First Measurement of $\sigma(gg \to t\bar{t})/\sigma(p\bar{p} \to t\bar{t})$

by

Shabnaz Pashapour Alamdari

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Physics
University of Toronto

Copyright © 2008 by Shabnaz Pashapour Alamdari

Abstract

First Measurement of $\sigma(gg \to t\bar{t})/\sigma(p\bar{p} \to t\bar{t})$

Shabnaz Pashapour Alamdari

Doctor of Philosophy

Graduate Department of Physics

University of Toronto

2008

The work presented here is the first measurement of the fraction of top quark pair production through gluon-gluon fusion. We use an integrated luminosity of 0.96±0.06 fb^{-1} of $p\bar{p}$ collisions at $\sqrt{s}$ of 1.96 TeV collected by the CDF II detector. We select $t\bar{t}$ candidates by identifying a high-p_T lepton candidate, a large missing E_T as evidence for a neutrino candidate and at least four high E_T jets, one of which has to be identified as originating from a b quark. The challenge is to discriminate between the two production processes with the identical final state, $gg \to t\bar{t}$ and $q\bar{q} \to t\bar{t}$. We take advantage of the fact that compared to a quark, a gluon is more likely to radiate a low momentum gluon and therefore, one expects a larger number of charged particles with low p_T in a process involving more gluons. Given the large uncertainties associated with the modeling of the low p_T charged particle multiplicity, a data-driven technique was employed. Using calibration data samples, we show there exists a clear correlation between the observed average number of low p_T charged particles and the average number of gluons involved in the production process predicted by Monte Carlo calculations. Given the correlation, one can identify low p_T charged particle multiplicity distributions associated with specific average number of gluons. The W+0 jet sample and dijets sample with leading jet E_T in the range of 80-100 GeV are used to find no-gluon and gluon-rich low p_T charged particle multiplicity distributions, respectively. Using these no-gluon and gluon-rich distributions in a likelihood fit, we find the fraction of gluon-rich events in $t\bar{t}$ candidates.

This fraction has contributions from the signal and background events. Taking into account these contributions and the $gg \to t\bar{t}$ and $q\bar{q} \to t\bar{t}$ production channel acceptances, we find $\sigma(gg \to t\bar{t})/\sigma(p\bar{p} \to t\bar{t}) = 0.07 \pm 0.14(stat) \pm 0.07(syst)$ in agreement with the theoretical predictions of 0.15 ± 0.05. This measurement is a stepping stone towards a better understanding of the production mechanism of top quark pairs and implies that at least 67% of the $t\bar{t}$ events are consistent with $q\bar{q}$ production at 95% C.L, leaving little room for non SM processes which are not similar to the $q\bar{q} \to t\bar{t}$ in their gluon radiation. The result confirms our current understanding of the SM high p_T production mechanism and the relative gluon PDFs at relatively high Q^2 and x.

Dedication

To my parents

for all I am...

Acknowledgements

I have been thinking about writing the acknowledgment of my thesis since almost the beginning of my PhD program, noting all the friendly gestures that brightened my days and all the interesting scientific, philosophical and social discussions that helped me to move forward. And yet, I can't find the words to adequately show my gratitude to those whose presence made my life as a graduate student pleasant, joyful and at times productive.

I am infinitely grateful to my PhD advisor, Pekka Sinervo, for all he has offered me. I have learned from Pekka not only how to face the challenges associated with physics but also how to deal with other obstacles one may find in life. Pekka's generous soul made my graduate experience a pleasant one and I treasure all the years we worked together. His intelligent approach to the scientific questions along with his fair attitude towards the efforts of others personifies the image I always had in mind of a true scientist. I wish I achieve his level of knowledge and grace in the field.

I would also like to thank John Martin and Michael Luke, the other members of my supervisory committee, for their insight and enthusiasm for my work. I have enjoyed the guidance and encouragements of many members of the CDF experiment specially members of the top quark group. I have had many interesting discussions with members of the CDF IPP Canada and enjoyed lots of joyful moments with them for which I am thankful. In particular, I enjoyed my friendship and a lot of stimulating physics discussions with Bernd Stelzer, Dan MacQueen, Jean-Francois Arguin, Oliver Stelzer-Chilton, Stan Lai, Kalen Martens, Andrew Hamilton, Ian Vollrath, Kostas Kordas, Michael Riveline, Reda Tafirout, Simon Sabik and Thorsten Koop.

I was lucky to find many good friends outside the IPP experimental high energy physics group with whom I shared many fond memories and thought-provoking discussions: Alejandro Ibarra, Ali Najmaie, Ardavan Darabi, Ashleigh Stelzer-Chilton, Betty Meng, Fazel Fallah-Tafti, Linda Karimi-Tabesh, Marcius Extavour, Pascal Vaudrevange,

Samansa Maneshi, Xiongfeng Ma and Xun Ma.

My closest friend, Elham Farahani, has been a source of inspiration. Her lively spirit, her passion for life and her honest soul have always been a confident presence to rely on in the wildest storms. I will always be indebted to Reza Sharifinia for his never-failing friendship, his carefree attitude and his caring frankness. I would like to thank Pascal Reviere for giving me the chance to joyfully test my philosophy in love and for those moments we were the power of our dreams.

I owe a huge thank you to my aunts and uncles whose thoughtfulness, consideration and affection have always paved my way and cheered up my moments. I would like to thank my cousins to whom I felt close, despite the geographical distances, and with whom I shared lots of laughters. The joy of looking at the cute pictures and cheerful smiles of my younger friends and cousins has been a constant, pleasant companion during these years. I am immensely grateful to them. My grandparents, Ali Baba with his sharp mind in solving problems and teasing me, Maman Sona with her enthusiasm for learning and her kind heart, Baba Ghatari with his notes and books always all around him were all part of me all these years. I will never forget them. Maman Setareh with her charming laughter and her readiness to play is the symbol of a fun grandmother. I hold her so dearly.

I cannot put in words how deeply I am grateful to my parents. Their devotion, love and care are beyond any expression. Their unshakable trust in my abilities has always inspired me to follow my dreams. My father's warm gaze and kind soul have always surrounded me with a feeling of safety. My mother's sweet smile and sincere nature have pervaded my life with beauty. Together, they have thought me to be truthful to myself and to reach for the unknown.

Contents

1 Introduction **1**

1.1 Our Current Understanding of Particle Physics in Brief 3

1.2 Top Quark 7

1.2.1 Production 9

1.2.2 Decay 12

1.3 Parton Branching 12

1.4 W Production 15

1.5 Dijet Production 16

1.6 Measurement 18

1.6.1 Motivation 19

1.6.2 Method 19

2 Experiment **22**

2.1 Accelerator 22

2.1.1 Proton Source 22

2.1.2 Main Injector 24

2.1.3 Antiproton Source 24

2.1.4 Tevatron 25

2.2 CDF II Detector 25

2.2.1 Tracking 28

2.2.2 Calorimeter System 30
2.2.3 Muon System 32
2.2.4 The Luminosity Counters 33
2.3 Trigger and Data Acquisition Systems 34
2.3.1 Level-1 Trigger 34
2.3.2 Level-2 Trigger 36
2.3.3 Level-3 Trigger 36
2.3.4 Trigger Paths for This Analysis 38
2.4 Reconstruction 39
2.4.1 Track Reconstruction 40
2.4.2 The z Vertex Reconstruction 41
2.4.3 Lepton Reconstruction 41
2.4.4 Jet Reconstruction 43
2.4.5 Secondary Vertex Reconstruction 44
2.4.6 Missing E_T Measurement 45
2.5 Monte Carlo Simulation 45
2.5.1 Monte Carlo Packages 46

3 Analysis **49**
3.1 Discriminator 49
3.1.1 Alternative Discriminators 51
3.2 Samples and Selections 53
3.2.1 W+n Jet Sample 54
3.2.2 Dijet Sample 55
3.2.3 Multijet Sample 56
3.2.4 $t\bar{t}$ Signal Sample 56
3.3 Background Processes 57
3.3.1 W+0 Jet Sample 57

3.3.2 $t\bar{t}$ Sample 59

3.4 The Observable Used to Discriminate 59

3.5 Correlation between $<N_{trk}>$ and $<N_g>$ 62

3.5.1 The $<N_g>$ Estimate 64

3.5.2 The Observed $<N_{trk}>$ 64

3.5.3 $<N_{trk}>$ vs. $<N_g>$ 69

3.5.4 Is There a Q^2 Effect? 71

4 Measurement **73**

4.1 Distribution Fits 73

4.1.1 No-gluon and Gluon-rich Parameterizations 73

4.1.2 Finding the Gluon-rich Fraction 75

4.2 Gluon-rich Fraction of $t\bar{t}$ Sample 83

4.3 $\sigma(gg \to t\bar{t})/\sigma(p\bar{p} \to t\bar{t})$ 85

4.4 Systematic Uncertainties 88

4.4.1 Sources of Uncertainy in Track Multiplicity Distribution 88

4.4.2 Other Systematic Uncertainies 90

5 Conclusion **93**

5.1 Possible Future Improvements 95

5.2 Future Prospects at the Tevatron and LHC 97

Bibliography **98**

List of Tables

1.1 The estimated or measured mass of the fundamental particles [5]. 6

1.2 The two-to-two parton subprocesses are listed along with their invariant matrix elements squared $\bar{\Sigma}|\mathcal{M}|^2$ averaged (summed) over the colour and spin indices of the initial (final) states. Also the numerical value of these $\bar{\Sigma}|\mathcal{M}|^2$ at 90° in the CM frame is given for comparison. The variables $\hat{s}^2$, $\hat{t}^2$ and $\hat{u}^2$ are defined in the text. 17

2.1 The parameters for different calorimeters. The unit λ is the attenuation length and χ_0 is a radiation length. The energy is measured in units of GeV. 31

3.1 The background processes for the W+0 jet sample separated in the two groups of similar and different production mechanisms. 58

3.2 The background processes for the $t\bar{t}$ sample separated in the three categories of LF, HF and non-W backgrounds. The non-W background has contributions from both LF and HF processes. 61

3.3 The average number of gluons in each sample as predicted by MC calculations and the average number of gluons as found using the correlation fit for data. All uncertainties are statistical. 70

4.1 The fraction of gluon-rich events in each sample as predicted by MC calculations and the fraction of gluon-rich events as found using the likelihood fit to track multiplicity distributions. Uncertainties for the MC fractions include both statistical and systematical contributions. The uncertainties on the fit results to the data are only statistical. 76

4.2 Gluon-rich fraction values from the likelihood fit to the low p_T track multiplicity distributions for W+1, 2 and 3 jet samples with no positive b-tag and with at least one positive b-tag, as well as the extrapolated gluon-rich fractions for both tagged and no-tag sets. 85

4.3 The background fractions used in the analysis, f_b, f_{bkg}^{HF} and f_{bkg}^{LF} in the tagged sample. 85

4.4 Sources of systematic effects and their effects on the measured values. . . 92

List of Figures

1.1 An illustration of the effect of the Higgs field in generating the mass of the quarks, leptons and force carriers. The three generations of the leptons and quarks are also shown separately. Fermilab Neg. #: 05-0440-01D. . . 5

1.2 The CTEQ6M PDFs at $Q = 100$ GeV as a function of x. The distributions for different partons are shown with different line styles as specified in the legend. 8

1.3 Leading-order Feynman diagrams corresponding to $t\bar{t}$ production through a) quark-antiquark annihilation and b) gluon-gluon fusion. 10

1.4 The hadronic $t\bar{t}$ production cross section as a function of the hadronic CM energy. The contribution from gg initial states (dotted line) and from $q\bar{q}$ initial states (dashed lines) are also shown. The upper (lower) dashed lines are for $p\bar{p}$ (pp) collisions. 11

1.5 Splitting function values (including the colour factors) for gluon splitting to two gluons (solid line) and quark radiating a gluon (dash line) calculated according to the Altarelli-Parisi equations as a function of fraction of energy carried away by the radiated gluon, $(1 - z)$. 14

1.6 Notation for $g \to gg$ branching, courtesy of [22]. 15

1.7 The plot shows the rate for the dijet process with gg, qg and qq as a function of the transverse energy of the jets, courtesy of [22]. 18

2.1 A schematic of the chain of accelerators at Fermilab. The direction of protons and antiprotons are shown. Different machines and experiments are also labeled. Fermilab Neg. #: 00-0635D. 23

2.2 The peak luminosities in stores and the percentage loss of the protons and antiprotons in the Tevatron for the data used in this analysis as a function of time. The red and black regions represent the loss of protons and antiprotons, respectively, during the injection to Tevatron. The green corresponds to the loss arising from proton-antiproton collisions where the beams pass through each other. 26

2.3 Elevation view of one-half of the CDF II detector. 27

2.4 The geometry of the CDF II tracking system. Different detectors are labeled. 28

2.5 1/6 section of the COT end plate. The total number of supercells, the wire orientation and the average radius of the superlayers are given. The enlargement shows the sense and field slot geometry. Dimensions are in cm. 30

2.6 A block diagram of the data flow at CDF. 35

2.7 The block diagram of the Level-1 and Level-2 trigger. 37

3.1 Comparison of the charged particle multiplicity distributions between $t\bar{t}$ events produced through gg fusion (red) and $q\bar{q}$ annihilation (blue) using PYTHIA MC calculations. The normalization is arbitrary, with each distribution having equal number of entries. 50

3.2 Comparison of the invariant mass of the $t\bar{t}$ system between $t\bar{t}$ events produced through gg fusion (red) and $q\bar{q}$ annihilation (blue) using PYTHIA MC calculations. The normalization is arbitrary, with each distribution having equal number of entries. 51

3.3 Comparison of the transverse momentum of the $t\bar{t}$ system between $t\bar{t}$ events produced through gg fusion (red) and $q\bar{q}$ annihilation (blue) using HERWIG MC calculations. The normalization is arbitrary, with each distribution having equal number of entries. 52

3.4 Comparison of the pseudorapidity of the $t\bar{t}$ system between $t\bar{t}$ events produced through gg fusion (red) and $q\bar{q}$ annihilation (blue) using HERWIG MC calculations. The normalization is arbitrary, with each distribution having equal number of entries. 53

3.5 The distribution of lepton isolation vs. $\not\!\!E_T$ for both electrons and muons used to estimate the QCD background in the W+0 jet sample. Number of events found in each region and the QCD background fraction found are also shown on the plot. 60

3.6 A schematics of the unrolled central calorimeter with a low E_T and a high E_T jet in the tracking region ($|\eta| \leq 1.1$). The area excluded from the tracking due to the presence of these jets is shown by the circles. Please note that this is not to scale. 63

3.7 The low p_T track multiplicity distribution in dijet 100-120 GeV data sample. 65

3.8 The low p_T track multiplicity distribution in dijet 160-180 GeV data sample. 66

3.9 The low p_T track multiplicity distribution in W+1 jet data sample. . . . 67

3.10 The low p_T track multiplicity distribution in W+2 jet data sample. . . . 68

3.11 The correlation between the average low p_T track multiplicity (data) and the average number of gluons (MC). The dotted line is from a linear fit to the points. 69

3.12 The $<N_{trk}>$ in the W+1 jet data sample as a function of the energy of the leading jet. 72

4.1 Comparison between the gluon-rich and no-gluon distributions and parameterizations. The vertical scale is normalized such that the distributions have unit area. 75
4.2 The fit result for the dijet sample with leading jet E_T of 80-100 GeV. . . 77
4.3 The fit result for the dijet sample with leading jet E_T of 100-120 GeV. The two components of the fit (gluon-rich and no-gluon) are also shown. 78
4.4 The fit result for the dijet sample with leading jet E_T of 120-140 GeV. The two components of the fit (gluon-rich and no-gluon) are also shown. 79
4.5 The fit result for the dijet sample with leading jet E_T of 140-160 GeV. The two components of the fit (gluon-rich and no-gluon) are also shown. 80
4.6 The fit result for the dijet sample with leading jet E_T of 160-180 GeV. The two components of the fit (gluon-rich and no-gluon) are also shown. 81
4.7 The fit result for the dijet sample with leading jet E_T of at least 180 GeV. The two components of the fit (gluon-rich and no-gluon) are also shown. 82
4.8 The fit result for the tagged $W+\geq 4$ jet sample. The two components of the fit (gluon-rich and no-gluon) are also shown. 84
4.9 The fit result for the tagged W+1 jet sample. The two components of the fit (gluon-rich and no-gluon) are also shown. 86
4.10 The fit result for the no-tag W+3 jet sample. The two components of the fit (gluon-rich and no-gluon) are also shown. 87

5.1 The expected statistical uncertainty in the measurement of the fraction of $t\bar{t}$ production through gg fusion at CDF as a function of integrated luminosity. 96

Chapter 1

Introduction

The human quest to understand nature has been a centuries-long journey. It must have been started when the primeval humans opened their eyes and were awed by nature in its purest, untouched state. Since then, we, mankind, have come a long way in understanding nature while the question concerning the very beginning of existence has been our faithful companion all along. We have developed and put to trial numerous theories that have finally evolved to our current state of knowledge. From Democritus's suggestion of the atom in 450 BCE to the experimental works of more modern scientists such as Rutherford [1] and to the global efforts of theorists and experimentalists alike using sophisticated tools the thirst for a deeper answer to how our universe came to existence has been a source of inspiration.

Only during the last century, through experimental observations and theoretical descriptions, have we learned about the fundamental constituents of matter. We understand now that an atom is made of electrons, neutrons and protons and that both proton and neutron are made up of fundamental particles called quarks. We observed another class of fundamental matter particles with properties similar to the electron, called leptons. We came to know three different types of forces that are responsible for the interactions between the matter particles. The well-known gravitational force is one of the funda-

mental forces, however it is too weak to have a noticeable effect on the interactions of the fundamental matter particles. In contrast to the gravitational force, the effects of the strong force are considerable, and perhaps best known as the force holding protons and neutrons together in the atom nuclei. The electromagnetic force, acting on particles with electric charge, and the weak force, observed in radiative decay of neutrons to protons, are considered two aspects of a fundamental force called the electroweak force.

We learned that the fundamental forces act on the fundamental particles through special particles called force carriers. The force carrier for the electromagnetic force (photon) and for the strong force (gluon) are massless whereas the weak force acts through massive particles called W and Z bosons. We have also learned that for any particle there is an antiparticle like a mirror image of the particle. These fundamental particles, their classifications and interactions are discussed in the next section. Each class of quarks and leptons consists of six members which differ in their masses and in some cases in their electric charges. The heaviest quark is called the top quark and has a mass about that of the gold atom. Because of its very large mass, top quark interactions are conceptually simpler to understand. The study of top quark properties has allowed us to explore different aspects of the quarks and their interactions.

Top quarks have only been observed in the highest energy matter-antimatter collisions possible in the laboratory, and then only through a pair-creation process where a top quark and its antimatter equivalent is created simultaneously[1]. However, the details of this production process are not well understood. We have made a first measurement of the fraction of top quark pair production through gluon-gluon fusion using data of proton-antiproton collisions at the center-of-momentum energy of 1.96 TeV collected by the CDF II detector. It is an interesting measurement as it can test our understanding of the top quark pair production. A deviation from the prediction for this fraction may

[1]There is also a strong evidence for the production of the single top quarks through electroweak interactions.

hint to the existence of unknown physics. Given that the collision of a proton and antiproton at such high energies breaks them into their constituents and results in the interaction between the quarks and gluons and that the prediction for top quark pair production depends on our knowledge of the momentum distribution of the quarks and gluons in the proton/antiproton just before they collide, this measurement could also help us to better understand the structure of the quarks and gluons inside the protons. In order to do this measurement, we have taken advantage of the fact that gluons are more likely than quarks to radiate a low momentum gluon. As such we expect to have more charged particles with low transverse momentum in events involving more gluons in the production process, as discussed later on in this chapter. We employ a data-driven method to define low-transverse-momentum charged particle multiplicity distributions for no-gluon and gluon-rich data samples and use these distributions to find the fraction of gluon-rich events in a top quark pair candidate sample. The theoretical aspects, the experimental apparatus, the method, the results and the implications of this measurement are described in this thesis. The work presented here makes a small contribution to the world-wide, historical quest to understand nature.

1.1 Our Current Understanding of Particle Physics in Brief

Physicists develop and take advantage of a theoretical model to understand and predict the behaviour of physical phenomena. In the case of particle physics, this model is called the standard model (SM) [2]. It is a framework formed in the 1960s and 1970s through both theoretical development and experimental observations. It has successfully described all particle physics phenomena, withstanding significant experimental tests. The SM was born from the combination of quantum mechanics and relativity theory, or quantum field theory. The SM is a theory with local gauge symmetries, specifically

$SU(3)_C \times SU(2)_W \times U(1)_Y$. The $SU(3)_C$ symmetry represents the quantum chromodynamics (QCD) force [3] explaining the strong interactions. The charge involved in strong interactions creates three possible states for the fundamental representation of $SU(3)_C$ symmetry and hence, analogous to the three basic colours red, green and blue, it is given the name *colour*. The strong interactions are mediated by the gluons, massless gauge bosons with spin 1 and 8 colour charges. The $SU(2)_W \times U(1)_Y$ symmetry represents the weak isospin and hypercharge[2] symmetry, respectively, defining the electroweak interactions [4]. It gives rise to the other four gauge bosons, photon (γ), $W^{\pm}$ and Z bosons. These bosons have spin 1, with γ and Z bosons being neutral while $W^{\pm}$ have electric charges of ± 1. The γ is massless, but the $W^{\pm}$ and Z bosons gain mass through a device known as the Higgs mechanism, which requires the existence of another particle called the Higgs boson, the only spin 0 particle of the SM and the only one that has not been observed experimentally. The Higgs mechanism is also responsible for the mass generation of the matter building blocks as illustrated in Fig. 1.1.

According to the SM, there exists two types of fundamental matter particles, leptons and quarks, each categorized into three generations containing two species. As shown in Fig. 1.1, the leptons consist of electron (e), muon (μ), tau (τ) and their corresponding neutrinos (ν_e, ν_μ, ν_τ). Similarly the quark section has three generations of two species, up (u), charm (c), top (t) quarks and their corresponding isospin partners, down (d), strange (s) and bottom (b). Leptons and quarks are spin one-half particles. The charged leptons (e, μ, τ) have negative unit electric charge, the neutrinos are neutral and the quarks have fractional electric charges of $+\frac{2}{3}$ for u, c and t quarks and $-\frac{1}{3}$ for the rest. In addition to the electric charge, quarks carry colour charge. The colour charge is associated with the strong interactions. The leptons do not have a colour charge and thus do not participate

[2]Weak isospin is a quantum number associated to the helicity, the direction of the spin with respect to the momentum, of fundamental fermions forming them into doublets if they couple with the charged currents and into singlets if they do not. The weak hypercharge is a quantum number of fundamental fermions relating their weak isospin with their electric charge.

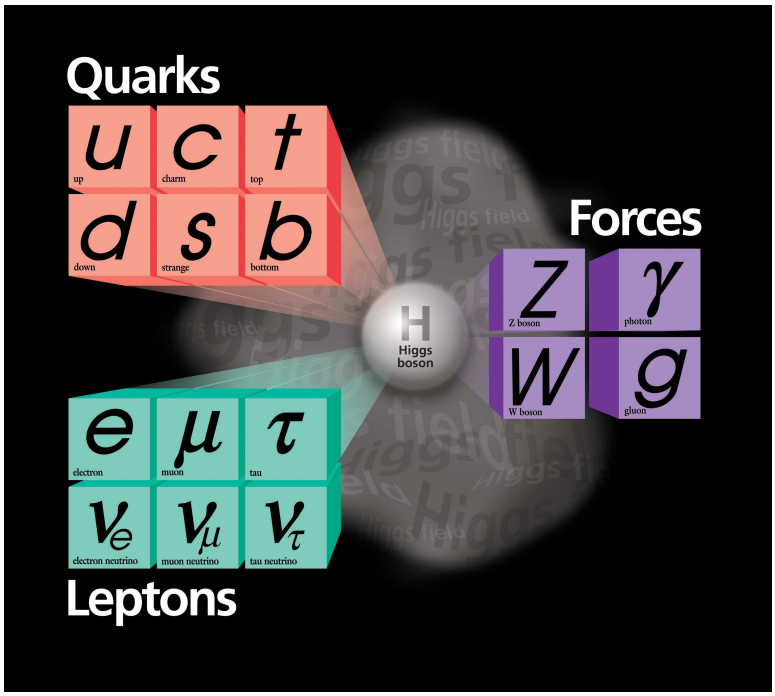

Figure 1.1: An illustration of the effect of the Higgs field in generating the mass of the quarks, leptons and force carriers. The three generations of the leptons and quarks are also shown separately. Fermilab Neg. #: 05-0440-01D.

in strong interactions. Also for each fundamental matter particle, there is an anti-matter counterpart, like a mirror image. It carries the same spin but opposite sign electric charge and anti-colour. The masses of the fundamental particles are summarized in Table 1.1 [5].

Particle	Mass
Electron	$0.51099892 \pm 0.00000004$ MeV
Muon	105.658369 ± 0.000009 MeV
Tau	$1777.99^{+0.29}_{-0.26}$ MeV
Neutrino	< 2 eV
Up quark	1.5-3.0 MeV
Down quark	3-7 MeV
Strange quark	95 ± 25 MeV
Charm quark	1.25 ± 0.09 GeV
Bottom quark	4.20 ± 0.07 GeV
Top quark	174.2 ± 3.3 GeV
Photon	$< 6 \times 10^{-17}$ eV
W boson	80.403 ± 0.029 GeV
Z boson	91.1876 ± 0.0021 GeV
H boson	> 114.4 GeV @ 95% C.L
Gluon	massless

Table 1.1: The estimated or measured mass of the fundamental particles [5].

The quarks and gluons, the coloured particles, are usually referred to as partons. Partons cannot exist as free particles (at least for more than 10^{-24} s, the time it takes for them to form a collective state). This feature is called confinement. Due to confinement, one cannot observe a free parton and as a result, two types of matter particles exist,

mesons and baryons, formed by colourless states of quarks and gluons. Mesons are formed by two (valance) quarks, such as $u\bar{d}$ forming a pion (π), and baryons are formed by three (valance) quarks, such as uud forming a proton (p). The constituents of the mesons and baryons, collectively known as hadrons, can act as free partons in interactions with very high momentum exchange, a phenomenon known as asymptotic freedom. The probability density of a parton carrying a specific fraction, x, of the momentum of the hadron in an interaction with four-momentum transfer Q^2 is given by the parton distribution functions (PDFs). The SM does not make precise predictions for the PDFs but they are rather estimated using experimental data fitted to a set of parameterizations [6]. Depending on the methods used for the PDF estimates, we get different sets of PDFs. The most commonly used sets are CTEQ [7] and MRST [8]. The distributions from the two sets are in very good agreement with each other for the quark distributions and show similar shapes for the gluon distributions. The CTEQ6M PDFs for the quarks and gluons in proton are plotted in Fig. 1.2.

This theoretical framework for the QCD and electroweak forces defines a set of rules through which one can predict the properties of particles and interactions. Among the most useful tools of the SM are the Feynman diagrams and rules that facilitate the calculations of the amplitude for a given interaction, examples of which are given in the following section. However, the mass of the fundamental matter particles and the gauge bosons as well as the coupling of the particles interacting through different forces are free parameters of the SM and have to be measured experimentally. This is one of the most significant shortcomings of the SM.

1.2 Top Quark

After the discovery of the b quark [9] in 1977, the existence of a heavier weak isospin partner, the t quark, was indicated by experimental and theoretical studies [10]. It

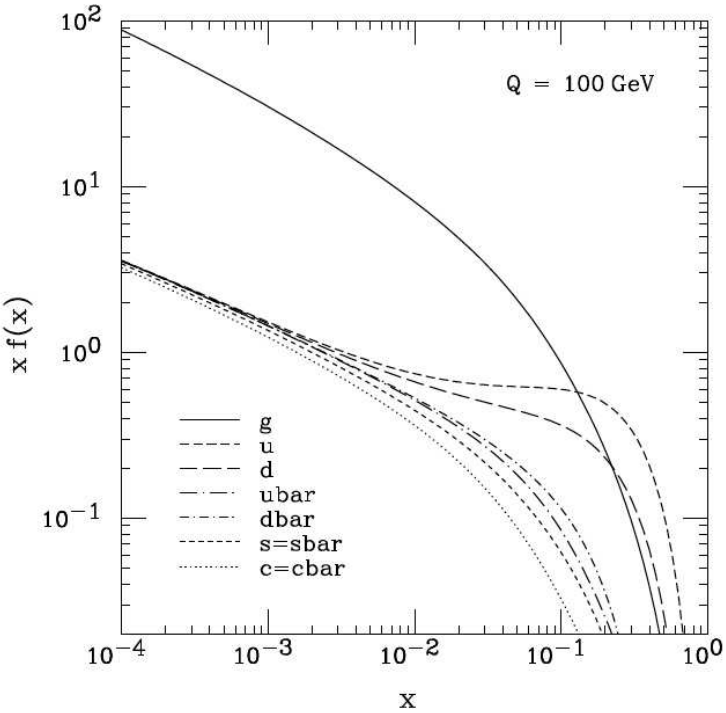

Figure 1.2: The CTEQ6M PDFs at $Q = 100$ GeV as a function of x. The distributions for different partons are shown with different line styles as specified in the legend.

was not till 1995 that the CDF and DØ collaborations reported conclusive evidence for top quark production at the Fermilab Tevatron [11]. Since then, many studies have been devoted to the understanding of this massive quark, examples of which are listed in [12]. Most notably its mass is now measured with a precision of about 1% [13] and the total cross section for its pair production is measured by many different methods and decay channels [14, 15, 16]. These are found to be in agreement with the theoretical predictions [17]. Top quark production at hadron colliders and top quark decay are discussed in the following subsections.

1.2.1 Production

According to the SM, in hadron collisions, top quarks can be produced in pairs through the strong interaction or as single top quarks through weak interactions. In proton-antiproton ($p\bar{p}$) collisions at center-of-momentum (CM) energies $\sqrt{s}$ =1.96 TeV, pair production is dominant. Here, we concentrate on the top quark pair production in hadron collisions. The $t\bar{t}$ production has a similar mechanism to the production of other heavy quarks. Using perturbative QCD (pQCD), one can calculate the cross section for heavy quark pair production [18]. The leading-order processes responsible for creation of a pair of heavy quarks Q with mass m in hadron collisions are

$$q + \bar{q} \longrightarrow Q + \bar{Q}, \tag{1.1}$$

$$g + g \longrightarrow Q + \bar{Q}, \tag{1.2}$$

where q is a light quark and g is a gluon, both being the constituents of the colliding hadrons. The Feynman diagrams of these processes are shown in Fig. 1.3. The corresponding partonic cross sections are [19]

$$\sigma_{q\bar{q}}(\hat{s}) = \frac{8\pi\alpha_s^2}{27\hat{s}}\beta\left[1 + \frac{\rho}{2}\right] \tag{1.3}$$

and

$$\sigma_{gg}(\hat{s}) = \frac{\pi\alpha_s^2}{3\hat{s}}\left[\left(1+\rho+\frac{\rho^2}{16}\right)\ln\left(\frac{1+\beta}{1-\beta}\right)-\beta\left(\frac{7}{4}+\frac{31}{16}\rho\right)\right], \tag{1.4}$$

where $\rho = 4m^2/\hat{s}$ and $\beta = \sqrt{1-\rho}$ is the velocity of the heavy quarks in the $Q\bar{Q}$ CM frame with energy $\sqrt{\hat{s}}$. In order to get the total hadronic cross section one should average the partonic cross sections over the $q\bar{q}$ and gg luminosities in hadronic collisions. This is where the PDF distributions play an important role in the prediction of the production cross sections. The more precisely the PDFs are known the more precise the prediction will be. As discussed in [20], at the Tevatron for CM energies of 1.96 TeV, using next-to-leading order (NLO) calculations we expect $(15 \pm 5)\%$ of the top quark pairs to be produced through gluon-gluon fusion and the rest through quark-antiquark annihilation. The 5% uncertainty is due to the PDF uncertainties associated with the gluons at high x. At the Large Hadron Collider (LHC) with proton-proton collisions at CM energies of 14 TeV, the gluon-gluon fusion is the dominant production mode with a rate of about 90%. Figure 1.4 shows the $t\bar{t}$ production cross section as a function of $\sqrt{s}$.

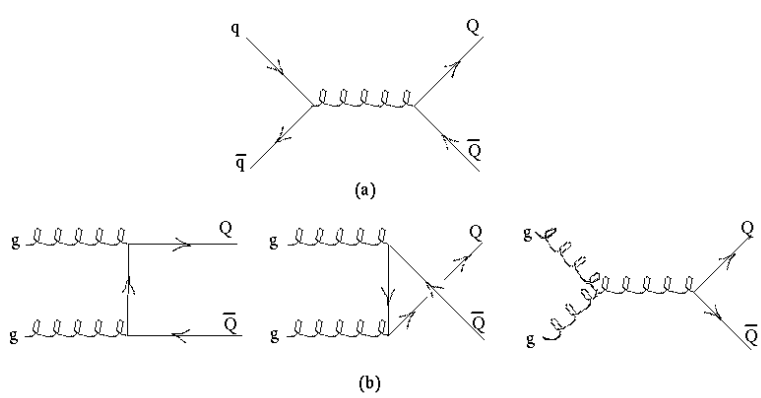

Figure 1.3: Leading-order Feynman diagrams corresponding to $t\bar{t}$ production through a) quark-antiquark annihilation and b) gluon-gluon fusion.

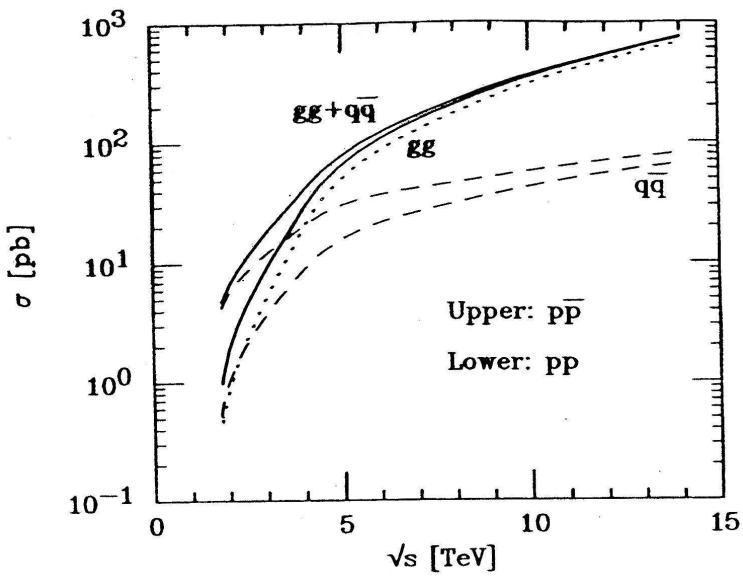

Figure 1.4: The hadronic $t\bar{t}$ production cross section as a function of the hadronic CM energy. The contribution from gg initial states (dotted line) and from $q\bar{q}$ initial states (dashed lines) are also shown. The upper (lower) dashed lines are for $p\bar{p}$ (pp) collisions.

1.2.2 Decay

The top quark is the heaviest fermion in the SM and due to its large mass, it decays before it hadronizes. Therefore, it must be observed and studied indirectly through its decay products. The SM predicts that a top quark almost always decays via the weak force to a W boson and a b quark. Then, for a $t\bar{t}$ system we have $t\bar{t} \longrightarrow bW^+ \bar{b}W^-$.

The $t\bar{t}$ decay channels are classified based on the W boson decays. A W boson can decay either to a lepton and the corresponding neutrino or to light quarks forming two hadronic jets. When both W bosons decay leptonically, we obtain a dilepton final state $t\bar{t} \longrightarrow b\ell^+\nu\bar{b}\ell^-\bar{\nu}$. In current experiments, we limit ℓ to be either an electron or a muon, given the difficulties of efficiently tagging τ leptons. When one of the W bosons decays leptonically and the other one decays hadronically, we obtain a lepton + jets final state $t\bar{t} \longrightarrow b\ell^+\nu\bar{b}q\bar{q}'$ or $t\bar{t} \longrightarrow bq\bar{q}'\bar{b}\ell^-\bar{\nu}$. Finally, when both W bosons decay hadronically, we obtain the all-hadronic final state $t\bar{t} \longrightarrow bq_1\bar{q}_1'\bar{b}q_2\bar{q}_2'$. For our analysis, we use the lepton+jet channel as described in section 3.2.4 as that provides the cleanest channel with relatively high efficiency. Please note that the dilepton channel has the lowest background, however its small branching ratio and our lower efficiency in observing leptons compared to jets statistically limit us in choosing this channel as our signal.

1.3 Parton Branching

One of the key concepts used in this analysis is the higher probability of a gluon to radiate a low momentum gluon compared to the probability of a quark doing so. This can be seen in Fig. 1.5, where the Altarelli-Parisi [21] equation, a theoretical formulation shown to effectively model this process, is used to find the splitting function values for a parton radiating a gluon as a function of the fraction of energy the radiated gluon will carry off the initial parton. The matrix element squared for $n + 1$ partons in the small

angle region[3] in terms of the n parton matrix element squared is given by

$$|\mathcal{M}_{n+1}|^2 \sim \frac{4g^2}{t} C_A F(z; \varepsilon_a, \varepsilon_b, \varepsilon_c) |\mathcal{M}_n|^2 \qquad \text{for} \quad g \to gg, \tag{1.5}$$

and

$$|\mathcal{M}_{n+1}|^2 \sim \frac{4g^2}{t} C_F F(z; \lambda_a, \lambda_b, \varepsilon_c) |\mathcal{M}_n|^2 \qquad \text{for} \quad q \to qg, \tag{1.6}$$

where g^2 is the strong coupling constant, z is the fraction of the energy the incoming parton carries after radiating the gluon, t is defined as $\frac{z}{1-z} E_a^2 \theta_b^2$, E_a is the energy of the incoming parton and θ_b is the change in the angle of the incoming parton after radiating a gluon. $\varepsilon_a(\lambda_a)$, $\varepsilon_b(\lambda_b)$ and ε_c are the polarization (helicity) of the incoming gluon (quark) before and after radiation and the polarization of the radiated gluon, respectively. The function $F(z; \varepsilon_a, \varepsilon_b, \varepsilon_c)$ can be calculated from the diagram shown in Fig. 1.6 which shows the splitting of a gluon to two gluons. Similarly the function $F(z; \lambda_a, \lambda_b, \varepsilon_c)$ can be found using similar calculations when a quark radiates a gluon.

One can define $\langle F \rangle$ by averaging $F(z; \varepsilon_a, \varepsilon_b, \varepsilon_c)$ with respect to the incoming parton polarization and summing over the polarization of the two outgoing partons. Then, according to the Altarelli-Parisi equation, the splitting function for the gluon, $\hat{P}_{gg}(z)$, is given by

$$C_A \langle F \rangle \equiv \hat{P}_{gg}(z) = C_A \Big[\frac{z}{1-z} + \frac{1-z}{z} + z(1-z) \Big] \tag{1.7}$$

and analogously for $q \to qg$, the spin-averaged splitting function for a quark radiating a gluon, $\hat{P}_{qq}(z)$, is given by

$$C_F \langle F \rangle \equiv \hat{P}_{qq}(z) = C_F \Big[\frac{1+z^2}{1-z} \Big], \tag{1.8}$$

where C_A and C_F are the colour factors equal to 3 for gluon and 4/3 for quark splitting, respectively [22].

[3]Where the angle between the radiated gluon and the original parton after radiating the gluon is small.

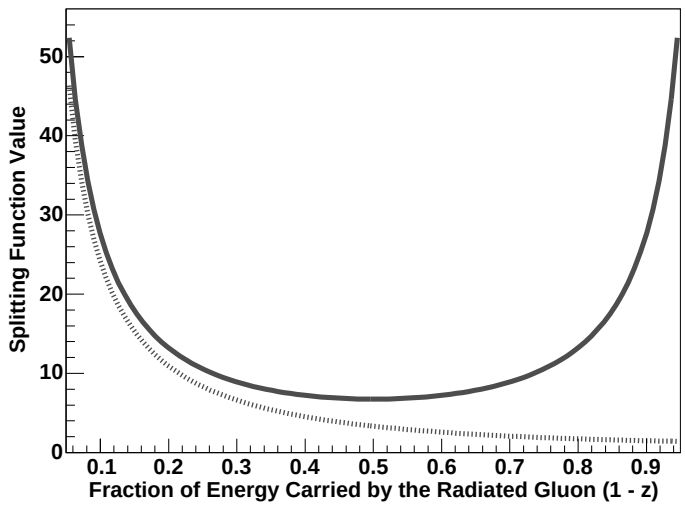

Figure 1.5: Splitting function values (including the colour factors) for gluon splitting to two gluons (solid line) and quark radiating a gluon (dash line) calculated according to the Altarelli-Parisi equations as a function of fraction of energy carried away by the radiated gluon, $(1 - z)$.

As shown above, the splitting function value is proportional to the probability of gluon radiation from a parton, with the proportionality factors common between quarks and gluons for any specific z and θ_b. Figure 1.5 shows the splitting function for gluon and quark as a function of the fraction of energy carried by the radiated gluon $(1-z)$. Given the symmetry in the case of g splitting, one expects low energy gluons to be present also when splitting results in a low z value. Therefore, it is more likely for a low energy gluon to be produced from a gluon than from a quark. As a result, one expects to see a larger number of low p_T particles in an event with a higher number of gluons, given that the gluons eventually hadronize into mesons or baryons. The larger number of low p_T particles is the way we can detect gluon-rich initial states.

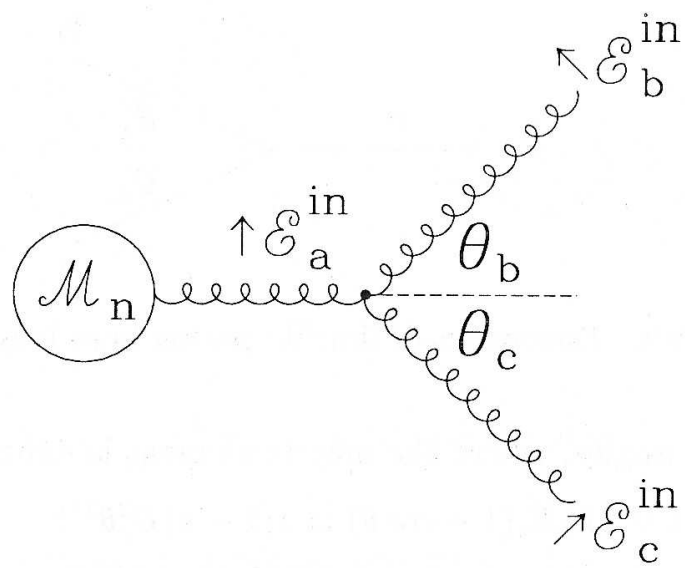

Figure 1.6: Notation for $g \to gg$ branching, courtesy of [22].

1.4 W Production

In this work, we take advantage of the difference in the gluon content of different data samples as discussed in section 1.6.2 in order to calibrate our technique, and inclusive W

boson production provides a key sample for this purpose. Here, we briefly discuss how the gluon content of the W+n jet samples changes with the number of jets. W production is a Drell-Yan process, $q\bar{q'} \rightarrow W$, with the W boson decaying to two fermions. In the case of our analysis, we require it to decay to a charged lepton and its associated neutrino. This is the production process for the W+0 parton final state. Once we have a parton produced in association with the W boson, the production process changes to $q\bar{q'} \rightarrow Wg$ or $gq \rightarrow Wq'$. As is clear from these processes, the W+1 parton process involves 1 gluon. Similarly, for the W+2, 3 and 4 partons, we expect larger number of gluons involved in the production process on average [23].

For the work presented here, we use Monte Carlo calculations to estimate the average number of gluons involved in each production process. It is worth noting that experimentally, one cannot always observe all the partons as jets and/or may observe a jet that does not come from the partons that are part of the production process. These effects are taken into account in the calculations and measurements of these processes, which are in very good agreement [24].

1.5 Dijet Production

In hadron collisions, when a parton from one of the hadrons scatters violently off a parton from the other hadron to produce two partons with high transverse momentum, the two final state particles are observed as jets, leading to a dijet process. These final state jets usually have equal and opposite momenta in their CM frame, if only two partons are produced and there is a relatively small transverse momentum associated with the initial state partons. The dijet processes provide a second calibration dataset for this analysis.

The invariant matrix element squared for $2 \rightarrow 2$ parton subprocesses with massless partons are given in Table 1.2 (taken from [25]). Also shown for comparison is the numerical value of these $\bar{\Sigma}|\mathcal{M}|^2$ at 90° in the CM frame calculated in [22]. The $\hat{s}^2$, $\hat{t}^2$ and

Process	$\bar{\Sigma}\lvert\mathcal{M}\rvert^2/g^4$	Numerical value for 90°
$qq' \to qq'$	$\frac{4}{9}\frac{\hat{s}^2+\hat{u}^2}{\hat{t}^2}$	2.22
$q\bar{q}' \to q\bar{q}'$	$\frac{4}{9}\frac{\hat{s}^2+\hat{u}^2}{\hat{t}^2}$	2.22
$qq \to qq$	$\frac{4}{9}(\frac{\hat{s}^2+\hat{u}^2}{\hat{t}^2} + \frac{\hat{s}^2+\hat{t}^2}{\hat{u}^2}) - \frac{8}{27}\frac{\hat{s}^2}{\hat{u}\hat{t}}$	3.26
$q\bar{q} \to q'\bar{q}'$	$\frac{4}{9}\frac{\hat{t}^2+\hat{u}^2}{\hat{s}^2}$	2.22
$q\bar{q} \to q\bar{q}$	$\frac{4}{9}(\frac{\hat{s}^2+\hat{u}^2}{\hat{t}^2} + \frac{\hat{t}^2+\hat{u}^2}{\hat{s}^2}) - \frac{8}{27}\frac{\hat{u}^2}{\hat{s}\hat{t}}$	2.59
$q\bar{q} \to gg$	$\frac{32}{27}\frac{\hat{t}^2+\hat{u}^2}{\hat{t}\hat{u}} - \frac{8}{3}\frac{\hat{t}^2+\hat{u}^2}{\hat{s}^2}$	1.04
$gg \to q\bar{q}$	$\frac{1}{6}\frac{\hat{t}^2+\hat{u}^2}{\hat{t}\hat{u}} - \frac{3}{8}\frac{\hat{t}^2+\hat{u}^2}{\hat{s}^2}$	0.15
$gq \to gq$	$\frac{4}{9}\frac{\hat{s}^2+\hat{u}^2}{\hat{s}\hat{u}} + \frac{\hat{u}^2+\hat{s}^2}{\hat{t}^2}$	6.11
$gg \to gg$	$\frac{9}{2}(3 - \frac{\hat{t}\hat{u}}{\hat{s}^2} - \frac{\hat{s}\hat{u}}{\hat{t}^2} - \frac{\hat{s}\hat{t}}{\hat{u}^2})$	30.4

Table 1.2: The two-to-two parton subprocesses are listed along with their invariant matrix elements squared $\bar{\Sigma}|\mathcal{M}|^2$ averaged (summed) over the colour and spin indices of the initial (final) states. Also the numerical value of these $\bar{\Sigma}|\mathcal{M}|^2$ at 90° in the CM frame is given for comparison. The variables $\hat{s}^2$, $\hat{t}^2$ and $\hat{u}^2$ are defined in the text.

$\hat{u}^2$ are defined as $(p_1 + p_2)^2$, $(p_1 - p_3)^2$ and $(p_2 - p_3)^2$ where p_i is the momentum carried by the ith parton where i =1 and 2 correspond to the two initial state partons and i =3 and 4 correspond to the final state partons. To see the effect of the jet transverse energy on the gluon content please refer to Fig. 1.7, where we show the composition of dijet events as a function of the jet transverse energy. The figure describes the results of $p\bar{p}$ collisons at 1.8 TeV, however, it is still valid as a qualitative description for CM energies of 1.96 TeV. The higher the jet transverse energy is, the lower the gluon contribution to the production process on average. This allows us to vary the gluon fraction in the calibration sample by varying the jet transverse energy threshold. As in the case of the W+n jet production, the dijet process has been extensively studied, and the theoretical description and measurements agree very well [26].

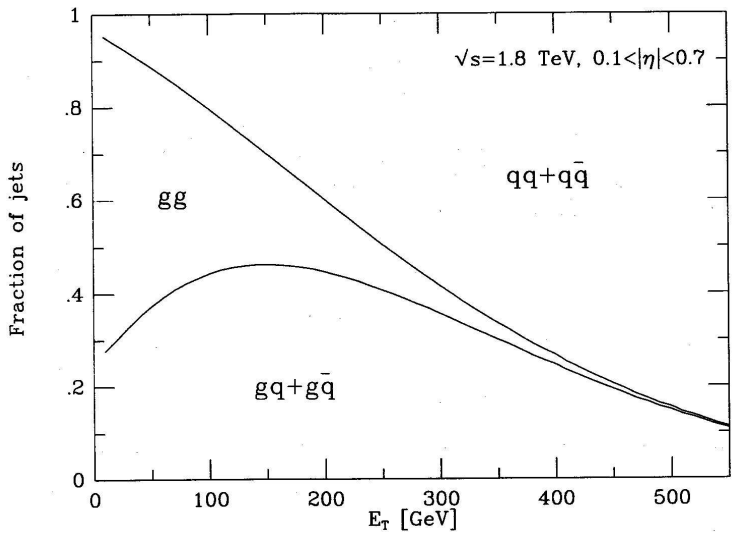

Figure 1.7: The plot shows the rate for the dijet process with gg, qg and qq as a function of the transverse energy of the jets, courtesy of [22].

1.6 Measurement

We have made a first measurement of the fraction of $t\bar{t}$ production through gluon-gluon fusion in $p\bar{p}$ collisions at $\sqrt{s}$ =1.96 TeV. We have used 0.96 fb^{-1} of data collected by the upgraded Collider Detector at Fermilab (CDF II).

After the discovery of the top quark, many studies have been dedicated to the understanding of this heavy particle. As discussed in section 1.2, the total cross section of $t\bar{t}$ production has been measured using different methods and different decay channels and the results are in good agreement with the SM predictions. However, there has been a suggestion [27] that there may exist production and decay mechanisms for the top quarks beyond the SM, where the excess in production is balanced by the non-SM decays of top quarks, with a final result in agreement with the SM. Our measurement is the first attempt to measure the fraction of $t\bar{t}$ production cross section through gluon-gluon fusion, thereby testing in more detail the SM predictions for $t\bar{t}$ production.

1.6.1 Motivation

A measurement of the fraction of $t\bar{t}$ cross section production through gluon-gluon fusion provides a test of the SM pQCD prediction. This quantity has been included in the calculations for the total $t\bar{t}$ cross section, however it had never been measured separately. Given the good agreement between the SM prediction and the experimental results for the total $t\bar{t}$ production cross section as discussed above, a deviation from the SM prediction for gluon-gluon production can point to the existence of top quark production and decay mechanisms beyond the SM.

Additionally, as discussed in section 1.2.1, the partonic $t\bar{t}$ production cross section is directly related to the partonic luminosities. Therefore, a precise measurement of the individual partonic cross section can provide information on our understanding of the partonic luminosities. According to the NLO pQCD calculations, we expect (15±5)% of the $t\bar{t}$ events to be produced through gluon-gluon fusion. The 5% uncertainty in this calculation is almost entirely due to the uncertainties in the gluon PDF. As such, one can use a precise measurement of this fraction to put stronger constraints on the momentum distribution of gluons at high x inside the proton.

1.6.2 Method

In order to make a measurement of the fraction of $t\bar{t}$ production cross section through gluon-gluon fusion, one needs to discriminate between the two $t\bar{t}$ production channels, $gg \to t\bar{t}$ and $q\bar{q} \to t\bar{t}$. Given the same final state, $t\bar{t}$, we have to find a suitable variable to discriminate between the initial state particles. We take advantage of the higher probability for a gluon than a quark to radiate a gluon which carries a small fraction of the original parton's momentum, as discussed and shown in section 1.3. We use the low energy charged particle multiplicity as our discriminating variable. As there are large uncertainties associated with the modeling and detection of low energy gluon radiation,

we employ a data-driven technique. We use two sets of calibration samples, namely the W+n jet and two-jet (dijet) samples. The larger the number of jets in the W+n jet samples and the lower the energy of the jet with the highest energy in the dijet sample, we expect a larger number of gluons involved in the production processes. We show that there is a correlation between the observed average number of charged particles with low transverse momentum and the average number of gluons involved in the production process in these calibration samples. Establishing this correlation, we then use W+0 jet and dijet events with a transverse energy of 80-100 GeV for its most energetic jet and define the no-gluon and gluon-rich distributions of the charged particle multiplicity with low transverse momentum from data, respectively. We use these two distributions in a likelihood fit to the observed multiplicity distributions in the top quark signal sample to find the fraction of gluon-rich events in our $t\bar{t}$ candidate sample.

For this analysis, we use $t\bar{t}$ events in which one of the W bosons decays to two hadrons and the other decays to a lepton and its associated neutrino. To reduce the background, we also require at least one of the jets to be identified as originating from a b quark. Once we find the fraction of gluon-rich events in the $t\bar{t}$ candidates we subtract the contribution to the gluon-rich fraction from background processes to find the gluon-rich fraction in the signal $t\bar{t}$ sample. Finally, we translate the fraction of the gluon-rich events to the fraction of $t\bar{t}$ production cross section using the acceptances for the $gg \to t\bar{t}$ and $q\bar{q} \to t\bar{t}$ processes.

We started with a brief theoretical overview of our current knowledge of the particle physics in general and the top quark in particular. We also discussed the theoretical aspects related to the analysis, namely the radiation of gluons and the production mechanisms for our calibration samples. At the end of this chapter, we motivated the measurement and gave a brief description of how the measurement is done. In chapter 2, we describe the experimental apparatus which includes the accelerator and the detector used for this analysis. We also discuss the process of data collection and reconstruction

to prepare data for the final analysis as well as the simulation tools. In chapter 3, the general analysis methodology is explained including the event selection for different signal and calibration samples. In chapter 4, we describe the method and the results of our measurement. The sources of the systematic uncertainties and their estimated associated uncertainties are also described. Finally, in chapter 5 we summarize the work presented here and briefly discuss its implications.

Chapter 2

Experiment

For this analysis, we have used 0.96 ± 0.06 fb^{-1} of data from proton-antiproton collisions at 1.96 TeV collected between March 2002 and February 2006. Protons and antiprotons are accelerated and collided in the Tevatron and the data from the collisions are collected by the CDF experiment. In this chapter, we describe the accelerator chains and the CDF II detector components used in this analysis. The trigger system, data acquisition and reconstruction as well as the data simulation are also summarized.

2.1 Accelerator

At Fermi National Laboratory (Fermilab) we produce, accelerate and collide protons and antiprotons. To do so, we use a chain of accelerators shown in Fig. 2.1. A description of the machines and methods used follows [28].

2.1.1 Proton Source

Production of protons starts with converting hydrogen gas to ionized, negatively charged, hydrogen gas (H^-) and accelerating the resultant gas to an energy of 750 KeV through a high voltage DC potential. Then the 750 KeV hydrogen ions are sent to the Linear

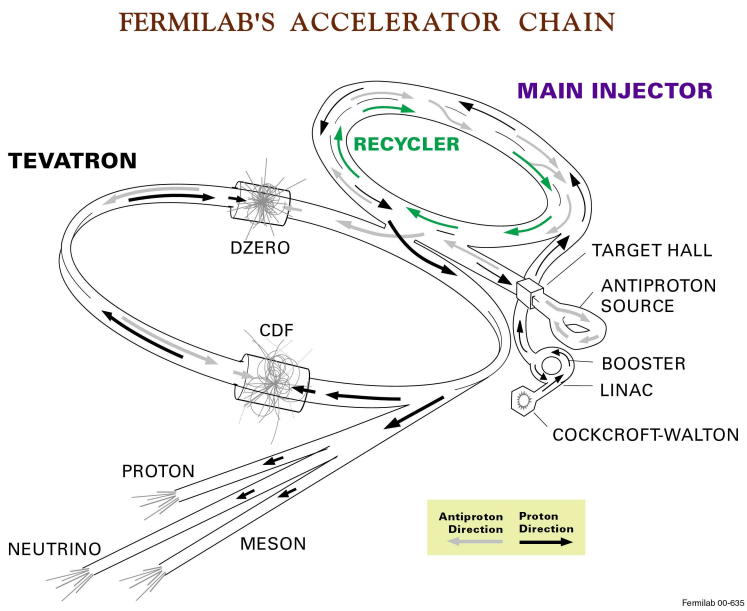

Figure 2.1: A schematic of the chain of accelerators at Fermilab. The direction of protons and antiprotons are shown. Different machines and experiments are also labeled. Fermilab Neg. #: 00-0635D.

Accelerator (Linac). Using a set of radio frequency (RF) cavities, the Linac accelerates the hydrogen ions to an energy of 400 MeV. The next level of acceleration is through the Booster, the first synchrotron in the chain of Fermilab accelerators with a radius of 75.47 m. It strips the electrons off the hydrogen ions arriving from the Linac, leaving only protons. The Booster then accelerates the protons to 8 GeV using a series of magnets and RF cavities.

2.1.2 Main Injector

The next device in the chain of accelerators is another synchrotron, the Main Injector (MI). At about 3.3 km, it has a circumference seven times the Booster and slightly more than half of the circumference of the Tevatron. The Main Injector can be used for several purposes. One is to accelerate protons from the booster to an energy of 120 GeV and send them to the Antiproton Source explained in the next subsection. Another is to accelerate either protons or antiprotons to an energy of 150 GeV and send them to the Tevatron. Protons and antiprotons arriving at the MI are formed into discrete packets of particles, called a bunch.

2.1.3 Antiproton Source

To produce antiprotons, a nickel target is struck by 120 GeV protons coming from the Main Injector. As a result a spray of secondary particles is produced. The negatively charged secondary particles are focused and rendered parallel by the use of a lithium lens. Using magnets, we collect 8 GeV antiprotons from the secondary particles. These antiprotons are then sent to a rounded triangle-shaped storage ring called the Debuncher. The primary purpose of the Debuncher is not to accelerate but to reduce the spread in momentum of the antiprotons using an RF manipulation and stochastic cooling [29]. The 8 GeV antiprotons are then transfered to the Accumulator. The Accumulator is another rounded triangle-shaped storage ring where the antiprotons are stored until being sent

to the Main Injector to be accelerated to 150 GeV and injected into the Tevatron.

2.1.4 Tevatron

The Tevatron is a synchrotron with a radius of 1 km making it the largest accelerator at Fermilab. Protons are injected into the Tevatron one bunch at a time in 3 trains of 12 bunches with 2.6 μs gaps between trains and 396 ns gaps between each bunch. Once protons are loaded, antiprotons are injected 4 bunches at a time in a similar 12-bunch-train structure. Typically, there are 20-80$\times 10^9$ antiprotons and 240-300$\times 10^9$ protons in a bunch. Electrostatic separators are used to keep proton and antiproton beams apart. A set of RF cavities are used to accelerate the protons and antiprotons to an energy of 980 GeV while the superconducting magnets are used to retain them in a circular orbit. Proton and antiproton beams are rotating in the opposite direction and once particles reach their final energy, at specific collision points, the beams pass through each other producing $p\bar{p}$ collisions at center-of-momentum energy of 1.96 TeV. The collisions continue as long as the density of particles in the beam is not too low or until there is a disruption in the process. The period after the collisions start till we drop the beam (or it is abnormally terminated) is called a store. A typical store can last about 20 hours and has an initial luminosity of about 10^{32} cm^{-2}s^{-1}. The Tevatron performance has been improved since the beginning of the data-taking period for this analysis, with one of the key challenges to achieve higher luminosities being the loss of protons and antiprotons in the process. Figure 2.2 shows the peak luminosities in stores and the percentage loss of the protons and antiprotons at Tevatron for the data used in this analysis [30].

2.2 CDF II Detector

During the mid-1980s, the Collider Detector at Fermilab (CDF), a multipurpose detector, was constructed and used to detect $p\bar{p}$ collisions at a center-of-momentum energy of

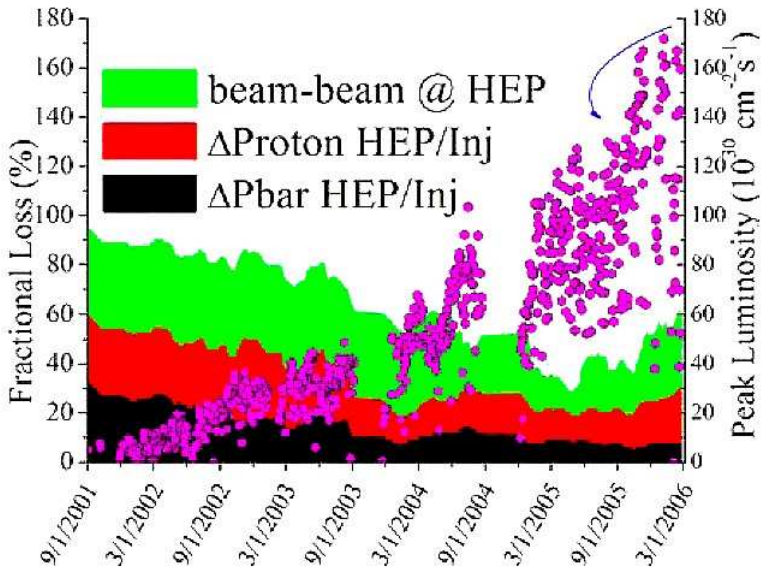

Figure 2.2: The peak luminosities in stores and the percentage loss of the protons and antiprotons in the Tevatron for the data used in this analysis as a function of time. The red and black regions represent the loss of protons and antiprotons, respectively, during the injection to Tevatron. The green corresponds to the loss arising from proton-antiproton collisions where the beams pass through each other.

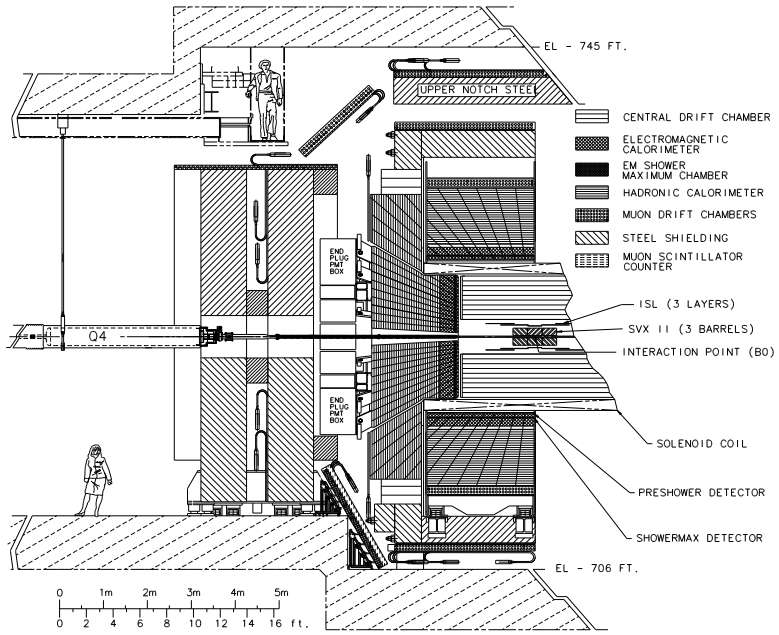

Figure 2.3: Elevation view of one-half of the CDF II detector.

1.8 TeV. It was in operation till 1996, the end of a data-taking period referred to as Run I. For the current data-taking period started in 2001, the CDF detector was upgraded and hence is called the CDF II detector. It is used to record 1.96 TeV $p\bar{p}$ collision data that are subsequently analyzed in various ways. Different components and characteristics of the detector are described in the following subsections and a detailed description of CDF II can be found in [31, 32]. The coordinate system used by the CDF experiment is a polar-cylindrical coordinate system, (z,θ,ϕ), defined such that the z axis is along the direction of protons, ϕ is the angle with respect to the x-axis in the x-y plane and θ is the angle with the z axis. The x axis is on the same plane as the Tevatron pointing outward from the center of the ring and the y-axis is perpendicular to the beam pointing upward. Other variables used very often at CDF are the pseudorapidity, η, defined as $-\log(\tan(\theta/2))$, momentum component transverse to the beam, p_T, defined as $p\sin(\theta)$, and the transverse energy, E_T, defined as $E\sin(\theta)$. Figure 2.3 shows an elevation view of one-half of the CDF II detector.

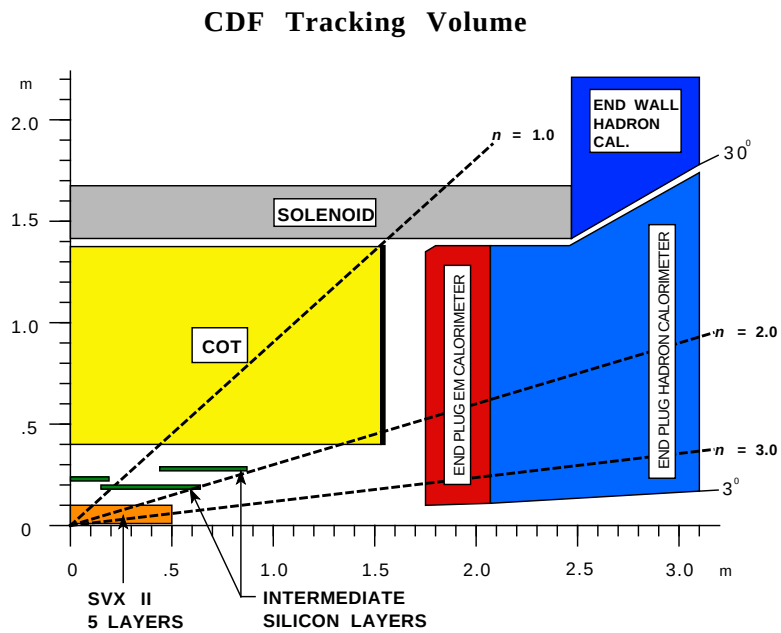

Figure 2.4: The geometry of the CDF II tracking system. Different detectors are labeled.

2.2.1 Tracking

The tracking system, which has a cylindrical structure, is the first system encountered by the particles produced in the $p\bar{p}$ collision. The tracking system consists of silicon detectors and a drift chamber. These are described in the following subsections. Figure 2.4 shows the geometry of the CDF tracking system. To make a momentum measurement of the charged particles, the tracking system is immersed in a 1.4 T magnetic field produced by a superconducting solenoid, also shown in Fig. 2.4.

Silicon Detectors

Silicon detectors are designed to precisely determine the three-dimensional impact parameter of tracks to be used for identifying short-lived hadrons using a set of silicon microstrip detectors. The central portion of the silicon detectors used to find the vertex is called the SVX II [33] with a length of 86 cm parallel to the beam. It consists of 5 layers of double-sided silicon sensors with a combination of axial (parallel to the beam)

and 90° and small-angle stereo (angles with respect to the beam) layers. The charged particles passage through the semiconductors leaves a group of ions and free electrons. The ionization in each silicon layer is sensed at the closest electrode and used to determine the position of the charged particle. The axial layers are used to measure the r-ϕ position, the 90° stereo layers are used to measure the z component and the small-angle stereo layers provide the r-z information when combined with the other measurements. The SVX II detector layers reside at a distance from 2.45 cm to 10.6 cm from the beam. There is a radiation hard silicon layer, L00 [34], consisting of one-sided axial layers which is located inside the SVX II right on the pipe in which the colliding beams pass through the detector. To increase the η coverage, an Intermediate Silicon Layers Detector (ISL) is used [35], consisting of a central cylindrical layer at a distance of 22 cm from the beamline and two cylindrical layers in the forward region at distances of 20 and 28 cm from the beam. The ISL also consists of double-sided silicon detectors. The L00, SVX II and ISL detectors together provide an impact parameter resolution of 40 μm for high energy tracks.

Central Outer Tracker

The Central Outer Tracker (COT) is a cylindrical open cell drift chamber 310 cm long along the beam direction and sitting within the radial interval of 40-140 cm from the beam right after the silicon detectors as shown in Fig. 2.4. The COT [36] consists of 96 sense wire layers that are grouped into 8 sets (superlayers). The superlayers are arranged radially parallel to the beam and are divided in ϕ to form supercells. Each supercell consists of 12 sense wires alternating with potential wires and a field sheet on each side. Approximately half of the sense wires are axial and half are stereo with $\pm 2^o$ angles with respect to the axial wires set in alternate superlayers. The wires are set such that the maximum drift distance is about 0.88 cm. Figure 2.5 shows a 1/6 section of the COT end plate and relevant numbers as well as an enlargement of the sense and field

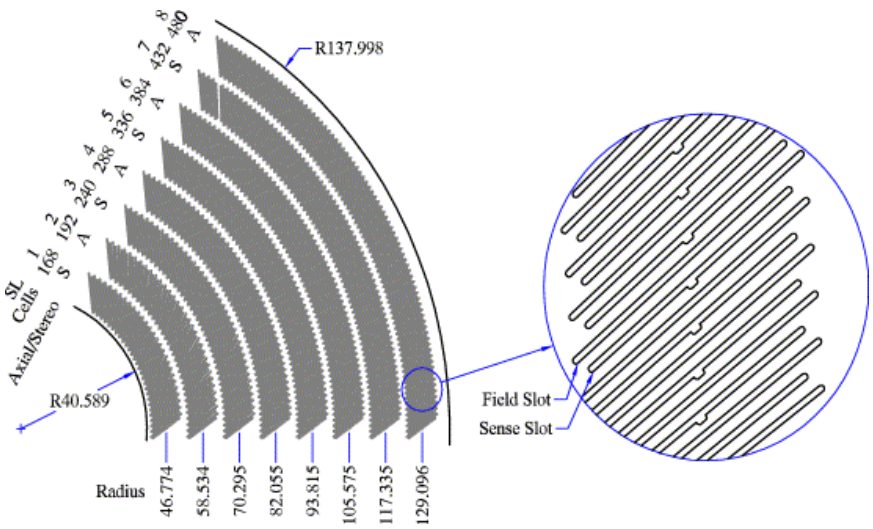

Figure 2.5: 1/6 section of the COT end plate. The total number of supercells, the wire orientation and the average radius of the superlayers are given. The enlargement shows the sense and field slot geometry. Dimensions are in cm.

slot geometry. The COT is filled with a gaseous mixture of 50% argon and 50% ethane providing a maximum drift time of 177 ns. Similar to the silicon detectors, the COT detects the ionization left by a charged particle, drifting the electrons to the closest sense wire. The time of drift is converted into a position of the ionization deposition. The COT has a hit position resolution of 140 μm and a momentum resolution, $\sigma(p_T)/p_T$, of $0.0015p_T$, where p_T is given in GeV/c. The tracking reconstruction is highly efficient for tracks with $p_T \geq 250$ MeV/c.

2.2.2 Calorimeter System

The calorimeter system is used to measure the energy of the particles produced by the $p\bar{p}$ collisions and to identify electrons, photons and hadrons. It also measures the overall energy flow of the $p\bar{p}$ interactions to infer the presence of noninteracting particles such as neutrinos and help to identify muons. The calorimeter system at CDF, formed by seven subsystems based on their η coverage and type of calorimeter, consists of projective towers

and covers the $|\eta| < 3.6$ region. The central calorimeter system, $|\eta| < 1.1$, includes the electromagnetic calorimeter (CEM) [37], the position detector located at shower maximum (CES), the hadronic calorimeter (CHA) and the endwall hadronic calorimeter (WHA) [38]. The WHA calorimeter extends the coverage to $|\eta| < 1.2$. The central calorimeter subsystems are cylindrical and are located after the solenoid as shown in Fig. 2.3. The coverage in the forward region, $1.1 < |\eta| < 3.6$, is provided by the forward calorimeter system consisting of the plug electromagnetic calorimeter (PEM) [39], the plug position detector at shower maximum (PES) [40] and the plug hadronic calorimeter (PHA) [32]. The general construction of the electromagnetic and hadron calorimeters as well as the shower maximum detectors are described below. The details for each system are given in Table 2.1.

Calorimeter	CEM	CHA (WHA)	PEM	PHA
Total Thickness	$19\chi_0$, 1λ	4.5λ	$21\chi_0$, 1λ	7λ
Lead/Iron Layer	$0.6\chi_0$	1 in. (2 in.)	$0.8\chi_0$	2 in.
Scintillating Layer	5 mm	10 mm	4.5 mm	6 mm
$\sigma(E)/E$	$13.5\%/\sqrt{E_T}$	$50\%/\sqrt{E}$	$14.4\%/\sqrt{E}\oplus$ 0.7%	$80\%/\sqrt{E}\oplus$ 5%

Table 2.1: The parameters for different calorimeters. The unit λ is the attenuation length and χ_0 is a radiation length. The energy is measured in units of GeV.

Electromagnetic Calorimeters

The electromagnetic calorimeters are made of alternating lead and scintillating sampling sheets. Towers are organized in η and ϕ. In the central region, the towers are segmented by 0.11×15° in η-ϕ and in the forward region they have granularities ranging from $\Delta\eta$ between 0.09 and 0.64 and $\Delta\phi$ of either 7.5^o or 15°. The scintillation light is lead to the photomultiplier tubes by wavelength shifters that are located along the tower wedges in ϕ in case of the central calorimeters. In the case of the plug calorimeters, the wavelength

shifting fibers are embedded in the scintillators and spliced into clear optical cables that carry the light to the photomultiplier tubes.

Hadronic Calorimeters

The hadronic calorimeters, which follow the electromagnetic calorimeters, consist of alternating iron and scintillating layers and are also divided into towers in η and ϕ. As much as possible, the same segmentations as for the electromagnetic calorimeters are used for the hadron calorimeters, however in some cases a hadronic tower is shared between two electromagnetic towers. Similar to the electromagnetic calorimeters, the light is transfered by the wavelength shifters or clear optical fibers to the photomultiplier tubes.

Shower Maximum Detectors

The CES detector [37] provides shower position measurement and transverse shower development at shower maximum, about $6\chi_0$, within the CEM calorimeter. This is done by measuring the charge deposition on orthogonal wires and strips in a gaseous mixture of 95% Ar and 5% CO_2. The CES has a position resolution of 2 mm at 50 GeV. Similar to the CES detector, a shower maximum position detector, PES, is located within the PEM calorimeter. The PES detector uses scintillating strips for the position measurements.

2.2.3 Muon System

The muon system, used to identify muon candidates, includes the central muon detector (CMU) [41], the central muon upgrade (CMP) and the central muon extension (CMX) [42]. All of these systems detect charged particles that have passed through significant amount of material. The CMU detector, a set of single wire drift cells, is located right after the CHA. The drift cells in the CMU are arranged such that they align with the calorimeter tower wedges. They cover a range of $|\eta|$ <0.6. The CMP,

the second set of drift chambers, sits behind extra shielding of steel and the CMU. Each of the CMU and CMP can provide four position measurements of a muon candidate penetrating through the solenoid, calorimeter, and shielding. The CMX is another set of the drift cells arranged in the intermediate angle region and extending the coverage to $|\eta| < 1$. The minimum detectable muon p_T is 1.4, 2.2 and 1.4 GeV/c for the CMU, CMP and CMX detectors, respectively [32].

2.2.4 The Luminosity Counters

The $p\bar{p}$ instantaneous luminosity, $\mathcal{L}$, is related to the frequency of the p and $\bar{p}$ bunch crossing, f_{BC}, average number of $p\bar{p}$ interactions per bunch crossing, μ, and the cross section for inelastic $p\bar{p}$ interactions, σ, as

$$\mathcal{L} = f_{BC}\mu/\sigma. \tag{2.1}$$

The inelastic $p\bar{p}$ cross section has been measured to be σ =59.3 mb, f_{BC} is 2.5 MHz and so if we measure μ, we can find $\mathcal{L}$. In order to do so, we use the Cherenkov Luminosity Counters (CLC) located on the two sides of the CDF II detector close to the beam. These counters cover the region $3.7 < |\eta| < 4.7$. The CLC consists of 48 conical Cherenkov counters on each side of the detector arranged in three concentric layers pointing to the center of the interaction region. The counters are filled with isobutane gas at atmospheric pressure [43]. The Cherenkov light produced by the charged particles passing through the gas in each counter is collected by a photomultiplier tube located at the end of the cone away from the center of the CDF II detector. A charged particle from a $p\bar{p}$ interaction passing through a cone results in about 100 photoelectrons. For the period of data taking for this analysis, μ at CDF was measured by counting the number of bunch crossings with no interactions. Taking advantage of the Poisson distribution of the number of interactions per bunch crossing corrected for the acceptance of the CLC, we then translate the number of empty bunch crossings to μ. The CLC acceptance and

operation introduces a 4.2% uncertainty into the luminosity measurement for the dataset used for this analysis.

2.3 Trigger and Data Acquisition Systems

The high rate of $p\bar{p}$ interactions in the Tevatron does not allow the experiments to permanently record the data from the detectors for every single interactions. The detectors recording speed is less than 50 Hz while the bunch crossing rate is about 2 MHz. Therefore, at CDF, we make use of a three-level trigger system to keep the most interesting interactions with high efficiency. Each level achieves a reduction in rate such that the next level has more time to process the data. The key performance goal is to minimize "deadtime." Deadtime occurs when the lower level has to reject candidate collisions without processing because the next level cannot accept an event. The Trigger System Interface (TSI) and clock systems synchronize the trigger and Data Acquisition (DAQ) systems. Each trigger level is explained in more detail below. The specific trigger paths for collecting the data samples used in this analysis are also described. Figure 2.6 shows a block diagram of the data flow at CDF. Please note that the clock values correspond to a 132 ns bunch crossing, whereas we used a 396 ns bunch crossing for this data. The corresponding 396 ns bunch crossing values are given in the text.

2.3.1 Level-1 Trigger

Using a subset of detector information, the Level-1 trigger finds physics objects and makes a decision based on the number and energies of these objects. The latency (*i.e*, the time available to make a decision) for the Level-1 trigger is about 5.5 μs, which requires a minimum of 14 beam-crossings of data to be stored in a local buffer for each detector while a decision is being made at Level-1. Upon an "accept" decision made by the Level-1 trigger, the information is passed to the Level-2 trigger. Once this is done,

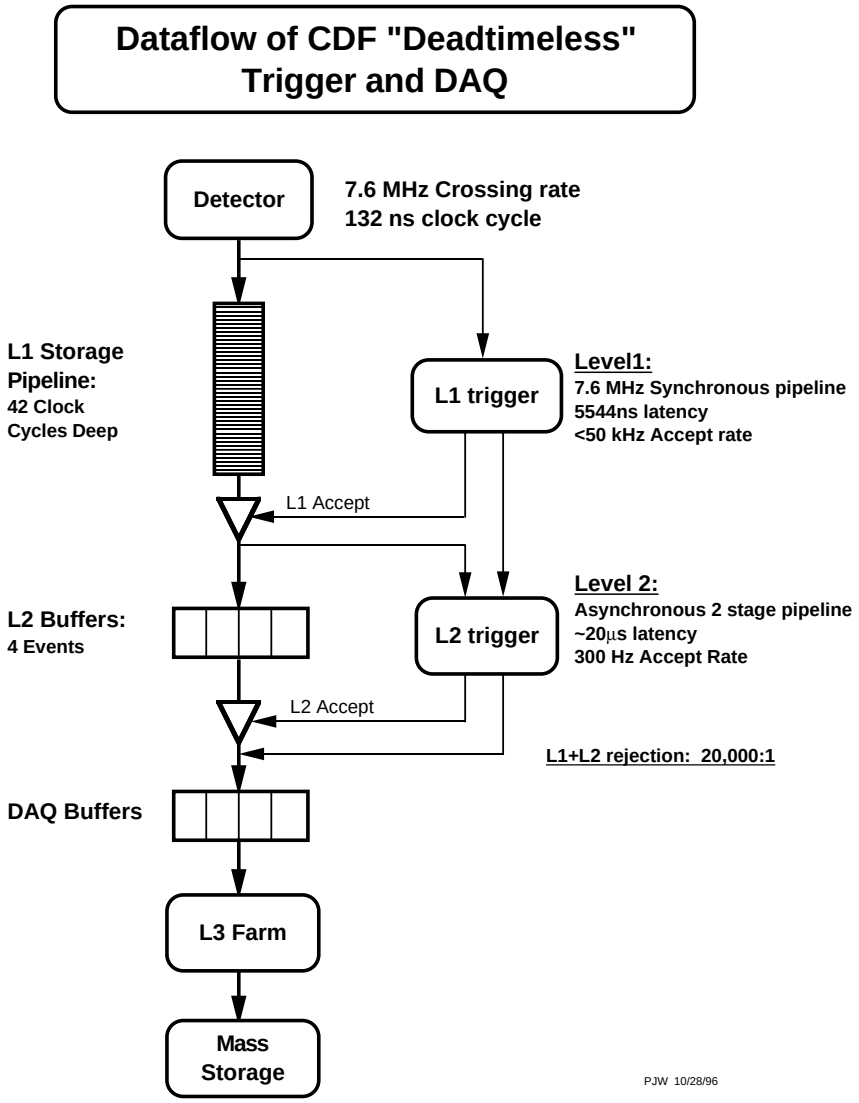

Figure 2.6: A block diagram of the data flow at CDF.

the local buffer is free to accept information for the next event. The processing time of the Level-2 trigger is on average about 20 μs and therefore the data from Level-1 trigger has to be stored in Level-2 trigger buffers. The procedure allows for a 40 kHz rate of data flow to the Level-2 trigger with a deadtime $\leq$10%. Figure 2.7 shows a block diagram of the information used by the Level-1 and Level-2 triggers. A processor, the eXtremely Fast Tracker (XFT), is used to quickly find track candidates at the Level-1 trigger rate and provide that to the Level-2 trigger. The XTRP unit extrapolates the XFT tracks to match the hits in the muon detectors.

2.3.2 Level-2 Trigger

The Level-2 trigger takes advantage of the same information available at the Level-1 with higher precision in addition to more information from a few other detectors to perform a limited event reconstruction, as shown in Fig. 2.7. The Level-2 buffer stores up to 4 events to be processed. Upon a Level-2 acceptance, the DAQ system collects the information from all the detectors for the event in the buffers to be transfered to the Level-3 trigger. The acceptance rate of the Level-2 trigger is limited to about 300 Hz.

2.3.3 Level-3 Trigger

The Level-3 trigger receives the complete information from all the detectors for the Level-2 accepted events and using a large computer farm reconstructs each event using the same algorithms employed for subsequent analysis. The reconstructed events are checked against a set of requirements which are used to categorize the reconstructed events for storage or monitoring purposes. Once satisfying the Level-3 trigger requirements, the reconstructed events are passed on to the Data Logger system. Based on the event categories, the Data Logger subsystem sends the data either for monitoring or for storage and subsequent processing. The monitoring system checks the integrity of the collected data and/or uses the data for calibration purposes.

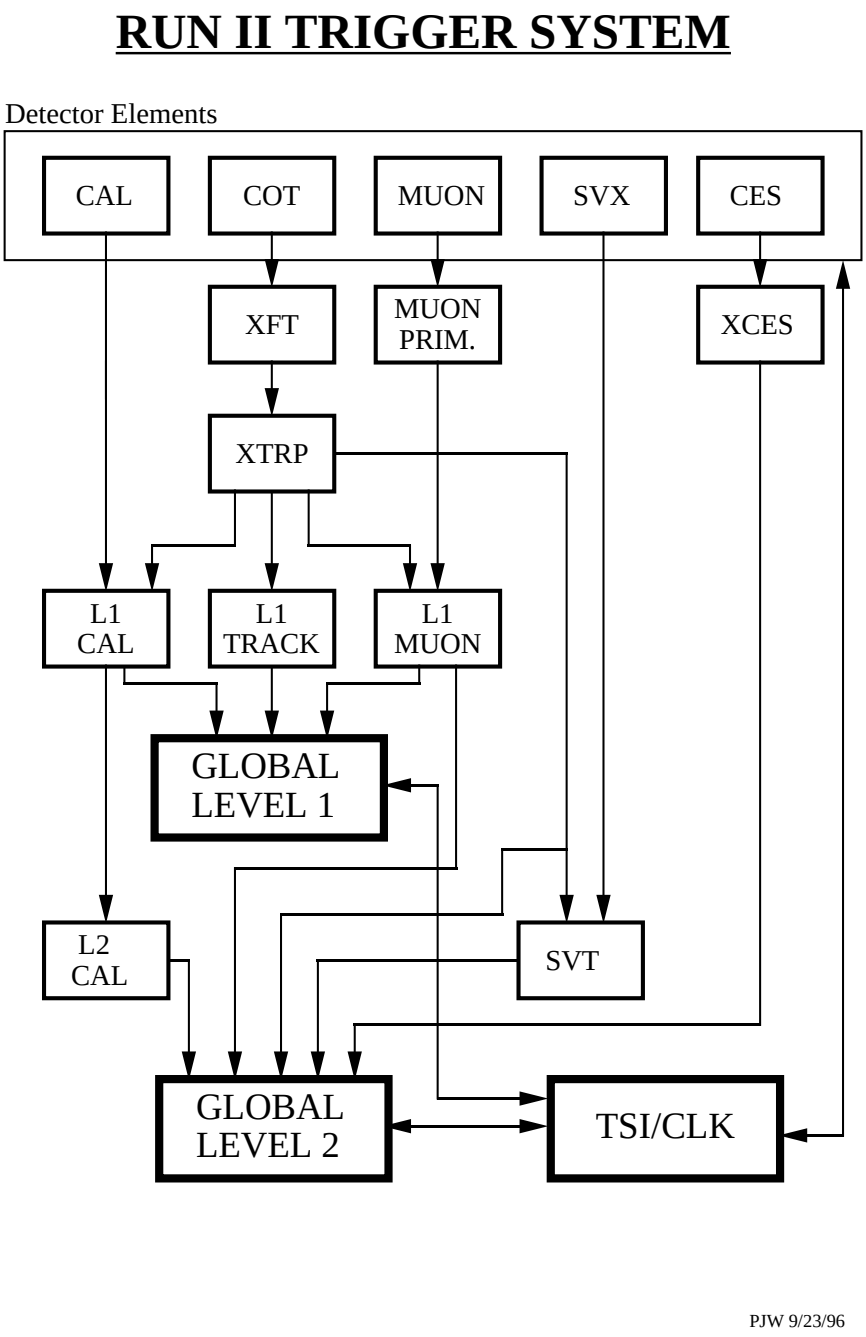

Figure 2.7: The block diagram of the Level-1 and Level-2 trigger.

2.3.4 Trigger Paths for This Analysis

The data samples used for this analysis, as detailed in section 3.2, are collected using different trigger requirements. These requirements are listed below for the three sets of triggers used for this measurement.

High-p_T Electron Trigger

The high-p_T lepton trigger includes high-p_T electron and high-p_T muon trigger paths. In this analysis, these are used for the selection of W boson and $t\bar{t}$ candidates. In the high-p_T electron trigger path, at Level-1, we require at least one calorimeter tower[1] with $E_T \geq 8$ GeV, with a fraction of energy deposition in the hadron tower over the energy deposition in the electromagnetic tower (E_{had}/E_{em}) of no more than 0.125 and an XFT track with p_T more than 8 GeV/c pointing toward the tower. At Level-2, the trigger towers with $E_T \geq 7.5$ GeV adjacent to the seed tower found at the Level-1 are clustered with the seed tower. We require a cluster of towers with $E_T \geq 16$ GeV and $|\eta| < 1.3$, $E_{had}/E_{em} \leq 0.125$ and an XFT track with p_T of at least 8 GeV/c pointing to the cluster. Finally at Level-3, we require a cluster of at least three towers[2] in the central calorimeter with $E_T \geq 18$ GeV, the shower profile consistent with that of a showering electron, $E_{had}/E_{em} \leq 0.125$, and a COT track pointing to the energy cluster with p_T of at least 9 GeV/c.

High-p_T Muon Trigger

The high-p_T muon trigger used for this analysis consists of two different trigger paths, one selects muons using the CMU/CMP (CMUP) detectors and the other using the CMX detector. For both CMUP and CMX paths, at Level-1, we require a muon "stub" with a minimum p_T of 6 GeV/c, where a stub is a set of hits in the muon chambers that match

[1] A tower in this context, for Level-1 and Level-2, is typically larger than the full segmentation.
[2] For Level-3, we use the full calorimeter segmentation.

each other and happen synchronously with the bunch crossings. For the CMUP path, we also require a stub in the CMP detector to match that of the CMU detector. An XFT track with a minimum p_T of 4 GeV/c and 8 GeV/c should also match the CMUP and CMX stub, respectively. At Level-2, the XFT track is required to have a higher quality and p_T of at least 8 GeV/c and 10 GeV/c for the CMUP and CMX path, respectively. At Level-3, for both paths, we require a match between a fully reconstructed COT track with a minimum p_T of 18 GeV/c and the stubs in the respective muon detectors, CMUP or CMX.

Jet Triggers

The two jet triggers used for this analysis are labelled Jet50 and Jet100. For the Jet50 (Jet100) path, at Level-1, we require at least one trigger tower with a minimum E_T of 5 (10) GeV in the central or plug calorimeter. At Level-2, calorimeter clusters are formed around a tower with $E_T \geq 3$ GeV by adding the adjacent towers with a minimum E_T of 1 GeV. We require at least one such cluster with $E_T \geq 40$ (90) GeV for the Jet50 (Jet100) trigger path. Finally at Level-3 a cone-based clustering algorithm, as discussed in section 2.4.4, with a cone size of 0.7 is used to form jet candidates. We require at least one jet with $E_T \geq 50$ (100) GeV for the Jet50 (Jet100) trigger path, which is where these triggers derive their names.

2.4 Reconstruction

A careful offline data processing is the last step before having the data ready for different studies. A set of algorithms are used to perform the task of translating the information received from the detector into specific physics objects that enable the physicists to find the signature of particles produced by each $p\bar{p}$ interaction. The methods used to reconstruct the physics objects employed in this analysis are described in the following

sections.

2.4.1 Track Reconstruction

The hit information collected by the silicon detectors and COT is used to reconstruct the paths of the charged particles (tracks) in the detector and measure their charge and momentum. The tracks are placed into three different categories, COT tracks, Silicon Standalone tracks and the Inside-Out tracks. To reconstruct the COT tracks, first we find axial and stereo track segments in each superlayer. A segment is a set of hits in one or few adjacent cells matched based on their time difference and the drift velocity. Then the axial segments are matched to other axial segments to form $r-\phi$ tracks with a common curvature, C. As an alternative way of finding tracks, the axial hits are filled in a histogram based on C. Then using the histogram-linked hits and segments, we make $r-\phi$ tracks. Next the stereo segments are attached to the $r-\phi$ tracks to make three-dimensional (3D) tracks. We now have two sets of 3D tracks and so as the last step, we merge the two sets together removing the duplicate tracks.

To find the Silicon Standalone tracks, we start with the $r-\phi$ hit selection. We require at least 4 $r-\phi$ hits and fit them with a curve to get the $r-\phi$ track parameters. Then we search for the corresponding $r-z$ hits, first in the small-angle stereo layers and then in the 90° stereo layers. We require the track to have at least one small angle and one 90° hit. Finally, a minimum χ^2 helical fit is used to find the track and its parameters.

In order to increase the efficiency for finding those tracks which may not pass through the first 4 layers of the COT, the Inside-Out tracks are defined. This tracking algorithm starts with the Silicon Standalone track and defines a road within the COT detector and identifies those hits within the road that form a short COT track. The hits found in this way are fit using the impact parameter and the z_0 parameters of the silicon track as constraints. The Inside-out track is defined by the parameters from this fit.

2.4.2 The z Vertex Reconstruction

Using the tracking information, one can find the $p\bar{p}$ collision point, the primary vertex, along the z axis. The algorithm starts with a candidate list of vertices, either from a pre-tracking vertex finder or from a COT-based vertex finder. The vertices from the list are then associated with reconstructed tracks in the event. We use a set of good quality tracks, having required a minimum number of hits in COT or silicon detectors, and assign to them specific quality values[3] to do so. The tracks should have an impact parameter ≤ 1 cm. A z vertex has a quality value defined as the sum of the quality values for each of its associated tracks. The vertices used in this analysis must have a minimum quality value of 12 to be considered as a good quality vertex. Virtually all the samples used in this analysis, discussed in section 3.2, have a very high vertex finding efficiency as in case of events with jets we have a high efficiency given the large number of tracks and in case of W+0 jet sample, if a z vertex with high quality is not found, we use the distance of closest approach of the track associated with the high p_T lepton in the event.

2.4.3 Lepton Reconstruction

In this section we describe the identification criteria for electrons and muons. At CDF, there are several different categories used for electron or muon identification that vary in their efficiency and purity. The main categories used in this analysis, tight electrons and tight muons, which have the most stringent requirements, are described in the following.

Tight Electrons

We start the search for candidate electrons in the events that are triggered with the high-p_T electron path explained in section 2.3.4. We search for an electromagnetic cluster formed by two adjacent towers in the central electromagnetic calorimeter region, $|\eta| \leq 1$,

[3] Assigning a value of 2, 4, 6 or 12, increasing in value for the tracks with the minimum requirements to the tracks with the most stringent requirements.

with a minimum electromagnetic E_T of 20 GeV. We require the CES hits associated with the cluster be matched to a good quality COT track with a minimum p_T of 10 GeV/c. The fraction of energy deposition in the hadronic calorimeter over the energy deposition in the electromagnetic calorimeter (E_{had}/E_{em}) must be $\leq 0.055 + (0.00045 \times E)$, where E is the energy of the cluster. The observed lateral energy profile is required to be consistent with that expected from an electron shower and the r-z profile of the shower should match that expected for electrons (CES $\chi^2 < 10$). For tracks with $p_T \leq 50$ GeV/c, the cluster E_T over the track p_T (E/p) must be ≤ 2. We also require the excess energy in towers within $\Delta R = \sqrt{(\eta_e - \eta_t)^2 + (\phi_e - \phi_t)^2}$ of 0.4 around the electromagnetic cluster to be less than 10% of the E_T of the cluster, where (η_e, ϕ_e) and (η_t, ϕ_t) are the $\eta - \phi$ coordinate of the most energetic tower in the cluster and the tower which is not part of the cluster, respectively. Additionally, we reject the electron candidate if it is consistent with coming from a photon conversion. The electron conversion algorithm rejects an electron candidate if it finds a track with opposite charge of the track associated to the electron when the two tracks are close in polar angle ($\Delta(\cot\theta) \leq 0.04$) and appear to come from a common point in the r-ϕ plane ($\Delta(xy) \leq 2$ mm).

Tight Muons

We start the search for candidate muons in the events that are triggered with the high-p_T muon paths explained in section 2.3.4. For both tight CMUP and tight CMX muons, we require a good quality COT track with $p_T \geq 20$ GeV/c matched to a muon stub in the CMUP and CMX detectors, respectively. After being extrapolated to the muon chambers, the track should fall within 7, 5 and 6 cm of the stub in the r-ϕ plane for the CMU, CMP and CMX detector, respectively. The energy deposited in the calorimeter towers associated with the track and stub must be consistent with that of a minimum ionizing particle. We also apply an isolation cut similar to that of the electron where the excess energy within 0.4 in η-ϕ of the muon towers must be less than 10% of the

track p_T. Additionally, the muon candidate should not be consistent with a penetrating cosmic ray. To reject such cases, we require that the track originates from the primary vertex and that the time for the energy deposition in the calorimeter towers associated with the muon be synchronized to the $p\bar{p}$ collision.

2.4.4 Jet Reconstruction

High-energy partons created in $p\bar{p}$ collisions "shower" into a collection of quarks and gluons that ultimately fragment into hadrons. These hadrons and particles will form a spray of particles called a jet, detected in the calorimeter systems. At CDF, jets are defined using a clustering algorithm with specific cone sizes known as JetClu [44]. The algorithm starts with geometrical "seed" towers with a minimum E_T of 1 GeV and adds the E_T of towers with at least 1 GeV which fall within a distance $\Delta R = \sqrt{(\eta_c - \eta_t)^2 + (\phi_c - \phi_t)^2}$ of the seed tower. The cone size, ΔR, is typically 0.4 in this analysis. The (η_c, ϕ_c) and (η_t, ϕ_t) are the $\eta - \phi$ coordinates of the center of the seed tower and the calorimeter tower to be clustered, respectively. The E_T of the tower is the sum of the electromagnetic and hadronic energy deposited in the tower multiplied by $\sin\theta$. After the cluster is defined, the E_T-weighted centroid of the cluster is calculated. The clustering procedure is iterated with the new centroid replacing the previous seed till the centroid coordinate is stable. It is worth mentioning that for this analysis the towers associated with tight electron candidates are excluded from the clustering procedure.

The jets reconstructed as the clusters of energy described above underestimate the E_T of the parton producing the jet. A set of energy corrections [44] are applied to the jets to estimate more accurately the true energy of the parton. The E_T of the jet before any correction is referred to as "raw E_T". The corrections, applied consecutively, are listed below.

- **A Relative Correction** is applied to raw jet energy to make the calorimeter response uniform as a function of pseudorapidity.

- **The Multiple Interaction Correction** is applied to subtract the energy deposited by the particles produced through extra $p\bar{p}$ interactions represented by the extra z vertices in the event.

- **The Absolute Energy Correction** is applied to correct for the non-linearity and energy loss in the un-instrumented regions of the calorimeter. The absolute energy scale of the electromagnetic calorimeters was set through an extensive study of $Z \rightarrow e^+e^-$ events in Run I and Run II, using inclusive electron candidates to determine any time-dependent shifts. The absolute energy scale of the hadronic calorimeter was established through a study of the minimum-ionizing energy deposition of muons from J/Psi decay, as well as higher momentum muons. The overall jet energy scale was then checked using jet-photon balancing.

We either correct the jets for the η-dependance and the multiple interactions, referred to as L4 corrections, or we also apply the absolute energy correction, which historically we refer to as L5 corrections.

2.4.5 Secondary Vertex Reconstruction

The $t\bar{t}$ signature, four jets and a W boson observed through finding an electron or a muon candidate and evidence for the presence of a neutrino, is the same signature for W boson production in association with multiple partons. In order to reduce this QCD background process, we require that at least one of the jets in $t\bar{t}$ candidate events be identified as originating from a b quark [14]. To do so, we take advantage of the long lifetime of the B-hadrons and search for the presence of a secondary vertex formed by the decay daughters of the B-hadron inside the jet. For each jet, we search for tracks with good quality, a minimum p_T and number of hits, and a large impact parameter d_0[4] associated with the jet. We require at least two good quality tracks to form a vertex

[4] $d_0 > 2\sigma_{d_0}$, where σ_{d_0} is the uncertainty in the impact parameter measurement.

within the jet with a two dimensional decay length $L_{2D} \geq 3\sigma_{L_{2D}}$ where $\sigma_{L_{2D}}$ is the total estimated uncertainty of L_{2D} and includes the uncertainty in the primary vertex. This b-tagging algorithm has an overall efficiency of about 45% in detecting at least one b quark candidate in a $t\bar{t}$ event.

2.4.6 Missing E_T Measurement

As apparent by their name, the non-interacting particles such as neutrinos escape the detectors without interacting with the material and as such create a momentum imbalance in the event. Therefore, to find evidence of their production and to measure their p_T, we apply the conservation of momentum and measure the amount of transverse energy missing in the event, hence called missing E_T. The missing E_T ($\not\!\!E_T$) is defined by $\vec{\not\!\!E}_T = -\sum_i E_T^i \hat{n}_i$, where i is the calorimeter tower number with $|\eta| < 3.6$, and $\hat{n}_i$ is a unit vector perpendicular to the beam axis pointing at the ith calorimeter tower. We also define $\not\!\!E_T = |\ \vec{\not\!\!E}_T|$. In the presence of muons or jets, the above calculation (raw $\not\!\!E_T$) overestimates the $\not\!\!E_T$ as muons do not deposit all their energy in the calorimeter and the raw jet energies are underestimated. We therefore correct for the jet energies and the muons by adjusting $\vec{\not\!\!E}_T$ by the corrected jet E_T values and the muon candidate p_T.

2.5 Monte Carlo Simulation

An important strategy in a particle physics analysis is the use of Monte Carlo (MC) simulation to define our expectations based on our understanding and assumptions of the nature of the particles and their interactions. Here, we briefly describe the MC simulation tools used in this analysis.

There are three main steps in a MC simulation program: event generation, showering/fragmentation and the detector simulation. The first step is to predict the nature of the interaction knowing the type and energy of the initial-state particles and the desired

final-state particles. For CDF, the initial-state particles are p and $\bar{p}$ each with an energy of 980 GeV. The event generator programs then predict the kinematics of the desired final-state particles using a set of random numbers and assumptions based on our understanding of the underlying physics along with probability distributions for the partons within the protons. At this stage the MC generator produces a list of initial and final state particles. These particles in turn have to be treated using different algorithms to reproduce the parton shower and fragmentation to form hadrons and create observable particles that in turn can interact with or decay within the detector. Different MC generator packages can be used alone or together with other programs to perform the first two steps.

The last step is a detailed simulation of the CDF II detector response. The CDF II detector simulation framework is based on GEANT3 [45]; however, other programs are also used for the simulation of specific components to either improve the simulation or decrease the processing time. The framework also includes decay packages to generate the decay products of particles that would decay within the detector. The CDF II detector simulation framework models the interaction of the particles with the detector and presents the simulated events in the same format as the data collected by the CDF II detector. Hence, the same reconstruction algorithms can be used for the samples produced by the MC simulation.

2.5.1 Monte Carlo Packages

In this section we introduce the specific MC simulation packages used to produce the MC samples used for this analysis. The detector simulation for all these MC samples is done using the CDF II detector simulation framework.

The $t\bar{t}$ MC Samples

The $t\bar{t}$ MC samples used for the studies discussed in the next chapter were made using either PYTHIA [46] or HERWIG [47] calculations for both the event generation and shower/fragmentation. The same PYTHIA sample is used to calculate the acceptance values for the $gg \to t\bar{t}$ and $q\bar{q} \to t\bar{t}$ production processes. In order to estimate the next-to-leading order effects on the average number of gluons in the $gg \to t\bar{t}$ process we use a $t\bar{t}$ MC sample produced by the MC@NLO [48] program.

The W + n Jet MC Samples

The MC samples used to calculate the average number of gluons for the W + n jet samples are generated with the ALPGEN [49] program and model the showering and fragmentation using PYTHIA algorithms. This sample is made by adding W+0, W+1, W+2, W+3 and W+4 parton MC samples weighted by their corresponding cross sections. It is possible, for example, to have a W+2 partons event coming from W+2 parton matrix element processes, or to have the W+2 parton event with exactly the same topology produced from a W+1 parton matrix element calculation with a second parton being produced through gluon radiation once the W+1 parton event is passed on to the showering program. Therefore, the MC samples are also treated to avoid such double-countings, using a process called "matching". The MLM [50] scheme is used for the matching of this sample.

To estimate the systematic uncertainties associated with this calculation, we use a different set of MC samples for these processes. The second sample is generated with the MADGRAPH [51] program and is treated for the showering and fragmentation using PYTHIA algorithms. A different matching process known as CKKW [52] scheme is used for this sample.

The Jet MC Samples

For the jet MC samples we have created $2 \rightarrow 2$ QCD processes using the PYTHIA program. Both the generation and showering/fragmentation are done by the same program. The second jet sample used for some of the systematic uncertainty estimates consists of the same $2 \rightarrow 2$ processes created using HERWIG calculations.

Chapter 3

Analysis

3.1 Discriminator

To make a measurement of $\sigma(gg \to t\bar{t})/\sigma(p\bar{p} \to t\bar{t})$, one needs to discriminate between the two $t\bar{t}$ production processes. Given their very similar final state, as discussed in section 1.6.2, the difference in these processes comes from the initial state partons. As described in section 1.3, a gluon is more likely to radiate additional gluons carrying a small fraction of its energy compared to a quark. Therefore, one expects to have a higher low p_T gluon radiation in the scattering processes involving gluons. As a result, one will have a larger number of low p_T particles and consequently, larger number of low p_T charged particles in processes involving more gluons.

Using $t\bar{t}$ PYTHIA MC calculations, we look at the distribution of the number of low p_T charged particles, N_{trk} , in events produced through gg fusion and $q\bar{q}$ annihilation. Figure 3.1 shows a comparison of the N_{trk} distributions between $t\bar{t}$ events produced through gg and $q\bar{q}$ processes. These are charged particles with p_T between 0.3 and 2.9 GeV/c and $|\eta| \leq 1$. The dashed red distribution comes from the $gg \to t\bar{t}$ process and the solid blue distribution corresponds to the $q\bar{q} \to t\bar{t}$ process.

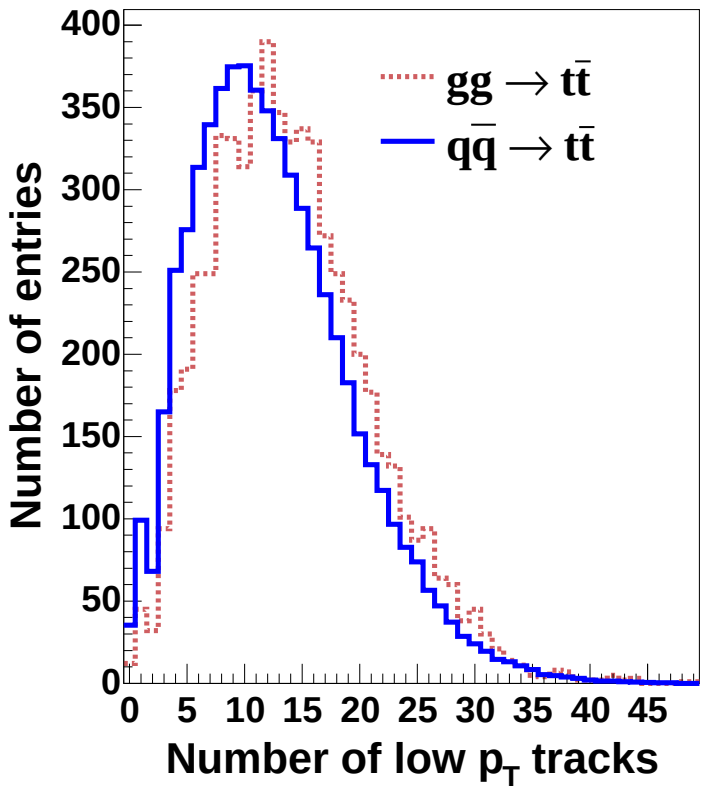

Figure 3.1: Comparison of the charged particle multiplicity distributions between $t\bar{t}$ events produced through gg fusion (red) and $q\bar{q}$ annihilation (blue) using PYTHIA MC calculations. The normalization is arbitrary, with each distribution having equal number of entries.

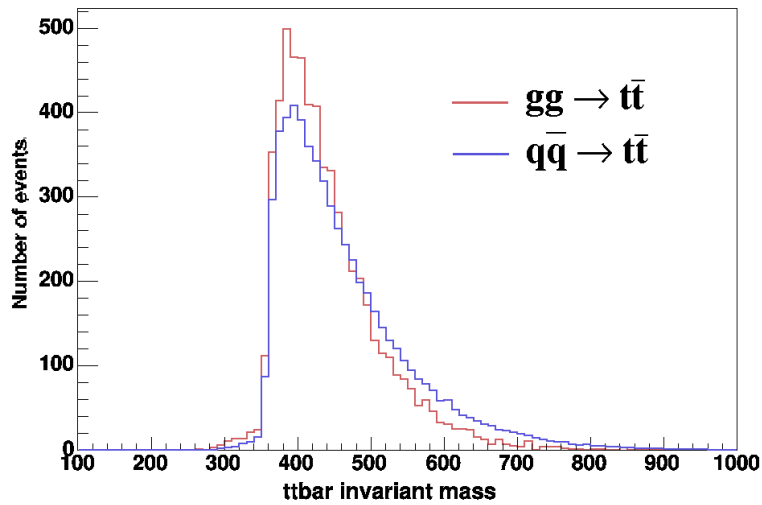

Figure 3.2: Comparison of the invariant mass of the $t\bar{t}$ system between $t\bar{t}$ events produced through gg fusion (red) and $q\bar{q}$ annihilation (blue) using PYTHIA MC calculations. The normalization is arbitrary, with each distribution having equal number of entries.

3.1.1 Alternative Discriminators

The N_{trk} parameter is not the only variable that differentiates the two $t\bar{t}$ production mechanisms. Any variable that is related to the initial state partons can potentially be used as a disciminant. The initial state partons mainly differ in their momentum distributions within protons. As can be seen in section 1.2.1, valance quarks are more likely to have enough energy to produce top quark pairs at the center of momentum energies available at the Tevatron. This difference can be seen in the invariant mass of the $t\bar{t}$ system as shown in Fig. 3.2. The dashed red distribution corresponds to the $gg \rightarrow t\bar{t}$ process and the solid blue distribution comes from the $q\bar{q} \rightarrow t\bar{t}$ process. The $q\bar{q} \rightarrow t\bar{t}$ distribution has more entries at higher mass values as expected.

In addition, the larger probability for gluon radiation from gluons provides higher transverse momentum for the $t\bar{t}$ events produced through the gg fusion process compared to those coming from $q\bar{q}$ annihilation. This is shown in Fig. 3.3 where the dashed red distribution comes from the $gg \rightarrow t\bar{t}$ process and the solid blue from $q\bar{q} \rightarrow t\bar{t}$ process.

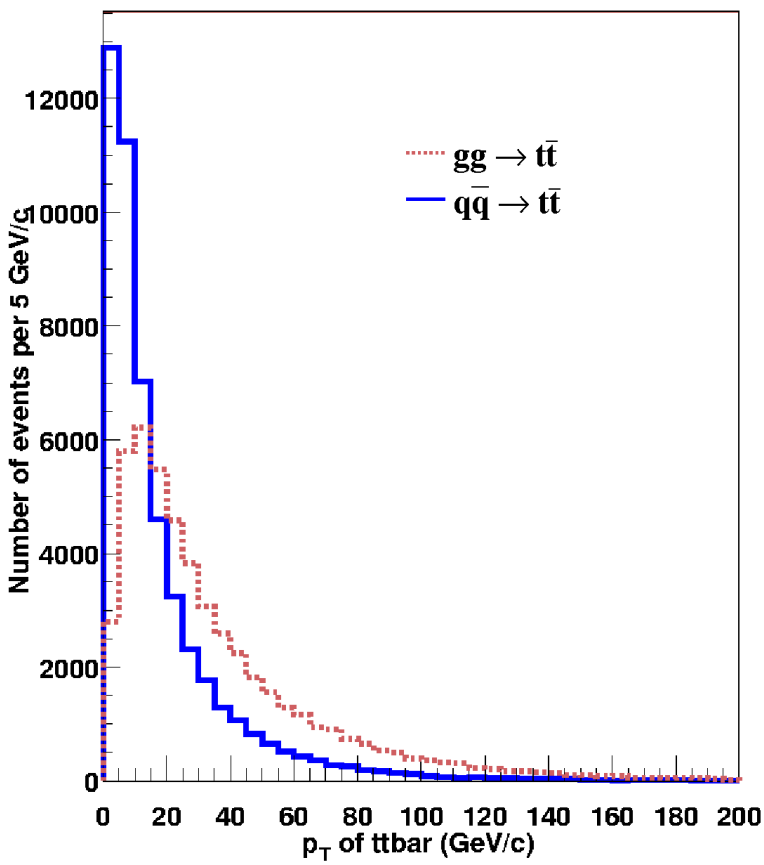

Figure 3.3: Comparison of the transverse momentum of the $t\bar{t}$ system between $t\bar{t}$ events produced through gg fusion (red) and $q\bar{q}$ annihilation (blue) using HERWIG MC calculations. The normalization is arbitrary, with each distribution having equal number of entries.

The two production processes also differ in their production kinematics, as discussed in section 1.2.1. These differences affect the angular distribution of the $t\bar{t}$ system. Figure 3.4 shows that the $gg \rightarrow t\bar{t}$ processes are more central compared to the $q\bar{q} \rightarrow t\bar{t}$ processes.

There are other variables that provide some discrimination. However, for our analysis, the N_{trk} is the most promising. All of the other variables suffer from significant construction inefficiencies, smearing and resolution effects, washing out the differences arising from the initial states. For example, the η distribution of the $t\bar{t}$ system, shown in Fig. 3.4, is limited by the reconstruction efficiency as the main difference lies in larger η

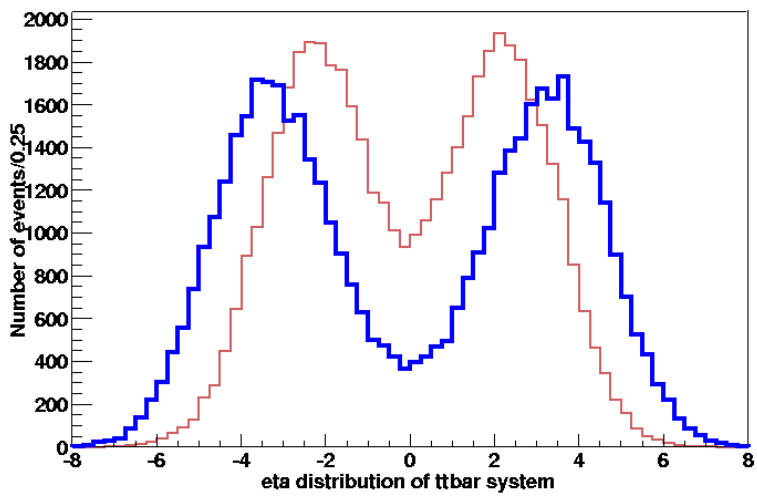

Figure 3.4: Comparison of the pseudorapidity of the $t\bar{t}$ system between $t\bar{t}$ events produced through gg fusion (red) and $q\bar{q}$ annihilation (blue) using HERWIG MC calculations. The normalization is arbitrary, with each distribution having equal number of entries.

where we reconstruct $t\bar{t}$ events that are produced more centrally or in case of the p_T of $t\bar{t}$ system, shown in Fig. 3.3, we may include a jet which is not part of the $t\bar{t}$ decay products, or may not have observed a jet which is. They therefore do not have comparable powers of discrimination, compared with the N_{trk} distribution, which does not suffer from these effects.

3.2 Samples and Selections

The N_{trk} variable, although a good discriminator, is not reliably modeled by MC calculations. In order to do so, one needs to model low p_T gluon radiation, one needs to model the effects of multiple interactions and underlying events in detail and one needs to model the effect of backgrounds. Given the large uncertainties associated with modeling of these effects, the use of MC calculations to define expectations for the signal and background processes were not used for this analysis. Instead, we take a data-driven approach and use data calibration samples with different gluon content to define the shape

of the relevent N_{trk} distributions. We do use MC calculations to model the gluon content of the different calibration processes and their signatures as we discuss in the following sections.

For the calibration of gluon-gluon processes, we use dijet samples with different leading jet E_T ranges. The lower the leading jet E_T, the higher the gluon content of the sample. We also use W+n jet, n=0, 1, 2 or 3, samples for calibrating the $q\bar{q}$ processes as well as processes involving gluons. The higher the number of jets, the higher the gluon content of the sample. Our signal sample consists of $t\bar{t}$ events where one of the W bosons decays hadronically and the other decays to a charged lepton and its corresponding neutrino. In the following, we will restrict our selections of leptons to either electron or muon candidates.

We have used the high quality data collected during the running period, corresponding to an integrated luminosity of 0.96±0.06 fb^{-1}. The selection requirements for each sample as well as the background estimates are described in the following subsections.

3.2.1 W+n Jet Sample

We consider those W+n jet events where the W boson decays to a charged lepton and its corresponding neutrino. The criteria explained here are applied to both data and MC calculations. For W+n jet MC samples, we use ALPGEN+PYTHIA calculations passed through the CDF II detector simulation, as discussed in section 2.5. To estimate the systematic uncertainties, we use MADGRAPH+PYTHIA calculations for W+n parton processes. Specifically, we require the following:

- One and only one tight electron or muon candidate as described in section 2.4.3.
- A minimum $\not\!\!E_T$ of 20 GeV, as indicative of the presence of a neutrino. The $\not\!\!E_T$ is corrected for the jet E_T corrections and is described in section 2.4.6.
- We remove any event that satisfies the requirements for a $t\bar{t}$ dilepton or a Z boson

candidate [14].

- We veto any event that is consistent with being from a cosmic ray or where the electron is consistent with coming from a photon conversion, as discussed in section 2.4.3.

- At least one primary vertex.

- The primary vertex of the event to be within ± 60 cm of z=0 and that the track associated with the tight lepton candidate originates within 5 cm of the primary vertex on the z axis. The primary vertex finding algorithm is described in section 2.4.2.

Jets are defined using the JetClu algorithm with a cone of size R =0.4 (for details refer to section 2.4.4) and are used to categorize the W events as W+0, 1, 2 and 3 jet events. Each jet is required to have a minimum corrected E_T of 15 GeV and $|\eta| \leq 2$. The jets are corrected for L4 corrections, as discussed in section 2.4.4.

3.2.2 Dijet Sample

The dijet calibration sample is defined by a selection that identifies the 2→2 scattering processes discussed in section 1.5. The higher the leading jet E_T, the lower the gluon content is expected to be. The criteria explained here is applied to both data and MC calculations. For the dijet MC samples, we use two sets of PYTHIA calculations passed through the CDF II detector simulation, one with a minimum parton p_T of 40 GeV/c (Jet40 MC), and the other one with a minimum p_T of 90 GeV/c (Jet90 MC). We impose the following additional criteria:

- To avoid any trigger bias, we require a minimum uncorrected leading jet with E_T of 75 and 130 GeV for Jet50 data (Jet40 MC) and Jet100 data (Jet90 MC), respectively.

- We remove any event that has any tight electron or muon candidate.
- We require 2 and only 2 jets within $|\eta| \leq 2$ with a minimum corrected E_T of $\geq$15 GeV.
- The two jets should be back-to-back in ϕ within 35°.
- We require at least one good primary vertex in the event.

Jets are defined using the JetClu algorithm with a cone size of R =0.4 and are corrected for L5 corrections. These events are separated into subsamples according to their leading jet E_T ranges in steps of 20 GeV, starting with a leading jet E_T of 80 GeV. We call the subsamples dijet 80-100 GeV, dijet 100-120 GeV, dijet 120-140 GeV, dijet 140-160 GeV and dijet 160-180 GeV. A final dijet subsample consisting of all events with E_T $\geq$180 GeV is also defined.

3.2.3 Multijet Sample

As discussed later in section 3.4, we make a correction based on the number of high E_T jets to find the correct number of low p_T tracks in a given event. To find the value for this correction, we remove the two and only two back-to-back jet requirements from dijet sample, defining a multijet sample. The correction is the only instance where we use the multijet sample, and as such, we do not use any MC calculations for multijet processes.

3.2.4 $t\bar{t}$ Signal Sample

For our signal sample, we use $t\bar{t}$ events where one of the W bosons decays to an electron or a muon and the corresponding neutrino and the other W boson decays to two quarks. Therefore, one expects to have a lepton candidate, large missing E_T and at least four jets in such events, two originating from b quarks. As such, the criteria for our signal sample is exactly as the W+n jets except that we require at least four jets. This criteria,

however, will introduce a large background coming from W+n jet processes. To reduce this background, we also require one of the four jets to be tagged as a b quark (b-tagged) as explained in section 2.4.5. For our signal $t\bar{t}$ modeling we use HERWIG and PYTHIA MC calculations passed through the CDF II detector simulation. The PYTHIA sample is used to estimate the $gg \to t\bar{t}$ and $q\bar{q} \to t\bar{t}$ acceptances.

3.3 Background Processes

It is important to know the processes that have similar signatures to that of our signal sample and our calibration samples and therefore pass our criteria. As shown later in section 4.1, we also need to know the backgrounds in our calibration samples so that we can estimate the average number of gluons involved in their production processes. In case of the signal sample, we also need to estimate the background rates to be able to correct the measurement for the effects introduced by the backgrounds, as discussed in section 4.2.

3.3.1 W+0 Jet Sample

For the purpose of this analysis, the background processes in the W+0 jet sample is divided into two groups. The first group consists of processes that have a similar production mechanism. The second group consists of processes that have different production mechanisms. Given the similar production mechanism, we do not need to correct for the first group of processes as they will have similar average number of gluons involved in their production. The second group of the processes are those that one needs to correct for their contribution as explained in section 4.1.1. These two background groups are listed in Table 3.1.

We use ALPGEN+PYTHIA MC calculations to find the average number of gluons as explained in section 3.5.1 in the W+n jet background. As we use the average number

Category	Background Processes	The Reason for its Acceptance
Similar Mechanism	$W \to \tau\nu_{tau}$	τ decays to e or μ
	$Z \to e^+e^-/\mu^+\mu-$	One of the leptons escapes detection
Different Mechanism	W+n jet, n>0	One or more jets escape detection
	QCD	A jet mimics an electron and at least one misses detection

Table 3.1: The background processes for the W+0 jet sample separated in the two groups of similar and different production mechanisms.

of gluons, we do not need to worry about the exact number of background events but just its fraction. We estimate that this background contributes 0.05 ± 0.10 gluons to the average number of gluons in W+0 jet sample. The uncertainties are systematic and are estimated from the difference between a similar calculation using MADGRAPH+PYTHIA MC calculations. The fraction of QCD background in the W+0 jet sample is estimated using data. To do so, we remove the $\not\!\!E_T$ and the lepton isolation cuts and define four different subsamples, A, B, C and D based on their isolation and $\not\!\!E_T$ values. The isolation vs. the $\not\!\!E_T$ distribution is shown in Fig. 3.3.1 for W decays to either electrons or muons and neutrinos. The areas occupied by the four are also shown on the plot. Assuming that the lepton isolation is uncorrelated with the $\not\!\!E_T$, we expect the number of QCD background events in the signal area D, N_D^{bkg}, to be related to the number of events observed in the other three areas, N_A, N_B and N_C as follows:

$$\frac{N_A}{N_B} = \frac{N_D^{bkg}}{N_C} \tag{3.1}$$

and therefore, the fraction of QCD background in the W+0 jet sample can be found as:

$$\frac{N_D^{bkg}}{N_D} = \frac{N_A}{N_B}\frac{N_C}{N_D}. \tag{3.2}$$

As the systematic uncertainty we take half of the difference between the background

fractions we find if we use W decays only to electrons or only to muons, which is a conservative estimate given that it is perhaps the most sensitive test of the assumption that the lepton isolation is uncorrelated with the $\not\!\!E_T$. The estimated QCD background for the W+0 jet sample is $(4.9 \pm 0.4)\%$. The statistical uncertainties are negligible compared to the systematic uncertainty estimated above.

3.3.2 $t\bar{t}$ Sample

In order to find the gluon-rich fraction of events in the $t\bar{t}$ events, we need to know the fraction of background events as well as the fraction of background events involving heavy-flavour (HF) or light-flavour (LF) quarks. We use the background calculations developed for the $t\bar{t}$ production cross section measurement using SecVtx b-tag lepton+jet $t\bar{t}$ events [14] with ~ 700 pb^{-1}. We do not need the number of background events but rather the fraction. We add a 10% normalization uncertainty to account for possible differences that may arise from differences in instantaneous luminosity between the 700 pb^{-1} sample and our 960 pb^{-1} sample. The sources of background and their fractional contribution to the $t\bar{t}$ sample are summarized in Table 3.2. The processes are categorized into HF, LF and non-W sources. The non-W background consists of processes involving either LF or HF final states that are not included in the LF and HF categories.

3.4 The Observable Used to Discriminate

As discussed in section 3.1, we take advantage of the charged particle multiplicity to discriminate between samples that contain large number of gluons in their production process and those with no or few gluons. As we are interested in the production process, we do not wish to be sensitive to the final state particles, underlying activity of partons that are not part of the initial production process, or any extra proton-antiproton interactions. The track selection requirements to achieve these goals are as follows.

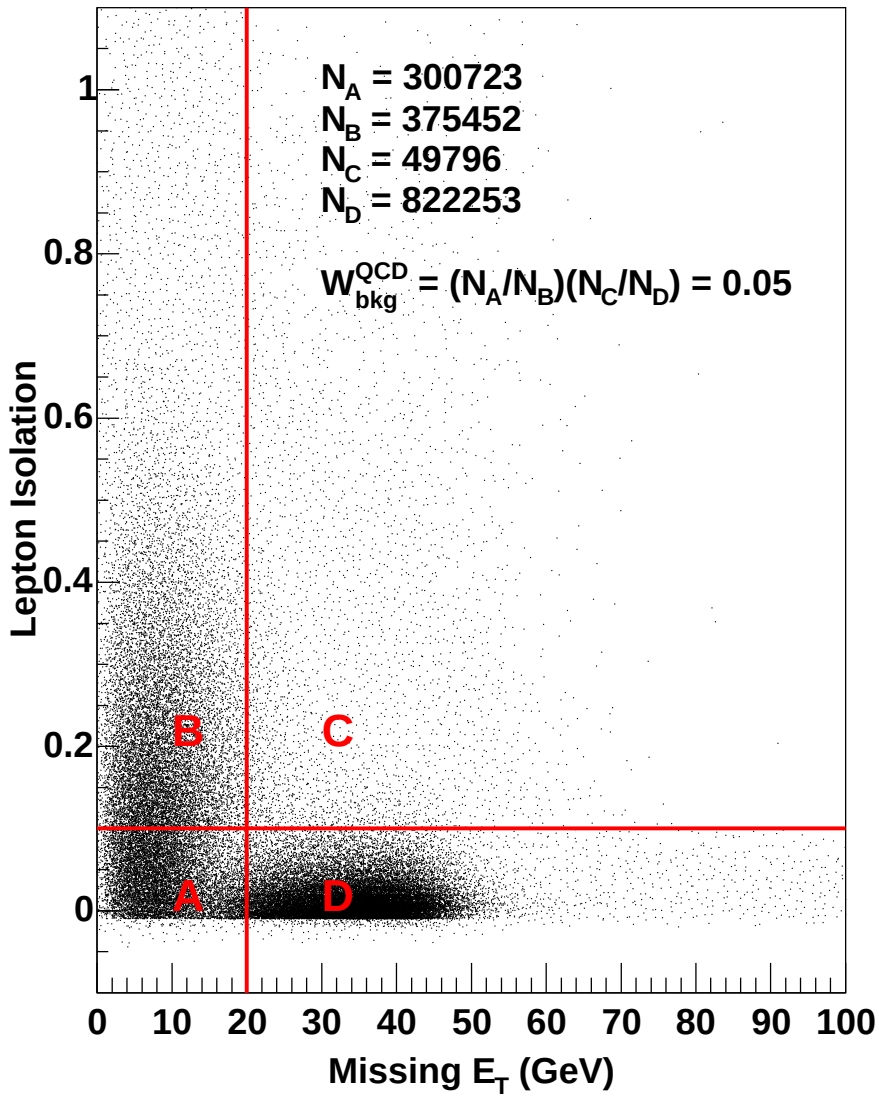

Figure 3.5: The distribution of lepton isolation vs. $\not\!\!E_T$ for both electrons and muons used to estimate the QCD background in the W+0 jet sample. Number of events found in each region and the QCD background fraction found are also shown on the plot.

Category	Process	Contribution (%)
LF	Mistag/W+LF	5.1±0.8
	Diboson	0.7±0.1
HF	$Wb\bar{b}$	2.7±0.8
	$Wc\bar{c}$	1.3±0.4
	Wc	0.2±0.1
	Single Top	0.3±0.1
non-W	QCD	2.8±0.5
Total		13.1±1.8

Table 3.2: The background processes for the $t\bar{t}$ sample separated in the three categories of LF, HF and non-W backgrounds. The non-W background has contributions from both LF and HF processes.

- We require at least 16 stereo and 20 axial COT hits to define a track.

- We use tracks with p_T in the range 0.3-2.9 GeV/c and $|\eta| \leq 1.1$, where we expect to have high tracking efficiency.

- To eliminate charged particles arising from final state jets, the tracks should not be part of the jets present in the event. We therefore require the tracks not to fall within $\Delta R = 0.6$ of the high E_T jets (15 GeV or more) and within $\Delta R = 0.4$ of the low E_T (6-15 GeV) jets in the event. The smaller cone size required for the low E_T jets is set as such due to the fact that the lower E_T jets may come from ISR/FSR[1] . Since we expect a higher number of low E_T ISR jets from $gg \rightarrow t\bar{t}$ events, we choose a smaller rejection cone for low E_T jets as a compromise between rejecting the common final state particles and taking into account the ISR particles.

[1]The ISR and FSR stand for initial state radiation and final state radiation, respectively. They refer to the gluons radiated from the partons in the initial or final state of the production process.

- The track should match the primary vertex of the event within ± 3 cm. This requirement reduces the contribution from other interactions as well as the tracks arising from multiple scattering of charged particles within the detector.

The fact that we exclude regions around the jets, depending on the number of jets in the tracking region in the event, provides different tracking area available for different events. Figure 3.6 shows an example schematics of the unrolled central calorimeter with a low E_T and a high E_T jet in the tracking region ($|\eta| \leq 1.1$). The area excluded from the tracking due to the presence of these jets is shown by the circles. To have a comparable track multiplicity, we find the track density for each event by dividing the observed track multiplicity by the tracking area available for tracking in the absence of jets. Then, we multiply this density with the total central area, 4.4π, to get a normalized track multiplicity.

Even though tracks from jets are excluded, the track multiplicity still has a dependency on the number of high E_T jets in the event as expected given the infrequent occurrence of high-angle tracks generated in the showering process. We therefore have further contributions from each high E_T jet present in the central ($|\eta| \leq 1.1$) region. We use the slope of the mean track multiplicity versus number of jets in the central region in the multijet data events as a correction for this contribution. For different periods of data-taking, the slopes are 0.90 ± 0.03, 0.97 ± 0.04 and0.96 ± 0.04. For this estimate, we require one and only one good primary vertex in multijet data events to be insensitive to the effects of multiple interactions. The correction differs based on the data-taking conditions and on average it is about 1 track per high E_T jet in the central region.

3.5 Correlation between $<N_{trk}>$ and $<N_g>$

As discussed in section 1.3, we expect larger number of charged particles to be produced in processes involving more gluons. To check this expectation, we look at the average

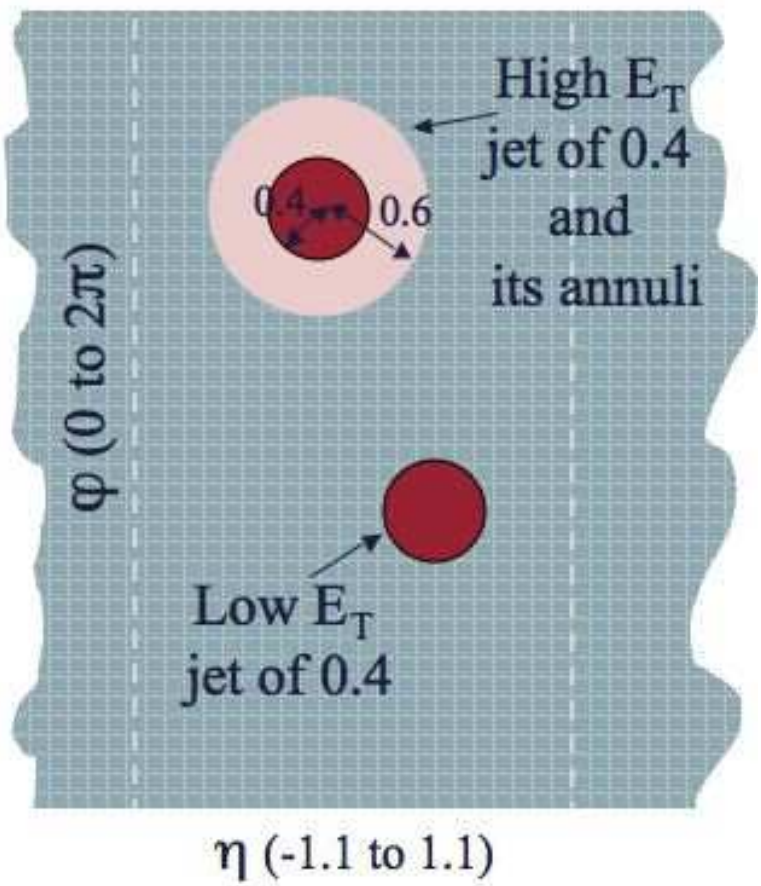

Figure 3.6: A schematics of the unrolled central calorimeter with a low E_T and a high E_T jet in the tracking region ($|\eta| \leq 1.1$). The area excluded from the tracking due to the presence of these jets is shown by the circles. Please note that this is not to scale.

number of low p_T tracks, $<N_{trk}>$, observed in our calibration data samples as a function of average number of gluons, $<N_g>$, involved in the production of these processes using MC calculations.

We define below how $<N_g>$ is estimated for each sample.

3.5.1 The $<N_g>$ Estimate

We apply the same event selection cuts to data and to MC samples. Using the generator-level information, we count the number of gluons in each event, taking into consideration the 2 incoming and all the outgoing partons. We define the outgoing partons as the immediate daughters of the 2 incoming partons. For all dijet samples, we have 2 incoming and 2 outgoing partons. In the W samples, depending on the type of generated event, we have 2 incoming and 0, 1, 2, 3 or 4 (excluding the W boson) outgoing partons corresponding to the W+0, 1, 2, 3 or 4 parton samples used to create the W+n jet MC sample. To get $<N_g>$ in a sample, we sum over the number of gluons in each event of our MC samples and divide the sum by the total number of events in the sample. Comparing the MC prediction for each sample from different MC calculations, an uncertainty of 0.1 is assigned to $< N_g >$ prediction for all datasets. In case of the dijet samples, this also includes the difference expected from leading order and the next-to-leading order calculations [53].

3.5.2 The Observed $<N_{trk}>$

We determine the track multiplicity in each event as described in the previous section. For each calibration sample we create the track multiplicity distribution. Figures 3.7, 3.8, 3.9 and 3.10 show the track multiplicity distributions for dijet 100-120 GeV, dijet 160-180 GeV, W+1 jet and W+2 jet samples, respectively. To characterize each distribution, we determine the arithmetic mean, $<N_{trk}>$.

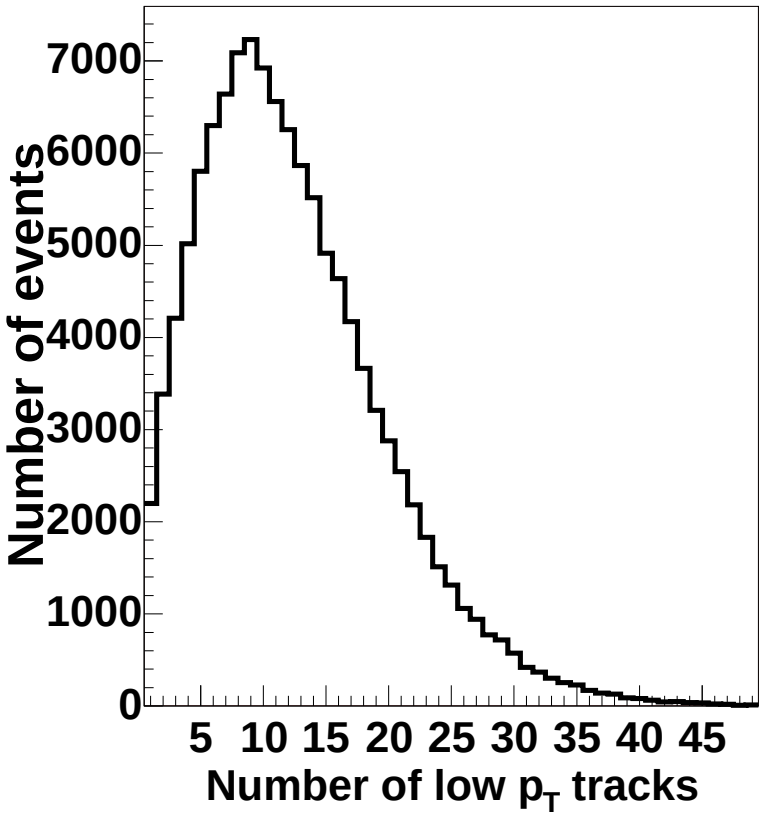

Figure 3.7: The low p_T track multiplicity distribution in dijet 100-120 GeV data sample.

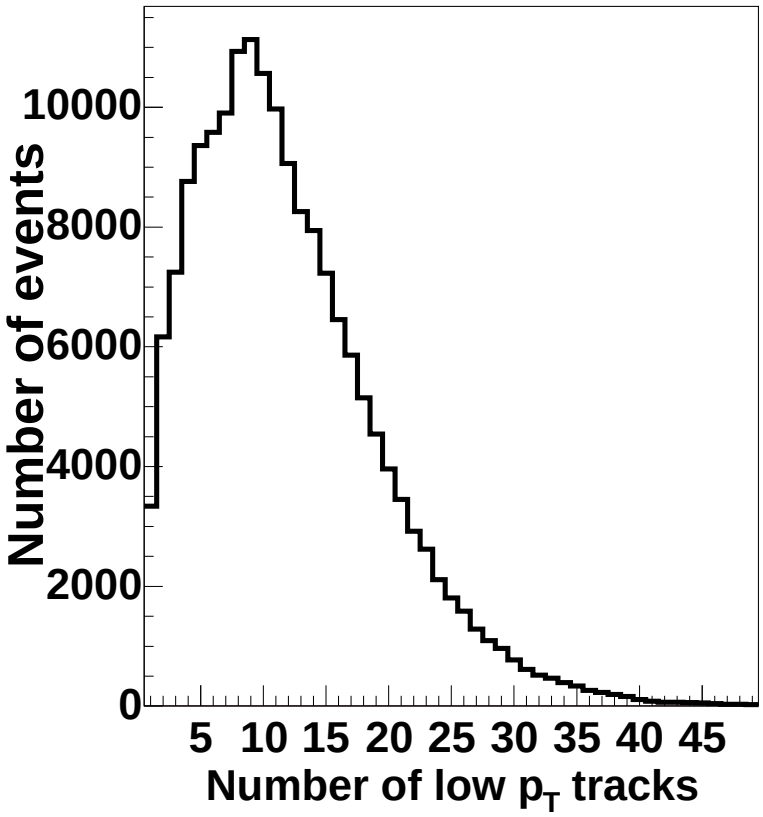

Figure 3.8: The low p_T track multiplicity distribution in dijet 160-180 GeV data sample.

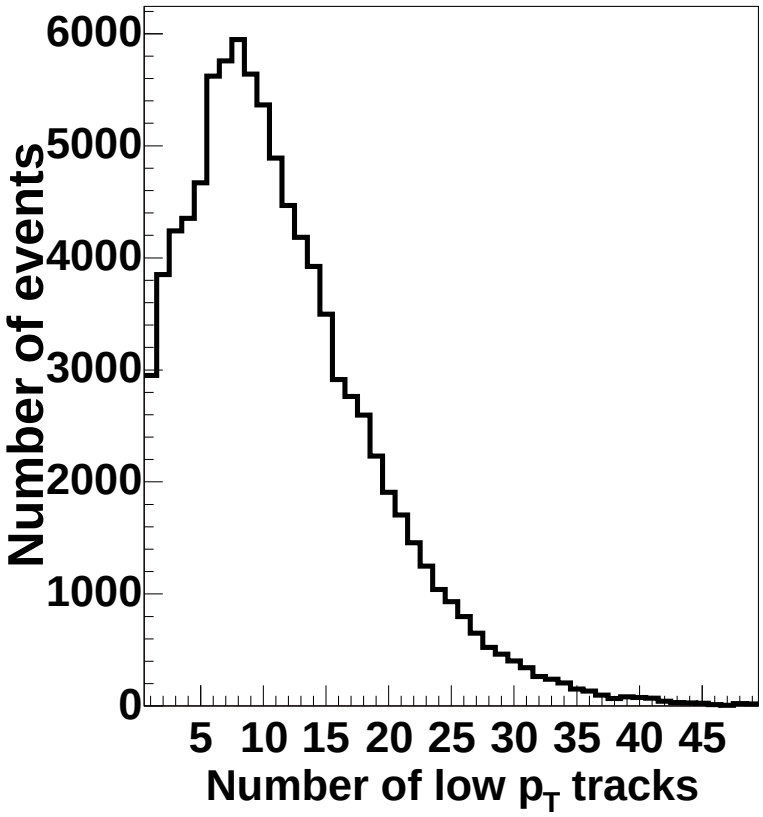

Figure 3.9: The low p_T track multiplicity distribution in W+1 jet data sample.

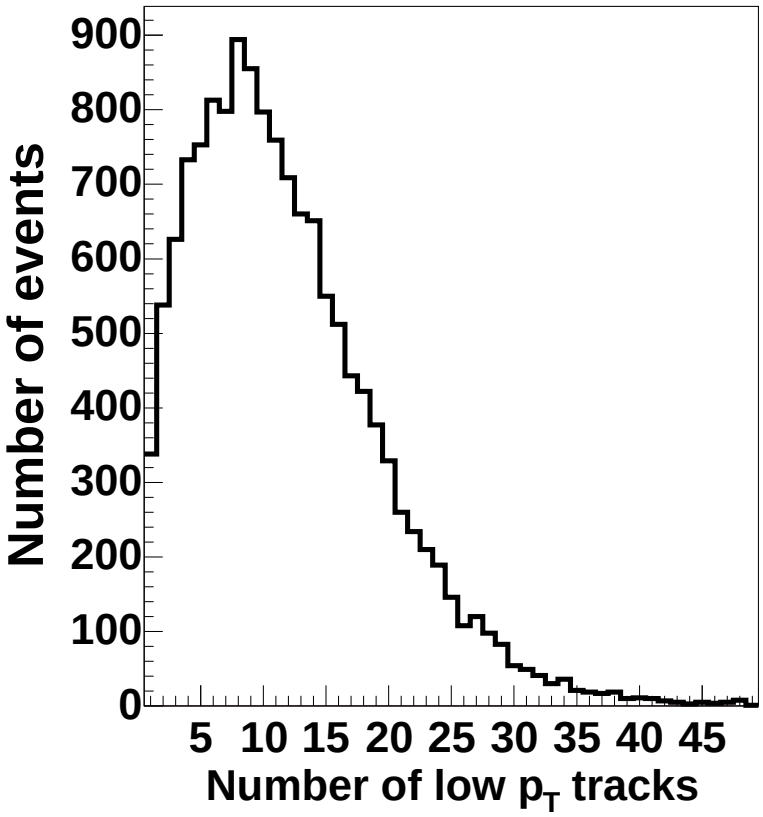

Figure 3.10: The low p_T track multiplicity distribution in W+2 jet data sample.

3.5.3 $<N_{trk}>$ vs. $<N_g>$

We use six different calibration samples to show that there is a correlation between $<N_{trk}>$ and $<N_g>$. Figure 3.11 shows the $<N_{trk}>$ vs. $<N_g>$ relationship in the W+0, 1 and 2 jet samples along with the dijet samples with leading jet E_T range 80-100, 100-120 and 120-140 GeV. A linear χ^2 fit to these data is also shown, quantizing what is a clear correlation. This linear fit can then be used, given a measured $<N_{trk}>$, to predict $<N_g>$ in samples not used in Fig. 3.11. The comparisons of the predicted and the observed $< N_g >$ based on MC calculations are shown in Table 3.3.

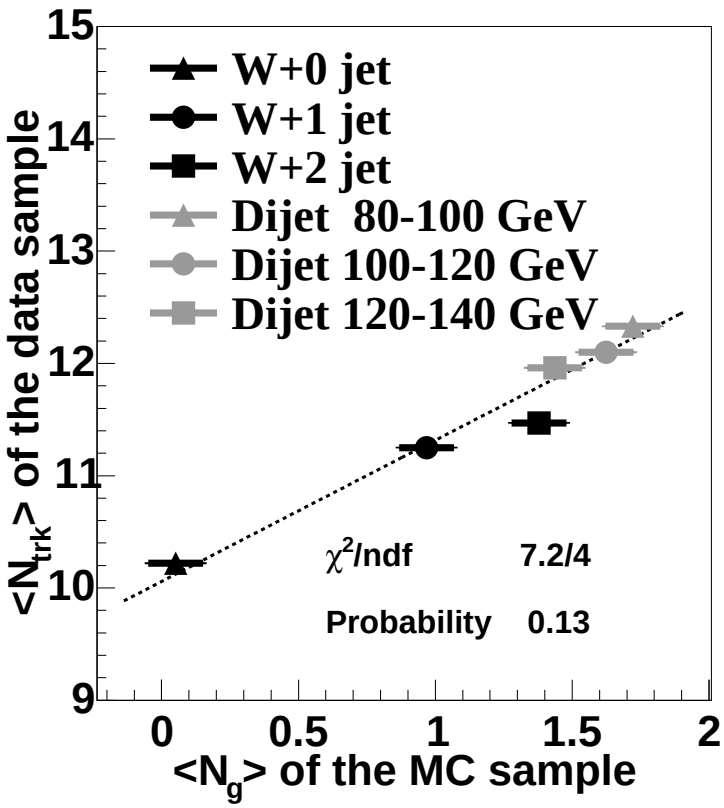

Figure 3.11: The correlation between the average low p_T track multiplicity (data) and the average number of gluons (MC). The dotted line is from a linear fit to the points.

The good agreement between the expected $<N_g>$ determined from the $< N_{trk} >$

Sample	MC Expectation	from Fit to Data
dijet 140-160 GeV	1.26 ± 0.04	$1.39^{+0.06}_{-0.05}$
dijet 160-180 GeV	1.13 ± 0.04	1.23 ± 0.05
dijet 180-200 GeV	0.99 ± 0.07	$1.08^{+0.05}_{-0.06}$
dijet 200-220 GeV	0.92 ± 0.10	$0.88^{+0.04}_{-0.07}$
dijet 220+ GeV	0.67 ± 0.10	$0.65^{+0.05}_{-0.10}$

Table 3.3: The average number of gluons in each sample as predicted by MC calculations and the average number of gluons as found using the correlation fit for data. All uncertainties are statistical.

correlation and the MC prediction confirm that the $<N_{trk}>$ and the $<N_g>$ in a given sample are correlated.

3.5.4 Is There a Q^2 Effect?

The processes used to show the above correlation also differ in their Q^2. This raises the question that the increase in the $<N_{trk}>$ as $<N_g>$ increases may also be affected by the difference in Q^2 of the processes. As later on we want to use the track multiplicity distributions from our calibration samples to model the distributions for the $t\bar{t}$ processes, we need to show the difference in Q^2 will not affect the distributions. To do so, we take advantage of the fact that the $<N_g>$ in the W+1 jet sample does not change as the jet energy (and therefore the Q^2) in the production process increases. Figure 3.12 shows the $<N_{trk}>$ vs. the E_T range of the jet in W+1 jet data sample, confirming that the $<N_{trk}>$ is largely independent of the jet E_T and the Q^2 of the process[2]. Therefore, we can use track multiplicity distributions with specific $<N_g>$ to model the track multiplicity distributions of any other sample with similar $<N_g>$.

[2]The drop in $<N_{trk}>$ over the full range is less than 10% of the overall change in $<N_{trk}>$ with $<N_g>$ ranging from 0 to 2.

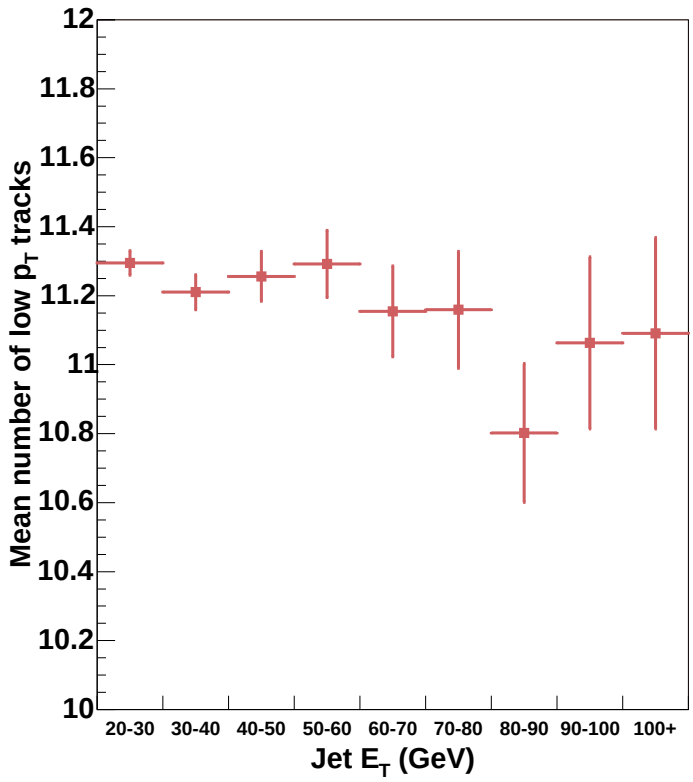

Figure 3.12: The $<N_{trk}>$ in the W+1 jet data sample as a function of the energy of the leading jet.

Chapter 4

Measurement

As shown in section 3.5, there is a correlation between $<N_{trk}>$ and $<N_g>$. This means that we can associate a specific $<N_g>$ to a given N_{trk} distribution. We take advantage of this correlation to define two different N_{trk} distributions, one associated with initial states with no gluon content and the other with a gluon content comparable to that of the $gg \rightarrow t\bar{t}X$ process, what we will call a "gluon-rich" process. Using these distributions in a likelihood fit to the observed N_{trk} distribution, as described later, one can find the fraction of gluon-rich events in any given data sample.

4.1 Distribution Fits

4.1.1 No-gluon and Gluon-rich Parameterizations

In order to obtain the gluon-rich distribution, we start with the dijet sample with leading jet E_T in the range 80-100 GeV. This sample consists almost entirely of real dijet final states. It has a $(27\pm3)\%$ contribution from $qq \rightarrow qq$ processes as predicted using PYTHIA MC calculations, with the rest arising from qg and gg processes as shown in Fig. 1.7. As there is no gluon contribution to this scattering subprocess, we assume that we can use the W+0 jet N_{trk} distribution to model the N_{trk} distribution for the $qq \rightarrow qq$ process,

as the W+0 jet sample is mainly a $q\bar{q}'$ process. To subtract the $qq \rightarrow qq$ contribution to the N_{trk} distribution of the dijet sample with E_T of 80-100 GeV, we normalize the W+0 jet N_{trk} distribution to that of the dijet 80-100 GeV process multiplied by the 0.27 factor and subtract the normalized W+0 jet N_{trk} distribution from the dijet N_{trk} distribution to get the gluon-rich N_{trk} distribution, where there are at least 2 gluons involved in the scattering process of the dijet samples. This sample has an $<N_g>$ of about 2.4, estimated using PYTHIA MC calculations.

The W+0 jet sample has two types of background processes as discussed in section 3.3.1. The background processes with similar production mechanism to the W+0 jet process do not contain any gluons in their production process and as such are of no concern in defining the no-gluon N_{trk} distribution. However, one needs to subtract the gluon contributions to the N_{trk} distribution of the W+0 jet sample from the background processes with different production mechanisms, namely the QCD and W+n(>0) jet processes. The QCD background is estimated to be $(4.9 \pm 0.4)\%$ of the W+0 jet sample. Assuming an $<N_g>$ of 2 for the QCD processes, this background contributes an $<N_g>$ of 0.10 ± 0.01 to the W+0 jet sample. The W+n(>0) jet processes are estimated to contribute an $<N_g>$ of 0.05 ± 0.10 to the W+0 jet sample. So, in total the $<N_g>$ contribution from the backgrounds to the W+0 jet sample is 0.15 ± 0.10. Given that the gluon-rich N_{trk} distribution has an $<N_g>$ of about 2.4, we estimate the W+0 jet sample has a $(6 \pm 4)\%$ contribution from the gluon-rich N_{trk} distribution. We therefore normalize the gluon-rich N_{trk} distribution to the W+0 jet sample N_{trk} distribution and subtract the result with a 0.06 factor from the W+0 jet sample N_{trk} distribution to get the no-gluon distribution.

We iterate on the above procedure to get our final gluon-rich and no-gluon N_{trk} distributions. The distributions converge after the first iteration. We model, for the convenience of parametrization, the no-gluon N_{trk} distribution with 2 Landau and 3 Gaussian distributions and the gluon-rich distribution with 2 Landau and 5 Gaussian distributions, the

components selected to obtain an accurate parameterization. The no-gluon and gluon-rich parameterizations are compared to the distributions and each other in Fig. 4.1.

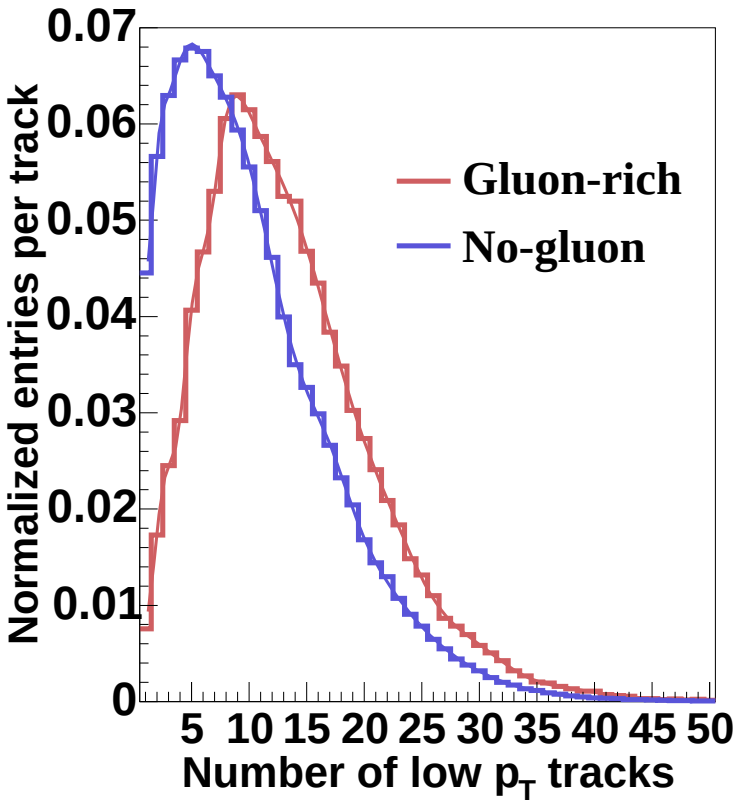

Figure 4.1: Comparison between the gluon-rich and no-gluon distributions and parameterizations. The vertical scale is normalized such that the distributions have unit area.

4.1.2 Finding the Gluon-rich Fraction

Using the no-gluon and gluon-rich parameterizations in a simple binned likelihood fit with two free parameters to the N_{trk} distribution of any given sample, we find the fraction of the gluon-rich events in the sample. The likelihood fit has the form

$$N[f_g\mathcal{F}_g(N_{trk}) + (1 - f_g)\mathcal{F}_q(N_{trk})], \tag{4.1}$$

where N is a normalization factor and f_g is the fraction of gluon-rich events. The functions $\mathcal{F}_g$ and $\mathcal{F}_q$ are the normalized parameterizations of the gluon-rich and no-gluon N_{trk} distributions, respectively.

It is worth mentioning that f_g will be the same as the fraction of gluon-rich events only if the $<N_g>$ in the gluon-rich events of the sample is approximately similar to the $<N_g>$ of our gluon-rich N_{trk} distribution, 2.4. This is the case for the $gg \rightarrow t\bar{t}$ process, where we expect to have 2 gluons in the initial state and some contributions from higher order processes, increasing the $<N_g>$.

As the other dijet samples have similar $<N_g>$, we measure f_g in these samples and compare them to the MC calculations. Figures 4.2 to 4.7 show the fits to the calibration samples. The fit results and the MC calculations are compared in Table 4.1. The good agreement between the fit result and MC calculations gives us confidence that we can use the same method to find the gluon-rich fraction in the $t\bar{t}$ sample.

Jet E_T Range (GeV)	MC Expectation	f_g from Fit to Data
80-100	0.73 ± 0.03	0.73 ± 0.01
100-120	0.69 ± 0.03	0.69 ± 0.01
120-140	0.63 ± 0.04	0.66 ± 0.01
140-160	0.57 ± 0.04	0.63 ± 0.01
160-180	0.52 ± 0.04	0.57 ± 0.01
$\geq$180	0.42 ± 0.05	0.49 ± 0.01

Table 4.1: The fraction of gluon-rich events in each sample as predicted by MC calculations and the fraction of gluon-rich events as found using the likelihood fit to track multiplicity distributions. Uncertainties for the MC fractions include both statistical and systematical contributions. The uncertainties on the fit results to the data are only statistical.

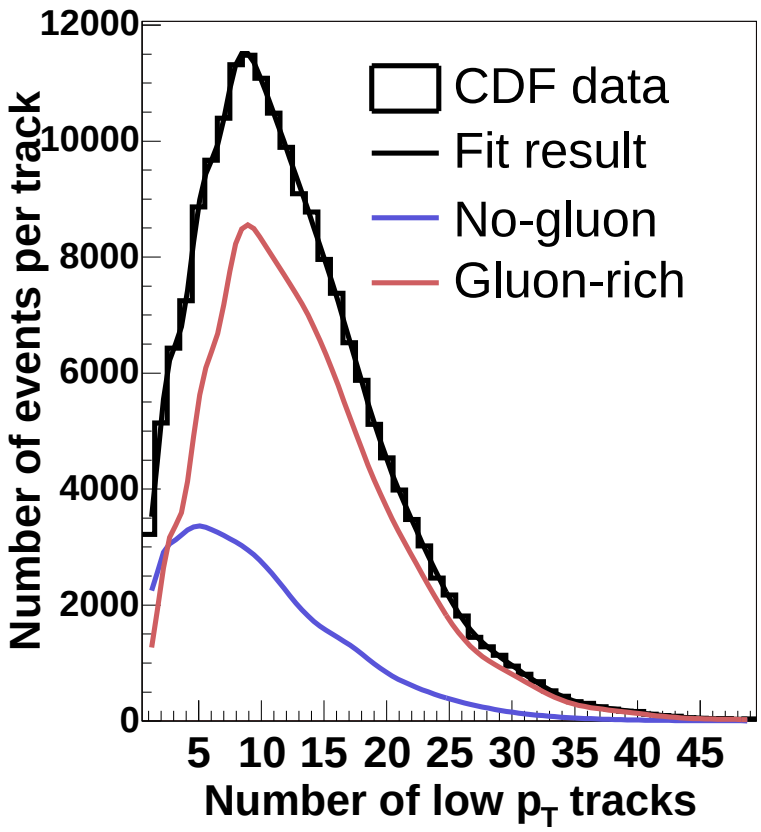

Figure 4.2: The fit result for the dijet sample with leading jet E_T of 80-100 GeV.

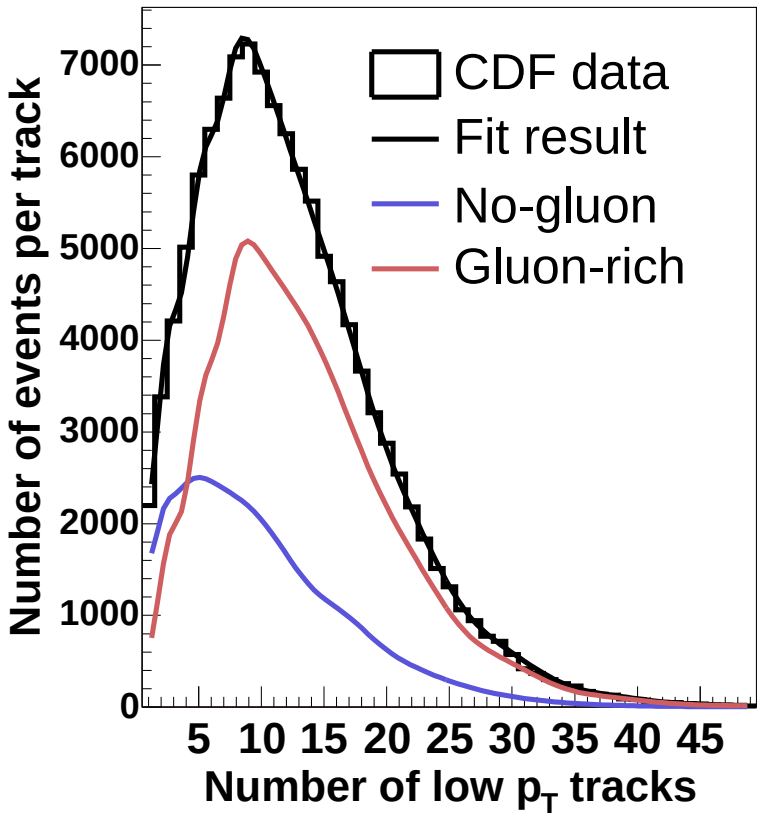

Figure 4.3: The fit result for the dijet sample with leading jet E_T of 100-120 GeV. The two components of the fit (gluon-rich and no-gluon) are also shown.

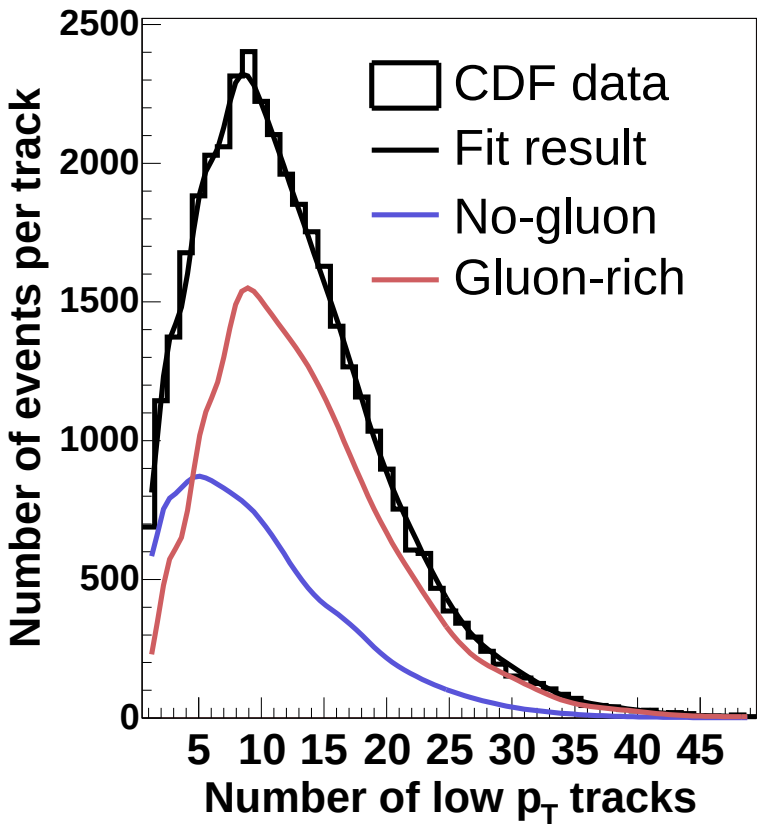

Figure 4.4: The fit result for the dijet sample with leading jet E_T of 120-140 GeV. The two components of the fit (gluon-rich and no-gluon) are also shown.

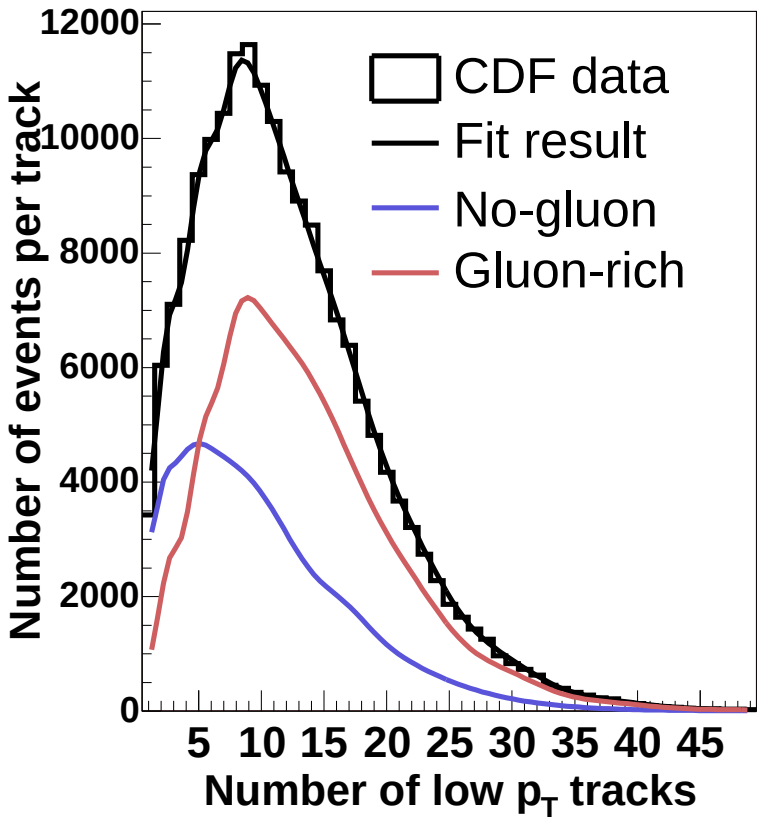

Figure 4.5: The fit result for the dijet sample with leading jet E_T of 140-160 GeV. The two components of the fit (gluon-rich and no-gluon) are also shown.

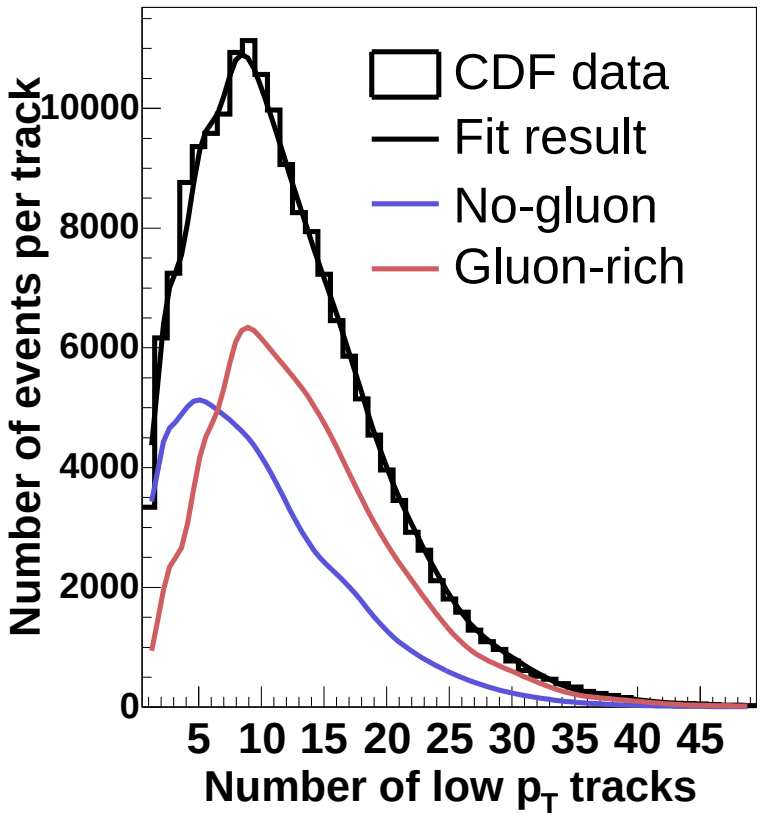

Figure 4.6: The fit result for the dijet sample with leading jet E_T of 160-180 GeV. The two components of the fit (gluon-rich and no-gluon) are also shown.

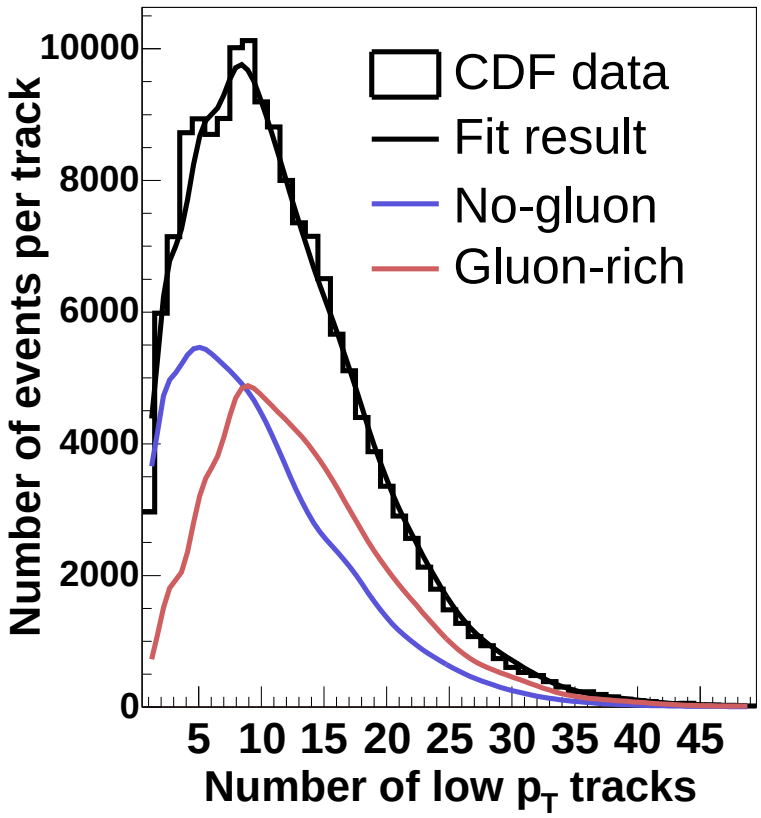

Figure 4.7: The fit result for the dijet sample with leading jet E_T of at least 180 GeV. The two components of the fit (gluon-rich and no-gluon) are also shown.

4.2 Gluon-rich Fraction of $t\bar{t}$ Sample

To measure the gluon-rich fraction in the $t\bar{t}$ candidate sample, we fit the $W+\geq 4$ jet N_{trk} distribution, as shown in Fig. 4.8. The f_g found in this sample reflects two components, the $t\bar{t}$ gluon-rich fraction and the gluon-rich fraction of the background events. Therefore, knowing the background fraction in our sample, f_b, and the measured f_g from the fit, we can write

$$f_g = f_b f_g^{bkg} + (1 - f_b) f_g^{t\bar{t}}, \tag{4.2}$$

where, f_g^{bkg} and $f_g^{t\bar{t}}$ are the gluon-rich fraction of the background and $t\bar{t}$ signal, respectively. The latter is what we want to measure, while f_b can be estimated for the $W+\geq 4$ jets sample as done in $t\bar{t}$ cross section measurements and briefly explained in section 3.3.

In order to find the fraction of gluon-rich components in the background, we measure f_g in $W+1$, $W+2$ and $W+3$ jet data samples with no positive SecVtx b-tag and with at least one tight SecVtx b-tag using a fit to the N_{trk} distribution for each sample. We then extrapolate the f_g values from the $W+1$, 2 and 3 jet samples to the $W+4$ or more jet sample for both the tagged sample, $f_g^{bkg^{Tagged}}$, and the no-tag sample, $f_g^{bkg^{Notag}}$. We consider the tagged sample as representative of the single top and heavy flavour backgrounds (HF), and the no-tag sample as representative of the light flavour background (LF). As the background coming from QCD processes (non-W) consists of both HF and LF events, we consider half of this background to contribute to HF and half to the LF background. We ignore the negligible contribution to HF from $Z \to b\bar{b}$ in the diboson background. Finally, one can get an estimate of f_g^{bkg} by using the $f_g^{bkg^{Tagged}}$ and $f_g^{bkg^{Notag}}$ weighted by the corresponding background fractions, f_{bkg}^{HF} and f_{bkg}^{LF}. This results in

$$f_g^{bkg} = f_g^{bkg^{Notag}} f_{bkg}^{LF} + f_g^{bkg^{Tagged}} f_{bkg}^{HF}. \tag{4.3}$$

We determine the f_g^{bkg} uncertainty assuming Gaussian distributions for the four variables used to define f_g^{bkg} and then numerically calculating a distribution for f_g^{bkg}. Using this method, we find a value of 0.53 ± 0.09 for the gluon-rich fraction of the background. The

gluon fractions found by the fit to the W+1, 2 and 3 jet samples in both tagged and no-tag samples as well as the extrapolated values are shown in Table 4.2. As examples, the fits to the tagged W+1 jet sample and the no-tag W+3 jet sample are shown in Fig. 4.9 and Fig. 4.10, respectively. The gluon-rich fraction in the no-tag sample increases with increasing jet multiplicity. The tagged sample has contributions from $t\bar{t}$ in the 2 and 3 jet bin and as such, assuming SM predictions for $t\bar{t}$, one would expect the actual gluon-rich fraction to be different from what we observe if we take into account the contribution from $t\bar{t}$. To correct for this, one needs to know $\sigma(gg \to t\bar{t})/\sigma(p\bar{p} \to t\bar{t})$. As this is the variable we measure and given the small contribution of tagged f_g to the calculation of $f_g^{t\bar{t}}$, we use the observed fractions with no correction.

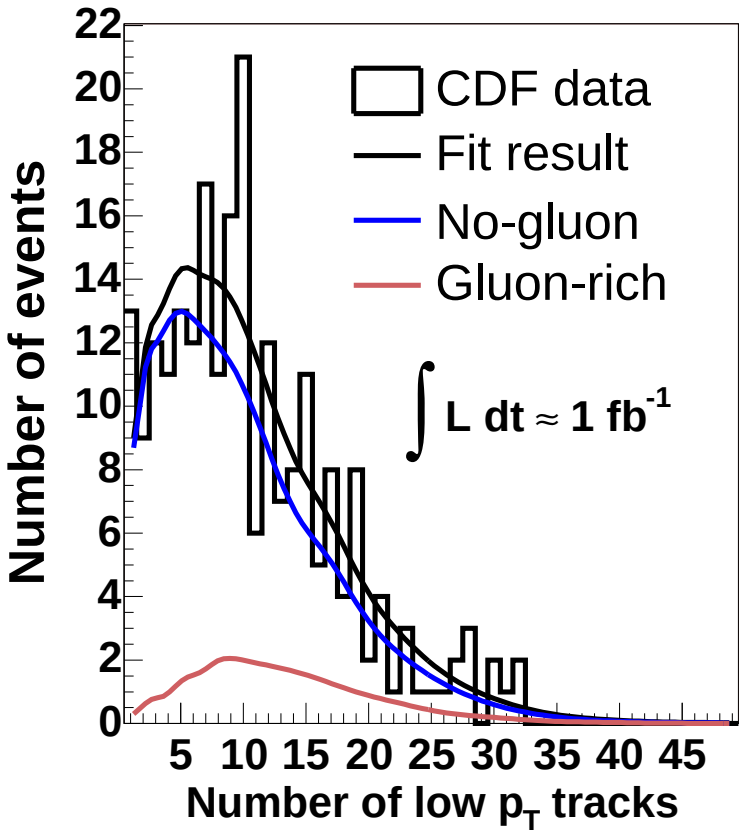

Figure 4.8: The fit result for the tagged W+≥4 jet sample. The two components of the fit (gluon-rich and no-gluon) are also shown.

Jet Multiplicity	f_g in the No-tag Sample	f_g in the Tagged Sample
W+1 jet	0.41 ± 0.01	0.56 ± 0.05
W+2 jet	0.51 ± 0.02	0.42 ± 0.08
W+3 jet	0.56 ± 0.04	0.44 ± 0.12
Extrapolated W+≥4 jets	0.72 ± 0.05	0.27 ± 0.19

Table 4.2: Gluon-rich fraction values from the likelihood fit to the low p_T track multiplicity distributions for W+1, 2 and 3 jet samples with no positive b-tag and with at least one positive b-tag, as well as the extrapolated gluon-rich fractions for both tagged and no-tag sets.

The background fractions used in the analysis, f_b, f_{bkg}^{HF} and f_{bkg}^{LF} in the tagged sample, are summarized in Table 4.3. Using $f_b = 0.13 \pm 0.02$, $f_g^{bkg} = 0.53 \pm 0.09$ and measured $f_g = 0.15 \pm 0.14$, we get $f_g^{t\bar{t}} = 0.09 \pm 0.16$. The systematic uncertainties will be discussed in a subsequent section.

	Tagged W+4 or more jet sample
HF/bkg	0.45 ± 0.11
LF/bkg	0.55 ± 0.13
bkg/S+bkg	0.13 ± 0.02

Table 4.3: The background fractions used in the analysis, f_b, f_{bkg}^{HF} and f_{bkg}^{LF} in the tagged sample.

4.3 $\sigma(gg \to t\bar{t})/\sigma(p\bar{p} \to t\bar{t})$

The last step to measure $\sigma(gg \to t\bar{t})/\sigma(p\bar{p} \to t\bar{t})$ is to estimate the relative acceptance of $gg \to t\bar{t}$ and $p\bar{p} \to t\bar{t}$. To do so, we use a PYTHIA MC calculations. We use about four million $t\bar{t}$ events, of which about 5×10^4 arise from gg fusion and about 8×10^5

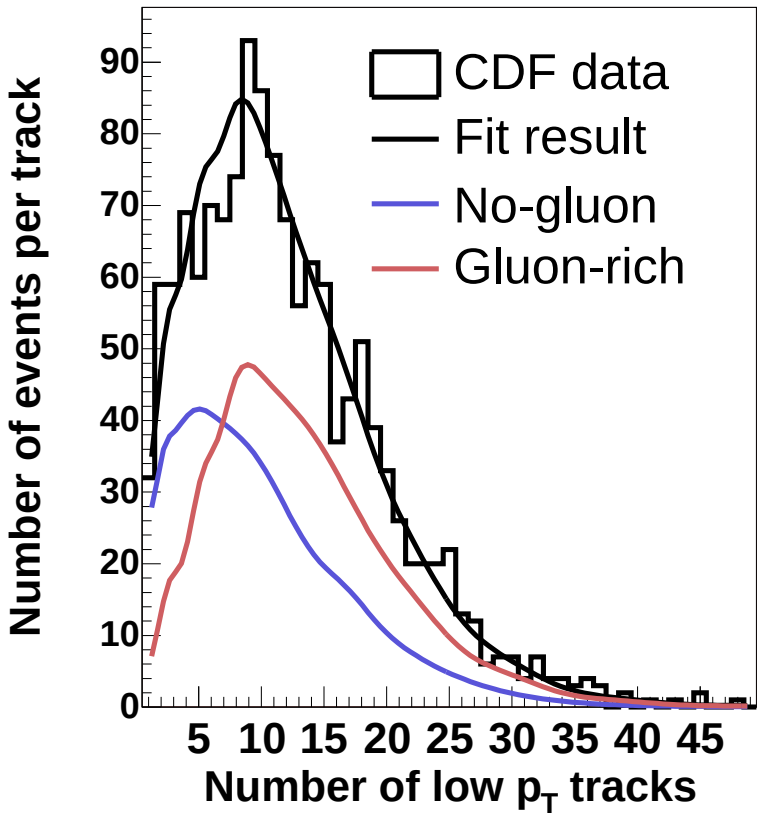

Figure 4.9: The fit result for the tagged W+1 jet sample. The two components of the fit (gluon-rich and no-gluon) are also shown.

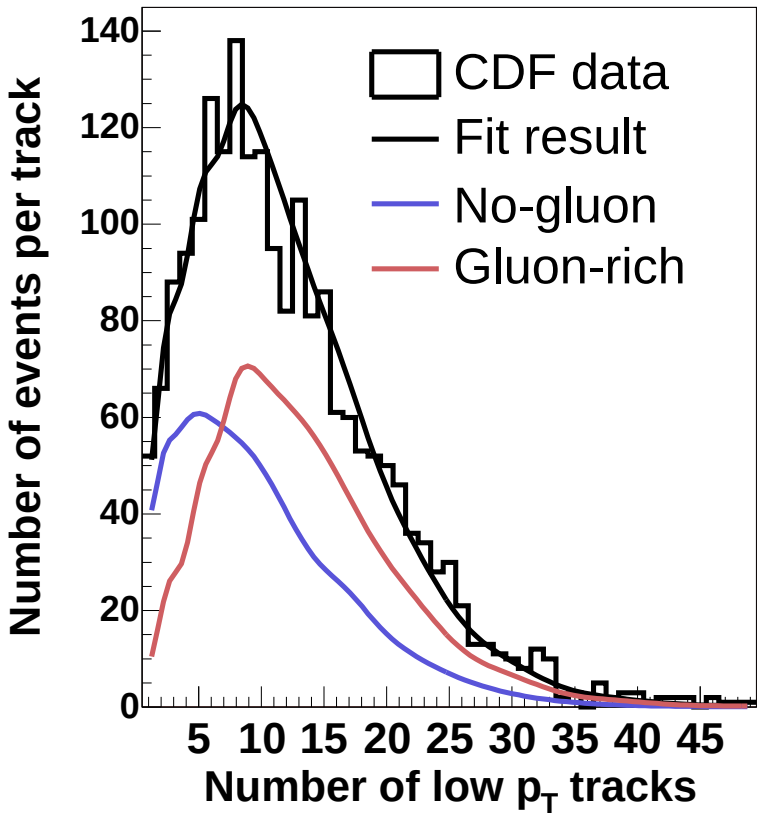

Figure 4.10: The fit result for the no-tag W+3 jet sample. The two components of the fit (gluon-rich and no-gluon) are also shown.

come from $q\bar{q}$ annihilation. The fraction of $gg \to t\bar{t}$ events that falls in the 4 or more jet bins is higher than that of the $q\bar{q} \to t\bar{t}$, as expected due to the higher gluon radiation probability for gluons. Using the MC calculations, we find $(14.1 \pm 0.5(stat + syst))\%$ of $gg \to t\bar{t}$ and $(11.5 \pm 0.4(stat + syst))\%$ of $q\bar{q} \to t\bar{t}$ events pass our tagged sample criteria. These numbers do not include a standard correction to the MC b-tag efficiency or the W boson branching fractions. As we are interested in the relative acceptance, the effects of these factors cancel out. We find

$$\frac{\sigma(gg \to t\bar{t})}{\sigma(p\bar{p} \to t\bar{t})} = \frac{1}{1 - (\mathcal{A}_{gg\to t\bar{t}}/\mathcal{A}_{q\bar{q}\to t\bar{t}}) + (\mathcal{A}_{gg\to t\bar{t}}/\mathcal{A}_{q\bar{q}\to t\bar{t}})(1/f_g^{t\bar{t}})} = 0.07 \pm 0.14(\text{stat}), \quad (4.4)$$

where $\mathcal{A}_{gg\to t\bar{t}}$ and $\mathcal{A}_{q\bar{q}\to t\bar{t}}$ are the acceptance for $gg \to t\bar{t}$ and $q\bar{q} \to t\bar{t}$, respectively.

4.4 Systematic Uncertainties

The systematic uncertainties of this measurement are estimated in a few steps. First, we identify and find the uncertainties affecting the track multiplity distributions. The estimates, in principle, are done by adjusting the central values and observing the changes in the relevant variables and distributions used in the calculation of $\sigma(gg \to t\bar{t})/\sigma(p\bar{p} \to t\bar{t})$. Second, we find the uncertainties for the measured gluon-rich fraction and the background gluon-rich fraction estimates. We then use these uncertainties and the background fraction uncertainties to find the uncertainty in the $t\bar{t}$ gluon-rich fraction and propogate these into a systematic uncertainty in $\sigma(gg \to t\bar{t})/\sigma(p\bar{p} \to t\bar{t})$.

4.4.1 Sources of Uncertainy in Track Multiplicity Distribution

The track selection requirements and the corrections applied to get the N_{trk} values, as decribed in section 3.4, as well as the estimated $<N_g>$ and background fractions used to obtain the no-gluon and gluon-rich distributions, as described in section 4.1.1, can systematically affect the shape of the N_{trk} distributions. These sources and their contribution to the estimated systematic uncertainties are described below.

- The process composition of W+0 jet and dijets with E_T of 80-100 GeV

  We have used ALPGEN+PYTHIA and PYTHIA MC calculations for the process composition of W+0 jet and dijet events with E_T of 80-100 GeV samples, respectively. We also had the MADGRAPH+PYTHIA estimates for the W+0 jet sample. We have used a central value of 0.27 $\pm$ 0.03 for the $qq \rightarrow qq$ process for the dijets with leading jet E_T in the range 80-100 GeV and 0.15 $\pm$ 0.10 for $< N_g >$ of W+0 jet sample. To find the systematic uncertainties due to the quark-gluon compositions used in the definition of gluon-rich and no-gluon distributions, we fluctuated the central values by one standard deviation.

- The choice of jet E_T threshold

  One expects higher numbers of jets coming from initial or final state gluon radiation in events with higher gluon content. As we exclude the low p_T tracks that fall within a radius of 0.4 from the centroid of low E_T jets (6-15 GeV), our low p_T track multiplicity distribution might change differently for the gluon-rich and no-gluon events. To estimate the effect of this cut, we measure f_g and estimate f_g^{bkg} using a low E_T cut of 8 GeV instead of 6 GeV resulting in a conservative estimate.

- The track multiplicity correction per high E_T jet

  To reduce contributions to $< N_{trk} >$ from the high E_T jets present in the event, we make additional corrections of 0.90 ± 0.03, 0.97 ± 0.04 and 0.96 ± 0.04 tracks per jet, discussed in section 3.4, to the track multiplicity of the event for each central high E_T jet in the three subsets of data, differing primarily in their average instantaneous luminosity. We estimate the uncertainty associated with this correction by making the correction of $\pm 1\sigma$ of the central value for each dataset before combining them.

4.4.2 Other Systematic Uncertainies

Apart from the sources described in the previous section affecting the N_{trk} distributions, there are other variables and procedures used to obtain the final result, namely the acceptance values and the estimated f_g^{bkg}. The choice of tools and methods used for these estimates can systematically affect the measurement. The effects of these choices are summarized below.

- The estimation of f_g^{bkg}

  As mentioned in section 3, we estimate this value by extrapolation in the no-tag and tagged samples weighted by the HF and LF background fractions. Therefore, the sources mentioned above change the estimate of f_g^{bkg}. The systematic uncertainty associated with this variable is the root-square sum of uncertainty in the central value, half of the difference in the values if we assign all non-W background to LF or to HF backgrounds and half of the difference of the low and high values of each of the above uncertainties, except for the low jet E_T cut. In the latter, we take the difference instead of half of the difference as we do not use the values for a lower E_T cut given the uncertainties associated with defining a jet with E_T less than 6 GeV.

- The acceptance for $t\bar{t}$ events

  We associate a systematic uncertainty of 3% for the acceptance due to the parton distribution function (PDF) and MC generator differences. This value is based on the uncertainties due to PDF (2%) and choice of MC generator (2%) in the $t\bar{t}$ production cross section measurement reported in [14].

These systematic uncertainties are summarized in Table 4.4. It is worth mentioning that there is no initial-state, final-state or multiple interaction uncertainties associated with the measurement, as we properly averaged over these variables in this measurement. We also do not rely on any modeling of the initial or final state radiation. Additionally, as the

same data taking period is used for the data used to produce the no-gluon, gluon-rich and the different samples N_{trk} distributions, no systematic effect due to possible differences arising from instantaneous luminosity effects is expected.

Taking into account these systematics effects, we find

- $f_g = 0.15 \pm 0.14 \pm 0.07$,
- $f_g^{bkg} = 0.53 \pm 0.11$,
- $f_g^{t\bar{t}} = 0.09 \pm 0.16 \pm 0.08$,
- $\frac{\mathcal{A}_{gg\to t\bar{t}}}{\mathcal{A}_{q\bar{q}\to t\bar{t}}} = 1.23 \pm 0.06$

and then determine $\sigma(gg \to t\bar{t})/\sigma(p\bar{p} \to t\bar{t})$ as

$$\frac{1}{1 - (\mathcal{A}_{gg\to t\bar{t}}/\mathcal{A}_{q\bar{q}\to t\bar{t}}) + (\mathcal{A}_{gg\to t\bar{t}}/\mathcal{A}_{q\bar{q}\to t\bar{t}})(1/f_g^{t\bar{t}})} = 0.07 \pm 0.14 \pm 0.07. \tag{4.5}$$

The result corresponds to an upper limit of 33% for the fraction of $gg \to t\bar{t}$ production cross section at 95% confidence level (CL), where the statistical and systematic uncertainties are included. To find the upper limit, for any specific true fraction we create 10^4 pseudoexperiments using the no-gluon and gluon-rich N_{trk} distributions, assuming 13% background and 53% gluon-rich fraction in the background. For each pseudoexperiment, we use the likelihood fit discussed in section 4.1.2 to find the gluon-rich fraction in the generated sample and follow the calculations to find the result for $\sigma(gg \to t\bar{t})/\sigma(p\bar{p} \to t\bar{t})$. The relevant variables involved in the calculation are drawn from a Gaussian probability distribution that takes into account the systematic uncertainties in the variables. To account for the remaining systematic uncertainties, we smear the result for $\sigma(gg \to t\bar{t})/\sigma(p\bar{p} \to t\bar{t})$ with a Gaussian distribution of mean 0.0 and width 0.7. We then find the percentage of pseudoexperiments with a final result greater than 0.07. In case of true fraction of 33%, 95% of the pseudoexperiments give such a result.

	f_g	f_g^{bkg}
Track/Jet Correction	± 0.051	± 0.001
Low E_T Jet Cut	± 0.021	± 0.035
$qq \to qq$ Contribution to the Dijet Sample	± 0.002	± 0.019
W+0 Jet $< N_g >$	± 0.039	± 0.007
non-W Variation	-	± 0.06
Modeling the f_g^{bkg} Distribution	-	± 0.09
Total	± 0.07	± 0.11
	$f_g^{t\bar{t}}$	
f_g	± 0.08	
f_g^{bkg}	± 0.02	
f_b	± 0.01	
Total	± 0.08	
	$\sigma(gg \to t\bar{t})/\sigma(p\bar{p} \to t\bar{t})$	
$f_g^{t\bar{t}}$	± 0.07	
$\mathcal{A}_{gg\to t\bar{t}}/\mathcal{A}_{q\bar{q}\to t\bar{t}}$	± 0.004	
Total	± 0.07	

Table 4.4: Sources of systematic effects and their effects on the measured values.

Chapter 5

Conclusion

We made the first measurement of the fraction of top quark pair production through gluon-gluon fusion. Even though the partonic cross sections for top quark pair production and hence $\sigma(gg \to t\bar{t})/\sigma(p\bar{p} \to t\bar{t})$ have been considered as part of the total top quark pair production cross section, there had never been an attempt to measure this quantity. A precise measurement of this quantity tests the pQCD calculations for the relative $t\bar{t}$ production cross section. A deviation from the pQCD prediction can hint to the existence of beyond SM production and decay mechanisms for the top quark. Although the total $t\bar{t}$ cross section in $p\bar{p}$ collisions is measured in both 1.8 and 1.96 TeV using many different methods and in different decay channels [14, 15, 16] and the results are in agreement with the standard model predictions, there have been suggestions that there could exist production mechanisms beyond the SM for top quarks whose effect is balanced with the beyond SM decay mechanisms for top quark decay [27]. Furthermore, such a measurement helps to place better constraints on the uncertainty of the momentum distribution of gluons in protons at high x.

We have studied the properties of the $t\bar{t}$ production through different intitial states, gg or $q\bar{q}$, using MC calculations and identified arguably the most sensitive discriminating variable, namely the low p_T track multiplicity. As discussed in section 1.3, the probability

for a gluon to radiate a gluon is higher than that for a quark. Therefore we expect to see larger number of low p_T charged particles in processes involving more gluons. One of the main advantages of using the low p_T track multiplicity is that this variable is observed experimentally with no need to distinguish between the decay products of the top and antitop quarks as would have been the case for some other variables such as p_T of the top or antitop quark. As such we do not expect to lose the discriminating power of this variable due to the smearing effects of incorrect final state reconstruction.

The large uncertainties involved in the modeling of low p_T gluon radiation makes MC calculations an undesirable method to predict the low p_T track multiplicity. Therefore we employ a data-driven method for the measurement, giving it a unique experimental characteristic. We use data calibration samples with well-known production processes, namely W+n jet and dijet samples. The larger the number of jets in the W+n jet samples and the lower the E_T range of the leading jet in the dijet sample, we expect larger number of gluons to be involved in the production processes. We use these calibration samples and show that there exists a correlation between the observed average number of low p_T tracks and the average number of gluons involved in the process calculated using MC calculations to predict the relative rate of the parton interactions. We then take advantage of this correlation and define no-gluon and gluon-rich low p_T track multiplicity distributions from the W+0 jet sample and from the dijet sample with leading jet E_T of 80-100 GeV as explained in section 4.1.1. We then use these distributions in a likelihood fit to find the fraction of gluon-rich events in the $t\bar{t}$ candidate sample.

This fraction consists of a signal and a background contribution. We estimate the contribution from the background by measuring the gluon-rich fractions of W+1, 2 and 3 jet samples with no b-tag and with at least one b-tag jet and extrapolate their values to the 4 or more jet sample. Assuming that the b-tag and no-tag extrapolations represent the backgrounds with and without real heavy quarks, we estimate the gluon-rich fraction of the background. Finally, we subtract this contribution from the fit result and translate

the observed fraction of gluon-rich events to a measurement of the fraction of $gg \to t\bar{t}$ production using the acceptance values of the $t\bar{t}$ events produced through gg and $q\bar{q}$ initial states.

We find a value of $\sigma(gg \to t\bar{t})/\sigma(p\bar{p} \to t\bar{t})$ =0.07±0.14(stat)±0.07(syst) in agreement with the NLO prediction of 0.15±0.05. We can conclude that the $q\bar{q}$ initial state is the dominant channel for $t\bar{t}$ production. The result shows that at least 67% of the $t\bar{t}$ events are consistent with $q\bar{q} \to t\bar{t}$ at 95% confidence level, leaving little room for non SM processes which are not similar to the $q\bar{q} \to t\bar{t}$ in their gluon radiation. Since the ttbar production process is sensitive to the modelling of interactions at high transverse momentum, it confirms our current theoretical understanding of this process. It also confirms that the relative parton distribution functions are correct in the less well-explored regime of high Q^2 and x.

5.1 Possible Future Improvements

The result is statistics-limited with a statistical uncertainty twice the size of the systematic uncertainty. The data-driven method used for the analysis is a major contributor to the relatively small systematic uncertainty. Using the data eliminates the uncertainties associated with the contribution to the low p_T track multiplicity from the underlying event, extra interactions and multiple scattering, as, on average, these sources contribute similarly to both the gluon-rich and no-gluon distributions defined using data. Figure 5.1 shows the expected statistical uncertainties as a function of integrated luminosity. The plot only takes into account the increase in the statistics in the candidate sample and considers no improvement in the method or combination with any other measurement. Another effect that has not been considered in this extrapolation is the possible smearing effect of higher instantaneous luminosity in the future data compared to the earlier data samples. However, we believe this would not compromise the technique and in the

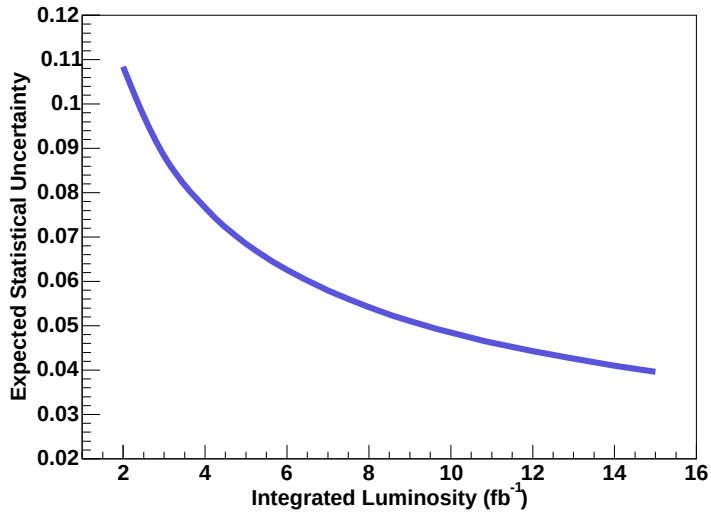

Figure 5.1: The expected statistical uncertainty in the measurement of the fraction of $t\bar{t}$ production through gg fusion at CDF as a function of integrated luminosity.

worst case scenario one could categorize the data samples based on their instantaneous luminosity if need be.

As can be seen in Fig. 5.1 with about 5 fb^{-1} of data we can get a measurement with an absolute statistical uncertainty of about 7%, comparable to the present systematic uncertainty. The largest source of systematic uncertainty for the measurement, about 5%, is due to the correction in the track multiplicity due to the presence of high E_T jets in the tracking region. This correction is estimated using multijet samples, and so higher statistics may help in reducing the uncertainty associated with this estimate and in turn result in a smaller systematic uncertainty. The second largest source of systematic uncertainty, contributing about 4%, is the gluon content of the W+0 jet sample. This value is estimated partly from MC calculations and partly from the QCD background estimates using data. The main uncertainty from this source comes from the MC calculation and the estimate of the average number of gluons involved in the QCD background. Therefore, an increase in statistics does not improve this uncertainty directly. The other

systematic uncertainty source is the low E_T jet cut, contributing about 2%. This is an important quantity as it can be considered the border of rejecting low p_T tracks common to the $t\bar{t}$ final state or low p_T tracks originating from initial state radiation. There is no trivial way of separating these tracks apart and therefore we do not expect to be able to reduce this uncertainty easily. The larger statistics allows us to have a better estimate of the background fraction and the gluon content of the background. However, this is a relatively small background, (13±2)%, and so we do not expect to see a large improvement in the overall systematic uncertainty arising from this source.

5.2 Future Prospects at the Tevatron and LHC

There is a growing interest in the study of $t\bar{t}$ production mechanism evident from other efforts at CDF to make a measurement of the same quantity, the fraction of $t\bar{t}$ production through gluon-gluon fusion, using a Neural Network (NN) analysis and a multi-variant technique [54].

Along with a more precise measurement at the Tevatron, it would be very interesting to measure the same fraction at the LHC. According to the SM, we expect gluon-gluon fusion to be the dominant production channel at the LHC, corresponding to about 90% of the total cross section. Given the high luminosity and therefore larger number of extra interactions at the LHC, the low p_T track multiplicity may not be a suitable variable to discriminate between the two SM production channels. To find the best method for the measurement at the LHC, one needs to study the characteristics of the $t\bar{t}$ production through different production channels as well as taking into account the strengths of the detectors.

As a final note, even though this first measurement, as with almost any other first measurements, does not have the desired precision to fully address all the relevant questions motivating the analysis, it provides the stepping stones for a better understanding

of the $t\bar{t}$ production mechanism and its underlying physics. Furthermore, its data-driven technique is a testimony to the experimental possibilities made available when one designs measurements that are, to the extent possible, self-calibrating.

Bibliography

[1] E. Rutherford, Philosophical Magazine Series 6 **vol. 21**, 669 (1911).

[2] S.L. Glashow, Nucl. Phys. **22**, 579 (1961);
S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967);
A. Salam, in *Elementary Particle Theory: Relativistic Groups and Analyticity* (Nobel Symposium No. 8), edited by N. Svartholm (Almqvist and Wiksell, sweden), 367 (1968).

[3] A review of QCD is provided in W. Marciano and H. Pagels, Phys. Rept. **36**, 137 (1978).

[4] A. Salam and J.C. Ward, Phys. Lett. **13**, 168 (1964);
S.L. Glashow, J. Iliopoulos and L. Maiani, Phys. Rev. D **2**, 1285 (1970);
Makoto Kobayashi and Toshihide Maskawa, Prog. Theor. Phys. **49**, 652 (1970).

[5] W.-M. Yao *et al.*, J. Phys. **G 33**, 1 (2006).

[6] Guthrie Miller *et al.* (SLAC-MIT Collaboration), Phys. Rev. D **5**, 528 (1972);
A.C. Benvenuti *et al.* (BCDMS Collaboration), Phys. Lett. B **223**, 485 (1989);
T. Ahmed *et al.* (H1 Collaboration), Nucl. Phys. **B439**, 471 (1995);
M. Derrick *et al.* (ZEUS Collaboration), Z. Phys. C **65**, 379 (1995);
A.D. Martin, W.J. Stirling and R.G. Roberts, Phys. Rev. D **50**, 6734 (1994).

[7] J. Pumplin, D.R. Stump, J. Huston, H.L. Lai, P. Nadolsky and W.K. Tung, J. High Energy Phys. **0207**, 012 (2002).

[8] A.D. Martin, W.J. Stirling, R.S. Thorne and G. Watt, Phys. Lett. B **652**, 292 (2007).

[9] S.W. Herb *et al.*, Phys. Rev. Lett. **39**, 252 (1977).

[10] H. Albrecht *et al.* (ARGUS Collaboration), Phys. Lett. B**192**, 245 (1987);
F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **68**, 447 (1992);
F. Abe *et al.* (CDF Collaboration), Phys. Rev. D **45**, 3921 (1992);
S. Abachi *et al.* (D0 Collaboration), Phys. Rev. Lett. **72**, 2138 (1994);
J. Alitti *et al.* (UA2 Collaboration), Phys. Lett. B**276**, 365 (1992);
F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **73**, 220 (1994);
LEP Collaborations (ALEPH, DELPHI, L3 and OPAL) and LEP Electroweak Working Group, CERN-PPE-93-157 (1993). J.R. Ellis, M.K. Gaillard, D.V. Nanopoulos and S. Rudaz, Nucl. Phys. **B131**, 285 (1977)[Erratum-ibid. **B132**, 541 (1978)]; A. Ali and Z.Z. Aydin, Nucl. Phys. **B148**, 165 (1979); U. Amaldi, A. Bohm, L.S. Durkin, P. Langacker, A.K. Mann, W.J. Marciano, A. Sirlin and H.H. Williams, Phys. Rev. D **36**, 1385 (1987).

[11] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **74**, 2626 (1995);
S. Abachi *et al.* (D0 Collaboration), Phys. Rev. Lett. **74**, 2632 (1995).

[12] A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. D **75**, 052001 (2007);
A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. Lett. **98**, 072001 (2007);
A. Abulencia *et al.* (CDF Collaboration), Phys. Lett. B **639**, 172 (2006);
A. Affolder *et al.* (CDF Collaboration), Phys. Rev. Lett. **87**, 102001 (2001);
V.M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **98**, 181802 (2007);
V.M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **98**, 041801 (2007);

V.M. Abazov *et al.* (D0 Collaboration), Phys. Rev. D **75**, 031102 (2007);
V.M. Abazov *et al.* (D0 Collaboration), Phys. Lett. B **639**, 616 (2006).

[13] CDF Collaboration and D0 Collaboration, FERMILAB-TM-2380-E, hep-ex/0703034 (2007).

[14] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. D **71**, 052003 (2005).

[15] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. D **71**, 072005 (2005);
D. Acosta *et al.* (CDF Collaboration), Phys. Rev. Lett **93**, 142001 (2004);
D. Acosta *et al.* (CDF Collaboration), Phys. Rev. D **72**, 032002 (2005);
A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. D **74**, 072005 (2006);
A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. D **74**, 072006 (2006);
A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. Lett. **96**, 202002 (2006).

[16] V.M. Abazov *et al.* (D0 Collaboration), Phys. Rev. D **76**, 092007 (2007);
V.M. Abazov *et al.* (D0 Collaboration), Phys. Rev. D **76**, 052006 (2007);
V.M. Abazov *et al.* (D0 Collaboration), Phys. Rev. D **74**, 112004 (2006);
V.M. Abazov *et al.* (D0 Collaboration), Phys. Rev. D **76**, 072007 (2007).

[17] N. Kidonakis and R. Vogt, Phys. Rev. D **68**, 114014 (2003);
E.L. Berger and H. Contopanagos, Phys. Lett. B **361**, 115 (1995);
W. Bernreuther *et al.*, Nucl. Phys. **B690**, 81 (2004).

[18] J.C. Collins, D.E. Soper and G. Sterman, Nucl. Phys. **B263** 37 (1986).

[19] J.H. Kühn, *Lectures presented at the XXIII SLAC Summer Institute on Particle Physics, "The Top Quark and the Electroweak Interaction", July 10-21, 1995, SLAC, Stanford, USA, hep-ph/9707321.*

[20] M. Cacciari, S. Frixione, M.L. Mangano, P. Nason and G. Ridolfi, J. High Energy Phys. **0404**, 068 (2004).

[21] V.N. Gribov and L.N. Lipatov, Sov. J. Nucl. Phys. **15**, 438 (1972);
L.N. Lipatov, Sov. J. Nucl. Phys. **20**, 94 (1975);
G. Altarelli and G. Parisi, Nucl. Phys. **B126**, 298 (1977);
Y.L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977).

[22] R.K. Ellis, W.J. Srirling and B.R. Webber, *QCD and Collider Physics*, Cambridge University Press (1996).

[23] F.A. Berends, H. Kuijf, B. Tausk, and W.T. Giele, Nucl. Phys. **B357**, 32 (1991).

[24] CDF Collaboration, FERMILAB-CONF-06-377-E (2006).

[25] B.L. Combridge, J. Kripfganz, and J. Ranft, Phys. Lett. B **70**, 234 (1977).

[26] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **71**, 2542 (1993). B. Abbott *et al.* (D0 Collaboration), Phys. Rev. Lett. **82**, 2457 (1999).

[27] G.L. Kane and S. Mrenna, Phys. Rev. Lett. **77**, 3502 (1996).

[28] `http://www-bdnew.fnal.gov/operations/rookie_books/rbooks.html`, Tevatron Rookie Book v1.0 (2006).

[29] `http://www-bdnew.fnal.gov/operations/rookie_books/rbooks.html`, Antiproton Source Rookie Book v1.1 (2006).

[30] V. Shiltsev, A. Jansson and R. Moore, FERMILAB-CONF-06-099 (2006).

[31] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. D **71**, 032001 (2005).

[32] The CDF II Detector Technical Design Report, (CDF Collaboration), FERMILAB-Pub-96/390-E.

[33] A. Sill *et al.*, Nucl. Instrum. Meth. **A447**, 1 (2000).

[34] C.S. Hill *et al.*, Nucl. Instrum. Meth. **A530**, 1 (2004).

[35] A. Affolder *et al.*, Nucl. Instrum. Meth. **A453**, 84 (2000).

[36] A. Affolder *et al.*, Nucl. Instrum. Meth. **A526**, 249 (2004).

[37] S. Balka *et al.*, Nucl. Instrum. Meth. **A267**, 272 (1988).

[38] S. Bertolucci *et al.*, Nucl. Instrum. Meth. **A267**, 301 (1988).

[39] M. Albrow *et al.*, Nucl. Instrum. Meth. **A480**, 524 (2002).

[40] G. Apollinari *et al.*, Nucl. Instrum. Meth. **A412**, 515 (1998).

[41] G. Ascoli *et al.*, Nucl. Instrum. Meth. **A268**, 33 (1988).

[42] T. Dorigo *et al.*, Nucl. Instrum. Meth. **A461**, 560 (2001).

[43] D. Acosta *et al.*, Nucl. Instrum. Meth. **A494**, 57 (2002).

[44] A. Bhatti *et al.*, Nucl. Instrum. Meth. **A566**, 375 (2006).

[45] R. Brun and F. Carminati, "GEANT Detector Description and Simulation Tools", CERN Programming Library Long Writeup **W5013** (1993).

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

[47] G. Corcella *et al.*, J. High Energy Phys. **01**, 10 (2001).

[48] S. Frixione *et al.*, J. High Energy Phys. **08**, 07 (2003).

[49] M.L. Mangano *et al.*, J. High Energy Phys. **0307**, 001 (2003).

[50] S. Hoche *et al.*, hep-ph/0602031 (2006).

[51] J. Alwall *et al.*, J. High Energy Phys. **0709**, 028 (2007).

[52] F. Krauss, J. High Energy Phys. **0208**, 015 (2002).

[53] Private communications with Joey Huston.

[54] R. Eusebi, E. Halkiadakis, S. Somalwar and J. Yamaoka, CDFNote 8811 (2007).