dealing with the experimental data. It is important to extract the best possible top mass value, utilizing all of the information about an event through the maximum of 24 solutions that are available for the event. It was also discussed that a simultaneous fit to the Jet Energy Scale (JES) is performed for the events under consideration by scaling all the jets in the event by a JES factor (α_{JES}). This has been the major source of systematic uncertainty in the top quark mass measurements and the idea behind doing a simultaneous fit is to absorb a part of this uncertainty into a statistical one. Since the same events are used to extract the fitted top mass (m_t) and the JES, these measurements are correlated. Also extracted from the kinematic fit is the uncertainty on the fitted top mass value for each possible solution (σ_i). The fit is repeated for different values of the α_{JES} which is varied in steps of 1% in an interval of $\pm$15% around 1. JES_{input}=1 corresponds to the jet energies that are corrected using the γ+jet events, de-

scribed in Section 3.6 along with the sample dependent corrections described in Section 3.6.1. These energies are further corrected to parton level using the PLC described in the previous chapter (depending on the flavor of the jet). Only those jet-parton assignments for which the kinematic fit converges at all values of the α_{JES}, are taken into account. The fitted mass $m_i(\alpha_{JES})$, the estimated uncertainty on the fit $\sigma_i(\alpha_{JES})$ and the goodness of fit $\chi_i^2(\alpha_{JES})$, all depend on the JES parameter. Since the fitter uses a constraint of $m_W = 80.4$ GeV, the χ^2 is expected to be the best, when the invariant mass of the hadronically decaying W is closest to the known W mass. Additional sensitivity to the fitted JES may come from the fitted mass distribution in the background $(W + jets)$ events. To better discriminate between signal and backgrounds, $b - tagging$ is used. The analysis is calibrated for events that have 1 b-tagged jets and for events that have$\geq$ 2 b-tagged jets separately and then combined at a later stage. The events which have no b-tagged jet among the 4-leading jets have poor mass resolution and hence do not contribute to the precision. b-tagging also helps in distinguishing between correct and wrong jet-parton assignment hypotheses for an event.

In order to obtain an optimal separation between $t\bar{t}$ signal and the predominant background $(W + jets)$, without biasing the top quark mass measurement, a likelihood discriminant based on certain topological variables, is constructed. This discriminant is used to determine the "purity" of an event to weigh the signal and background probabilities of the event accordingly. These variables are listed as follows :

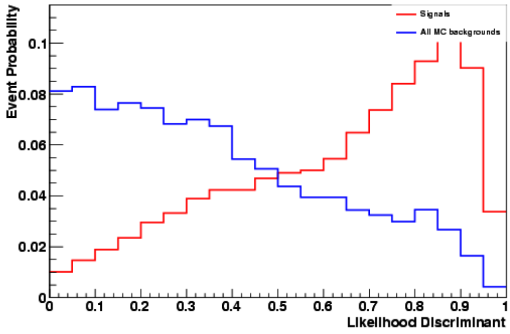
- **Aplanarity**, which is defined as $3/2$ times the smallest eigenvalue of the normalized laboratory-frame momentum tensor of the jets and the charged lepton.

- $\not\!\!E_T$**,** missing transverse energy

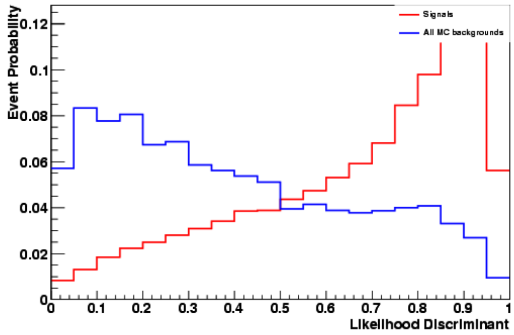
- **Centrality,** $H'_{T_2} \equiv {}^{H_{T_2}}/_{H_{\|}}$**,** where H_{T_2} is the scalar sum of the transverse momenta of the jets excluding the leading jet and $H_{\|}$ is the sum of the magnitudes of the

momentum components along the beam line of jets, isolated charged lepton, neutrino. The neutrino parallel component ($p_{\|}^{\nu}$) is computed by requiring that the invariant mass of the lepton-neutrino system is equal to the known W mass. In case of two real solutions for the above, the smallest one is taken.

- $K'_{T_{min}}$, defined as $(\triangle R_{ij}^{min} \cdot E_T^{lesser\ j})/(E_T^W)$. This quantity is the measure of the jet separation normalized to the transverse energy of the reconstructed W boson. $\triangle R_{ij}^{min}$ is the smallest distance in $\eta - \phi$ space, between any two of the 4 leading jets. $E_T^{lesser\ j}$ is the smaller of the two jet E_T. The transverse energy of the W boson is defined as $E_T^W \equiv |p_T^{lepton}| + |\not{E}_T|$.



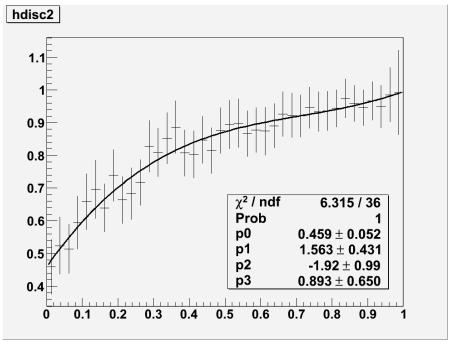
a) Topological discriminant distribution for $t\bar{t}$ and $W + jets$ backgrounds in $e + jets$ channel.



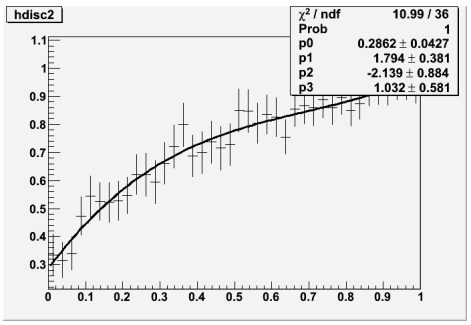
b) Topological discriminant distribution for $t\bar{t}$ and $W + jets$ backgrounds in $\mu + jets$ channel.

Figure 6.1: Comparison of the "topological discriminant" values for signal (red) and backgrounds (blue).

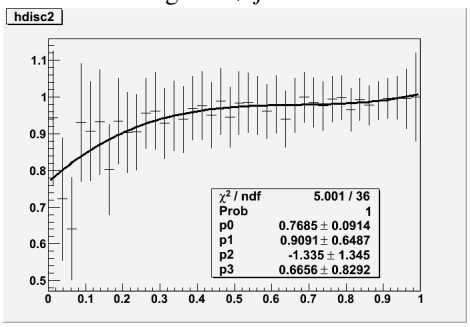
The discriminant (formed out of the topological variables) shown in Fig. 6.1 can be used to map to the signal purity of the event defined as $\left(\frac{S}{S+B}\right)$, where S is the normalized discriminant distribution of the signal while B is the same for backgrounds (mainly, $W + jets$). Fig. 6.2 shows such a mapping for different tagging bins in both e+jets and μ+jets channel.



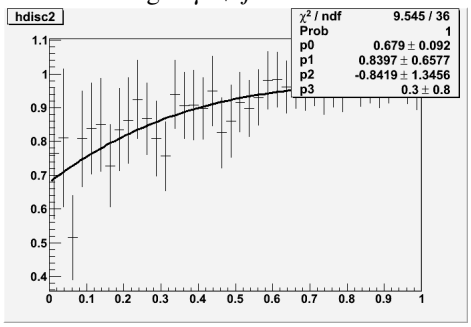
(a) Purity versus discriminant for events with 1 b-tag in $e+jets$ events.



(b) Purity versus discriminant for events with 1 b-tag in $\mu+jets$ events.



(a) Purity versus discriminant for events with $\geq$2 b-tag in $e+jets$ events.



(b) Purity versus discriminant for events with $\geq$2 b-tag in $\mu+jets$ events.

Figure 6.2: Purity versus Discriminant fits for e+jets, μ+jets events in two different tagging bins.

6.2 Likelihood Construction

Given an event sample, a most likely top mass is extracted by calculating a likelihood for each event (in the sample) as a function of assumed top quark mass (m_{top}), the jet energy scale factor (JES) and the fraction of $t\bar{t}$ events in the event sample (f_{top}). The signal fraction, f_{top} is used here as a nuisance parameter. Details of the purity fit can be found in Ref.[62, 63].The event likelihood $(\mathscr{L}_{event})$, is composed of two terms, describing the hypotheses that the event was $t\bar{t}$ signal, or an event originating from a background process :

$$\mathscr{L}_{event}(x; m_{top}, \alpha_{JES}, f_{top}) = f_{top} \times \mathscr{P}_{sgn}(x; m_{top}, \alpha_{JES}) + (1 - f_{top}) \times \mathscr{P}_{bkg}(x; \alpha_{JES}) \tag{6.1}$$

where x denotes the full set of observables that characterize the event, f_{top} is the signal fraction of the event sample and $\mathscr{P}_{sgn}$, $\mathscr{P}_{bkg}$ are the probabilities for $t\bar{t}$ and $W + jets$ production, respectively. The backgrounds here mainly comprise of W+jets, while the other backgrounds viz. QCD multijet are not modeled explicitly. The event observable x, can be thought of as a vector, containing all relevant data about the event topology and kinematics. The topological part is not correlated with the mass and is the probability of the observed discriminant value to occur given the event is signal or background like. The kinematic component (x_{fit}) contains the mass information, extracted from the constrained kinematic fit which provides the sensitivity to the top quark mass and the jet energy scale. Assuming that the two probabilities, described above, are independent of each other, one can write :

$$\mathscr{P}_{sgn}(x; m_{top}, \alpha_{JES}) \equiv \mathscr{P}_{sgn}(D) \times \mathscr{P}_{sgn}(x_{fit}; m_{top}, \alpha_{JES}) \tag{6.2}$$

and

$$\mathscr{P}_{bkg}(x; \alpha_{JES}) \equiv \mathscr{P}_{bkg}(D) \times \mathscr{P}_{bkg}(x_{fit}; \alpha_{JES}) \tag{6.3}$$

where D (discriminant) is evaluated at $JES_{input} = 1$. The normalized probability distributions of the discriminant ($\mathscr{P}_{sgn}(D)$ & $\mathscr{P}_{bkg}(D)$) are obtained from Monte Carlo simulations as shown in Fig. 6.2 and are assumed to be independent of JES. The mass information (x_{fit}), consisting of all the fitted masses $m_i(\alpha_{JES})$, estimated uncertainty $\sigma_i(\alpha_{JES})$ and the measure of goodness-of-fit $\chi^2_i(\alpha_{JES})$is obtained from the kinematic fitter as explained below.

6.2.0.1 Top Mass templates

In this section, the construction of $\mathscr{P}_{sgn}(x_{fit};; m_{top}, \alpha_{JES})$ is discussed. The output of the kinematic fitter can be used to calculate the signal probability of an event, as the sum over the 24 solutions corresponding to the jet-parton assignment and the two neutrino solutions. The relative probability for each of the solution i to be correct is estimated through a weight function w_i. This weight function can be resolved into two parts. Without b-tagging, w_i purely represents the χ_i^2 of the corresponding kinematic fit and is defined with a Gaussian probability as $w_i = exp(-\frac{1}{2}\chi_i^2)$. To weight the combinations which are more likely to be correct due to the presence of b-tagged jets, an additional relative weight of $w_{btag,i}$ is applied. This relative weight, represents the probability that observed b-tags are compatible with the assumed jet-parton assignment.

$$w_{btag,i} = \prod_{j=1,n_{jet}} p_i^j \tag{6.4}$$

wherein, p_i^j can assume a value of $\varepsilon_{l,}\varepsilon_b, (1-\varepsilon_l), (1-\varepsilon_b)$, depending on the assumed flavor of the jet (light or b) and whether or not the particular jet was tagged. The tagging rates for light and b-quark jets $\varepsilon_l \, and \, \varepsilon_b$ are determined from data as parametrized functions of jet p_T and η , where the jet p_T is based on the reconstructed jet energy for $JES_{input} = 1$. n_{jet}, here represents the 4-leading jets that are fed into the kinematic fitter. Hence, the total event permutation weight can be written as :

$$w_i = \exp(-\frac{1}{2}\chi_i^2) \prod_{j=1,n_{jet}} p_i^j \tag{6.5}$$

The mass dependent signal probability is modeled as :

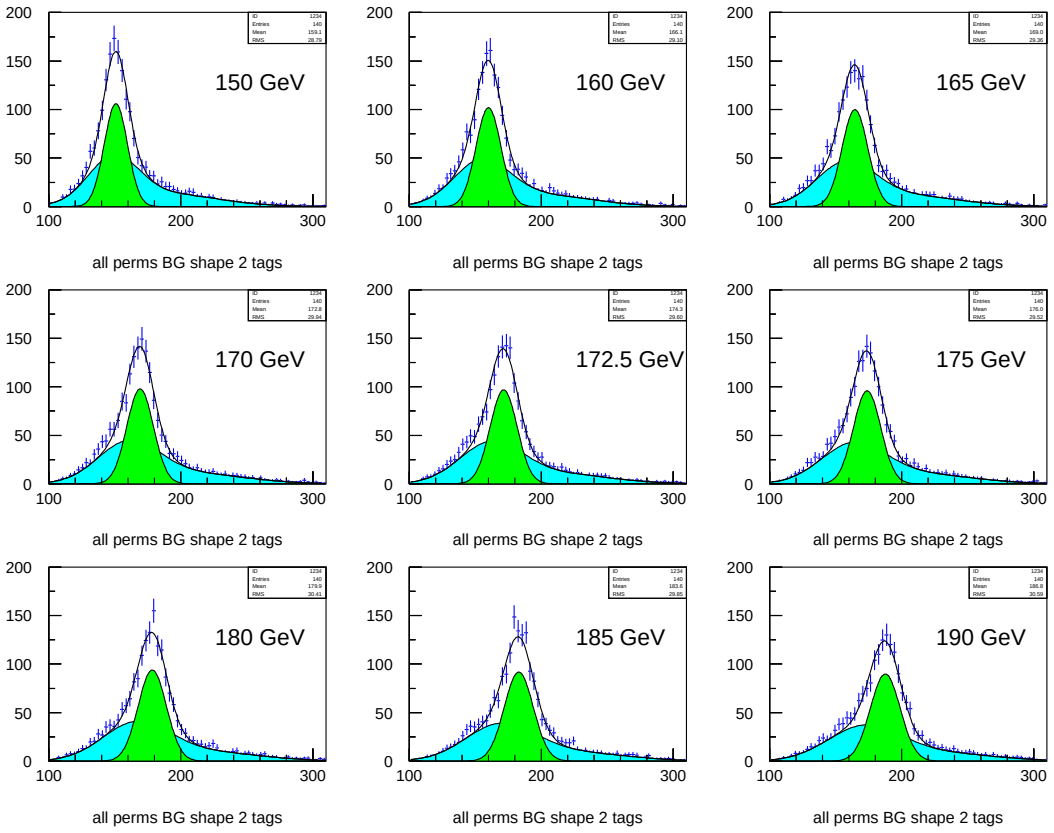
$$\mathscr{P}_{sgn}(x_{fit};m_{top},\alpha_{JES}) = \sum_{i=1}^{24} w_i \left[f_{correct}^{ntag} \cdot \int_{100}^{300} G(m_i,m^{'},\sigma_i) \cdot BW(m^{'},m_{top}) dm^{'} \right] + \sum_{i=1}^{24} w_i \left[(1-f_{correct}^{ntag}) \cdot S_{wrong}^{ntag}(m_i,m_{top}) \right] \quad (6.6)$$

In the above expression, the two terms correspond to the shapes of correct solutions and the wrong solutions weighted with a fraction that depends on the number of b-tags in the event and is estimated from the Monte Carlo simulations. These shapes are shown in Fig. 6.3-6.4, while the $f_{correct}^{ntag}$ values for e+jets and μ+jets is listed in Table. 6.2.0.1. The term compatible with the correct solutions is modeled as a Gaussian, centered at the fitted top mass value m_i and width corresponding to its uncertainty σ_i, convoluted with a Breit-Wigner line shape of width equivalent to the known top quark width of 2 GeV. The Gaussian part describes the experimental resolution while the relativistic Breit-Wigner part describes the the expected distribution of the line shape of the top and the anti-top quarks in the event, for a given top quark mass m_t.The BW is normalized to an interval of 100 to 300 GeV in order for it to not bias the top mass measurement in the region of interest.

The second term in the signal probability, represents the shape of those jet-parton assignments that are not correct combinations. These shapes for a given m_t are obtained from the Monte Carlo simulation of $t\bar{t}$ generated at different m_t. Only combinations that are known to be "wrong" are filled in histogram with each entry weighted with its coorresponding weight w_i. These shapes are then fitted with double Gaussian as shown in Fig. 6.5. The fit parameters show a nice linear dependence on m_t. This parametrization is shown in Table. 6.1.

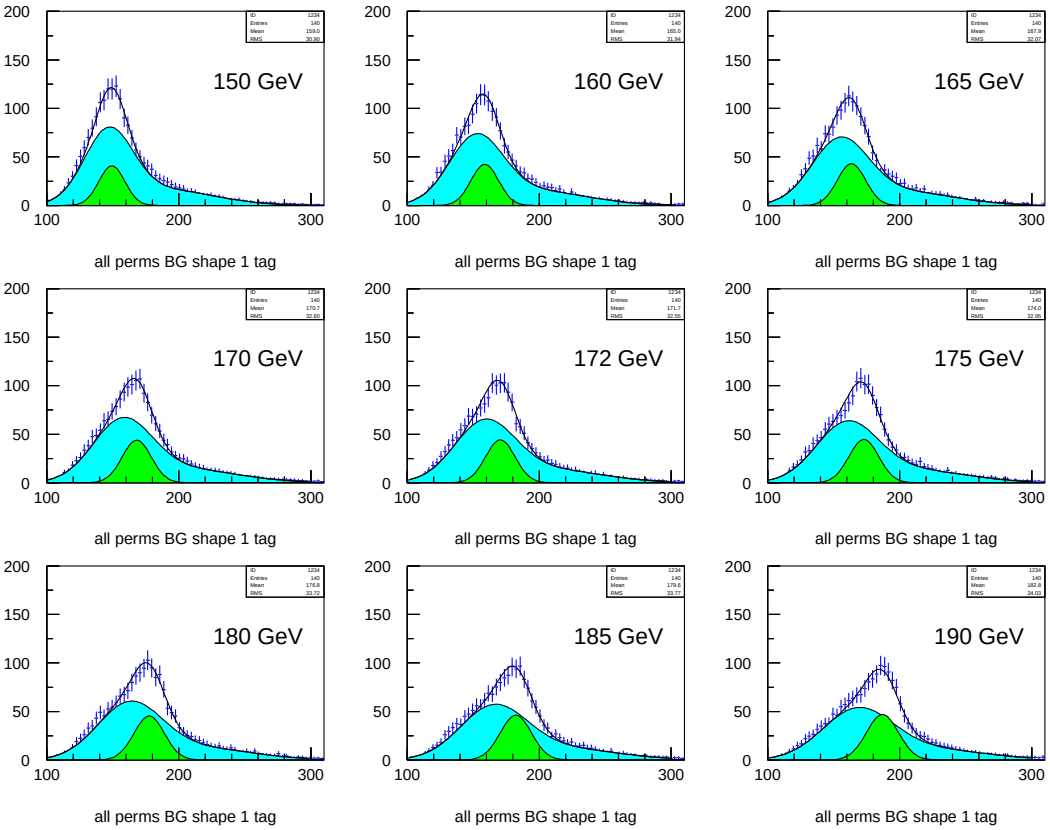
# tags	$e+jets$	$\mu+jets$
1	$f_{correct}=0.2708$	$f_{correct}=0.2702$
≥ 2	$f_{correct}=0.743$	$f_{correct}=0.7322$

Table A: $f_{correct}^{ntag}$ for $e+jets$ and $\mu+jets$ channels. The value decreases with lesser tags.



Sum (black) of weighted right combination shape (green) and wrong combination shape (blue), overlay-ed on the $t\bar{t}$ MC simulation in $e+jets$ channel for events with $\geqq$2 b-tags

Figure 6.3: Right and wrong combination shapes for different generator level top quark masses.



Sum (black) of weighted right combination shape (green) and wrong combination shape (blue), overlay-ed on the $t\bar{t}$ MC simulation in $e + jets$ channel for events with exactly 1 b-tag

Figure 6.4: Right and wrong combination shapes for different generator level top quark masses or 1 b-tagged events.

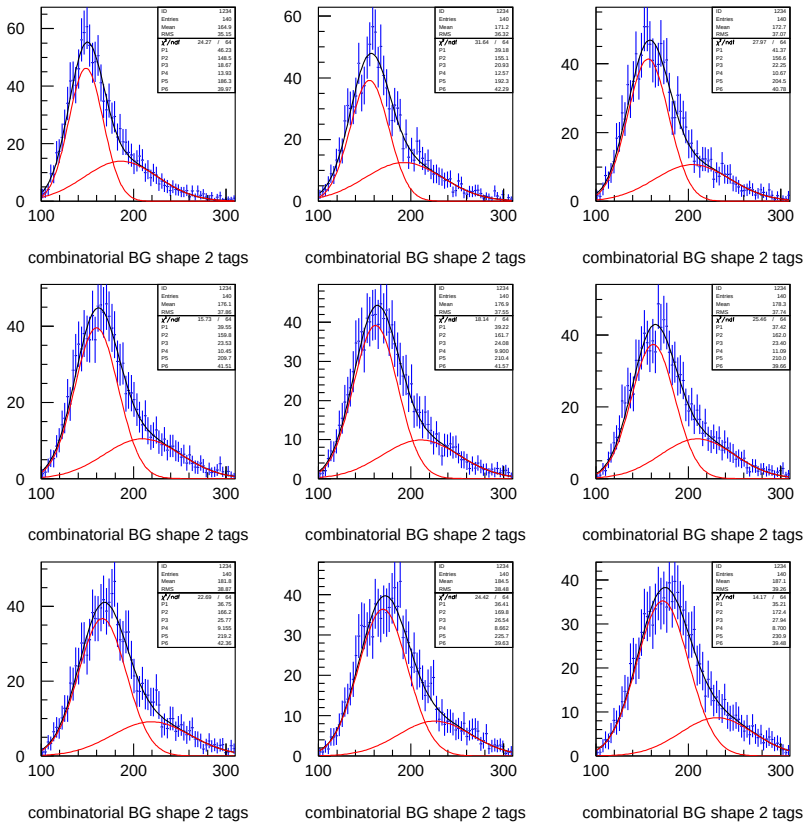


Figure 6.5: Weighted wrong combination shapes fitted with double Gaussian.

	1tag				≧2 tags			
	Gauss1		Gauss2		Gauss 1		Gauss 2	
	p0	p1	p0	p1	p0	p1	p0	p1
a	57.08	-0.5468	13.2	-0.134	38.23	-0.264	10.09	-0.13
mean	159.8	0.5241	207.9	0.8694	163.0	0.5679	212.6	1.06
σ	23.29	0.2058	41.66	0.084	24.33	0.221	40.8	0.0037

Parameters used to describe the combinatorial background shapes (arbitrary normalization). The shape is that of double Gaussian $G(m_{fit}) = a \times \exp\left(-(mean - m_{fit})^2/2\sigma^2\right)$, where the 3 parameters a, mean and σ are parametrized linearly w.r.t. the generator level top mass m_t as $p0 + p1\,.\,(m_t - 175)$ GeV for $e + jets$ channel (above) and $\mu + jets$ channel (below)

	1tag				≧2 tags			
	Gauss1		Gauss2		Gauss 1		Gauss 2	
	p0	p1	p0	p1	p0	p1	p0	p1
a	28.54	-0.229	6.19	-0.075	38.52	-0.313	9.584	-0.108
mean	159.7	0.605	208.7	1.075	163.5	0.577	214.8	1.014
σ	23.49	0.20341	42.67	0.072	24.78	0.224	42.40	0.047

Table 6.1: Parametrization of combinatorial shapes with respect to generator level top mass.

6.3 W+jets Background shapes

In this section, $\mathscr{P}_{bkg}(x_{fit;JES})$ in Eq. 6.3 is discussed. The background probability of an event is estimated from $W + jets$ Monte Carlo simulations. It is taken to be :

$$\mathscr{P}_{bkg}(x_{fit;JES}) \equiv \sum_{i=1}^{24} w_i\,.\,\mathscr{BG}(m_i) \tag{6.7}$$

where, $\mathscr{BG}$ is the background shape estimated from fitted top mass in $W + jets$ sample. Such a shape is shown in Fig. 6.6. A common shape is used for both 1 and $\geqq$ 2 b-tags. This is done to avoid unscrupulous features arising from statistical fluctuations. The shape is also smoothed by calculating the average value in a sliding window of $\pm$10 GeV around each fitted mass. These shapes are evaluated at every JES factor.

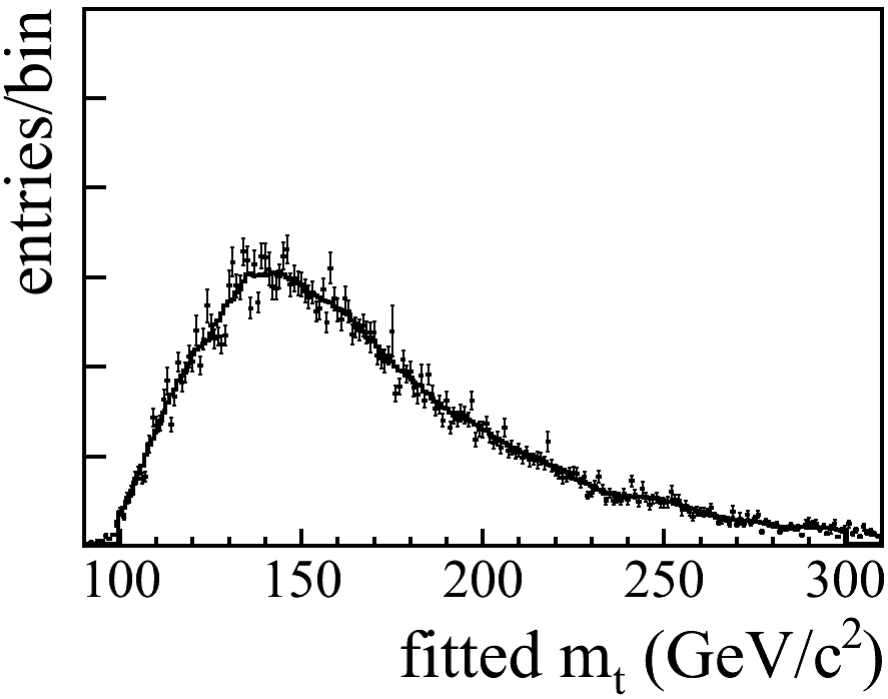


Figure 6.6: A sample $\mathscr{BG}$ shape for $\mu + jets$ events at $JES = 1$

6.4 Likelihood of a sample

Since each event is independent, the combined likelihood for the whole sample is calculated as the product of single event likelihoods. This likelihood is maximized with respect to $m_t, JES \& f_{top}$.

$$\mathscr{L}_{samp}(m_t, \alpha_{JES}, f_{top}) = \prod_j \mathscr{L}_{evtj}(m_t, \alpha_{JES}, f_{top}) \tag{6.8}$$

The likelihood ($\mathscr{L}$) is computed for all values of f_{top} between 0 and 1, in the steps of 0.05. The parameter m_t, varies between $125\,GeV/c^2$ to $225\,GeV/c^2$ in steps of $1\,GeV/c^2$.The third free parameter, α_{JES}, is varied from $+15\%$ to -15% around the nominal value of JES, in the steps of 1%. The range is chosen so that the measurement is not affected by boundary effects. First, for each value of α_{JES} and m_t, the likelihood is maximized with respect to f_{top}. This leads to a likelihood distribution in the two dimensional grid of m_t and JES (Fig.

6.13). A fit to the global minimum of this distribution yields extracted values of m_t^{fitted} and JES^{fitted} for a sample. This procedure is first applied to the simulated MC events to check the consistency of the input and output values of two of the freely floating parameters (m_t and JES). This is achieved through constructing pseudo-experiments and calibration, that are described in the next chapter.

6.5 Calibration

The likelihood expressed in Eq. 6.6 entails only an approximate representation of the true detector resolution effects and the complex physics involved in the production and reconstruction of top quark candidate events. The four leading jet hypothesis, described in the chapter 5 does not fully describe an event, for example those events in which the jets arise from radiation effects. Approximations of this form lead to a loss of sensitivity in the determination of mass and introduces a *bias* on the estimated parameter (m_t). For this reason, the analysis is first calibrated using Monte Carlo (MC) simulated events. The bias on the measured mass and the verity of the estimated statistical uncertainty can be achieved through *ensemble testing*. Each ensemble corresponds to a simulated experiment that is conducted to emulate the size and composition of the actual data observed in the detector. Hence, an ensemble is also referred to as a *pseudo-dataset*. To have maximum sensitivity and to capture the detailed structure of likelihood, the ensemble tests are performed in four different channels separately. These four channels are :

- $e + jets, \geq 2 tags$
- $\mu + jets, \geq 2 tags$
- $e + jets, \; exactly \, 1 tag$
- $\mu + jets, \; exactly \, 1 tag$

The fractions of $t\bar{t}$, $W + jets$ in each channel are allowed to fluctuate according to the estimated fractions in the actual data sample. These fractions were derived in Sec. 4.4 and have been tabulated in Table. 6.2. 3000 such ensembles are constructed in each channel with the amount of $t\bar{t}$ and $W + jets$, as shown in the table. The number of effectively independent ensembles varies from channel to channel and thus affects the uncertainty on each of the measured values. The MC statistics available in $e + jets$, $\mu + jets$ channels (without splitting into different tag bins) varies from approximately, 1 million to 50,000 events for $t\bar{t}$ and W+jets respectively. Each of these events are weighted with event weights (w_i). The effective number of independent ensembles (N_{eff}) that can be constructed from this pool of MC events (N_{samp}) , is crudely given by :

$$N_{eff} = \frac{N_{samp}}{N} \tag{6.9}$$

In the Eq. 6.9, N represents that number of events that are picked in an ensemble. If the ensemble is an admixture of signal and background events, $N = N_{t\bar{t}} + N_{W+jets}$ would be the right approximation. N_{eff} is used in assigning the uncertainty on the extracted value of the parameters, for each channel. To make optimal use of the available MC statistics, standard re-sampling techniques are used, allowing for multiple use of MC events when constructing the ensembles. If we represent w_i as the event weight of the i^{th} event in the sample, then the re-sampling involves, throwing a random number and comparing it to the $\frac{w_i}{\sum w_i}$ (sum in the denominator runs over all events in the sample under consideration). The i^{th} event is selected only if the value of this fraction is greater than the random number value. The process is carried until N number of events are selected. Since N_{samp} is quite large, compared to N (eg. first two rows in Table. 6.2), it leads to a negligible correlation due to the multiply selected events, within an individual ensemble.

Ensemble tests are performed for nine input top quark masses (150, 160, 165, 170, 172.5, 175, 180, 185, 190 GeV/c^2) . The JES factor (α_{JES}) is varied by multiplying, all the measured jet energies by a factor and feeding the events back into the kinematic fitter to recalculate the likelihood. The reference JES, which corresponds to jet energy scale corrections described in Section. 3.6 is often referred to as the *nominal* JES with $\alpha_{JES} = 1.0$. The two dimensional (2D) likelihood, thus obtained is searched for a global minimum in the values of $-2\ln(\mathscr{L})$, at which location a two dimensional function is fitted :

$$\mathscr{F}(m_t, \alpha_{JES}) = p_0 \times m_t^2 + p_1 \times \alpha_{JES}^2 + p_2 \times m_t + p_3 \times \alpha_{JES} + p_4 \times m_t \times \alpha_{JES} + p_5 \quad (6.10)$$

The Eq. 6.10, implies a basic assumption that the distribution of $-2\ln(\mathscr{L})$ values is Gaussian along the m_t and JES axes. Since the signal shapes involved in calculating the signal likelihoods have been shown to be of Gaussian nature (Fig. 6.3), it corroborates the above assumption. Any deviations would automatically introduce biases and this provides the purpose of achieving accurate calibrations. The standard deviation is the width of the Gaussian, which is represented by σ for which $-2\ln(\mathscr{L})$ increases by one unit. This is represented as :

$$-2\ln((\mathscr{L}(m_t^{fitted} + \sigma)) + 2\ln(\mathscr{L}(m_t^{fitted})) = 1 \quad (6.11)$$

Eq. 6.6 is similar for the other extracted parameter JES^{fitted}. The m_t^{fitted} and JES^{fitted} are obtained from minimum of the function $\mathscr{F}(m_t, JES)$, represented by Eq. 6.10:

$$(m_t^{fitted}, JES^{fitted}) \equiv (-\frac{2\times p_1\times p_2 - p_3\times p_4}{4\times p_0\times p_1 - p_4^2}, -\frac{2\times p_0\times p_3 - p_2\times p_4}{4\times p_0\times p_1 - p_4^2}) \qquad (6.12)$$

The function $\mathscr{F}$ is fitted to the bins near the minimum, in the range of $\pm 20\ GeV/c^2$ in the m_t direction and ± 0.09 in the α_{JES} factor dimension. σ is scaled by $\sqrt{N_{eff}}$ to provide, a rough approximation of the statistical uncertainty on the corresponding fitted parameters. In the next sections, another way of determining the uncertainty on the measurement is mentioned, that includes the effect of the slope of calibration curves.

In order to gauge the effect on calibrations by increasing the impurities of the samples, it is important to check the method with the purest sample first. In the next section, the calibrations for only parton matched $t\bar{t}$ events are described, followed by full fledged calibrations of signal and background samples in Section 6.7.

source	$e+jets, \geq 2tags$	$\mu+jets, \geqq 2tags$	$e+jets,\ 1tag$	$\mu+jets, 1tag$
$t\bar{t}$ production process	125	87	242	167
W+jets	4	6	37	45
Z+jets	1	0	3	3
WW, WZ, ZZ + jets	0	0	2	2
single t production	2	1	4	3
multijets	2	0	19	6
DATA	128	120	286	240

Table 6.2: Event yields in different channels

6.6 Ensemble tests with parton matched $t\bar{t}$ events

In this study, ensemble tests are performed using only the parton matched $t\bar{t}$ events. In addition, only the correct solutions corresponding to the right jet-parton assignments are used to construct the likelihood. Hence, only the first term in Eq. 6.6 is considered, while the

combinatorial and $\mathscr{BG}$ shapes are turned off, while evaluating the likelihood. This section mainly describes, how well the input parameters can be reproduced using the method.

Fig. 6.7, shows the distribution of m_t^{fitted} when the input mass of 172.5 GeV/c^2 is used in the four different channels. The *bias* , which is the difference between the fitted mass and the generator level mass (input), is plotted in Fig. 6.11. This is done for various values of input top mass and is fit linearly, yielding an offset and a slope. As can be seen from the figure, the dependence of the fitted mass on the input mass is stronger for channels with fewer tags. The offset is close to the ideal value for channels with $\geqq 2 tags$. This is somewhat expected as the probability of finding the correct solution in the events for which two of the leading four jets are known to come from the b-quarks. The value of m_t^{fitted} corresponds to the point in 2D likelihood space where $-2\ln(\mathscr{L})$ is minimized with respect to JES and m_t. The value of JES^{fitted} is a few percent away from the input value. The distribution of JES^{fitted} for various channels in which the input value is the nominal value, is shown in Fig. 6.9. To understand the behavior of JES^{fitted} with respect to the input JES, the full 2D fit is performed for two other scenarios :

- All the jets in the events are scaled down by 6% of their values and an appropriate adjustment to the $\not\!\!E_T$ is made. The event selection is done again and the selected events are reconstructed using the same procedure as for the nominal JES. This corresponds to $JES_{input} = 0.94$. The distribution of JES^{fitted} for various channels, in this scenario is shown in Fig. 6.8.

- $JES_{input} = 1.06$ corresponds to scaling the jet energies up by 6% and following the above procedure. The distribution of JES^{fitted} for various channels, in this scenario is shown in Fig. 6.10.

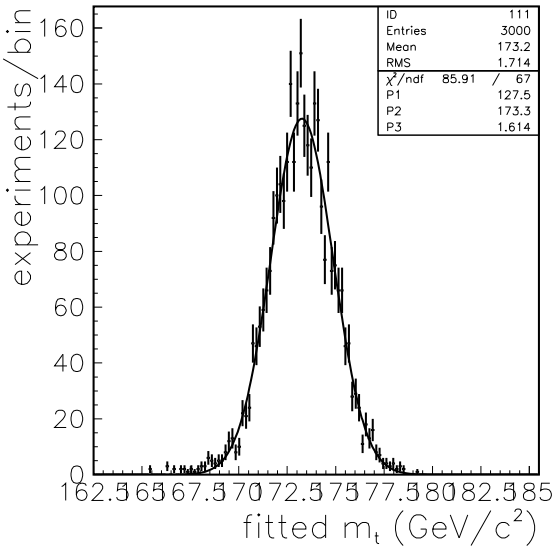
Putting it all together, JES^{fitted} for different input JES is plotted in Fig. 6.12, for different input top masses. Two things follow from these plots :

- JES^{fitted} is independent of input top mass for almost all values of input JES.
- Although the offsets in the plots do not exactly come out to be at the true value, the behavior of JES^{fitted} is as expected, i.e. the fitted values scale up, when the JES_{input} is scaled up and vice versa. This is shown in Table. 6.3.

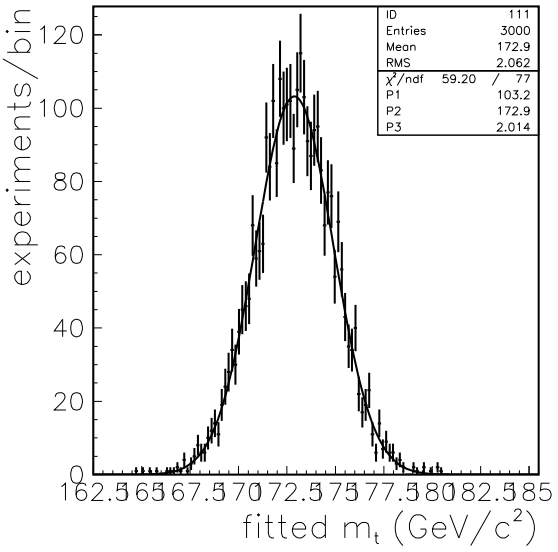
Also, the difference between electron and muon channels is negligible, as expected.

JES_{input}	$e+jets, \geq 2tags$	$\mu+jets, \geq 2tags$	$e+jets, 1tag$	$\mu+jets, 1tag$
0.94	0.922	0.924	0.933	0.935
1	0.97	0.97	0.973	0.974
1.06	1.039	1.04	1.049	1.052

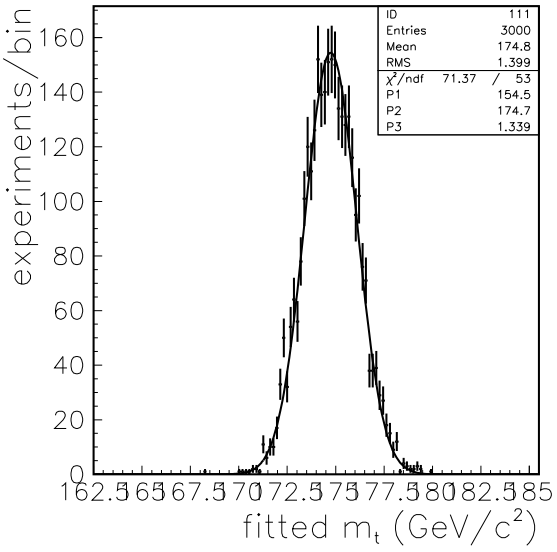
Table 6.3: Offset in JES^{fitted} for three different values of JES_{input} in the four channels



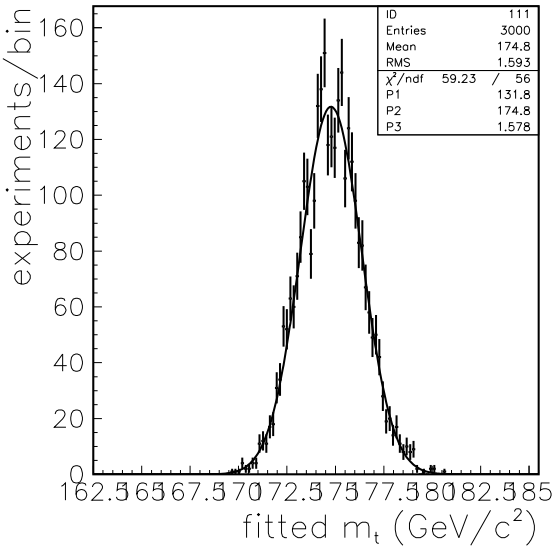
(a) $e + jets, \geq 2tags$



(b) $\mu + jets, \geq 2tags$

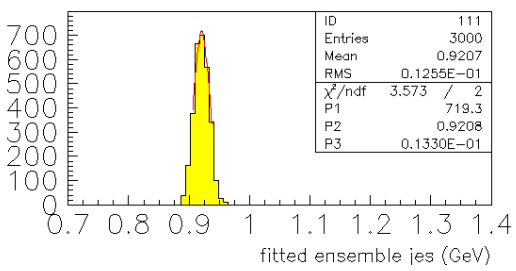


(c) $e + jets, 1tag$

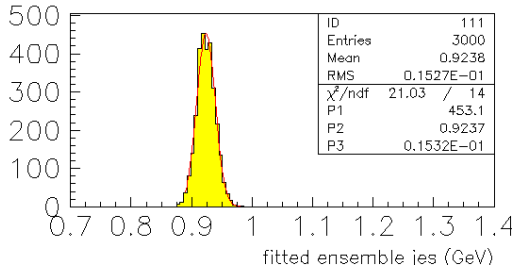


(d) $\mu + jets, 1tag$

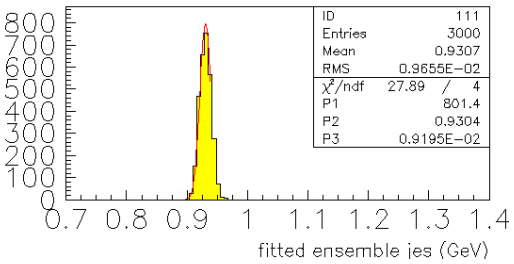
Figure 6.7: m_t^{fitted} distributions with input top quark mass of $172.5\,GeV/c^2$ at JES^{fitted} for $JES_{input} = 1.0$, for different channels (parton matched $t\bar{t}$ events)



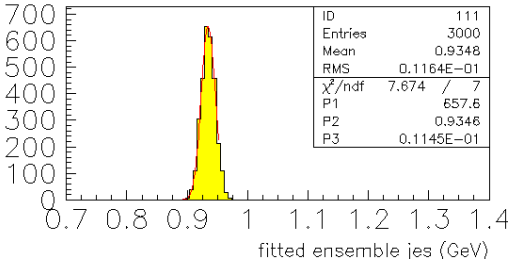
(a) $e + jets, \geq 2tags$



(b) $\mu + jets, \geq 2tags$

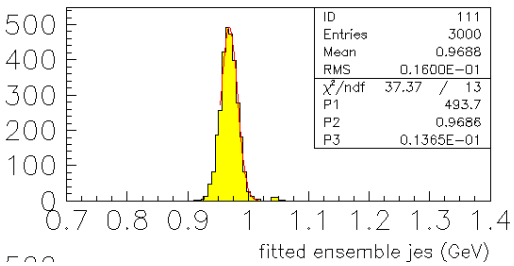


(c) $e + jets, 1tag$

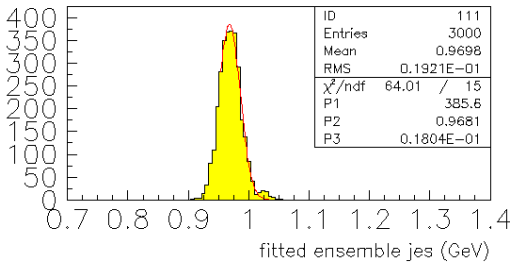


(d) $\mu + jets, 1tag$

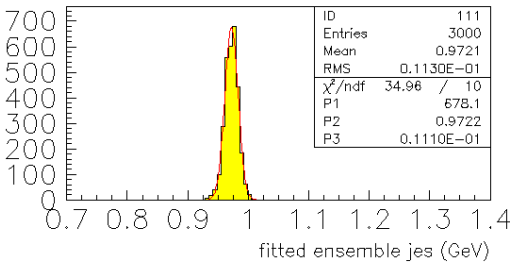
Figure 6.8: JES^{fitted} distribution for $JES_{input} = 0.94$ in parton matched $t\bar{t}$ events with input mass of $172.5\,GeV/c^2$, in different channels



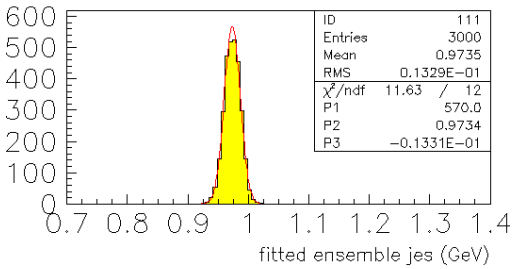
(a) $e + jets, \geq 2tags$



(b) $\mu + jets, \geq 2tags$

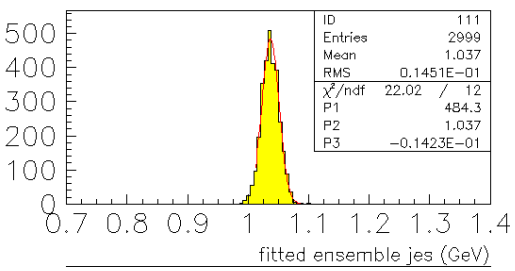


(c) $e + jets, 1tag$

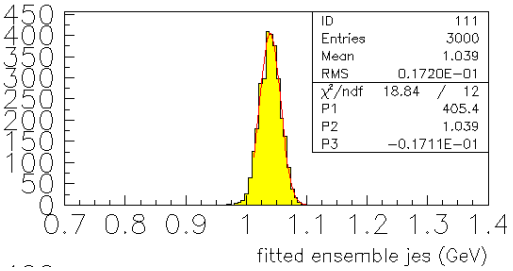


(d) $\mu + jets, 1tag$

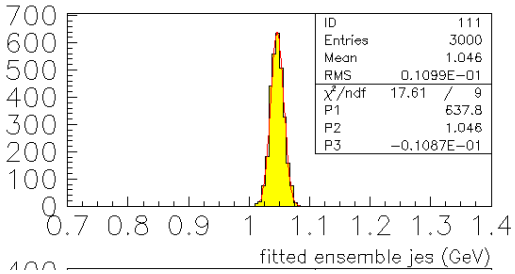
Figure 6.9: JES^{fitted} distribution for $JES_{input} = 1.0$ in parton matched $t\bar{t}$ events with input mass of $172.5\,GeV/c^2$, in different channels



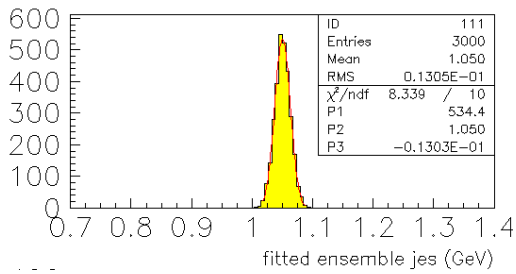
(a) $e + jets, \geq 2tags$



(b) $\mu + jets, \geq 2tags$

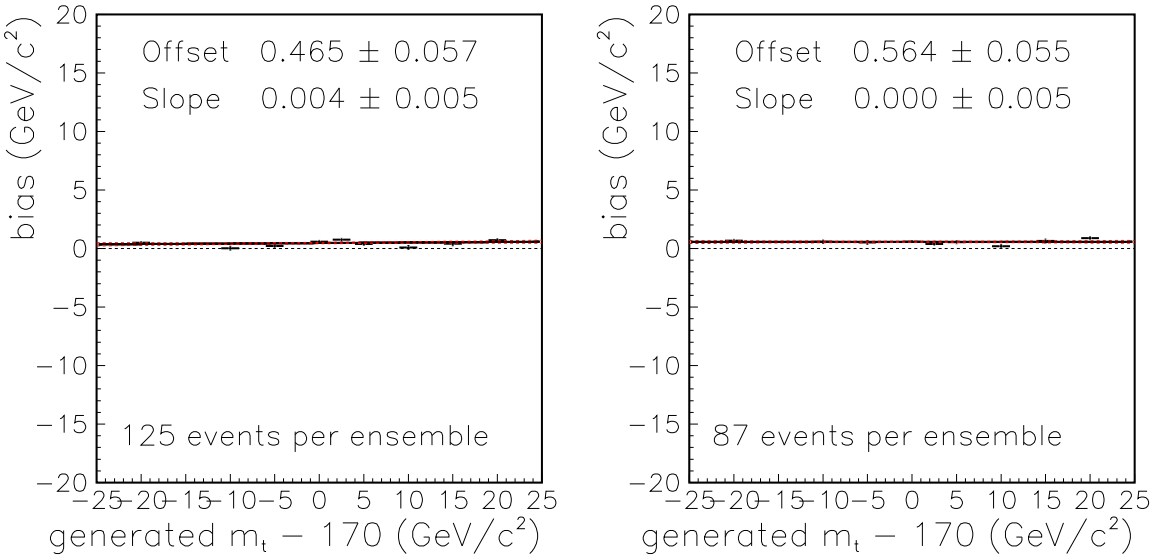


(c) $e + jets, 1tag$

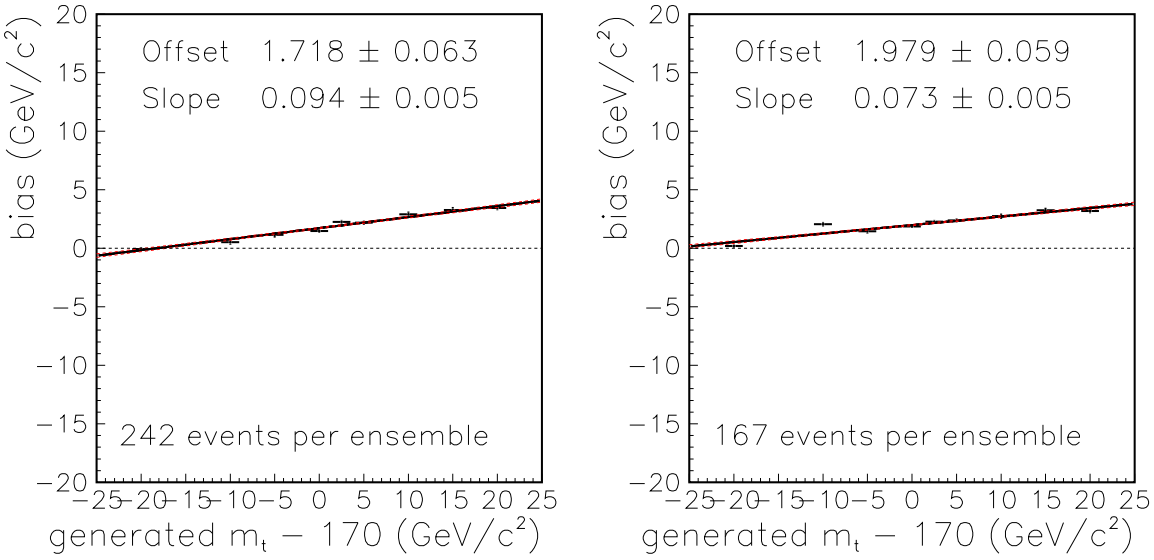


(d) $\mu + jets, 1tag$

Figure 6.10: JES^{fitted} distribution for $JES_{input} = 1.06$ in parton matched $t\bar{t}$ events with input mass of $172.5\,GeV/c^2$, in different channels

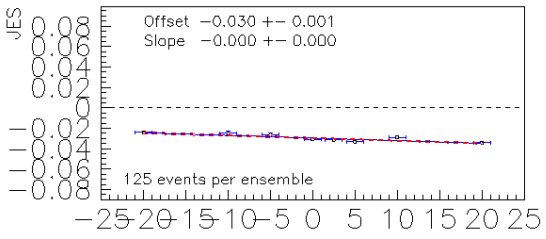


(a) $e + jets, \geq 2tags$ (b) $\mu + jets, \geq 2tags$

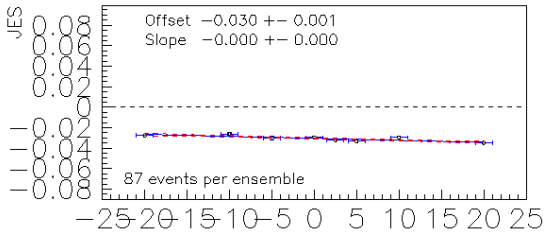


(c) $e + jets$, $1tag$ (d) $\mu + jets$, $1tag$

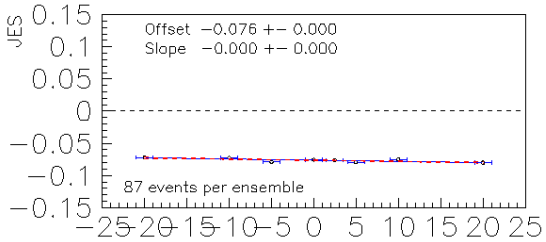
Figure 6.11: m_t^{fitted} calibration at JES^{fitted} for the nominal JES as input in the parton matched $t\bar{t}$ events, for different channels



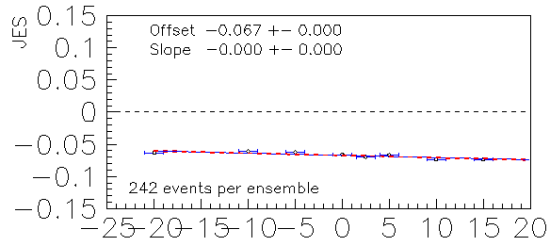
(a) $e + jets, \geq 2tags$



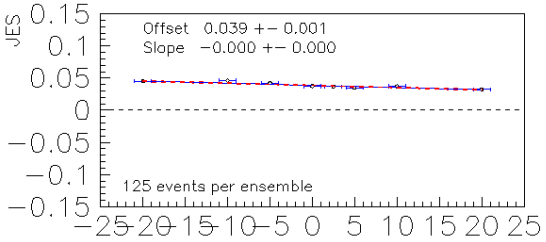
(b) $\mu + jets,\ 1tag$



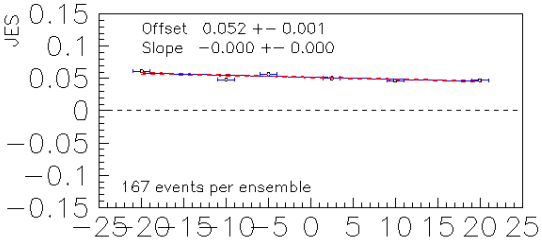
(c) $\mu + jets, \geq 2tags$



(d) $e + jets,\ 1tag$



(e) $e + jets, \geq 2tags$



(f) $\mu + jets,\ 1tag$

Figure 6.12: JES^{fitted} calibrations with JES_{input} scaled by 0% (top), −6% (middle) and +6% (bottom) for parton matched $t\bar{t}$ events with different input masses. The y-axis of all the plots represents $(JES^{fitted} - 1.0)$, while the x-axis represents $(m_t^{generated} - 170)\,GeV/c^2$.

6.7 Ensemble tests and calibration with signal and background events

In this section, the ensemble tests are conducted for samples that are admixtures of $t\bar{t}$ and $W/Z + jets$ events. The sample composition for each sample, closely follows the one

detailed in Table. 6.2, for each of the four channels. The multijet event fractions are low and the shapes do not affect the likelihood and hence they are not included in the ensemble testing procedure. The combinatorial and the $\mathscr{BG}$ shapes are turned on. The $\mathscr{BG}$ shapes are also varied with scaling of JES and hence the likelihood in this case, represents full likelihood. The calibrations are carried out in three stages :

- Single channel JES^{fitted} and m_t^{fitted} calibrations are carried out
- A residual calibration is derived for both the parameters after re-assigning the likelihood values from the calibrated values to the original 2D grid of m_t and JES.
- All the four channels are combined by adding the $-2\ln\mathscr{L}$ values of all the four channels in each calibrated ensemble and m_t and JES are determined from the minimum of the combined likelihood.

The calibrations in these three stages are detailed in the following sections.

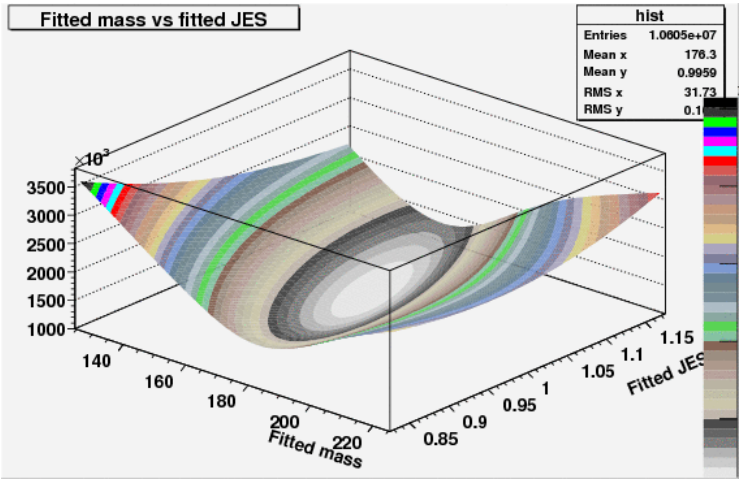


Figure 6.13: A 2D representation of the sum of $-2\ln(Likelihood)$ for 3000 ensembles in one channel

6.7.1 Channel-by-channel calibrations

Ensemble tests in this category involve, construction of pseudo-datasets with signal ($t\bar{t}$) and background ($W/Z + jets$) events in the four channels separately with the likelihood entailing the signal and background shapes. Each ensemble is fitted with the two dimensional function, $\mathscr{F}(m_t, \alpha_{JES})$, whose minimum value yields a m_t^{fitted} and $JES^{fitted.}$. The distributions of these two quantities for the 3000 ensembles that are constructed, are Gaussian. The peak values of m_t^{fitted} distributions for different input top quark masses are tabulated in Tables 6.4-6.7, corresponding to the four channels. Similar to the calibration curves, shown in Fig. 6.11, the m_t^{fitted} peak values are plotted against the generator level mass in Fig. 6.14 in each of the four channels. The scale of these plots is zoomed in near 172.5 GeV, for convenience. Also, the RMS of the m_t^{fitted} distributions are entered in the table, which are multiplied by $\frac{1}{\sqrt{N_{eff}}}$ to represent the approximate statistical error on the peak value of the m_t^{fitted} distribution. These errors are shown as the error bars in Fig. 6.14. To account for the correlation of statistical uncertainties, a 95% confidence interval band (blue shaded area) is drawn along a linear fit to the representative points. The parameters of these fits represents the mass calibration, which are used in the next section. The linear fit parameters (p_0, p_1) are tabulated in Table. 6.9. Also, shown in the plots is the ideal scenario where the input mass is equal to the output mass. This is drawn as a dotted blue line in the plots and is closer to the actual calibration for channels with $\geq 2 tags$.

Similarly, the parameter, JES^{fitted} is calibrated by repeating the the whole procedure at six different values of input JES (0.94, 0.97, 1.0, 1.03, 1.06). The mass calibrations, mentioned in the above paragraph are obtained for $JES_{input} = 1.0$. The six values of JES_{input} are obtained, not just by scaling the jets with the appropriate factor, but also by repeating the event selection and adjustment of the $\not\!\!E_T$. The peak values of the JES^{fitted} distributions for the four channels at the input top quark mass of 172.5 GeV are tabulated in Table.

6.8. These are plotted as $(JES^{fitted} - 1.0)\, vs\, (JES_{input} - 1.0)$ in Fig. 6.15, with the 95% confidence interval on the linear fit,shown as the shaded yellow band, while the ideal scenario depicted as the red dotted line. The parameters of the linear fits, which is the JES calibration, are tabulated in Table 6.9.

$Generator Mass$ (GeV/c^2)	$Fitted mass$ (GeV/c^2)	RMS GeV/c^2	$Scale = \frac{1}{\sqrt{N_{eff}}}$ (GeV/c^2)	$Error$ (GeV/c^2)
150	148.8	2.083	0.086	0.179138
160	157.1	2.363	0.108	0.255204
165	161.9	2.515	0.109	0.274135
170	166.3	2.705	0.068	0.18394
172.5	168.5	2.805	0.072	0.20196
175	170.9	2.855	0.0796	0.227258
180	175.5	2.803	0.079	0.274694
185	179.9	3.028	0.0978	0.2961384
190	184.4	3.412	0.0728	0.2483936

Table 6.4: m_t^{fitted} peak values for different input top quark mass for ensembles with both $t\bar{t}$ and $W/Z + jets$ events in $e + jets, \geqq 2tags$ channel.

$Generator Mass$ (GeV/c^2)	$Fitted mass$ (GeV/c^2)	RMS (GeV/c^2)	$Scale = \frac{1}{\sqrt{N_{eff}}}$	$Error$ (GeV/c^2)
150	148.9	2.678	0.0709	0.1898702
160	157.1	2.962	0.0885	0.262137
165	161.7	2.951	0.0896	0.2644096
170	165.6	3.235	0.0563	0.1821305
172.5	167.7	3.305	0.0605	0.1999525
175	169.8	3.39	0.06623	0.2245197
180	174	3.398	0.0819	0.2782962
185	179	3.781	0.0813	0.3073953
190	184.1	3.991	0.0606	0.2418546

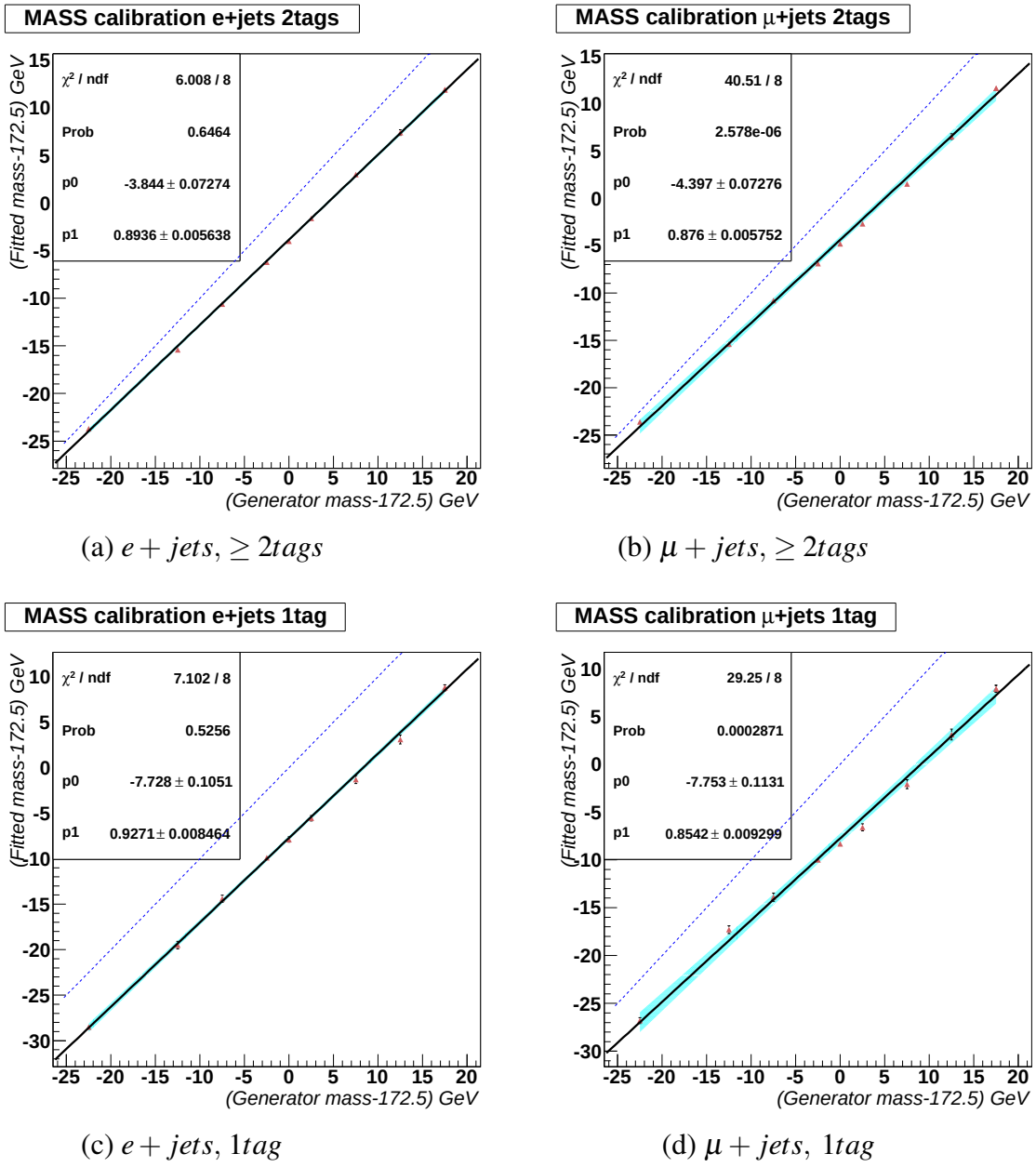
Table 6.5: m_t^{fitted} peak values for different input top quark mass for ensembles with both $t\bar{t}$ and $W/Z + jets$ events in $\mu + jets, \geqq 2tags$ channel.

$Generator Mass$ (GeV/c^2)	$Fitted mass$ (GeV/c^2)	RMS (GeV/c^2)	$Scale = \frac{1}{\sqrt{N_{eff}}}$	$Error$ (GeV/c^2)
150	144	2.764	0.1062	0.2935368
160	153	3.091	0.137	0.423467
165	158.1	2.821	0.1392	0.3926832
170	162.6	2.941	0.0842	0.2476322
172.5	164.6	3.276	0.0904	0.2961504
175	166.94	3.019	0.1	0.3019
180	171.19	3.435	0.126	0.43281
185	175.6	3.977	0.125	0.497125
190	181.28	3.587	0.091	0.326417

Table 6.6: m_t^{fitted} peak values for different input top quark mass for ensembles with both $t\bar{t}$ and $W/Z+jets$ events in $e+jets$, $1tag$ channel.

$Generator Mass$ (GeV/c^2)	$Fitted mass$ (GeV/c^2)	RMS (GeV/c^2)	$Scale = \frac{1}{\sqrt{N_{eff}}}$	$Error$ (GeV/c^2)
150	145.7	3.498	0.0905	0.316569
160	155.2	3.494	0.1169	0.4084486
165	158.6	3.634	0.1177	0.4277218
170	162.5	3.649	0.0723	0.2638227
172.5	164.16	3.596	0.0775	0.27869
175	165.9	4.25	0.0858	0.36465
180	170.4	4.32	0.1081	0.466992
185	175.6	5.206	0.1083	0.5638098
190	180.36	4.75	0.07855	0.3731125

Table 6.7: m_t^{fitted} peak values for different input top quark mass for ensembles with both $t\bar{t}$ and $W/Z+jets$ events in $\mu+jets$, $1tag$ channel.



(a) $e + jets, \geq 2tags$ (b) $\mu + jets, \geq 2tags$

(c) $e + jets$, $1tag$ (d) $\mu + jets$, $1tag$

Figure 6.14: $(Fitted\ mass - 172.5)\, GeV/c^2$ vs. $(Generated\ mass - 172.5)\, GeV/c^2$ for different channels for ensembles consisting of signal and background events.

JES_{input}	JES^{fitted}	RMS
0.94	0.976	0.024
0.97	1.003	0.024
1	1.029	0.023
1.03	1.055	0.022
1.06	1.081	0.022

(a) $e + jets, \geqq 2 tags$

JES_{input}	JES^{fitted}	RMS
0.94	0.981	0.029
0.97	1.009	0.027
1	1.034	0.027
1.03	1.061	0.026
1.06	1.084	0.026

(b) $\mu + jets, \geqq 2 tags$

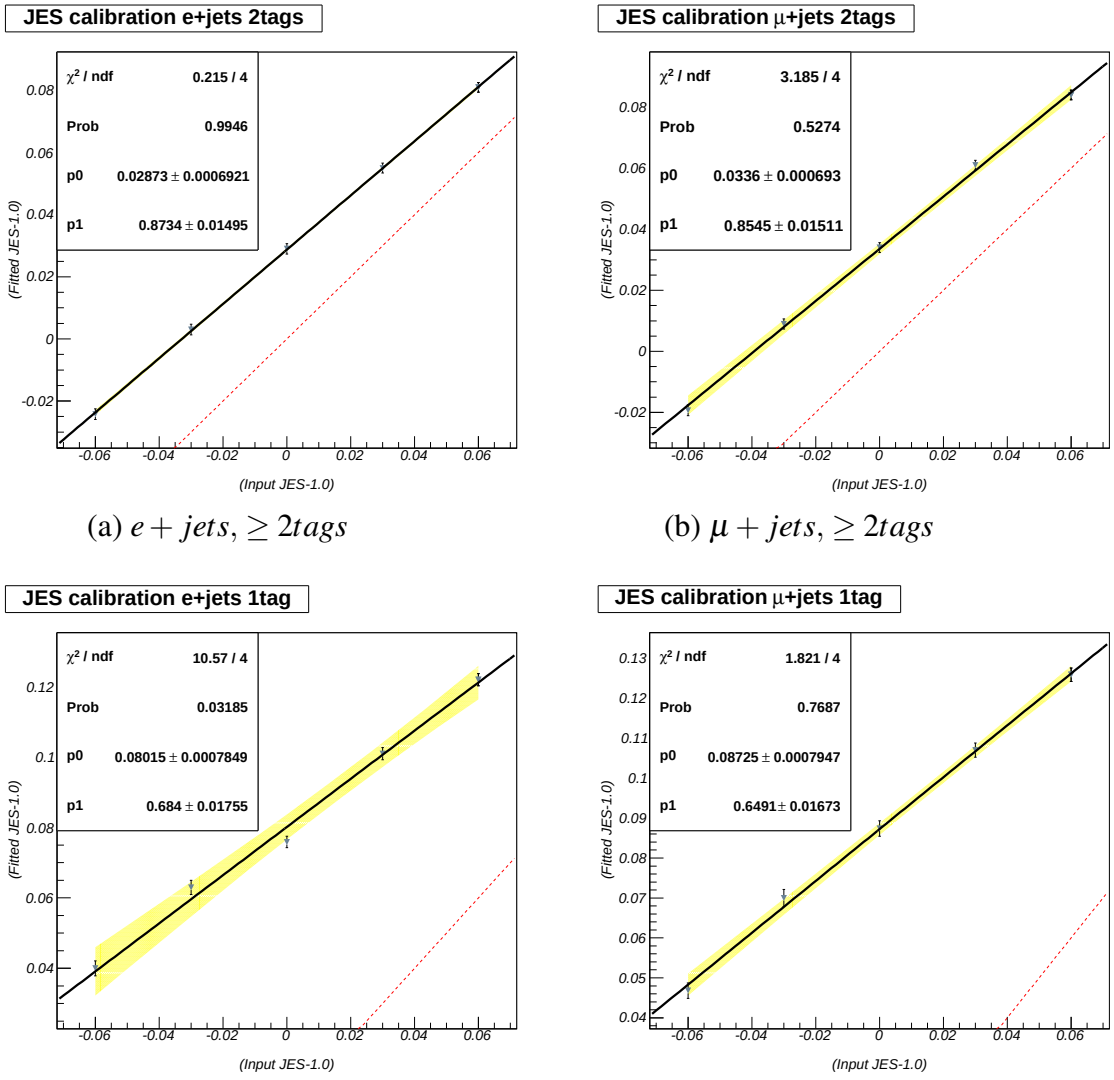
JES_{input}	JES^{fitted}	RMS
0.94	1.04	0.024
0.97	1.063	0.022
1	1.076	0.018
1.03	1.101	0.019
1.06	1.122	0.019

(c) $e + jets, 1 tag$

JES_{input}	JES^{fitted}	RMS
0.94	1.047	0.025
0.97	1.07	0.027
1	1.087	0.025
1.03	1.107	0.023
1.06	1.126	0.022

(d) $\mu + jets, 1 tag$

Table 6.8: JES^{fitted} peak values for six different values of JES_{input} for an input top quark mass of 172.5 GeV, across the four channels.



(a) $e + jets, \geq 2tags$ (b) $\mu + jets, \geq 2tags$

(c) $e + jets,\ 1tag$ (d) $\mu + jets,\ 1tag$

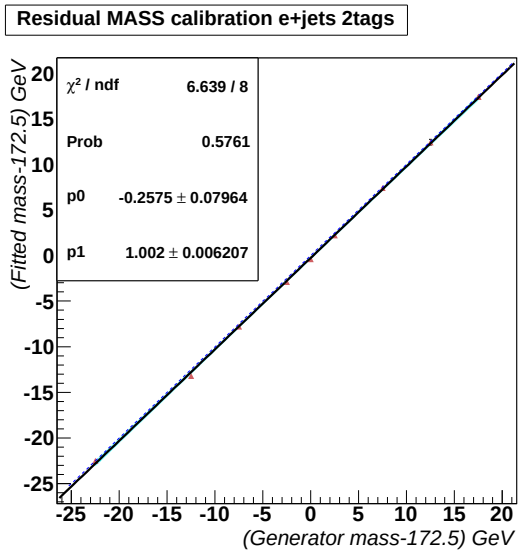
Figure 6.15: $(Fitted\ JES - 1.0)$ vs. $(Input\ JES - 1.0)$ for different channels

Channel	$(m_t^{fitted} - 172.5) = p_1 \times (m_t^{input} - 172.5) + p_0$		$(JES^{fitted} - 1.0) = p_1 \times (JES_{input} - 1.0) + p_0$	
	$p_0\,(offset)$	$p_1\,(slope)$	$p_0\,(offset)$	$p_1\,(slope)$
$e + jets, \geqq 2tags$	-3.844	0.894	0.0287	0.873
$\mu + jets, \geqq 2tags$	-4.397	0.876	0.0334	0.855
$e + jets,\ 1tag$	-7.728	0.927	0.0802	0.684
$\mu + jets,\ 1tag$	-7.753	0.854	0.0873	0.649

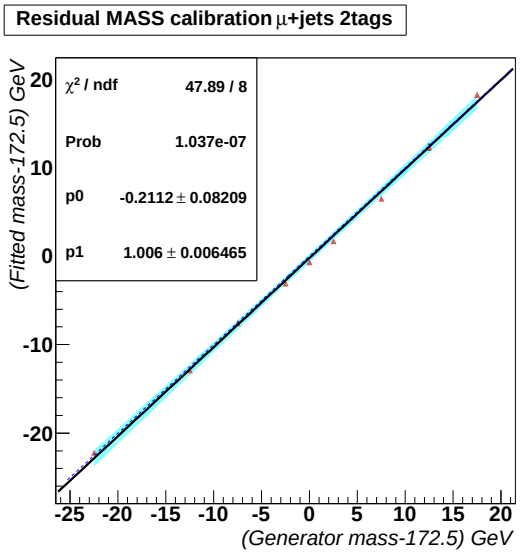
Table 6.9: First Calibration parameters for m_t^{fitted} and JES^{fitted} in the four channels.

6.7.2 Residual Calibrations for each channel

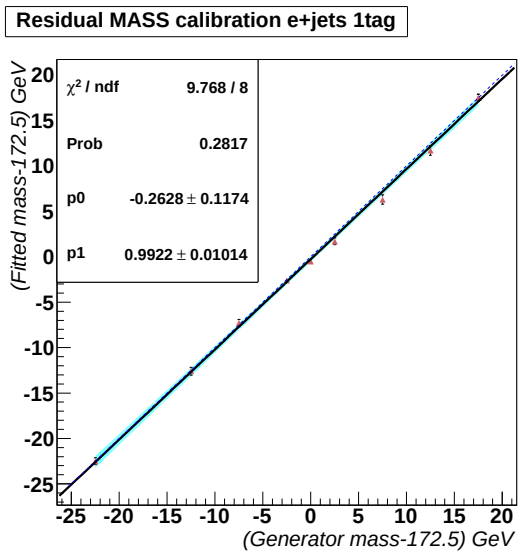
The calibrations obtained in the previous section are used to reproduce the input values of top mass and JES. First, the likelihood values in the slices of the 2D grid along the JES direction is fitted with a sum of quadratic and quartic functions in each of the mass bins for every ensemble. This allows one to re-evaluate the likelihood at JES^{fitted}, using the linear transformation, whose parameters are listed in Table 6.9. The likelihood thus evaluated is re-assigned to the bin in the original 2D grid. This completes the calibration along the JES axis. m_t is calibrated in a similar way through the linear transformation of m_t^{fitted}, listed in the Table. 6.9. The likelihood value is read off from the bin that contains m_t^{fitted} and is re-assigned to the bin that contains m_t^{input}. This is done for all the values of JES in the 2D grid. It is to be noted that the bin width of $1\,GeV/c^2$, is the limitation in the precise knowledge of m_t^{fitted}. Each ensemble likelihood distribution is again fitted with the two dimensional function of $\mathscr{F}(m_t, \alpha_{JES})$ and the peak values of the new m_t^{fitted} and JES^{fitted}, turn out to be Gaussian in nature. These are tabulated in Tables 6.10-6.13 for m_t^{fitted} and Table 6.14 for JES^{fitted}. The new fitted values are again put on a calibration curve, which is referred to as the *residual calibration*. The residual calibration for the m_t^{fitted} peak values vs. the generator level top quark mass is shown in Fig. 6.16 and the residual JES calibration for different channels is shown in Fig. 6.17. As can been seen from both the figures, the dotted lines that represent the ideal scenario of input being equivalent to the output, are eclipsed by the linear fit to the points, within the 95% confidence interval band. This procedure provides a cross check on the calibrations derived in the previous section. It is also worthwhile to note that the offsets in the residual calibrations are within the limits of the bin size and hence further applying these calibrations would require some interpolation within inside the bin. This is discussed in the next section that also discusses the combination of the four channels.



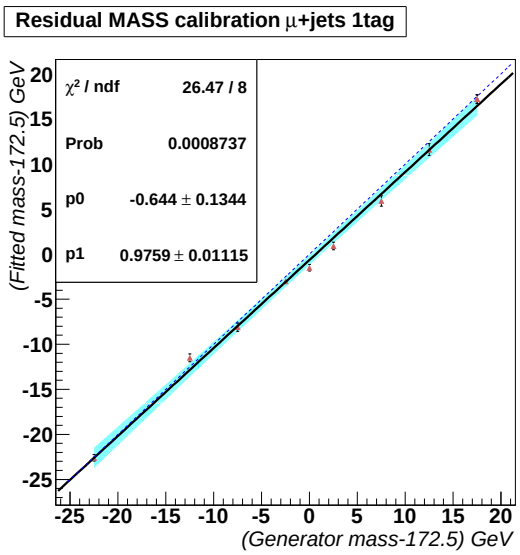
(a) $e + jets, \geq 2tags$



(b) $\mu + jets, \geq 2tags$



(c) $e + jets, 1tag$



(d) $\mu + jets, 1tag$

Figure 6.16: $(Fitted\ mass - 172.5)\,GeV$ vs. $(Generated\ mass - 172.5)\,GeV$ for different channels at $JES = 1$

$Generator\,Mass$ (GeV/c^2)	$Fitted\,mass$ (GeV/c^2)	RMS GeV/c^2	$Scale = \frac{1}{\sqrt{N_{eff}}}$ (GeV/c^2)	$Error$ (GeV/c^2)
150	150	2.326	0.086	0.2
160	159.3	2.363	0.108	0.2552
165	164.7	2.515	0.109	0.2741
170	169.6	3.016	0.068	0.2051
172.5	172.1	3.128	0.072	0.2252
175	174.7	3.174	0.0796	0.2527
180	179.9	3.12	0.098	0.3058
185	184.9	3.339	0.0978	0.3266
190	189.9	3.772	0.0728	0.2746

Table 6.10: m_t^{fitted} peak values (second calibration) for different input top quark mass for ensembles with both $t\bar{t}$ and $W/Z + jets$ events in $e + jets, \geqq 2tags$ channel.

$Generator\,Mass$ (GeV/c^2)	$Fitted\,mass$ (GeV/c^2)	RMS GeV/c^2	$Scale = \frac{1}{\sqrt{N_{eff}}}$ (GeV/c^2)	$Error$ (GeV/c^2)
150	150.3	3.020	0.0709	0.2141
160	159.6	3.323	0.0885	0.2941
165	164.9	3.336	0.0896	0.2989
170	169.4	3.670	0.0563	0.2066
172.5	171.8	3.739	0.0605	0.2262
175	174.2	3.864	0.06623	0.25591
180	179	3.830	0.0819	0.31367
185	184.8	4.246	0.0813	0.3452
190	190.7	4.466	0.0606	0.27064

Table 6.11: m_t^{fitted} peak values (second calibration) for different input top quark mass for ensembles with both $t\bar{t}$ and $W/Z + jets$ events in $\mu + jets, \geqq 2tags$ channel.

$Generator\,Mass$ (GeV/c^2)	$Fitted\,mass$ (GeV/c^2)	RMS GeV/c^2	$Scale = \frac{1}{\sqrt{N_{eff}}}$ (GeV/c^2)	$Error$ (GeV/c^2)
150	150	3.637	0.1062	0.38625
160	159.9	3.161	0.137	0.43306
165	165.2	2.904	0.1392	0.40423
170	169.9	3.004	0.0842	0.2529
172.5	172	3.436	0.0904	0.3106
175	174.2	3.728	0.1	0.3728
180	178.8	4.072	0.126	0.5131
185	184.2	4.132	0.125	0.5165
190	190	4.180	0.091	0.3804

Table 6.12: m_t^{fitted} peak values (second calibration) for different input top quark mass for ensembles with both $t\bar{t}$ and $W/Z+jets$ events in $e+jets$, $1\,tag$ channel.

$Generator\,Mass$ (GeV/c^2)	$Fitted\,mass$ (GeV/c^2)	RMS GeV/c^2	$Scale = \frac{1}{\sqrt{N_{eff}}}$ (GeV/c^2)	$Error$ (GeV/c^2)
150	149.9	4.000	0.0905	0.362
160	161	3.755	0.1169	0.439
165	164.4	3.959	0.1177	0.466
170	169.5	4.039	0.0723	0.292
172.5	171	4.723	0.0775	0.366
175	173.4	4.825	0.0858	0.414
180	178.4	5.328	0.1081	0.576
185	184.1	6.085	0.1083	0.659
190	189.7	6.170	0.0786	0.485

Table 6.13: m_t^{fitted} peak values (second calibration) for different input top quark mass for ensembles with both $t\bar{t}$ and $W/Z+jets$ events in $\mu+jets$, $1\,tag$ channel.

$Channel$	$(m_t^{fitted}-172.5)=p_1\times(m_t^{input}-172.5)+p_0$		$(JES^{fitted}-1.0)=p_1\times(JES_{input}-1.0)+p_0$	
	$p_0\,(offset)$	$p_1\,(slope)$	$p_0\,(offset)$	$p_1\,(slope)$
$e+jets, \geqq 2tags$	-0.258	1.002	-2.38×10^{-4}	1.008
$\mu+jets, \geqq 2tags$	-0.211	1.006	-3.41×10^{-4}	1.006
$e+jets, 1tag$	-0.263	0.992	-1.67×10^{-4}	0.978
$\mu+jets, 1tag$	-0.644	0.976	0.0025	0.982

Table 6.15: Residual Calibration parameters for m_t^{fitted} and JES^{fitted} in the four channels.

JES_{input}	JES^{fitted}	RMS
0.94	0.938	0.024
0.97	0.969	0.024
1	1.001	0.023
1.03	1.030	0.022
1.06	1.060	0.022

(a) $e + jets, \geqq 2 tags$

JES_{input}	JES^{fitted}	RMS
0.94	0.937	0.029
0.97	0.971	0.027
1	1.000	0.027
1.03	1.032	0.026
1.06	1.059	0.026

(b) $\mu + jets, \geqq 2 tags$

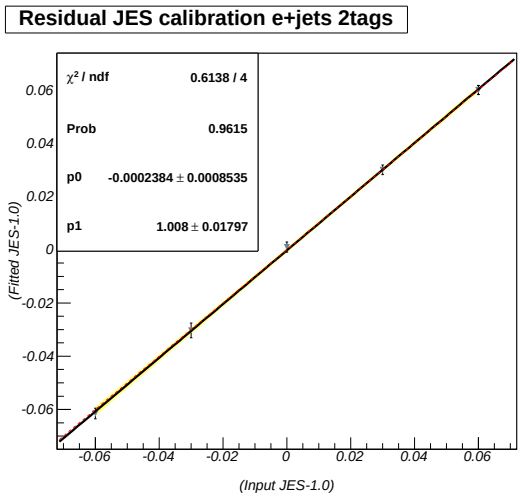
JES_{input}	JES^{fitted}	RMS
0.94	0.94	0.024
0.97	0.974	0.022
1	0.997	0.018
1.03	1.029	0.019
1.06	1.059	0.019

(c) $e + jets, 1 tag$

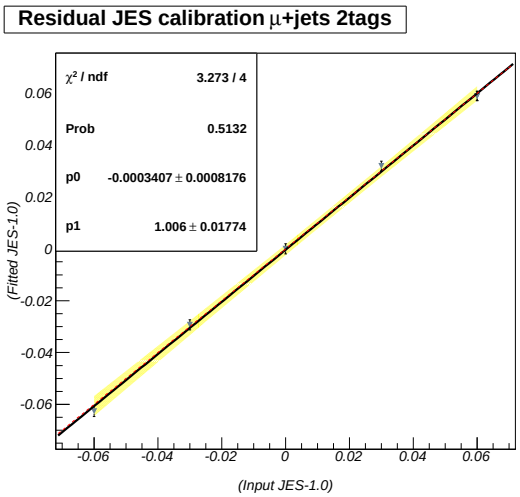
JES_{input}	JES^{fitted}	RMS
0.94	0.942	0.025
0.97	0.973	0.027
1	1.004	0.025
1.03	1.032	0.023
1.06	1.061	0.022

(d) $\mu + jets, 1 tag$

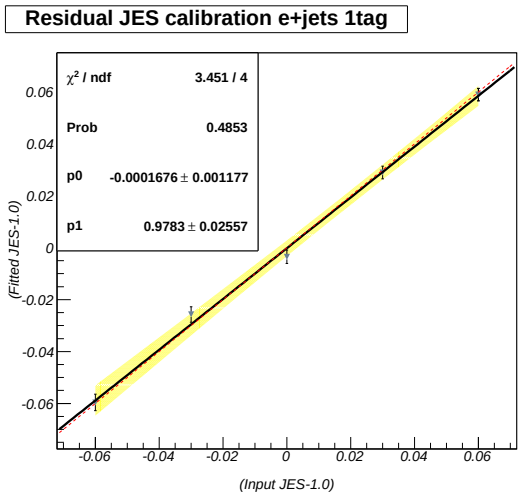
Table 6.14: JES^{fitted} (second calibration) peak values for six different values of JES_{input} for an input top quark mass of 172.5 GeV, across the four channels.



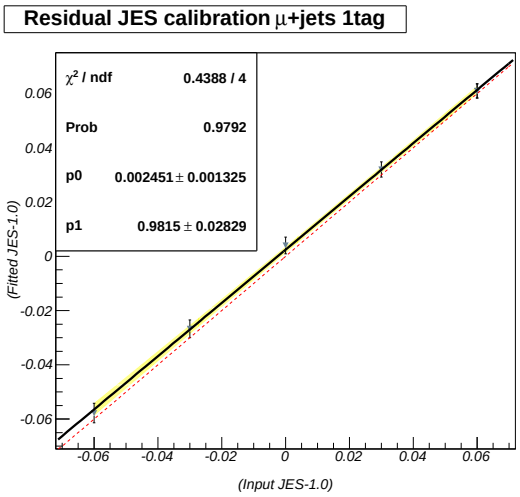
(a) $e + jets, \geq 2tags$



(b) $\mu + jets, \geq 2tags$



(c) $e + jets, 1tag$



(d) $\mu + jets, 1tag$

Figure 6.17: $(Fitted\ JES - 1.0)$ vs. $(Input\ JES - 1.0)$ for different channels

6.7.3 Combination

The combination of four channels proceeds through two stages. First, each channel is calibrated according to the linear transformation presented in Table 6.15. This procedure was described in the previous section and the only difference lies in the fact that the likelihood value at m_t^{fitted} is obtained through bi-linear transformation, demonstrated in Fig. 6.18. For this, the quadrant of the bin containing m_t^{fitted} is identified and the likelihood values of its nearest neighboring points are used in a bi-linear transformation to obtain the likelihood value at m_t^{fitted}. Unlike, using the bin center closest to m_t^{fiited} (as done in previous calibrations), this procedure gives a more accurate local characterization of the likelihood. Every ensemble in each channel is thus calibrated in the above fashion. The 2D negative log-likelihood $(-2\ln\mathscr{L})$ distribution of each calibrated ensemble is then added across all channels, before fitting it again with the two dimensional function $\mathscr{F}(m_t, \alpha_{JES})$. The resulting m_t^{fitted} and JES^{fitted} values are tabulated in Tables 6.16-6.18.

In the discussion that follows, two m_t^{fitted} values are quoted :

- The '2D' fitted mass refers to the peak value of the m_t^{fitted} distribution obtained from the minimum of $\mathscr{F}(m_t, JES)$ and hence it is the fitted mass at JES^{fitted}.
- The '1D' fitted mass refers to the peak value of m_t^{fitted} distribution obtained from the minimum of the 2D likelihood distribution when sliced at JES=1

When the JES^{fitted} comes out to be equal to one, both the definitions of m_t^{fitted} yield the same value but the 2D fitted mass uncertainty has a component of the uncertainty of JES^{fitted}, while the 1D fitted mass uncertainty is only statistical in nature. The errors (σ) on the fitted parameters are re-estimated using the Eq. 6.11, which in the two dimensional case represents a rectangle bounding the elliptical cross-section of the $-2\ln\mathscr{L}$ distribution when its value increases by one, with respect to the minimum value. These errors are used

in defining the *pull* distribution of the estimated parameters. Pull is the deviation of the fitted parameter from its average value divided by the uncertainty. When a large number of pseudo-experiments are used for parameter estimation, the pull distribution is expected to be a Gaussian, centered at zero and with unit width. In this analysis, pulls are defined as follows :

$$pull_{mass} = \frac{m_t^{fitted} - m_t^{generator}}{\sigma_{mass}} \tag{6.13}$$

$$pull_{JES} = \frac{JES^{fitted} - JES^{input}}{\sigma_{JES}} \tag{6.14}$$

The distribution of pulls is shown in Fig. 6.20. σ_{mass} and σ_{JES} are estimated using Eq. 6.11 and are scaled by a constant factor of 1.24 and 1.62 respectively to allow for the widths of the pulls to be close to one. This is demonstrated in Fig. 6.21 which emphasizes the fact that the pull widths are constant over the entire mass range for both the 2D and 1D fitted mass. The distribution of the fitted parameters for the nominal input JES and an input top quark mass of 172.5 GeV are shown in Fig. 6.19. When put on a calibration curve, the 2D and 1D fitted masses lie within 0.5 GeV of the true (input) value and the 95% confidence interval bands (2σ) around the linear fit to the points, are used in assessing the systematic uncertainty on the measurement, as shown in Fig. 6.22. More distributions of fitted parameters and their respective pulls at three different input top quark mass and three different input JES factor, are shown in Appendix C.

The procedure described in this section is applied to the data and the results with systematic errors are detailed in the next chapter.

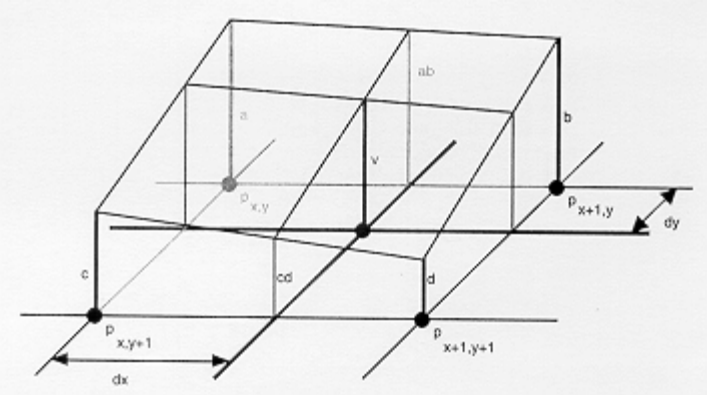


Figure 6.18: Interpolation of likelihood values inside a bin through bi-linear transformation or nearest neighbor sampling. {Ref.[6] }

$Generator Mass$ (GeV/c^2)	$2D\ Fitted mass$ (GeV/c^2)	RMS GeV/c^2	$Scale = \frac{1}{\sqrt{N_{eff}}}$ (GeV/c^2)	$Error$ (GeV/c^2)
150	150.4	1.491	0.05103	0.07608573
160	160.2	1.555	0.065	0.101075
165	165.2	1.576	0.0657	0.1035432
170	170	1.675	0.0405	0.0678375
172.5	172.3	1.782	0.04336	0.07726752
175	174.7	1.83	0.0478	0.087474
180	179.7	1.886	0.05975	0.1126885
185	185	2.001	0.0595	0.1190595
190	190.5	2.149	0.0437	0.0939113

Table 6.16: '2D' fitted mass for the combination of four channels.

JES_{input}	JES^{fitted}	RMS
0.94	0.938	0.017
0.97	0.971	0.016
1	0.999	0.015
1.03	1.030	0.014
1.06	1.059	0.014

Table 6.18: Fitted JES for the combination of four channels.

$Generator Mass$ (GeV/c^2)	$2D\ Fitted mass$ (GeV/c^2)	RMS GeV/c^2	$Scale = \frac{1}{\sqrt{N_{eff}}}$ (GeV/c^2)	$Error$ (GeV/c^2)
150	150.1	1.135	0.05103	0.05791905
160	160.2	1.132	0.065	0.07358
165	165.1	1.12	0.0657	0.073584
170	169.8	1.141	0.0405	0.0462105
172.5	172.3	1.185	0.04336	0.0513816
175	175	1.154	0.0478	0.0551612
180	180	1.229	0.05975	0.07343275
185	185.3	1.255	0.0595	0.0746725
190	190.5	1.317	0.0437	0.0575529

Table 6.17: '1D' fitted mass for the combination of four channels.

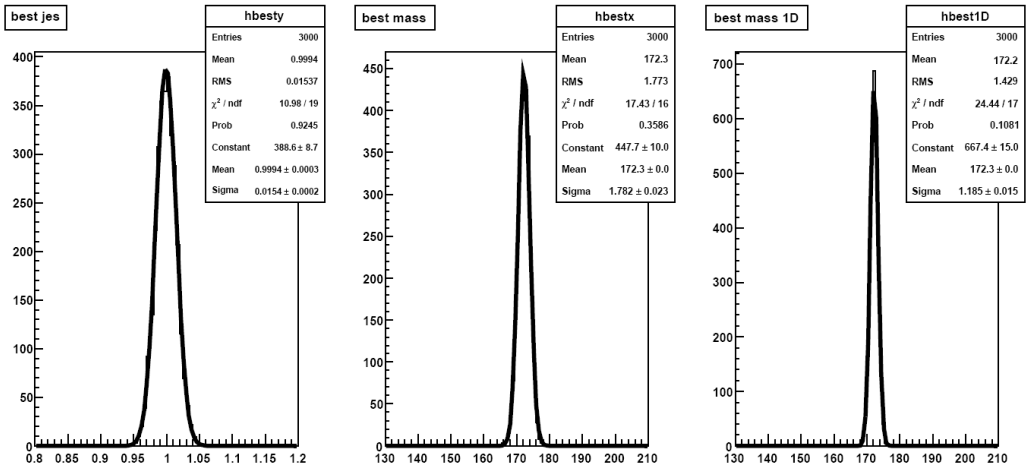


Figure 6.19: JES^{fitted}, 2D-m_t^{fitted} and 1D-m_t^{fitted} distributions for nominal input JES and an input top quark mass of 172.5 GeV.

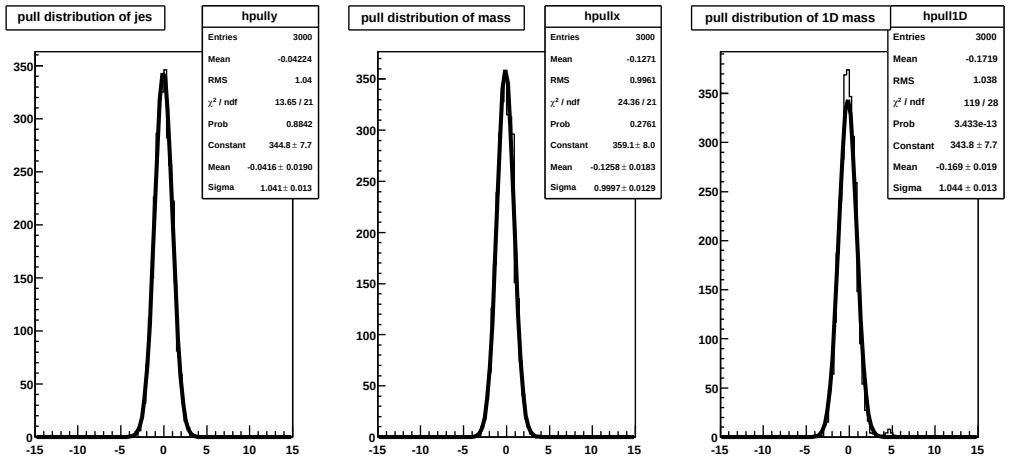


Figure 6.20: Pull distributions of JES^{fitted}, m_t^{fitted}(2D) and m_t^{fitted} (1D) for nominal input JES and an input top quark mass of 172.5 GeV.

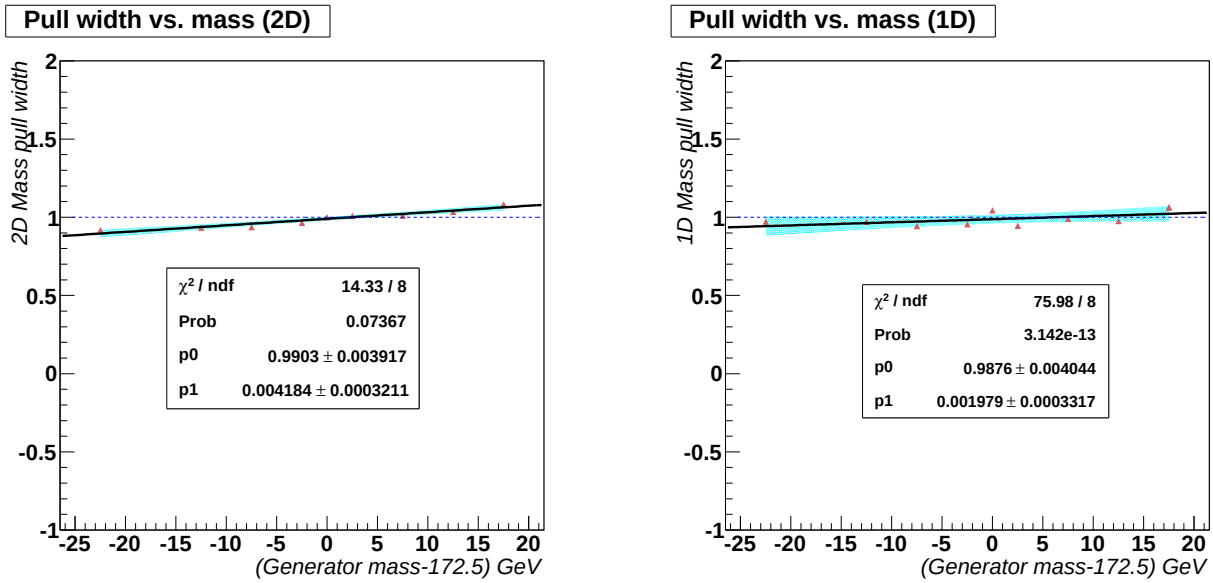


Figure 6.21: Pull width vs. generator level top mass for the '2D' and '1D' cases ($Input\,JES = nominal\,JES$).

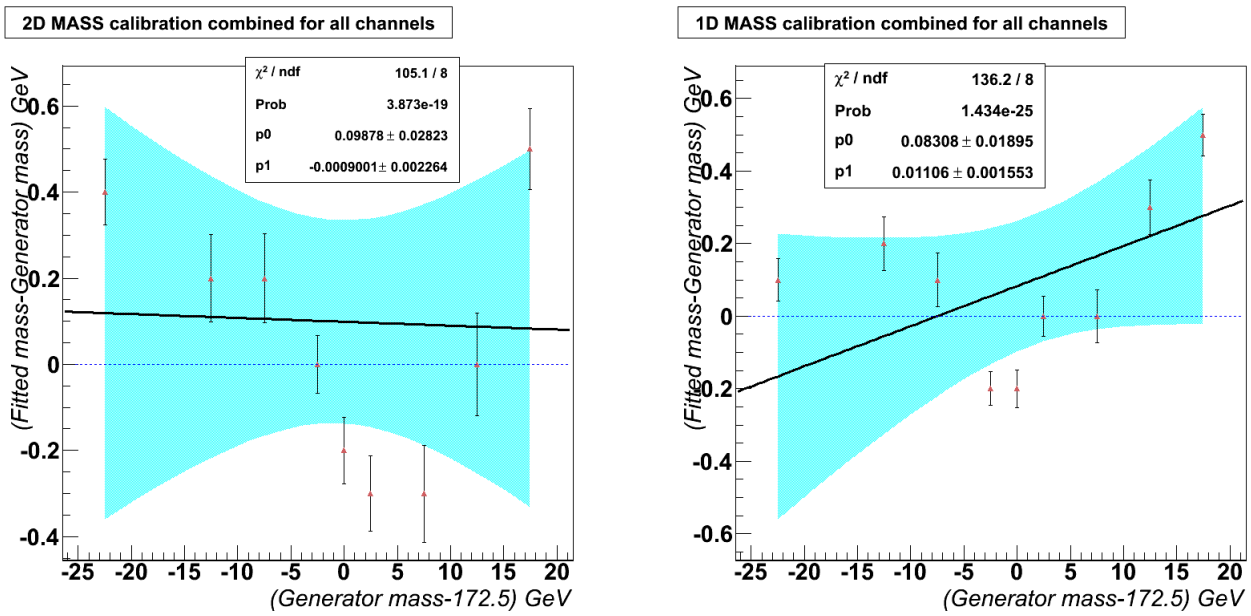


Figure 6.22: Final mass calibrations after combining the four channels.

CHAPTER **Seven**

Results and Systematic uncertainties

7.1 Data Results

The calibrations deduced in the previous chapter are applied to the data events to obtain the observed fitted mass. The results are tabulated in Table. 7.1. To estimate the uncertainties that are statistical and JES related, three measurements are quoted in the table. The full two dimensional (2D) result includes both statistical and JES components in the quoted uncertainty and is derived from the two dimensional likelihood, that is obtained in the same way as described in the previous chapter. The one dimensional (1D) measurement refers to the resultant fitted mass and its uncertainty for the case when the fitted JES is equal to unity. If the Monte Carlo JES was assumed to be true JES, as observed in data, a 1D measurement would result with an uncertainty that is purely statistical in nature. The sample used to derive the JES, described in Section. 3.6, has an uncertainty associated with it, depicted by 1 σ bands that are shown in Fig. 3.6. In this way, fitted JES could be constrained to unity by addition of a prior to $-2\ln(\mathscr{L})$ with a width that is compatible with the difference in the fitted JES values for the $+1\,\sigma$ and $-1\,\sigma$ values for the input JES. The resulting measurement is referred to as, 2D with *prior*. The prior used in this analysis is

given as follows :

$$prior = -2 \times \ln\left\{ \frac{1}{\sqrt{2\pi}} \times \frac{1}{0.012} \times \exp\left(-\frac{(y-1)^2}{2\times(0.012)^2} \right) \right\}$$

Channel	Fitted Mass GeV/c^2	Fitted JES
$e+jets, \geqq 2tags$	170.0±3.12 $(stat+JES)$	1.06±0.022
$\mu+jets, \geqq 2tags$	177.1±3.17 $(stat+JES)$	1.00 ± 0.028
$e+jets, 1tag$	175.0±3.84 $(stat+JES)$	1.01 ± 0.028
$\mu+jets, 1tag$	168.1±3.96 $(stat+JES)$	1.07 ± 0.028
$All\ 4\ channels\,(2D)$	172.6±1.75 $(stat+JES)$	1.03 ± 0.013
$All\,4\,channels\,(2D\,with\,prior)$	173.6±1.67 $(stat+JES)$	1.02 ± 0.011
$e+jets, \geqq 2tags\,(1D)$	174.8±2.17 $(stat)$	n/a
$\mu+jets, \geqq 2tags\,(1D)$	177.0±2.15 $(stat)$	n/a
$e+jets, 1tag\,(1D)$	$176.0 \pm 2.65\,(stat)$	n/a
$\mu+jets, 1tag\,(1D)$	174.8±2.81 $(stat)$	n/a
$All\,4\,channels\,(1D)$	175.7±1.19 $(stat)$	n/a

Table 7.1: Summary of m_t measurements in different channels along with the combination of all 4 channels.

7.2 Systematics

The evaluation of systematic uncertainties in the top quark mass measurement, involves accounting for insufficient understanding of signal and background modeling through various theoretical processes and the lack of complete knowledge of detector response to various objects used in reconstruction of top quark events.

The list of systematics is tabulated in Table 7.2. A detailed description of the systematics can be found in Ref.[64],[7] and is briefly summarized below.

The uncertainties corresponding to the detector response are evaluated by shifting the central values of the respective parameters viz. JES, lepton momentum scale, trigger efficiency, by ± 1 standard deviations. The top events in the MC for $m_t = 172.5\,GeV/c^2$ are

re-evaluated with the shift in these parameter values and the systematic uncertainty on the value of m_t from these shifts is defined as $\pm|m_t^+ - m_t^-|/2$, unless both shifts are in the same direction, relative to the nominal value, in which case the uncertainty is defined as $max\{|m_t^+ - m_t|/2, |m_t^- - m_t|/2\}$. When comparing between two models, the largest difference among the resulting fitted masses is taken as the systematic uncertainty.

7.2.1 Detector Response

Jet Energy Scale

The uncertainties on the global jet energy scale factor, described in Sec. 3.6, are propagated to the m_t measurement. It is found that the obatined top quark mass is shifted by $1.2\,GeV/c^2$ up ($1.2\,GeV/c^2$ down) when we vary the carrection factors by one σ up (down). This shift is also taken as the width of the prior, as defined in the previous section. For the '2D' top quark mass measurement, the corresponding uncertainty is included in the quoted (*stat+JES*) uncertainty.

Sample Dependent JES

The response of the detector to various flavors of partons (b, $light$, $gluon$) was described in Sec.3.6.1 and the corresponding uncertainties were depicted in Fig.3.9. A variation of this correction factor to the $\pm 1\,\sigma$ values does not shift the m_t or the fitted JES value.

Residual JES

The in-situ JES calibration entails only the global scale difference in JES between data and MC. The dependence of this scale factor on the p_T, η of the jets can introduce a systematic shift to the m_t measurement. To evaluate this, the energy of each jet in a $t\bar{t}$ MC sample is scaled by a factor that is parameterized as a function of p_T and η. This parameterization

corresponds to the quadratic sum of the uncertainties of the jet energy scale in data and MC. This parameterization is shown in Fig. 7.1. Re-evaluating m_t by correcting all the jets in MC (both $t\bar{t}$, W+jets) using this parametrized uncertainty as a function or the kinematic quantities, allows for determination of the systematic uncertainty. The correction factors shown in the figure are relative to the average jet transverse energy correction factor.

Jet Energy Resolution

The modeling of energy resolution for jets can affect the parton level corrections (PLC) and hence a variation in the resolution parameter by one standard deviation is studied for $t\bar{t}$ events for $m_t = 172.5\,GeV/c^2$ and the shift in the fitted value is used to assign the final systematic uncertainty in this category.

Jet Identification Efficiency

The uncertainties associated with the scale factors used to arrive at the data to Monte Carlo agreement in the jet identification efficiencies are propagated to the measurement of m_t by decreasing the efficiencies in $m_t^{generated} = 172.5\,GeV/c^2$ $t\bar{t}$ MC sample. The systematic uncertainty corresponds to one sigma deviation uncertainties associated with the scale factors used to achieve data and MC agreement in jet identification efficiencies.

JSSR Shifting

In general, the jets have better resolution in simulation than in the data. To account for this effect, simulated jets are randomly removed based on data jet identification. In addidtion, the energy scale of jets in the simulation is shifted to match the mean value of transverse momentum imbalance of jets in data and simulation. This process is called $shifting$ and the full procedure is termed as *JSSR* (jet shifting, smearing and removal, Ref.[65]). In this analysis, the shifting is turned off for the $t\bar{t}$ MC. A systematic uncertainty is assigned

toaccomodatethis effect by extracting m_t for a sample of $t\bar{t} \rightarrow l + jets$ with the mass of 172.5 GeV.

b-tagging efficiency

A discrepancy between data and MC in the modeling of the efficiency to tag b-jets can result in a systematic shift of the extracted m_t value. This effect is evaluated by varying the tag rate functions (TRF) and mistag rate function for light quarks by 5% Ref.[66] and 20% respectively. Redoing the ensemble tests for the $t\bar{t}$ MC with $m_t = 172.5\,GeV/c^2$ shifts the extracted top mass by 0.08 GeV, as quoted in Ref.[64]. The TRFs used in this analysis are a part of the top mass extraction algorithm and the only efficiency of the b-jets that could affect the outcome of this analysis is that of the Neural Network b-tagging output that is used to split the sample in one tag and greater than two tag channels, as descibed in Sec. 4.4.1. Since we combine the channels to obtain the final result, the effect of any small change of number of events in individual channels would only be a second order effect to the assigned systematic uncertainty in this category.

Trigger Efficiency

The Monte Carlo events used in this analysis are weighted to account for the trigger efficiency in accordance with the data. To evaluate the systematics effect in the measurement of top mass, all the event weights corresponding to trigger are set to unity and the ensemble tests are repeated to observe any shift in the extracted mass. This shift is found to be $0.05\,GeV/c^2$ and is tabulated in the Table 7.2.

Lepton Momentum Scale efficiency

Differences in lepton momentum scale in data and MC can introduce systematic shift in m_t measurement. To evaluate this, the peak of dilepton invariant mass in $J/\psi \rightarrow ll$ and $Z \rightarrow ll$

decays are compared between data and MC and an absolute momentum scale for electrons and muons is determined. A fit to the two mass points of 3.09 GeV/c^2 and 91.18 GeV/c^2 (corresponding to the J/ψ and Z invariant masses) is performed as a function of the mean value of the transverse momentum of the leptons. The m_t is measured for $t\bar{t}$ MC with a generator level mass of 172.5 GeV/c^2 without the rescaling of lepton p_T and with lepton p_T values rescaled using the fit described above. Half of the largest difference in extracted m_t is taken as its systematic uncertainty in this category.

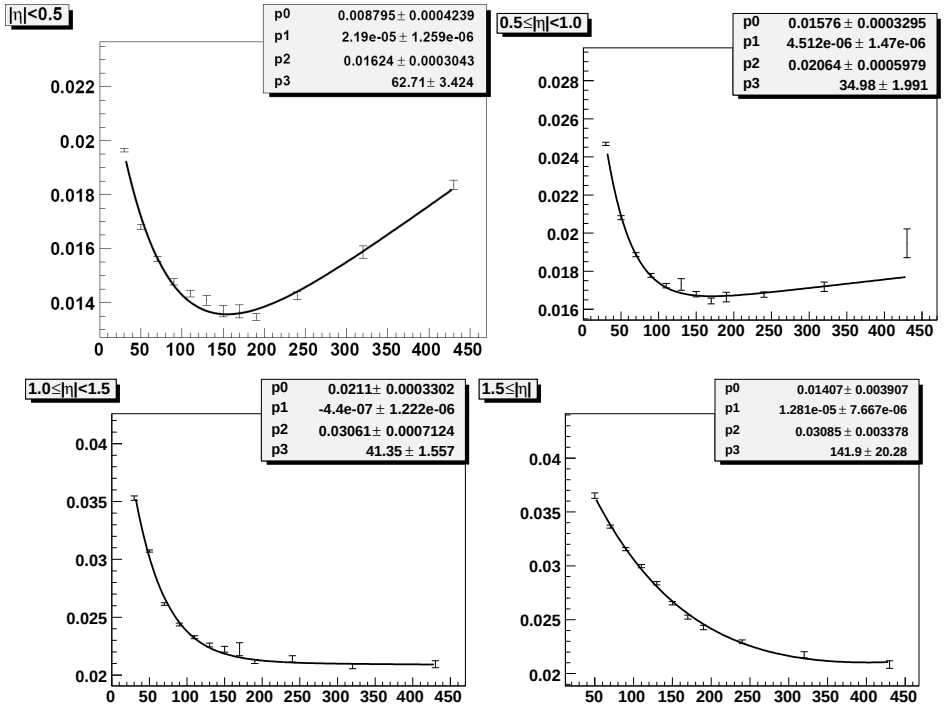


Figure 7.1: Residual JES parametrization as a function of p_T of jets in various $|\eta^{det}|$ for the $t\bar{t}$ MC.

7.2.2 Production Systematics

7.2.2.1 Signal Modeling

ISR/FSR

The incoming and outgoing partons may give rise to additional jets which are referred to as the initial and final state radiation. These jets may well be misidentified as the product of $t\bar{t}$ decay and hence bias the reconstruction of the event. This effect is studied using Drell Yan dilepton events which are produced from a $q\bar{q}$ initial state similar to $t\bar{t}$ production (Ref.[67]). The p_T of dilepton pairs as a function of their invariant mass is compared between data and MC and the best value for the hadronization scale (Λ_{QCD}) and the virtuality scale (Q^2), extrapolated to $t\bar{t}$ mass region, is extracted. The values obtained for $\pm 1\,\sigma$ deviations around the mean lepton pair p_T are used to evaluate the ISR systematics. The QCD evolution equation is the same for both sets of scales and hence the above obtained sets of Λ_{QCD} and Q^2 are used for FSR systematics too. This amounts to $\pm 0.65\,GeV/c^2$ uncertainty on the extracted m_t when the ensemble tests are done without the prior.

Higher Order QCD effects

The m_t value is extracted for two models of the hard scattering and shower evolution process. ALPGEN interfaced with HERWIG (for parton showering) is compared with MC@NLO Ref.[68] interfaced with HERWIG. This includes the effects of $q\bar{q}/gg$ fraction in the initial state. The JES^{fitted} for the full two dimensional fit yields 1.02 for ALPGEN+HERWIG and 1.05 for MC@NLO+HERWIG , corresponding to a systematic uncertainty of $2.6\,GeV/c^2$ in the extracted m_t.

Hadronization and UE effects

Keeping the same set of hard-scattering processes through ALPGEN, two different models of hadronization are interfaced with ALPGEN. These are HERWIG and PYTHIA. The difference in the extracted m_t for these two models yields systematic uncertainty of $\pm 1.7\,GeV/c^2$ for the '2D' measurement.

Color reconnection

The colored objects (quarks), which are produced as the decay product of $t\bar{t}$, besides the initial and final state radiations interact with each other and also with the remnants of the $p\bar{p}$ collisions. This can change the kinematic and topological distributions of jets and the systematic uncertainty corresponding to this effect is termed color reconnection. To evaluate this, $t\bar{t}$ MC is generated with two PYTHIA tunes that involve angular ordering of showers (similar to the nominal sample). These tunes are APRO Ref.[69] and ACR-PRO. The difference in the extracted m_t provides the uncertainty of $\pm 0.7\,GeV/c^2$ for the '2D' measurement without a prior.

Multiple $p\bar{p}$ interactions

Luminosity effects from additional $p\bar{p}$ interactions are simulated by overlaying on MC events, the unbiased triggers from random $p\bar{p}$ crossings. These overlaid events are then reweighted to the number of interaction vertices, to assure that the simulation reflects the instantaneous luminosity profile of the data. The ensemble studies are repeated and m_t is extracted to find a shift of $\pm 0.07\,GeV/c^2$, as quoted in Ref.[64].

Parton distribution functions

$t\bar{t}$ MC generated with PYTHIA for a mass of 172.5 GeV is reweighted to match possible excursions in the PDF parameters, represented by two sets of 20 eigenvectors of CTEQ6M uncertainty PDFs Ref.[11]. Ensemble studies are repeated for each of the variants and uncertainties are added in quadrature as follows :

$$\delta m_t^{PDF} = \frac{1}{2}\left(\sum_{i=1}^{20}\{\Delta M(S_i^+) - \Delta M(S_i^-)\}^2\right)^{1/2} \tag{7.1}$$

The sum runs over PDF excursions in the positive (S_i^+) and negative (S_i^-) directions.

7.2.2.2 Background Modeling

W+jets heavy flavor scale factor

The normalization of the leading order (LO) ALPGEN MC for the W+jets sample is increased by a factor of 1.47 for W+heavy flavor jets contribution to provide agreement with higher order (NLO) calculation of cross-sections, that include NLL corrections based on the MCFM MC generator Ref.[57]. This is done before normalizing the backgrounds to the data. To account for the systematics due to this scaling, the analysis is repeated by taking the scale factors as 1.97 and 0.97. The difference in the extracted m_t with the nominal is used as systematic uncertainty on the measurement.

Modeling of b-quark

The default modeling of b-quark fragmentation used in the $t\bar{t}$ MC corresponds to that of PYTHIA, which is based on the Bowler scheme Ref.[70], whereby the fragmentation parameters are tuned to the LEP e^+e^- data Ref.[71]. To assess the systematics, the events are reweighted to account for the differences between LEP and SLAC data Ref.[72]. The

ensemble studies are repeated and the difference in the extracted m_t is determined. Also the b-jets calorimeter response is different in jets that decay semileptonically than for those that don't. The effect on m_t due to this was studied in Ref.[72] and was determined to be $\pm 0.05\, GeV/c^2$. This uncertainty is added in quadrature to the one, derived for fragmentation function, mentioned above. The net uncertainty is tabulated in Table7.2.

Factorization and Renormalization scales

The W+jets MC samples used in this analysis are generated by employing the identical factorization and renormalization scales such that $Q^2 = M_W^2 + \sum p_T^2$, where the sum runs over all the jets in an event. To determine the uncertainty on m_t due to the uncertainty on this scale, W+jets samples are generated with the scale variation of $(Q^2/2)$ and $(2Q^2)$. This affects the transverse momentum distribution of jets while keeping the normalization factor intact. The $\mathscr{BG}$ shapes are re-evaluated and the likelihood re-determined for the ensemble studies. Half of the largest excursion in the extracted m_t is taken as the systematic uncertainty.

7.2.3 Method

MC calibration

The statistical uncertainties associated with the slope and offset of the final calibration curve is shown in Fig.6.22. The 68% confidence interval (half of the width of the band shown in the Fig. 6.22)is used to assign the systematic uncertainty on the extracted m_t at 172.5 GeV.

Multjet contamination

Multijet events are not used in deriving the calibrations. The effect on calibration is studied by selecting a multijet-enriched sample of events from data (obtained by using the lepton reverse isolation criteria in selection). The ensemble studies were repeated and the uncertainty was found to be $\pm 0.14\, GeV/c^2$ in the extracted m_t as shown in Ref.[73].

Systematics Source	*Uncertainty (2D)* GeV/c^2	*Uncertainty (1D)* GeV/c^2	*Uncertainty (2D with prior)* GeV/c^2
DETECTOR RESPONSE			
Jet Energy Scale (JES)	n/a	± 1.2	n/a
Sample Dependent JES	0	0	0
Residual JES	± 0.2	0	± 0.2
Jet Energy Resolution	± 0.4	± 0.75	0
Jet Identification efficiency	± 0.1	± 0.1	± 0.1
JSSR shifting	± 0.3	± 0.2	0
b-tagging efficiency	± 0.08	± 0.08	± 0.08
Trigger efficiency	± 0.05	± 0.05	± 0.05
Lepton scale efficiency	± 0.1	± 0.1	± 0.1
PRODUCTION PROCESS			
Signal Modeling :			
ISR/FSR	± 0.65	± 0.25	± 0.35
Higher order effects	± 2.6	± 0.6	± 1.4
Hadronization, UE effects	± 1.7	± 0.1	± 1.0
Color Reconnection	± 0.7	± 0.1	± 0.4
Multiple $p\bar{p}$ interactions	± 0.07	± 0.07	± 0.07
Choice of PDF	± 0.35	± 0.12	± 0.35
Background Modeling :			
W+jets heavy flavor scale factor	± 0.05	± 0.05	± 0.05
b-quark fragmentation modeling	± 0.1	0	± 0.1
Factorization & renormalization scales	± 0.05	± 0.05	± 0.05
METHOD			
MC calibration	± 0.12	± 0.09	± 0.12
Multijet contamination	± 0.1	± 0.1	± 0.1
TOTAL	± 3.32	± 1.59	± 1.86

Table 7.2: Various sources of systematic uncertainties to the top quark mass measurement.

To summarize, the measurement of the top quark mass using a full two dimensional likelihood fit with in-situ JES yields :

$$m_t = 172.6 \pm 1.75\,(stat + JES) \pm 3.32\,(sys)\ GeV/c^2 = 172.6 \pm 3.75\ GeV/c^2 \qquad (7.2)$$

Constraining the measured jet energy scale to the MC jet energy scale within a known uncertainty, by incorporating a *prior* yields :

$$m_t = 173.6 \pm 1.67\,(stat + JES) \pm 1.86\,(sys)\ GeV/c^2 = 173.6 \pm 2.49\ GeV/c^2 \qquad (7.3)$$

Using only the apriori JES from $\gamma +$ *jets* events, the one dimensional measurement yields :

$$m_t = 175.7 \pm 1.19\,(stat) \pm 1.2\,(JES) \pm 1.04\,(sys)\ GeV/c^2 = 175.7 \pm 1.98\ GeV/c^2 \quad (7.4)$$

All the above measurements are consistent with each other, within the assigned uncertainties and emphasize the strong dependence of the precision of top quark mass measurement on the hadronic jet energy scale. Radiation and higher order effects dominate the systematic uncertainty on the '2D' measurement.

CHAPTER **Eight**

Conclusions and Perspectives

We have measured the top quark mass using 4.3 fb^{-1} of RunII data collected at DØ . The dominant systematic uncertainty on the top quark mass measurements done till date have been attributed to the jet energy scale. Hence, two different approaches with complimentary techniques for jet energy scale were used for measuring the top quark mass in this dissertation. First, a simultaneous fit of the top quark mass with the jet energy scale is performed and the measurement yields a top quark mass value of **172.6** GeV/c^2 with an associated uncertainty of 2.17 %. In the second approach, the top quark mass is measured using the *apriori* jet energy scale calibration derived from $\gamma + jets$ events. This method yields a value of the top quark mass to be 175.7 GeV/c^2. The associated uncertainty obtained via this approach is 1.12 %. It is thus found that the simultaneous measurement is not as precise as the latter because of the important systematic uncertainties that affect the jet energy scale measurement and thus the top quark mass measurement. So we try a third approach which tries to combine the two techniques by applying a jet energy scale prior to the simultaneous measurement. This approach yields a value of top quark mass to be **173.6** GeV/c^2 with an associated uncertainty of 1.43 %. This approach only partially mediates the effect and the measurement is still dominated with significant systematic uncertainties

in the jet energy scale. Hence the precision on the value of top quark mass is best obtainted with apriori jet energy scale. Comparing the measurement of the most precise value of top quark mass obtained in this dissertation (Eq. 7.4) with the most recent Tevatron combination for the top quark mass, which is reportedly found to be 173.18 GeV/c^2 with an associated uncertainty of 0.54% (Ref. [7]), it can be concluded that that they are compatible with each other. A summary of measurements of top quark mass corresponding to various run periods and different channels is shown in Fig. 8.1. A direct comparison of the measurements in this dissertation with those of Fig. 8.1 is not completely trivial due to the presence of various correlations (statistical and systematics related).

Top quark mass interpretation

It is still evident that the '2D' measurement (Eq. 7.2) does not correspond to the best precision with systematic uncertainties corresponding to gluon radiation and higher order effects dominating the jet energy scale uncertainty and hence the top quark mass measurement.This brings us to the topic of interpretation of top quark mass. For a quark, the mass parameter is introduced in the QCD Lagrangian whose value depends on the renormalization scheme and the renormalization scale. At high energies, the QCD coupling constant (α_s) is small and the observables are typically calculated in perturbation theory, commonly applying $\overline{MS}$ renormalization scheme (Ref. [74]) . For an observed, non-colored particle, the position of the pole in the propagator defines the mass. However, the pole mass cannot be determined to arbitrary accuracy owing to the non-perturbative effects (long distance physics a.k.a. confinement) which is limited by the hadronization scale Λ_{QCD}. The relation between the $\overline{MS}$ mass and the pole mass is known to three loops (Ref. [75]). In principle, it is possible to absorb higher order corrections into the pole mass definition via line shape with constant width. This has been done for extracting the Z boson mass (Ref. [76]).

It is also to be noted that the top quark mass measurements rely on the comparison of the data with the simulated events. The simulation used in this analysis to obtain the calibration is ALPGEN interfaced with PYTHIA. ALPGEN has fixed width in the quark propagator and PYTHIA uses a factor dependent on α_s to approximate the loop corrections. In this way a simulation of $t\bar{t}+nlp,\ n=0,1,2$ is obtained. It has been argued that the Monte Carlo mass (m_t^{MC}) is closer to the pole mass within 1 GeV/c^2 (Ref. [77, 78]). Mass extraction for the inclusive $t\bar{t}+X$ cross section has been reported recently by DØ (Ref. [1]). Fig. 8.2 shows the dependence of $\sigma_{t\bar{t}}$ on the top quark pole mass in various higher order schemes and soft gluon resummations. This is also tabulated in Table 8.1. Although the mass from extracted cross section would never be as precise as that of the direct measurements, the '2D' measurement (Eq. 7.2) is compatible with all of the extracted masses shown in the table and is closest to the Approximate NNLO scheme (Ref. [13, 79]).

Top mass measurements at hadron colliders are reaching the level of theoretical uncertainties. In future, the top quark mass measurement at the $t\bar{t}$ threshold including bound state effects (Coulomb summations, Ref. [80]) would be a subject matter of study. Also, it would be desirable to determine the top quark mass that is not limited by the uncertainty on the pole mass. This could possibly be achieved by performing $t\bar{t}$ threshold scans and doing measurements of the $\sigma_{t\bar{t}}$ near threshold at a future e^+e^- collider. The top quark mass parameter measured this way could be translated to the $\overline{MS}$ mass with a smaller uncertainty due to a better knowledge of the initial state. This has been discussed in literature (e.g. Ref. [81] and references therein) and the technique is similar to the determination of W boson mass from WW cross-section, which was performed at LEP 2.

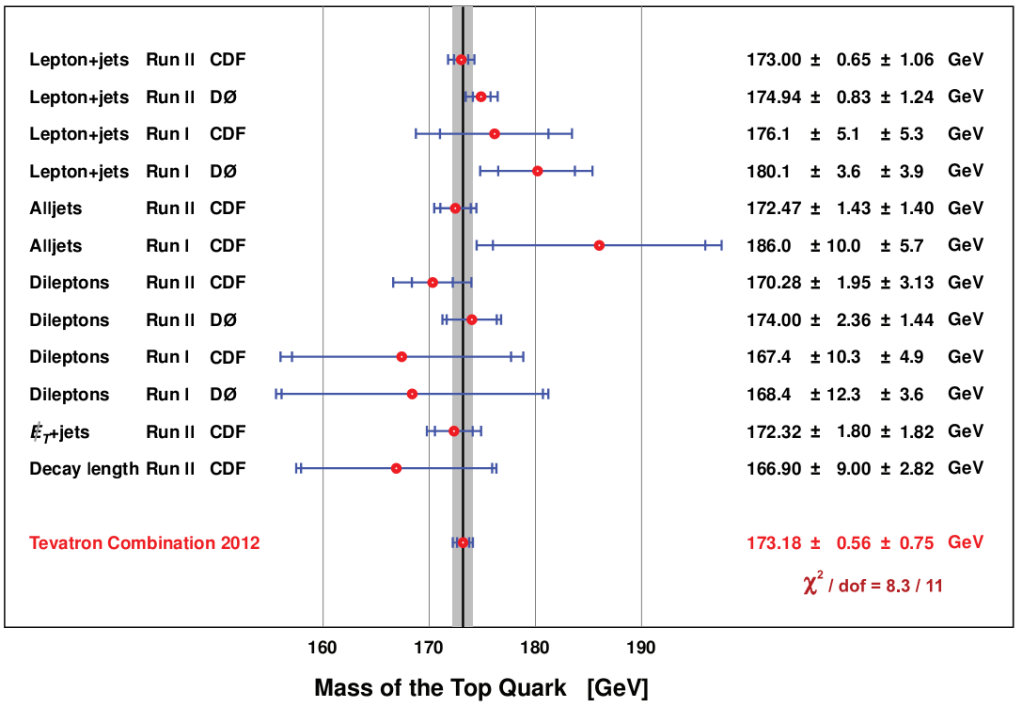


Figure 8.1: The twelve input measurements of the top quark mass from the Tevatron collider experiments along with the resulting combined value. The grey region corresponds to $\pm 0.94\,GeV/c^2$. (Ref. [7])

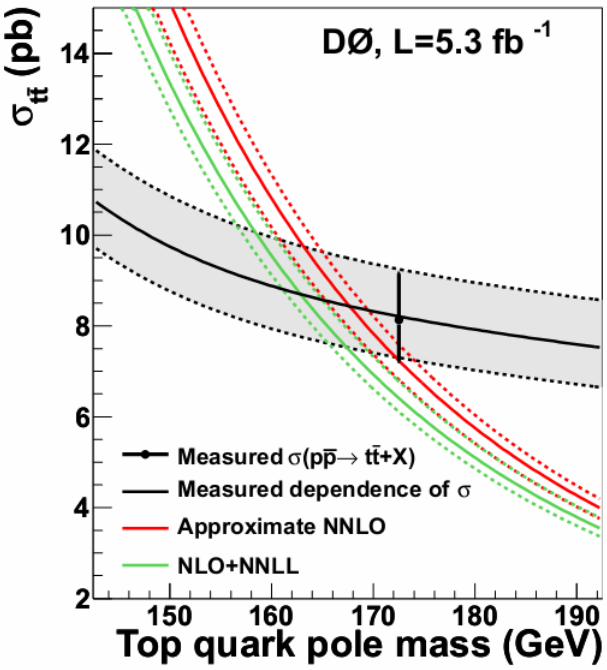


Figure 8.2: Measured $\sigma_{t\bar{t}}$ and theoretical NLO+NNLL and aproximate NNLO calculations of $\sigma_{t\bar{t}}$ as a function of top quark pole mass, assuming that the mass of top quark in simulation is equal to the pole mass. (Ref. [1])

Theoretical Prediction	$m_t^{pole}(GeV/c^2)$, given $m_t^{MC} = m_t^{pole}$
NLO [82]	$164.8^{+5.7}_{-5.4}$
NLO+NLL [83]	$166.5^{+5.5}_{-4.8}$
NLO+NNLL [84]	$163.0^{+5.1}_{-4.6}$
Approximate NNLO [13]	$167.5^{+5.2}_{-4.7}$
Approximate NNLO [79]	$166.7^{+5.2}_{-4.5}$

Table 8.1: Values of m_t^{pole}, with their 68% confidence interval uncertainties, extracted for different predictions of $\sigma_{t\bar{t}}$. The result assumes that the top quark mass in the simulation is equal to the pole mass of the top quark propagator. (Ref. [1])

APPENDIX **A**

Parton Energy distributions for light quarks

The Parton Level Corrections and corresponding resolution functions for light quarks are obtained by using matched jet-parton pairs for (u,d,c,s) quarks in $t\bar{t} \rightarrow l + jets$ events. The figures below show the parton energy distribution in various jet energy bins for 4 detector regions. The distributions are close to Gaussian and are fitted with the Gaussian function near the peak. The mean value gives the PLC, while the RMS of the Gaussian provides with the parton energy resolution for that particular jet energy bin.

- Region 1

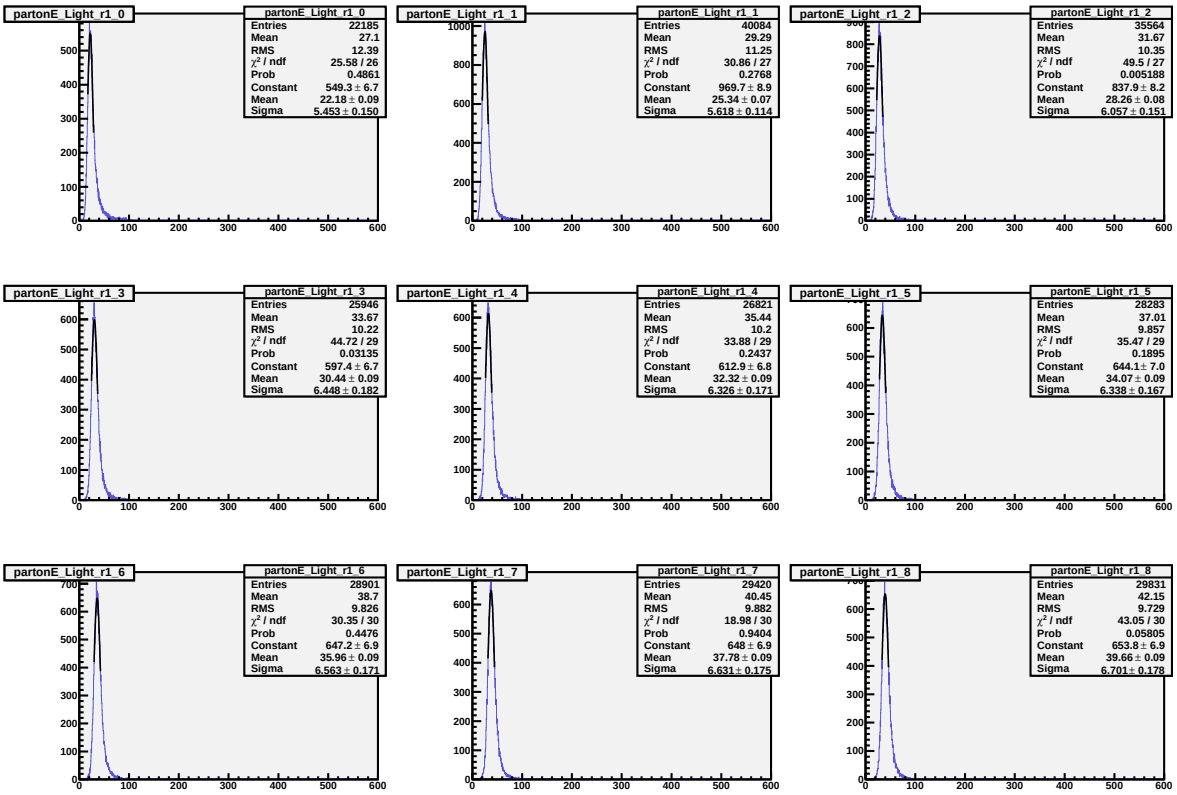


Figure A.1: E_{parton} distribution for light quarks $\{|\eta| \in [0,0.5)\}$ in E_{jet} bins with bin boundaries $0, 19, 23, 26, 28, 30, 32, 34, 36, 38$ *GeV*

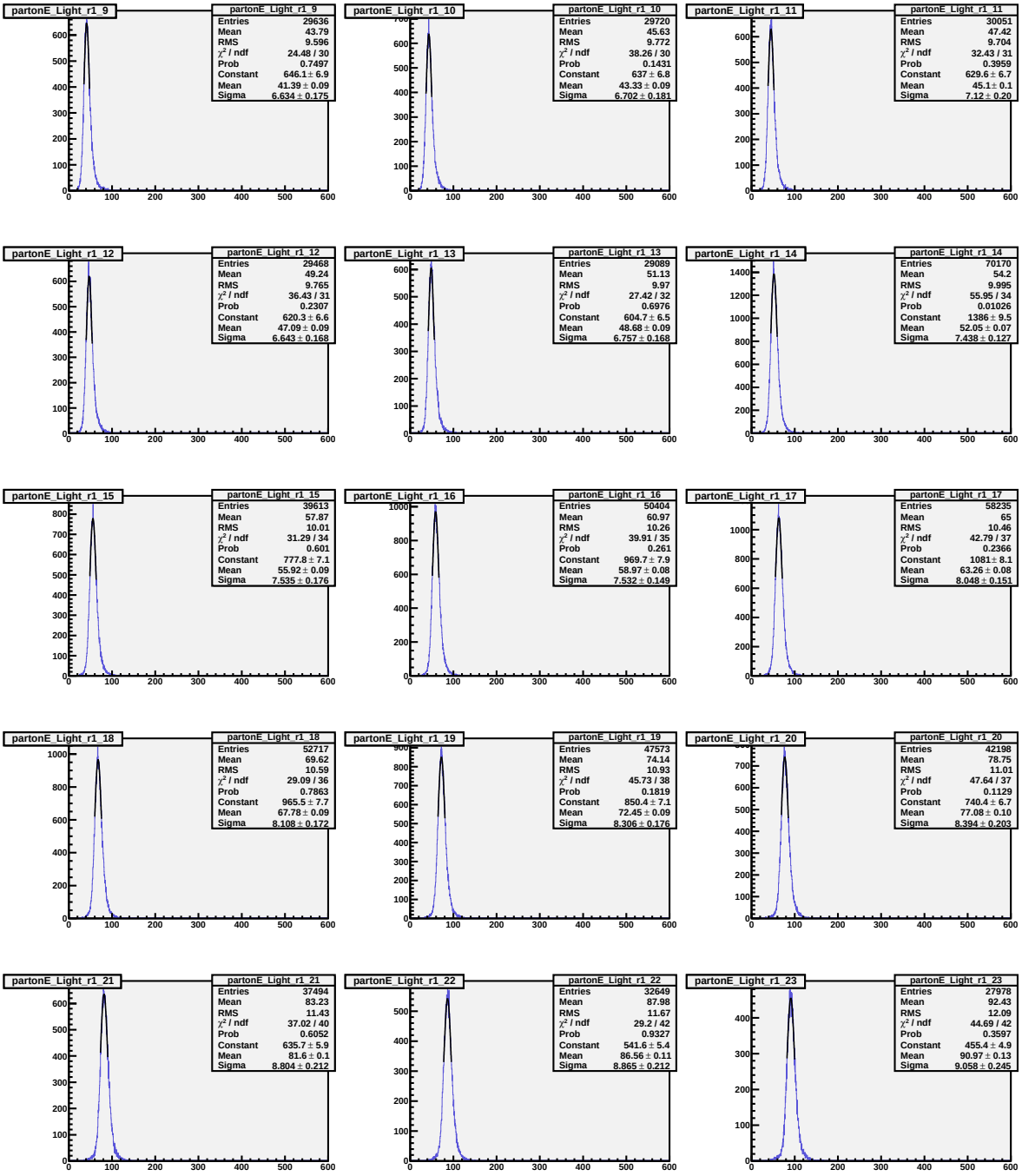


Figure A.2: E_{parton} distribution for light quarks $\{|\eta| \in [0,0.5)\}$ in E_{jet} bins with bin boundaries $38,40,42,44,46,48,53,56,60,65,70,75,80,85,90,95$ *GeV*

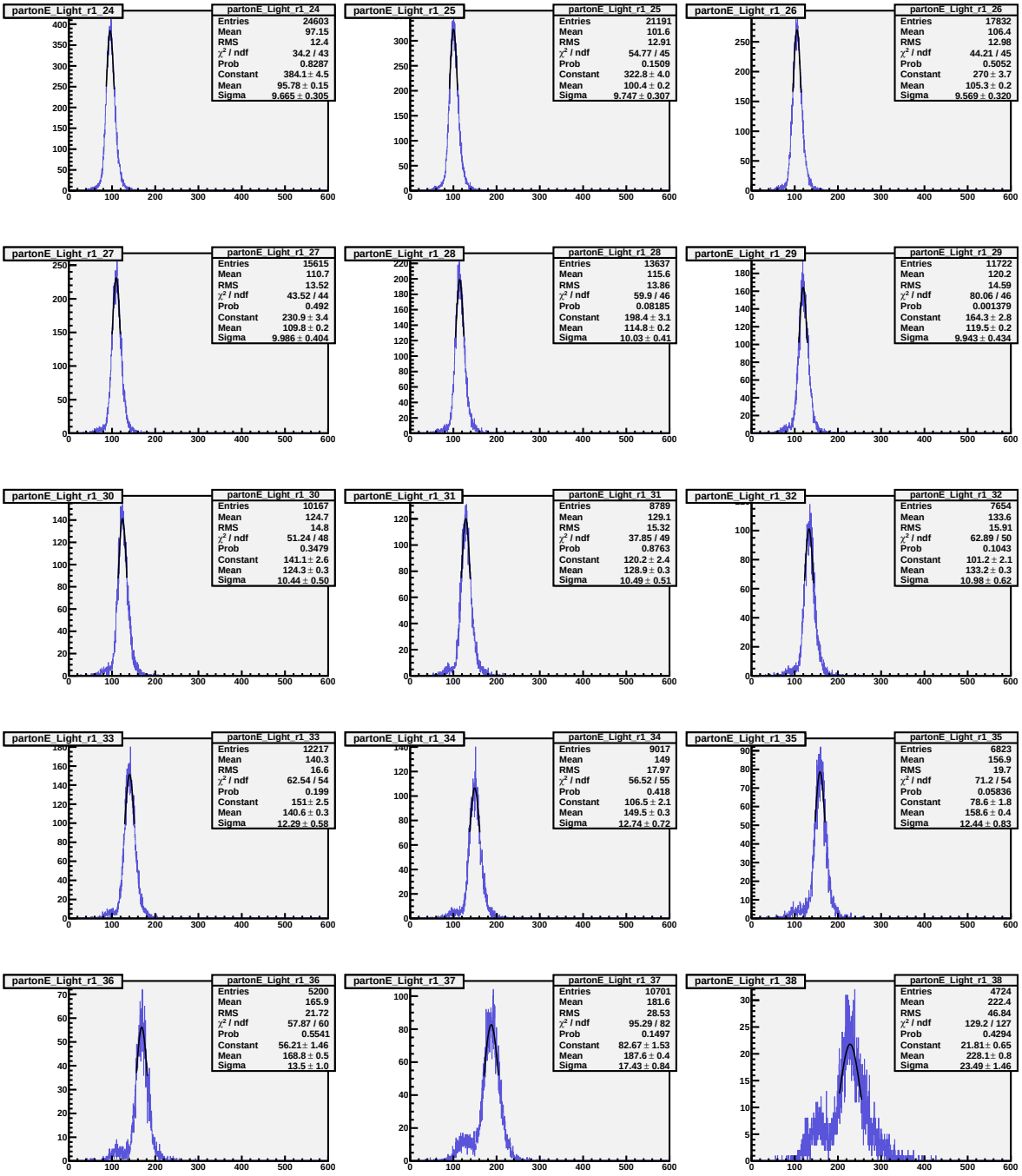


Figure A.3: E_{parton} distribution for light quarks $\{|\eta| \in [0, 0.5)\}$ in E_{jet} bins with bin boundaries $95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 150, 160, 170, 180, 220, 400$ *GeV*

- Region 2

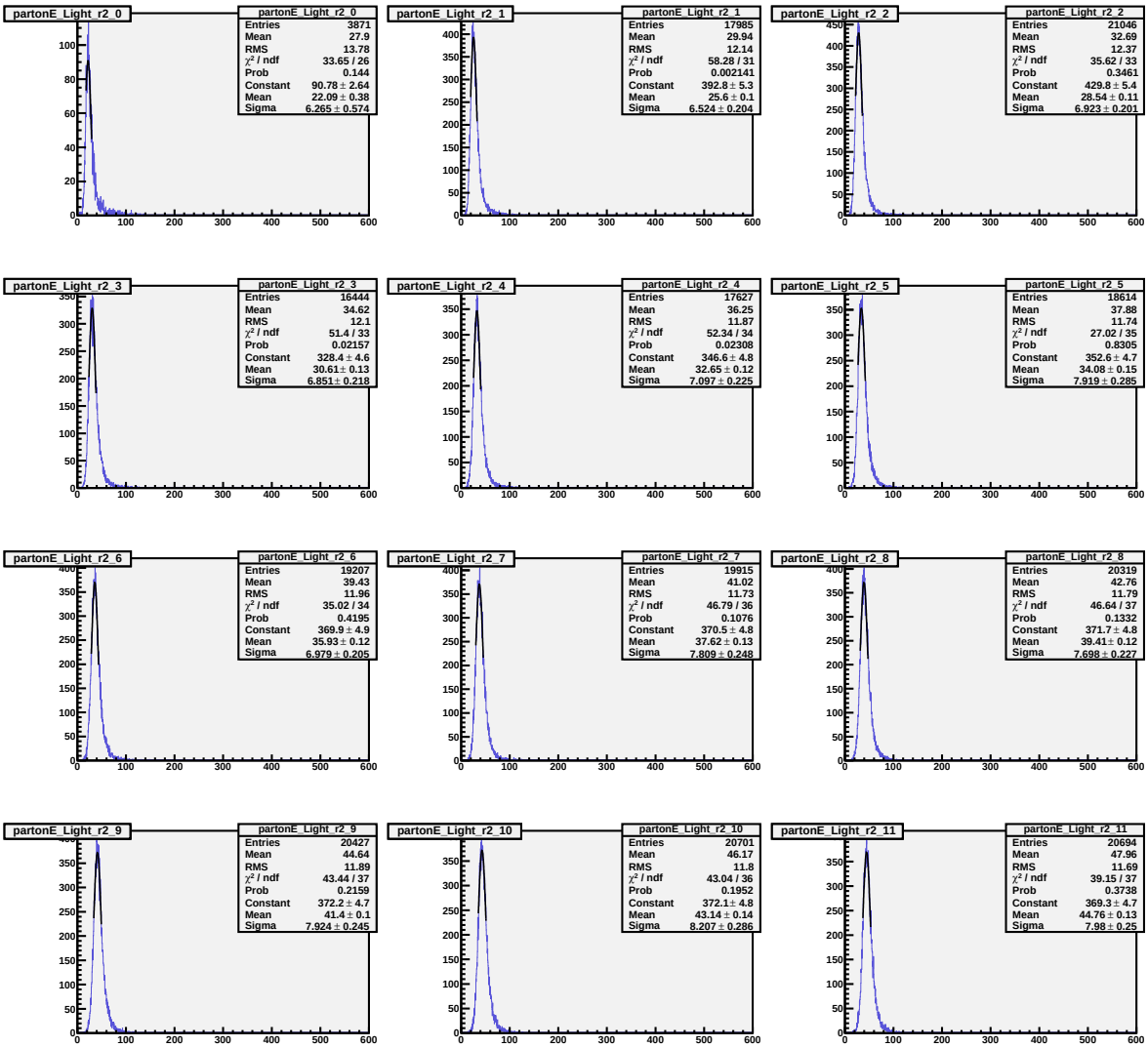


Figure A.4: E_{parton} distribution for light quarks $\{|\eta| \in [0.5, 1.0)\}$ in E_{jet} bins with bin boundaries $0, 19, 23, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44$ *GeV*

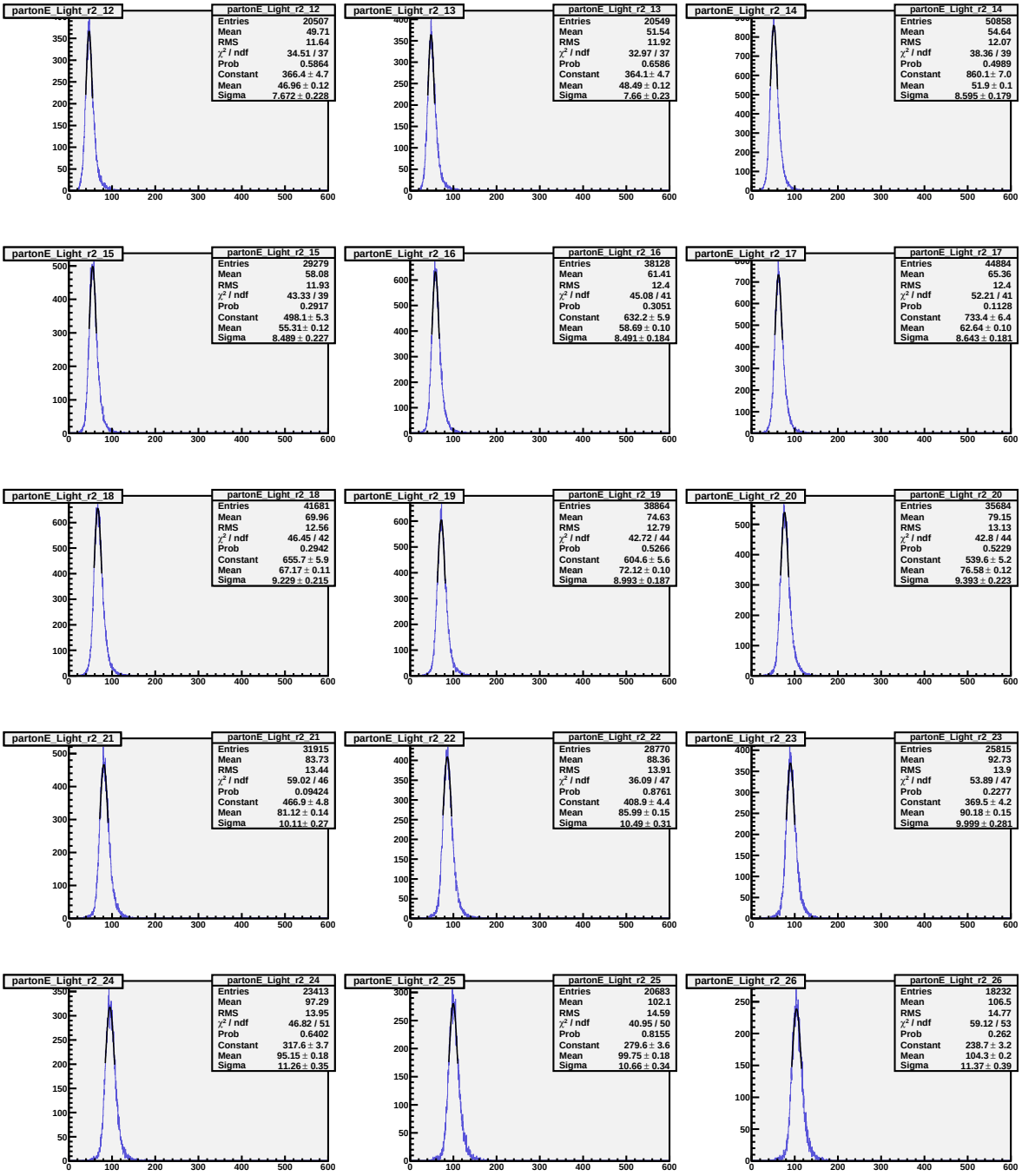


Figure A.5: E_{parton} distribution for light quarks $\{|\eta| \in [0.5, 1.0)\}$ in E_{jet} bins with bin boundaries $44, 46, 48, 53, 56, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110$ *GeV*

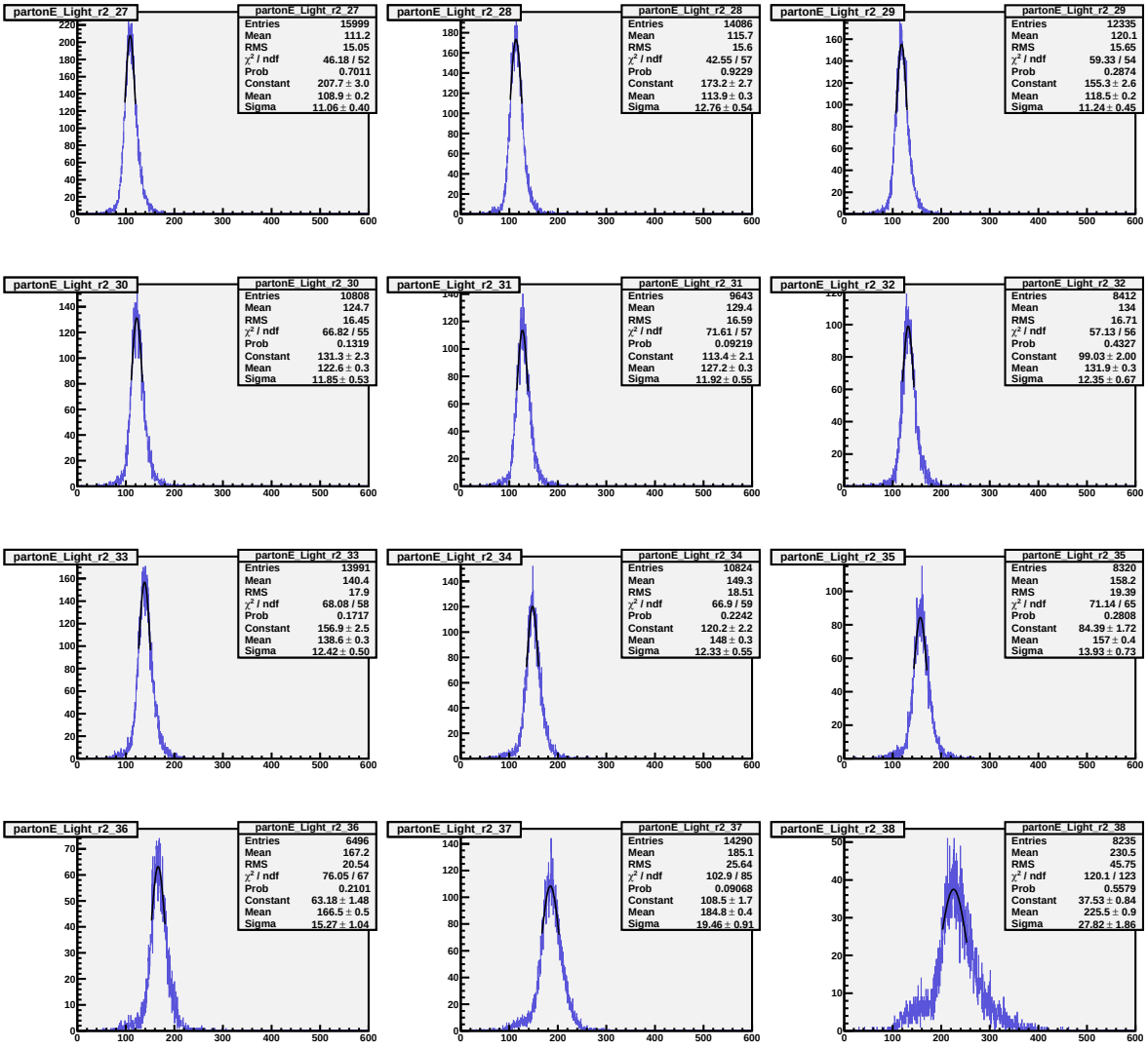


Figure A.6: E_{parton} distribution for light quarks $\{|\eta| \in [0.5, 1.0)\}$ in E_{jet} bins with bin boundaries $110, 115, 120, 125, 130, 135, 140, 150, 160, 170, 180, 220, 400$ *GeV*

- Region 3

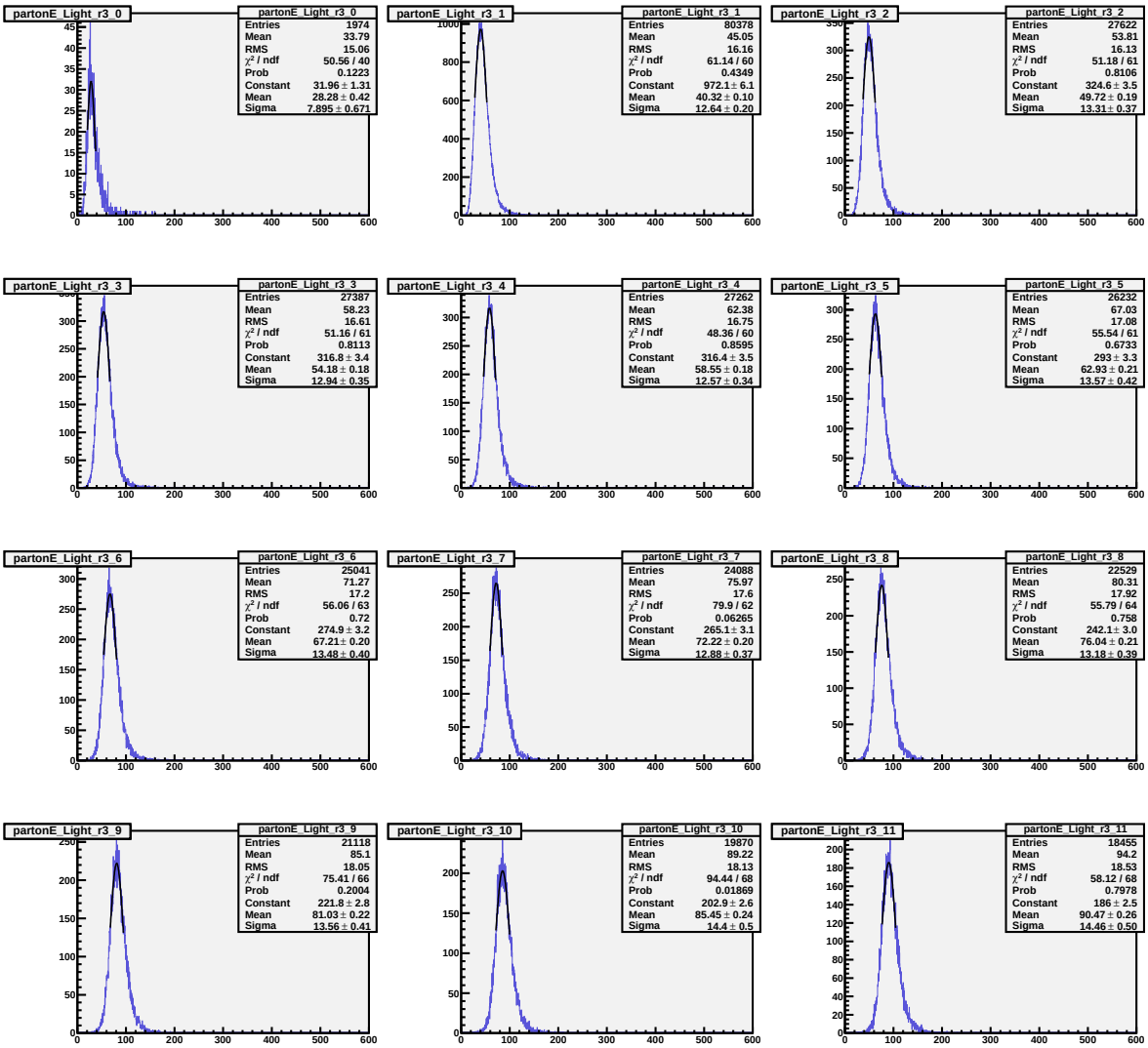


Figure A.7: E_{parton} distribution for light quarks $\{|\eta| \in [1.0, 1.5)\}$ in E_{jet} bins with bin boundaries $0, 25, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95$ *GeV*

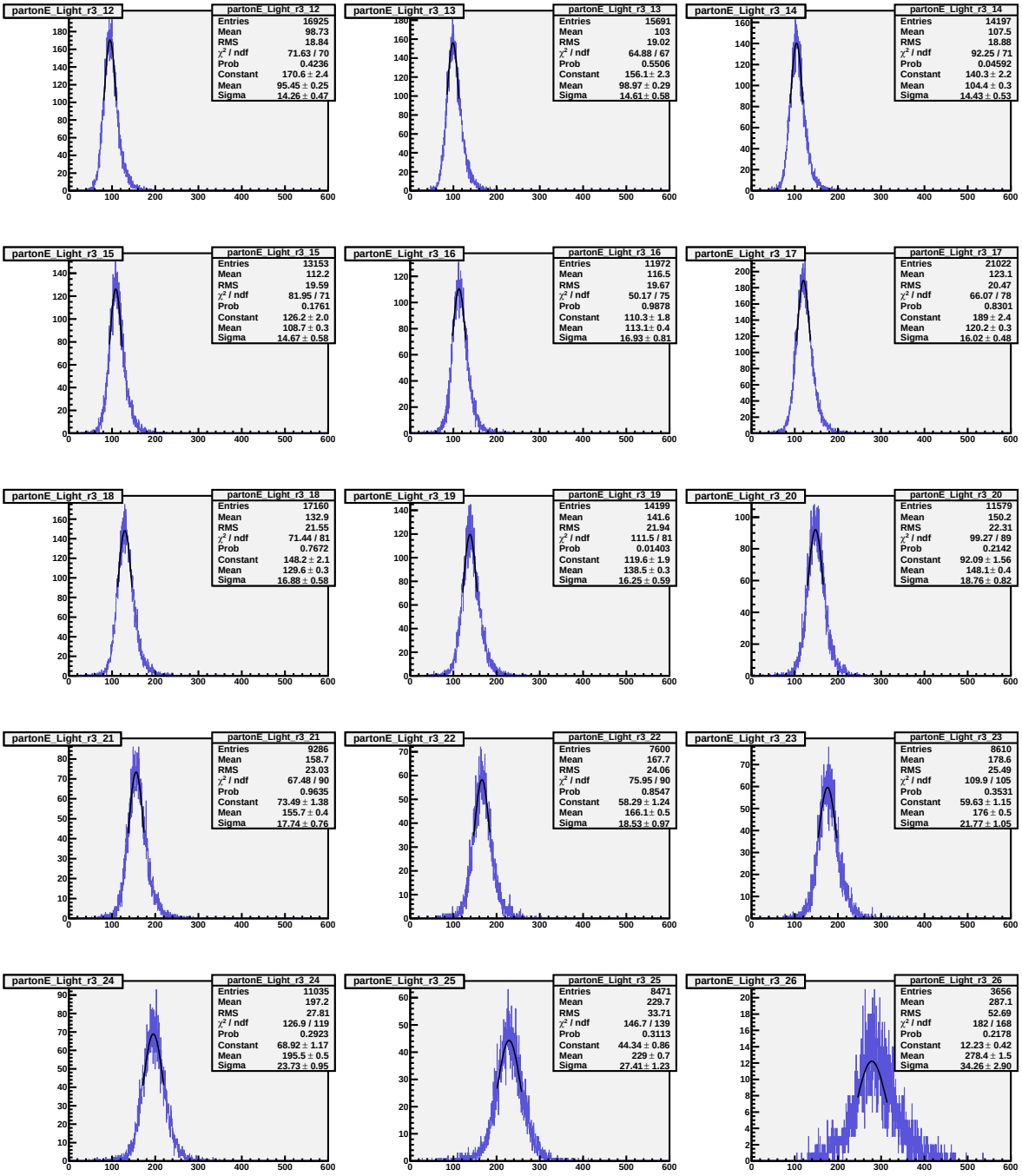


Figure A.8: E_{parton} distribution for light quarks $\{|\eta| \in [1.0, 1.5)\}$ in E_{jet} bins with bin boundaries $95, 100, 105, 110, 115, 120, 130, 140, 150, 160, 170, 180, 195, 225, 280, 500$ *GeV*

- Region 4

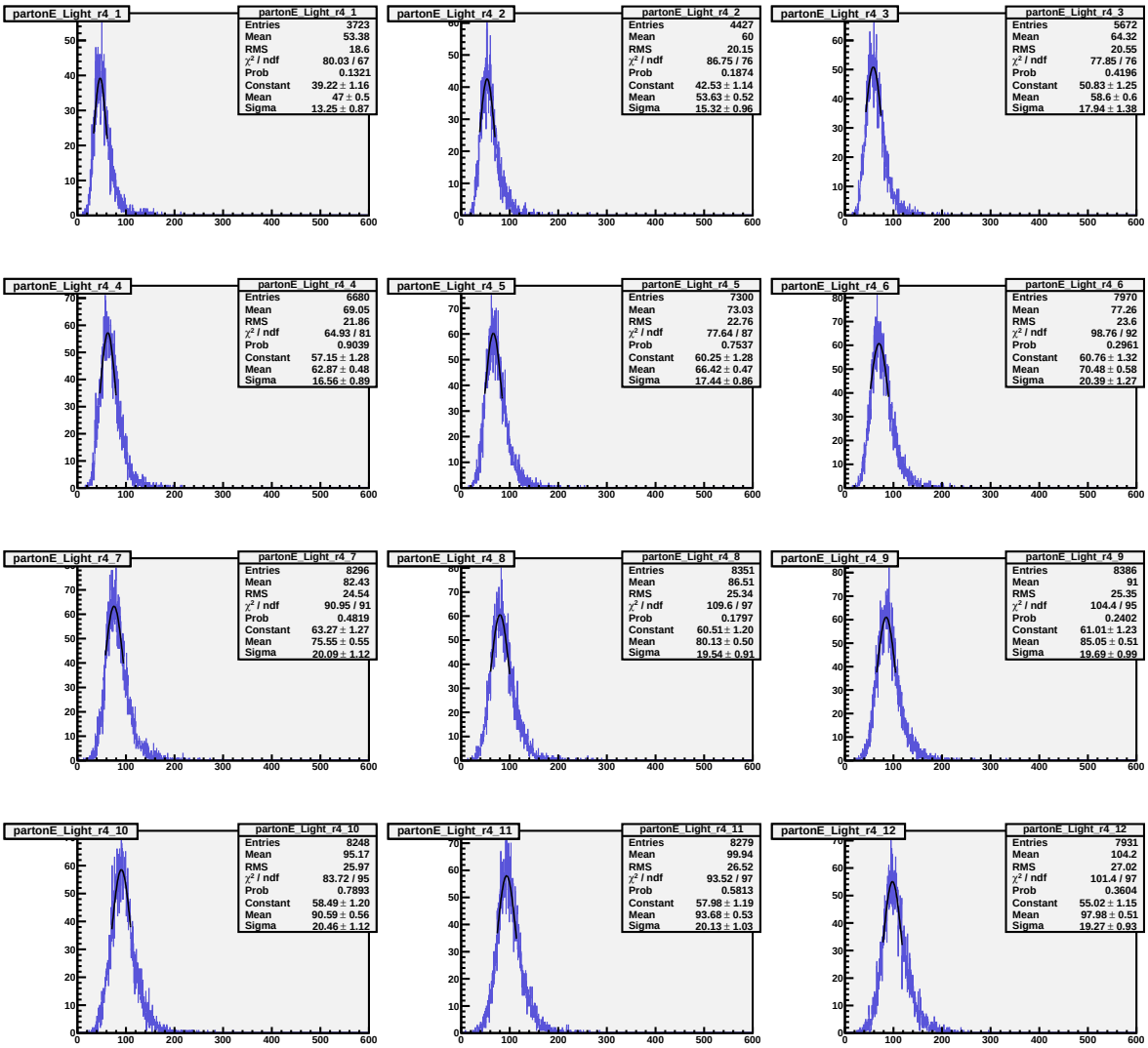


Figure A.9: E_{parton} distribution for light quarks $\{|\eta| \in [1.5, 2.5)\}$ in E_{jet} bins with bin boundaries $25, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100\ GeV$

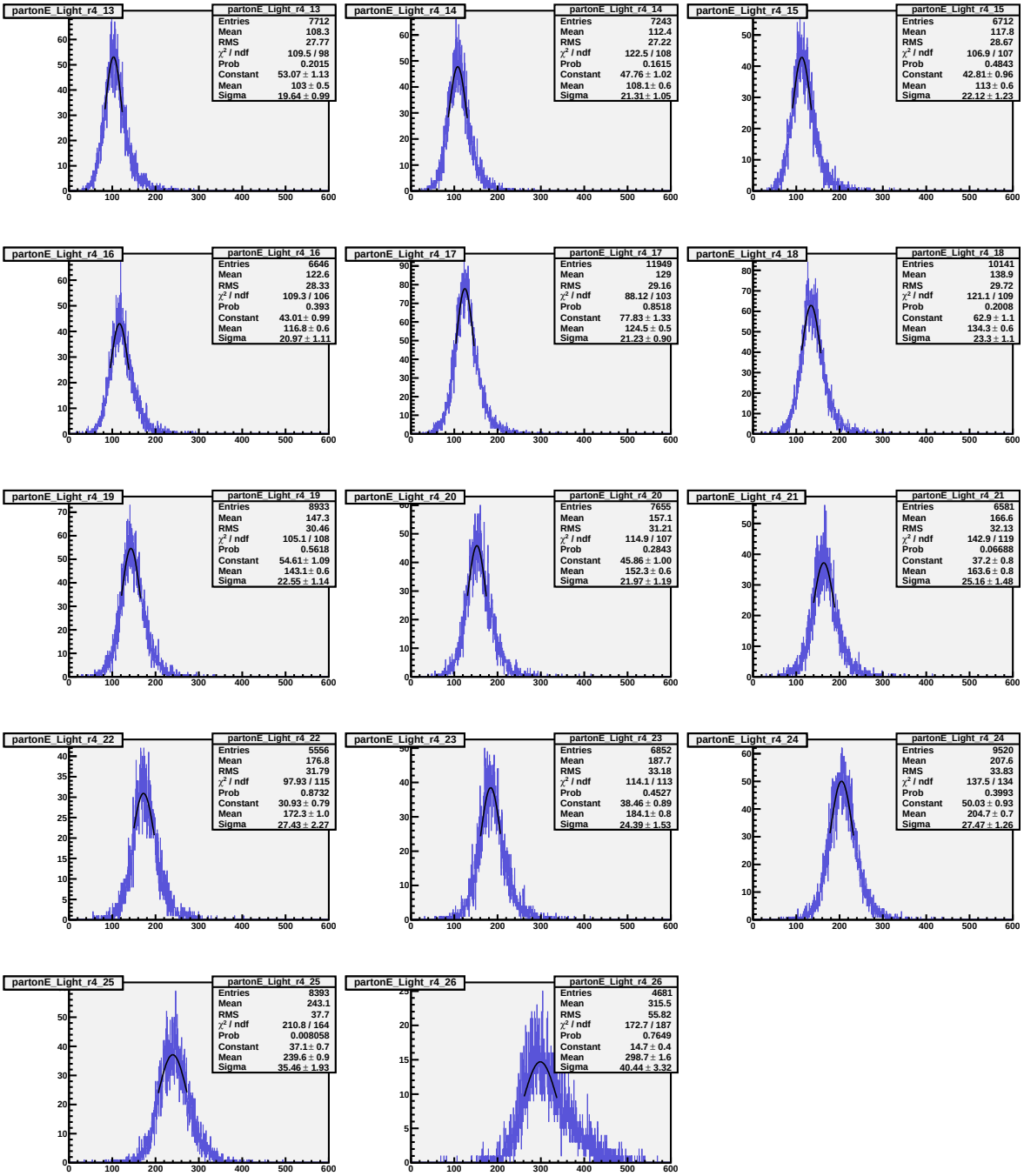


Figure A.10: E_{parton} distribution for light quarks $\{|\eta| \in [1.5, 2.5)\}$ in E_{jet} bins with bin boundaries $100, 105, 110, 115, 120, 130, 140, 150, 160, 170, 180, 195, 225, 280, 500$ *GeV*

APPENDIX **B**

Parton Energy distributions for b-quarks

The Parton Level Corrections and corresponding resolution functions for $b-$quarks are obtained by using matched jet-parton pairs for $b-$quarks in $t\bar{t} \rightarrow l + jets$ events. The figures below show the parton energy distribution in various jet energy bins for 4 detector regions. The distributions are close to Gaussian and are fitted with the Gaussian function near the peak. The mean value gives the PLC, while the RMS of the Gaussian provides with the parton energy resolution for that particular jet energy bin.

- Region 1

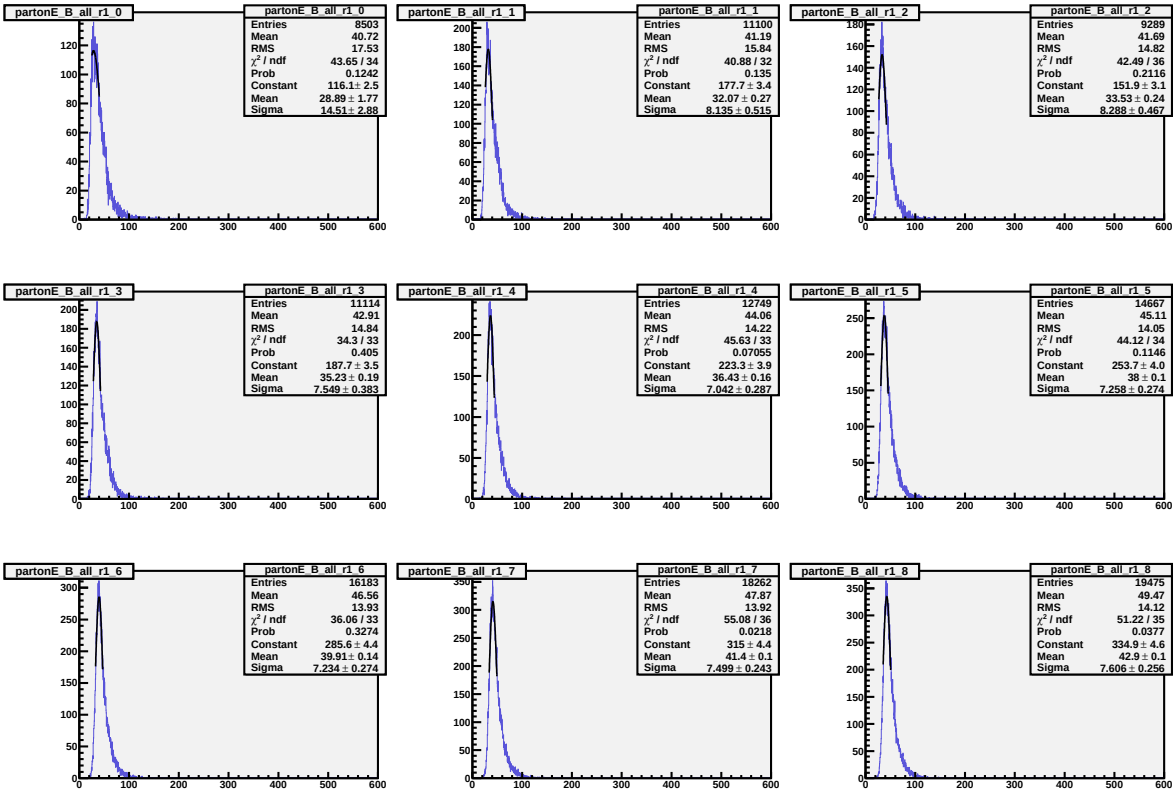


Figure B.1: E_{parton} distribution for b-quarks $\{|\eta| \in [0, 0.5)\}$ in E_{jet} bins with bin boundaries $0, 19, 22, 24, 26, 28, 30, 32, 34, 36\ GeV$

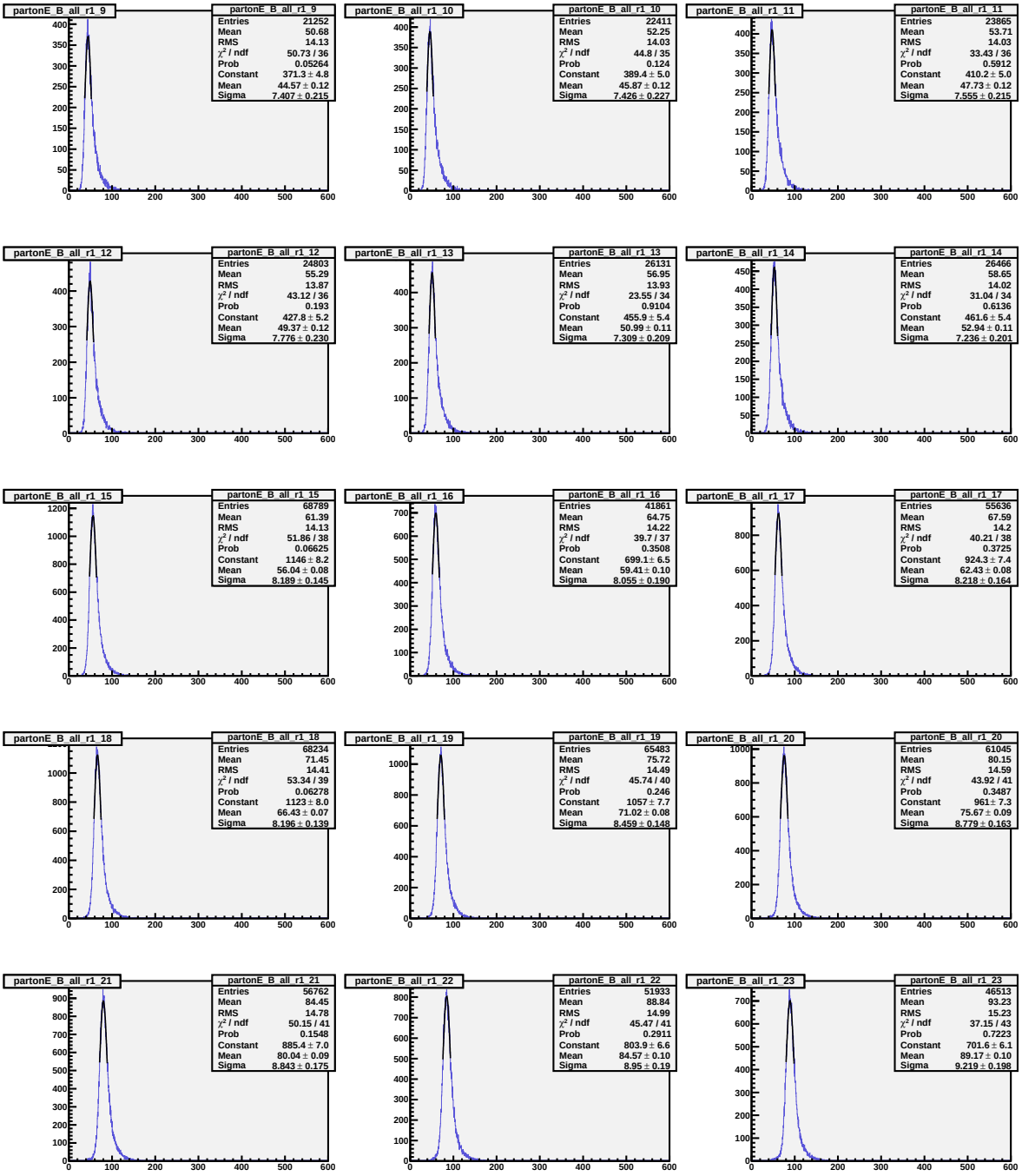


Figure B.2: E_{parton} distribution for b-quarks $\{|\eta| \in [0,0.5)\}$ in E_{jet} bins with bin boundaries $36, 38, 40, 42, 44, 46, 48, 53, 56, 60, 65, 70, 75, 80, 85, 90$ *GeV*

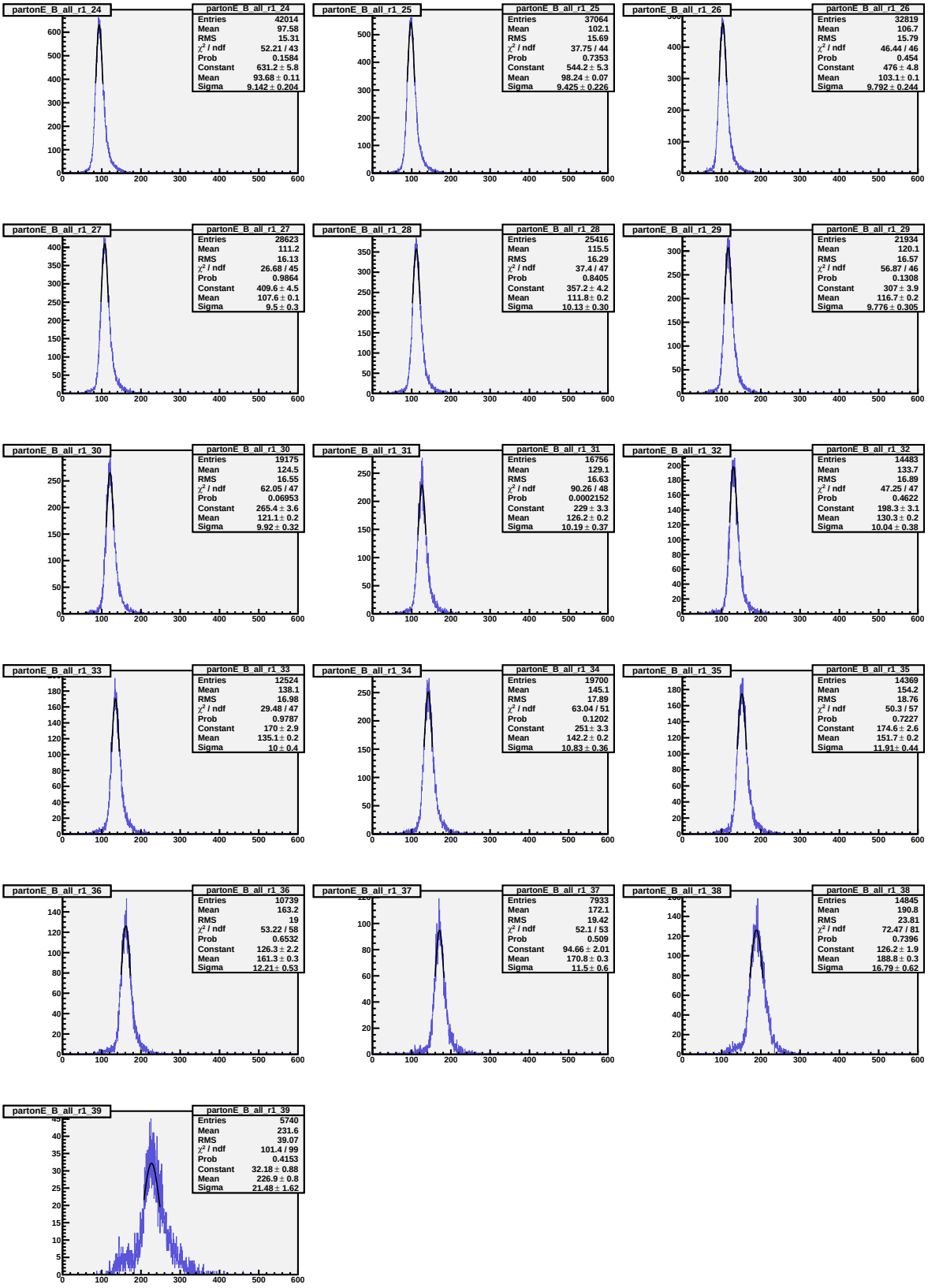


Figure B.3: E_{parton} distribution for b-quarks $\{|\eta| \in [0,0.5)\}$ in E_{jet} bins with bin boundaries $90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 150, 160, 170, 180, 220, 400$ *GeV*

- Region 2

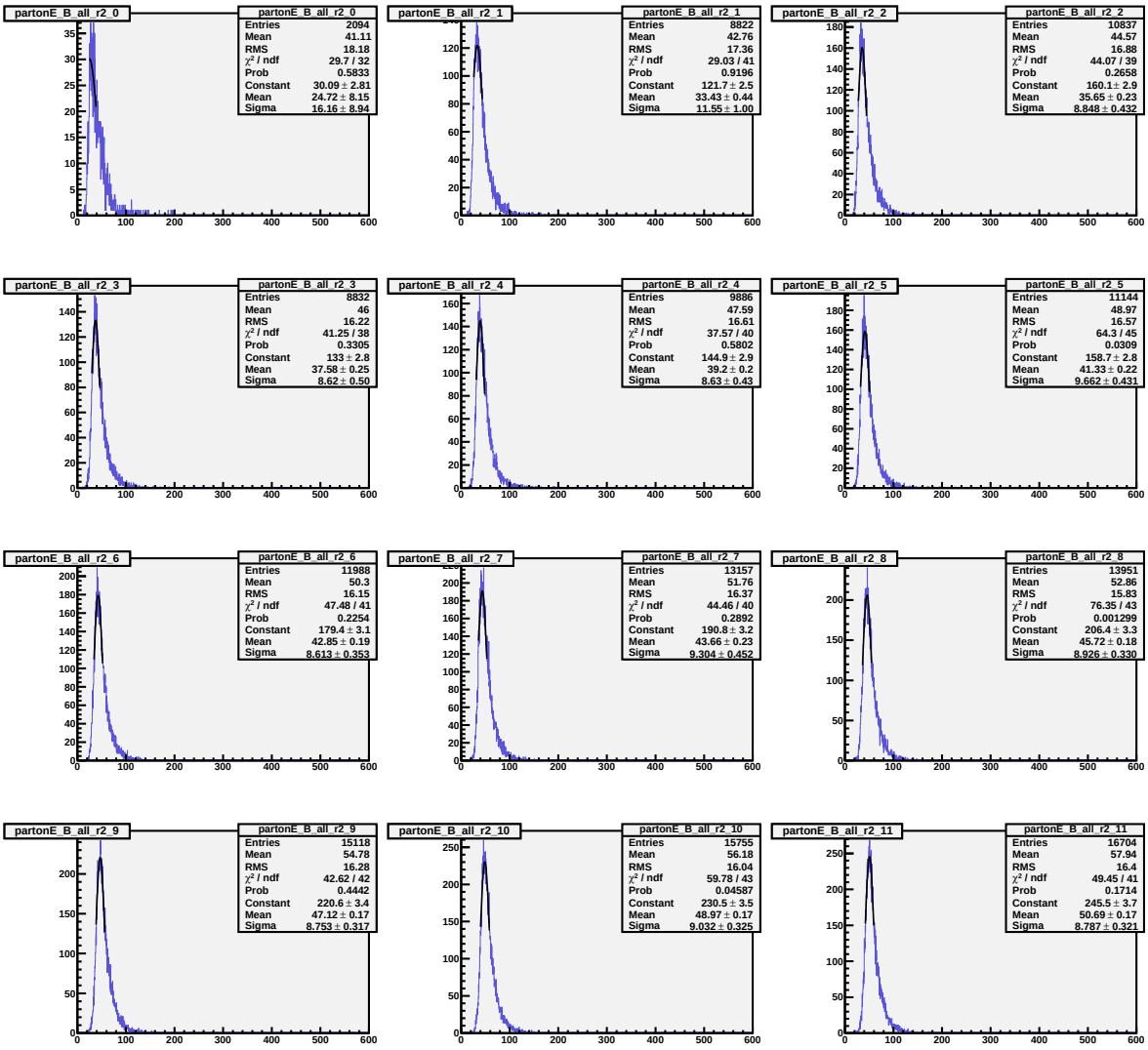


Figure B.4: E_{parton} distribution for b-quarks $\{|\eta| \in [0.5, 1.0)\}$ in E_{jet} bins with bin boundaries $0, 19, 23, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44$ *GeV*

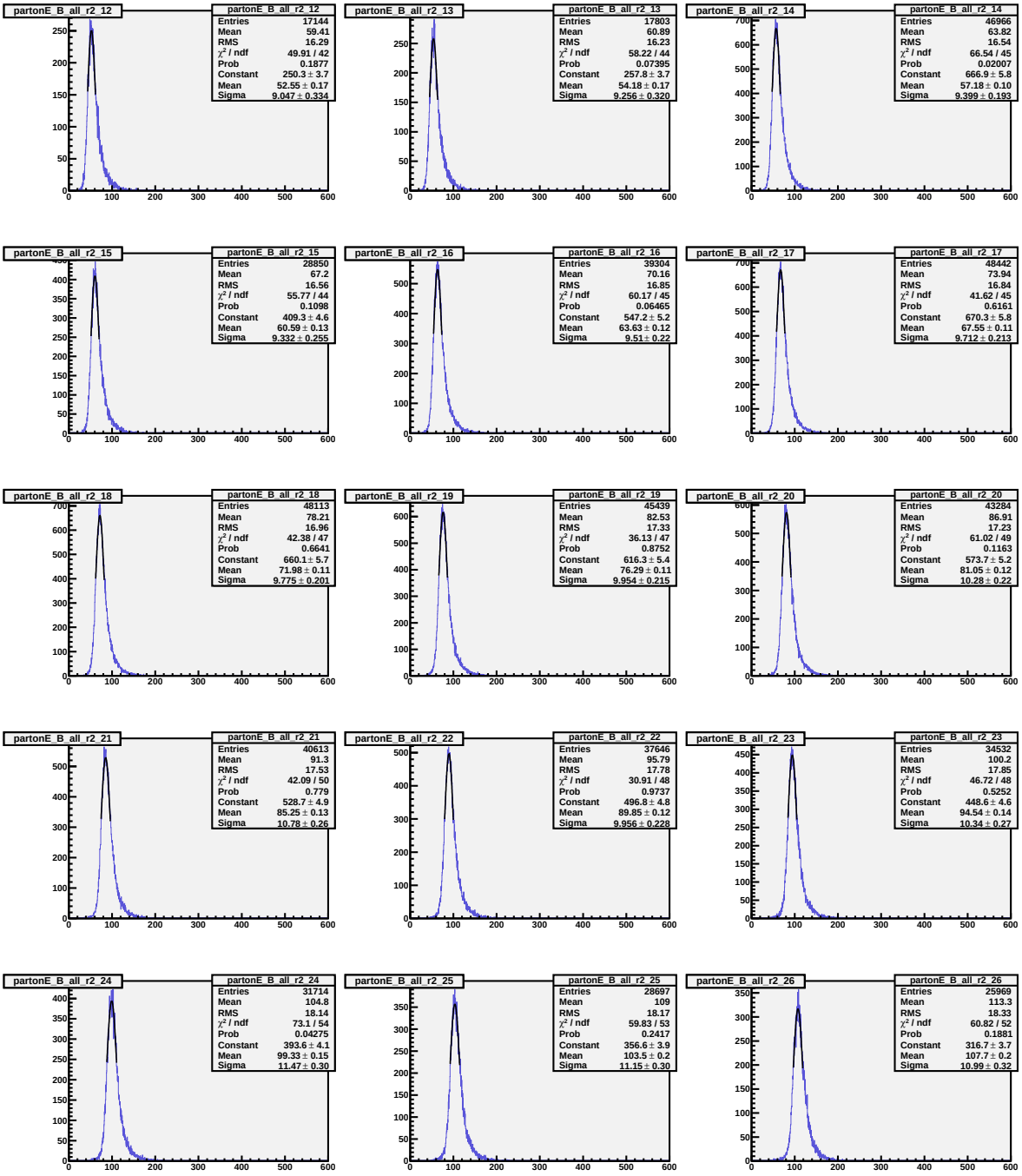


Figure B.5: E_{parton} distribution for b-quarks $\{|\eta| \in [0.5, 1.0)\}$ in E_{jet} bins with bin boundaries $44, 46, 48, 53, 56, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110$ *GeV*

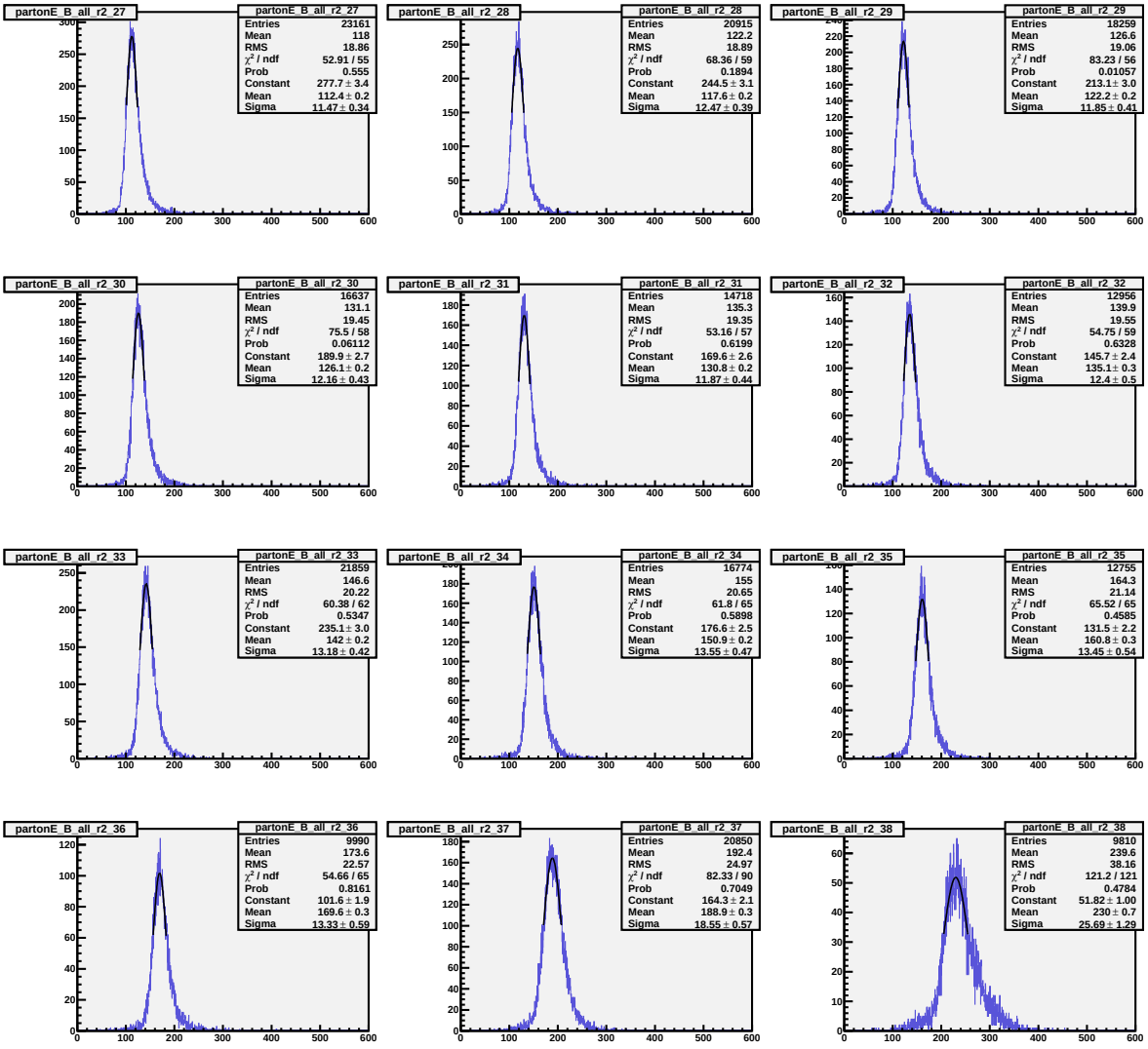


Figure B.6: E_{parton} distribution for b-quarks $\{|\eta| \in [0.5, 1.0)\}$ in E_{jet} bins with bin boundaries $110, 115, 120, 125, 130, 135, 140, 150, 160, 170, 180, 220, 400$ *GeV*

- Region 3

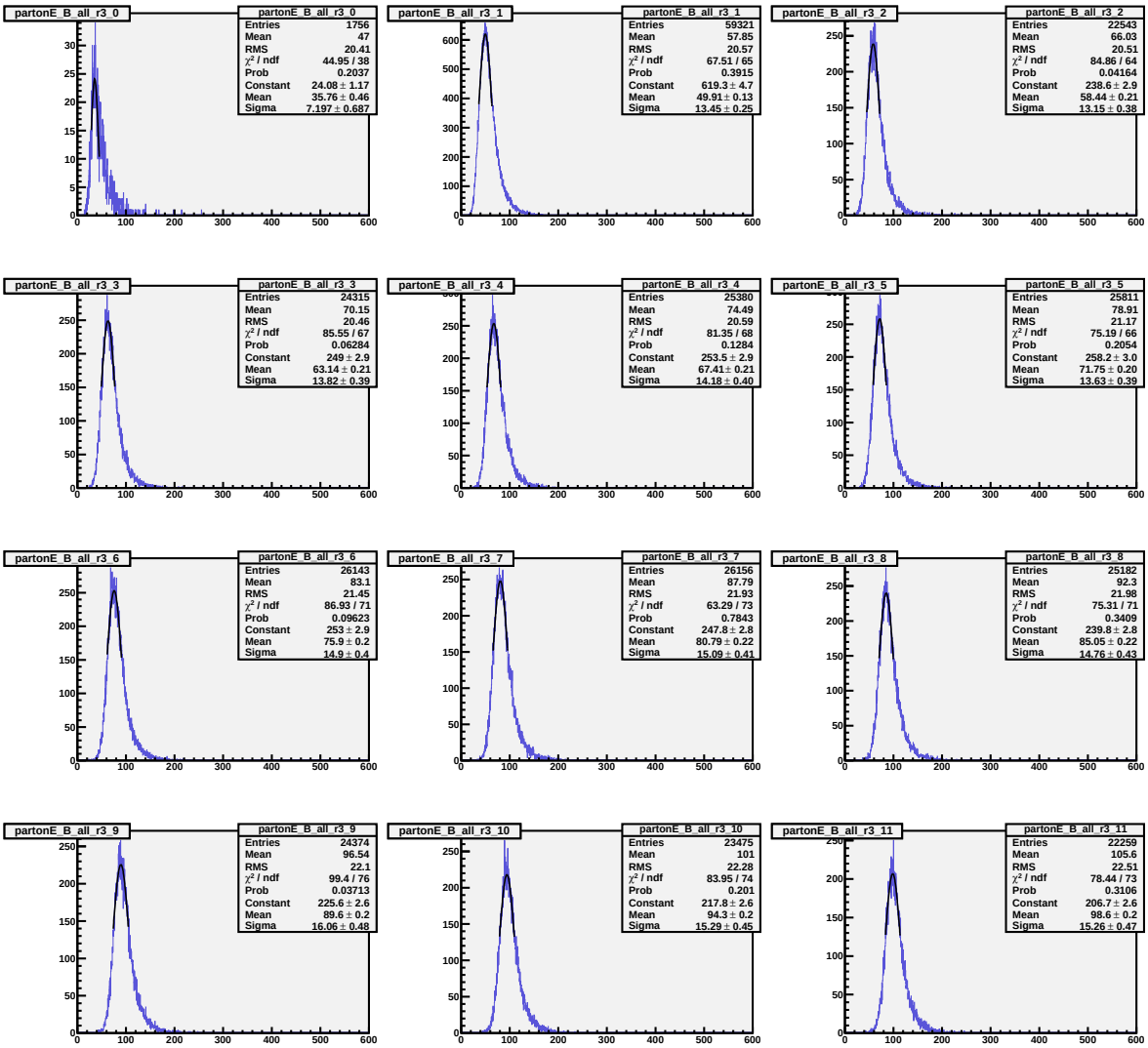


Figure B.7: E_{parton} distribution for b-quarks $\{|\eta| \in [1.0, 1.5)\}$ in E_{jet} bins with bin boundaries $0, 25, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95$ *GeV*

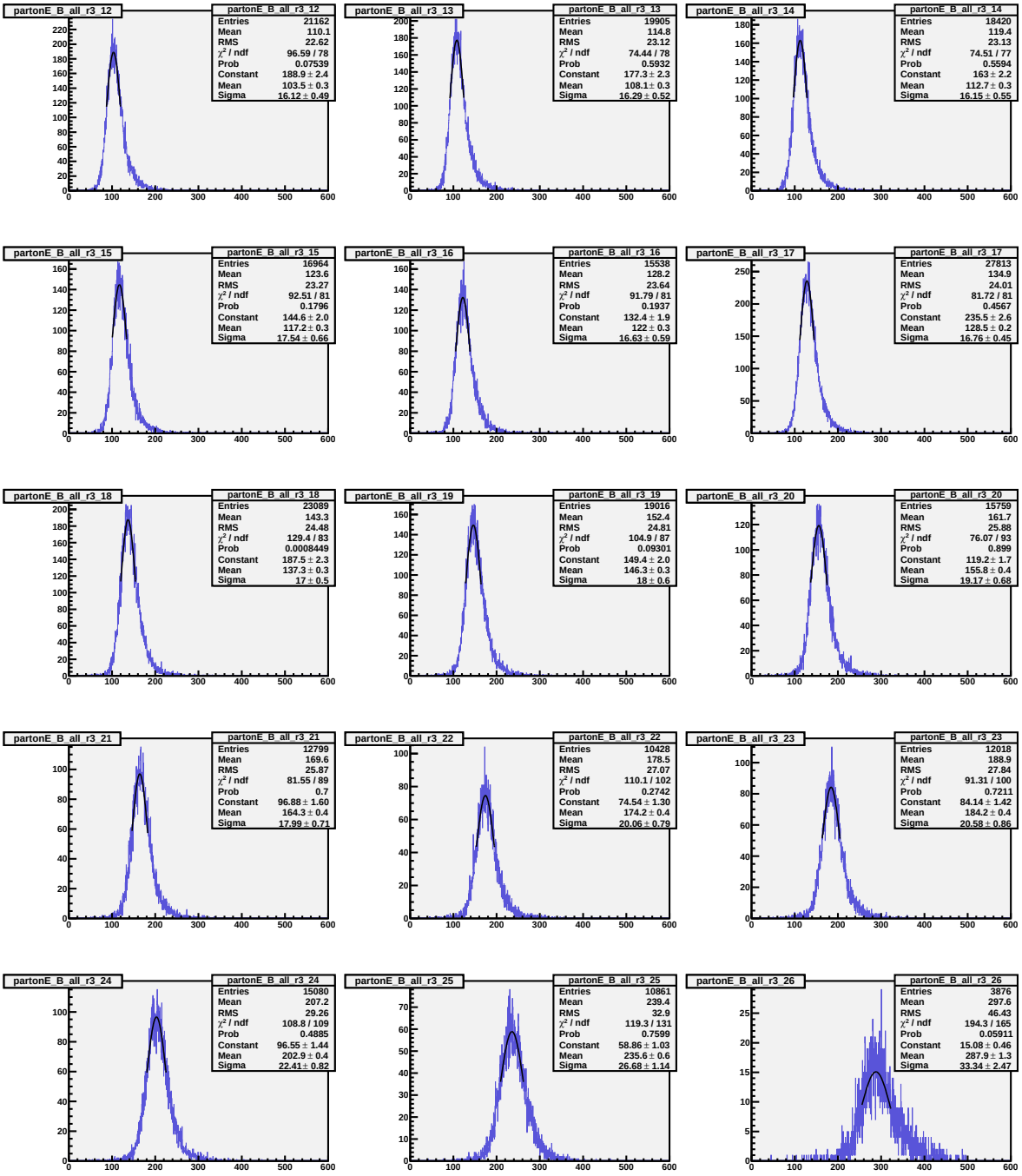


Figure B.8: E_{parton} distribution for b-quarks $\{|\eta| \in [1.0, 1.5)\}$ in E_{jet} bins with bin boundaries $95, 100, 105, 110, 115, 120, 130, 140, 150, 160, 170, 180, 195, 225, 280, 500$ *GeV*

- Region 4

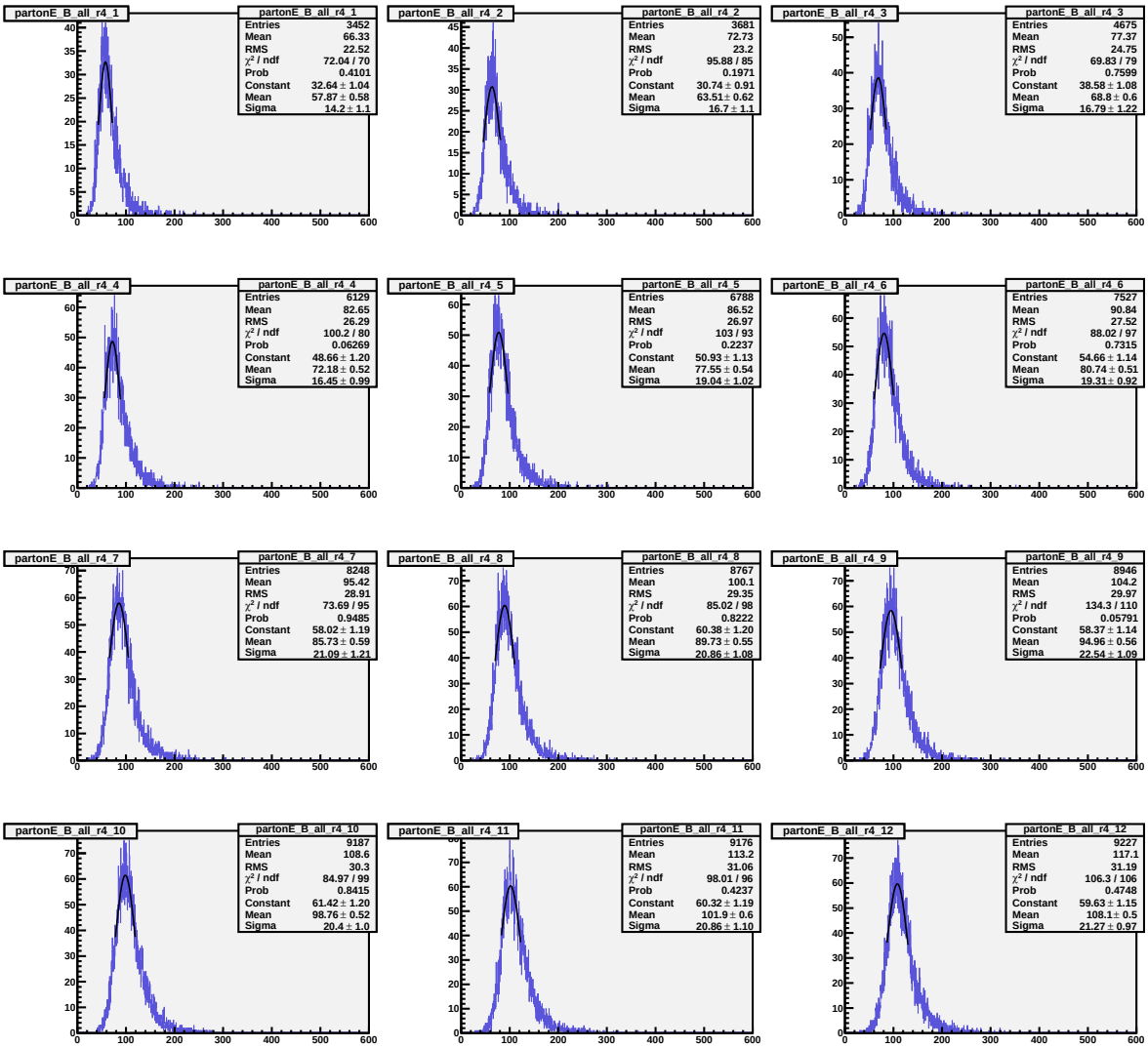


Figure B.9: E_{parton} distribution for b-quarks $\{|\eta| \in [1.5, 2.5)\}$ in E_{jet} bins with bin boundaries $25, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100\ GeV$

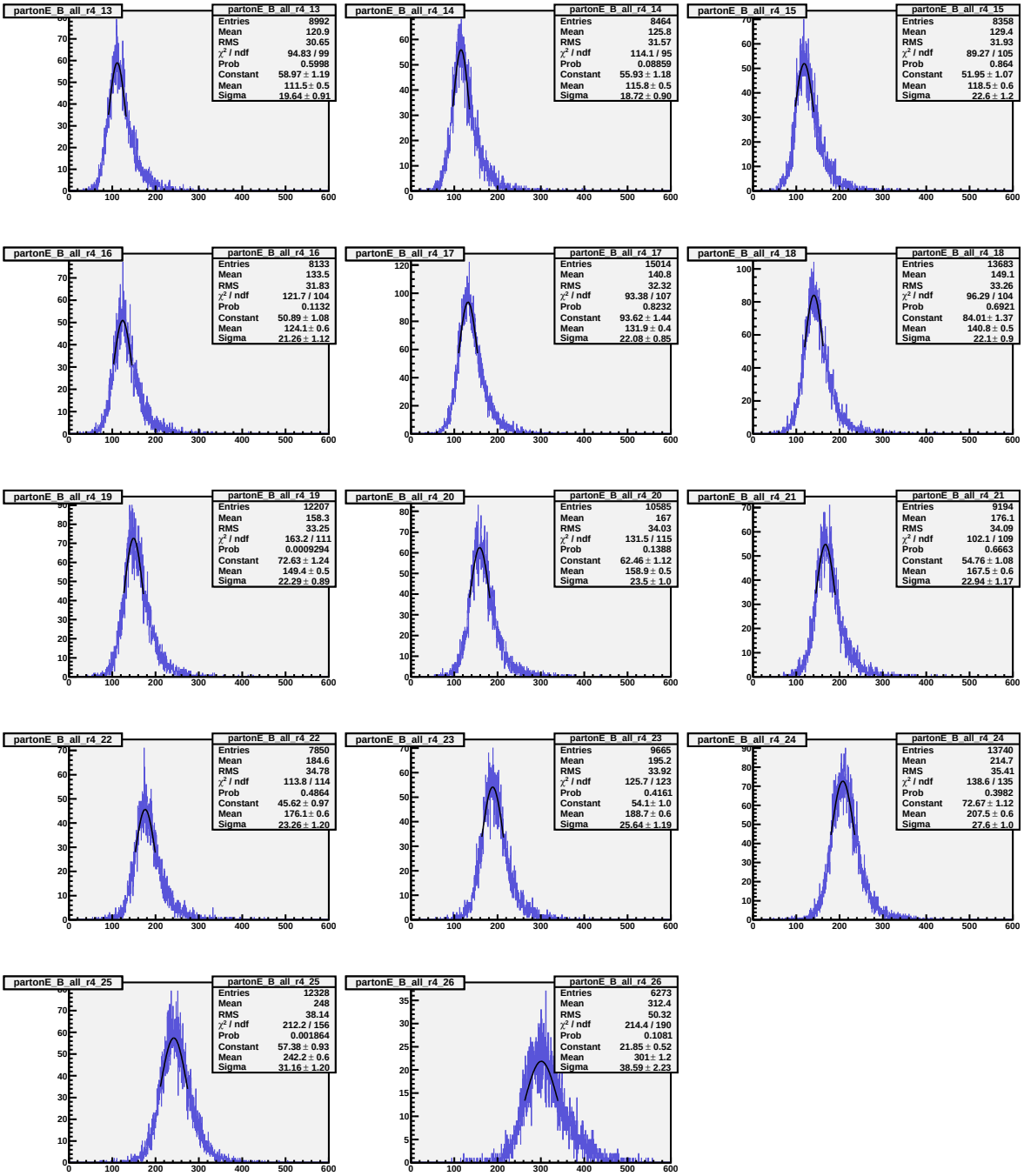


Figure B.10: E_{parton} distribution for b-quarks $\{|\eta| \in [1.5, 2.5)\}$ in E_{jet} bins with bin boundaries $100, 105, 110, 115, 120, 130, 140, 150, 160, 170, 180, 195, 225, 280, 500$ *GeV*

APPENDIX **C**

Pull distributions

- *Input JES = nominal JES*

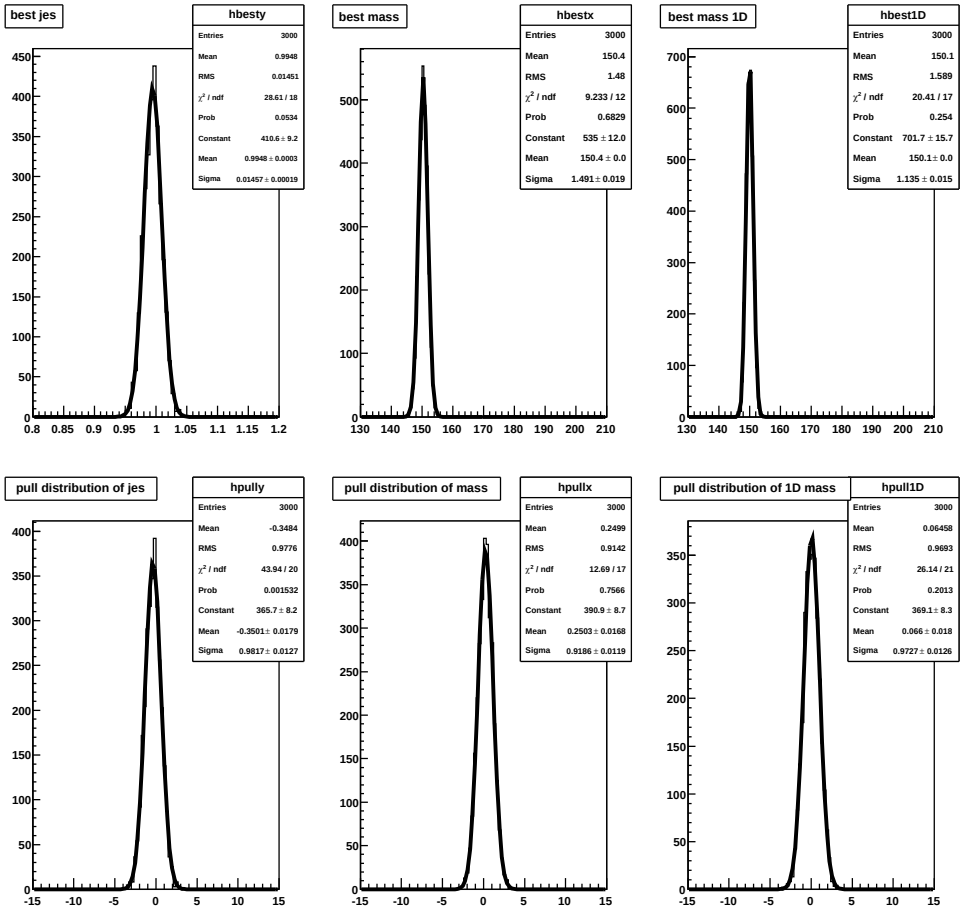


Figure C.1: Fitted parameters and their respective pulls for input top quark mass of $150\,GeV/c^2$

- $Input\,JES = nominal\,JES$

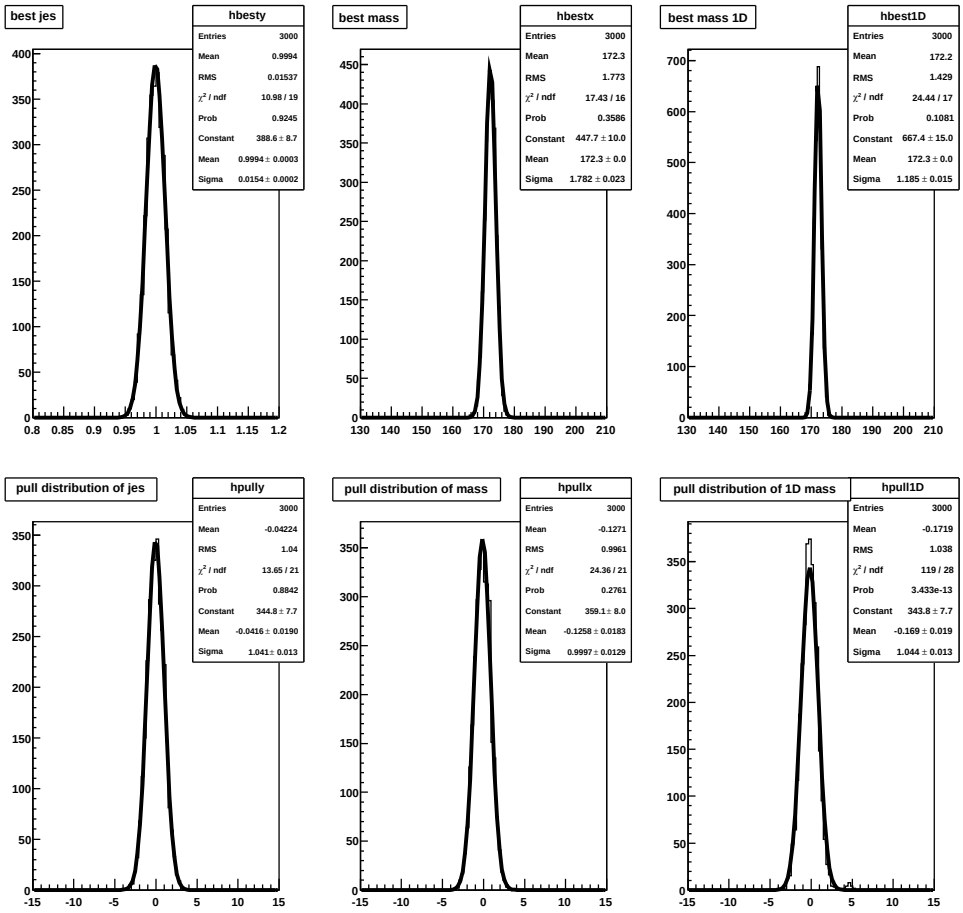


Figure C.2: Fitted parameters and their respective pulls for input top quark mass of $172.5\,GeV/c^2$

- $Input\,JES = nominal\,JES$

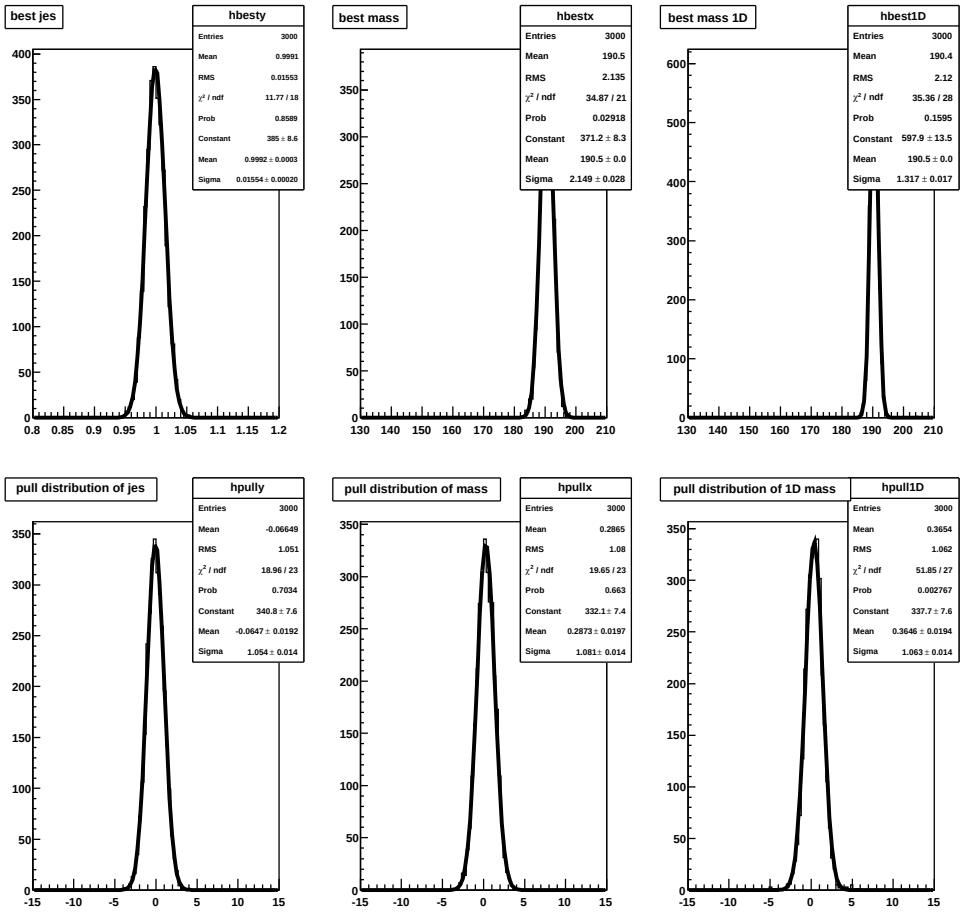


Figure C.3: Fitted parameters and their respective pulls for input top quark mass of $190\,GeV/c^2$

- $Input\,JES = 0.94$

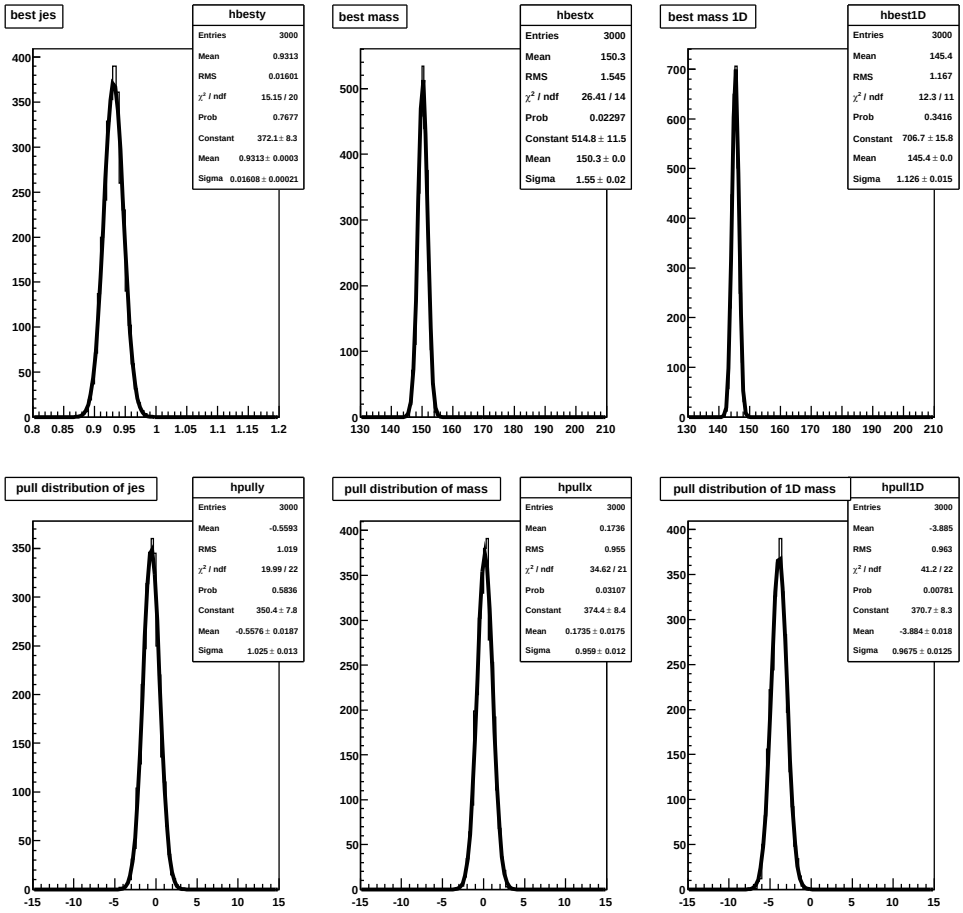


Figure C.4: Fitted parameters and their respective pulls for input top quark mass of $150\,GeV/c^2$

- $Input\,JES = 0.94$

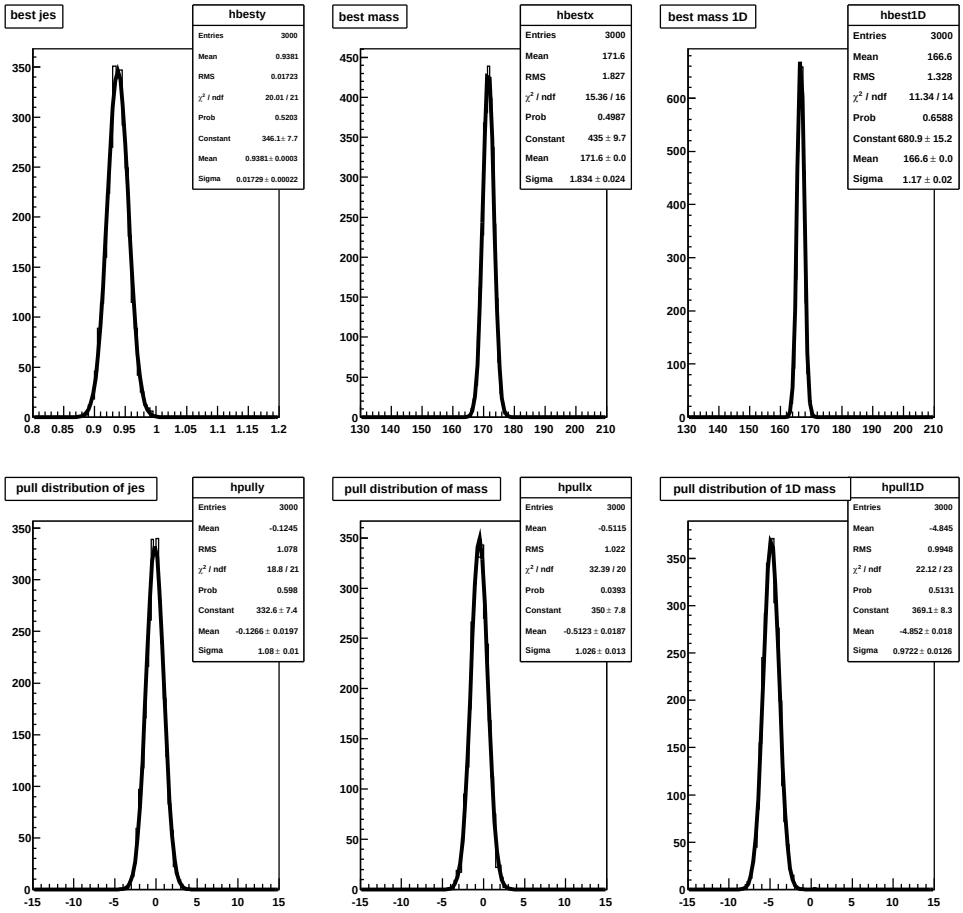


Figure C.5: Fitted parameters and their respective pulls for input top quark mass of $172.5\,GeV/c^2$

- $Input\,JES = 0.94$

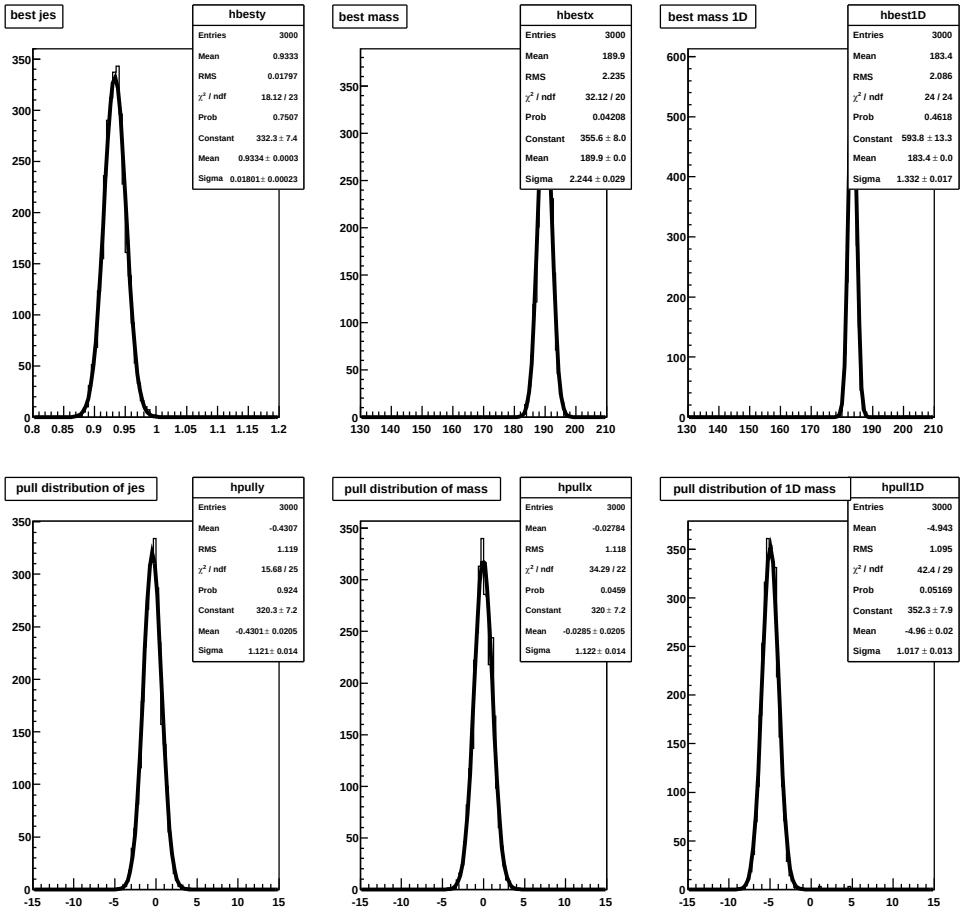


Figure C.6: Fitted parameters and their respective pulls for input top quark mass of $190\,GeV/c^2$

- $Input\,JES = 1.06$

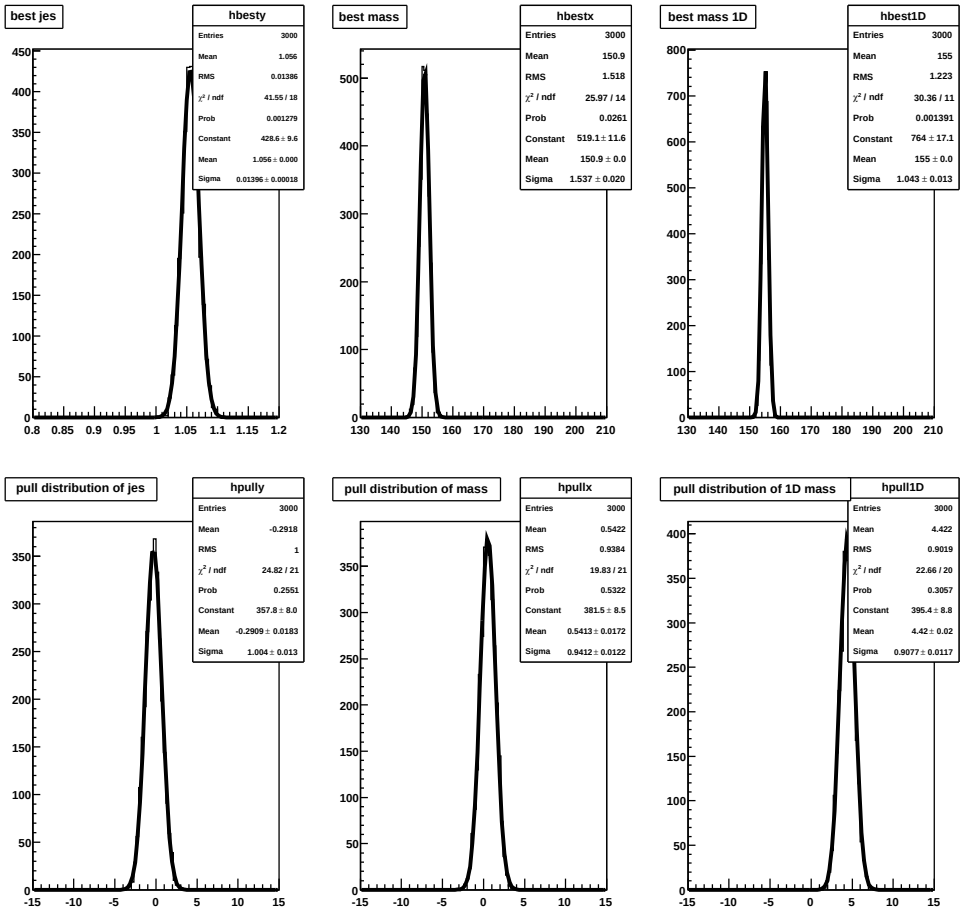


Figure C.7: Fitted parameters and their respective pulls for input top quark mass of $150\,GeV/c^2$

- $Input\,JES = 1.06$

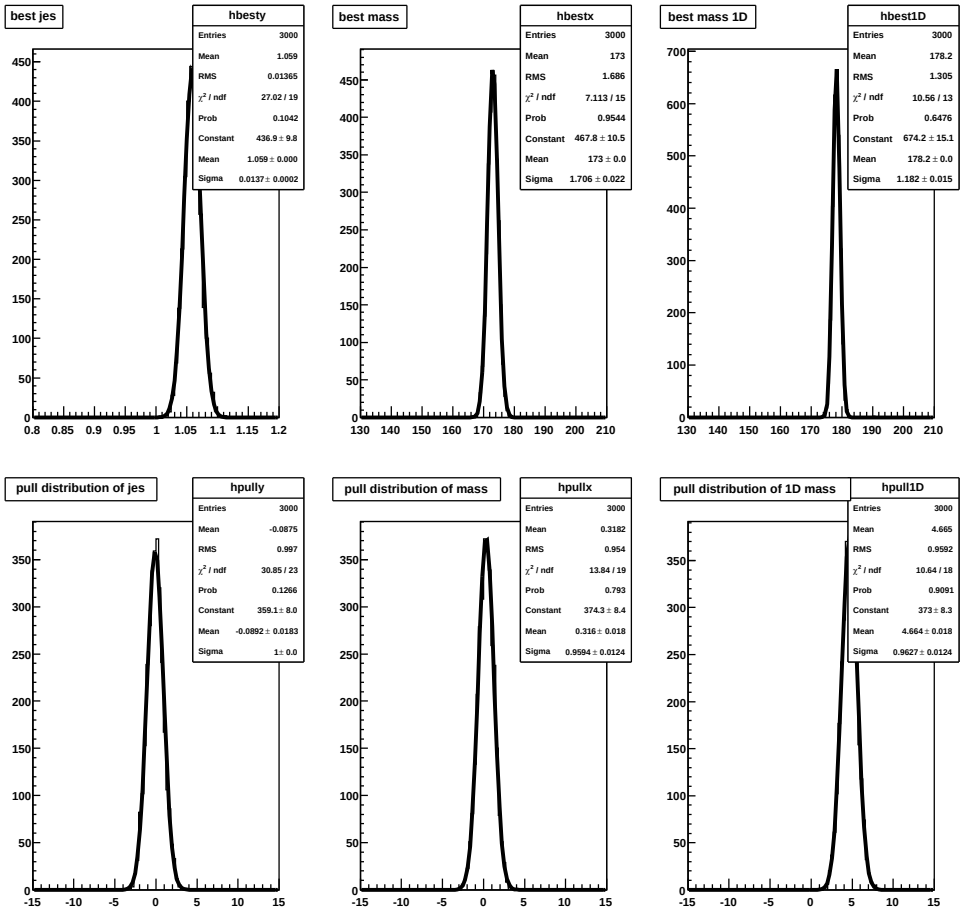


Figure C.8: Fitted parameters and their respective pulls for input top quark mass of $172.5\,GeV/c^2$

- $Input\,JES = 1.06$

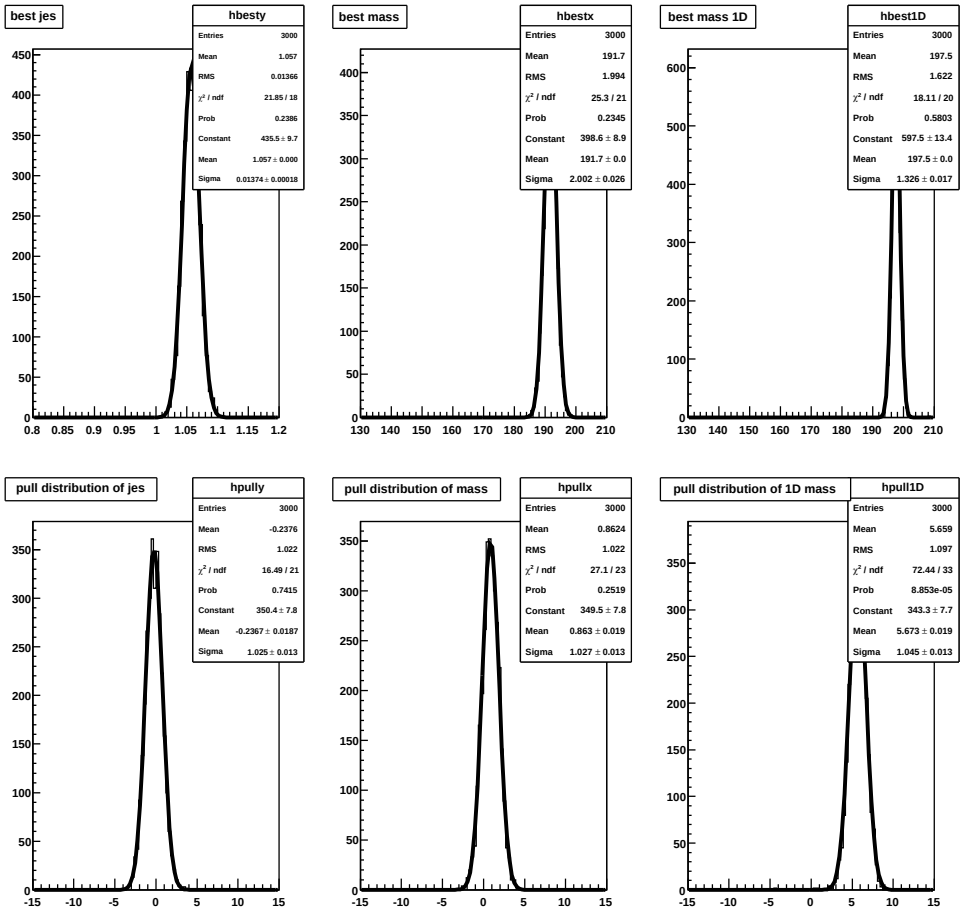


Figure C.9: Fitted parameters and their respective pulls for input top quark mass of $190\,GeV/c^2$

BIBLIOGRAPHY

[1] V. M. Abazov *et al.*, "Determination of the pole and MSbar masses of the top quark from the $t\bar{t}$ cross section," *Phys.Lett.*, vol. B703, pp. 422–427, 2011.

[2] V. Abazov *et al.*, "Determination of the strong coupling constant from the inclusive jet cross section in $p\bar{p}$ collisions at sqrt(s)=1.96 TeV," *Phys.Rev.*, vol. D80, p. 111107, 2009.

[3] "Tevatron Electroweak Working Group," *http://tevewwg.fnal.gov/*.

[4] M. Mulhearn, "The Level 2 Trigger," *D0 Collaboration meeting*, 2011.

[5] B. Webber, "Parton shower Monte Carlo event generators," *Scholarpedia*, vol. 6, no. 12, p. 10662, 2011.

[6] *http://www.viz.tamu.edu/faculty/parke/ends489f00/notes/sec1_8.html*.

[7] T. Aaltonen *et al.*, "Combination of the top-quark mass measurements from the Tevatron collider," *FERMILAB-PUB-12-336-E-TD, submitted to Phys. Rev. D*, 2012.

[8] F. Englert and R. Brout, "Broken symmetry and the mass of gauge vector mesons," *Phys. Rev. Lett.*, vol. 13, pp. 321–323, Aug 1964.

[9] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global conservation laws and massless particles," *Phys. Rev. Lett.*, vol. 13, pp. 585–587, Nov 1964.

[10] P. W. Higgs, "Broken symmetries and the masses of gauge bosons," *Phys. Rev. Lett.*, vol. 13, pp. 508–509, Oct 1964.

[11] J. Pumplin, D. Stump, J. Huston, H. Lai, P. M. Nadolsky, *et al.*, "New generation of parton distributions with uncertainties from global QCD analysis," *JHEP*, vol. 0207, p. 012, 2002.

[12] V. M. Abazov *et al.*, "Measurement of the top quark pair production cross section in the lepton+jets channel in proton-antiproton collisions at $\sqrt{s}$=1.96 TeV," *Phys.Rev.*, vol. D84, p. 012008, 2011.

[13] S. Moch and P. Uwer, "Theoretical status and prospects for top-quark pair production at hadron colliders," *Phys.Rev.*, vol. D78, p. 034003, 2008.

[14] V. Ahrens, A. Ferroglia, M. Neubert, B. D. Pecjak, and L. L. Yang, "Precision predictions for the t+t(bar) production cross section at hadron colliders," *Phys.Lett.*, vol. B703, pp. 135–141, 2011.

[15] M. Beneke, P. Falgari, S. Klein, and C. Schwinn, "Hadronic top-quark pair production with NNLL threshold resummation," *Nucl.Phys.*, vol. B855, pp. 695–741, 2012.

[16] M. Cacciari, M. Czakon, M. Mangano, A. Mitov, and P. Nason, "Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation," *Phys.Lett.*, vol. B710, pp. 612–622, 2012.

[17] V. M. Abazov *et al.*, "Determination of the width of the top quark," *Phys.Rev.Lett.*, vol. 106, p. 022001, 2011.

[18] R. R. Wilson, "The Tevatron," *Phys.Today*, vol. 30N10, pp. 23–30, 1977.

[19] *www-d0.fnal.gov*.

[20] *www-cdf.fnal.gov*.

[21] V. Abazov *et al.*, "The Upgraded D0 detector," *Nucl.Instrum.Meth.*, vol. A565, pp. 463–537, 2006.

[22] R. Ruchti, "The use of scintillating fibers for charged-particle tracking," *Ann.Rev.Nucl.Part.Sci.*, vol. 46, pp. 281–319, 1996.

[23] P. S. Baringer *et al.*, "Cosmic ray tests of the D0 preshower detector," *Nucl.Instrum.Meth.*, vol. A469, pp. 295–310, 2001.

[24] S. Abachi *et al.*, "Search for high mass top quark production in $p\bar{p}$ collisions at $\sqrt{s} =$ 1.8 TeV," *Phys.Rev.Lett.*, vol. 74, pp. 2422–2426, 1995.

[25] S. J. Wimpenny, "THE HADRON AND ELECTRON RESPONSE OF THE URANIUM / LIQUID ARGON CAORIMETER MODULES FOR THE D0 DETECTOR," *Nucl.Instrum.Meth.*, vol. A279, pp. 107–113, 1989.

[26] A. L. Spadafora, "Test beam results from the D0 liquid argon end calorimeter electromagnetic module," *Nucl.Instrum.Meth.*, vol. A315, pp. 279–284, 1992.

[27] P. Franzini, "Performance of the D0 uranium - liquid argon calorimeter modules. D0 Calorimeter Group," *Nucl.Instrum.Meth.*, vol. A289, pp. 438–445, 1990.

[28] H. Aihara, "Performance of the D0 end calorimeter electromagnetic module," *IEEE Trans.Nucl.Sci.*, vol. 38, pp. 398–402, 1991.

[29] S. Abachi *et al.*, "The D0 Detector," *Nucl.Instrum.Meth.*, vol. A338, pp. 185–253, 1994.

[30] S. Eidelman, K. Hayes, K. Olive, M. Aguilar-Benitez, C. Amsler, D. Asner, K. Babu, R. Barnett, J. Beringer, P. Burchat, C. Carone, C. Caso, G. Conforto, O. Dahl, G. D'Ambrosio, M. Doser, J. Feng, T. Gherghetta, L. Gibbons, M. Goodman, C. Grab, D. Groom, A. Gurtu, K. Hagiwara, J. Hernández-Rey, K. Hikasa, K. Honscheid, H. Jawahery, C. Kolda, K. Y., M. Mangano, A. Manohar, J. March-Russell, A. Masoni, R. Miquel, K. Mönig, H. Murayama, K. Nakamura, S. Navas, L. Pape, C. Patrignani, A. Piepke, G. Raffelt, M. Roos, M. Tanabashi, J. Terning, N. Törnqvist, T. Trippe, P. Vogel, C. Wohl, R. Workman, W.-M. Yao, P. Zyla, B. Armstrong, P. Gee, G. Harper, K. Lugovsky, S. Lugovsky, V. Lugovsky, A. Rom, M. Artuso, E. Barberio, M. Battaglia, H. Bichsel, O. Biebel, P. Bloch, R. Cahn, D. Casper, A. Cattai, R. Chivukula, G. Cowan, T. Damour, K. Desler, M. Dobbs, M. Drees, A. Edwards, D. Edwards, V. Elvira, J. Erler, V. Ezhela, W. Fetscher, B. Fields, B. Foster, D. Froidevaux, M. Fukugita, T. Gaisser, L. Garren, H.-J. Gerber, G. Gerbier, F. Gilman, H. Haber, C. Hagmann, J. Hewett, I. Hinchliffe, C. Hogan, G. Höhler, P. Igo-Kemenes, J. Jackson, K. Johnson, D. Karlen, B. Kayser, D. Kirkby, S. Klein, K. Kleinknecht, I. Knowles, P. Kreitz, Y. Kuyanov, O. Lahav, P. Langacker, A. Liddle, L. Littenberg, D. Manley, A. Martin, M. Narain, P. Nason, Y. Nir, J. Peacock, H. Quinn, S. Raby, B. Ratcliff, E. Razuvaev, B. Renk, G. Rolandi, M. Ronan, L. Rosenberg, C. Sachrajda, Y. Sakai, A. Sanda, S. Sarkar, M. Schmitt, O. Schneider, D. Scott, W. Seligman, M. Shaevitz, T. Sjöstrand, G. Smoot, S. Spanier, H. Spieler, N. Spooner, M. Srednicki, A. Stahl, T. Stanev, M. Suzuki, N. Tkachenko, G. Trilling, G. Valencia, K. van Bibber, M. Vincter, D. Ward, B. Webber, M. Whalley, L. Wolfenstein, J. Womersley, C. Woody, O. Zenin, and R.-Y. Zhu, "Review of Particle Physics," *Physics Letters B*, vol. 592, pp. 1+, 2004.

[31] B. Casey, M. Corcoran, Y. Enari, R. Partridge, M. Prewitt, and G. Snow, "The D0 Run

2b Luminosity Constant," *D0Note 5945*, 2009.

[32] B. Casey *et al.*, "The D0 Run IIb Luminosity Measurement," *FERMILAB-TM-2529-E*, 2012.

[33] T. Adams, Q. An, K. Black, T. Bose, N. Buchanan, *et al.*, "The D0 Run II impact parameter trigger," *Nucl.Instrum.Meth.A*, 2007.

[34] R. D. Angstadt, G. Brooijmans, M. Johnson, A. Kulyavtsev, M. Mulders, *et al.*, "The DZERO level 3 data acquisition system," *IEEE Trans.Nucl.Sci.*, vol. 51, pp. 445–450, 2004.

[35] A. Khanov, "HTF: histogramming method for finding tracks. The algorithm description," *D0Note 3778*, 2000.

[36] G. Borissov, "Ordering a Chaos or ... Technical Details of AA Tracking," *http://www-d0.fnal.gov/global_tracking/talks/20030228/talk-adm-030228.ps*, vol. All D0 Meeting, February 28 (2003).

[37] H. Greenlee, "The D0 Kalman Track Fit," *D0Note 4303*, 2003.

[38] S. Abachi *et al.*, "Beam tests of the D0 uranium liquid argon end calorimeters," *Nucl.Instrum.Meth.*, vol. A324, pp. 53–76, 1993.

[39] A. Schwartzman and C. Tully, "Primary Vertex Reconstruction by Means of Adaptive Vertex Fitting," *D0Note 4918*, 2005.

[40] A. Schwartzman and M. Narain, "Probabilistic Primary Vertex Selection," *D0Note 4042*, 2007.

[41] G. Steinbruek, "Measurement of the Angular Distribution of Electrons from W boson decays at D0," *Ph.D. Thesis, University Of Oklahoma (1999)*, 1999.

[42] J. Kozminski, R. Kehoe, H. Weerts, S.-J. Park, A. Quadt, J. Gardner, and S. Jabeen, "Electron Likelihood in p14," *D0Note 4449*, 2004.

[43] U. Bassler and G. Bernardi, "Towards a Coherent Treatment of Calorimetric Energies: Missing Transverse Energy, Jets, E.M. Objects and the T42 Algorithm," *D0Note 4124*, 2003.

[44] G. C. Blazey, J. R. Dittmann, S. D. Ellis, V. D. Elvira, K. Frame, *et al.*, "Run II jet physics," *FERMILAB-CONF-00-092-E*, pp. 47–77, 2000.

[45] B. Abbott *et al.*, "Determination of the absolute jet energy scale in the D0 calorimeters," *Nucl.Instrum.Meth.*, vol. A424, pp. 352–394, 1999.

[46] V. M. Abazov *et al.*, "Measurement of the inclusive jet cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," *Phys.Rev.*, vol. D85, p. 052006, 2012.

[47] K. DeVaughan, G. Golovanov, U. Heintz, S. Jabeen, A. Kupco, *et al.*, "Jet Energy Scale Determination for D0 Run IIb," *D0Note 5801*, 2012.

[48] Z. Ye, S. Atkins, K. Augsten, S. Evstyukhin, A. Juste, *et al.*, "Correction for the MC-Data Difference in the Jet Response at D0 for Run IIB," *D0Note 6144*, 2011.

[49] Z. Ye, S. Atkins, K. Augsten, S. Evstyukhin, A. Juste, *et al.*, "Correction for the MC-Data Difference in the Jet Response at D0," *D0Note 6143*, 2011.

[50] V. Abazov *et al.*, "b-Jet Identification in the D0 Experiment," *Nucl.Instrum.Meth.*, vol. A620, pp. 490–517, 2010.

[51] T. Sjostrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual," *JHEP*, vol. 0605, p. 026, 2006.

[52] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, "ALPGEN, a generator for hard multiparton processes in hadronic collisions," *JHEP*, vol. 0307, p. 001, 2003.

[53] S. Hoeche, F. Krauss, N. Lavesson, L. Lonnblad, M. Mangano, *et al.*, "Matching parton showers and matrix elements," *arXiv:hep-ph/0602031*, 2006.

[54] R. Hamberg, W. van Neerven, and T. Matsuura, "A Complete calculation of the order $\alpha - s^2$ correction to the Drell-Yan K factor," *Nucl.Phys.*, vol. B359, pp. 343–405, 1991.

[55] E. Boos, V. Bunichev, L. Dudko, V. Savrin, and A. Sherstnev, "Method for simulating electroweak top-quark production events in the NLO approximation: SingleTop event generator," *Phys.Atom.Nucl.*, vol. 69, pp. 1317–1329, 2006.

[56] N. Kidonakis, "Single top production at the Tevatron: Threshold resummation and finite-order soft gluon corrections," *Phys.Rev.*, vol. D74, p. 114012, 2006.

[57] J. M. Campbell and R. Ellis, "MCFM for the Tevatron and the LHC," *Nucl.Phys.Proc.Suppl.*, vol. 205-206, pp. 10–15, 2010.

[58] R. Brun and F. Carminati, "CERN Program Library Long Writeup," *unpublished*, vol. W5013, 1993.

[59] S. Fatakia and U. Heintz, "Correcting the reconstructed jet's 4-vector to the parton level," *D0Note 4247*, 2003.

[60] M. Demarteau, P. Houben, M. Vreeswijk, and M. Weber, "p17 parton level corrections and resolutions for hitfit," *D0Note 5418*, 2007.

[61] S. S. Snyder, "Measurement of the top quark mass at D0," *FERMILAB-THESIS-1995-27*, 1995.

[62] M. Mulders and M. Weber, "Top Mass Measurement with b-tagging and Jet Energy Scale Fit in the Lepton+Jets Channel using the Ideogram Method," *D0 Note 4802*, 2006.

[63] M. Demarteau, P. Houben, *et al.*, "Top Mass Measurement in the Lepton+Jets Channel using the Ideogram Method," *D0 Note 5099*, 2006.

[64] V. M. Abazov *et al.*, "Precise measurement of the top-quark mass from lepton + jets events at D0," *Phys. Rev. D*, vol. 84, p. 032004, Aug 2011.

[65] N. Makovec and J.-F. Grivaz, "Shifting, Smearing and Removing Simulated Jets," *D0 Note 4914*, 2005.

[66] V. Abazov *et al.*, "b-Jet Identification in the D0 Experiment," *Nucl.Instrum.Meth.*, vol. A620, pp. 490–517, 2010.

[67] A. Abulencia *et al.*, "Top quark mass measurement using the template method in the lepton + jets channel at CDF II," *Phys.Rev.*, vol. D73, p. 032003, 2006.

[68] S. Frixione and B. R. Webber, "The MC and NLO 3.4 Event Generator," *arXiv:0812.0770 [hep-ph]*, 2008.

[69] P. Abreu *et al.*, "Tuning and test of fragmentation models based on identified particles and precision event shape data," *Z.Phys.*, vol. C73, pp. 11–60, 1996.

[70] M. Bowler, "e+ e- Production of Heavy Quarks in the String Model," *Z.Phys.*, vol. C11, p. 169, 1981.

[71] Y. Peters, K. Hamacher, and D. Wicke, "Precise tuning of the b fragmentation for the D0 Monte Carlo," *FERMILAB-TM-2425-E*, 2006.

[72] V. Abazov *et al.*, "Measurement of the top quark mass in the lepton+jets final state with the matrix element method," *Phys.Rev.*, vol. D74, p. 092005, 2006.

[73] V. Abazov *et al.*, "Precise measurement of the top quark mass from lepton+jets events at D0," *Phys.Rev.Lett.*, vol. 101, p. 182001, 2008.

[74] (Particle Data Group), J. Beringer, *et al.*, "The Review of Particle Physics," *Phys. Rev. D*, vol. D86, p. 010001, 2012.

[75] W. Yao *et al.*, "Review of Particle Physics," *J.Phys.G*, vol. G33, pp. 1–1232, 2006.

[76] J. Alcaraz *et al.*, "A Combination of preliminary electroweak measurements and constraints on the standard model," 2006.

[77] A. H. Hoang and I. W. Stewart, "Top Mass Measurements from Jets and the Tevatron Top-Quark Mass," *Nucl.Phys.Proc.Suppl.*, vol. 185, pp. 220–226, 2008.

[78] S. Fleming, A. H. Hoang, S. Mantry, and I. W. Stewart, "Jets from massive unstable particles: Top-mass determination," *Phys.Rev.*, vol. D77, p. 074010, 2008.

[79] N. Kidonakis and R. Vogt, "Next-to-next-to-leading order soft gluon corrections in top quark hadroproduction," *Phys.Rev.*, vol. D68, p. 114014, 2003.

[80] S. Catani, M. L. Mangano, P. Nason, and L. Trentadue, "The Resummation of soft gluons in hadronic collisions," *Nucl.Phys.*, vol. B478, pp. 273–310, 1996.

[81] A. Hoang, A. Manohar, I. W. Stewart, and T. Teubner, "The Threshold t anti-t cross-section at NNLL order," *Phys.Rev.*, vol. D65, p. 014014, 2002.

[82] P. M. Nadolsky, H.-L. Lai, Q.-H. Cao, J. Huston, J. Pumplin, *et al.*, "Implications of CTEQ global analysis for collider observables," *Phys.Rev.*, vol. D78, p. 013004, 2008.

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

[84] V. Ahrens, A. Ferroglia, M. Neubert, B. D. Pecjak, and L. L. Yang, "Renormalization-Group Improved Predictions for Top-Quark Pair Production at Hadron Colliders," *JHEP*, vol. 1009, p. 097, 2010.