Studies of High pT jets in CMS

A Thesis

Submitted to the
Tata Institute of Fundamental Research, Mumbai
for the degree of Master of Science
in Physics

by

Suvadeep Bose

Department of High Energy Physics
Tata Institute of Fundamental Research
Mumbai
January, 2007

To
My Parents

Acknowledgement

I would like to take the opportunity to acknowledge the invaluable and indispensable guidance of Prof. Sunanda Banerjee, without which this project would not have materialised. It was not only physics but also software techniques that I learnt from him. He showed me how to write a program in a neat and compact way. He gave me enough freedom to learn the collider physics at LHC in my own way but simultaneously guided me all along whenever I had problems.

The acknowledgement will be incomplete without the mention of the support I got from Dr. Monoranjan Guchait and Dr. Gobinda Majumder. They were the people who taught me the experimental and theoretical aspects of particle physics in my graduate course at TIFR. A substantial part of this project has been carried out in collaboration with Dr. Guchait. I took help from Dr. Majumder whenever I needed for any sort of discussion from detector physics to computing techniques.

During the course of my work TIFR always offered a pleasant atmosphere and all the facilities I needed. I express my gratitude to all the professors in my graduate school at TIFR and all the members of the Department of High Energy Physics. I remember the occasional discussions I had with Prof. Atul Gurtu, Prof. Probir Roy and Dr. Amol Dighe on physics-related issues. The project work I carried out during my first year of graduate course under the guidance of Prof. N. K. Mondal was my first exposure to experimental particle physics. I learnt many basic things from that project.

My warm thanks go to Mr. P. V. Deshpande and Mr. C. Niranjan for maintaining the computing system in our group in good shape all along. They always extended their helping hands whenever I had problems with the computers. I learnt a bit of linux programming from them as well.

My seniors, Seema and Shubhendu, helped me on numerous occasions during the course of this work. I truly appreciate their help since the very first days of my joining the group. At a time when I was getting started with this subject they were the sources of constant support which made me have a feel for my work.

I would like to thank those closest to me, whose presence helped make the completion of my graduate work possible. This is to thank all the fellow students in the Experimental High Energy Physics group - Nikhil, Arun, Abinash, Garima, Pratima, Sangeeta.

I thank my friends in TIFR for their smiles and company in hard times - Ajay, Amit, Aniket, Manjusha, Suresh, Basudev, Navodit, Kadir, Rakesh, Vandna, Shanta, Shamik, to name a few.

I also enjoyed the cultural atmosphere at TIFR that helped me to breathe out from the academic world. The annual events organised by the students and the musical evenings that we have in our campus not only were welcoming experiences for me but they also opened the door for new friendships.

The road to my graduate degree has been long and winding, so I would also like to thank some people from the early days. My undergraduate teachers in Jadavpur University and the great friends there; the people associated to my stay in Cambridge as an affiliated student during my 2nd BA; the good old friends from school days in Kolkata – I remember you all and thank you for offering me such nice memories which remain as assets for a lifetime.

Last but not the least I thank my parents for their understanding, love and encouragement.

Contents

Synopsis

Title of the Thesis	:	Studies of high p_T jets in CMS
Name of the Candidate	:	Mr. Suvadeep Bose
Registration No.	:	PHYS–080 dated 27.06.2006
Degree	:	Master of Science
Subject	:	Physics
Thesis Supervisor	:	Prof. Sunanda Banerjee
Institution	:	Tata Institute of Fundamental Research.

Studies of High p_T Jets in CMS

The signature of hadronic events is a large number of particles observed in the final states and large visible energy measured in the detector. The hadronic events are characterised by two, three or more jet topology, corresponding to zero, one or more hard quark/gluon radiations. The jet properties of hadronic events have manifestly direct correspondence with the global event shape variables of the hadronic events. Hadronic event shape variables are tools to study both the amount of gluon radiation and the details of the hadronisation process. Some of these variables, e.g.. *thrust* (T), can be used to measure the geometric properties of the energy flow in QCD final states.

The behaviour of QCD di-jet events has been studied at the generator level and also at the reconstructed level. To generate events packages like PYTHIA has been used in the framework of CMS specific software like CMKIN. In order to reconstruct events CMS fast simulation package like FAMOS has been used. The standard work has also been reproduced in the framework of CMSSW package which is a self-independent package for CMS. This software is going to be used in CMS experiment for generating and analysing events.

The $2 \rightarrow 2$ processes are used in general as a basis of the study and only hard-scattering events which are expected in the CMS experiment are selected in particular. The CMS detector is one of the four detectors employed in the Large Hadron Collider (LHC) experiment at CERN to verify the Standard Model and to look for physics beyond it, if there is any. LHC, a proton-proton collider, will operate at centre-of-mass energy of $\sqrt{s} = 14$TeV and will begin with the luminosity $\sim 10^{33}\text{cm}^{-2}\text{s}^{-1}$. Jet production at this energy has a very steep dependence on transverse momentum p_T. On the other hand large number of high p_T jets are expected to be seen at this luminosity. Creation of such a high statistics jet sample takes long time even at the generator level and hence it is not possible to produce events for the entire range of transverse momentum energy. The procedure followed here is to break the whole momentum range in small parts (called *bins*) and then to integrate all the bins to get a full distribution. For dividing the whole range into bins the same energy bins with appropriate trigger biases (for L1 and HLT) as well as the prescaling factors as proposed by the CMS JetMet group has been used.

The event shape variables in di-jet event sample are analysed taking into account the effect of trigger biases. These distributions will be compared with the expectations from

resummed calculations. This will give the sensitivity to measure α_s at the LHC and also the structure of underlying events.

The jets will be reconstructed at the detector using the information of the energy measurement with the electromagnetic and hadron calorimeters. The response at low energy is of a big concern and one needs to calibrate the jet energy scale with some correction from Monte Carlo. Currently there has been a considerate effort going on with the tuning of the Monte Carlo programme with test beam data using electron and hadron beams at a variety of energies. The energy response in the test beam is studied in order to understand the performance of the CMS calorimeter system.

List of Figures

List of Tables

Chapter 1

Introduction

Jet production is the dominant process in high-transverse-energy (high p_T) hadron-hadron collisions. This process is well described by perturbative QCD in terms of a point-like scattering cross section convoluted with a pair of parton distribution functions that express the momentum distribution of partons within the proton. The hard-scattering cross section itself can be written as an expansion in the strong coupling constant $\alpha_s(Q^2)$. The leading term in this expansion corresponds to the emission of two partons. The next term includes diagrams where an additional parton is observed in the final state due to gluon radiation (e.g. $q\bar{q} \to q\bar{q}g$). Such diagrams diverge when any of the three partons become soft or when two of the partons become collinear. Tree-level expressions can be used and compared directly to experiment for configurations where partons or jets are required to be energetic and well separated. These requirements avoid regions where singularities dominate the cross sections. Nevertheless, artifacts of the singularities in the theory can be found in the behaviour of the three-jet differential cross section.

The huge QCD cross section ensures that jets will dominate high-p_T physics at the LHC. Jets will not only provide a benchmark for understanding the detector, but will also serve as an important tool in the search for physics beyond the Standard Model. Event signatures for SUSY, Higgs boson production, compositeness, and other new physics processes require accurate reconstruction and measurement of jets coming from high-p_T quarks and gluons. The problems with associating a jet measured in a calorimeter with a scattered parton is an old, persistent problem in hadron collisions. Jet energy resolution and linearity are key factors in separating signal events from backgrounds. Missing transverse energy resolution is closely related to the calorimeter jet energy response.

The second chapter includes a brief introduction to the subject of perturbative QCD in the framework of the Standard Model with special emphasis on the parton model and calorimetric jet triggers. The jet algorithms, topology of jets and the event shape variables for inclusive dijet production are discussed in the third chapter. The fourth chapter gives an outline of the CMS calorimeter system and the response of the CMS hadron calorimeter for the test beam set up. The behaviour of QCD dijets at generator level and reconstructed level is presented in the fifth chapter. An analysis of the dijet mass distribution for a proposed trigger table is also presented in the same chapter.

Chapter 2

Theoretical background

A brief overview of Quantum Chromodynamics (QCD) is presented in this chapter with its place within the Standard Model.

2.1 Elements of the Standard Model

In the present understanding of particle physics, the elementary building-blocks of matter comprise spin-$\frac{1}{2}$ fermions and their antiparticles. The particles can be grouped into three families, each containing two quark flavours, one charged lepton and one neutrino, as shown in Table 2.1. In the absence of interactions, each fermion is represented by the quanta of a field ψ, satisfying the Dirac Equation

$$(i\gamma^\mu \partial_\mu - m)\,\psi \;\; = \;\; 0 \quad , \tag{2.1}$$

which corresponds to the Lagrangian density

$$\mathcal{L}_{\rm D} \;\; = \;\; \bar{\psi}\,(i\gamma^\mu \partial_\mu - m)\,\psi \quad . \tag{2.2}$$

The spinor ψ has four components, which collectively correspond to the helicity states of the particle and antiparticle.

Interactions are explained in the Standard Model by imposing *local gauge symmetries* on the fields: an example will be described in the next section. These symmetries require the existence of four vector fields, whose quanta are the spin-1 gauge bosons listed in

Quarks		Leptons	
Down (d)	Up (u)	Electron (e^-)	Electron neutrino (ν_e)
$Q = -1/3$	$Q = +2/3$	$Q = -1$	$Q = 0$
$m = 5.0$–8.5 MeV	$m = 1.5$–4.5 MeV	$m = 0.511$ MeV	$m < 3$ eV (95% CL)
Strange (s)	Charm (c)	Muon (μ^-)	Muon neutrino (ν_μ)
$Q = -1/3$	$Q = +2/3$	$Q = -1$	$Q = 0$
$m = 80$–155 MeV	$m = 1.0$–1.4 GeV	$m = 106$ MeV	$m < 0.19$ MeV (90% CL)
Bottom (b)	Top (t)	Tau (τ^-)	Tau neutrino (ν_τ)
$Q = -1/3$	$Q = +2/3$	$Q = -1$	$Q = 0$
$m = 4.0$–4.5 GeV	$m = 174 \pm 5$ GeV	$m = 1.78$ GeV	$m < 18$ MeV (95% CL)

Table 2.1: The elementary fermions of the Standard Model. The charge Q for each particle is expressed in units of the proton charge. The masses m are taken from Ref. [1].

	Charge	Mass (GeV)	Direct couplings to other particles						
			Quarks	$e^\pm$, $\mu^\pm$, $\tau^\pm$	ν_e, ν_μ, ν_τ	γ	$W^\pm$	Z^0	g
Photon (γ)	0	0	•	•			• †	†	
$W^\pm$ bosons	±1	80.4	•	•	•	• †	†	• †	
Z^0 boson	0	91.2	•	•	•	†	• †		
Gluon (g)	0	0	•						• †

Table 2.2: The gauge bosons of the Standard Model, and a summary of their interactions. Couplings indicated with bullets (•) arise from trilinear terms in the Lagrangian, while those with daggers (†) are quartic; when expressed in Feynman diagram notation, these correspond respectively to 'three-point' and 'four-point' vertices. The direct interaction of the photon and Z^0 boson is possible only via the quartic WWZγ vertex.

Table 2.2. Additional terms are introduced into the Dirac Lagrangian, leading to interactions between the fermions and gauge fields, and between the gauge fields themselves. The photon and the $W^\pm$ and Z^0 bosons are responsible for the electroweak interaction, and the gluon is the carrier of the strong interaction. Although a further gauge boson, the 'graviton' has been postulated, no complete theory of gravity currently exists within the Standard Model.

One additional particle, the scalar Higgs boson (H^0), is predicted to exist, but has not yet been observed conclusively. The Higgs field introduces a *spontaneously broken symmetry* into the Standard Model, thereby offering an explanation for the non-zero masses of the $W^\pm$ and Z^0 bosons. It also accounts for the mass terms in the Dirac Lagrangian, but does not predict the masses of the individual fermions.

Comprehensive discussions of the Standard Model can be found in many textbooks, such as Refs. [2, 3]. A more detailed treatment of QCD is given in Ref. [4].

2.2 The Lagrangian of the QCD

In QCD the six quark flavours are represented by quantum fields $q = \{u, d, s, c, b, t\}$, which behave identically, apart from their differing masses, and do not directly interact with one another. The quark fields have an extra degree of freedom known as colour; each of the three components q_a $(a = 1, 2, 3)$ is a Dirac spinor. Treating them as non-interacting fermion fields, the Dirac Lagrangian would therefore become

$$\mathcal{L} \;=\; \sum_a \bar{q}_a \left(i\gamma^\mu \partial_\mu - m\right) q_a \quad . \tag{2.3}$$

Under a unitary "phase transformation" applied to the three-component colour vector q, it becomes

$$q_a \;\rightarrow\; q'_a \;=\; \sum_b \Omega_{ab} q_b \;\equiv\; \sum_b \exp\left[i \sum_A \alpha^A \lambda^A_{ab}\right] q_b \;\; , \tag{2.4}$$

where the 3×3 Hermitian matrices λ^A $(A = 1, 2, \ldots 8)$ are the generators of the Lie group SU(3), and α^A are eight arbitrary constants. The Lagrangian given in Equation (2.3) is invariant under this global transformation, due to the unitary property of the matrices; this is analogous to the invariance of the Dirac Lagrangian under the phase transformation $\psi \rightarrow \psi' = \psi e^{i\phi}$.

The global colour transformation demonstrates the conservation of colour in a non-interacting theory, but does not introduce any physical dynamics. The theory of QCD is derived by requiring the invariance of the Lagrangian under *local* SU(3) colour transformations: instead of choosing the same unitary matrix, $\Omega = \exp[i \sum_A \alpha^A \lambda^A]$, at all points in space and time, the coefficients α^A are allowed to vary, giving

$$q_a \;\rightarrow\; q'_a \;=\; \sum_b \Omega_{ab}(x) q_b \;\equiv\; \sum_b \exp\left[i \sum_A \alpha^A(x) \lambda^A_{ab}\right] q_b \;\; . \tag{2.5}$$

Substituting this transformed quark field into Equation (2.3), the Lagrangian is found no longer to be invariant, because the space-time derivatives act on the coefficients $\alpha^A(x)$. To restore the invariance of the Lagrangian, the partial derivative ∂_μ should be first replaced

with a *covariant derivative*

$$(D_\mu)_{ab} = \partial_\mu \delta_{ab} + \frac{ig}{2} \sum_A \mathcal{A}_\mu^A \lambda_{ab}^A \quad , \tag{2.6}$$

where the eight *gauge fields* $\mathcal{A}^A$ are introduced, each with four space-time components μ; the free parameter g is a universal coupling constant. The Lagrangian now becomes

$$\begin{aligned} \mathcal{L} &= \sum_{a,b} \bar{q}_a \left(i\gamma^\mu D_\mu - m \right)_{ab} q_b & (2.7) \\ &\equiv \sum_a \bar{q}_a \left(i\gamma^\mu \partial_\mu - m \right) q_a + \frac{ig}{2} \sum_{a,b} \sum_A \bar{q}_a \left(\gamma^\mu \mathcal{A}_\mu^A \right) \lambda_{ab}^A q_b \quad . & (2.8) \end{aligned}$$

In the last line, $\mathcal{L}$ has been decomposed into two contributions: the first is the Dirac Lagrangian for three non-interacting components of a fermion field, and the second introduces interactions between the gauge fields and the quarks. The quanta of the eight fields $\mathcal{A}^A$ are called gluons, and are responsible for the observed strong interactions of quarks. To complete the process of establishing local gauge invariance, the transformation properties of the gluon fields must be chosen such that the covariant derivative $\sum_b (D_\mu)_{ab} q_b$ transforms in the same way as the quark field itself,

$$\sum_b (D'_\mu)_{ab} q'_b = \sum_{b,c} \Omega_{ab}(x) \, (D_\mu)_{bc} \, q_c \quad . \tag{2.9}$$

This is achieved with the relationship

$$\sum_A \mathcal{A}'^{\,A}_\mu \lambda^A = \Omega(x) \left[\sum_A \mathcal{A}_\mu^A \lambda^A \right] \Omega^{-1}(x) + \frac{2i}{g} (\partial_\mu \Omega(x)) \, \Omega^{-1}(x) \quad , \tag{2.10}$$

where the colour indices of the λ^A and $\Omega(x)$ matrices are suppressed.[1]

One further contribution must be inserted in the Lagrangian, to specify the equations of motion for the gluon fields. In quantum electrodynamics, the Lagrangian for the photon field A is given by

$$\mathcal{L}_{\text{photon}} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \quad , \tag{2.11}$$

where F is simply a quantised form of Maxwell's electromagnetic field strength tensor

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad . \tag{2.12}$$

[1] A simpler transformation law, of the form $\mathcal{A}'^{\,A} = \mathcal{A}^A + \delta \mathcal{A}^A$, exists when the gauge transformation $\Omega(x)$ differs only infinitesimally from from the identity matrix.

Applying the Euler-Lagrange Equations to $\mathcal{L}_{\text{photon}}$ gives the familiar Maxwell Equations, governing the internal dynamics of the field. An analogous term appears in the Lagrangian of QCD,

$$\mathcal{L}_{\text{gluon}} = -\frac{1}{4}\sum_{A} F^{A}_{\mu\nu}F^{\mu\nu}_{A} \quad , \tag{2.13}$$

but here the eight field strength tensors for the gluons are

$$F^{A}_{\mu\nu} = \partial_{\mu}\mathcal{A}^{A}_{\nu} - \partial_{\nu}\mathcal{A}^{A}_{\mu} - g\sum_{B,C} f^{ABC}\mathcal{A}^{B}_{\mu}\mathcal{A}^{C}_{\nu} \quad , \tag{2.14}$$

where the structure constants f^{ABC} are defined by the commutation relations of the SU(3) generators, $\left[\lambda^{A}, \lambda^{B}\right] = 2if^{ABC}\lambda^{C}$. The last term of Equation (2.14), which is derived by imposing local SU(3) gauge symmetry on the octet of gluon fields, arises because the gauge transformations of QCD do not commute. The expansion of the product $F^{A}_{\mu\nu}F^{\mu\nu}_{A}$ in Equation (2.13) gives rise to an array of terms containing products of two, three and four gluon fields. The three- and four-gluon terms in the Lagrangian are due to the self-interaction of the gluon field, which has no analogue in QED.

Collection of all terms together lead to the complete Lagrangian density of QCD: [2]

$$\mathcal{L}_{\text{QCD}} \;=\; \sum_{a}\bar{q}_{a}\left(i\gamma^{\mu}\partial_{\mu} - m\right)q_{a} \;+\; \frac{ig}{2}\sum_{a,b}\sum_{A}\bar{q}_{a}\left(\gamma^{\mu}\mathcal{A}^{A}_{\mu}\right)\lambda^{A}_{ab}\,q_{b} \;-\; \frac{1}{4}\sum_{A}F^{A}_{\mu\nu}F^{\mu\nu}_{A} \quad . \tag{2.15}$$

It is beyond the scope of this work to derive the Feynman rules associated with the QCD Lagrangian, or indeed to discuss the formal interpretation of Feynman diagrams; a list of the QCD Feynman rules can be found in Ref. [4]. However, a quick examination of the terms in the Lagrangian shows that the permitted vertices are as shown in Figure 2.1.

Up to this point, the coupling constant of QCD has been denoted g. From now on, however, the related quantity $\alpha_{\text{S}} = g^2/4\pi$will be used.

The idea of local gauge invariance under non-Abelian transformations was first proposed by Yang and Mills in 1954 [5], and has subsequently provided the foundation for both the strong and electroweak field theories of the Standard Model; the simpler Abelian case gave us quantum electrodynamics.

[2]When performing practical calculations, some further terms need to be inserted to fix the gauge. These are discussed in Ref. [4].

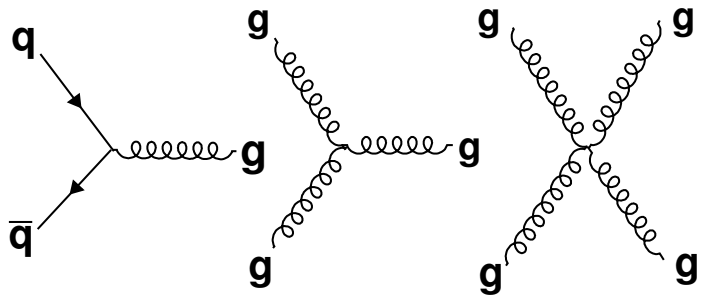


Figure 2.1: Feynman vertices in QCD

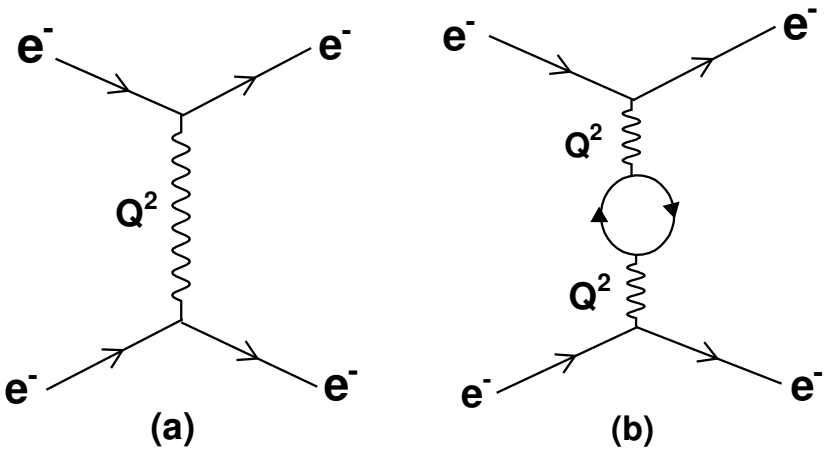


Figure 2.2: Leading QED contributions to the scattering of two electrons

2.3 Renormalisation and the Running Coupling

When using perturbation theory to calculate predictions for a physical observable in a quantum field theory, one must typically sum a series of Feynman diagrams corresponding to the appropriate initial and final states. In Figure 2.2, for example, diagram (b) is a higher-order contribution to the electromagnetic scattering process shown in diagram (a). Unfortunately, however, most Feynman diagrams with loops lead to divergent integrals; whereas diagram (b) should introduce a small correction to the cross section, it appears that it will instead cause a finite cross section to become infinite! This problem was first identified in QED, and was solved by the principle of renormalisation.

The apparent paradox of the infinite loop integrals can be addressed in QED by a

redefinition of the fine-structure constant, $\alpha_{\rm em}$. One must remember that no laboratory experiment can distinguish between the two processes shown in Figure 2.2, and therefore any observations of electromagnetic scattering will already 'include' the higher-order contributions. Remarkably enough, it can be shown that the effect of these corrections is to modify the effective *coupling strength* of the photon (albeit by an infinite factor), without altering its detailed properties as a photon. Hence no attempt is made to build Feynman diagrams out of the original 'bare' photons, but instead the 'ready-assembled' photons provided by nature are used, which include all of the loop diagrams such as Figure 2.2(b).

However, the loop integrals cannot be ignored altogether. Even though the higher-order diagrams contain divergent integrals, they remarkably have a *finite* dependence on the squared four-momentum Q^2 of the exchanged photon. Consequently, if two experiments study photons of different Q^2, they will observe different effective coupling strengths: photons of higher virtuality interact more strongly.

An analogous situation arises in QCD: the two loop diagrams in Figure 2.3 contain divergent integrals, which can be absorbed into a redefinition of the gluon coupling strength $\alpha_{\rm S}$. There is a very significant difference between the QCD and QED cases, however. The sign of the Q^2-dependence for the gluon self-interaction loop in Figure 2.3(c) is opposite to that for the fermion loops in Figures 2.2(b) and 2.3(b). As a consequence of this extra diagram, which is not present in QED, the strong coupling $\alpha_{\rm S}$ *decreases* at high virtualities. Conversely, $\alpha_{\rm S}$ grows without limit as the virtuality of the gluon approaches zero. Therefore quarks will appear to be "asymptotically free" in very high energy interactions, while at low energy scales their interactions are so strong that free quarks are never observed. This low-energy behaviour of QCD prevents the reliable use of perturbation theory at scales below a few GeV; in particular, the process of *hadronisation*, by which free quarks are converted into observable hadrons, is not well understood.

The scale-dependence of the "running coupling" $\alpha_{\rm S}(Q)$ is given by the Renormalisation Group Equation (RGE):

$$Q\frac{\partial\alpha_{\rm S}}{\partial Q} \;=\; 2\beta(\alpha_{\rm S}) \quad , \tag{2.16}$$

where the β-function is of the form

$$\beta(\alpha_{\rm S}) \;=\; -\frac{\beta_0}{4\pi}\alpha_{\rm S}^2 - \frac{\beta_1}{8\pi^2}\alpha_{\rm S}^3 - \frac{\beta_2}{128\pi^3}\alpha_{\rm S}^4 + \mathcal{O}(\alpha_{\rm S}^5) \quad . \tag{2.17}$$

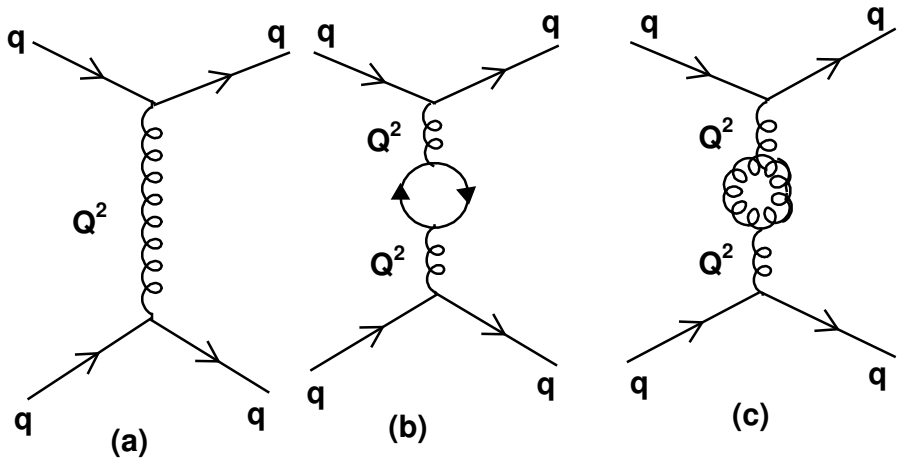


Figure 2.3: Leading QCD contributions to the scattering of two quarks

The coefficients β_n are functions of the number of kinematically accessible quark flavours, and are listed in Ref. [1]. To first order, the solution of the RGE for $N_{\rm f}$ flavours[3] can be written as

$$\alpha_{\rm S}(Q) = \alpha_{\rm S}(Q_0)\left[1 - \frac{\alpha_{\rm S}(Q_0)}{12\pi}(33 - 2N_{\rm f})\ln\left(\frac{Q^2}{Q_0^2}\right)\right]^{-1} \quad , \tag{2.18}$$

where $\alpha_{\rm S}(Q)$ and $\alpha_{\rm S}(Q_0)$ are the values of $\alpha_{\rm S}$ at two different scales; for the purposes of our analysis, however, we will use numerical solutions of the RGE with coefficients up to β_2 included.

When calculating the effects of divergent loop diagrams in a physical process, one must choose an energy scale μ at which to 'renormalise' the diagrams. In QED, for example,the Q^2-dependence of a divergent diagram can be calculated by comparing it with $Q = 0$ case. For the strong interaction, however, the scale $Q = 0$ would be an inappropriate reference point. Instead we choose a scale $Q = \mu$, close to the characteristic energy scale of the physical process. This "renormalisation scale" is an unphysical parameter, and plays no part in the QCD Lagrangian; hence physical observables are not expected to depend on it. However, as can be seen later in this chapter, the cancellation of μ from physical predictions will be incomplete unless the calculation itself is complete. Schematically, the prediction for an observable R can be written as

$$R(\alpha_{\rm S}) = R_{\rm known}(\alpha_{\rm S}, \mu) + R_{\rm unknown}(\alpha_{\rm S}, \mu) \tag{2.19}$$

[3]All of the QCD calculations applied in our analysis will use $N_{\rm f} = 5$.

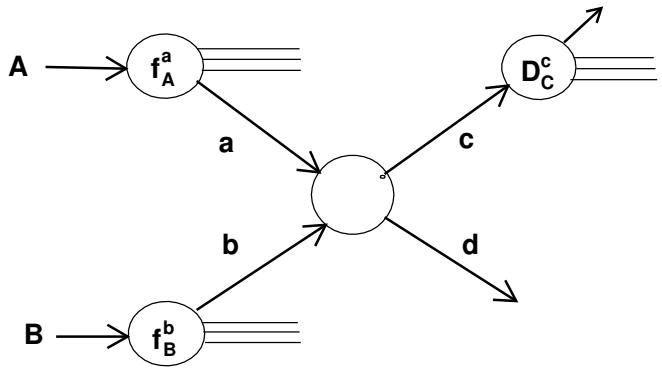


Figure 2.4: The hadronic interaction $AB \to CX$ at large p_T in terms of the parton sub-process $ab \to cd$, the structure functions f_A^a, f_B^b and the fragmentation function D_c^C.

where

$$\frac{\mathrm{d}R}{\mathrm{d}\mu} = 0 \ , \quad \text{but} \quad \frac{\mathrm{d}R_{\text{known}}}{\mathrm{d}\mu} \neq 0 \tag{2.20}$$

The magnitude of the derivative $\mathrm{d}R_{\text{known}}/\mathrm{d}\mu$ is often used to estimate the size of the unknown term, since it should approach zero as $R_{\text{known}} \to R$.

2.4 Parton Model for Large p_T Scattering

If hadrons are made up of more fundamental parton constituents (i.e. quarks and gluons) it must be possible to describe any hadronic reaction in terms of the interactions of these constituents. But this view point will obviously be most useful for those reactions in which the basic parton scattering process is fairly well separated in time from the more complex confinement effect, which prevents the partons from escaping as free particles and causes them re-assemble in hadrons. The production of hadrons which have a large momentum component (p_T) transverse to the beam direction is a good example of such a reaction [6].

The basic diagram for $AB \to CX$, where C is the hadron with large $\|p_T\|$ (say $p_T > 2$ GeV/c) while X represents all the other particles in the final state.

The incoming particles A and B contain, *inter alia*, partons a and b (respectively) which scatter, producing partons labelled c and d (which may often be same as a and

b) which have a large transverse momentum component q_T. Subsequently hadron C is produced from c via the confinement mechanism. Since q_T is the conjugate variable to the impact parameter of the parton scattering process, large q_T implies that the partons have scattered at a small distance where, following the theory of asymptotic freedom, α_s is small.

The cross section for the inclusive process $AB \to CX$ can be estimated with the following consideration. Let $f_A^a(x_a)$ be the probability that hadron A contains a parton a which is carrying a fraction $x_a = q_a/p_A$ of its momentum, $0 \leq x_a \leq 1$. Neglecting masses of both the hadrons and partons as being small compared to the momenta under consideration and also presuming that a does not have any momentum in directions transverse to the beam direction (the z axis), the four-momenta may be approximated as:

$$\begin{aligned} p_A &\simeq (p_A, 0, p_A) \\ q_a \simeq (q_a, 0, q_a) &= x_a(p_A, 0, p_A) \end{aligned}$$

since $p_A \gg m_A$, etc. The functions $f_A^a(x_a)$ are called the **structure functions** of A. Similarly the **fragmentation function** $D_c^C(z_c)$ can be defined to represent the probability that the outgoing parton c produces a hadron C carrying a momentum fraction $z_c = p_C/q_c$, where $0 \ll z_c \ll 1$. It has been assumed that C is produced collinearly with c and that the fragmentation process depends only on z_c and is independent of the nature of the initial state.

If the particle masses are neglected, the invariant variables for $AB \to CX$ are

$$\begin{aligned} s &= (p_A + p_B)^2 \simeq 2p_A \cdot p_B \\ t &= (p_A - p_C)^2 \simeq 2p_A \cdot p_C \end{aligned}$$

where $\sqrt{s}$ is the total centre-of-mass energy. The corresponding variables for the parton sub-process, $ab \to cd$, are

$$\begin{aligned} \bar{s} &= (q_a + q_b)^2 \simeq 2q_a \cdot q_a = 2x_a x_b p_A \cdot p_B \simeq x_a x_b s \\ \bar{t} &= (q_a - q_c)^2 \simeq -2q_a \cdot q_c = 2x_a p_A \cdot p_C/z_c \simeq x_a t/z_c \end{aligned}$$

The invariant cross section for $AB \to CX$ can be expressed as the weighted sum of

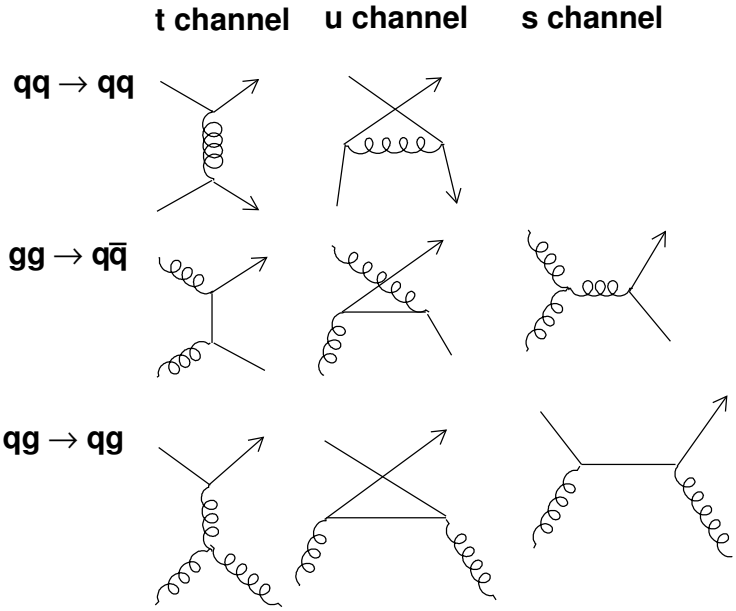


Figure 2.5: Basic QCD diagrams. Quarks, q, and gluons, g, are shown by straight and curly lines, respectively.

differential cross sections, $d\sigma/d\bar{t}$, of all possible parton scatterings that can contribute:

$$E_C \frac{d\sigma}{d^3 p_C}(AB \to CX) = \Sigma_{abcd} \int_0^1 dx_a \int_0^1 dx_b f_A^a(x_a) f_B^b(x_b) \frac{1}{\pi z_c} \frac{d\sigma}{d\bar{t}}(ab \to cd) D_c^C(z_c).$$

2.4.1 Parton Scattering Cross Section

The leading-order QCD diagrams for the basic $ab \to cd$ sub-process are shown in Figure 2.5. They are very similar to QED diagrams except for the replacement of α by α_s multiplied by the appropriate SU(3) colour factor.

Subprocess	$\lvert A\rvert^2$
$q_1 q_2 \to q_1 q_2$	$4/9(s^2+u^2)/t^2$
$q_1 \bar{q}_2 \to q_1 \bar{q}_2$	$4/9(s^2+u^2)/t^2$
$q_1 \bar{q}_1 \to q_2 \bar{q}_2$	$4/9(t^2+u^2)/s^2$
$q_1 q_1 \to q_2 q_2$	$4/9(\frac{s^2+u^2}{t^2} + \frac{s^2+t^2}{u^2}) - \frac{8}{27}\frac{s^2}{ut}$
$q_1 \bar{q}_1 \to q_2 \bar{q}_2$	as above with $s \leftrightarrow u$
$q\bar{q} \to gg$	$\frac{32}{27}(\frac{u^2+t^2}{ut}) - \frac{8}{3}(\frac{u^2+t^2}{s^2})$
$gg \to q\bar{q}$	$\frac{1}{6}(\frac{u^2+t^2}{ut}) - \frac{3}{8}(\frac{u^2+t^2}{s^2})$
$qg \to qg$	$-\frac{4}{9}(\frac{u^2+s^2}{us}) + (\frac{u^2+s^2}{t^2})$
$gg \to gg$	$\frac{9}{2}(3 - \frac{ut}{s^2} - \frac{us}{t^2} - \frac{st}{u^2})$

Table 2.3: Expressions for cross section for various partonic sub-processes

The cross sections listed in the Table 2.3, together with the structure and fragmentation functions, provide all the ingredients needed to obtain the cross sections for the hadronic process $AB \to CX$ based on Equation 2.21.

2.5 Monte Carlo Models

The QCD predictions, discussed so far, have been concerned with free quarks and gluons. Before reaching the detector, however, these partons must 'fragment' into bound colourless states. This final phase of the QCD interaction, which occurs at a low characteristic energy scale, cannot be predicted by perturbation theory. Numerical simulations based on semi-empirical models are used instead. The Monte Carlo program used in this analysis to simulate multi-hadron events is PYTHIA [7].

The simulation of each event proceeds in four distinct stages. The procedure is illustrated schematically in Figure 2.6:

- generation of a pair of partons through a hard scattering process from initial state partons which are taken from the protons using the parton density function, possibly with initial-state radiation
- a "parton shower", in which gluons are radiated from the parton pair; the gluons may then radiate other gluons, or split to form new $q\bar{q}$ pairs. This stage should reproduce as closely as possible the predictions of perturbative QCD.
- fragmentation of the parton system into hadrons
- decays of short-lived hadrons such as the π^0, K^0_S, $\cdots$.

The implementation of these stages in the Monte Carlo programs are briefly described below; not all stages are implemented by all programs.

2.5.1 The Parton Shower

In the PYTHIA [7] program, the perturbative stage of the event is simulated using a numerical implementation of the DGLAP equations [8] which are used in the derivation of the NLLA analytical predictions as well. The progress of each parton in the cascade

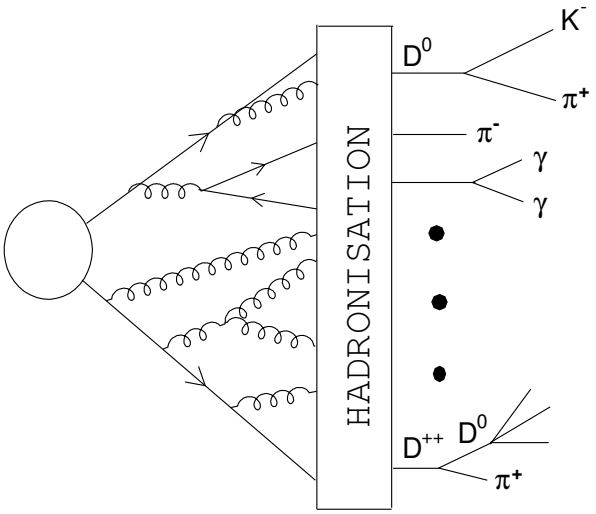


Figure 2.6: A schematic view of the stages to be simulated after a hard scattering process between two partons.

is parameterised by an "evolution variable", $t = \ln(Q^2/\Lambda^2)$, where Q^2 is related to the virtuality of the parton, and Λ is some arbitrary fixed scale. The probability $\mathcal{P}$ for a parton a of virtuality Q^2 to split into two partons b and c, carrying momentum fractions z and $(1-z)$ respectively, can then be written in the form

$$\frac{\mathrm{d}\mathcal{P}}{\mathrm{d}t} \propto \int_0^1 \mathrm{d}z \, \frac{\alpha_\mathrm{S}}{2\pi} \, P_{a\to bc}(z) \qquad [0 < t < t_0] \quad , \tag{2.21}$$

where t_0 and t are the evolution variables before and after the splitting, and $P_a \to bc$ are the Altarelli-Parisi splitting kernels for the processes $q \to qg$, $g \to gg$, $g \to q\bar{q}$ and $q \to q\gamma$. The precise definition of the scale Q^2 varies between models: PYTHIA uses $Q^2 = m_a^2$ where m_a^2 is the invariant mass-squared of the parton a.

The models include various features to incorporate coherence effects,[4] the simplest of which is the property of *angular ordering*: the opening angle θ_{bc} of each branching $a \to bc$ is required to be less than that of the previous branch in the cascade. This result can be explained [4] in terms of the Uncertainty Principle. Other effects include correlations between the azimuthal branching angles, due to gluon polarisation.

The shower is continued until the virtuality of the partons reaches some lower limit $Q_0 \sim 1$ GeV, which is a tunable parameter of the model. The final parton configuration

[4]The acronym 'HERWIG' stands for "Hadron Emission Reactions With *Interfering Gluons*".

is then passed to the non-perturbative hadronisation stage. More detailed discussions of parton shower physics are given in Refs. [4] and [9].

2.5.2 Fragmentation

Hadronisation is simulated in PYTHIA using the *Lund string model* [10]. Unlike the electromagnetic field patterns formed by distributions of charges and currents, the corresponding fields in QCD are expected to be confined in narrow regions stretched between the colour charges; this is a result of the gluon's self-coupling property. According to the string model, the field lines will eventually 'break' at several points to form new $q\bar{q}$ or diquark-antidiquark pairs which lead to meson and baryon production. The model has many tunable parameters, which have been chosen to give optimal agreement with the LEP data.

2.5.3 Decays

Some of the hadrons produced in these collisions are expected to decay very close the interaction point; the experiments will then measure event shape observables based on the daughter particles. For convenience, all particles with lifetimes less than 3×10^{-10} s are considered to be unstable. PYTHIA simulates the decays of these particles based on standard branching ratios supplied by the Particle Data Group [1].

2.6 Jet Production

2.6.1 Jets in Hadron Scattering

The large p_T hadron scattering processes are predominantly 4-jet events. Two of the jets contain fragments of the incoming hadrons and hence continue more or less in the forward direction, while the other two wide-angle back-to-back jets, stemming from the hard scattered partons, contain particles with large p_T. Thus accompanying the large p_T particle which usually triggers the detectors to notify that a large p_T event has occurred, there are other particles, travelling essentially in the same direction, forming a "same-side" jet, and in the opposite hemisphere an "away-side" jet. The longitudinal jets appear

to be the same as those observed in ordinary hadron scattering when no large p_T particle is produced.

2.6.2 Calorimetric Jet Triggers

As mentioned above in a 4-jet event, large p_T hadrons are expected to be produced. Two of these jets result from hard, large-angle, parton scattering while the others contain the fragments of the beam particles. Due to the fragmentation process, the parton energy is shared by a large number of particles (average multiplicity increasing with parton energy). As a result, a high-p_T jet has only a small probability of producing a high p_T particle. This is called the 'trigger bias' effect. So if one identifies jets by looking for large p_T particles one will inevitably miss most of the interesting jet cross section.

A more effective way of finding these high p_T jets at high energy is to trigger instead on the occurrence of a large amount of energy (summed over all particles) deposited in calorimeter at a fairly large angle to the beam direction. It is useful to introduce the transverse energy of particle i defined by

$$E_{iT} = E_i \sin\theta_i$$

where E_i is its energy and θ_i is the centre-of-mass angle between its direction of motion and the beam axis. Then the total transverse energy of an event is given by

$$E_T \;=\; \Sigma_i E_{Ti}$$

summed over all the particles in the event.

Chapter 3

Observables

3.1 Jets in Hadron Colliders

Hadron collisions at the Tevatron are a prolific source of high transverse momentum jets. The collider experiments, CDF and DØ, have made very high precision measurements of the inclusive jet rates [11, 12], as well as the rate of n-jet events, $n \geq 2$ [13, 14]. At present, perturbative calculations exist to next-to-leading order ($\mathcal{O}(\alpha_s^2)$) only for the inclusive [15–17] and dijet [18,19] rates. The next-to-leading order (NLO) calculation for three-jet rate is close to completion [20]. These allow quite precise studies of QCD, albeit for very inclusive quantities.

At NLO, the cross sections begin to depend on the exact definition of the jets, since the jets begin to develop internal structure. The experiments have used measures of this internal structure as a cross-check on the reliability of the calculations, as well as as a study of QCD in its own right. In fact these more exclusive event properties contain considerably more information about QCD dynamics, and make an ideal place to study QCD.

One of the most popular measures of jet structure in hadron collisions is the 'jet shape': the fraction of the jet's energy within a cone of a given size, centred on the jet direction. This has been measured by both CDF and DØ over wide ranges of phase space [21, 22].

3.1.1 Jets

All the algorithms discussed here define the momentum of a jet in terms of the momenta of its constituent particles in the same way, inspired by the Snowmass accord [23]. The transverse energy, $E_{T\text{jet}}$, pseudo rapidity, η_{jet}, and azimuth, ϕ_{jet}, of the jet are given by:

$$\begin{aligned} E_{T\text{jet}} &= \sum_{i \in \text{jet}} E_{Ti}, \\ \eta_{\text{jet}} &= \sum_{i \in \text{jet}} E_{Ti}\,\eta_i / E_{T\text{jet}}, \\ \phi_{\text{jet}} &= \sum_{i \in \text{jet}} E_{Ti}\,\phi_i / E_{T\text{jet}}. \end{aligned} \tag{3.1}$$

where E_{Ti}, η_i, ϕ_i are transverse energy, pseudo rapidity and azimuthal angle of the constituent particle i. Boost-invariant variables are only used here, so whenever the variable 'angle' is mentioned, this will refer to the Lorentz-invariant opening angle $R_{ij} = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}$. Also, the term 'energy' will refer to the mean transverse energy, $E_T = E \sin\theta$, where θ is the polar angle.

3.1.2 Jet Algorithms

The first jet algorithms for hadron physics were simple cones. Over the last two decades, clustering techniques have greatly improved in sophistication. Three principal jet reconstruction algorithms have been coded and studied for CMS: the iterative cone, the midpoint cone, and the inclusive k_T jet algorithms. The midpoint-cone and k_T algorithms are widely used in offline analysis in current hadron collider experiments, while the iterative cone algorithm is simpler and faster and commonly used for jet reconstruction in trigger systems.

The jet algorithms are used with a specific recombination scheme for adding the four-momenta of the constituents. In the energy scheme, four-momenta of the constituents are simply added as four-vectors. This produces massive jets. In the E_T scheme, jets are assumed massless. Here the jet transverse momentum is obtained from the ΣE_T of the constituents and then the direction of the jet is fixed in one of the two ways: (i) $\sin\theta = \Sigma E_T / E$ where E is the jet energy (usually used with cone algorithms), or (ii) $\eta = \Sigma E_{Ti}\eta_i / \Sigma E_T$ and $\phi = \Sigma E_{Ti}\phi_i / \Sigma E_T$ (usually used with the k_T algorithm). In all cases

the jet E_T is equal to $c \cdot p_T$.

The inclusive k_T algorithm merges, in each iteration step, input objects into possible final jets and so the new jet quantities, the jet direction and energy, have to be calculated directly during the clustering. The cone jet algorithms, iterative and midpoint, group the input objects together as an intermediate stage and the final determination of the jet quantities (recombination) is done in one step at the end of the jet finding.

Both algorithms, applied for LHC, use an angular distance measure based on the azimuthal angle ϕ and, instead of the pseudo-rapidity η, the true rapidity $y = 0.5 \ln(E + p_z)/(E - p_z))$ which has become an established standard in recent publications. The distance between two objects i and j hence reads

$$\Delta R_{ij} = \sqrt{(\Delta_{ij}\phi)^2 + (\Delta_{ij}y)^2}$$

In addition, the most frequently used recombination scheme, the E scheme, implying a simple four-momentum addition, is employed.

3.1.2.1 k_T Algorithm

The fully inclusive k_T algorithm including an R parameter [24] is discussed here. It clusters particles (partons or calorimeter cells) according to the following iterative steps:

1. for every pair of particles, a closeness variable is defined

$$d_{ij} = \min(E_{Ti}, E_{Tj})^2 R_{ij}^2 \left(\approx \min(E_i, E_j)^2 \theta_{ij}^2 \approx k_T^2\right). \tag{3.2}$$

2. for every particle, a closeness variable to the beam particles is defined

$$d_{ib} = E_{Ti}^2 R^2. \tag{3.3}$$

3. if $\min\{d_{ij}\} < \min\{d_{ib}\}$, particles i and j are *merged* according to Equation (3.1) (other merging schemes are also possible [25]).

4. if $\min\{d_{ib}\} < \min\{d_{ij}\}$, jet i is declared *complete*.

These steps are iterated until all jets are complete. In this case, all opening angles within each jet are $< R$ and all opening angles between jets are $> R$.

3.1.2.2 Iterative Cone Algorithm

In the iterative-cone concept with cone radius R, particles are clustered into jets according to the following steps:

1. the particles are passed through a calorimeter with cell size $\delta\eta \times \delta\phi$ (in LHC, $\delta\eta = \delta\phi = 0.087$). In the parton-level algorithm, this is simulated by clustering together all partons lying within a cone of size $\delta\eta \times \delta\phi$ of each other.

2. every calorimeter cell (cluster) with energy above E_0, is considered as a 'seed cell' for the following step (in LHC, $E_0 = 1$ GeV).

3. a jet is defined by summing all cells within an angle R of the seed cell according to Equation (3.1).

4. if the jet direction does not coincide with the seed cell, step 3 is reiterated, replacing the seed cell by the current jet direction, until a stable jet direction is achieved.

5. a long list of jets is thus obtained, one for each seed cell. Many entries in this list are duplicates: these are removed[1].

6. some jets could be overlapping. Any jet that has more than 50% of its energy in common with a higher-energy jet is merged with that jet: all the cells in the lower-energy jet are considered part of the higher-energy jet, whose direction is again recalculated according to Equation (3.1).

7. any jet that has less than 50% of its energy in common with a higher-energy jet is split from that jet: each cell is considered part only of the jet to which it is the nearest.

It is to be noted that despite the use of a fixed cone of radius R, jets can contain energy at angles greater than R from their direction, because of step 6. This is not a particular problem. This is essentially also the algorithm used by ZEUS [26], except that their merging/splitting threshold is 75% instead of 50%.

[1] In DØ, any candidate jet with energy below 8 GeV are also thrown away. For jets above 16 GeV, this makes only a small numerical difference, which is not important and so these jets are kept.

3.1.3 Infrared and Collinear Safe

The formal definition of "infrared safe"s is as follows:

> *An observable is infrared safe if, for* any *n-parton configuration, adding an infinitely soft parton does not affect the observable at all.* (3.4)

The formal definition of "collinear safe" is similar:

> *An observable is collinear safe if, for* any *n-parton configuration, replacing any massless parton by an exactly collinear pair of massless partons does not affect the observable at all.* (3.5)

Depending on its behaviour under these conditions, three types of observable can be classified:

Safe: Both the above conditions are fulfilled. Cross sections are calculable order by order in perturbation theory, with hadronisation effects resulting in only power-suppressed corrections.

Unsafe: One or both the conditions is violated. Cross sections are infinite in perturbation theory, signalling that they are fundamentally non-perturbative.

Almost unsafe: At first sight it appears that one or both the conditions is violated, but in fact close inspection of the algorithm shows that seemingly-minor details actually rescue its safeness. Cross sections are calculable order by order in perturbation theory, but can be rather unreliable, owing to the reliance on (typically) small parameters.

3.1.4 Jet Cross Sections

The fully inclusive jet cross section is one of the simplest quantity to be calculated:

$$\sigma(p_{T_{\min}}) \equiv \int_{p_{T_{\min}}}^{\infty} dp_T \left.\frac{d\sigma}{dp_T\, d\eta}\right|_{\eta=0}. \tag{3.6}$$

Defining a jet through iterative cone algorithm may give rise to ambiguity if the cones from the two partons overlap. In the case of two hard partons (as shown in Figure 3.1a),

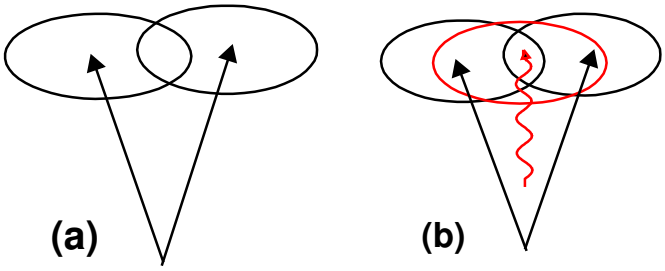


Figure 3.1: Illustration of the problem region for the iterative cone algorithm. In (a), there are two hard partons, with overlapping cones. In (b) there is an additional soft parton in the overlap region.

the two cones overlap, but there is no energy in the overlap region. So the splitting procedure is trivial, and it is classified as a two-jet configuration.

If the same event is considered with the addition of a soft parton, close to the energy threshold E_0, illustrated in Figure 3.1b and if the soft jet is marginally below threshold, the classification will be as above, with the soft parton being merged with whichever hard parton it is nearest. If on the other hand it is marginally above threshold, there is an additional seed cell. The cone around this seed encloses both the hard partons and thus a third stable solution is reached. Now the merging and splitting procedure produces completely different results. In either the CDF or DØ variants the result is the same: each of the outer jets overlaps with the central one, with the overlap region containing 100% of the outer one's energy. Thus each is merged with the central one, and it is classified as a one-jet configuration.

3.2 Event Shape Variables

Many variables can be defined which are sensitive to the radiation of hard gluons from quarks. Observables which describe the topology of hadronic events are known as event shape variables. In order to be finite order-by-order in perturbation theory, after renormalisation of ultraviolet divergences, an observable should be both infrared and collinear safe. This implies that the observable should be insensitive to the emission of soft particles and to the splitting of one particle into collinear ones. In other words, the observable

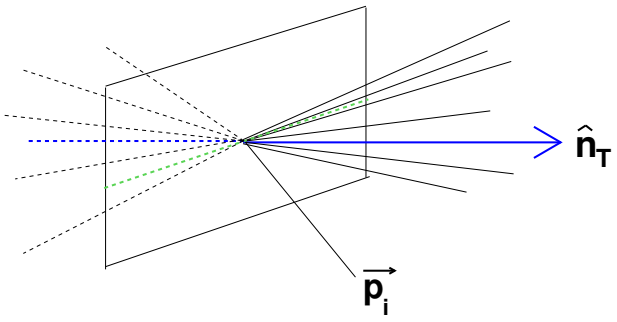


Figure 3.2: Orientation of the two hemispheres with respect to the thrust axis.

should depend linearly on final state momenta. Apart from the theoretical arguments, infrared and collinear safe variables are also preferred for experimental reasons. They allow calorimetric measurements, and adding a soft particle or splitting a particle into two with half the energy, changes the measurement in a continuous way.

Different event shape variables have different sensitivities to fragmentation and higher order perturbative effects. Therefore, an important estimate of theoretical uncertainties can be made by comparing the values of a physical parameter, $\alpha_S$for instance, derived from different observables.

3.2.1 Thrust, Major, Minor and Oblateness

Thrust (T) is defined by

$$T = \max_{\hat{\boldsymbol{n}}} \left(\frac{\sum_i |\boldsymbol{p}_i \cdot \hat{\boldsymbol{n}}|}{\sum_i |\boldsymbol{p}_i|} \right) \quad , \tag{3.7}$$

where the thrust axis $\hat{\boldsymbol{n}}_T$ is defined as the unit 3-vector $\hat{\boldsymbol{n}}$ which maximises the expression.

For a perfectly 'pencil-like' two-jet event, the thrust axis lies parallel to the jets, so $|\boldsymbol{p}_i \cdot \hat{\boldsymbol{n}}_T| = |\boldsymbol{p}_i|$, yielding $T = 1$. In the case of a 'spherical' event, with an infinite number of particles distributed isotropically in the final state, the thrust becomes a ratio of solid angle integrals:

$$T = \frac{\int |\cos\theta| \, \mathrm{d}\Omega}{\int \mathrm{d}\Omega} = \frac{2\pi}{4\pi} = \frac{1}{2} \quad . \tag{3.8}$$

It can be shown that all events satisfy $\frac{1}{2} < T < 1$. All the other event shapes

considered here approach zero in the two-jet limit; for consistency, we will therefore define the observable $y = 1 - T$ which shares this property.

The concept of thrust was already in use before the advent of QCD. In 1964 [27], a "principal axis" equivalent to $\hat{\boldsymbol{n}}_{\boldsymbol{T}}$ was proposed for the analysis of jets observed in hadron collisions, though the origin of the jets was hitherto unexplained. Later, in 1977 [28], it was recognised that this "maximum directed momentum" represented a calculable quantity in perturbative QCD.

Thrust major ($\boldsymbol{T_{\text{maj.}}}$) is defined in the same way as thrust, except that the axis $\hat{\boldsymbol{n}}$ is constrained to be orthogonal to the thrust axis:

$$T_{\text{maj.}} \;=\; \max_{\hat{\boldsymbol{n}} \perp \hat{\boldsymbol{n}}_T} \left(\frac{\sum_i |\boldsymbol{p}_i \cdot \hat{\boldsymbol{n}}|}{\sum_i |\boldsymbol{p}_i|} \right) \quad . \tag{3.9}$$

The axis which maximises the quantity in parentheses is $\hat{\boldsymbol{n}}_{\boldsymbol{T}_{\text{maj.}}}$.

Thrust minor ($\boldsymbol{T_{\text{min.}}}$) is analogous to T and $T_{\text{maj.}}$ except that $\hat{\boldsymbol{n}}$ is orthogonal to both $\hat{\boldsymbol{n}}_{\boldsymbol{T}}$ and $\hat{\boldsymbol{n}}_{\boldsymbol{T}_{\text{maj.}}}$:

$$T_{\text{min.}} \;=\; \frac{\sum_i |\,\boldsymbol{p}_i \cdot \hat{\boldsymbol{n}}_{\boldsymbol{T}_{\text{min.}}}|}{\sum_i |\boldsymbol{p}_i|} \quad , \quad \text{where } \hat{\boldsymbol{n}}_{\boldsymbol{T}_{\text{min.}}} \;=\; \frac{\hat{\boldsymbol{n}}_{\boldsymbol{T}} \times \hat{\boldsymbol{n}}_{\boldsymbol{T}_{\text{maj.}}}}{\left| \hat{\boldsymbol{n}}_{\boldsymbol{T}} \times \hat{\boldsymbol{n}}_{\boldsymbol{T}_{\text{maj.}}} \right|} \tag{3.10}$$

Oblateness ($\boldsymbol{O}$) is simply the difference between thrust major and thrust minor:

$$O \;=\; T_{\text{maj.}} - T_{\text{min.}} \tag{3.11}$$

The thrust major and thrust minor are both zero for a perfect two-jet event, since all particle momenta are parallel to the thrust axis, and hence orthogonal to $\hat{\boldsymbol{n}}_{\boldsymbol{T}_{\text{maj.}}}$ and $\hat{\boldsymbol{n}}_{\boldsymbol{T}_{\text{min.}}}$. Furthermore, the thrust minor is zero for a three-jet event,[2] since momentum conservation dictates that all particles must lie in the plane orthogonal to $\hat{\boldsymbol{n}}_{\boldsymbol{T}_{\text{min.}}}$. For a spherical event, both $T_{\text{maj.}}$ and $T_{\text{min.}}$ approach a maximum value of 1/2. The oblateness is unusual, in that it vanishes for both two-jet events and spherical events.

When a three-jet event approaches the two-jet limit, as depicted in Figure 3.3, the four observables have differing sensitivities to the opening angle θ:

[2]These statements are only valid for perfectly narrow jets, or partons. In reality, hadronisation and particle decays introduce some transverse momentum within the jet.

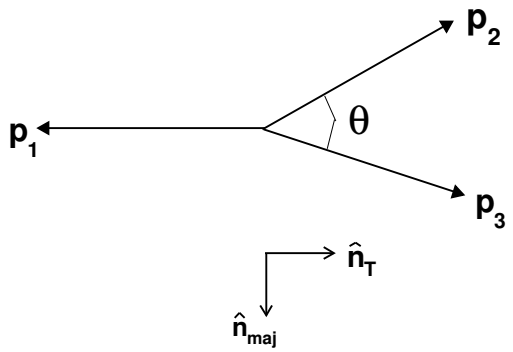


Figure 3.3: A three-jet event approaching the two-jet limit

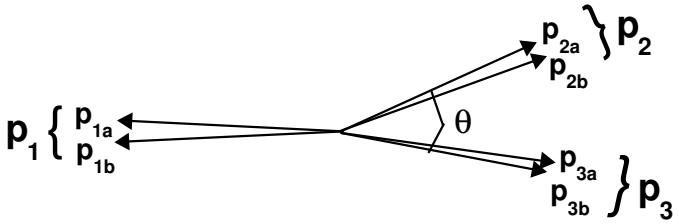


Figure 3.4: Collinear branching of partons within a three-jet event

$$
\begin{aligned}
1 - T &\sim 1 - \cos\theta &&\sim \theta^2 \\
T_{\text{maj.}} &\sim \sin\theta &&\sim \theta \\
T_{\text{min.}} &= 0 \\
O &= T_{\text{maj.}} &&\sim \theta
\end{aligned}
$$

Hence the thrust major is expected to be affected to a greater extent than the thrust by the effects of hadronisation and particle decays.

All four observables are linear with respect to the momenta $\boldsymbol{p}_i$. Hence if one particle splits into two, such that both final state particles continue on the same trajectory as shown in Figure 3.4, the observables do not change. In perturbative QCD, the matrix element for a process involving soft or collinear gluon emission becomes infinite in the limit of small momentum transfers. This leads to "infrared divergences" in quantities which are sensitive to such processes. Thus distributions for observables such as those considered above, which are "infrared safe" can be reliably calculated.

3.3 Hadronic Event Shape and Jet Productions

Hadronic event shapes are tools to study both the amount of gluon radiation and the details of hadronisation process. Events from $q\bar{q}$ final states without hard gluon radiation result in two collimated back-to-back jets of hadrons. Emission of one hard gluon leads to planar 3-jet events. Emission of two or more energetic gluons can cause non-planar multi-jet event structures.

Scaling violations for observables that do not depend on absolute energies or momenta, e.g., *thrust*, can be calculated using perturbative QCD to a high degree of accuracy. Scaling violations caused by the energy dependence of α_s which determines the amount of gluon radiation. In the leading order (LO) the probability of gluon radiation is proportional to α_s.

3.4 Relative Contributions from n-jet Configuration

Hadronic event topologies are characterised by hard gluon radiations. These different jet topologies appear differently in the distributions of the event shape variables. Relative

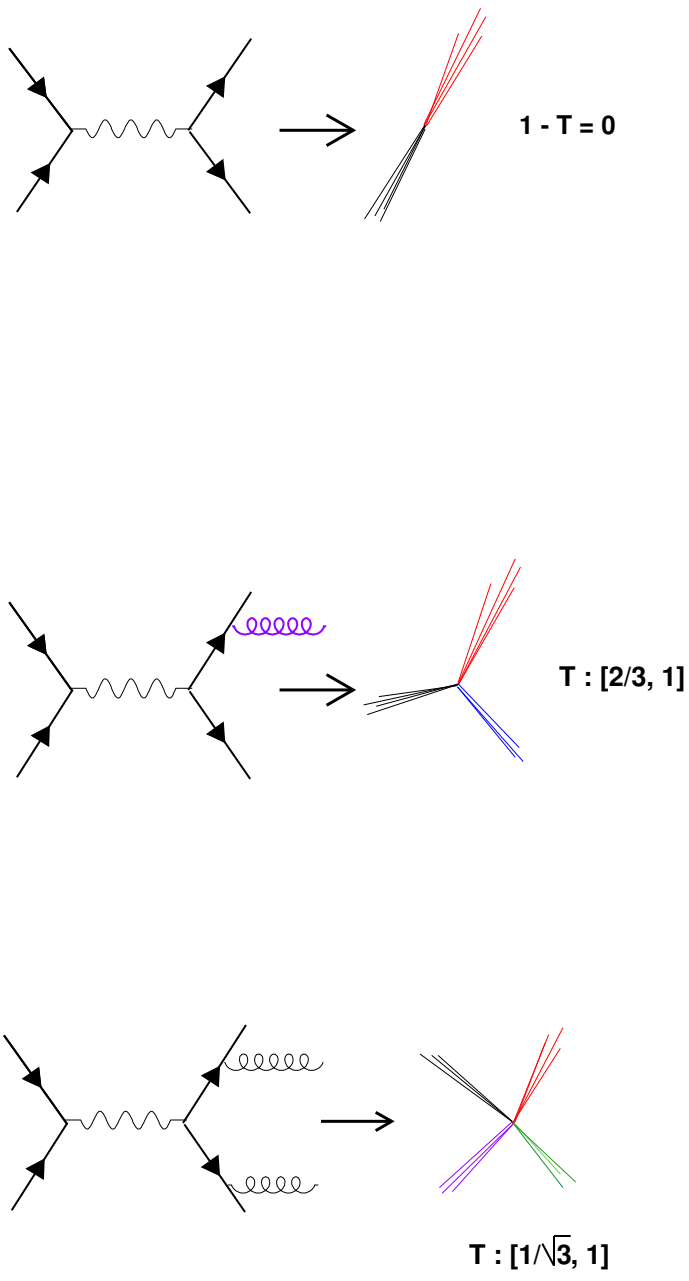


Figure 3.5: Two, three and four jets contributions to event shapes

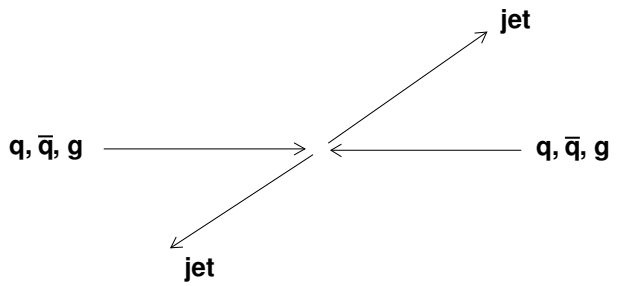


Figure 3.6: Possible $2 \rightarrow 2$ processes.

contributions to the event shapes for two-jet, three-jet and four-jet topologies are shown in Figures 3.5.

3.5 $2 \rightarrow 2$ Hard Scattering Processes

Figure 3.6 presents the decomposition of the total jet cross section into the partonic processes for $p\bar{p}$ collisions at the Tevatron and pp collisions at the LHC in dependence of the scaling variable $x_T = 2p_T/\sqrt{s}$, and illustrates the difference in cross section contributions of the PDF's compared to measurements possible today.

These kinds of two going to two processes (Figure 3.6) are being analysed:

$$\begin{aligned} f_i f_j &\rightarrow f_i f_j \\ f_i \bar{f}_i &\rightarrow f_k \bar{f}_k \\ f_i \bar{f}_i &\rightarrow gg \\ f_i g &\rightarrow f_i g \\ gg &\rightarrow f_k \bar{f}_k \\ gg &\rightarrow gg \end{aligned}$$

here f's represent the leptons and g's the gluons. These are all possible reactions in such process. These are also the processes where hard scattering processes are studied.

Chapter 4

CMS Detector

4.1 The CMS Calorimeter System

The Compact Muon Solenoid (CMS) is a general purpose detector at the Large Hadron Collider (LHC) capable of discovering, and studying, the nature of electroweak symmetry breaking for which the Higgs mechanism is presumed to be responsible and to shed light on the issue of the mathematical consistency of the Standard Model at energy scales above 1 TeV.

The overall layout of CMS is shown in Figure 4.1. The overall dimensions of the CMS detector are a length of 21.6 m, a diameter of 14.6 m and a total weight of 12500 tons. At the heart of CMS [29] sits a 13 m-long, 5.9 m diameter, 4 T superconducting solenoid. In order to achieve good momentum resolution within a compact spectrometer without making stringent demands on muon-chamber resolution and alignment, a high magnetic field has been chosen. The return field is large enough to saturate 1.5 m of iron, allowing four muon stations to be integrated to ensure robustness and full geometric coverage. The whole detector has the structure of a cylinder (called the barrel region) and two end-caps in both end of the cylinder. The bore of the magnet coil is large enough to accommodate the inner tracker and the calorimeters inside.

CMS calorimeter system [30] consists of two kinds of detectors - a crystal calorimeter to detect and measure energies of electromagnetic particles (photons and electrons) and a sampling calorimeter to absorb the hadrons and measure their energies.

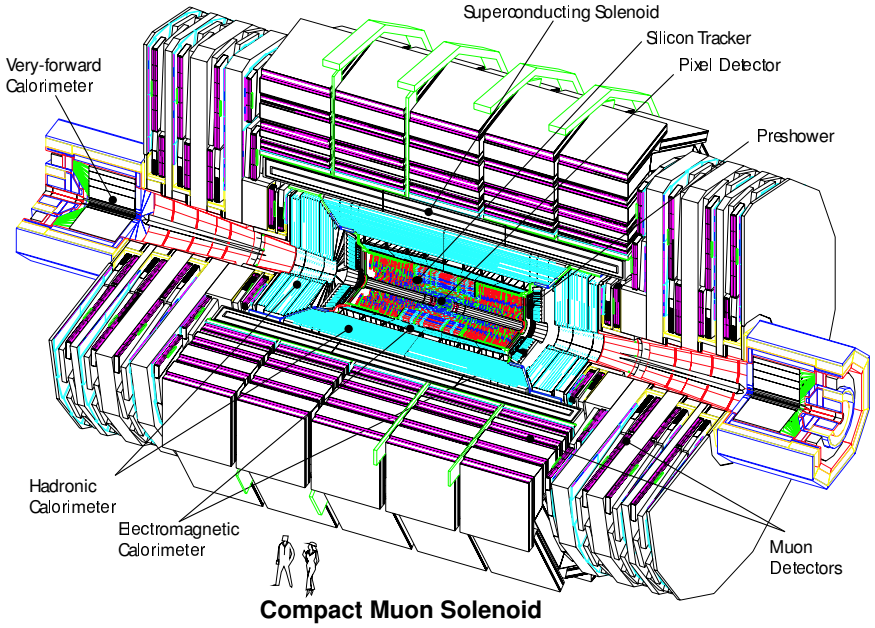


Figure 4.1: A view of the CMS detector

The CMS Electromagnetic Calorimeter [31] consists of lead tungstate ($PbWO_4$) crystals which provide a good energy and position resolution for electrons and photons.

The CMS Hadron Calorimeter [32] consists of a barrel detector (HB) and an endcap (HE) detector. It uses plastic scintillator as the active material and a copper alloy (90% Cu and 10% Zn) as the absorbing material. The alloy has an interaction length of $\sim 15cm$. Mechanically HB is a cylindrical structure made of connecting 18 calorimeter wedges into half barrels. Each wedge subtends an angle of 20° in ϕ and extends in z by 4.3 meters from the CMS mid-plane. There are 17 active layers of scintillator trays. Layer 1 and layer 17 are 8 mm thick and the remaining layers are each 4 mm thick. Two layers of scintillator trays, each 10 mm thick are added as an outer hadron calorimeter (HO)..

The scintillator light is collected from the trays using wavelength shifting fibres and is transmitted to Hybrid Photo diodes (HPDs) using optical fibres. HB/HO is divided into two longitudinal readout segments. Layers 1 to 17 are added to H1 and HO layers (18-19) to H2. Granularity of readout is $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$. The digitisation of the analog signal is done at the beam crossing frequency of 40 MHz by QIE chips (Charge(Q) Integration(I) range Encoding(E)).

4.2 The 2004 Test Setup for CMS HCal

The 2004 CMS HCal test beam experiment is a detector setup designed to study the performance of the CMS hadronic calorimeter. The detector is exposed to beams of pions (3, 5, 7, 9, 20, 30, 50, 100, 300 GeV/c), electrons (, 20, 30, 50, 100 GeV/c), and muons (80, 120 GeV/c) over a large energy range. The TB2004 setup consists of:

- 2 wedges of HCal Barrel (HB1 and HB2)
- 2 sectors of HCal Endcap
- mock-up of cables going within the crack between barrel and endcap.
- 6 trays of HO detectors in each of the 3 mock-up rings
- tail catcher iron
- 7×7 crystal matrix
- mock-up of material between ECal and HCal
- beam line trigger counters.

4.2.1 Detector Configuration

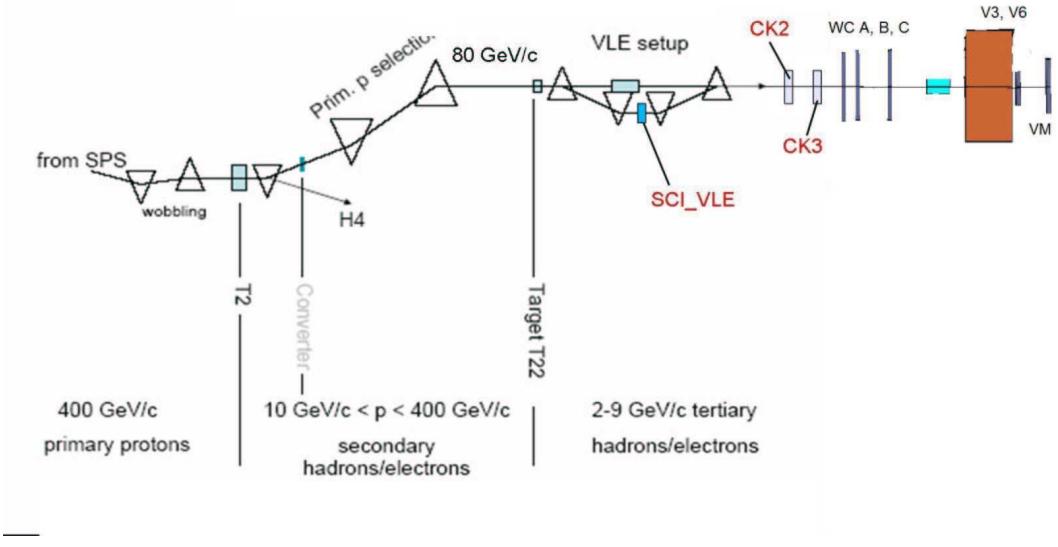


Figure 4.2: The set up for the test beam 2004

A schematic picture of experimental setup is shown in Figure 4.2. The primary beam from SPS is let fall on the target T2 which produces secondary beam (10-300 GeV). For high energy runs this secondary beam is used. In order to produce low energy beam the secondary beam is made to fall on the target T22, as shown in the figure. T22 produces tertiary hadrons/electrons in an energy range 2-9 GeV/c. The shielding is kept for checking the high energy beams while taking the Very Low Energy (VLE) runs. To ensure the purity of the beam, various veto counters and Cerenkov counters, CHN(1-3), are kept. CHN2 is used to distinguish between electrons and pions at low energies. It operates with a certain threshold energy. CHN3 is used for separating out pions from muons when the beam coming is of low energy. Three downstream veto counters (V3, V6, VM) are kept to distinguish muons. Several trigger counters (S1, S2, S3, S4) are present in the setup to decide the beam profile. The Hadron Barrel (HB) is made of calorimeter wedges connected to each other to form the shape of a half barrel, namely HB1 and HB2. Grouping of layers and towers are done as follows:

- HB1 ($\phi = 1 - 4$) has default grouping (Group 17, 19) keeping the lateral towers intact and only two longitudinal sampling (one for HB and one for HO).
- HB2 ($\phi = 5 - 8$) provides readout of all individual layers of HB (Group [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17]) summing all η towers in these layers but lateral profile in ϕ still available.

The downstream veto counters, V3 and V6 are kept behind HB1 and HB2 respectively during the low energy runs. The VM is kept after HO and is useful for tagging the high energy muon beams.

4.3 Oscar-Geant4 Simulation of HCal TB2004

4.3.1 Geant4

Detector simulation package for CMS has been designed in the Object Oriented framework using the GEANT4 toolkit [33] developed in a worldwide collaboration (refer [34]).

GEANT4 [35] is a simulation toolkit for the physics of electromagnetic and hadronic showers in any material and is accepted to be used for the detailed simulation of CMS

detector. It includes solid modelling. tracking of particles through matte handling field effects and boundaries, detector response, run and event management, object storage, visualisation and graphic user interface.

The CMS geometry is constructed in GEANT4 using the mother-daughter relation hierarchy where the complete CMS detector is treated as the mother volume [36]. The physical volume is described as a logical volume (shape + material) positioned in a mother volume with a proper translation and rotation matrix. A particle deposits energy in the parts which are declared as *sensitive detectors*. Each particle undergoes one of the processes which are broadly categorised as *continuous processes* and *discrete processes*. The particles move in a straight line (neutral particles) or curved path (charged particles in magnetic field) crossing the boundary in transportation while in continuous processes, the kinematics get modified but the particle retains its identity. The particle might lose its identity in discrete processes producing new particles. At the time of tracking,one of these possibilities is considered before taking a *step*.

The physics of electromagnetic showers is well understood and GEANT4 models used for this purpose are well tested. A hadronic shower simulation toolkit includes the interactions and cross sections of the scattering of all the mesons and baryons off any stable or long lived isotope. It includes the physics processes of elastic scattering, inelastic scattering, absorption, annihilation, charge exchange, radiative capture of neutral particles and the particle induced fission. The time structure of the events is also modelled. According to LHC requirements, the hadronic shower simulation shall provide means to extrapolate beyond the energies available at test-beams in a manner that is consistent with the theoretical knowledge. GEANT4 provides three models for the hadronic shower simulation : data driven models, parameterised models and theory-based models [37].

4.3.2 Oscar

The CMS simulation package, OSCAR [38], is based on the GEANT4 simulation toolkit and the CMS object-oriented framework for simulation and reconstruction. OSCAR manages all CMS detectors, both central (Tracker, Calorimeters and Muons), forward, as well several test beam layouts and prototypes. It implements their sensitive detector behaviour, track selections mechanisms, hit collections and numbering schemes.

4.3.3 Geometry

For the TB2004 simulations, OSCAR 3_7_0 framework is used for the geometry setup of CMS detector. A single particle beam of pions is directed to the centre of an HCal tower and energy deposition in the 3×3 tower around it is considered to find out the response and resolution of the calorimeter. Two cases of grouping (Figure 4.3) the layers longitudinally are considered:

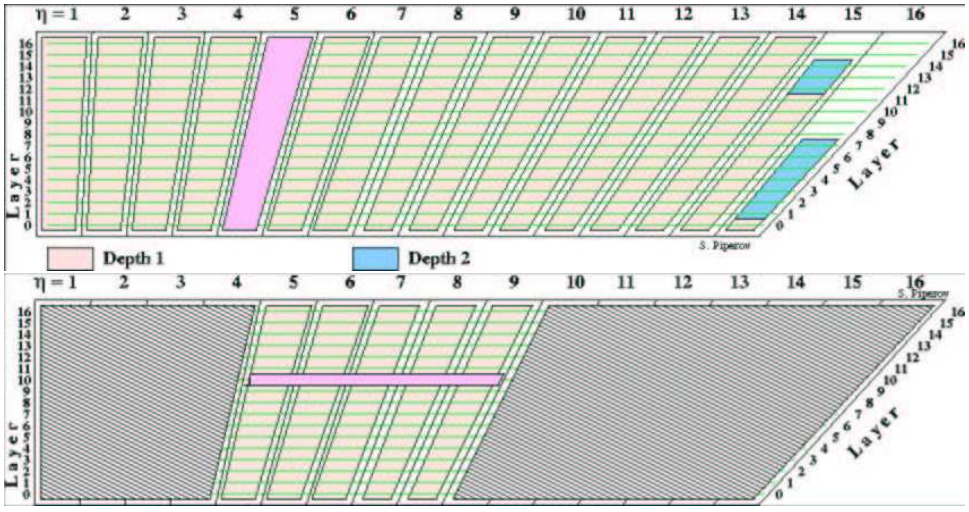


Figure 4.3: Grouping of Layers

- Default grouping groups layers $1-17$, $18-19$. Lateral grouping in both η and ϕ directions is also carried out.
- HB2 reads out 19 longitudinal segments individually. So they are kept separately but no division is done along the η direction. Grouping in ϕ has however been carried out.

Each segment forms a readout channel, therefore, there are 98 for the first case, 133 for the second case. The results at the SimHit[1] level always give an idealistic picture of a physical process. But in a real experiment contributions from electronic noise come at every level of readout system. To simulate the effect of digitisation on the resolution and linearity, the pulse shape is parametrised as a convolution of two contributions.

[1]A SimHit is a GEANT4 output which is characterised by deposited energy and arrival time.

4.3.4 Simulation Geometry

GEANT4 provides a set of models for the hadronic shower simulations. These models have to be validated with the experimental data to make them acceptable to be used in further experiments. One of the main aims of TB2004 is to make a study of shower profiles and linearity for the validation of these models. Two models, QGSP [39] and LHEP [39], provided by GEANT4 are considered for this study. QGSP is a theory driven (microscopic) model based on the quark-gluon string fragmentation. LHEP is a parametrised model for inelastic scattering.

4.3.5 Selection of Time Slices

LHC is designed to operate with the luminosity of 10^{34} cm^{-2} s^{-1} with a bunch crossing frequency of 40 MHz and inelastic pp cross section at $\sqrt{s} = 14$ TeV will be 55 mb. With this given bunch crossing frequency the bunch crossing interval turns out to be 25 ns. The readout of the calorimeter is integrated in 25 ns samples. For a given trigger in the test beam setup, 20 such time slices are read out. But among them only three to four time slices contain 99% of the signal. This can be illustrated through the following plots. The first one shows the charge distribution in the 20 time slices. In the Figure 4.4 the distribution in the central tower is shown where the beam was shot at together with the ones of the neighbouring towers.

The second plot (Figure 4.5) shows the signal across the 20 time slices where the individual contribution of each of the 3×3 towers is summed up. One clearly sees that it is sufficient to add up 4 time slices to get almost the full signal.

4.4 Analysis of TB2004 Data

TB2004 data are taken for electrons, pions and muons of different energies. A 7×7 crystal matrix is kept in front of HB wedges. Shower generated by an electron of a few hundreds of GeV energy is contained almost completely in ECal (which provides radiation). Hadronic showers can be characterised by *interaction length* (λ_{int}). ECal provides only 1 λ_{int} to hadrons, therefore, a hadron passing through ECal starts showering and continues to do so in HCal (9 λ_{int}). Very low energy pion beams (3-9 GeV) are highly contaminated with

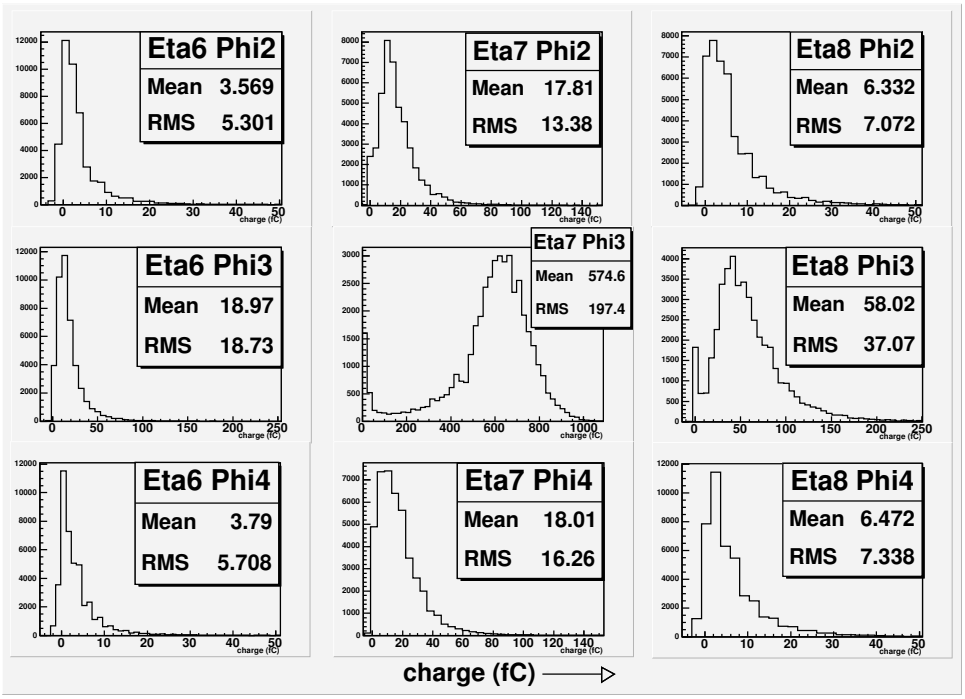


Figure 4.4: Plots of charge distributions in the central tower and the neighbouring 3 towers.

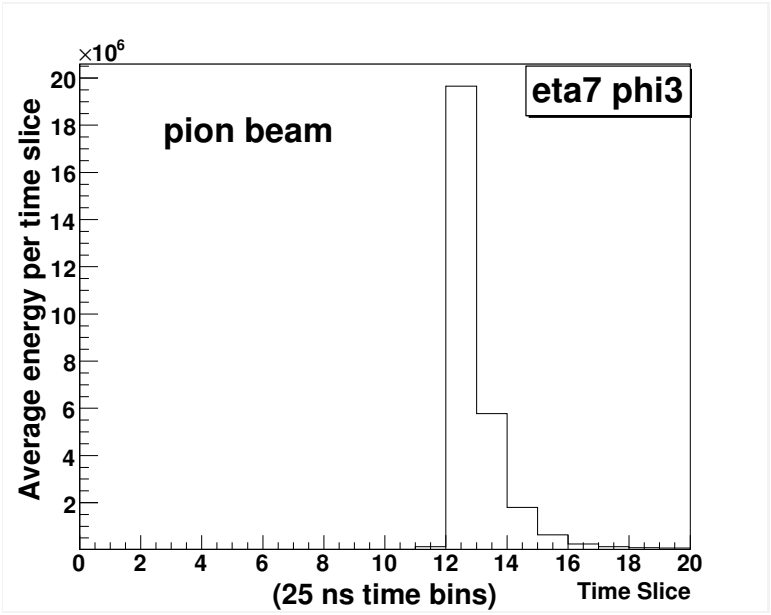


Figure 4.5: Signal pulse across 20 time slices, the contributions from all the 3×3 towers around the central tower being summed up.

electrons and muons and it is important to veto these particles effectively to study the true response of calorimeter to hadrons. The cuts are decided by looking at the QADC distribution of various veto counters. The energy deposited in HCal for different pion beam energies are shown before and after applying these cuts. In this report only the HCal alone set up is considered.

4.4.1 Beam Contamination

Here the data with the HCal *alone* setup (in absence of the crystal matrix) are examined. HCal is a sampling calorimeter, therefore, only a fraction of initial energy of particles is converted to visible signal. Pion beams with different energies are passed through and the energy deposition in HCal are shown in Figure 4.6. The energy deposit in the central tower, where the beam was shot at, is combined with those in the neighbouring cells (3×3 matrix). The energy deposition in the HB1 for pion beam with 100 GeV energy is taken as the reference and all other energies are scaled up with respect to that.

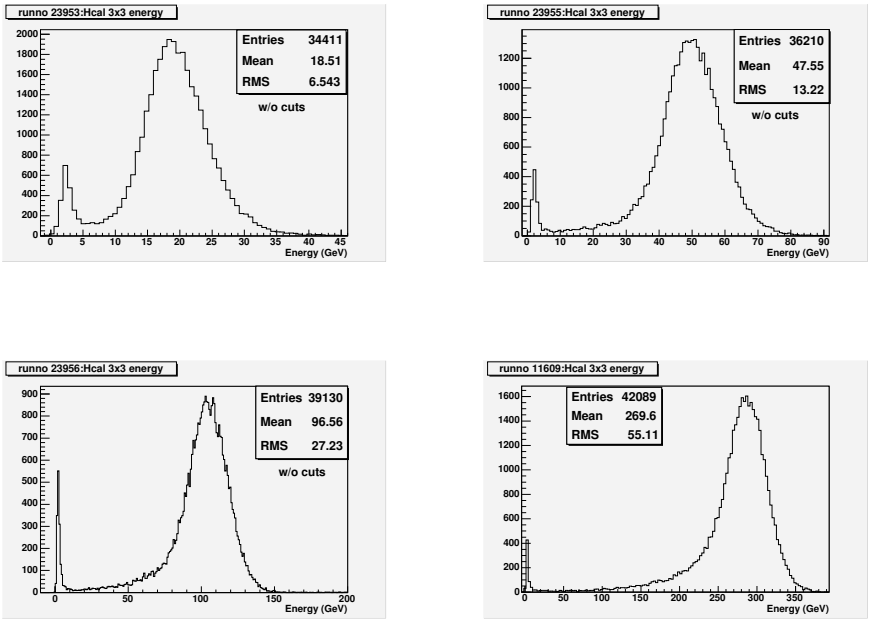
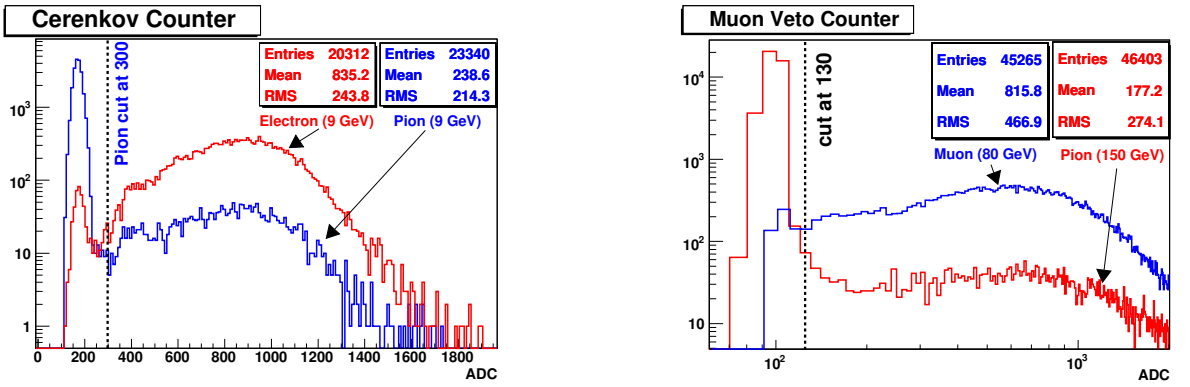


Figure 4.6: Plots of total energy deposition in HB1 for different pion energies before applying the cuts (energies: 20, 50, 100, 300 GeV).

As can be seen from the plots in Figure 4.6 that the low energy tails of such energy distributions contain small peaks which represent *contamination* by muons. To get rid of contamination from electrons and muon, signal in the beam counters are examined and suitable **cuts** are applied in their distributions.

4.4.2 Application of Cuts



(a) CK2 cut (b) VM cut

Figure 4.7: Determining the selection criteria.

The aim of this analysis is to plot energy response, resolution and the longitudinal shower profiles with pion beams of different energies in the HCal alone setup. Therefore, the beam has to be made free from all sorts of contamination as much as possible. This is done by using *Veto counters* and *Cerenkov counters*. Muons, being penetrating particles, will give signals in the veto counters. The Cerenkov detectors, on the other hand, are operated below pion or muon threshold (for low energy beams), so that signal in this detector will signal that the beam going through this detector to be electron. Examining the signal distributions in one particular counter, the threshold value, below/ above which the probability of getting one particular type of beam is very small, is estimated. For example, Figure 4.7(b) shows how to put a cut on VM for high energy pion runs. Here the QADC counts in the VM counter are plotted for an 80 GeV/c muon run with a 150 GeV/c pion run. From the figure it is evident that the presence of muons corresponds to a QADC count > 120. Hence in order to veto muons in pion beams a *high cut* on VM is applied at 120 as shown in Figure 4.7(b).

Figure 4.8 shows the energy deposition for pion runs at different beam energies and here the small peaks (as observed in Figure 4.6) due to contamination of muons has disappeared.

4.4.3 Energy Distribution in ECal+HCal Setup

The scatter plots of energy deposition in the ECal against that in the HCal clearly demonstrate the nature of the beam and also any contanimation in it. This is evident

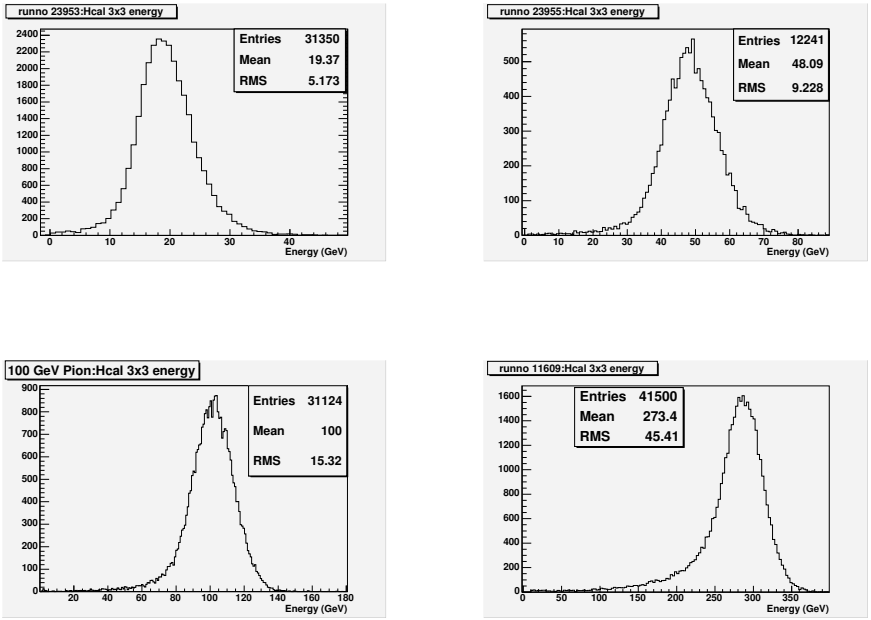


Figure 4.8: Plots of total energy deposition in HB1 for different pion energies after applying the cuts (energies: 20, 50, 100, 300 GeV).

from Figure 4.9 which shows scatter plots for electron beam at 50 GeV/c and pion beam at 100 GeV/c.

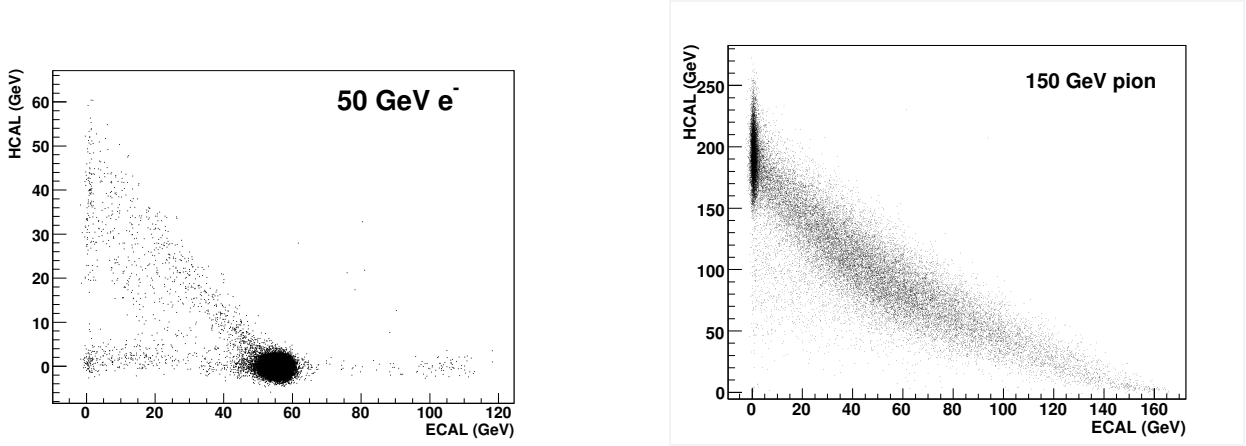


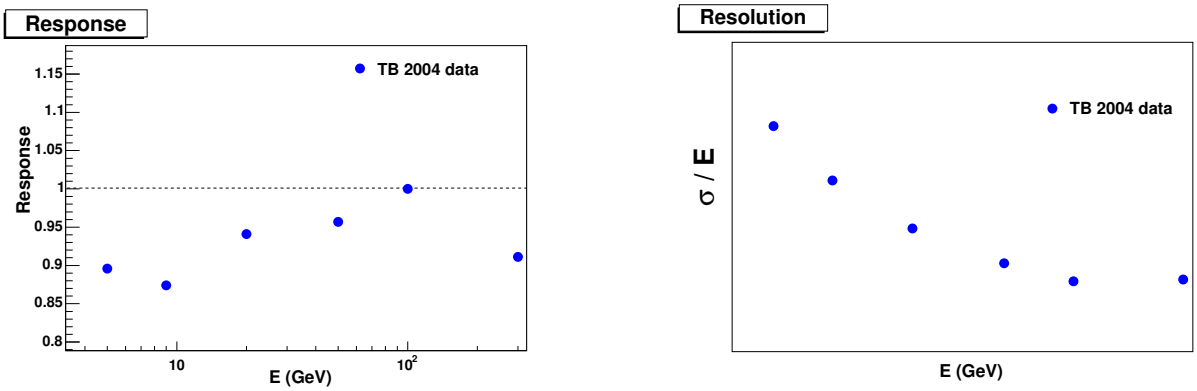
Figure 4.9: Scatter plots of energy depositions in HCal vs ECal for different beam energies.

4.4.4 Results

4.4.4.1 Linearity of Response

A particle loses all its energy in calorimeter and it is utilised in generating shower. The total number of particles produced in the medium must be, on average, proportional to

the initial energy of the particle and the response of detector is expected to be linear with energy. Default grouping is used to study the variation of response of HCal with energy. Since only a fraction of energy is deposited in the CMS HCal, the energy measured is normalised with respect to 100 GeV pions.



(a) Response (b) Resolution

Figure 4.10: Plots for response and resolution with test beam data (2004) as a function of beam energy for pions.

Figure 4.10(a) shows the **response**, measured energy divided by the nominal beam energy, as a function of nominal energy.

4.4.4.2 Resolution

Whatever may be the initial energy of the incident particle, it degrades down to the level of atomic excitations and deposits its entire energy. Because of the probabilistic nature of the processes involved in the production of shower particles, there is an intrinsic fluctuation in the number of particles. This intrinsic fluctuation gives an ultimate limit on the performance of the calorimeter. CMS HCal being a sampling calorimeter, only a fraction of energy is deposited in sensitive detector. Since there is always a fluctuation in the number of particles produced at a given depth of calorimeter, the sampling fluctuation is larger at lower energies as compared to the higher energies. Therefore, the resolution is expected to improve with the energy of the incident particle.

The resolution is calculated as the ratio of measured RMS to mean of the energy distribution. Figure 4.10(b) shows the resolution as a function of nominal energy.

4.4.4.3 Longitudinal Shower Profiles

The longitudinal shower profiles with energies 5, 9, 30, 100, 300 GeV are studied using the HB2 alone configuration. The mean fractional energy deposited in each layer is plotted as a function of layer number: 1–17. Figure 4.11 shows the shower profiles for pion beams at 50 and 100 GeV/c. It is observed that the profile has a sharply rising and a gradually falling part which ensures that the pion beam begins showering after entering the HCal. As the energy increases the tail becomes more and more prominent toward the further layers which can be explained by the fact that a low energy pion beam deposits most of its energy in the first a few layers of the absorber. In contrast a high energy pion beam continues to the further layers gradually depositing its energy as the shower development continues in the latter layers.

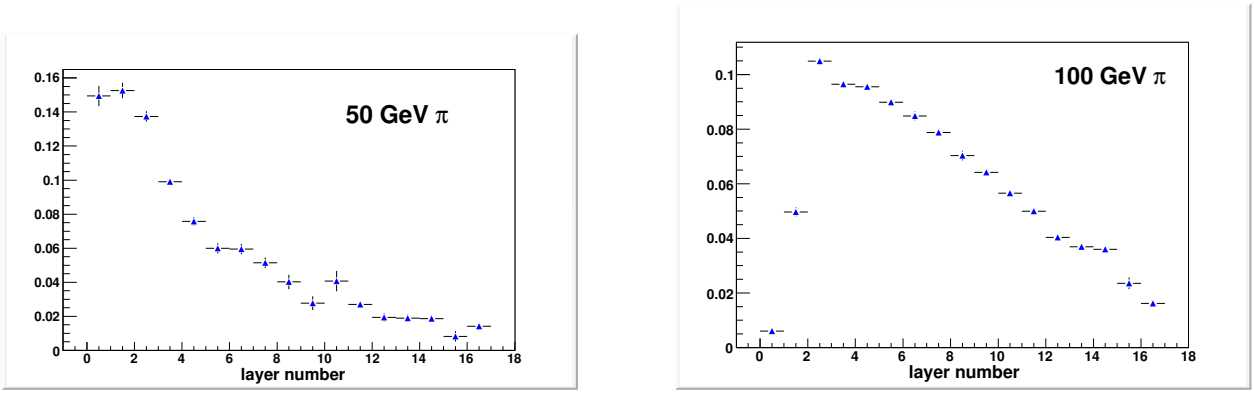


Figure 4.11: Longitudinal shower profiles for different energies of pions generated by TB2004 data.

4.5 Validation of Monte Carlo Simulation

OSCAR version 3_7_0 has been used for simulating events in the HCal alone setup. Two models, QGSP and LHEP, provided by GEANT4 are considered for this study.

Pions(π^-) are generated at the energies of 5, 10, 20, 50, 100, and 300 GeV. The particles are incident on the calorimeter at $\eta = 0.5655$ and $\phi = 0.1309$. These values correspond to the centre of the tower $\eta \times \phi$ index $= 7 \times 6$. The interaction of the particles with detector is simulated in the *absence of magnetic field*.

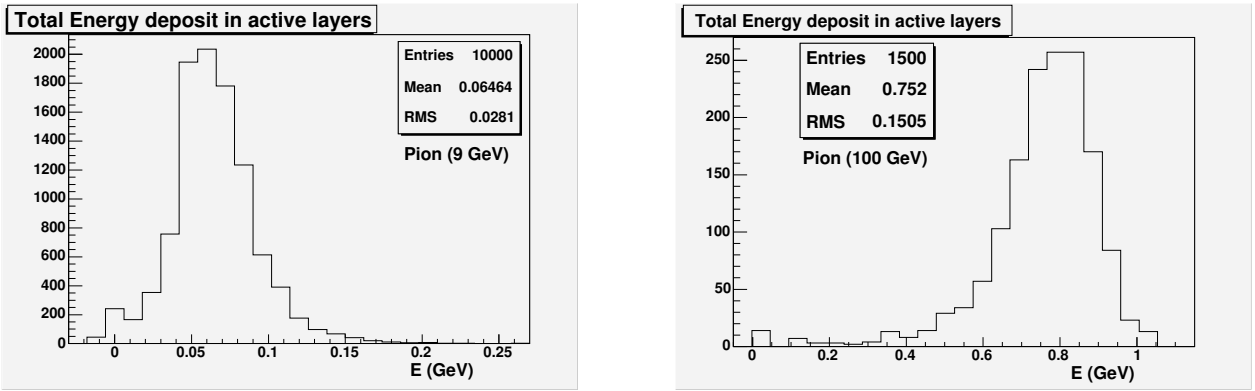
The number events generated for different energies is shown in Table 4.1.

Energy (GeV)	events generated (QGSP)	events generated (LHEP)
5	10000	10000
9	10000	10000
20	10000	10000
50	5000	5000
100	1500	1500

Table 4.1: Number of events generated for various energies of pion-beam

4.5.1 Energy Deposit in HCal Alone Setup

Typical distributions of energy deposited in HCal for the HCal alone setup for 9 GeV and 100 GeV pion beam are shown in Figure 4.12 for the QGSP model. These energy deposits can be translated to the energy of the incident pion through a set of calibration constants as mentioned in the analysis of test beam data.



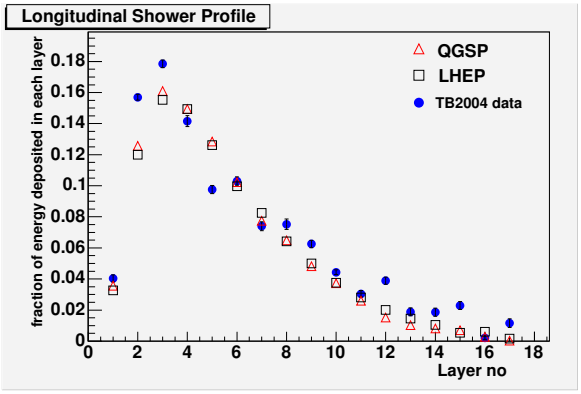
(a) 9 GeV pion (b) 100 GeV pion

Figure 4.12: Total energy deposition plots with digitised data.

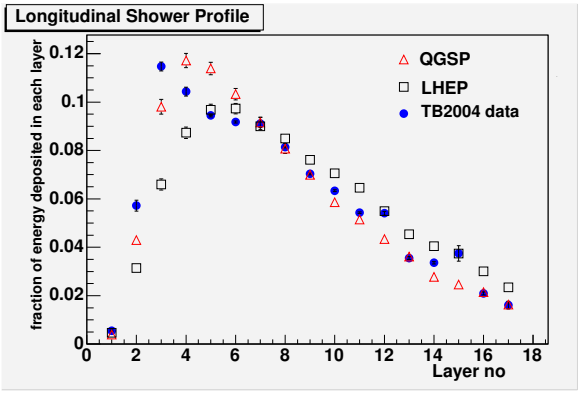
4.5.2 Longitudinal Shower Profiles

Longitudinal shower profiles for HB2 (17 layers) are created using the digitised data for pions (π^-: PDG Particle Id: -211 as given in the standard input file .orcarc) with different energies (5, 9, 20, 50, 100, 300 GeV).Two of them (9 and 100 GeV) are shown here in Figure 4.13 plotted with the test beam data.

The longitudinal shower profiles in Figure 4.13 resemble each other well for the two particular energies shown. The digitised data from QGSP model is seen to match the test beam data more than the LHEP model.



(a) 9 GeV pion

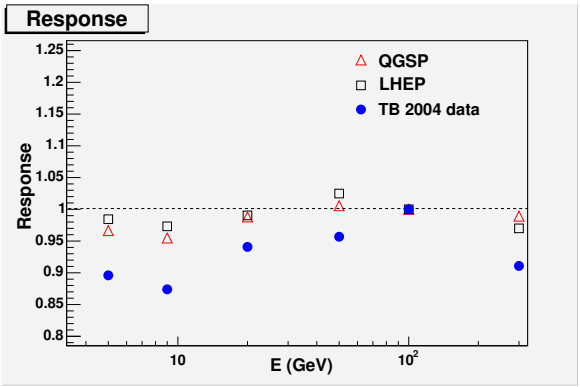


(b) 100 GeV pion

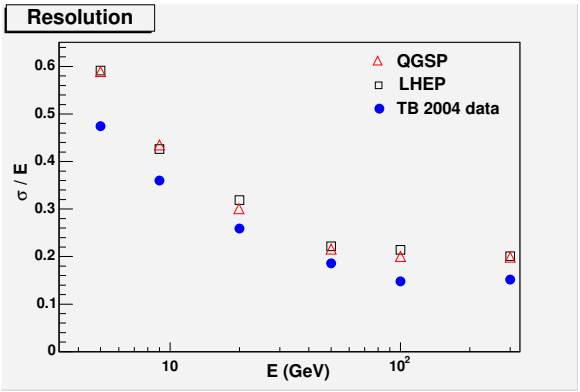
Figure 4.13: Validation of GEANT4 models with test beam data on the longitudinal shower profiles.

4.5.3 Response and Resolution

Figure 4.14(a) shows plot for response for HCal alone setup with the digitised data from the two physics models along with the test beam data. The deviation from linearity is seen to be more than for LHEP, especially at low energies.



(a) Response



(b) Resolution

Figure 4.14: Validation of GEANT4 models with test beam data on plots for response and resolution.

The resolution simulated by the physics models - QGSP and LHEP (under OSCAR 3_7_0 framework) for HCal (HB1 to be precise) is given in Figure 4.14(b) as a function of energy. The resolution is calculated as the measured rms per mean of the energy distribution of pions. At higher energies, the shower leakage fluctuations start contributing significantly in the HCal alone set up.

Figure 4.14 shows a comparison between real and simulated data (GEANT4 generated) in the context of *response* and *resolution*. It is seen that both the TB2004 data as well

as the digitised data simulated by the two physics models do not deviate much from linearity (within $\sim 4\%$) in the response plot. Energy response from digitised data has a larger deviation from linearity at very low energies. Between the two models LHEP shows a slightly better performance over QGSP. The resolution plots for both real data and simulated data show an improvement in resolution (decreasing σ/E) with increase of energy.

Chapter 5

Dijet Studies

5.1 Motivation

Study of jets is an important part of LHC physics. The 'planned' discovery of Higgs and SUSY depends on good jet measurement. Any unexpected discovery such as compositeness depends on measurement and understanding of the QCD jet rates. Detector calibration, synchronisation and triggering will be challenging to meet these demands.

The experimental motivation for dijet study is that LHC is a parton-parton collider in a previously unexplored energy region. If new parton-parton resonances exist then the LHC will produce them copiously (see Figure 5.1). These resonances should also decay to partons giving 2 jets in the final state. The experimental motivation to search for dijet and its resonance is intuitively obvious.

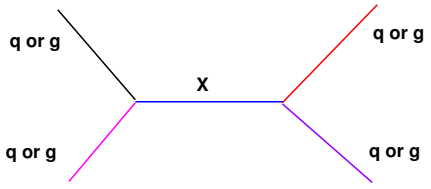


Figure 5.1: Feynman diagram for resonance production in dijet channel

Dijet resonance is of critical importance to theoretical models as well. Such resonances are found in models that address some of the most profound questions of particle physics such as the compositeness of quarks and unification of gravity with other fundamental forces of nature.

Inclusive dijet production ($pp \rightarrow X \rightarrow$ jet + jet) is the dominant LHC hard scattering process. Simple to observe, and rich in potential signals of new physics, dijets are expected to be one of the earliest CMS measurements.

5.2 Jet Generation and Reconstruction

Inclusive dijet production at LHC is dominated by the $2 \rightarrow 2$ QCD scattering of partons (quarks, antiquarks and gluons) in which there are only partons in the initial, intermediate and final states.

QCD jet production has, by several orders of magnitude, the largest cross section. The entire range of momenta is divided into 21 bins for minimising the computing time. A total of 210,000 events are generated from 21 contiguous p_T bins spanning from 0 to 4000 GeV/c (0-15, 15-20, 20-30, 30-50, 50-80, 80-20, 120-170, 170-230, 230-300, 300-380, 380-470, 470-600, 600-800, 800-1000, 1000-1400, 1400-1800, 1800-2200, 2200-2600, 2600-3000, 3000-3500, and 3500-4000) each consisting of 10,000 events. The focus has been kept mostly on the central rapidity region up to 2.5 in rapidity. The cross section for the above mentioned bins is tabulated in Table 5.1.

p_T Range (GeV/c)	cross section (mb)	p_T Range (GeV/c)	cross section (mb)
$0-15$	$5.522E+01$	$470-600$	$6.845E-07$
$15-20$	$1.450E+01$	$600-800$	$2.032E-07$
$20-30$	$6.264E-01$	$800-1000$	$3.549E-08$
$30-50$	$1.546E-01$	$1000-1400$	$1.079E-08$
$50-80$	$2.075E-02$	$1400-1800$	$1.049E-09$
$80-120$	$2.941E-03$	$1800-2200$	$1.446E-10$
$120-170$	$5.002E-04$	$2200-2600$	$2.378E-11$
$170-230$	$1.006E-04$	$2600-3000$	$4.268E-12$
$230-300$	$2.393E-05$	$3000-3500$	$8.444E-13$
$300-380$	$6.381E-06$	$3500-4000$	$9.718E-14$
$380-470$	$1.885E-06$		

Table 5.1: Cross section table

The cross sections are obtained by running stand-alone PYTHIA for those given ranges and for the above mentioned hard QCD processes using the CTEQ 5L for PDF.

5.2.1 Event Generation

The analysed events are generated with PYTHIA and subsequently subjected to the GEANT4 based CMS detector simulation and reconstruction programmes. The jet energy calibration has been performed with a Monte Carlo calibration method implying calibration factors that are applied *on a jet by jet basis to the calorimeter jets* depending on pseudo-rapidity η and transverse momentum p_T. The alternative data based technique of *gamma-jet calibration* where jet transverse energies are measured against recoiling high energetic photons could not yet be employed for his study. The effects of soft physics modelled in the form of parton showers and hadronisation models with subsequent decays has been taken into account.

The following things are considered for the event generation:

- only the hard scattering events, e.g., $ff \to ff$, $f\bar{f} \to f\bar{f}$, $f\bar{f} \to gg$, $fg \to fg$, $gg \to f\bar{f}$, $gg \to gg$.
- choice of fragmentation function: diquark-antidiquark pair production is allowed.
- decay of unstable particles: a particle is allowed to decay only if the average proper lifetime is smaller than $c\tau = 10$ mm.
- for the calculation of α_s at hard interactions only first order running of α_s is considered.
- the K factors are not included in hard cross sections for parton-parton interactions (i.e. for incoming quarks and gluons).
- the structure function is chosen up to *leading order* (PDF: CTEQ 5L).
- multiple interactions are allowed assuming a varying impact parameter and a hadronic matter overlap consistent with a double Gaussian matter distribution, with a core radius of 0.4 fm, and with a continuous turn-off of the cross section at a p_T cut off of 1.9 TeV/c.

- the Q^2 scale of the hard scattering is multiplied by unity to define the maximum parton virtuality allowed in the space-like showers.

- probability that an additional interaction in the multiple interaction formalism gives two gluons, with colour connections to 'nearest neighbours' in momentum space, is chosen to be $\frac{1}{3}$. The remaining fraction is supposed to consist of quark-antiquark pairs.

- in order to account for an assumed dominance of valence quarks at low transverse momentum scales, a probability is introduced that a gg scattering according to naive cross section is replaced by a $q\bar{q}$ one. That probability is parametrised with a regularisation scale and for the present calculation expressed as $P = 0.5(1 - \frac{p_T^2}{(p_T^2+b^2)^2)})$.

- the initial and final state QCD & QED radiations are switched on.

- fragmentation & decay are activated

- maximum absolute pseudo-rapidity value used for detector is 4 (i.e. $-4 \leq \eta \leq 4$).

- minimum E_T for a cell to be considered as a potential jet initiator is taken to be 1.5.

- no threshold is kept for the calorimeter, i.e. all cells are allowed to be clustered.

- minimum summed E_T for a collection of cells (i.e. a cluster) to be accepted as a *jet* is set to be 7 GeV.

- the radius of the cone used in the cone algorithm is kept at either 0.5 or 0.7. This means that the maximum distance in $R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ from cell initiator when grouping cells whether they qualify as a jet is 0.5 or 0.7.

- the transverse energy in a cell, E_T, is smeared according to a Gaussian distribution with standard deviation of $0.5 \times \sqrt{E_T}$, to simulate calorimeter resolution effects.

- number of bins in pseudo-rapidity range $-4 \leq \eta \leq 4$ is 91 to match with the definite cell size for the CMS hadron calorimeter (transverse granularity, $\Delta\eta = 0.087$, hence $8.0/0.087 = 91$).

- number of azimuthal bins is chosen to be 36 to match the azimuthal granularity of the CMS hadron calorimeter ($\Delta\phi = 0.087$).

For the present analysis the dijet system is defined as the two jets with the highest p_T in an event (leading jets).

From the number jets produced during a Monte Carlo process the two highest p_T jets are selected for determining the dijet mass. The p_T distribution and the η distribution of these two jets are studied here. Figure 5.2 shows the p_T and η distributions for the highest and the next highest momentum jets for the p_T range 230−300 GeV/c of the partonic system. Also shown here is the dijet mass distribution.

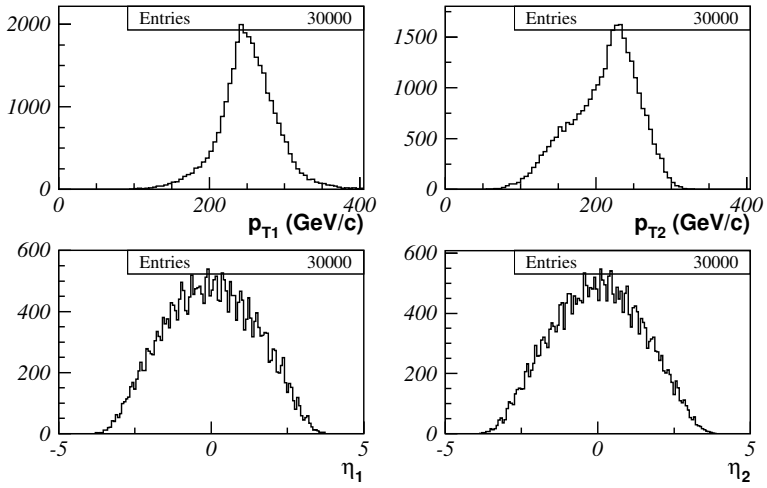


Figure 5.2: p_T and η distributions of the leading and next leading jets for the p_T bin 230−300 GeV/c.

The events where each of the leading jets has $|\eta| < 1$ are selected. This keeps practically all of the jet energies within the barrel region of the CMS calorimeter. This cut also enhances sensitivity to new physics, produced at low $|\eta|$, compared to the pre-dominantly t-channel processes from the QCD background. The same distribution is shown in Figure 5.3 after applying $|\eta| \leq 1$ and $p_T > 90$ GeV.

5.2.2 Jet Reconstruction

Detailed simulation of CMS detector response can be done using the GEANT4 toolkit. However, this process is rather CPU intensive. Even using today's fastest CPU's, simulation of a single pp interaction at LHC energies take several minutes for a high p_T event. On the other hand several tens of thousands of events are required for this study. On the other hand, a simple parameterisation can give rise to significant improvements concerning speed at the price of precision losses.

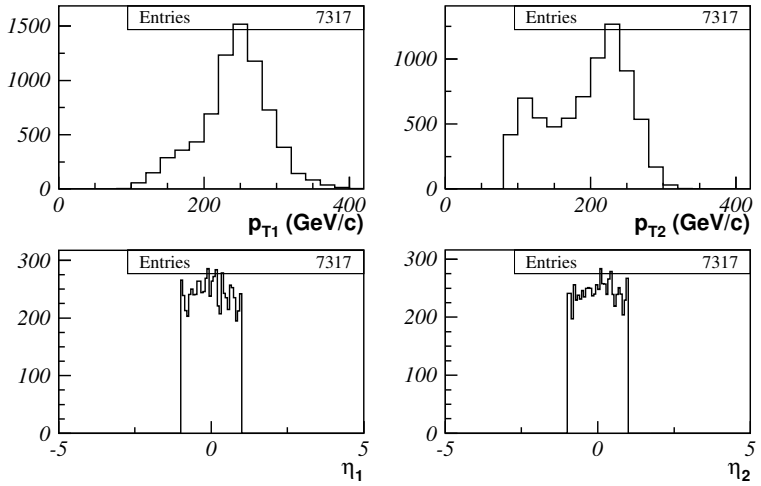


Figure 5.3: p_T and η distributions of the leading and the next leading jets with $|\eta| < 1$ for the p_T bin 230−300 GeV/c.

Therefore, a Fast Monte Carlo Simulation technique (FAMOS) is adopted for reconstruction of jets. This allows fast studies of large samples of Monte Carlo events. Nonetheless FAMOS is fully integrated in the general CMS framework but the components can be used independently.

5.3 Jet Triggers and Trigger Scheme

In order to construct the single jet trigger table for CMS one needs to define the trigger paths from level one (L1) through high level trigger (HLT). This includes reasonable thresholds, prescales, estimates of the resulting rates at L1 and HLT, and evolution of the trigger table with increasing luminosity. For inclusive jet and dijet analysis, to measure jet spectra down to low jet E_T and dijet mass requires multiple triggers, of roughly equal total rate, and with appropriately chosen E_T thresholds and prescales. The measurement down to low dijet mass is important for searches for both dijet resonances and for contact interactions, because it anchors the search distribution in the well understood mass regions previously explored by the Tevatron and understood to be dominated by QCD dijet production. The measurement down to low mass or E_T allows control over QCD background for searches of new physics.

The trigger tables required for the analysis are the dijet triggers. QCD jet events are the signal here as opposed to most other analysis where it is treated as a background.

HLT table provides tighter cuts on the distribution and is considered in this analysis. The trigger table used is given in Table 5.2

Jet Trigger	L1 Jet Trigger			HLT Jet Trigger			Total Prescale
	$p_T(E_T)$ (GeV)	Prescale (1/N)	Rate (KHz)	$p_T(E_T)$ (GeV)	Prescale (1/N)	Rate (Hz)	
High	177	1	~ 1	657	1	~ 1	1
Med	177	1	~ 1	350	30	~ 1	30
Low	177	1	~ 1	180	600	~ 1	600
Tiny	90	20	~ 1	90	600	~ 1	12000

Table 5.2: QCD jet trigger table for the CMS experiment used in this analysis

5.3.1 Method of Designing the Trigger Table

Presently the trigger table is constructed for the instantaneous luminosity of $\mathcal{L} = 2 \times 10^{33}$ $\text{cm}^{-2}\text{s}^{-1}$ and the jet E_T thresholds are chosen to saturate the allowed HLT bandwidth for jets.

The trigger works in a path that starts at a specific L1 trigger and flow to a specific HLT trigger. For example, the HLT trigger in the *High* path only looks at events that have passed the L1 trigger in the *High* path.

One starts with the *High* path by defining HLT E_T threshold of 657 GeV, which satisfies a rate requirement of roughly 1 kHz for this luminosity. The L1 trigger threshold for the *High* path is then chosen at 177 GeV, low enough so the HLT trigger is fully efficient at HLT E_T of 657 GeV. This L1 choice allows subsequent analysis to make full use of the HLT bandwidth, which is more precious than L1 bandwidth for jet analysis. The HLT trigger in the *Med* path is then chosen at an E_T threshold of 350 GeV, roughly half the value of the HLT threshold in the *High* path, and again the L1 threshold of 177 GeV is chosen so that HLT threshold is fully efficient. Since the unprescaled rate for this trigger would be too high, flooding the HLT bandwidth by a factor of 30 higher. Hence a prescale factor is introduced at HLT, reducing the HLT rate by a factor of 30. Similarly, for the *Low* path, the HLT threshold is kept at 180 GeV, from half the value of the HLT threshold in the *Med* path, and then the scale factor is increased to a value 600. For the *Tiny* path the values of both the L1 and the HLT threshold is kept at 90 GeV and the prescale is adjusted accordingly for L1 which keeps the total rate from that path at

roughly 1 kHz.

In practice the dijet mass distribution is obtained by weighting distribution from a given p_T bin with the appropriate pre-scaling factor to match the given luminosity. The *tiny* trigger provides sample of low E_T jets to study calibration and fake rates.

5.4 Dijet Analysis

5.4.1 Generated and Reconstructed Distribution

Figure 5.4 shows the p_T and η distributions of the two most energetic jets at generation and reconstruction level. The iterative cone algorithm is used with different cone radii, namely, 0.5 to 0.7. In the Figure 5.4 GenJetIC5A, JetIC5A, JetIC5C and JetIC7A represent generator level jet, reconstructed level jet (with cone radius 0.5), reconstructed jet with γ-jet calibration and reconstructed jet with cone radius 0.7 respectively. The jets at the detector level shows a significant shift towards smaller value in the p_T distribution of the leading jet as well as next-to-leading jet. This effect gets partially corrected after using the γ-jet calibration. The η distributions at the generator and reconstructed level are seen to match with each other.

5.4.2 Dijet Mass

Jets are reconstructed as localised energy depositions in the CMS calorimeters arranged in a projective tower geometry (EcalPlusHcalTowers). The jet energy E is defined as the scalar sum of the calorimeter tower energies inside a cone of radius $R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ = 0.5, centred on the jet direction. The jet momentum $\vec{p}$ is the corresponding vector sum of energies, with the vector pointing in the tower direction. The jet transverse energy is $E_T = E \sin\theta$, and the proton beam. Both the jet energy and momentum are corrected back to the particles in the jet cone originating from the hard interaction excluding pileup.

The dijet mass is defined as

$$M_{jj} = \sqrt{(E_1 + E_2)^2 - (p_x^1 + p_x^2)^2 - (p_y^1 + p_y^2)^2 - (p_z^1 + p_z^2)^2}$$

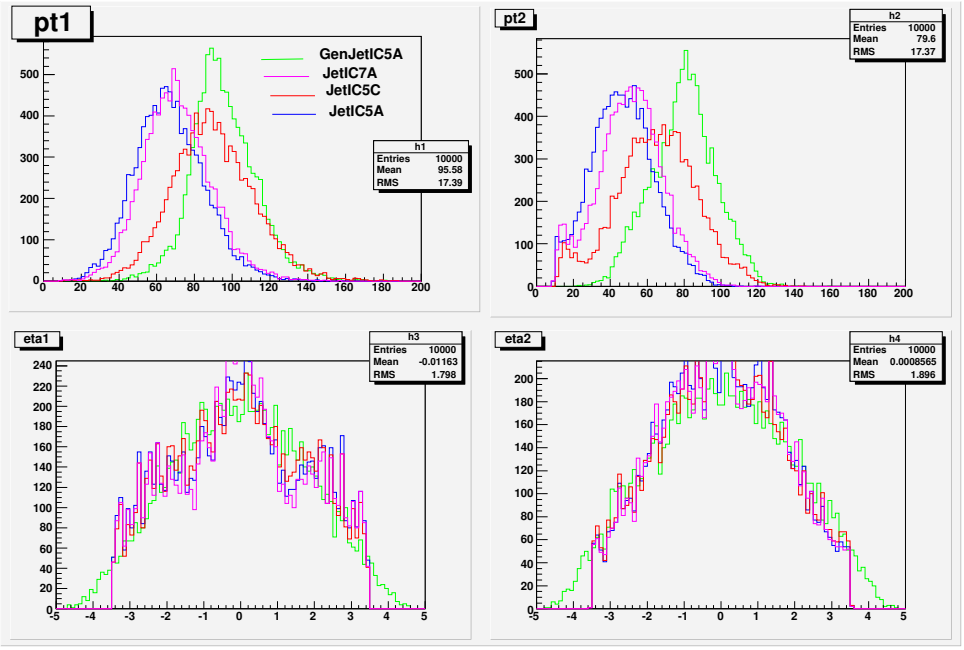


Figure 5.4: p_T and η distributions of the most and the next to most energetic distributions at generator and reconstructed level.

The events in which the leading jets are each has $|\eta| < 1$ are selected. Figure 5.5 shows dijet mass distribution for a given p_T bin (all events and restricting only those events where both the jets are restricted in the barrel, i.e., $|\eta| < 1$)

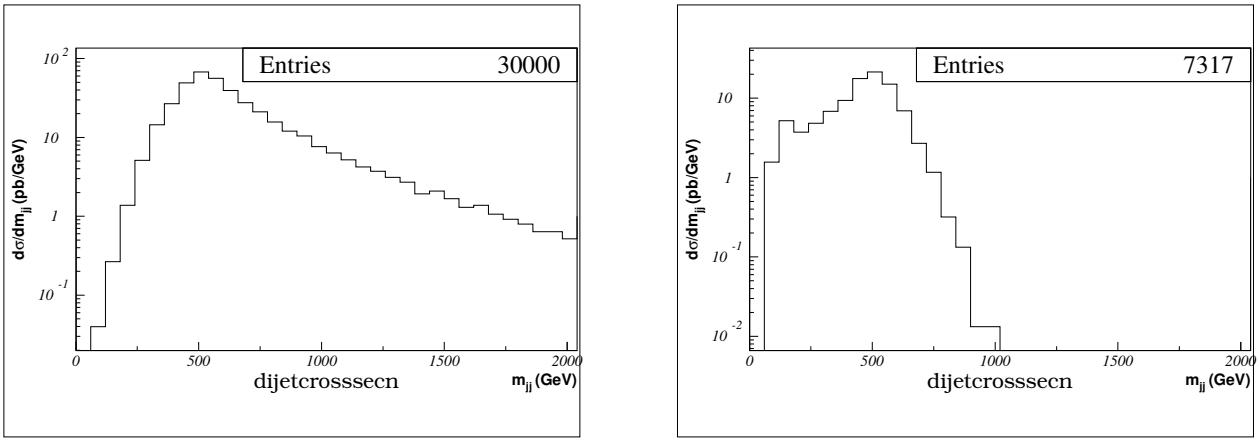


Figure 5.5: Dijet mass distribution for one p_T bin, without and with the $|\eta| < 1$ cut.

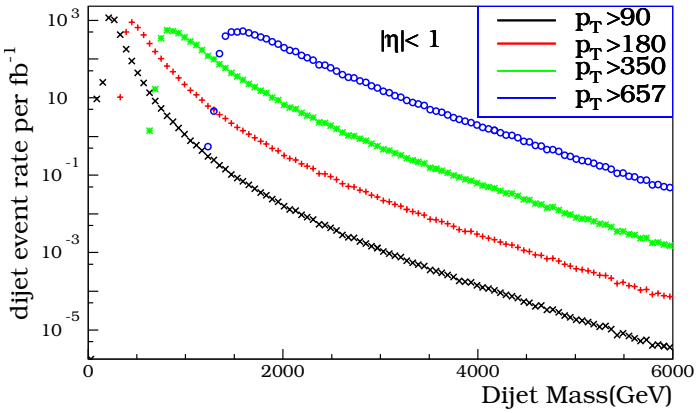


Figure 5.6: Trigger rate estimated here from single sample with varying p_T cut applied offline and considering the effect of HLT only.

5.4.3 Rates and Efficiencies

The rate expected from the the triggers (as presented in trigger table) as a function of dijet mass is plotted at the generator level using PYTHIA. The event rate distribution is done by generating events in CMKIN using PYHTIA.

L1 efficiency is expected to be at least 95% at the p_T thresholds of the HLT. So it is not necessary to include L1 in the rate plots. The turn-on of each trigger in Figure 5.6 is entirely due to the difference between the HLT p_T threshold listed and the dijet mass variable. The dijet mass distribution of events with a leading jet that passes the p_T threshold of HLT is plotted here. By construction there are comparable statistics in each trigger, and a high statistics overlap between adjacent triggers in the trigger table. It is to be noted that in this simulation, unlike the real experiment, a single sample is used to estimate the rate for all the triggers, so the statistics in the overlap regions among the triggers have identical fluctuations: they come from the same sample and are just scaled down in rate by the *prescaling factors*.

5.4.4 Dijet Mass Distribution

The triggers are combined to produce a differential cross section across the full mass spectrum resulting from QCD dijet production. The prescaled triggers allow to measure dijet mass down to 330 GeV, or even smaller if the efficiency of the lowest threshold

trigger can be fully understood. The mass measured with the prescaled triggers will enable connecting the mass distribution to that measured at the Tevatron. Since there has been no new physics beyond the Standard Model at the Tevatron, this mass region can be a control region of the CMS measurement which defines the QCD background to searches for new physics with dijets.

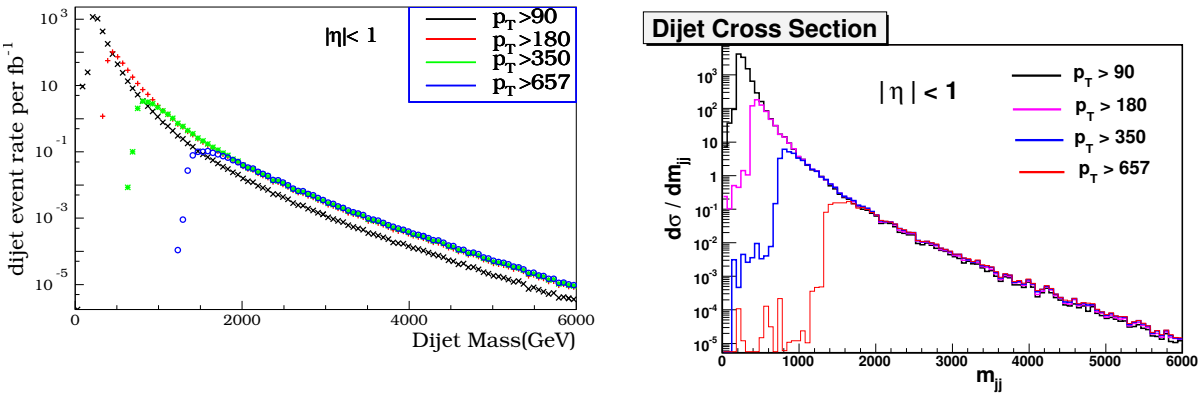


Figure 5.7: Dijet Cross section - generator level using PYTHIA and reconstructed level using FAMOS.

In the Figure(5.7) the cross sections from each trigger sample are combined to form the overall spectrum.

In the actual experiment the data samples come from the different trigger paths (as presented in trigger table), each with different prescales, and there is always the question of how reliably the paths have operated and how well the prescales are known. If the trigger does not always fire when it should, or if the prescales are wrong, then the cross sections from each trigger sample may not combine to form the correct spectrum, artificially introducing structures that could be misinterpreted as new physics. In the real experiment, to ensure the integrity of the trigger and prescale, the cross section in the overlap region will be measured where both the triggers are fully efficient, and the cross section at a given mass from two triggers will be compared. The overlap among triggers is being used by Tevatron experiments to measure the prescales and constrain their systematic uncertainty.

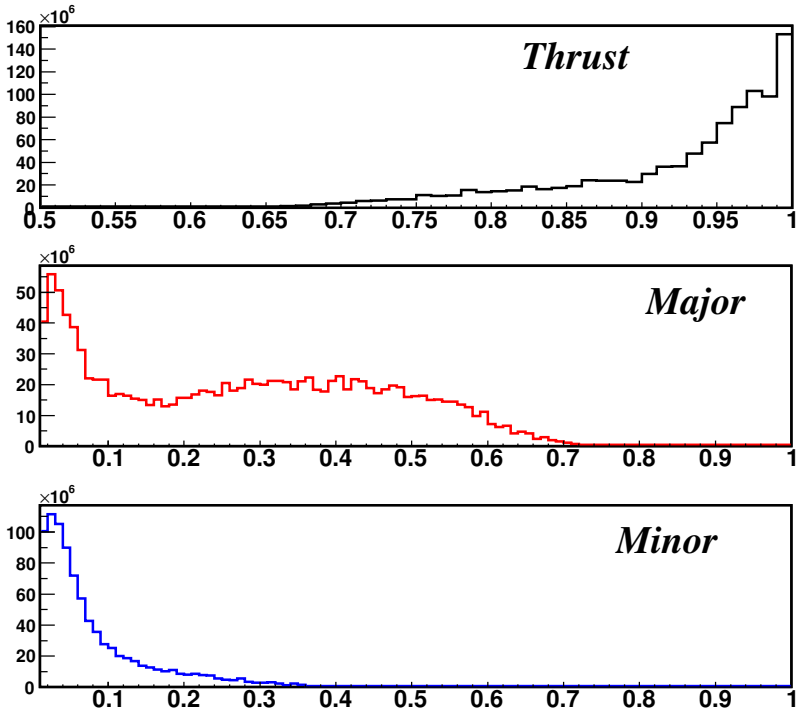


Figure 5.8: Distributions of thrust, major and minor from the high p_T events

5.4.5 Event Shape Distributions

QCD predicts scaling violations for observables that do not depend on absolute energies or momenta, as for example the thrust distribution. These scaling violations are caused by the energy dependence of α_s, which determines the amount of gluon radiation. In leading order the probability of gluon radiation is proportional to α_s. It is therefore expected that fewer 3-jet and multi-jet events will be observed at higher centre-of-mass energies, and that event shape distributions evolve toward the 2-jet limit.

The event shape variables - thrust, thrust-major and thrust-minor are evaluated and shown in Figure 5.8 at the reconstructed level for the hard scattering process. Had it been an ideal hard scattering with the production of two collimated back to back jets, the thrust distribution would have just peaked at unity. Instead there is a clear spread at small thrust values. The fractional deviation from unity qualitatively hints at production of more than two back to back jets and quantitatively measures the amount of gluon production. This will eventually give a handle to measure the running of strong coupling constant for the QCD process involved.

Chapter 6

Summary and Outlook

Inclusive dijet production ($pp \rightarrow X \rightarrow$ jet + jet) is the dominant LHC hard scattering process. Simple to observe, and rich in potential signals of new physics, dijets are expected to be one of the earliest CMS measurements.

The behaviour of QCD dijet events at a generator level and reconstructed level was presented.

The table for single jet triggers for CMS is discussed. The trigger table includes p_T thresholds, prescales, and rate estimates at L1 and HLT for the luminosity of 2×10^{33} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

An analysis of the dijet mass distribution from QCD was presented which would result from the proposed trigger table. To facilitate searches for new physics on this background the two leading jets were constrained to the central pseudo-rapidity region. QCD events can be analysed from a lowest dijet mass of 0.3 TeV to the highest dijet mass expected, which is around 4 TeV. The triggers used have sufficient overlap as well as adequate number of events for the analysis.

The hadronic jet topology, containing the information of quark/gluon radiations, was reflected in the event shape distributions (thrust) presented in the last section.

The calorimeter response to pion beams, in the context of the test beam experiment, has been discussed as well. The use of appropriate selection criteria for obtaining a pure beam was elucidated. The GEANT4 models are validated against the test beam data.

The QCD jets are reconstructed at the detector using the information of the energy

measurement with the electromagnetic and hadron calorimeters. Studying the energy response of the calorimeter in the test beam helps in understanding the performance of the CMS calorimeter system. This also validates the jet reconstruction algorithms applied to the CMS calorimeter system.

Bibliography

[1] Particle Data Group (K. Hagiwara *et al.*), *Review of particle physics*, Physical Review **D66** (2002) 010001.

[2] B. R. Martin and G. Shaw, *Particle Physics*. Chichester, UK: Wiley (1997).

[3] I. J. R. Aitchison and A. J. G. Hey, *Gauge Theories in Particle Physics*. Bristol, UK: IOP (1989).

[4] R. K. Ellis, W. J. Stirling, and B. R. Webber, *QCD and Collider Physics*. Cambridge, UK: CUP (1998).

[5] C.-N. Yang and R. L. Mills, *Conservation of isotopic spin and isotopic gauge invariance*, Physical Review **96** (1954) 191.

[6] P. D. B.Collins and A. D. Martin, *Hadron Interactions*, Adam Hilger Ltd, Bristol (1984).

[7] T. Sjöstrand *et al.*, *High-energy-physics event generation with PYTHIA 6.1*, Computer Physics Communication **135** (2001) 238. http://www.arXiv.org/abs/hep-ph/0010017 (E-print number: `hep-ph/0010017`).

[8] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*, Nuclear Physics **B126** (1977) 298.

[9] T. Sjöstrand, L. Lönnblad, and S. Mrenna, *PYTHIA 6.2: Physics and manual* (2001). http://www.arXiv.org/abs/hep-ph/0108264 (E-print number: `hep-ph/0108264`).

[10] B. Andersson, G. Gustafson, and B. Söderberg, *A general model for jet fragmentation*, Zeitschrift für Physik **C20** (1983) 317.

[11] F. Abe *et al.*, CDF Collaboration, Physical Review Letters **77** (1996) 438.

[12] G. C. Blazey, for the DØ Collaboration, in Proceedings of the 31st Rencontres de Moriond: QCD and High-energy Hadronic Interactions, Les Arcs, France, 1996, p. 155.

[13] F. Abe *et al.*, CDF Collaboration, Physical Review Letters **77** (1996) 5336, erratum *ibid.* **78** (1997) 4307; Physical Review **D54** (1996) 4221.

[14] S. Abachi *et al.*, DØ Collaboration, Physical Review **D53** (1996) 6000.

[15] F. Aversa, P. Chiappetta, M. Greco and J.P. Guillet, Physical Review Letters **65** (1990) 401; Zeitschrift für Physik **C46** (1990) 253;
F. Aversa, P. Chiappetta, L. Gonzales, M. Greco and J.P. Guillet, Zeitschrift für Physik **C49** (1991) 459.

[16] S. D. Ellis, Z. Kunszt and D.E. Soper, Physical Review Letters **62** (1989) 726; Physical Review **D40** (1989) 2188; Physical Review Letters **64** (1990) 2121.

[17] W. T. Giele, E. W. N. Glover and D. A. Kosower, Nuclear Physics **B403** (1993) 633; Physical Review Letters **73** (1994) 2019.

[18] S. D. Ellis, Z. Kunszt and D. E. Soper, Physical Review Letters **69** (1992) 1496;
S. D. Ellis and D. E. Soper, Physical Review Letters **74** (1995) 5182.

[19] W. T. Giele, E. W. N. Glover and D. A. Kosower, Physical Review **D52** (1995) 1486.

[20] W. T. Giele and W. B. Kilgore, Physical Review **D55** (1997) 7183;
W. B. Kilgore, 'Next-to-leading Order Three Jet Production At Hadron Colliders', hep-ph/9705384.

[21] F. Abe *et al.*, CDF Collaboration, Physical Review Letters **70** (1993) 713.

[22] S. Abachi *et al.*, DØ Collaboration, Physics Letters **B357** (1995) 500.

[23] J. E. Huth *et al.*, in *Research Directions for the Decade*, Proceedings of the Summer Study on High Energy Physics, Snowmass, Colorado, 1990, p. 134.

[24] S. D. Ellis and D. E. Soper, Physical Review **D48** (1993) 3160.

[25] S. Catani, Yu. L. Dokshitzer, M. H. Seymour, B. R. Webber, Nuclear Physics **B406** (1993) 187.

[26] M. Derrick *et al.*, ZEUS Collaboration, contribution to the XXVIII ICHEP, Warsaw, 1996, pa 02-043; DESY preprint in preparation.

[27] S. Brandt, C. Peyrou, R. Sosnowski, and A. Wroblewski, *The principal axis of jets—an attempt to analyze high-energy collisions as two-body processes*, Physics Letters **12** (1964) 57.

[28] E. Farhi, *A QCD test for jets*, Physical Review Letters **39** (1977) 1587.

[29] http://cmsinfo.cern.ch/ and references to Technical Design Reports therein.

[30] CMS Collaboration, "The Compact Muon Solenoid - Technical Proposal", CERN/LHCC 94-38, 1994.

[31] CMS Collaboration, "The Electromagnetic Calorimeter Technical Design Report", CERN/LHCC 97-033, 1997.

[32] The Hadron Calorimeter Project, Technical Design Report, CERN/LHCC 97-31, June, 1997.

[33] GEANT4 - a simulation toolkit, Nuclear Instruments and Methods **A506** (2003), 250-303.

[34] Sunanda Banerjee: Simulation of the CMS calorimeter system, CMS-IN-2006/002.

[35] see http://geant4.web.cern.ch/geant4/.

[36] P.Arce *et al.*, Detector description of CMS using GEANT4, CMS-IN-1999/026.

[37] J. P. Wellisch : String Parton Models in GEANT4, CHEP03, La Jolla, California, 24-28 March, 2003.

[38] see http://cmsdoc.cern.ch/OSCAR.

[39] GEANT4 Physics Reference Manual; http://geant4.web.cern.ch/geant4/UserDocumentation/UsersGuides/PhysicsReferenceManual/html81/PhysicsReferenceManual.html