

Measurement of the top pair-production cross
section in the dilepton channel in
proton-antiproton collisions at $\sqrt{s}=1.96$ TeV

by

Mircea Norocel Coca

Submitted in Partial Fulfillment

of the

Requirements for the Degree

Doctor of Philosophy

Supervised by

Professor Paul Tipton

Department of Physics and Astronomy

The College

Arts and Sciences

University of Rochester

Rochester, New York

2004

Dedicated to my wonderful parents, Viorica and Mihai
Dedicated to the memory of my grandparents, Stelian and Ana

Curriculum Vitae

Mircea Norocel Coca was born in Iasi, Romania on September 27, 1973. He completed his high-school education at "Emil Racovita" High-School, in Iasi, in 1988. He continued at the "Al. I Cuza" University from Iasi, graduating with a Bachelor of Arts degree in Physics, in 1997. From 1997 to 1998 he studied in Groningen University, The Netherlands, as a Socrate fellow, doing research in Theoretical Nonlinear Optical Spectroscopy. In August 1998 he began the graduate studies at University of Rochester, NY, in the Department of Physics and Astronomy. He received the Master of Arts degree in Fall 1999. Between 1998 and 2000 worked in Theoretical Particle Physics under the direction of Professor Sarada Rajeev. He joined the CDF Collaboration in September 2000 and pursued high physics energy research as a student of Professor Paul Tipton.

Acknowledgments

I owe much to countless people who encouraged, supported and dedicated much of their time to me, over the years. Unfortunately I can only mention few of them.

No ship can survive the storms and reach the shore without a master, and Paul was everything I could ask from an advisor, encouraging me to always pursue my interests, always succeeding in showing me the paths to solve problems, always showing patience with me. I am indebted to Paul for offering me the chance to enjoy a wonderful Summer School in Virgin Islands and “une vie de bohème” in Paris for two weeks. Thanks for giving my half the chance to meet and work with Dr. Stecy, Andreea is so grateful for this. Thanks so much for everything you have done for me.

I want to thank to Sarada Rajeev, my first advisor, when I arrived in Rochester in 1998, for opening my horizon to so many aspects of the theoretical realm of physics and mathematics.

I owe a great appreciation to Mark Kruse for his constant guidance, advice and friendship. We worked extremely well together on the first iterations of the analysis, and since then he was a constant source of knowledge I benefited from so many times.

Willis Sakumoto, many thanks for always answering my questions, teaching me so much about every single detail of the CDF detector and for showing me how exciting and rewarding can be analyzing the first $3 pb^{-1}$ of Run II data; for reading and carefully reviewing this thesis.

Thanks to Steve Blusk, my boss as I liked to called, my officemate for only an year, for introducing me to the interlock and cooling system, and silicon detector, and constant guidance and friendship after leaving the Rochester HEP Group.

Thanks to Eva Halkiadakis for the fruitful collaboration we had and for teaching me so many qualities a leader must have.

Many thanks to our university administrative assistants, especially Barbara, Connie and Sue, for always being there ready to help.

I want to thank to Jaco, Pat, Andy and many colleagues in the Top and Top Dilepton Group, for providing support and help to understand and solve all the aspects of my analysis, and maintaining an enjoyable team spirit, even in the exhausting “blessing” times.

Kind thanks to Young Kee Kim, always ready to spend few minutes when I needed the advice the most.

Thank to Henry Frisch, for supporting me over the time and conveying me a part of his constant excitement for science and life.

Special thanks to two of many of my great high-school teachers, who influenced my interest and the desire to pursue physics, Cozma Violeta, at “V. Alecsandri” High-School and Talpalaru Seryl, at “E. Racovita” High-School, Iasi. Thanks to many of my teachers at “Al. I. Cuza” University, Iasi, and especially two of them, Gheorghe Popa (classical mechanics and plasma physics) and Gheorghe Singurel, for their encouragement, friendship and constant guidance during my undergraduate years spent in Romania.

Thanks to my amazing and loving wife, Andreea, you are my best friend, my loving angel and the love of my heart. None of this would have been possible without you.

Table of Contents

Curriculum Vitae	iii
Acknowledgments	iv
List of Tables	x
List of Figures	xiv
Abstract	xviii
1 Introduction	1
2 Standard Model Theory and The Top Quark Physics	6
2.1 The Standard Model	6
2.2 Top Quark Physics	12
2.3 Top quark production and decay	13
2.4 Top Pair Production Cross-Section	18
3 The Collider Detector at Fermilab	25
3.1 The 1992-1995 Run I and 2000-2003 Run II	26
3.2 The CDF Coordinate System	29
3.3 The CDF Calorimeters	32
3.4 The CDF Tracking Detectors	38

3.5	The CDF Muon Detectors	47
3.6	Cerenkov Luminosity Counters	49
3.7	CDF Time of Flight System	50
3.8	The CDF Trigger System	51
4	Event Reconstruction	60
4.1	Energy reconstruction	61
4.2	Track reconstruction	64
4.3	Vertex finding	68
4.4	Electron reconstruction	68
4.5	Muon reconstruction	72
4.6	Jet reconstruction	73
4.7	Missing transverse energy (\cancel{E}_T) reconstruction	76
5	Data Samples and Triggers	77
5.1	Good Run Criteria	77
5.2	Luminosity Measurement	78
5.3	Inclusive Electron Sample	80
5.4	Plug Electron+Met Sample	81
5.5	Inclusive Muon Sample	82
5.6	QCD Jet Samples	84
6	Top Dilepton Event Selection	86
6.1	Geometrical and kinematic requirements	87
6.2	Lepton Identification	89
6.3	Conversion and Cosmic removal	104
6.4	Isolation cut	104

6.5	Dilepton Event Classification	105
6.6	Invariant Mass cut	106
6.7	\cancel{E}_T cut	109
6.8	$\Delta\phi$ or “L” cut	110
6.9	At least 2 Jets	110
6.10	H_T requirement	112
6.11	Opposite Sign	115
6.12	Trilepton Events	116
7	Top Dilepton Acceptance	117
7.1	Effects of the selection requirements on the acceptance	118
7.2	Results	121
7.3	Systematic uncertainties in the acceptance	124
7.4	Acceptance summary	128
8	Top Dilepton Backgrounds	129
8.1	QCD Fake Leptons	130
8.2	Drell-Yan lepton pair production	138
8.3	Di-tau production	145
8.4	Vector boson pair production	148
8.5	Other small sources of background	154
8.6	Summary of backgrounds	154
9	Unveiling the detector data	157
9.1	The top dilepton candidates	157
9.2	b-tags in the dilepton candidates	164
9.3	$t\bar{t}$ cross section estimation	166

10 Conclusions and Future Prospects	169
10.1 Conclusions	169
10.2 Future Prospects	170
Bibliography	180
Appendices	180
A The Fermilab Accelerator	181
A.1 A proton-antiproton collider	181
A.2 The Production of Protons and Antiprotons	183
A.3 The Tevatron	184
B MC Breakdown Tables	185

List of Tables

2.1	The $t\bar{t}$ decay channels and their corresponding branching ratios.	17
2.2	$t\bar{t}$ cross-section at the Tevatron, at $\sqrt{s} = 1.96$ TeV, for different top masses.	24
3.1	Data taking periods.	27
3.2	Specifications of the CDF calorimeter detectors.	36
3.3	The nominal transverse tower segmentation of the end plug calorimeter.	37
3.4	Overview of the Plug Calorimeter.	38
3.5	Few characteristics of the COT at CDF.	40
3.6	Design parameters of the SVX II detector.	44
3.7	Design Parameters of the CDF II Muon Detectors.	47
3.8	The cross-sections and event rates for few of the interesting physics processes at a hadron collider, compared with the inelastic $p\bar{p}$ cross-section.	52
4.1	The characteristics of the tower types.	63
4.2	The correspondence between η and IETA.	64
5.1	Systematic uncertainties for the luminosity calculation	79
5.2	Summary of luminosities used for the cross measurement.	79
5.3	Parent datasets and associated L3 trigger paths for the tight electron data sample.	80
5.4	Parent datasets and associated L3 trigger paths for the plug data sample.	81

5.5	Parent datasets and associated L3 trigger paths for the tight muon data sample.	82
5.6	The dilepton categories and the triggers used to collect the events.	84
5.7	Associated trigger names at each trigger level for the trigger paths jet20, jet50, jet70 or jet100, used to collect QCD enriched datasets.	85
6.1	Selection requirements for CEM electrons.	91
6.2	Selection requirements for PHX electrons.	95
6.3	Electron identification efficiencies.	96
6.4	Selection requirements for muons.	97
6.5	Muon identification efficiencies.	104
6.6	The dilepton categories used in the analysis.	106
6.7	Invariant mass selection requirements applied only for ee, $\mu\mu$ events.	108
6.8	Invariant mass selection requirements applied only for ee, $\mu\mu$ events.	109
7.1	Individual cut efficiencies and the total dilepton acceptance.	122
7.2	Fraction of total dilepton acceptance corresponding to various luminosities. .	122
7.3	Contribution to the $t\bar{t}$ signal acceptance from various event topologies. . . .	123
7.4	Contribution to the $t\bar{t}$ signal acceptance from various event topologies. . . .	123
7.5	Contribution to the $t\bar{t}$ signal acceptance from generator level lepton pairs . .	123
7.6	Uncertainties of the $t\bar{t}$ signal acceptance.	127
7.7	$t\bar{t}$ dilepton acceptance summary.	128
8.1	Jet fakeable object definitions.	134
8.2	The number of observed and predicted fakes for electron categories.	135
8.3	Fake background estimate in 193 pb^{-1}	136
8.4	Fake background prediction for same-sign events, in different jet bins.	136
8.5	Fake background systematic uncertainties.	138
8.6	AlpGen DY+2p MC samples used to determine the H_T cut efficiency	144

8.7	H_T cut efficiencies taken from AlpGen DY+2p MC.	144
8.8	Expected number of non-DY events inside the Z window.	145
8.9	Summary of DY background.	145
8.10	Efficiencies of the selection cuts applied to a Pythia $Z \rightarrow \tau\tau$ sample.	147
8.11	Fraction of events with 1 or ≥ 2 jets for data and Pythia Monte Carlo.	147
8.12	$Z \rightarrow \tau\tau$ background estimate for 193 pb^{-1}	148
8.13	WW background estimate for 193 pb^{-1}	151
8.14	Efficiencies for the selection cuts applied to a Pythia WZ sample.	152
8.15	Efficiencies for the selection cuts applied to a Pythia ZZ sample.	152
8.16	WZ background estimate for 193 pb^{-1}	153
8.17	ZZ background estimate for 193 pb^{-1}	153
8.18	Grand Table of systematic uncertainties.	155
8.19	Summary of the background contributions to the $t\bar{t}$ dilepton channels.	156
9.1	Breakdown of the ee top dilepton channel analysis of the 193 pb^{-1} of data.	158
9.2	Breakdown of the $\mu\mu$ top dilepton channel analysis of the 193 pb^{-1} of data.	158
9.3	Breakdown of the $e\mu$ top dilepton channel analysis of the 193 pb^{-1} of data.	159
9.4	Multilepton and summary of the analysis of the 193 pb^{-1} of data.	159
9.5	Run and event numbers for the candidate events.	160
9.6	The variables used to determine the SECVTX b-tagging efficiencies.	165
9.7	The b tagging information for the top dilepton candidates.	166
9.8	The number of observed and expected SECVTX tagged jets.	166
9.9	The values used to calculate the top dilepton cross-section.	167
10.1	Few Top Physics Prospects for 2 fb^{-1}	170
B.1	Breakdown of the e/e final states of the $t\bar{t}$ acceptance using Pythia MC.	185
B.2	Breakdown of the μ/μ final states of the $t\bar{t}$ acceptance using Pythia MC.	186

B.3	Breakdown of the e/μ final states of the $t\bar{t}$ acceptance using Pythia MC. . .	187
B.4	Multilepton and summary of the $t\bar{t}$ acceptance using Pythia MC.	187
B.5	Breakdown of the e/e final states of the $Z \rightarrow \tau\tau$ acceptance using Pythia MC.	188
B.6	Breakdown of the μ/μ final states of the $Z \rightarrow \tau\tau$ acceptance using Pythia MC.	188
B.7	Breakdown of the e/μ final states of the $Z \rightarrow \tau\tau$ acceptance using Pythia MC.	189
B.8	Multilepton and summary of the $Z \rightarrow \tau\tau$ acceptance using Pythia MC. . . .	189
B.9	Breakdown of the e/e final states of the WW acceptance using Pythia MC. .	190
B.10	Breakdown of the μ/μ final states of the WW acceptance using Pythia MC.	190
B.11	Breakdown of the e/μ final states of the WW acceptance using Pythia MC. .	191
B.12	Multilepton and summary of the WW acceptance using Pythia MC.	191
B.13	Breakdown of the e/e final states of the WZ acceptance using Pythia MC. .	192
B.14	Breakdown of the μ/μ final states of the WZ acceptance using Pythia MC. .	192
B.15	Breakdown of the e/μ final states of the WZ acceptance using Pythia MC. .	193
B.16	Multilepton and summary of the WZ acceptance using Pythia MC.	193

List of Figures

2.1	WW scattering Standard Model Feynman diagrams.	11
2.2	The 1-loop correction diagram to the Higgs mass.	12
2.3	Lowest order (α_s^2) Feynman diagrams for heavy quark production.	13
2.4	Single top production at the Tevatron.	15
2.5	The proton parton distribution functions for gluon, compared with the up and down quark.	21
2.6	Some representative Feynman diagrams contributing to the next to leading order (NLO) corrections $O(\alpha_s^3)$ to the process $q\bar{q} \rightarrow t\bar{t}$	23
3.1	Longitudinal view of half of the CDF detector.	26
3.2	The weekly and per run integrated luminosity between March 5, 2001 and March 2, 2004.	28
3.3	Run II instantaneous luminosity.	29
3.4	CDF Run II coordinate system.	30
3.5	Schematic view of one wedge of the Central Electromagnetic Calorimeter. . .	34
3.6	Cross section of one quadrant of the plug calorimeter.	38
3.7	Longitudinal View of the CDF II Tracking System.	39
3.8	On the left, three supercells for SL2 are shown, looking along the beam (z) direction. On the right, the endplate slots Sense and field planes are shown.	41
3.9	COT hit width for SL2, run 179643. The aging shows ϕ and Z dependence. .	42

3.10	The r-z view of the CDF II Silicon Detector.	43
3.11	An r-z view of the ISL layers radius. SVXII and COT are also shown.	45
3.12	SVX Cooling System.	46
3.13	Location of the central muon upgrade components in azimuth ϕ and pseudo-rapidity η for Run II.	49
3.14	Time of flight difference as a function of particle type and momentum.	51
3.15	The readout functional block diagram in Run II.	53
3.16	The trigger-system block diagram in Run II.	54
3.17	Schematic of the SVX II Data Acquisition System	56
3.18	Event Builder block diagram.	58
4.1	The logical flow of CDF event reconstruction in Run II.	61
4.2	An example of a Phoenix track reconstruction.	67
6.1	Tight central electron ID variables E_{had}/E_{em} , E/P , L_{shr} , ΔZ , χ^2_{strip} and $Q * \Delta X$	92
6.2	Tight central electron ID variables iso_{frac} , z_0 , and the dielectron invariant mass	93
6.3	CMUP electron ID variables from the data compared with Monte Carlo expectations.	98
6.4	CMU electron ID variables from the data compared with Monte Carlo expectations.	99
6.5	CMP electron ID variables from the data compared with Monte Carlo expectations.	100
6.6	CMX electron ID variables from the data compared with Monte Carlo expectations.	101
6.7	Invariant dimuon mass using at least a CMU or CMP.	102
6.8	The $\eta - \phi$ map of CMU-only and CMP-only muons	103

6.9	Invariant mass for ee and $\mu\mu$ top dilepton events, after lepton ID and isolation cuts.	107
6.10	H_T distributions for $t\bar{t}$ signal and WW, WZ and $Z \rightarrow \tau\tau$	113
6.11	Signal and background as a function of the cut placed on H_T	114
6.12	Various ratios of signal-to-background as a function of cut placed on H_T . . .	114
6.13	Invariant mass distribution for CP ee events in the data.	116
7.1	The pseudorapidity and the transverse momentum of the leptons from W decays, at the generator level.	119
7.2	Invariant mass $M_{\ell\ell}$, the missing transverse energy, the number of jets and the scalar transverse sum H_T distributions are shown for $t\bar{t}$ events.	120
7.3	Total dilepton acceptance versus top mass	124
7.4	Lepton ID data to Monte Carlo scale factor, as a function of jet multiplicities.	126
8.1	Example of a W boson production in association with 3 jets.	131
8.2	Missing energy for events with only one isolated or non-isolated central electron, from inclusive sample jet50.	133
8.3	Missing energy for events with only one isolated or non-isolated muon, from the inclusive sample jet 50.	133
8.4	H_T distributions in which each W+ ≥ 3 jets event is weighted by the probability for it to yield a fake.	138
8.5	Example of a DY production in association with two jets.	139
8.6	\cancel{E}_T distribution for Z loose events in the data.	141
8.7	The fractional jet multiplicity distributions for Pythia Monte Carlo DY events.	142
8.8	Dilepton invariant mass distributions from the Pythia DY Monte Carlo. \cancel{E}_T cut is applied. Left plots are for electrons, right plots are for muons. Top to bottom are for 0, 1, and ≥ 2 jets bins. Also shown is the ratio of events outside/inside the Z window. Overflows are shown in the last bin	143

8.9	Example of a di-tau production Feynman diagram in association with two jets.	146
8.10	Leading order Feynman diagrams for the vector boson pair production. . . .	149
8.11	Feynman diagrams for the tree level processes contributing to $p\bar{p} \rightarrow ZZ$ in SM.	149
8.12	Example of a WW production in association with two jets.	150
9.1	Lepton momenta distribution of the 13 candidates, the expected background and $t\bar{t}$ Pythia Monte Carlo signal for a 175 GeV top mass are shown.	161
9.2	The angle $\Delta\Phi_{ll}$, between the leptons in transverse plane for the 13 candidates	162
9.3	The dilepton invariant mass distribution, for the 13 candidates	162
9.4	Missing transverse energy for the data compared with the SM expectation. The azimuthal angle between the \cancel{E}_T and the nearest lepton or jet versus the \cancel{E}_T , for the 13 candidates and $t\bar{t}$ Pythia Monte Carlo with $m_t = 175$ GeV. . .	163
9.5	The left plot is the jet multiplicity distribution for the 13 candidates, compared with SM expectation, while the right plot shows the scalar transverse energy sum, $H_T(GeV)$	164
9.6	Distribution of the cross-section when varying the background and the acceptance.	168
A.1	Diagram of the Tevatron accelerator complex.	182

Abstract

The top quark was discovered by CDF and D0 in 1995, in Run I. This thesis presents the results of the top production cross-section measurement, in the dilepton channel, using data taken at the Collider Detector at Fermilab (CDF) in $p\bar{p}$ collisions, at a center-of-mass energy of $\sqrt{s}=1.96$ TeV. The dataset represents an integrated luminosity of 193 pb^{-1} and was collected between the period March 2002 - September 2003. Thirteen events were observed (1 e/e , 9 e/μ , 3 μ/μ), passing the selection requirements, with an estimated background of 2.8 ± 0.7 events. These are used to measure a $t\bar{t}$ cross-section of $\sigma_{t\bar{t}} = 8.4^{+3.2}_{-2.7}$ (stat) $^{+1.5}_{-1.1}$ (syst) ± 0.5 (lum)pb. This is in a good agreement with the Standard Model prediction of $\sigma_{t\bar{t}} = 6.7^{+0.71}_{-0.88}$ pb, for a top quark mass of 175 GeV/c^2 . Also few kinematical distributions are compared with the Standard Model and found to agree well.

1 Introduction

The ultimate goal of the high-energy quest is an exhaustive view of the underlying principles that govern our Universe. That is, one longs to identify those constituents of matter which are indivisible and to describe completely all the interactions between them. To understand the progress of our field, one has to go back to the beginnings. Chronologically one could talk of four major periods in the history of physics, including particle physics:

- Pre - 1550 AD: The Ancients

The story of particle physics goes back 2000 years to the Greeks. They gave much to the world of physics by developing the basis of fundamental modern principles such as the conservation of matter and atomic theory. Very few new developments occurred in the centuries following the Greek period. However, as the intense intellectual force of the Renaissance entered the field of physics, Nicolaus Copernicus (1473 - 1543 AD) and other great thinkers began to reject the Greek ideas in favor of new ones based on empirical methods. Copernicus' theories¹ ended the old era of scientific understanding as much as began the new scientific revolution. It was apparent that scientific theories could not be accepted without rigorous testing. Communication among scientists increased and was a stimulus for more discoveries.

¹Nicolaus Copernicus set forth the theory that the earth revolves around the sun, in his book "On the Revolutions of the Heavenly Bodies". *This heliocentric model was revolutionary in that it challenged the previous dogma of scientific authority of Aristotle and Ptolemy, and caused a complete scientific and philosophical upheaval.*

- 1550 - 1900 AD: The Scientific Revolution and Classical Mechanics

During the second period many discoveries were made, but I will mention only a few below. Sir Isaac Newton developed the laws of mechanics which explains object motion in a mathematical framework. James Clerk Maxwell showed that a few relatively simple mathematical equations could express the behavior of electric and magnetic fields and their interrelated nature; that is, an oscillating electric charge produces an electromagnetic field. Joseph Thompson (1897) was investigating a long-standing puzzle known as "cathode rays. He argued that the mysterious rays are streams of particles much smaller than atoms (known later as electrons), and put forth his "plum pudding" model of the atom. He imagined the atom as a slightly positive sphere with small, raisin-like negative electrons inside.

- 1900 - 1964 AD: Relativity and Quantum Revolution

At the start of the twentieth century, scientists believed that they understood the most fundamental principles of nature. Atoms were solid building blocks of nature, Newtonian laws of motion were well accepted and most of the problems of physics seemed to be solved. But in a few years the vision of our world changed. Max Planck suggested in 1905 that radiation is quantized (it comes in discrete amounts). Albert Einstein, one of the few scientists to take Planck's ideas seriously, proposed in 1905 that a quantum of light (the photon) which behaves like a particle. In 1909, Hans Geiger and Ernest Marsden, supervised by Ernest Rutherford, scattered alpha particles off a gold foil and observed large angles of scattering, suggesting that atoms have a small, dense, positively charged core, named "nucleus". Starting with Einstein's theory of relativity which replaced Newtonian mechanics, scientists gradually realized that their knowledge was far from complete. The growing field of quantum mechanics completely altered the fundamental precepts of physics. In 1913, Bohr combined the nuclear model of the atom with the quantum theory of light to put the basis of his famous theory of the hydrogen atom.

In 1915, Einstein formulated the general theory of relativity, one of the most important elements in today's understanding of cosmology.

- 1964 - Present: The Modern View or the Standard Model

By the mid-1960's, physicists realized that their previous understanding, where all matter is composed of the fundamental protons, neutrons, and electron, was insufficient to explain the myriad new particles being discovered. Gell-Mann's and Zweig's quark theory solved these problems. Over the last thirty years, the theory that is now called the Standard Model of particles and interactions has gradually grown and gained increasing acceptance with new evidence from particle accelerators.

In summary, currently we know that matter is composed of chemical elements, organized in the Mendeleev table based on their chemical behavior². Each element is made of indistinguishable atoms. Each atom has a nucleus made of neutrons and positively charged protons, surrounded by a cloud of electrons which make an atom neutral. Protons and neutrons are still not fundamental, and are found to be composed of fundamental³ particles, known as quarks⁴. In fact today there have been identified in nature three generations of two types of fundamental constituents, the leptons,

$$\begin{pmatrix} e \\ \nu_e \end{pmatrix}, \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}, \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix} \quad (1.1)$$

and the quarks,

$$\begin{pmatrix} u \\ d \end{pmatrix}, \begin{pmatrix} c \\ s \end{pmatrix}, \begin{pmatrix} t \\ b \end{pmatrix} \quad (1.2)$$

which make up all the objects of our universe.

The particles interact through forces, which allow the transformation of matter into other matter. At the basic level, the forces transform particles into other particles, or change their properties. There are four known forces in nature: electromagnetic, weak, strong and gravitational force. Today we have a very successful theory of the strong and electroweak forces, known as the Standard Model of Particle Physics, describing the properties and interac-

²Recognizing different patterns, Mendeleev created a table organized by placing elements with similar combining ratios in the same group. He arranged the elements within a group in order of their atomic mass.

³By fundamental one means "having no smaller constituent parts" or "indivisible"; in time, with more knowledge, what is today believed to be fundamental might not be tomorrow.

⁴Only the "up" and "down" quarks enter in the composition of protons and neutrons.

tions of the fundamental particles: leptons and quarks. The force of gravity is negligible compared with the strengths of other forces and becomes important only for interactions between massive bodies separated by large distances.

The heaviest quark, the top quark, was discovered only recently, in 1994 at Fermilab. In the last few years only a limited number of its properties were studied and this thesis checks if the SM prediction of the production rate of top quarks agrees with the observations.

This thesis' goal is to measure the top pair production cross-section in the dilepton decay channel, a process which can be represented as:

$$p\bar{p} \rightarrow t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow \ell\nu_\ell b\ell\nu_\ell\bar{b}, \quad (1.3)$$

where ℓ is either an electron or a muon. Chapter 2 summarizes the good and bad of the Standard Model (SM), and also discusses the importance of the top quark in the SM framework. The cross-section is a simple counting experiment, whose mathematical expression is

$$\sigma_{t\bar{t}} = \frac{N^{obs} - N^{bak}}{\epsilon_{dil} \cdot \mathcal{L}} \quad (1.4)$$

where

- N^{obs} is the number of events passing the selection cuts,
- N^{bak} is the expected number of Standard Model background events,
- ϵ_{dil} is the top dilepton acceptance, and
- \mathcal{L} is the total integrated luminosity of the dataset, $\mathcal{L} = \int L dt$.

The CDF detector, used to reconstruct and measure the energy or momentum of the particles resulting from a proton-antiproton collision, is described in Chapter 3. This includes the system used to select the few interesting events out of the many uninteresting collisions, known as the trigger system. The details of how various event pieces get reconstructed are unveiled in Chapter 4. Chapter 5 describes the data samples used for this analysis and the

total integrated luminosities of each dataset. The criteria used to select the top dilepton events, while trying to eliminate the backgrounds, are described in Chapter 6. The top dilepton acceptance (ϵ_{dil}), which is the fraction of top dilepton events left after the selection requirements, is determined in Chapter 6. The contribution to the expected number of dilepton events from backgrounds (non- $t\bar{t}$ processes) (N^{bak}) is presented in Chapter 8. The observed number of events passing the selection criteria (N^{obs}) is presented in Chapter 9; also, all the pieces are used to calculate the dilepton cross-section and the shapes of few kinematic distributions are compared with the expectation from Standard Model.

2 Standard Model Theory and The Top Quark Physics

The Standard Model (SM) is a theoretical framework, built from observations and gauge symmetry principles, describing the fundamental particles and their interactions. It has been amazingly successful in predicting a wide range of phenomena ¹. Its success strongly suggests that the SM will remain an excellent approximation to nature down to distances as small as 10^{-18} m (energy scales $Q \sim 100$ GeV). There is just one missing part, the conclusive test of the Electroweak Symmetry Breaking Mechanism (EWSB). However, there are strong reasons to believe that the SM is a *low energy effective approximation* of a more fundamental theory at higher energies. The goal of our field, high-energy physics, is to reveal this underlying theory.

2.1 The Standard Model

The Standard Model is a quantum field theory, described by the following Lagrangian ([53],[97]):

¹The Particle Data Book[87] is a testimony for its outstanding agreement with experimental data, summarizing the results of high-energy experiments over the last 5 decades.

$$\begin{aligned}
\mathcal{L}_{SM} &= -\frac{1}{4}W_i^{\mu\nu}W_{\mu\nu}^i - \frac{1}{4}B^{\mu\nu}B_{\mu\nu} - \frac{1}{4}G_a^{\mu\nu}G_{\mu\nu}^a \\
&\quad + i\bar{\psi}_L\gamma^\mu(\partial_\mu + igW_i^\mu T^i + ig'\frac{1}{2}B_\mu Y)\psi_L \\
&\quad + i\bar{\psi}_R\gamma^\mu(\partial_\mu + ig'\frac{1}{2}B_\mu Y)\psi_R \\
&\quad - g''(\bar{q}\gamma^\mu T_a q)G_\mu^a
\end{aligned} \tag{2.1}$$

$$\begin{aligned}
&\quad + i\frac{1}{2}\bar{\phi}(\partial_\mu + igW_i^\mu T^i + ig'\frac{1}{2}B_\mu Y)\phi - V(\phi) \\
&\quad + (G_1\psi_L\phi\psi_R + G_2\psi_L\phi_c\psi_R + h.c.)
\end{aligned} \tag{2.2}$$

where

$$V(\phi) = \mu^2\bar{\phi}\phi + \lambda|\phi|^4 \tag{2.3}$$

is the Higgs potential, with $\mu^2 < 0$. Left-handed and right-handed fields are defined as

$$\psi_L = \frac{1}{2}(1 - \gamma_5)\psi, \quad \psi_R = \frac{1}{2}(1 + \gamma_5)\psi. \tag{2.4}$$

Right-handed fermions transform under $U(1)_Y$; no right-handed neutrinos are introduced.

The gauge fields enter naturally into the theory by requiring local gauge invariance. The gauge groups are $SU(2)_L \times U(1)_Y$ for the electroweak interaction and $SU(3)_c$ for the strong interaction. The first line in the Lagrangian expression consists of the kinetic energy terms of the electroweak and strong gauge fields. The electroweak couplings of the gauge fields to the left-handed and right-handed components of the matter fields are shown on the second and third line. The next line is the strong coupling of the quarks to gluons. In the fifth, the Higgs boson field kinetic and potential energy terms are shown, while the terms in the last line generate the fermion masses. One element of the SM, the *Higgs mechanism*, provides a simple and elegant solution to the problem of electroweak and flavor symmetry breaking, generating masses for leptons, quarks, and massive gauge bosons (W^\pm , Z). However, the

Higgs particle is yet to be observed and its mass is sensitive to the scale of any new physics beyond the SM.

The minimal version of the standard model, with no right-handed neutrinos and the simplest possible electroweak symmetry-breaking mechanism has 19 arbitrary parameters: nine fermion masses (six quarks and three charged leptons), three angles and one phase specifying the CKM quark mixing matrix, three gauge coupling constants (α_{EM} , G_F and α_s), two parameters to specify the Higgs potential (v and m_H), and an additional phase θ that characterizes the QCD vacuum state. Among other parameters which could be included in the above list are the neutrinos' masses², the elements of a leptonic mixing matrix (analogue to the CKM matrix), or a possible strong CP violating phase (experimentally determined to be less than 10^{-9} , if existing).

However, even knowing all the arbitrary parameters, to test the SM predictions one *needs to measure a few more numbers*, such as parton distribution functions, particle fragmentation functions, and various properties of hadrons, in order to predict cross sections for real particles. In principle they could be calculable from the Standard Model Lagrangian (Eq. 2.1), but practically it is not yet possible. The challenge comes from Quantum Chromodynamics (QCD), the theory of the strong interactions, that explains quark confinement or asymptotic freedom. In fact, due to asymptotic freedom, perturbation theory works only at high energies. At sufficiently low energies the coupling constant is large, the non-perturbative effects become increasingly important and the calculations become intractable. The best hope we have is *lattice QCD*[84], an approach which relies on discretization of space-time and uses powerful computational tools to calculate parameters like masses of bound states of bottom and charm quarks, or to try to understand the behavior of matter at ultra high temperature and/or density.

Standard Model predictions have been very successful at accounting for the overwhelming amount of experimental data, but there are still many unsolved issues. To give the reader a flavor, I will mention few of the outstanding questions in our field today:

- What are dark matter and dark energy made of?

²There is evidence ([63],[81]) that the neutrinos may in fact have mass. This was not accounted for in the minimal version of the SM.

- How can neutrino oscillations be accommodated in the current standard model picture? What physics is responsible for neutrino masses and mixing? Do neutrino interactions violate CP?
- What is the origin of CP violation in the SM? Do CP violating terms play any role in an electroweak mechanism for baryogenesis?
- What is the cause of electroweak symmetry breaking (EWSB)? Is it the Higgs boson or is it something else?

The top quark could play an important role in the understanding of the generation of mass in the SM and in this way helps to answer the last question. This thesis will try to determine if the top quark production rate is in agreement with the standard model predictions, as new physics could affect it in many ways. It is worth mentioning that in Run I a few of the dilepton events looked only $\sim 1\%$ consistent with the SM description[6].

Electroweak Symmetry Breaking in the Standard Model

In the 1970's Glashow, Weinberg, and Salam unified the electromagnetic theory with the weak interaction model of Fermi into the *electroweak theory*. In fact, Glashow(1961) was the first to come up with the idea that the electromagnetic and weak interactions may be unified in a gauge theory, based on the $SU(2)_L \times U(1)_Y$ symmetry. The problem of how to generate masses and preserve the gauge invariance was solved later by Weinberg(1967), and Salam(1968), using the idea of *spontaneous symmetry breaking*. The resulting theory, referred to as the new *Weinberg-Salam model*³, was shown to be a renormalizable quantum field theory by t'Hooft(1971) ([52],[51]). This unification predicts the existence of three new particles, the W^\pm and Z bosons, discovered at the Sp \bar{p} S ([49],[50]) collider by the UA1 and UA2 experiments in 1983.

The requirement that the Lagrangian remain invariant under $SU(2)_L$ local transformations did not allow one to just add mass terms for the W and Z bosons, so a new mechanism had to be invented. In quantum field theory, the simplest way was to introduce a new scalar field ϕ , known as the *Higgs field*. Thus, the gauge symmetry is spontaneously broken; that

³The quarks were included into the theory by Glashow, Iliopoulos and Maiani (1970)[88].

is, the Lagrangian is still gauge invariant, but the vacuum state and spectrum of particles are not. The Higgs sector of the SM Lagrangian is the fifth term of Eq. 2.1. The potential energy is minimized at a value $\langle \phi \rangle \neq 0$ and has its minimum at $v = \mu/\sqrt{\lambda} = \sqrt{2} \langle |\phi| \rangle$. The two free parameters, μ and λ , are usually traded for the Higgs vacuum expectation value (vev) v and the physical Higgs mass (m_H) obtained after the renormalization analysis. The vacuum expectation value is known to be $v = 246$ GeV, but the Higgs mass is still unknown⁴.

The importance of the 1 TeV scale

In spite of its success, the current paradigm of particle physics leaves many unanswered questions and there are good reasons to believe that a thorough study of 1-TeV physics will yield some answers. The most serious structural problem of the SM is associated with the electroweak theory, and I will argue why understanding the Higgs sector might be the key toward new physics beyond the SM.

Let's consider for example the scattering of the W^+W^- bosons. At tree level, if only the diagrams shown in Figures 2.1(a)-(c) contribute, the scattering violates unitarity at the 1-1.5 TeV scale, unless the Higgs boson or other new physics exists to quench the high-energy behavior (Fig. 2.1(d)-(e)). Therefore, at about 1 TeV there must be either a Higgs-like boson or new physics, otherwise the Standard Model cannot make any predictions.

⁴The current best limit is $m_H > 114.4$ GeV, at 95% confidence limit (CL), established by direct searches, through the process $e^+e^- \rightarrow HZ$, by the LEP experiments[82]. The Higgs mass determined from the fit to the known data is 117^{+67}_{-45} GeV, and $m_H < 251$ GeV at 95% CL.

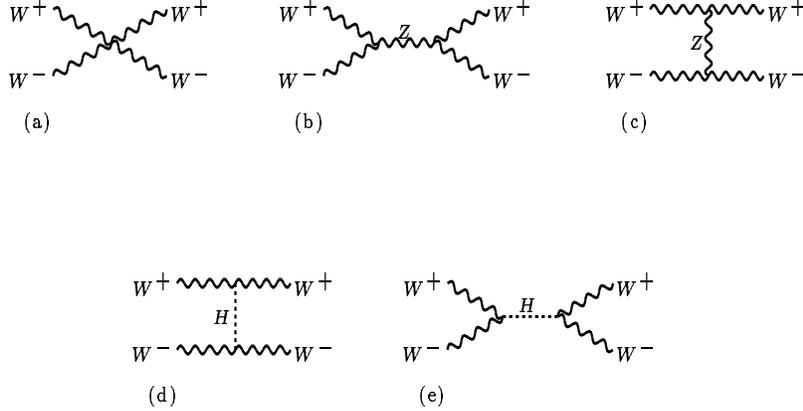


Figure 2.1: WW scattering Standard Model Feynman diagrams. If there is no Higgs boson, then (d) and (e) do not exist. In such a case, the WW scattering cross-section, with diagrams (a)-(c) contributing at tree level, grows without limits at large energies, violating the *unitarity*.

Also, even if the Higgs boson exists, the model will still suffer from what is known as the *hierarchy problem*. This refers to the fact that the Higgs mass is sensitive to the scale of any new physics beyond the Standard Model. To understand this problem, think of a fermion with Yukawa interaction $\lambda_f H \bar{f} f$. The 1-loop correction (Fig. 2.2) to the Higgs mass squared is quadratically divergent:

$$[H] \Delta m_H^2 = \frac{|\lambda_f|^2}{16\pi^2} [-2\Lambda_{UV}^2 + 6m_f^2 \ln(\frac{\Lambda_{UV}}{m_f}) + \dots] \quad (2.5)$$

where the cut-off Λ_{UV} could be as large as the Planck scale, $M_{Pl} = \frac{1}{\sqrt{(8\pi G_{Newton})}} = 2.4 \times 10^{18} \text{ GeV}$. However the physical Higgs mass is constrained to be close to the electroweak scale, $M_{EW} \approx 246 \text{ GeV}$. Thus, one needs an enormous cancellation of up to 32 orders of magnitude for the mass-squared between the 1-loop and tree-level to preserve a physical Higgs mass of about 1 TeV in the presence of new physics at Planck scale. Such *fine-tuning* is highly unnatural and one natural solution is to require new physics at the electroweak scale.

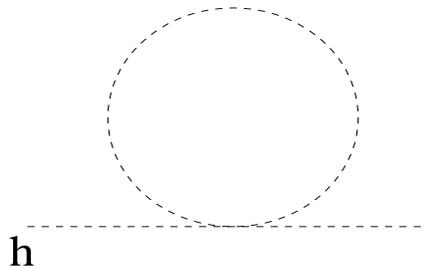


Figure 2.2: The 1-loop correction diagram to the Higgs mass.

There are three possible ways to solve this problem: to assume a new symmetry which will allow the cancellation of the quadratic divergences leaving only renormalizable divergences (for example, supersymmetry), to assume that the Higgs is not an elementary particle but rather a composite object (for example, a new strong force exists, such as in technicolor models) leaving no scalar field in the theory to deal with, or to argue that the integral limit is about 1 TeV rather than the Planck scale (for example, if large extra dimensions exist).

2.2 Top Quark Physics

The first particle of the third generation to be discovered was the tau lepton(τ), in 1975 [35]. Shortly after, in 1977, the Υ was discovered at Fermilab [40], [41], as a resonance in the $\mu^+\mu^-$ invariant mass spectrum. This was interpreted as a $b\bar{b}$ bound state. In the past twenty years, a plethora of experimental data[19] has established that bottom quarks have charge $Q_b=-1/3$ and $T_3=-1/2$. Therefore, searches were initiated for a partner to complete the third family of quarks in the Standard Model. These searches ended in 1994, when the CDF Collaboration of Fermilab announced the first evidence of top production [31], which was followed soon after in 1995 by the announcement of the top quark discovery by both the CDF and DØ Collaborations [32], [30].

2.3 Top quark production and decay

This section introduces and explains the top production and decay mechanisms at a hadron collider. The section ends with the theoretical calculation progress and importance of a cross section measurement.

In Standard Model, the top quark is the $Q=+2/3$, $T_3=+1/2$ member of the weak-isospin doublet containing the bottom quark. The SM structure of the top quark interactions is very rich, ranging from the color structure in the top-gluon interaction, $g_s[T_a]^{ij}\bar{t}_i\gamma^\mu t_j A_\mu^i$, the chirality structure, parity violation and flavor mixing in the top-W boson coupling $(g/\sqrt{2})V_{tq}(\bar{t}_L\gamma^\mu q_L)W_\mu^+$, the parity violation in the top-Z coupling $(g/4\cos\theta_W)(\bar{t}[1-(8/3)\sin^2\theta_W]\gamma^\mu - \gamma^\mu\gamma^5)Z_\mu q$, where $q = u$ or c , to the large strength, but simple structure, found in the top-Higgs Yukawa interaction $y_t h t \bar{t}$.

2.3.1 Strong and weak production channel

The top quark was observed in 1995 in the *strong production* mechanism, where top is produced in pairs, t and \bar{t} , via quark-antiquark annihilation and gluon-gluon fusion. The leading order (LO) diagrams are shown in Figure 2.3, and a detailed discussion in the context of the cross-section calculation will come in Section 2.4.

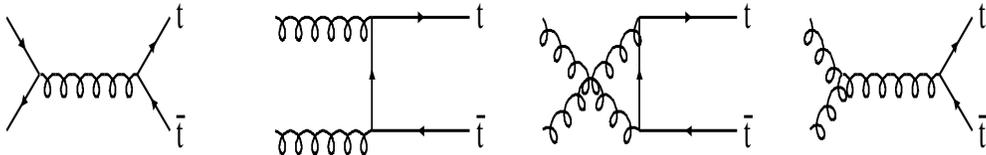


Figure 2.3: Lowest order (α_s^2) Feynman diagrams for heavy quark production. The first diagram represents the $q\bar{q}$ annihilation, while the remaining three represent the gg fusion.

However, the Standard Model allows for a second way to produce top, *singly*, along

with a b quark, through the *electroweak interaction*, mechanism referred to as *single top* production. There are three channels of single top production:

- *s-channel production*, $q\bar{q}' \rightarrow t\bar{b}$, usually referred to as W^* production,
- *t-channel production*, $qg \rightarrow t\bar{b}q'$, known as W -gluon fusion production,
- *associated production*, with a W boson $bg \rightarrow tW$, known as the Wt production.

The Wt process proceeds via gluon- b quark interaction, which makes the cross section negligible at the Tevatron (about 7% of the single top inclusive cross-section).⁵

A few Feynman diagrams for the W^* and W -gluon channels are shown in Figure 2.4.

⁵However at the LHC it will account for about 30% of the single top cross-section.

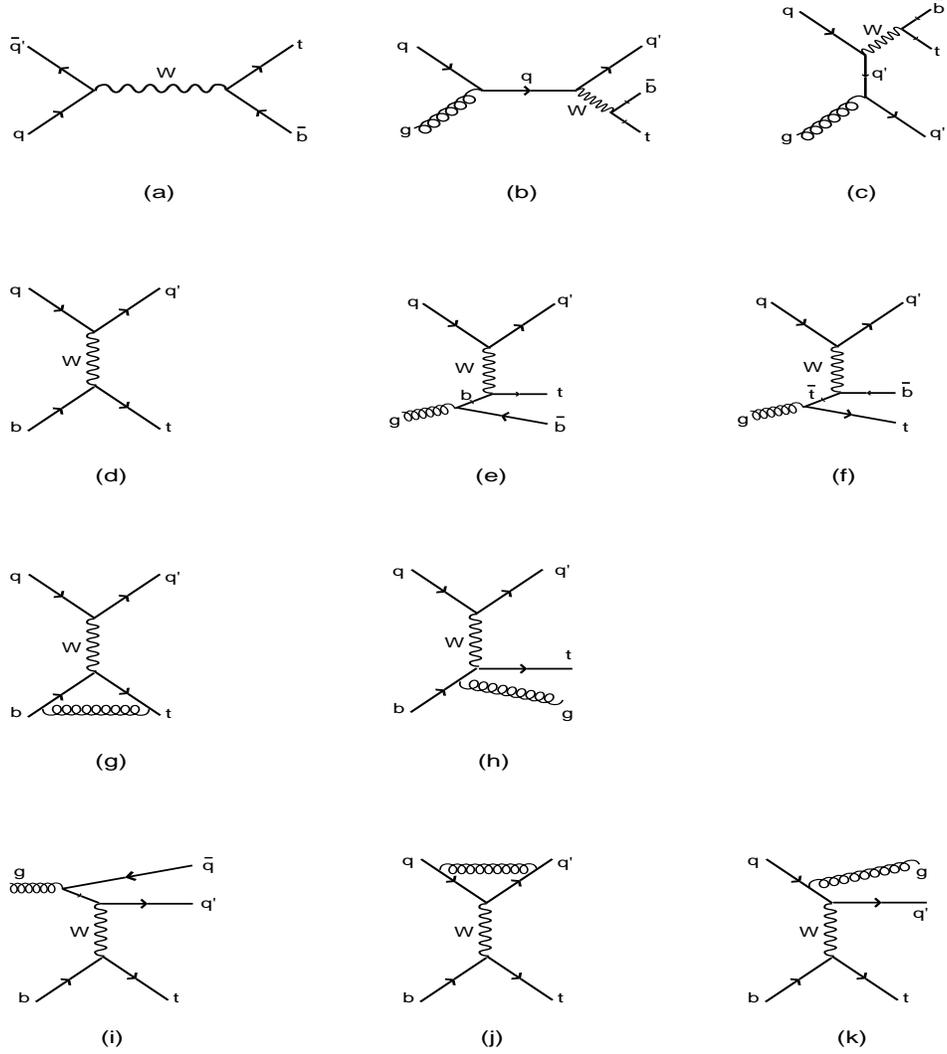


Figure 2.4: Single top production at the Tevatron: (a) *LO* diagram for W^* single top, (b)-(c) *NLO* corrections to W^* , (d) *LO* W -gluon single top, (e)-(f) *NLO* corrections of order $1/\ln(m_t^2/m_b^2)$, (g)-(k) *NLO* corrections of order α_s . Note that for the W^* channel (a)-(c) the W boson has $q^2 > (m_t + m_b)^2$, while for the W -gluon channel (d)-(k) the W has $q^2 < 0$.

The next-to-leading (NLO) cross-sections for the two main processes, at the center-of-mass energy $\sqrt{s} = 1.96$ TeV, are [15]:

$$\sigma_{NLO}(p\bar{p} \rightarrow t\bar{b} + X) = 0.884 \pm 0.004(stat) \pm 0.050(NLO \text{ scale}) \text{ pb}, \quad (2.6)$$

$$\sigma_{NLO}(p\bar{p} \rightarrow qt\bar{b} + X) = 1.980 \pm 0.004(stat) \pm 0.113(NLO \text{ scale}) \text{ pb}, \quad (2.7)$$

Even if the combined cross-section for single top (about 2.9 pb) is only twice smaller than $t\bar{t}$

cross-section, single top production has not been observed, because the final state signature is hidden under large backgrounds and, it also has lower detection efficiency. The current best limits are $\sigma_{s+t \text{ combined}} < 13.7$ pb at 95% CL and $\sigma_t < 8.5$ pb at 95% CL [17], both obtained using 162 pb^{-1} of Run II data.

2.3.2 Top quark decay and final state signatures

In the Standard Model, the top quark, as all the other fermions, has a V-A charged-current weak interaction. Assuming that there are just three generations of quarks, the CKM mixing parameter V_{tb} is close to unity.⁶ Therefore the top quark decays into a real W boson and a b quark almost all the time. Ignoring the b quark mass, at LO, the top quark (bW) partial width [55] is:

$$\Gamma(t \rightarrow bW^+) = \frac{G_F}{8\pi\sqrt{2}} m_t^3 |V_{tb}|^2 \left(1 - 3\frac{m_W^4}{m_t^4} + 2\frac{m_W^6}{m_t^6}\right) \quad (2.8)$$

Assuming that $|V_{tb}| = 1$ and $m_t, m_W \gg m_b$, one gets for the partial width:

$$\Gamma(t \rightarrow bW) \approx 175 \text{ MeV} \left(\frac{m_t}{m_W}\right)^3 \quad (2.9)$$

Using the known masses of W, b and top quarks, the top decay rate and the top lifetime are:

$$\Gamma(t \rightarrow bW) \approx 1.56 \text{ GeV} \quad (2.10)$$

$$\tau_{top} = 1/\Gamma \approx 4 \times 10^{-25} \text{ s}. \quad (2.11)$$

The NLO calculation of the partial width is 1.42 GeV [71]. Just as the (bW) partial width is proportional to $|V_{tb}|^2$, the widths for the decays $t \rightarrow sW$, dW , are proportional to $|V_{ts}|^2$ and $|V_{td}|^2$, and are a correction of about 0.2 % to the total top decay width $\Gamma = \sum_q V_{tq}$.

⁶ Using the unitarity constraint and the very precised measured values for $|V_{ub}|$ and $|V_{cb}|$, one gets $0.9991 < |V_{tb}| < 0.9994$.

Given that the hadronization time scale is $\tau_{had} = 1/(\Lambda_{QCD}) \approx 10^{-23} \text{ s}$ ⁷, this implies that the top quark is produced and decays as a *free* quark. Thus, the theoretical calculations are free of the complications due to chromodynamic forces. There are no top mesons, hadrons or toponium spectroscopy. Another benefit of being a free quark is that the top decay products carry information about its original spin and allows us to study $t\bar{t}$ spin correlations.

As top decays almost exclusively to a W and b, the final states can be classified based on how the W decays. Explicitly, W could decay:

- leptonically: $W^+ \rightarrow \ell^+ \nu_\ell$, $W^- \rightarrow \ell^- \bar{\nu}_\ell$, where $\ell = e, \mu$ or τ , or
- hadronically: $W^+ \rightarrow q\bar{q}'$, where $q\bar{q}' = u\bar{d}', c\bar{s}'$; $W^- \rightarrow \bar{q}q'$, where $\bar{q}q' = \bar{u}d', \bar{c}s'$.

Based on this, there are three final state categories:

1. dilepton, with *both* W bosons decaying leptonically,
2. lepton+jets, with *one* W boson decaying leptonically and the *other* hadronically, and
3. all-hadronic, with *both* W bosons decaying hadronically .

Table 2.1 summarizes all the decay modes and branching ratios.

W^+	W^-				
	$e^- \bar{\nu}_e$	$\mu^- \bar{\nu}_\mu$	$\tau^- \bar{\nu}_\tau$	$\bar{u}d$	$\bar{c}s$
$e^+ \nu_e$	1/81	1/81	1/81	3/81	3/81
$\mu^+ \nu_\mu$	1/81	1/81	1/81	3/81	3/81
$\tau^+ \nu_\tau$	1/81	1/81	1/81	3/81	3/81
$u\bar{d}$	3/81	3/81	3/81	9/81	9/81
$c\bar{s}$	3/81	3/81	3/81	9/81	9/81

Table 2.1: The $t\bar{t}$ decay channels and their corresponding branching ratios.

The *all-hadronic channel* has the largest branching ratio and 6 jets in the final state. However, it suffers from a large QCD background, therefore the capability to identify one or both jets coming from b quarks is necessary to extract the $t\bar{t}$ signal.

⁷ $\Lambda_{QCD} = 200 \text{ MeV}$

The *lepton+jets channel* was the best choice in Run I, because it has a large branching ratio, a high- P_T lepton and missing energy, which allows one to separate the signal from backgrounds more efficiently. Also, in the final state there are 4 jets, two of which are b jets and usually at least one is required to be b tagged.

Dilepton channel signature

The dilepton channel has the *smallest branching ratio* of all, about 5%, considering the final states with leptonic decays of Ws to e and μ only.⁸ However, its signature allows for a *very good separation between signal and backgrounds*, of about 3 to 1. The final state consists of two high- P_T leptons, large missing energy, \cancel{E}_T , from the undetected neutrinos, and at least two jets from the fragmentation and hadronization of the b quarks. In this thesis, the selection criteria for this channel are discussed in Chapter 6, while the signal contribution and the backgrounds estimates are presented in Chapters 7 and 8, respectively.

2.4 Top Pair Production Cross-Section

The high-energy interactions of hadrons, with a momentum transfer Q^2 , larger than the QCD energy scale, Λ_{QCD}^2 , are described by the *QCD improved parton model*. The heavy top quark production involves a large Q^2 , and therefore a small α_s , so a perturbative expansion in the coupling constant is valid. The parton model, originally envisioned by Feynman[9], provides a physics picture of a hard scattering event in a frame in which the hadron is rapidly moving. In such a frame the hard interaction occurs on a *time scale short compared to the scale which controls the evolution of the parton system*. During the hard scattering, the partons, which make up the incoming hadrons, can be treated as if they were effectively free. The incoming hadrons provide broad-band beams of partons, which carry a fraction of the parent hadrons momenta. The total inclusive heavy quarks production cross-section, for a hard scattering with characteristic momentum scale Q^2 , is expressed as:

⁸In practice, the leptonic decays of τ 's do not contribute much to the $t\bar{t}$ dilepton signal, because the p_T spectrum of the leptons from τ 's is much softer than the one from direct leptonic decays of W bosons to e or μ .

$$\sigma(P_1, P_2) = \sum_{i,j} \int_0^1 dx_1 \int_0^1 dx_2 f_i(x_1, \mu^2) f_j(x_2, \mu^2) \hat{\sigma}_{ij}(\hat{s}, Q^2/\mu^2, \alpha_s(\mu^2)) \quad (2.12)$$

where the sum is over the initial partons, and the integrations are over the two interacting parton momentum fractions. The quantities appearing in the equation above represent the following:

- $f_i(x_1, \mu^2) dx_1$ is the probability that the parton i carries a fraction between x_1 and $x_1 + dx_1$, of the incident hadron (p or \bar{p} at the Tevatron) momentum, P_1 . $f_i(x_1, \mu^2)$ is called *parton distribution function (PDF)*, and is determined experimentally, mainly from deep inelastic scattering experiments;
- $\sqrt{\hat{s}}$ is the center-of-mass energy of the (i, j) parton system; it is related to the $p\bar{p}$ center-of-mass, \sqrt{s} , by the formula $\hat{s} = x_1 x_2 s$;
- Q^2 is the characteristic momentum scale of the hard scattering. For $t\bar{t}$ pair production, $Q^2 \approx m_t^2$;
- μ^2 is the renormalization and factorization scale; in fact there are *two scales*, μ_F and μ_R , but a common choice is to assume them equal, $\mu_F = \mu_R = \mu$;
- α_s is the strong running coupling constant;
- $\hat{\sigma}_{ij}$ is the short-distance cross-section for the scattering of partons of types i and j , to produce a heavy quark pair. Since the coupling is small at high Q^2 , $\hat{\sigma}_{ij}$ can be calculated in the *perturbation series expansion*, in the powers of the running coupling, α_s .

The impulse approximation is used in equation 2.12. Interference terms involving more than one active parton per hadron are not included. These terms are suppressed by the powers of the large scale, Q , but they will become important for the center-of-mass energy of the LHC. The *factorization procedure* is used to remove the long distance pieces from

the hard scattering cross section, by placing them in the PDFs. The short distance cross-section contains only the physics of the hard scattering. To lowest order in α_s , the Feynman diagrams for $t\bar{t}$ production are shown in Figure 2.3.

The $t\bar{t}$ production cross-section depends on two complementary ingredients, as can be seen from Eq. 2.12, *the parton-level cross-section* and *the proton parton densities*. To understand which production mechanism dominates, $q\bar{q}$ vs gg , both factors have to be weighted. The partonic cross-section can be expanded in the α_s powers as

$$\hat{\sigma}_{ij}(\hat{s}, Q^2/\mu^2, \alpha_s(\mu^2)) = \frac{\alpha_s^2(\mu^2)}{m_t^2} f_{ij}(\rho, \frac{\mu^2}{m_t^2}) \quad (2.13)$$

with

$$f_{ij}(\rho, \frac{\mu^2}{m_t^2}) = f_{ij}^0(\rho) + O(\alpha_s(\mu^2)) \quad (2.14)$$

where $\rho = 4 m_t^2/\hat{s}$. The zeroth order term, corresponding to the diagrams from Figure 2.3, yields [80]:

- for $q\bar{q}$ annihilation:

$$\hat{\sigma}_{q\bar{q}} = \frac{\alpha_s^2}{m_t^2} \frac{\pi\beta\rho}{27} [2 + \rho] \quad (2.15)$$

- for gg fusion:

$$\hat{\sigma}_{gg} = \frac{\alpha_s^2}{m_t^2} \frac{\pi\beta\rho}{192} \left[\frac{1}{\beta} (\rho^2 + 16\rho + 16) \ln\left(\frac{1+\beta}{1-\beta}\right) - 28 - 31\rho \right] \quad (2.16)$$

where

$$\beta = \sqrt{1 - \rho} \quad (2.17)$$

For the production close to the threshold (top is produced almost at rest), $\hat{s} = 4 m_t^2 + \epsilon$, where ϵ is small, and the partonic cross-sections take a simple form:

$$\hat{\sigma}_{q\bar{q}} = \frac{4\pi\alpha_s^2}{9\hat{s}} \beta \quad (2.18)$$

$$\hat{\sigma}_{gg} = \frac{59\pi\alpha_s^2}{48\hat{s}}\beta \quad (2.19)$$

Then

$$\hat{\sigma}_{q\bar{q}} : \hat{\sigma}_{gg} \approx 1 : 2.7. \quad (2.20)$$

At the Tevatron, the $t\bar{t}$ production occurs close to the threshold. For the threshold, let's set $x_i \approx x_j = x_{thr}$, then, from the formula $\hat{s} = x_1x_2s$, we obtain $x_{thr} \approx \frac{2m_t}{\sqrt{s}}$. In Run II, the center of mass energy is $\sqrt{s} = 1.96$ TeV, so $x_{thr} \approx 0.18$. At such large x , there are significantly more quarks than gluons inside the hadrons, to hard scatter (Figure 2.5). Thus, even if for the partonic cross-section, gg reigns over $q\bar{q}$, *the parton densities are the dominant effect*. At the Tevatron, the $t\bar{t}$ pairs are produced from $q\bar{q}$ annihilation 85% of the time, and 15% from gluon fusion. At the LHC, the situation is reversed, $x_{thr} \approx 0.025$, a regime where the gluons dominate, so $\sigma_{q\bar{q}} : \sigma_{gg} \approx 1:5.7$.

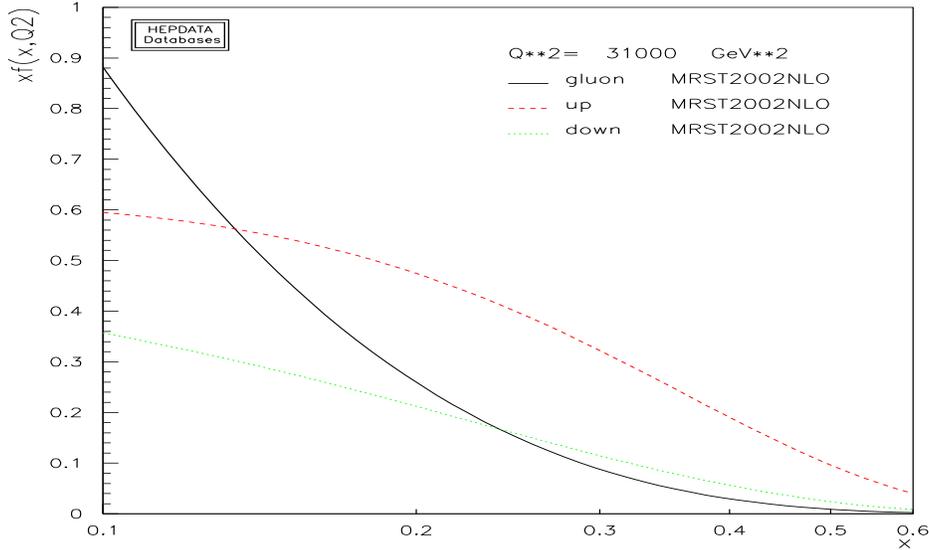


Figure 2.5: The proton parton distribution functions for gluon, compared with the up and down quarks.

The uncertainty in $\sigma_{t\bar{t}}^{LO}$ at the leading order Born approximation is $\sim 50\%$, and is mainly due to the renormalization and factorization scale choices, μ_f and μ_r , and their effects on α_s . Moreover, α_s is relatively large, therefore the next-to-leading order (NLO) diagrams,

of order $O(\alpha_s^3)$ were found to be important, being of the order of 30% [100; 11]. Some representative graphs for next to leading order contributions to $q\bar{q}$ annihilation are shown in Figure 2.6. Similar corrections exist for the gg fusion channel. At NLO, the corrections due to quark-gluon fusion processes, with a gluon splitting to $t\bar{t}$, contribute about 1% to the cross section. The largest contribution to NLO comes from *initial state gluon bremsstrahlung* diagrams. Electroweak 1-loop contributions to the lowest order matrix element, due to couplings to W, Z, γ or H, are found to be less than 2%, for $60 < m_H < 1000$ GeV [99]. Large logarithms, proportional with $(\alpha_s \ln^2 \beta)$ arise from *real emission of soft gluons*, where β measures the distance from the $t\bar{t}$ threshold production. As β approaches 0, the calculation becomes unstable. Terms, such as, $(1/n!)(\alpha_s \ln^2 \beta)^n$, appear to all orders in perturbation theory, resulting in a series which sums up to an exponential. The summation of such terms is referred to as *soft gluon resummation* and is part of the improved cross-section of heavy-quarks production. A leading-log(LL) resummation handles $(\alpha_s \ln^2 \beta)^n$ series, a next-to-leading-log(NLL) resums the $(\alpha_s(\alpha_s \ln^2 \beta))^n$ terms, and so on.

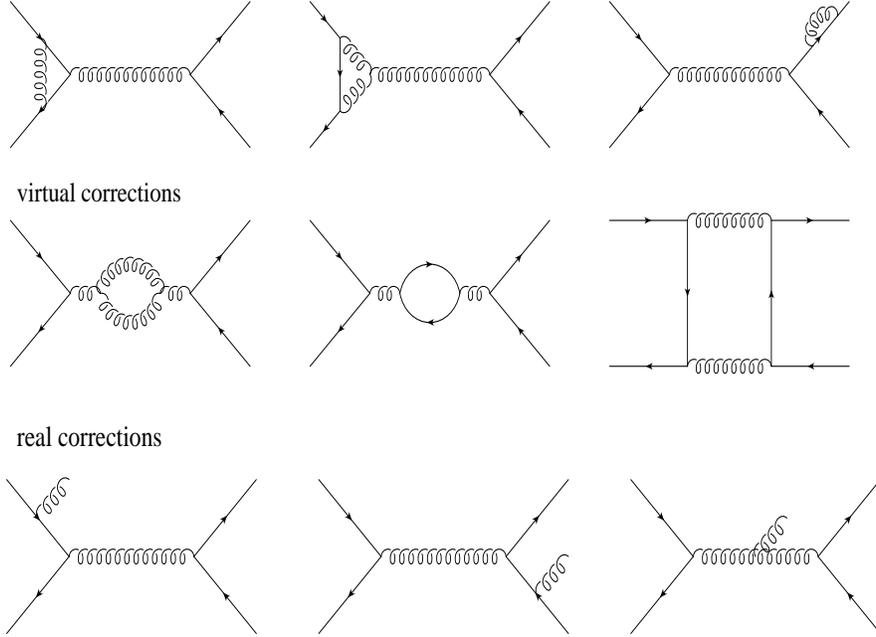


Figure 2.6: Some representative Feynman diagrams contributing to next to leading order (NLO) corrections $O(\alpha_s^3)$ to the process $q\bar{q} \rightarrow t\bar{t}$. The real corrections consist of gluon bremsstrahlung diagrams.

Beyond NLO, two main calculation techniques have been used. The first approach ([7],[5]) uses a double differential NNLO estimate and the cross-section found for a top mass of 175 GeV is

$$\sigma_{t\bar{t}} = 7.4 \pm 0.5 \pm 0.1 \text{ pb}, \quad (2.21)$$

at a $\sqrt{s} = 1.96$ TeV. The first uncertainty is due to an ambiguity in kinematics, while the second one is an estimate of the scale uncertainty⁹. A second technique uses a full next-to-leading soft log (NLL) resummation ([8],[68]), to compute the inclusive cross-section. For a top mass of 175 GeV, the calculation yields:

$$\sigma_{t\bar{t}} = 6.70_{-0.88}^{+0.71} \text{ pb}, \quad (2.22)$$

⁹Possible uncertainties due to PDFs are not included.

at a $\sqrt{s} = 1.96$ TeV. Table 2.2 shows the cross-section for various top masses. All theoretical uncertainties, including the latest PDFs¹⁰, are counted and it is the most complete theoretical estimate to date. The main sources of uncertainties are the PDFs and α_s . The leading source of PDF uncertainty comes from the lack of knowledge of the gluon luminosity at large values of x . In fact, the gg can vary up to a factor of 2, when varying the PDFs by $\pm 1\sigma$ (from 10% to 20% for $\sqrt{s} = 1.96$ TeV). The scale uncertainty¹¹ is about 5%, by inclusion of NLL resummation corrections as compared with 10%, when only NLO terms are used.

$\sqrt{s}(\text{GeV})$	$m_{top} (\text{GeV}/c^2)$	$\sigma_{min}(\text{pb})$	$\sigma_{ref}(6M)(\text{pb})$	$\sigma_{max}(\text{pb})$
1960	170	6.79	7.83	8.69
1960	175	5.82	6.70	7.41
1960	180	5.00	5.75	6.34

Table 2.2: $t\bar{t}$ cross-section at the Tevatron, at $\sqrt{s} = 1.96$ TeV, for different top masses. The central values are taken from CTEQ6M. The lower bound is determined using CTEQ6, while the upper bound arises from MRST set **J01** (extracted from a fit to the Tevatron jet data, $\alpha_s(M_Z) = 0.121$)[68]. All uncertainties (due to parton densities, α_s , scale dependence μ_f, μ_r) are included.

¹⁰The parton distribution functions are typically extracted comparing the existing data with NLO calculations for the relevant processes and extrapolated to the relevant Q^2 region using NLO evolution equations.

¹¹The scale uncertainty is a purely theoretical error due to the dependence of the cross-section on the renormalization (μ_R) and factorization (μ_F) scales used in the perturbative calculation.

3 The Collider Detector at Fermilab

The Collider Detector at Fermilab (CDF) is a multipurpose experiment for the study of the $p\bar{p}$ collisions at the center of mass energy $\sqrt{s} = 1.96$ TeV at the Fermilab Tevatron Collider, near Chicago, Illinois. The goal of this thesis is to *measure the $t\bar{t}$ production cross-section*, which requires that we are able to detect leptons¹, to precisely measure the missing energy due to undetected neutrinos and reconstruct the hard-scattered partons that, after fragmentation and hadronization, give rise to roughly collinear sprays of hadrons known as *jets*.

Historically the *first collisions* at CDF were produced in October of 1985 and the consecutive upgrades improved significantly the performance of both the accelerator² and the detector.

The current incarnation of CDF, shown schematically in Figure 3.1, is a 5000-ton cylindrical detector with a central barrel region, two end-cap (plug) regions enclosing the barrel, and two very-forward electromagnetic calorimeters located beyond the end-caps. It consists of silicon tracking detectors near beampipe, charged particle tracking chambers surrounding them, electromagnetic (EM) and hadronic (HAD) shower counters arranged in projective geometry, and drift chambers outside the hadron calorimeters, for muon detection. The

¹By leptons I will refer to muons or electrons in the final state, unless explicitly stated otherwise. Tau leptons decay quickly inside the detector, and only their products are detected. This analysis is only sensitive to leptonic τ decays.

²The accelerator complex will be briefly described in Appendix A.

tracking volume is immersed in a 1.4 T magnetic field, oriented in the *opposite sense* as the proton motion and generated by a 3 m diameter, 5 m long superconducting magnet coil. A detailed description of the CDF detector is available in the Technical Design Report (TDR)[39].

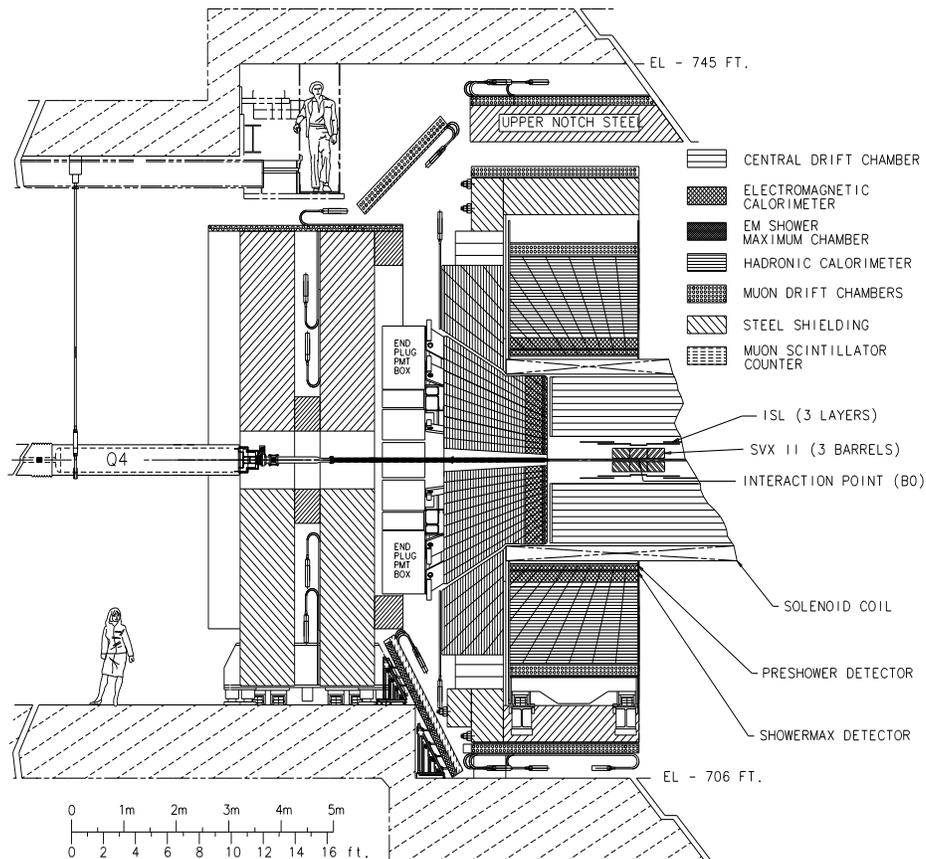


Figure 3.1: Longitudinal view of half of the Run II CDF detector.

3.1 The 1992-1995 Run I and 2000-2003 Run II

The knowledge we have about the top quark came from the data collected in Run I, from $p\bar{p}$ collisions at the Tevatron at a center-of-mass energy of $\sqrt{s}=1.8$ TeV. The full dataset had 109 pb^{-1} luminosity and was collected in two periods, between August 1992 and July 1995 (Table 3.1). The top discovery was announced in April, 1995, using only the first 67 pb^{-1}

of integrated luminosity.

Period of operation	Run	Integrated Luminosity (pb^{-1})
1987	-	0.025
1988-1989	Run 0	4.5
1992-1993	Run IA	19.4
1994-1996	Run IB	90.4
2001-2003	Run II	193.0

Table 3.1: Data taking periods, including RunIA and RunIB.

Run II started in March 2001, after six years of *major upgrades* both for accelerator complex and CDF detector. From *Run I* CDF inherited:

- The Superconducting Solenoid,
- The Central Muon System, and
- The Central Calorimeter.

The following subsystems are *new in Run II*:

- Front-End Electronics and DAQ System,
- Trigger System, with two new remarkable triggers, not available in Run I:
 - XFT in Level 1, and
 - SVT in Level 2
- Silicon Tracker (8 layers, $|\eta| < 2.0$),
- Central Outer Tracker ($|\eta| < 1.0$),
- Plug Calorimeter ($1.1 < |\eta| < 3.6$), and
- Extended Muon Coverage ($|\eta| < 1.5$, with gaps filled).

Descriptions of the majority of these systems are presented in the next sections.

Figures 3.2 and 3.3 show the integrated and peak luminosity, respectively, for the time period between March 5, 2001 to March 2, 2004. The first year of data taken was used for calibration purposes and to understand the subdetectors performance very well. For this analysis data collected between March 2002 until September 2003 is used.

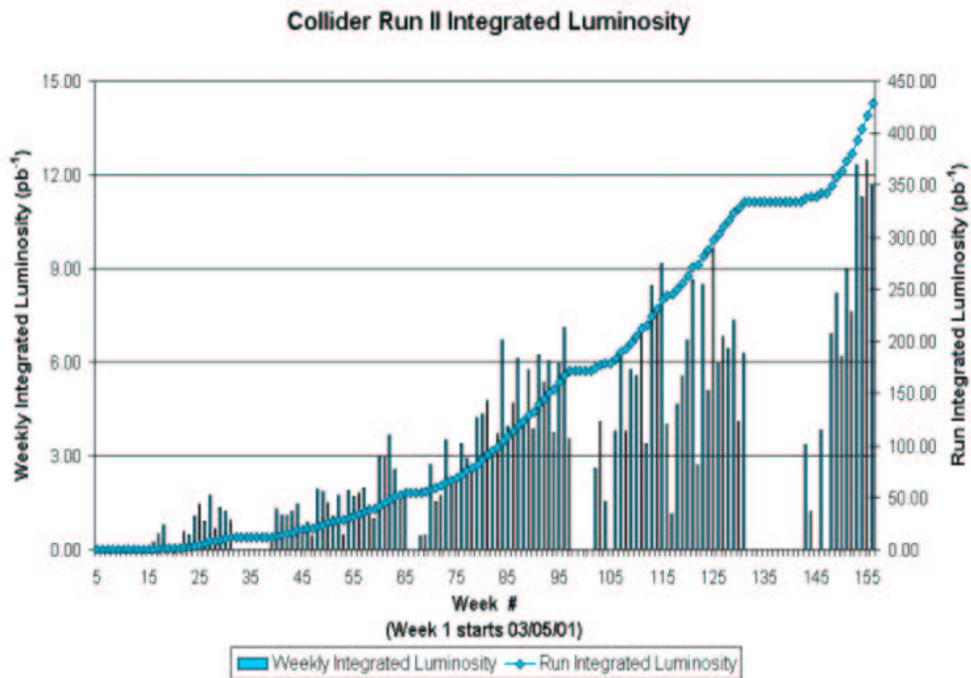


Figure 3.2: The weekly and per run integrated luminosity between March 5, 2001 and March 2, 2004 is shown. This analysis uses 193 pb^{-1} data taken between March 2002 and September 2003. Details about the luminosity measurement are given in Chapter 5.

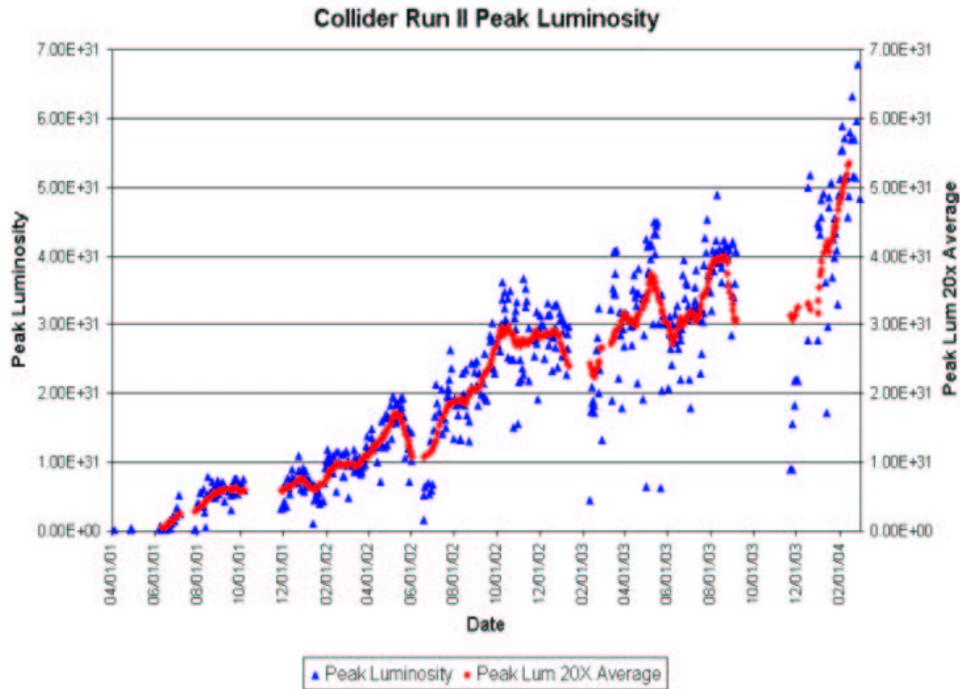


Figure 3.3: Run II instantaneous luminosity.

3.2 The CDF Coordinate System

The origin of the CDF coordinate system $(x,y,z) = (0,0,0)$ cm (Figure 3.4) is at the nominal point of collision, in the center of the detector. The positive z direction points in the direction of the proton beam (west to east), the positive y direction points upward (south to north) and the positive x direction points out of the ring.

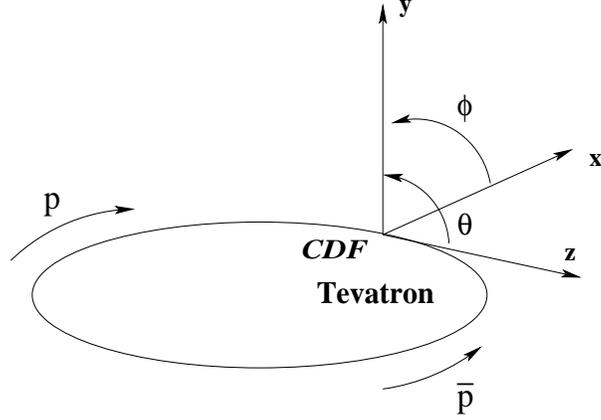


Figure 3.4: CDF Run II coordinate system.

Given that the energy spectrum of quarks inside the protons is very broad, the hard collision rest frame will be boosted, in general, along the beam direction with respect to the lab frame. Therefore it is appropriate to use *variables invariant under boosts along z direction*. The detector solid angle segmentation, described by the angular coordinates η and ϕ , satisfies this requirement. ϕ is the azimuthal angle about the z-axis. η is the pseudorapidity and is related to the polar angle θ through the relation:

$$\eta \equiv -\ln\left(\tan\frac{\theta}{2}\right) \quad (3.1)$$

Based on this definition, negative η corresponds to the west side of the detector, positive η to the east side of the detector, while $\eta = 0$ is the transverse x-y plane.

The *pseudorapidity* η is precisely the rapidity of a particle in the limit of $m \ll p$, where m is the particle rest mass and p its momentum magnitude. The *rapidity* is defined as

$$y \equiv \frac{1}{2} \ln \frac{E + p_z}{E - p_z} \quad (3.2)$$

Using the relations between the hyperbolic and trigonometric functions,

$$\sinh \eta = \cot \frac{\theta}{2} \quad (3.3)$$

$$\cosh \eta = \frac{1}{\sin \theta} \quad (3.4)$$

and the notation $\alpha \equiv m/p_T$, where $p_T = p \sin \theta$, one can rewrite the energy as

$$E = \sqrt{m^2 + \frac{p_T^2}{\sin^2 \theta}} \quad (3.5)$$

and further the rapidity as

$$y = \frac{1}{2} \ln \frac{\sqrt{\cosh^2 \eta + \alpha^2} + \sinh \eta}{\sqrt{\cosh^2 \eta + \alpha^2} - \sinh \eta} \quad (3.6)$$

If $m \ll p$, then as $\alpha \rightarrow 0$, and the expansion on y in terms of α becomes

$$y \approx \eta - \frac{1}{2} \alpha^2 \tanh \eta + O(\alpha^3) \quad (3.7)$$

So for $\alpha = 0$, $y = \eta$.

Under a Lorentz transformation to another frame moving at velocity β , y transforms as

$$y \rightarrow y + \frac{1}{2} \ln \frac{1 - \beta}{1 + \beta} = y + \text{constant} \quad (3.8)$$

This implies that the segmentation in rapidity is Lorentz invariant, $dy \rightarrow dy$ under a boost along beam direction. Still, the rapidity is a function of the particle's mass and polar angle. The pseudorapidity is used to define the angular segmentation. It depends only on the polar angle and is approximately Lorentz invariant under z boosts for high p_T particles. Also ϕ is invariant under z boosts, as it is a transverse plane variable.

3.3 The CDF Calorimeters

Outside the solenoid there are both electromagnetic and hadronic calorimeters that detect particle showers. Only the particles with transverse momentum greater than approximately 350 MeV/c will escape the magnetic field and reach the calorimeter detectors. These detectors offer close to 4π solid angle enclosure of the interaction point, a 2π azimuthal coverage and the polar angle extending down to about 3° from the beamline. They are finely segmented radially outwards, in pseudorapidity and azimuth, forming a *projective tower geometry* that points back to the interaction point.

The calorimeters are key in *measuring electron, photon, jet energies and net transverse energy flow* through the detector (allowing the extraction of the missing energy). Both electromagnetic and hadronic regions use *shower sampling* to measure the particle energy and position. The calorimeters are constructed of many layers of dense absorber materials (steel or lead) interleaved with scintillator material sensitive to ionization. Particles striking heavy absorber nuclei generate more particles which give rise to a *shower of particles*. In the *electromagnetic part*, for electrons and photons, the *phenomena of bremsstrahlung and pair production* combine to give rise to cascade showers as following: a parent electron will radiate photons, which convert to electron-positron pairs, which radiate and produce new pairs in turn, the number of particles increasing exponentially with depth in the medium. The shower shape is roughly a Gamma function $t^\alpha \exp(-\beta t)$, where t is the radiation length for EM and nuclear absorption length for HAD showers. A *hadron shower* appears when an incident hadron undergoes an inelastic nuclear collision producing secondary hadrons, which further multiply by inelastic interactions with the nuclei. The sum of the signals from all the scintillator layers is proportional to the energy of the incident particle.

The calorimeters of CDF closest to the $p\bar{p}$ interaction region consist of a central barrel region and two end-cap regions that enclose the barrel region. These calorimeters have a tower geometry that projects back to the center of CDF. The calorimeter regions consist on an inner electromagnetic (EM) section and an out hadronic (HAD) section. The central barrel region extends to $|\eta_{\text{det}}| = 1.1$, and the EM and HAD sections are called the *CEM and CHA*, respectively. The end-cap regions consist of two calorimeter systems: the “wall”

hadronic calorimeter (WHA) and the end plug calorimeters. The WHA provides coverage over $0.62 < |\eta_{\text{det}}| < 1.32$, complementing the CHA coverage. For this reason, the WHA, is included in CDFs “central calorimeter” category, along with the CEM and CHA. The plug calorimeters cover $1.1 < |\eta_{\text{det}}| < 3.5$, and its EM and HAD sections are called the *PEM and PHA* respectively. The PHA coverage begins from $|\eta_{\text{det}}| = 1.2$. The very-forward EM calorimeters, the mini-plug (MNP) calorimeters, are located beyond the outer ends of the PHA calorimeters. They cover the region, $3.5 < |\eta_{\text{det}}| < 5.5$.

Central Calorimeters

The central region ($0 < |\eta| < 1.1$ or $143^\circ < \theta < 37^\circ$) is the *most important to high transverse momentum physics*, being at large polar angle with respect to the beamline. For example top pair-production events tend to be central, top and anti-top being produced almost at rest. *The central calorimeter (the electromagnetic (CEM) and hadronic (CHA, WHA))* is retained largely unchanged from Run I other than the electronics, new in Run II. It consists of 2 barrels (one for the positive and one for negative η range), which are divided azimuthally into 24 wedges, each covering 15° in ϕ and extending 2.5 m along the beam axis on either side of the center of the detector. The wedge modules are stacked into four freestanding “C”-shaped arches to allow easy access to the inner components. One module is notched to allow access to the superconducting magnet. This affects tower 9, which is not a full size tower and is not used for electron identification in the dilepton analysis. Each wedge module is divided transversely into 10 projective towers, each subtending $\Delta\eta = 0.1$ units in pseudorapidity. Towers are segmented in depth, each depth being read out by separate electronics channels. A CEM module (Fig 3.5) is composed of 31 layers of 3.175 mm thick lead absorber interleaved with 5 mm thick layers of polystyrene scintillator. For each tower there are two wavelength shifters (WLS), one on each side in azimuth ϕ , which guide the green (490 nm) waveshifted light to photomultiplier tubes (PMT). Each tower is read out by 2 PMTs. The signal balance between PMTs allows further determination of ϕ for a single particle to 1° precision. In the Level 1 trigger, the energy is calculated as the average of the two tower energies, while in Level3 and offline, the tower energy is a geometric mean of the two PMT energies.

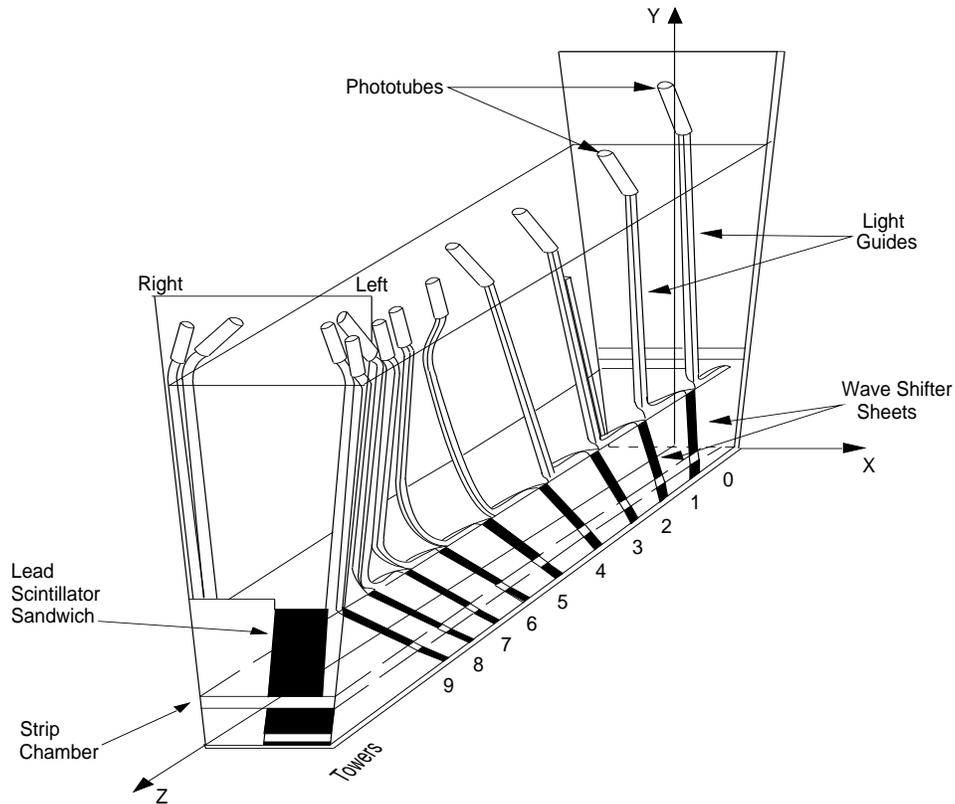


Figure 3.5: Schematic view of one wedge of the Central Electromagnetic Calorimeter (CEM). The Central Electromagnetic Strip Chamber (CES) is embedded at the maximum shower development point. Each wedge module is divided transversely into 10 projective towers.

The Central Electromagnetic Strip Chamber (CES) (Fig. 3.5) is embedded between the eighth lead layer and ninth scintillator layer. It is a proportional wire chamber that measures the positions and the transverse shower shapes of electromagnetic clusters in both r - z and r - ϕ planes. The CES is positioned at the average maximum longitudinal development of an electromagnetic shower, at about 5.9 radiation lengths from the inner radius of the CEM. There are 128 cathode strips that lie perpendicular to the beam direction measuring the z position of the shower. There are 64 anode wires, grouped in pairs, that lie parallel to the proton beam and measure the x coordinate. The position resolution is about 2 mm. The CES also provides position information for the identification of photons within particle showers. The detector is mostly unchanged from Run I, but the readout has been modified to accommodate the higher Run II collision rates. The detector has many cracks, or regions with low response. The region $-0.05 < \eta < 0.05$ near the 90° crack where the detector halves

meet, is not used, as the chambers are not fully efficient near the edges. This results in a loss in acceptance of about 5.0 %. For the same reason the showers with a distance less than 1° from the ϕ boundary between wedges are not considered. This translates into the requirement that the CES wire cluster closest to the extrapolated electron track be located at less than 21 cm from the center of the strip chamber in the r - ϕ plane. The loss in the acceptance is $24 \times 2^\circ / 360^\circ = 13.3\%$. Also the region $0.82 < \eta < 1.0$, $75^\circ < \phi < 90^\circ$ is explicitly excluded as it is uninstrumented. This is known as *chimney module* and is the access point for the cryogenic supply of the superconducting coil. The loss is 0.4 %. *The total acceptance loss due to the fiducial requirements is 18.7% of the geometrical acceptance for the central electrons.*

The Central Pre-Radiator (CPR) is situated in the gap between the solenoid coil and CEM, at a radius of about 168.0 cm ([91]). It is a single plane of multiwire proportional chamber with 32 sense wires running along the beam direction, providing 32 readout channels per wedge. The readout is split between two chambers, 16 channels at low z ($7.9 \text{ cm} < |z| < 119.7 \text{ cm}$) and 16 channels at high z ($123.5 \text{ cm} < |z| < 235.3 \text{ cm}$). It provides the measurement of the x coordinate only. CPR provides very good differentiation between electrons and minimum ionizing particles, like muons or hadrons. It allows help to identify electrons from conversions ($\gamma \rightarrow e^+e^-$), taking place inside the coil. However CPR information is not used in this analysis.

Parameter	CEM	CHA/WHA	PEM	PHA	
Angular η coverage	$ \eta < 1.1$	$ \eta < 0.9 / 0.7 < \eta < 1.3$	$1.1 < \eta < 3.6$	$1.2 < \eta < 3.6$	
Segmentation (in $ \eta < \text{range}$) $\Delta\eta \times \Delta\phi$		$ \eta < 1.1$ $0.1 \times 15^\circ$	$1.1-1.8$ $0.1 \times 7.5^\circ$	$1.8-2.1$ $0.16 \times 7.5^\circ$	$2.1-3.6$ $0.2-0.6 \times 15^\circ$
Absorber	lead	iron	lead	iron	
Active medium	scintillator	scintillator	scintillator	scintillator	
Energy Resolution	$14\%/\sqrt{E}$	$75\%/\sqrt{E} \oplus 3\%$	$16\%/\sqrt{E}$	$80\%/\sqrt{E}$	
Position Resolution	$0.2\text{cm} \times 0.2\text{cm}$	$10\text{cm} \times 5\text{cm}$			
Longitudinal Depth	$19 X_0, 1\lambda$	4.5λ	$21X_0, 1\lambda$	7λ	

Table 3.2: Specifications of the CDF calorimeter detectors. The energy resolutions are for incident electrons and photons (EM calorimeters) and isolated pions (HAD calorimeters). The position resolutions are averages for the calorimeter subsystems. X_0 refers to radiation lengths (EM calorimeters) and λ_0 refers to interaction lengths or attenuation (HAD calorimeters).

The *Central Hadron (CHA)* and *Wall Hadron (WHA)* calorimeters lie beyond the CEM. Each tower in the CEM is matched by a hadronic tower. The coverage for these detectors and the depth are given in Table 3.2. The CHA is constructed of 32 layers of 2.5 cm thick steel absorber alternating with 1.0 cm thick plastic scintillator. The WHA is constructed of 15 layers of 5.1 cm thick steel absorber alternating with 1.0 cm thick plastic scintillator.

Each gap between adjacent wedge modules is covered by a $12 X_0$ tungsten bar backed by a wire chamber. The tungsten serves to recover some of the response for particles, particularly photons, which would otherwise escape or produce unwanted light in the calorimetry readout wavelength shifters and light guides in this region. The chamber is known as *Crack Chambers*, and it is not used in this measurement.

Plug Calorimeters

In the region $|\eta| > 1.1$ new scintillating tile calorimeters replace the Run I gas calorimeters, which would have a time response incompatible with the crossing rate for Run II. The plug calorimeters (PEM, PHA, covering the region $1.1 < |\eta| < 3.64$) are closer to the beamline and subject to higher particle rates and energies. They consist of an electromagnetic (PEM) followed by a hadronic (PHA) section as shown in Figure 3.6. Both sections have as active elements scintillator tiles read out by wavelength shifting (WLS)

fibers embedded in the scintillator.

The EM section is a 23 layer lead and scintillator sampling device, with each layer made of 4.5 mm lead and 4 mm scintillator. The HAD section consist of 23 layers of iron and scintillator with each layer being composed of 51 mm of iron and 6 mm of scintillator. The detecting elements are arranged in a tower geometry pointing back towards the interaction region. The tower segmentation is shown in Table 3.3.

Tower Number	Tile ID	$\Delta\eta$	$\Delta\theta$	$\Delta\phi$
10	EM only	1.10-1.20	33-37 °	7.5 °
11	17,18	1.20-1.32	30-33 °	7.5 °
12	15,16	1.32-1.41	27-30 °	7.5 °
13	13,14	1.41-1.52	25-27 °	7.5 °
14	11,12	1.52-1.64	22-25 °	7.5 °
15	9,10	1.64-1.78	19-22 °	7.5 °
16	7,8	1.78-1.93	16-19 °	7.5 °
17	5,6	1.93-2.11	14-16 °	7.5 °
18	4	2.11-2.33	11-14 °	15 °
19	3	2.33-2.60	8-11 °	15 °
20	2	2.61-3.00	6-8 °	15 °
21	1	3.00-3.64	3-6 °	15 °

Table 3.3: The nominal transverse tower segmentation of the end plug calorimeter. The tower numbering is an extension of the CHA/WHA scheme.

The scintillator tiles of the first layer of the EM section are made out of 10 mm thick scintillator and are read out by Multi-Anode Photomultiplier tubes (MAPMT). They act as a preshower detector, known as *PPR (Plug Preradiator)*. A position detector is located at the depth of the EM shower maximum ($\approx 6 X_0$). It is known as the *PES (Plug Shower Max)* detector and is made of plastic scintillator strips read out by WLS fibers, followed by clear fibers that carry the light to MAPMTs. The energy resolution of the plug calorimeters is shown in Table 3.4.

	EM	HAD
Segmentation	$\sim 8 \times 8 \text{ cm}^2$	$\sim 24 \times 24 \text{ cm}^2$
Total Channels	960	864
Thickness	$21 X_0, 1 \lambda_0$	$7 \lambda_0$
Samples	22 + Preshower	23
Active	4 mm Scint	6 mm Scint
Passive	4.5 mm Pb	2 inch Fe
Light Yield (pe/MIP/tile)	5	5
Resolution	$16 \text{ \%}/\sqrt{E} \oplus 1\%$	$80 \text{ \%}/\sqrt{E} \oplus 5\%$

Table 3.4: Overview of the Plug Calorimeter. The EM (HAD) resolution is for a single electron (pion). X_0 is the radiation length and λ is the hadronic interaction length.

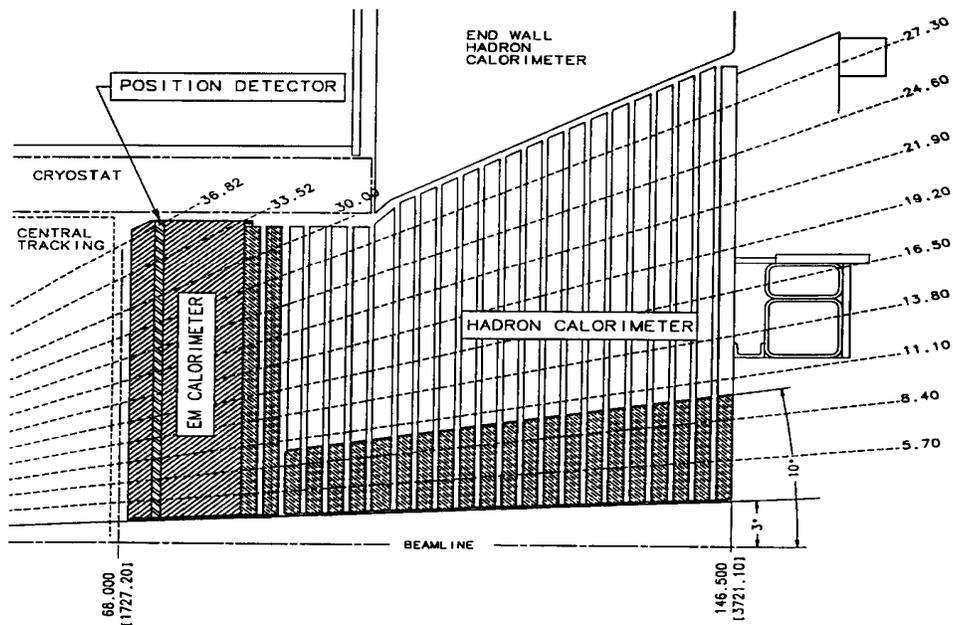


Figure 3.6: Cross section of one quadrant of the plug calorimeter.

3.4 The CDF Tracking Detectors

To achieve the objectives of the rich physics program planned for Run II, CDF tracking detectors were designed to be able to:

- Reconstruct charged particle tracks with efficiency greater than 95%, over the full range $|\eta| \leq 2.0$, and
- Measure charged particle momenta with precision of $\delta p_T/p_T = 0.1\%$ over the range $|\eta| \leq 1.0$, and with precision adequate for lepton identification, $\delta p_T/p_T \sim 0.4\%$, over the range $1.0 \leq |\eta| \leq 2.0$.

The tracking systems, shown schematically in Figure 3.7, are confined in a superconducting solenoid, which generates a 1.4 T magnetic field parallel to the beamline. At large radii, a new *cell drift chamber*, the COT (Central Outer Tracker) replaces the functionality of Run 0 and I CTC tracker over the region $|\eta| \leq 1.0$ ³. Inside the COT there is a *silicon inner tracker* built from three components: ISL, SVX and L00.

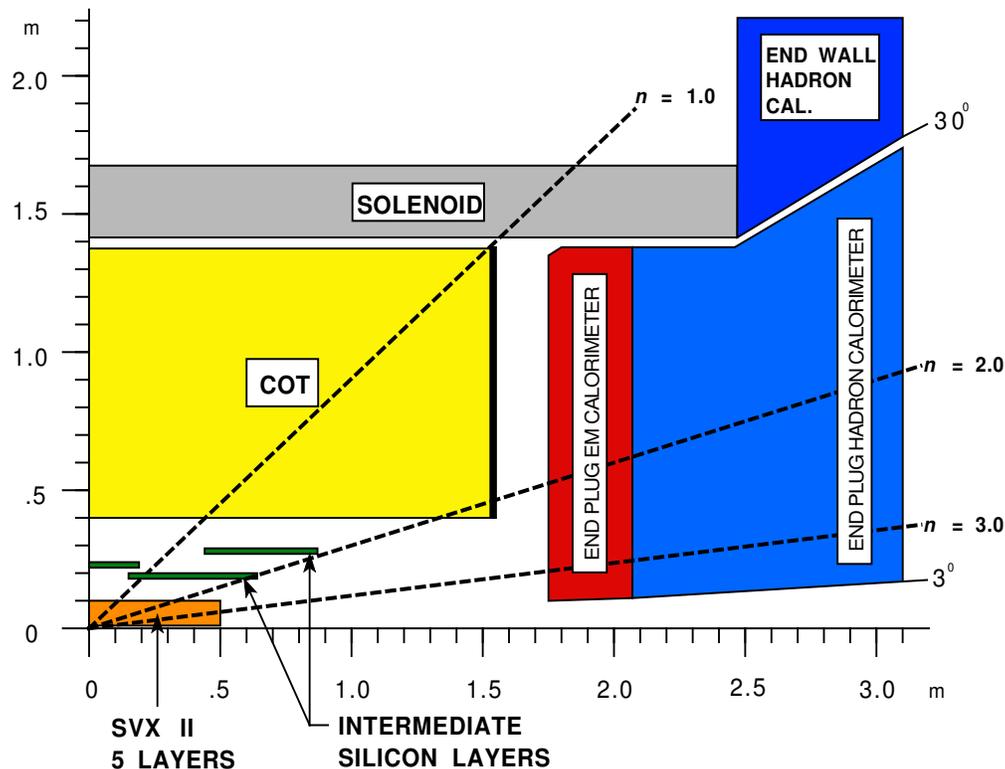


Figure 3.7: Longitudinal View of the CDF II Tracking System.

³In Run II CTC would suffer from severe occupancy problems at instantaneous luminosities $L \geq 1 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$

Central Outer Tracker

The COT is a large open cell, cylindrical drift chamber. It provides reliable tracking in the central pseudorapidity region ($|\eta| < 1.0$). The active volume of the COT spans 310 cm along the beam axis, z , between 43.4 cm and 132.3 cm in radius, and a complete 2π azimuthal coverage. The chamber contains 30240 sense wires that run the length of the chamber, grouped in 96 layers in radius, further grouped into eight “superlayers” as seen from the end plate section, shown in Figure 3.8. For operation with 396 ns bunch spacing, the drift gas consists of a mixture of Argon and Ethane in the proportions 50:50. Approximately half of the wires are axial (run along the z direction) and half are small angle (2°) stereo (titled at $\pm 2^\circ$ with respect to the beam direction). The superlayers alternate, starting with a stereo superlayer, nearest to the beampipe. The axial wires provide accurate tracking in the r - ϕ plane for the measurement of the transverse momentum, p_T , while the tracking information in r - z view is less accurate. Table 3.5 lists a few of the design parameters of the COT drift chamber.

COT	Parameters
Rapidity Coverage	$\eta < 1.0$
Radial Coverage	44 to 132 cm
Number of Layers	96
Number of Superlayers	8
Readout Coordinates per SLs	$+3^\circ 0 -3^\circ 0 +3^\circ 0 -3^\circ 0$
Cells/Layer	168 192 240 288 336 384 432 480
Sense wires/Cell	12 12 12 12 12 12 12 12
Radius of Center of SL	46 58 70 82 94 106 119 131 cm
Length of Active Region	310 cm
Material Thickness	$1.6\%X_0$
Maximum Drift Time	100 ns

Table 3.5: Few characteristics of the Central Outer Tracker at CDF.

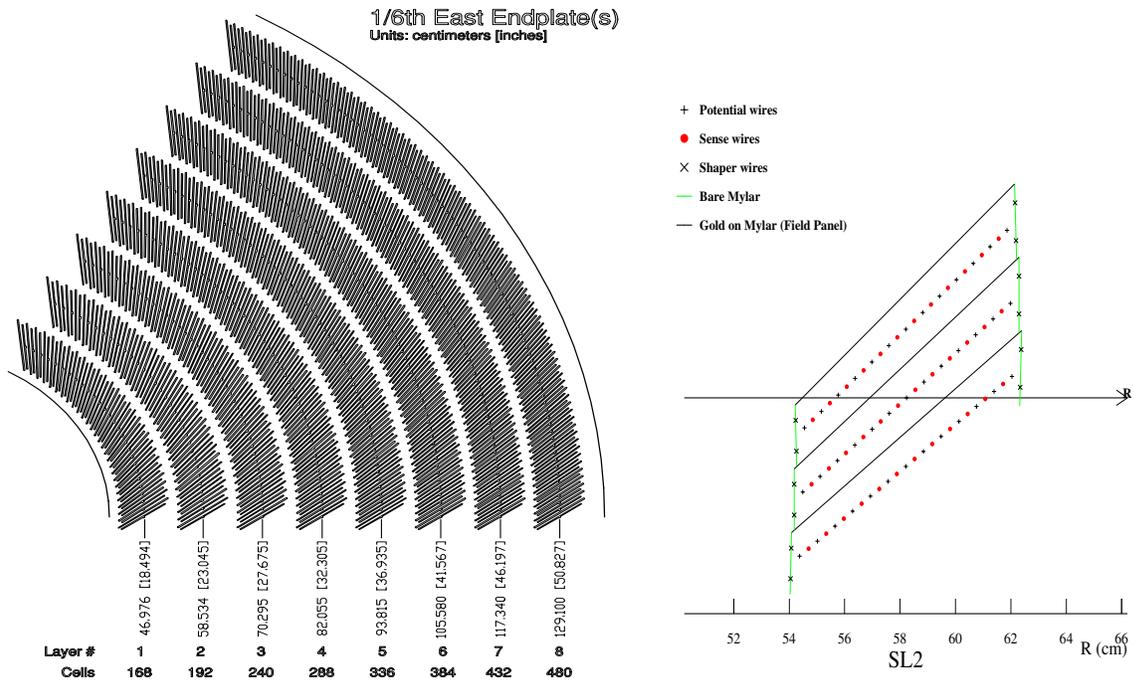


Figure 3.8: On the left, three supercells for SL2 are shown, looking along the beam (z) direction. The other superlayers are similar. On the right, the endplate slots Sense and field planes are shown. The 8 superlayers alternate, the even are axial, while the odd ones are stereo.

COT aging problems

The first evidence of the aging was realized from the drop in the wire gains, in the Fall 2003. The drop was seen to be ϕ and Z dependent, as shown in Figure 3.9. There are two things learned from the various studies performed in the following months: the hydrocarbon growths were found on the wires, and the Z dependence traced to the gas flow stagnation. A fix was implemented in mid-June 2004 which consisted of: increase in the gas flow for better recirculation of the gas, and the addition of 100 ppm oxygen to reverse the hydrocarbon growth. Now the COT wire gains have been recovered, they are back to the 2002 level, and ϕ and Z variations are nearly gone.

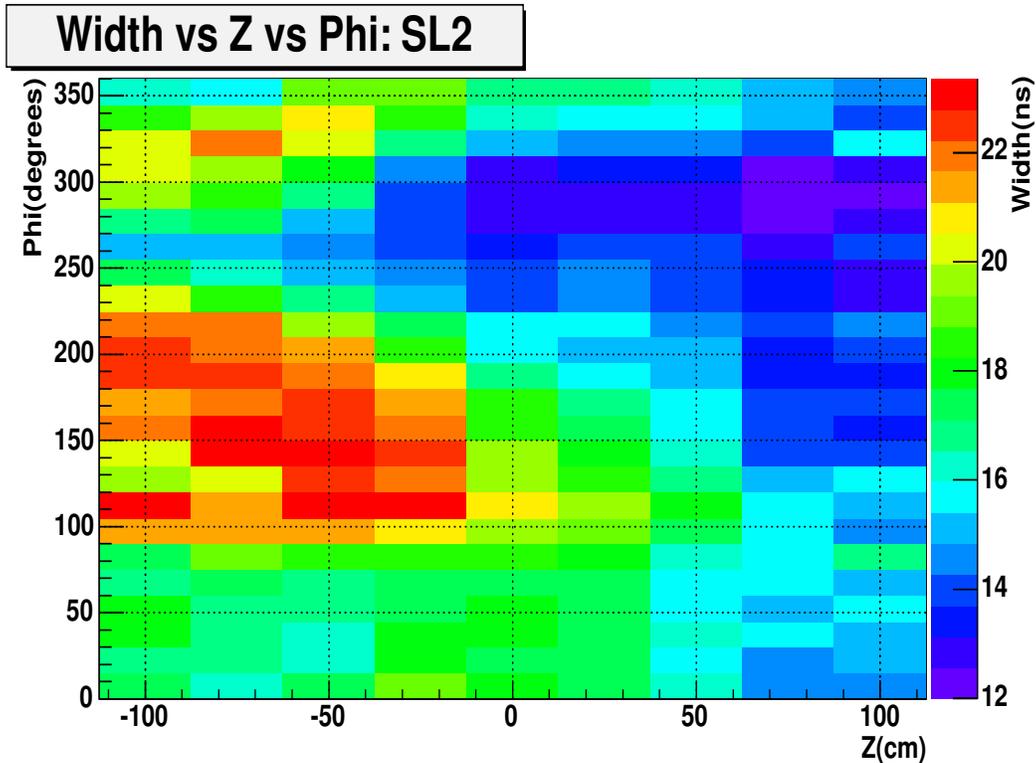


Figure 3.9: COT hit width for SL2, run 179643. The aging effect is seen to be dependent of ϕ and Z . With ϕ it gets worse near the bottom ($\phi = 180$ degrees), and it gets worse with increasing Z , from East to West.)

Silicon Detectors

The silicon inner tracking consists of three concentric silicon detectors located near the beampipe (Figure 3.7). During the operation, due to high doses of radiation, the silicon bulk changes from n-type to p-type. This is called *type inversion*. The rate of which type inversion occurs depends on the operation temperature, being slower at lower temperatures. Therefore a cooling system keeps the silicon detector at low temperature during the normal operation. Details about the Cooling and Interlock System are given later in this Section.

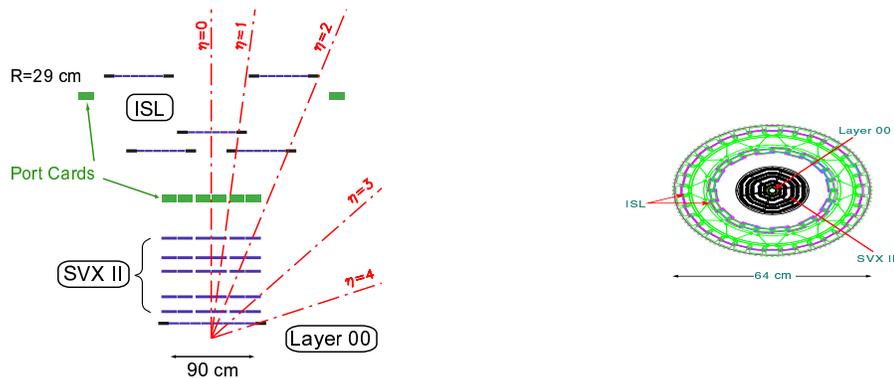


Figure 3.10: The r-z view of the CDF II Silicon Detector.

L00 Silicon Detector

The innermost silicon detector, Layer 00 (L00) consists of single-sided, radiation-tolerant, silicon strips, at very small radius (1.35 cm), just outside the beampipe (located between the radii of 0.83 and 1.25 cm). Layer 00 ([75; 38]) provides excellent coverage with minimal material inside the tracking volume, improving impact parameter resolutions and the b-tagging efficiencies.

Silicon Vertex Detector

SVX is situated outside L00, at the radius of 1.6 cm and consist of three cylindrical barrels, each 29 cm long. The barrels extend about 48 cm along z direction on each side of the interaction point, covering $\sim 1.5\sigma$ of the $p\bar{p}$ luminous region in Z ($\sigma_z \sim 28$ cm). Each barrel is divided into 12 wedges in ϕ and each wedge supports five layers of double-sided silicon micro-strip detectors between the radii of 2.4 and 10.7 cm from the beamline. SVX provides tracking coverage for the region $-2.0 < \eta < 2.0$. Of the five layers, three have 90° stereo design (Layer 0, 1 and 3), while two have 1.2° small-angle stereo design (Layer 2 and 4). The small angle information from all the layers is combined to form a three dimensional track. The impact parameter resolution is $\sigma_\phi < 30 \mu\text{m}$ and $\sigma_z < 60 \mu\text{m}$, for *central high-momentum tracks*. A few of the design parameters are shown in Table 3.6.

Detector parameter	SVX II
Readout coordinates	r- ϕ ,r-z
Number of layers per barrel	5
Ladder length	29.0 cm
r- ϕ readout channels	211,968
r-z readout channels	193,536
Total number of channels	405,504
Total number of readout chips	3168
Total number of ladders	180

Table 3.6: Design parameters of the SVX II detector.

Intermediate Layer Detectors

ISL consist of three silicon layers placed at radii of 20, 22 and 28 cm respectively from the beamline. As seen in Figure 3.10 the central layer covers the region $|\eta| < 1.0$, while the two outer layers cover the forward region $1.0 < |\eta| < 2.0$, where the COT coverage is incomplete (See Figure 3.11). ISL incorporates many features of the SVX II design: each layer is double sided and mounted in ladder assemblies, in a similar way. The combined information of the SVX and ISL allows the reconstruction of three dimensional tracks independently (standalone mode) of the COT, thus providing a simple way to measure the COT tracking efficiency. The ISL allows us to extend the tracking, lepton identification, and b-tagging capabilities in the plug region ($1.0 < |\eta| < 2.0$), the net effect for the dilepton analysis being an increase in the acceptance by about 30%. The resolution of the standalone tracking (SVX+ISL) has been estimated to be

$$\sigma_{p_T}/p_T^2 \approx 0.4 \%; \sigma_{d_0} = 15 \mu\text{m} ; \sigma_{\phi_0} = 0.3 \text{ mrad} ([39]).$$

where d_0 is the impact parameter of the track and ϕ_0 is the angle that the track forms with the x-axis.

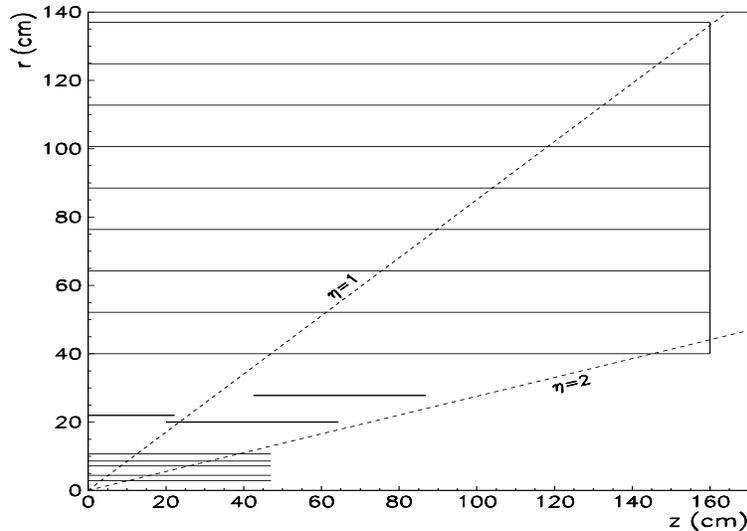


Figure 3.11: An r-z view of the ISL layers radius. SVXII and COT are also shown.

Silicon Interlock and Cooling System (SICS)

The Silicon Cooling System (3.12), is designed to maintain the Silicon detector at a low temperature during the data-taking. Its *brain*, the Interlock System, protects against any failure. The need for cooling is *two-fold*, to protect the on-board electronics from overheating and to keep the silicon ladders cold, since colder operation means less damage from the radiation, by minimizing the reverse annealing effect⁴. Therefore a cold liquid flows through the bulkheads to carry away the heat generated inside the detector. The coolant was designed to be a mixture of water with ethylene glycol (30%), which allows temperatures down to -10 Celsius.

The Silicon Cooling System consists of two customized chillers, one which cools the SVX and L00 ladders, running at -6 C, and a second, which cools the ISL ladders and all the portcards, running at +6 C. The heat exchangers use the B0 chilled water system. Each chiller has supply and return manifolds on the east and west sides of the detector. From the

⁴Radiation-induced defects are introduced into the crystal lattice, under high fluences of neutrons or high-energy hadrons. Defects can be electrically active, leading to increased space-charge, leakage current and charge trapping. Increased space-charge prevents the electric field from penetrating the material unless high bias voltages are used. Moreover, radiation-induced space-charge can increase after the radiation source is removed, a phenomenon called *reverse annealing*.

manifold, individual lines carry coolant to the silicon. To protect against any possible leaks inside the Silicon detector, we keep the pressure subatmospheric in the cooling lines going inside the detector. This is done by pulling vacuum on the return side. The air separator is the point of minimum pressure in the system, ~ 2 psia.

There are two types of interlocks, an interlock on the flow, and an interlock on the power. The logic is performed by a Siemens Programmable Logic Controllers (PLC), which also reads the temperatures, flows, pressures, conductivity and sends commands to valves and relays. The logic consists of two main steps:

- Check if it is safe to begin flowing, and
- Check if the system is stable and cold so one could turn on the power on the detector.

Because the Siemens PLC has no industry-standard safety rating, a second safety-rated PLC, Quadlog, is used, as a parallel backup.

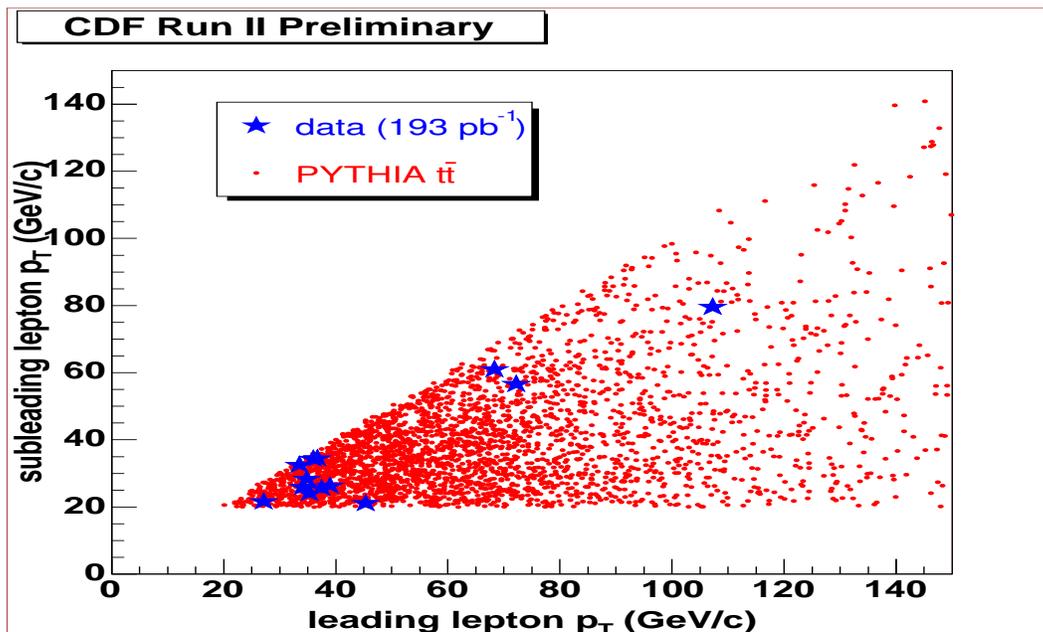


Figure 3.12: SVX Cooling System.

A third component of the SICS, iFIX, a commercial software, is used for monitoring and control of the SICS.

	CMU	CMP/CSP	CMX/CSX	BMU
Pseudorapidity coverage	$ \eta \leq 0.6$	$ \eta \leq 0.6$	$0.6 \leq \eta \leq 1.0$	$1.0 \leq \eta \leq 1.5$
$\Delta\phi^\circ$	360	360	360	270
Drift tube length (cm)	226	640	180	363
Max drift time (ns)	800	1400	1400	800
Number of channels	2304	1076	2208	1728
Pion interaction lengths	5.5	7.8	6.2	6.2-20
Minimum detectable μp_T	1.4 GeV/c	2.2 GeV/c	1.4 GeV/c	1.4-2.0 GeV/c
Multiple scattering resolution (GeV/p)	12 cm/p	15 cm/p	13 cm/p	13-25 cm/p

Table 3.7: Design Parameters of the CDF II Muon Detectors. Pion Interaction lengths and multiple scattering are calculated at a reference angle of $\theta = 90^\circ$ in CMU and CMP/CSP, at an angle of $\theta = 55^\circ$ in CMX/CSX, and show the range of values for the IMU.

3.5 The CDF Muon Detectors

Muon identification is possible because muons are the only charged particles which can penetrate large amounts of material. As described in 3.3, electrons and hadrons interact destructively in material. Muons, due to their large mass relative to the electron, do not suffer from such catastrophic interactions. Therefore the method for detecting muons is to place charged particle detectors behind a large amount of shielding material (typically steel).

In Run II there are four systems of scintillators and proportional chambers which allow muon detection over the region $|\eta| \leq 1.5$. The absorbers for these systems are the calorimeter steel, the magnet return yoke, additional steel walls, and the steel from the Run I forward muon toroids. The geometric ($\eta \times \phi$) coverage for muon detectors: CMU (Central Muon Detector), CMP (Central Muon Upgrade) and CMX (Central Muon Extension) is shown in Figure 3.13.

Central Muon Detector

The CMU is the original muon chamber used in Run I and consists of 144 modules with 16 rectangular cells per module, located behind 5.5 λ of absorber (CHA). The 16 cells in a module are stacked four deep in the radial direction, with a small ϕ offset between the first

and the third, and the second and fourth.

Central Muon Upgrade

Outside of CMU, there is an additional 60 cm of steel in the region 55 - 125 degrees, for further hadron absorption, and then a second set of muon chambers, which make up the central muon upgrade. The chambers form a box around the central detector. CMP chambers are rectangular, single-wire drift tubes configured in four layers with alternate half-cell staggering. The chambers are run in proportional mode. Table 3.7 lists few of design parameters of CMP.

Central Muon Extension

The CMX consists of conical sections of drift tubes and scintillation counters (CSX), located at each end of the central detector and extending in polar angle from 42° to 55° and 180-55 to 180-42 degrees. Additional steel absorbers were installed between the beampipe and CMX, to reduce beam-splash backgrounds ([18]).

Intermediate Muon Detector

The IMU (Intermediate Muon Detector) is designed to trigger on muons with $1.0 \leq |\eta| \leq 1.5$ and to identify offline isolated high- p_T tracks (muons, hadrons) over $1.5 \leq |\eta| \leq 2.0$. The offline identification benefits from the new ISL silicon detector, which allows to match the track matching to a muon stub. This analysis does not use the IMU chambers, but extending the muon identification up to $|\eta| \leq 2.0$ will be beneficial to top analyses.

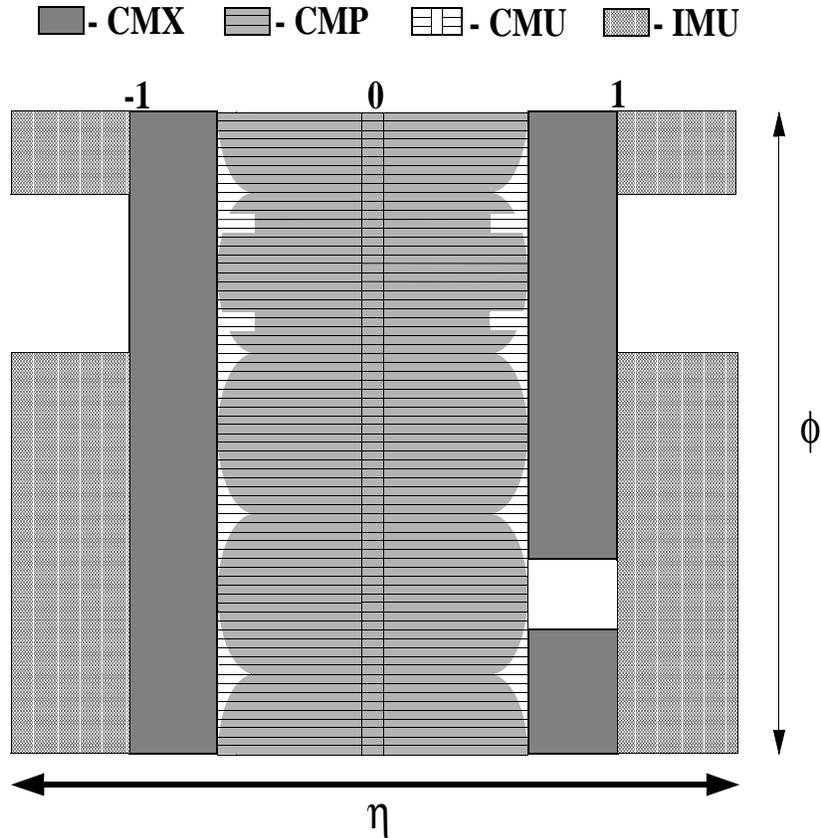


Figure 3.13:

Location of the central muon upgrade components in azimuth ϕ and pseudorapidity η for Run II.

3.6 Cerenkov Luminosity Counters

The integrated luminosity of the data samples, another piece needed to measure the dilepton cross-section, is determined in CDF using the Cerenkov Luminosity Counters (CLC)[42].

CLC consists of two modules, installed around the beampipe, at both ends of the CDF detector, inside the end-plug calorimeters, in pseudorapidity range of $3.7 < |\eta| < 4.7$. Each module is made of 48 thin, long, conical, gaseous Cerenkov counters, pointing towards the center of the interaction region. There are three concentric layers of counters, each layer having 16 counters, surrounding the beampipe. Each counter is made of highly reflective

aluminized mylar, with a light collector (conical mirror) at the large aperture of the mylar cones. The light collector gathers the Cerenkov radiation to a fast and radiation hard photomultiplier, with good ultra-violet (UV) efficiency. The modules are enclosed in a pressure vessel, filled with isobutane gas at about 22 psi, which is a very good radiator (refraction index is 1.00143) and has good transparency for UV radiation⁵.

CLC allows to determine the average number of $p\bar{p}$ interactions, by measuring the number of particles originating from the collisions within a bunch crossing, and their arrival time. *Prompt particles* will transverse the full length of the counter and generate a large PMT signal (~ 100 photoelectrons), while the non-primary and other backgrounds yield little signals. Appropriate amplitude thresholds are applied to discriminate the primary particles against the backgrounds.

The details of the offline luminosity measurement are given in Section 5.2.

3.7 CDF Time of Flight System

The Time of Flight (TOF) detector[92] is made of 216 bars of plastic scintillator (4 cm \times 4 cm \times 280 cm), with photomultipliers attached at both ends of each bar. It is located between the COT and the cryostat of the solenoid, at the radius of 140 cm and covers an $|\eta| < 1.0$ pseudorapidity window.

As plotted in Fig. 3.14, an accurate measurement of a particle's time of flight can be used quite effectively in particle identification. For example, TOF allows the separation of kaons and pions with a 2σ statistical precision, up to a momentum of 1.6 GeV/c. This will enhance the b flavor tagging and it is crucial in precise measurements of neutral B meson flavor oscillations.

TOF is also used in the trigger system, for example, to check for abnormally large pulses caused by highly ionizing particles, such as a magnetic monopoles.

⁵Most of the Cerenkov radiation is emitted in the ultra-violet region of the spectrum.

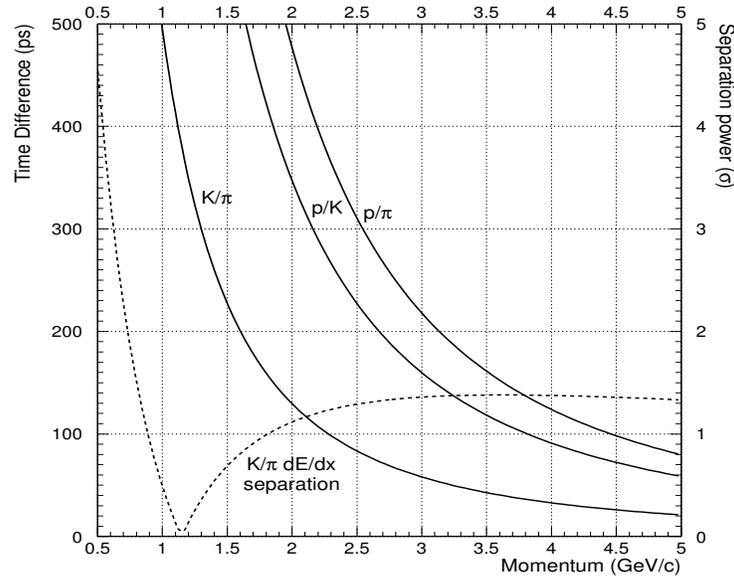


Figure 3.14: Time of flight difference as a function of particle type and momentum. The dashed line indicates the COT dE/dx separation power for $K\pi$: TOF and COT complement each other in different momentum regions.

3.8 The CDF Trigger System

The physics at a hadron collider requires a sophisticated trigger, because the collision rate is so much higher than the rate one can afford to write the data on tape. In Run II the collision rate is approximately equal to the crossing rate of 7.6 MHz, while the writing speed is about 50 Hz. So we keep an event every 150,000 collisions. The trigger system is designed to be able to efficiently select the most interesting physics events from the numerous minimum bias events.⁶ For example, the total $t\bar{t}$ cross-section is approximately nine orders of magnitude smaller than the minimum bias cross section. Table 3.8 shows the rates for a few of the physics processes of interest, compared with the inelastic $p\bar{p}$ collision rate.

⁶A minimum bias event is required to satisfy some minimal trigger conditions, usually some activity at small polar angle with respect to the beamline.

Process	Cross-section	Event rate
inelastic $p\bar{p}$	60 mb	6 MHz
$p\bar{p} \rightarrow b\bar{b}$	10 μ b	1kHz
$p\bar{p} \rightarrow WX$	5 nb	0.5 Hz
$p\bar{p} \rightarrow ZX$	0.5 nb	0.05 Hz
$p\bar{p} \rightarrow t\bar{t}$	6.7 pb	0.00067 Hz
$p\bar{p} \rightarrow WW b\bar{b}$	2 pb	0.0002 Hz
$p\bar{p} \rightarrow WH b\bar{b}$	15 fb	0.0000015 Hz

Table 3.8: The cross-sections and event rates for few of the interesting physics process at a hadron collider, compared with the inelastic $p\bar{p}$ cross-section. Instantaneous luminosity was assumed to be $10 \times 10^{31} cm^{-2}s^{-1}$.

Due to mainly the reduction in the accelerator bunch spacing from $3.5 \mu s$ to 132-396 ns, but also many new subdetectors, the trigger electronics along with all CDF front-end electronics is new for Run II. The CDF trigger and readout system is designed to operate with a beam crossing interval of 132 nsec. Therefore, the readout and trigger clock cycle has a period of 132 nsec.

CDF has a *three level trigger system*, each level reducing the rate enough to allow the next level of processing the data with *minimal deadtime*. Figure 3.15 shows the *Data Acquisition System (DAQ)* functional block, while the Run II *trigger system* is sketched in Figure 3.16 and few details are presented below.

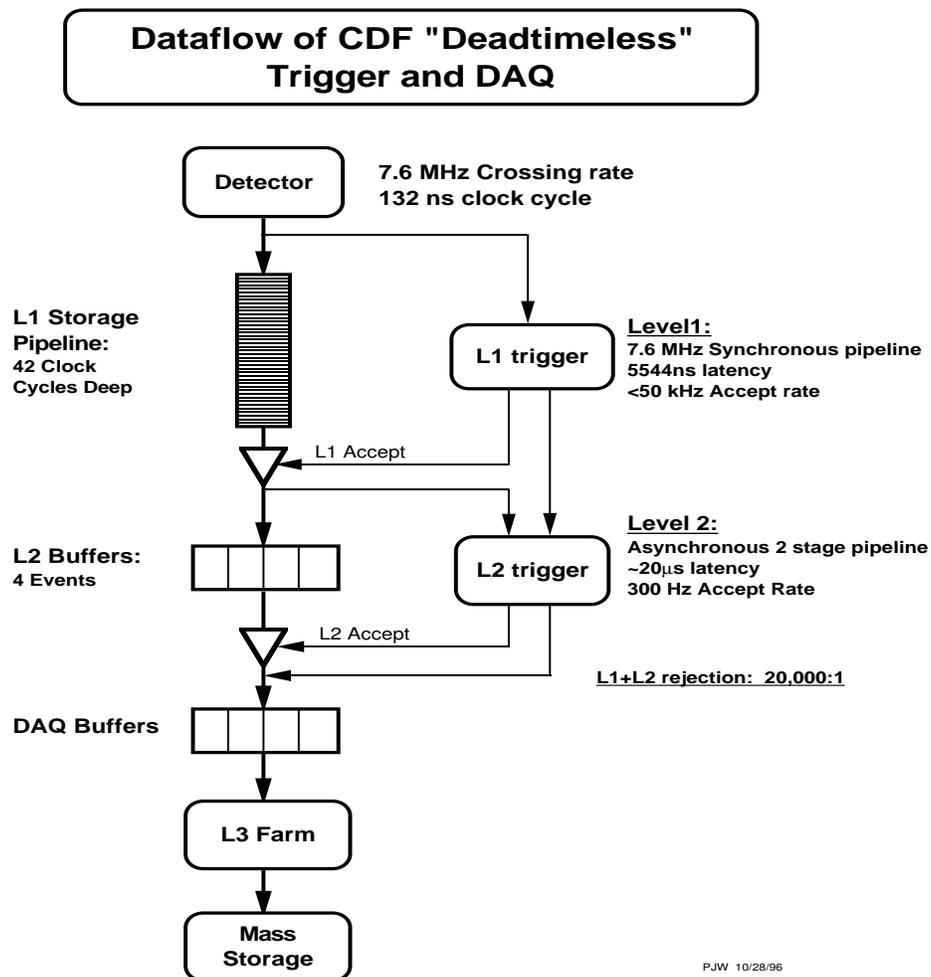
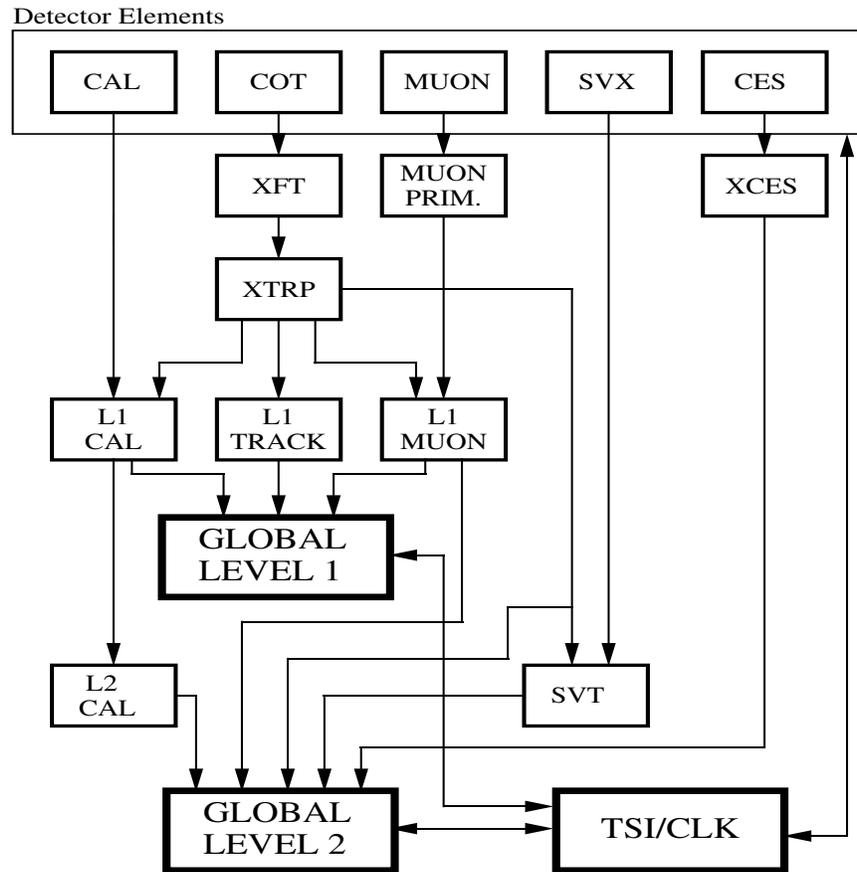


Figure 3.15: The readout functional block diagram in Run II.

RUN II TRIGGER SYSTEM



PJW 9/23/96

Figure 3.16: The trigger-system block diagram in Run II.

Level-1

Level-1 (L1) trigger uses custom designed hardware to find physics objects based on a subset of the detector information and makes a decision based on a simple object counting. Level-1 latency to make a trigger decision is $5.5 \mu\text{s}$. This requires that every subdetector have local data buffers for the 42 beam crossings (assuming a 132 ns bunch crossings separation) that occur during the decision taking latency.

The most significant trigger change compared with Run I is the addition of the *Track Finding Trigger (XFT)* in Level-1. Tracks from the drift chamber will be reconstructed in

2.7 μs after a $p\bar{p}$ collision. This opens up a lot of trigger possibilities: a track can be matched to an electromagnetic-calorimeter cluster for improved electron identification, or a track can be matched to a stub in the muon system for better muon identification and momentum resolution.

The Level-1 hardware consists of *three parallel synchronous processing streams* which feed inputs to the single Global Level-1 decision unit. One stream finds calorimeter based objects (L1CAL), another finds muons (MUON PRIM-L1MUON), while the third finds tracks in the central tracking chamber (XFT-XTRP-L1 TRACK). The tracks are sent to the calorimeter and muon streams as well as the track only stream, because the electron and muon triggers require a track pointing to an outer subdetector. Using simple logical operations (*ANDs* and *ORs*) of objects, up to 64 triggers can be formed from these streams. The elements of the Level-1 trigger are synchronized to the same 132 ns clock with a decision made every 132 ns by the Global Level-1.

The Level-1 calorimetry trigger allows CDF to trigger on electrons, photons, jets, total event energy and missing transverse energy, while the Level-1 muon trigger provides single and dimuon objects at the Level-1. The maximum L1 accept rate is ~ 20 kHz.

Level-2

Once an event is accepted by the Level-1 trigger, the front-end electronics move the data to one of the four on-board Level-2 buffers on all front-end and trigger modules. If all the four L2 buffers are full, then the experiment starts to incur deadtime. The most significant addition to the Level-2 in Run II is the *Silicon Vertex Tracker (SVT)* trigger, which allows for the first time in a hadron-collider experiment the ability to trigger on secondary vertices. This makes accessible a large number of processes involving hadronic b-quark decays. The highly parallel and complex silicon data acquisition system (DAQ), shown in 3.17, reads the entire silicon detector ($\approx 406\,000$ channels) in about 10 μs .

Level-2 is pipelined in two stages, each taking approximately $10 \mu s$. The latency remains $20 \mu s$, however the time between L2 decisions will be $10 \mu s$, resulting in a minimal deadtime for a L1 accept of 50kHz.

The L2 trigger performs a limited event reconstruction using all the L1 information, but with higher precision. In addition, at L2, data from the central shower-max and the SVX detectors is used, which improves the identification of electrons and photons and the secondary vertices reconstruction. Also, a cluster-finding algorithms is available to trigger more efficiently on jets. The current L2 maximum output rate is 380 Hz.

Level-3

The Level-3 (L3) trigger subsystem [61] has two main components, the *Event Builder (EVB)*, shown in Figure 3.18, and a *Level-3 processors farm*.

The initial rate of events depends upon the luminosity. The Level-1 and Level-2 physics cuts and prescales are tuned so that the output rate is about 300Hz, which is the designed input rate of the Event Builder (EVB). L1 and L2 systems need to make their decisions at a very high rate which makes it impossible to fully reconstruct each event. While L1 and L2 are making their decision, the event pieces are stored in the buffers of numerous (≈ 140) Front End (FE) crates. After a Level-2 decision is made, the *Event Builder* assembles all of the event fragments from the Front End crates into one data block. The assembling is a two stage process (See Figure 3.18):

- Event pieces from the FE crates are concentrated in 15 EVB crates, also known as *SCPU or VRB* crates.
- Event fragments from the EVB crates are sent to one of 16 Level-3 subfarms, where the event is finally assembled in one data block, and processed.

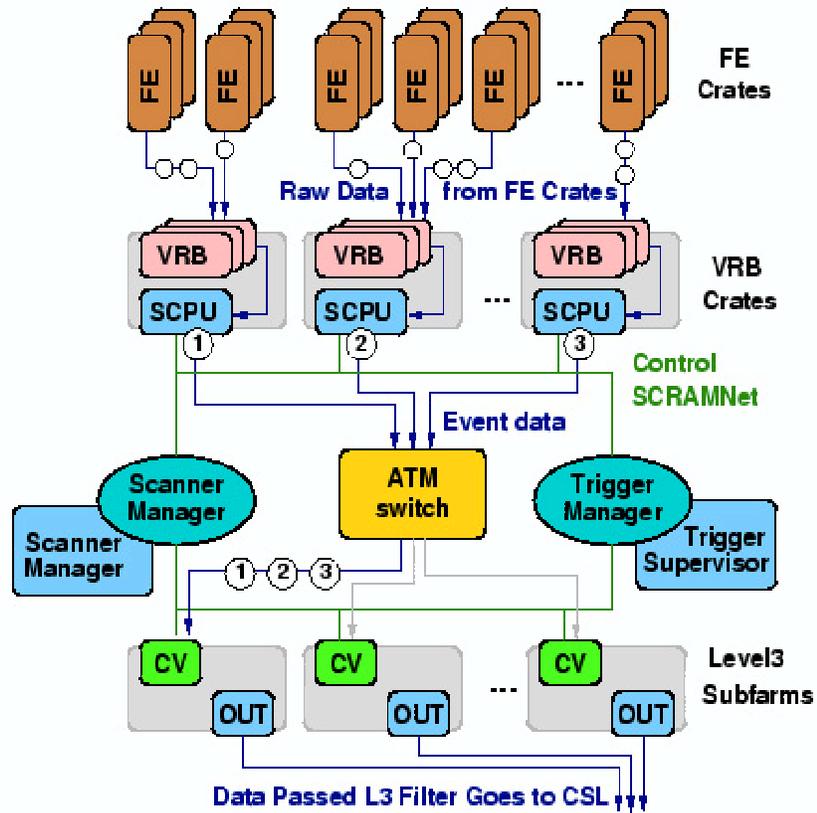


Figure 3.18: Event Builder block diagram.

The Level-3 Farm reconstructs the event with a simplified version of the CDF offline reconstruction software, and applies a software trigger. The designed Level-3 input and output rates have been exceeded, and are now 380 Hz and 85 Hz, respectively [65]. Each event passes several stages:

- Event data is reformatted ⁷ for Level-3 Filter,
- Level-3 Filter reconstructs the event,
- Level-3 Filter makes a trigger decision, and
- Events which pass the Level 3 trigger are sent to the *Consumer Server Logger (CSL)*.

⁷ The event is rearranged in a format readable by Level3 Filter

It takes about a second for one Linux PC to make a trigger decision on one event. It requires a large farm of 250 Dual Pentium III PCs to deal with the required input rate. Each of the 250 Processor nodes runs an instance of Level-3 Filter code on each of two CPUs and reconstructs two events in parallel. 16 nodes called *Converter nodes* serve as the Farm input. Converters distribute the data flow coming from the EVB between 16 subdivisions of the Level3 Farm called *Subfarms*. Each subfarm contains 14-18 Processor nodes and has one Converter node as an input. Eight PCs called *Output nodes* are dedicated to forward the events passing L3 Filter, from the Level-3 Farm to the CSL. Each Output node is shared by two subfarms. Events passed Level-3 are sent to CSL.

The CSL writes event data to disk and distributes a small fraction of events to *Consumer* programs that are used by the shift crew to monitor the quality of incoming data from the Control Room. The L3 accept rate is now about 85 Hz.

4 Event Reconstruction

This chapter describes the reconstruction of physics objects for a proton-antiproton collision, referred to as *an event*. The Run II reconstruction package uses the C++ Object Oriented programming language¹, and is constructed from a number of independent software modules, each handling the data from a subdetector. Each module communicates with the others through an interface. The modules are executed *sequentially*, in an order specified at run-time, in a *tcl* file². The behavior of a module can also be controlled at run-time through a number of parameters. The logical flow of event reconstruction in CDF is shown in 4.1. The reconstruction process starts with the raw data recorded by the CDF detectors. Then the calibration constants, noise suppression and various corrections are applied, before proceeding to construct higher-level objects, such as *Electrons*, *Muons*, *Jets* or *Missing Energy* .

¹Remember that in Run I, CDF used Fortran code for event reconstruction. The choice of C++ was made based on its wide use in scientific fields, including most of the new HEP projects, and its object-orientation strengths.

²A *tcl file* consists of commands which allows user to specify which modules will be executed, in which order and also the input parameters passed to the modules.

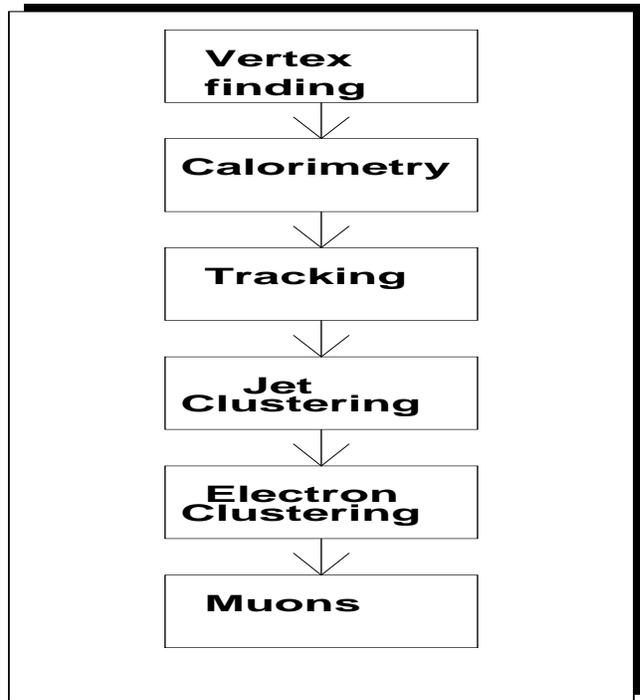


Figure 4.1: The logical flow of CDF event reconstruction in Run II.

4.1 Energy reconstruction

The *calorimeters* are used to measure the electron, photon, jet and net transverse energies. The raw energy in the calorimeter is measured in the ADC counts of each photomultiplier tube (PMT), corresponding to a calorimeter electronic channel. The wavelength shifters convert the blue scintillator light into green wavelengths which have a larger attenuation length. The green light is transported to the photomultiplier tube via clear light guide bars in the central calorimeters and plastic optical fibers in the plug calorimeters. For each central calorimeter tower, there are two PMTs. For each physical plug calorimeter tower, there is only a single PMT. The ADC counts are converted to energy (GeV)³, using detector dependent scale factors, determined using either the test beam data or specific data samples

³The actual conversion values are stored in the Offline Database Table CALDIGITOGEV3, and they are 3.199 MeV/ADC count for CEM and 4.440 MeV/ADC count for PEM.

collected during special runs. During the calibration runs, problem channels, such as dead or hot ones, are identified and processed separately. For each event, an $\eta - \phi$ array of tower energies is obtained, after the noisy channels are suppressed and spurious sources of energy are properly removed. This $\eta - \phi$ array is used to construct *an array of transverse energies*, E_T , using the polar angle θ of each tower center, with respect to the event's primary vertex. The event vertex is determined by extrapolating the particle tracks back to the beamline, the details for which are given later in this chapter. The transverse energy in a tower is $E_T = E \sin(\theta)$, where E is the energy measured in the tower.

There are two C++ classes of towers in the CDF software, *CalTower* and *PhysicsTower*. The CalTowers are the *raw experimental* towers, i.e. they contain the information that was read by the photomultiplier tubes, while the PhysicsTowers have their transverse energies calculated from a specific vertex and have a four vector associated to them. Therefore, a CalTower will contain, in particular, the energy information, but not the E_T information, which depends on the vertex knowledge. On the other hand, a PhysicsTower will contain the E_T information, and will also have a 4-momentum vector associated to them. The CalTowers are used to create the PhysicsTowers via CalData, and the PhysicsTowers are used to construct *electron, jets and missing transverse energy* objects. It is possible to create a set of PhysicsTowers (PhysicsTowerData) from Monte Carlo particles or just plain four-vectors. The electromagnetic and hadronic E_T in each tower of the CDF detector can be accessed using an *integer representation (ieta, iphi)* for the $\eta - \phi$ segmentation of the detectors. There are 9 types of towers in the calorimeter, depending on the rapidity of the tower. The classification of the TowerType is based upon:

1. The number of detectors per tower, e.g. TowerType 1 has CHA and WHA detectors, while TowerType 0 has only a CHA hadronic detector.
2. Different (η, ϕ) granularity of the towers.

The table 4.1 summarizes the characteristics of the tower types [3].

Tower Type	eta tower	EM phi granularity	HAD phi granularity	EM calors	HAD calors	EM+PPR pmt	HAD pmt	IETA index WEST	IETA index EAST
0	0-5	24	24	1CEM	1CHA	2	2+0	25-20	26-31
1	6-7	24	24	1CEM	2CHA 2WHA	2+0	2CHA 2WHA	19-18	32-33
2	8-9	24	24	1CEM	1WHA	2+0	2	17-16	34-35
3	10	2×24	1×24	1PEM	1WHA	2+2	2	15	36
4	11	2×24	1×24	1PEM	2WHA 2PHA	2+2	2WHA 2 PHA	14	37
5	12-17	2×24	2×24	1PEM	1PHA	1+1	1	13-8	38-43
6	18-20	24	24	1PEM	1PHA	1+1	1	7-5	44-46
7	21	24	24	1PEM	2PHA 2MPC	1+1	1	4	47
8	22-25	24	24	1MPC	1MPC	0	0	3-0	48-51

Table 4.1: The characteristics of the tower types. A tower type with a granularity of 2×24 in the electromagnetic (EM) section of the calorimeter means there are two towers per 15 degrees opening angle in phi. In case of the tower type 3 and 4, they have only one hadronic section, that is shared between the two towers.

The correspondence between η and IETA is given in the Table 4.2. The tower with IPHI = 0 starts at $\phi = 0$ and IPHI increase as the azimuthal coordinate ϕ . The segmentation in phi depends on the tower type, as shown in Table 4.1.

IETA(E +)	IETA(W -)	$ \eta $ range
26	25	-0.0000-0.1308
27	24	0.1308-0.2595
28	23	0.2595-0.3841
29	22	0.3841-0.5033
30	21	0.5033-0.6162
31	20	0.6162-0.7226
32	19	0.7226-0.8225
33	18	0.8225-0.9160
34	17	0.9160-1.0036
35	16	1.0036-1.1000
36	15	1.1000-1.2000
37	14	1.2000-1.3170
38	13	1.3170-1.4153
39	12	1.4153-1.5231
40	11	1.5231-1.6426
41	10	1.6426-1.7770
42	9	1.7770-1.9311
43	8	1.9311-2.1119
44	7	2.1119-2.3313
45	6	2.3313-2.6113
46	5	2.6113-3.0001
47	4	3.0001-3.6425

Table 4.2: The correspondence between η and IETA.

4.2 Track reconstruction

Track reconstruction in CDF is performed using the information provided by the tracking detectors, which are placed in a magnetic field. Via curvature, one can determine the charged particle momenta and discriminate between positively and negatively charged particles. The neutral particles do not leave traces in the tracking detectors. The algorithms used to reconstruct the tracks depend on the tracking detectors used (COT-only, Silicon-only or both), on the pseudorapidity region (central region, with $|\eta| < 1.0$ or forward region) and on the requirement that the track finding is *seeded*⁴ or not. Thus, the following tracking

⁴The tracking basically consist in grouping detector hits into tracks. *Seeded tracking* means that there are already available a set of paths or extrapolated tracks (seed tracks), along which the tracking algorithm will

algorithms were developed in CDF:

- COT-only: histogram tracking (HL)[76] and segment linking tracking (SL) [79];
- Silicon-only: standalone tracking [16] and Phoenix tracking (PHX) [96], the last used only in the forward region, for electrons only;
- COT+Silicon: outside-in [62] and inside-out [102].

For this dilepton cross-section measurement, *COT-only tracks* are used in the central region of the detector, a combination of histogram and segment linking, hit-based, unseeded algorithms; in the forward region, Phoenix tracking, a seeded algorithm, is used. The usage of COT-only tracks in the central region, without silicon information, was preferred, because over time, various beam incidents limited the performance of the silicon detectors. In both cases, to get the reconstructed tracks, a three-dimensional, five parameter fit to a track-helix is performed. Each track is characterized by five track parameters, described below:

- curvature: the 2-D curvature of the track, transverse to the beamline, which is inversely proportional to the transverse momentum, p_T , of the track.
- d_0 , impact parameter: the distance of closest approach of the track to the interaction vertex, in the transverse plane.
- $\cot \theta$: the cotangent of the polar angle θ .
- z_0 : the z coordinate of the point of the closest approach of the track to the interaction vertex, in the transverse plane.
- ϕ_0 : the ϕ direction of the track at the point of closest approach.

In this thesis, all track parameters, but $\cot \theta$, are used in identification of various objects. The details of the tracking algorithms used in this paper are discussed next.

The COT Pattern Recognition Algorithm

try to find hits and fit them to tracks. This improves the tracking reconstruction time in the very busy environments.

The track reconstruction starts from the hits in the tracking chamber. Then a pattern recognition algorithm[21] is used, in the following order:

- Find the *track segments* for axial and stereo superlayers: First, the 3-hits from consecutive wires are used to make *segment seeds*. All good segment seeds from an initial hit are used to search for hits on the remaining wires, and only the best segment, with the largest number of hits, associated with a hit, is kept.
- Link the axial segments in *r- ϕ tracks*: The axial pattern recognition consists of two algorithms, Segment Linking (SL) and Histogram Linking (HL), which complement each other very well, to give a high reconstruction efficiency. The axial SL links the axial superlayer segments into 2-D tracks, in r- ϕ plane. HL begins from a seed segment position and the beam position, and histograms the hits from the other superlayers as a function of the curvature. Therefore, the hits corresponding to a track will lie in the same bin and are identified as part of a track.
- Attach the stereo information to make *3-d tracks*: The stereo pattern recognition consists of two steps: Segment Linking and Hit Linking. The first one, Segment Linking consist of attaching stereo segments to 2-D axial tracks. Once this is done, the event vertices are reconstructed using z_0 of the event tracks. Hit Linking uses the vertices as seeds to try associating stereo hits to axial tracks which failed the Segment Linking step. At the end, a full 3-D 5-parameter fit is performed for the tracks with at least 2 stereo and 12 axial hits.

The COT tracking efficiency, determined using a W sample, triggered by the central electron calorimetry cluster and missing energy, was found to be $(98.3 \pm 0.12 \%)$ [60].

The Phoenix Algorithm

The Phoenix tracking algorithm was designed for forward ($|\eta| > 1.2$), high- p_T electrons, in a region where the COT provide very limited coverage. It makes full use of calorimeter and silicon information to produce charge-identified tracks used either to select electrons or veto electromagnetic jets which could fake a photon. The algorithm makes use of three elements: the 3-D event primary vertex, discussed later; the calorimeter cluster position of

the electromagnetic object, determined using the shower max information, and the transverse electromagnetic energy of the object, which determines the absolute value of the curvature. These two points and a curvature define two helices, due to the curvature sign ambiguity. The two possible paths are used as seeds in a pattern recognition algorithms, which will try to associate them silicon hits (Figure 4.2). The outcome could be 0, 1 or 2 tracks, known as *Phoenix tracks*.

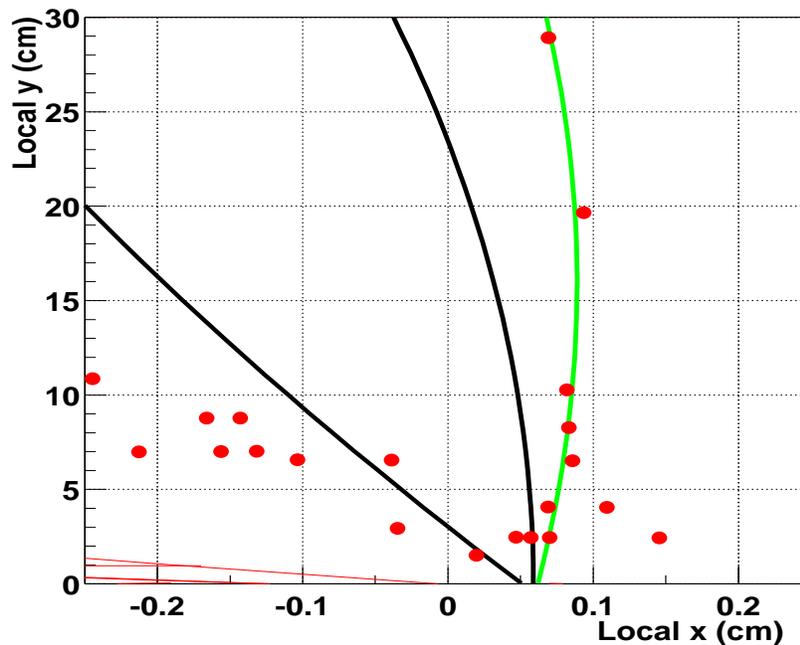


Figure 4.2: An example of a Phoenix track reconstruction. The black tracks are the seed tracks used as input for the pattern recognition. The red dots are the silicon hits, while the green track is the resulting silicon track.

When two tracks are found, the one with the best χ^2 is chosen as the best Phoenix track associated with the cluster. Due to the busy environment in the forward region, the track curvature sign could be misidentified. The track charge misidentification rate is dependent on the track's η , and also depends on the track quality requirements. For this analysis, the average track charge fake rate over $1.1 < \eta < 2.0$ is about 15%, as discussed in Section 6.11.

4.3 Vertex finding

A top dilepton event has two high- p_T leptons in the final state, *i.e.* two high- p_T tracks. Therefore, for the event primary vertex, the z_0 associated with the highest p_T lepton is used. Every event is required to have at least one lepton in the central region. In case one lepton is in the forward region, the central lepton is used to determine the event vertex, as the forward leptons have a larger uncertainty of the z_0 . The difference $|z_{01} - z_{02}|$ between the two leptons in each event is required to be less than 5 cm.

4.4 Electron reconstruction

An electron candidate[85] is built in a few steps, as described in the following:

- Make clusters from the individual towers;
- Make an electromagnetic object, basically a photon or electron candidate, by using additional information from various detectors.

Electromagnetic Clustering

The clustering begins with a grid of (η, ϕ) towers, described in Section 4.1. The various corrections are already applied at this time, and they are discussed later in this section. The EM clustering begins with all the *clusterable towers*, with a transverse energy greater than 100 MeV, sorted by the EM E_T . Out of this, the towers which have more than 2 GeV of EM transverse energy are defined as *seed towers*, ordered by the EM E_T . A cluster is formed from a seed tower and neighboring clusterable towers, known as *shoulder towers*, satisfying few extra requirement, depending on the pseudorapidity region (CEM or PEM). Also the shoulder towers must be located within the same detector as the seed tower, so a PEM tower cannot be part of a cluster seeded by a CEM tower and vice versa. Clustering approach varies between the two detectors as their geometry is different.

In the Central Electromagnetic Calorimeter, the shoulder towers must be within the same wedge ($\Delta\phi = 0$) as the seed tower. The maximum cluster size is 3 towers in η ($\Delta\eta = 1$)

and 15° in ϕ , and can at most consist of the seed and its nearest neighbors in η on either side of it. The clustering begins from the highest E_T seed, so any lower E_T seed tower candidate could become a shoulder tower, in which case it is being removed from the seed tower list. Also a cluster has to be highly electromagnetic; the hadronic energy must be less than 12.5% of the electromagnetic energy⁵, to be considered as an EM candidate. This helps reduce the clusters produced by hadronic jets. In case the cluster energy is above 100 GeV, it is accepted without these additional requirements.

In the Plug Electromagnetic Calorimeter, the correspondence between physical and logical towers is not fixed. Most of the CalData towers in the plug have 48 fold phi symmetry (which means 48 towers in ϕ , each of 7.5° , see Table 4.1 for details). Only for η towers 10 and 11, the logical ϕ segmentation is 24, so that it matches the WHA tower segmentation behind it. Having the same electromagnetic and hadronic tower segmentation helps with electron identification. So, most of the plug towers have the same logical and physical segmentation.

There are up to 8 neighboring physical towers to a seed tower, each of which required to have $E_T > 100$ MeV, to be part of a cluster. The largest E_T daughter tower is added to the seed tower, to form a cluster. Next, all possible pairs of neighboring towers are considered to be combined with the already existing cluster, to form a 2×2 cluster. The pair with the largest E_T is selected for the EM cluster. So, if no pair with towers $E_T > 100$ MeV is found, the cluster will have only two towers. However, this happens rarely, as the typical cluster has 2×2 configuration.

Electromagnetic Object Creation

Once an EM cluster is found, the next step is to build a CdfEmObject, by adding a track matching the cluster (for CEM clusters only), pre-radiator (CPR or PPR) and shower-maximum calorimeter information (PES or CES) for the photon/electron candidate. For the PEM clusters, there is no track associated with the cluster, by default. There are two algorithms used to construct a CdfEmObject, namely the *CalDrivenEmObject* and *CalOnlyCdfEmObject*, their names being suggestive for what each does. For this analysis, the CalDrivenEmObject algorithm is used, and it will be briefly detailed below.

⁵The cluster EM energy is used, but only the hadron energy in the seed tower and its pair tower (the pair forms a *trigger tower*) is used to calculate the hadronic energy.

For each CdfEmObject with a CEM EM cluster, each *track* is iteratively extrapolated to the plane of the CES, for the wedge containing the associated cluster. The extrapolation assumes the particle follows a helix, in the CDF solenoidal magnetic field. The final extrapolated track is required to be within 25 cm of the tower center in r - ϕ and within 38 cm of the seed tower center in z direction. This requirements correspond slightly to more than three physical towers in CEM. The cut in z allows about 2 cm beyond the boundary of a cluster shoulder tower. Also a matching track must transverse a CEM tower in the wedge of interest. In case more than one track fulfills the above criteria, the highest p_T track located within the seed tower or not further than 5 cm in z , beyond the seed tower boundary, is elected as the “best matching” track. Also some track quality criteria must be satisfied for the “best matching” track, such as a minimum number of COT axial and stereo hits.

For the PEM CdfEmObjects, a track is associated using the Phoenix algorithm, described before. This algorithm starts from the EM cluster and the primary vertex, to look for silicon hits along the road between the two points.

Next the algorithm attempts to associate *shower max clusters* with the CdfEmObject. In CEM, there are two collections of CES clusters: *an unbiased collection*, created from a list of wires or strip seeds over a threshold energy and used in the creation of photon objects, and *a track based collection*, that is seeded by the nearest wire or strip to an extrapolated track and it is used for electrons. If the CES cluster is track-based, the track seeding the CES cluster is required to be one of the tracks previously associated with this CdfEmObject. The best matching CES cluster is chosen as the one seeded by the best matching track. This CES cluster will be used in the selection of an CEM electron, for the current analysis. The *pre-shower clusters* are associated in a similar way as the CES clusters, however their are not used in the electron selection criteria.

Electron Corrections

Various corrections have to be applied to electron variables during the process of taking the raw data and constructing higher level variables. They are briefly described below.

A. Event Vertex Correction

The default variables dependent on θ assume that the interaction point is $z=0$ cm. We reprocess data offline and recalculate all the θ dependent variables, using the primary vertex described in Section 4.3.

B. CEM Energy Corrections

The raw CEM energy needs to be corrected for various inefficiencies before being used to select electrons/photons. The following corrections are applied:

- *face corrections*: The calorimeter response depends on the local x and z position within a tower. These corrections were extracted using the test-beam data taken in 1994 [64] and gets applied to both data and simulation. Also, extra corrections due to the attenuation of light passing through the scintillator in its way toward the wavelength shifters, or other smaller corrections, are applied.
- *tower-to-tower gain variations corrections*: These corrections maintain the calibrations to better than a percent, and significantly improve the energy resolution. These corrections have been determined using a central electron sample, with trigger threshold at $E_T > 8$ GeV[69]. The goal is to have an uniform energy scale over all the CEM towers. The gain correction factor $corr_i$, for tower i , is defined as the average $\langle E/p \rangle$, over the window 0.8-1.25. Then, each electron transverse energy, for tower i , becomes $E'_i = E_i/corr_i$, and the process is repeated until it converges. This correction improves the energy resolution by about 5%.
- *time-dependent gain corrections*: These corrections make the energy scale uniform in time. The correction factors[69] are extracted from the average $\langle E/p \rangle_j$, in the range 0.9-1.1, over all the towers, such as $\langle E/p \rangle$ becomes the same over all the runs j . There is still a degree of freedom left, the overall energy scale. This is chosen, such as the dielectron invariant mass peaks at the PDG value, 91.0 GeV.

The above corrections set the CEM energy scale.

C. COT Beam-Constrained Tracking (BC)

The COT resolution of the raw tracks is significantly improved by beam constraining the tracks, that is refitting them, such as they pass through the beam-spot[77]. All of the tracks used in this paper are COT-only, BC tracks.

Even after the tracks are beam-constrained, there is a systematic bias in momentum measurement of positively versus negatively charged particles, as a function of the azimuthal angle, ϕ . A correction known as *false-curvature* correction[67] is applied to the signed curvature as following:

$$\frac{q}{p_T^{corr}} = \frac{q}{p_T} - 0.00037 - 0.00110 \times \sin(\phi + .28) \quad (4.1)$$

where q is the charge of the particle (electron or positron), p_T is the COT BC transverse momentum. The correction is in fact a multiplicative factor for the total momentum, p . The track angles, θ and ϕ , are left unchanged.

The identification requirements of an electron, for the current analysis, are detailed in Section 6.2.1.

4.5 Muon reconstruction

A muon candidate[57] is created using the following procedure:

- Make stub,⁶ using the hits in the muon chambers;
- Make a muon object, CdfMuon, by matching the stubs to tracks.

Stub Finding

The reconstruction of muons begins with hits in the muon chambers.

The chamber's TDCs measure a drift time, which gets converted to a drift distance using *D to E algorithms*.⁷ Further, to relate the drift distance to a particle position, a *drift model*

⁶A stub is a muon track segment, in the muon chambers.

⁷The drift distance is the distance from the wire to location the muon occupied in its flight.

is used. The CMU and BMU measure both the drift distance and the position of the muon along the sense wire, “z” distance, while CMP and CMX only measure the drift distance.

Also the run-dependent calibration constants⁸, detector-position constants, and noise suppression are applied to the raw data from the detector. Once the clean up is over, the hits are used in a pattern-recognition algorithm to find stubs in the muon chambers. The output of the fit is a stub position and the direction of the stub.

Muon Object Construction

Once the stubs are found, the next step is to construct a CdfMuon object by adding the tracks matching a stub. This is done by extrapolating the tracks (COT-only, in our case) to the stubs, taking into account the inhomogeneous magnetic field and dense material. The best matching track is the closest one in $r\text{-}\phi$. A CdfMuon object will also carry information about the energy deposition in the calorimeter. However, there could be CdfMuons, with a track, but without a stub associated with it. These are known as *stubless muon* candidates. They are later required to be minimum ionizing, and are used in the present analysis.

Muons are also corrected by using the beam-constrained tracking, and the calorimeter transverse energies use the event primary vertex.

4.6 Jet reconstruction

Once the event primary vertex is known, the *hot towers*⁹ removed, and the tower level energy corrections are applied, we are left with an $(\eta - \phi)$ array of tower energies (accessed in PhysicsTowerData). Starting from this, a *clustering* process is used to reconstruct the jets. The jets are the cluster of hadrons arising from the fragmentation of final state “partons” from hard collisions within the $p\bar{p}$ interaction. Clustering algorithms selects a set of calorimeter towers, that are taken to be *particles*, which are typically emitted close to each other in angle, and combines their momenta to form the momentum of a jet. The selection process is

⁸The muon-chamber constants are drift-model parameters, used to determine a corrected drift time, drift distance, and, when possible, a hit z-position.

⁹A hot tower is a noisy electronic channel.

called the *jet algorithm* and the momentum addition rule is called the *recombination scheme*. The main goal for the clustering is to “undo” the parton fragmentation and associate the clusters of energy deposition in the calorimeter with jets, such that the kinematic properties of the jets (e.g. 4-momenta) can be related to the corresponding properties of the energetic partons produced in the hard scattering. The 4-momentum $p^\mu(\mathbf{E}, \mathbf{p})$ of each jet is computed assuming that the energy in a tower comes from *massless particles* striking the center of the tower, and summing over all the towers i , 4-momenta, that are part of the jet cluster:

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

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

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

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

The cluster of towers associated with a jet are typically selected to be within a fixed size cone in $\eta - \phi$ space from a point within the cluster. The distance in $\eta - \phi$ space from this point is defined as:

$$R = \sqrt{\Delta\eta^2 + \Delta\phi^2}. \quad (4.6)$$

A typical cone size is $R = 0.4$.

Using the variables defined in equations 4.2-4.5, various quantities, such as the jet $p_T \equiv \sqrt{p_x^2 + p_y^2}$, $E_T \equiv p_T \frac{E}{p}$ and $\eta \equiv \frac{1}{2} \log \frac{E+p_z}{E-p_z}$, are easily extracted. All of the above quantities are based on the raw calorimeter energies; the only corrections applied are at the tower level, before any clustering. The jets used with the raw energy and 4-momentum are referred as *uncorrected* jets. However, there are a variety of effects which cause the jets to be mismeasured, such as:

- calorimeter non-linear response;
- reduced calorimeter response near the boundaries of the different calorimeter towers;

- loss of low momentum particles inside the magnetic field;
- energy deposited in the towers outside the clustered jet (eg: outside the cone $R < 0.4$);
- contribution from the underlying event (beam remnants) or additional interactions, and
- energy loss due to muons and neutrinos (eg: semileptonic decays of b quarks).

The standard corrections are meant to reproduce the average E_T correctly, but they cannot reduce the jet fluctuations around the mean E_T . In analyses trying to reconstruct heavier particle masses which produce jets, there are extra corrections applied. Typical jet correction increases the jet energy by 20-30 %. The jet corrections applied to the jets used in this analysis are discussed in section 6.9.

In Run II, the jet clustering could be done by using a variety of *cone algorithms* like JetClu (the main algorithm used in Run I, cone-based algorithm), SeedlessCone, MidPoint or *K_T algorithms*, like K_T Clus, all already implemented. This analysis uses JetClu, which is described below.

The CDF JetClu clustering algorithm in Run II

The CDF jet clustering algorithm, being a cone algorithm, forms the jets by collecting particles whose trajectories (i.e. towers whose centers) lie within a circle of *fixed radius* R in η - ϕ space, around a selected *seed* point. Initial seed points are selected as centers of regions of the size of a tower (0.1 x 15 deg in η - ϕ coordinates) of high E_T , in the calorimeters. The minimum energy of a seed, $E_T(EM + HAD)$, is required to be above a few hundred MeV. This procedure has the advantage of saving computing time. On the other hand, the inclusion of the seeds procedure, along with merging/splitting jet rules, causes difficulties in comparisons with QCD theoretical predictions related to jets at fixed perturbative order. These are the infrared sensitiveness and collinear problems([33]). The jet selection for the top dilepton analysis is discussed in Section 6.9.

4.7 Missing transverse energy (\cancel{E}_T) reconstruction

The neutrinos are very weakly interacting particles and since there is not enough material, they escape the detector. Because the protons and anti-protons are moving along the beam-line (z axis), we assume that the system of hard scattered partons has zero transverse momentum. This is a very good approximation, even if is not always true. Since the longitudinal component of the colliding partons is unknown, only the neutrinos transverse momentum is measured from transverse momentum conservation, which is initially zero. The transverse energy imbalance in the detector is proportional to the neutrino momentum and it is called *missing* E_T or \cancel{E}_T . The raw \cancel{E}_T is defined as the negative vector sum of the transverse energy in the calorimeter, both electromagnetic and hadronic:

$$\vec{\cancel{E}}_T (\cancel{E}_{T_x}, \cancel{E}_{T_y}) = - \sum_i [(E_i^{em} + E_i^{had}) \sin\theta_i] \times \hat{n}_i \quad (4.7)$$

where E_i^{em}, E_i^{had} are the electromagnetic, respectively hadronic energy of the i^{th} tower, and \hat{n}_i is a transverse unit vector pointing to the center of each tower. The θ_i is the polar angle of the i^{th} tower with respect to the z=0 coordinate frame. The sum in the above equation extends to $|\eta| < 3.6$, to avoid the forward calorimeter region. The magnitude of the \cancel{E}_T is defined as usual:

$$\cancel{E}_T \equiv \cancel{E}_T |\vec{\cancel{E}}_T| = \sqrt{\cancel{E}_{T_x}^2 + \cancel{E}_{T_y}^2} \quad (4.8)$$

The \cancel{E}_T defined before assumes that the collision took place at z=0 and it is known as *raw* \cancel{E}_T . There are few corrections which are applied, before using \cancel{E}_T for an event selection, as described in Section 6.7. For example, if there are high- p_T muons in the event, they are part of the \cancel{E}_T correction

The resolution of \cancel{E}_T depends on the response of the calorimeter to the total energy deposited in the event, and therefore to the event topology. So, the \cancel{E}_T depends, to first order, on the total scalar sum E_T of the event, $\sum E_T$.

5 Data Samples and Triggers

This section describes the samples used by the top dilepton analysis, including the data clean-up criteria, the luminosity calculation and the associated triggers.

For this analysis, three main datasets are used:

- Tight electron dataset, stripped [95] from the inclusive electron dataset,
- Tight muon dataset, stripped [95] from the inclusive muon dataset, and
- Plug electron+missing energy dataset.

The parent datasets of these samples and the associated trigger paths are described in Sections 5.3 to 5.5. They represent data taken between March 2002-September 2003, with a luminosity between 150-193 pb^{-1} .

5.1 Good Run Criteria

Out of all the data written to tape, only a subset of runs, passing the so-called *Good Run Criteria* are used for analysis. Only runs with at least $10nb^{-1}$ are included. CDF has a database which keeps track of the status of all the detector components during a run. Using the database information, the Data Quality Monitoring Group (DQM) provides the final list

of good runs for physics. We use DQM version 4 good run list (GRL4) (See [36]), selected from the full list of runs taken between March 2002-September 2003. The following runs are also excluded from the default GRL4:

- 163463 and 163474: They had wrong SVX BeamLine in calibration pass 9, and
- 164844, 164870, 164871 and 164872: The raw data partially disappeared due to a CSL problem.

Dilepton Working Group decided to use the runs 146805 - 148157, even if DQM Web Page mentions that in this run range some cables were swapped in CES. If we were to exclude them, their effect on the cross-section result is negligible. No candidate events are in this run range.

5.2 Luminosity Measurement

The luminosity is measured using the Cerenkov Luminosity Counters (CLC). To find the luminosity for a data sample, the relevant offline Data File Catalog (DFC) entries for each of the good runs are added.

The total integrated luminosity is calculated using the formula:

$$\mathcal{L} = \frac{R_{p\bar{p}}}{\sigma_{in} \times \epsilon_{clc}}, \quad (5.1)$$

where $R_{p\bar{p}}$ is the rate of the $p\bar{p}$ inelastic collisions measured with the luminosity monitor (CLC) [43], ϵ_{clc} is the CLC acceptance and σ_{in} is the $p\bar{p}$ inelastic cross section.

We have used $\sigma_{in} = 61.7\text{mb}$ at $\sqrt{s} = 1.96\text{ TeV}$ for the luminosity estimation [90], obtained by scaling the CDF measurement $\sigma_{in} = 60.4 \pm 2.3\text{ mb}$, at 1.8 TeV to 1.96 TeV.

The CLC acceptance, ϵ_{clc} , is estimated using data and simulation using:

$$\epsilon_{clc} = \frac{N_{CLC+Plug\ tagged\ inelastic}}{N_{inelastic}} \cdot \frac{N_{CLC\ East-West\ Coincidence}}{N_{CLC+Plug\ tagged\ inelastic}} \quad (5.2)$$

where $N_{inelastic}$ is the number of inelastic collisions, $N_{CLC+Plug\ tagged\ inelastic}$ is the subset of inelastic collisions tagged by the CLC and the plug calorimeter, $N_{CLC\ East-West\ Coincidence}$ represents the subset of events which are simultaneously tagged by both East and West counters. The fraction $\frac{N_{CLC\ East-West\ Coincidence}}{N_{CLC+Plug\ tagged\ inelastic}}$ is determined from the data, while the other factor in ϵ_{clc} is extracted from the simulation. The acceptance is estimated to be $(60.2 \pm 2.4)\%$.

The quoted uncertainty is dominated by the uncertainty in the absolute normalization of the CLC acceptance (ϵ_{CLC}) for a single $p\bar{p}$ inelastic collision and by the inelastic cross section. The complete list of systematic uncertainties is shown in Table 5.1.

Effects	Uncertainty Estimate ()
CLC Acceptance	4.4
Inelastic cross-section (CDF)	2.5
Detector instability	2.0
Detector calibration	2.5
Online to offline transfer	1.0
Total uncertainty	5.7

Table 5.1: Systematic uncertainties for the luminosity calculation.

By requiring that the run belongs to the “Good Run List” described in section 5.1, the dilepton categories will fall in one of the four luminosity classes, shown in Table 5.2.

Sample	Luminosity (pb^{-1})
CEM/CMUP	193
CEM/CMUP and CMX	175
CEM/CMUP and Si	162
CEM/CMUP and CMX and Si	150

Table 5.2: Summary of luminosities used for the cross measurement. A 6 % systematic uncertainty is assumed for each luminosity.

5.3 Inclusive Electron Sample

The inclusive high- p_T central electron sample (known as $btop0g/0j$, see Table 5.3) consists mainly of the leptonic decays of the W and Z bosons. Also this sample will have events with high p_T jets, produced in QCD processes, which could mimic a high energy electromagnetic cluster. Out of the inclusive sample, a smaller subsample, called *Tight Central Electron Sample* (TCES) is stripped, by requiring a good electron candidate in each event, selected with tight requirements. There are 1119266 events in the tight sample.

Inclusive Central Electron Sample		
DATASET TYPE	NAME	FILTERING/REQUIREMENTS
L3 Trigger	<i>bhel08/09</i>	ELECTRON70_L2_JET ELECTRON_CENTRAL_18 ELECTRON_CENTRAL_18_NO_L2 W_NOTRACK W_NOTRACK_NO_L2 Z_NOTRACK
Primary	<i>btop0g/0j</i>	loose CEM electron cuts
Secondary	Tight Electron	≥ 1 tight electron (baseline CEM e cuts)

Table 5.3: Parent datasets and associated L3 trigger paths for the tight electron data sample.

Only the events passing at least one of the trigger requirements described below will be selected for our analysis.

5.3.1 Central Electron Trigger: CEM_18

This electron trigger selects events with a high transverse energy electromagnetic cluster, matched to a track (potentially an electron). The trigger path consists of the following requirements:

Level-1

A trigger tower is $0.2 \times 15^\circ$ in η - ϕ space. It is required that a trigger tower has a transverse energy, $E_T > 8\text{GeV}$ and the ratio $E_{had}/E_{em} < 0.125$, in case E_T is less than 14 GeV. Also

a COT track, reconstructed by XFT[25], has to match the trigger tower, with transverse momentum, $p_T > 8\text{GeV}/c$.

Level-2

The clustering is expanded and the adjacent towers with $E_T > 7.5$ GeV are added to the trigger tower found in Level 1 (also known as the “seed” tower.). The total energy of the *new cluster* has to be above 16 GeV, and also $E_{had}/E_{em} < 0.125$. A XFT track is again required to match the seed tower of the cluster and have $p_T > 8$ GeV/c.

Level-3

After full event reconstruction, an electromagnetic cluster with $E_T > 18$ GeV and $E_{had}/E_{em} < 0.125$ is required. A 3-D COT track, matching the cluster, with $p_T > 18$ GeV/c, is required.

5.4 Plug Electron+Met Sample

The inclusive high- p_T plug electron sample consists mainly of W -boson events, with the W decaying leptonically and QCD multijet events. The QCD multijet events dominate the sample. They consist of a jet passing the plug electron requirements and a second mismeasured jet that gives the missing E_T . The trigger paths associated with this dataset are shown in Table 5.4.

Out of the total sample, the events passing the trigger requirements described below will be selected for our analysis.

Plug Electron Sample		
DATASET TYPE	NAME	FILTERING/REQUIREMENTS
L3 Trigger	<i>bpel08/09</i>	MET_PEM_L1_EM8.&_MET15_v4 MET_PEM_v10
		PLUG_ELECTRON_20_v5
		PLUG_Z_v3
Primary	<i>unnamed</i>	MET_PEM_L1_EM8.&_MET15_v4 MET_PEM_v10

Table 5.4: Parent datasets and associated L3 trigger paths for the plug data sample.

5.4.1 Plug Electron Trigger: MET_PEM

This plug electron trigger select events with a high transverse energy electromagnetic cluster (potentially an electron or photon) and large transverse missing energy. In the current analysis, this trigger is only used to collect the $e\text{-}\mu$ dilepton events with a Phoenix electron and a non-CMUP or non-CMX muon. The trigger path consists of the following requirements:

Level-1

A trigger tower with a transverse energy, $E_T > 8$ GeV and the ratio $E_{had}/E_{em} < 0.125$, in case E_T is less than 14 GeV, is required. The raw missing energy, calculated assuming the interaction point at $z=0$ cm, has to be greater than 15 GeV.

Level-2

A forward cluster is required, with the electromagnetic transverse energy $E_T > 20$ GeV and $E_{had}/E_{em} < 0.125$.

Level-3

After the full reconstruction, a PEM cluster is required to have $E_T > 20$ GeV and $E_{had}/E_{em} < 0.125$. Also the raw missing energy $\cancel{E}_T > 15$ GeV is required.

5.5 Inclusive Muon Sample

Inclusive high p_T muons (known as *btop1j/1g*, see Table 5.5) come mainly from leptonic decays of W or Z bosons, or rarely from a QCD jet faking a muon object.

Inclusive Muon Sample		
DATASET TYPE	NAME	FILTERING/REQUIREMENTS
L3 Trigger	<i>bhmu08/09</i>	MUON_CMUP18 MUON_CMX18
Primary	<i>btop1g/1j</i>	loose muon cuts
Secondary	Tight Muon	≥ 1 tight muon (baseline μ cuts)

Table 5.5: Parent datasets and associated L3 trigger paths for the tight muon data sample.

5.5.1 Central Muon Triggers: CMUP_18 and CMX_18

This muon triggers selects events with a high transverse momentum track matched with a stub in the muon chambers (potentially a muon). There are *two triggers* which collect these muons, one requiring one muon to have stubs both in CMU and CMP (a CMUP muon, $|\eta| < 0.6$), the other requiring a stub only in CMX ($0.6 < |\eta| < 1.0$).

The CMUP trigger path (CMUP_18) consists of the following requirements:

Level-1

Hits in the CMU and CMP detectors, with arrival times within 124 ns, are required. Also an XFT track with $p_T > 4\text{GeV}/c$ must be matched to the hits.

Level-2

In the earlier periods, no L2 requirements were made, as no L2 muon trigger was available. Lately an XFT track with $p_T > 8\text{GeV}/c$ has been required, but not required to match the CMU or CMP hits.

Level-3

The CMU and CMP hits are required to match a 3-D reconstructed COT track, with $p_T > 18\text{GeV}/c$.

The CMX trigger path (CMX_18) consists of the following requirements:

Level-1

The hits in the CMX, with arrival times within 124 ns, are required. Also an XFT track with $p_T > 4\text{GeV}/c$ must be matched to the hits. For the later runs, a matching hit in the CMX scintillator plane is also required.

Level-2

No additional requirements at this trigger level.

Level-3

The CMX hits are required to match a 3-D reconstructed COT track, with $p_T > 18\text{GeV}/c$.

5.5.2 Top Dilepton Triggers Summary

The triggers used to collect each dilepton category are shown in Table 5.6.

Category	Trigger Path Required
ee categories	
CEM-CEM	CEM_18
CEM-PHX	CEM_18
$e\mu$ categories	
CEM-CMUP	CEM_18 CMUP_18
CEM-CMIO/U/P	CEM_18
CEM-CMX	CEM_18 CMX_18
PHX-CMUP	CMUP_18
PHX-CMX	CMX_18
PHX-CMIO/U/P	MET_PEM
$\mu\mu$ categories	
CMUP-CMUP	CMUP_18
CMUP-CMIO/U/P	CMUP_18
CMX-CMIO/U/P	CMX_18
CMX-CMX	CMX_18
CMX-CMUP	CMX_18 CMUP_18

Table 5.6: The dilepton categories and the triggers used to collect the events. The details on the selection requirements for each category will be given in Chapter 6.

5.6 QCD Jet Samples

In addition to the primary datasets used to collect the dilepton candidates, QCD enriched datasets are used for background estimates. These QCD samples are needed to estimate the probability that a jet will mimic a lepton (e or μ). There are four such datasets, which we will denote by *jet20*, *jet50*, *jet70* or *jet100*, depending on the trigger jet thresholds used to collect them.

These triggers have similar trigger paths, except for the jet energy thresholds or the trigger prescales¹ used. They are summarized in Table 5.7.

¹A trigger is said to have a prescale of N , where N is an integer, if 1 of every N events passing some criteria is accepted, while the rest are rejected. A prescale is used to limit the number of events recorded for processes with very large cross sections, as the trigger bandwidth is limited.

Trigger Path	Trigger Name		
	Level 1	Level 2	Level 3
<i>jet20</i>	L1_JET5_PS20	L2_JET15_PS12	L3_JET_20
<i>jet50</i>	L1_JET5_PS20	L2_JET40	L3_JET_50
<i>jet70</i>	L1_JET10	L2_JET60_PS8	L3_JET_70
<i>jet100</i>	L1_JET10	L2_JET90	L3_JET_100

Table 5.7: Associated trigger names at each trigger level for the trigger paths *jet20*, *jet50*, *jet70* or *jet100*, used to collect QCD enriched datasets.

Level-1

At L1, the *jet20* and *jet50* triggers require a trigger tower with the total transverse energy E_T ² greater than 5 GeV. Both are prescaled, *jet20* at 20, while *jet50* at 12. The *jet70* and *jet100* triggers require a tower with total transverse energy greater than 10 GeV with no prescale.

Level-2

At L2, a more sophisticated jet clustering is performed. A jet with the total transverse energy E_T greater than 15, 40, 60, and 90 GeV for *jet20*, *jet 50*, *jet70* and *jet100*, respectively, is required. Also for *jet20* and *jet 70* triggers prescales of 12 and 8, respectively, are imposed in L2.

Level-3

In L3, the event is fully reconstructed and a jet clustering with a cone size of 0.7 is performed. The event's vertex is assumed to be at $z_{vertex}=0$ cm, for simplicity. A reconstructed jet is required, with E_T greater than 20, 50, 70 and 100 GeV, for *jet20*, *jet 50*, *jet70* and *jet100*, respectively.

²Total transverse energy E_T means in the context of QCD enriched triggers the sum of electromagnetic and hadronic energies.

6 Top Dilepton Event Selection

The key in measuring a cross-section is to separate the signal events from the rest, while efficiently rejecting the background events. The signature of a top dilepton event, as was already shown in Section 11, consists of two high- p_T leptons, two high- E_T jets, large missing energy (as the two neutrinos are undetected) and a large amount of energy flowing in the transverse plane, as a result of the top quark being so heavy. This chapter describes the dilepton selection step by step. The dilepton event selection begins by looking for events with two well identified high- p_T leptons, with at least one of them isolated from nearby calorimetry activity. Next, the missing energy related cuts are applied to further improve the separation of signal from backgrounds. The top dilepton events are required to have at least two jets. Events with lower jet multiplicities are used *as a control sample* as they are background dominated. Finally dilepton events are required to have large transverse energy flow, H_T . The signal acceptance is estimated in Chapter 7 for a top mass of 175 GeV¹, while the contribution from the background processes is discussed in Chapter 8.

¹This particular top mass was chosen based on the fact the combined Run I Tevatron top mass was 174.3 ± 5.1 GeV[94]. Recently a new best world top mass average of 178.0 ± 4.3 GeV[93] was presented.

6.1 Geometrical and kinematic requirements

The starting point in the event selection is to impose detector fiducial and kinematic constraints such as requiring the lepton energies or momentum to be above some threshold. This is to insure that the leptons and jets are in active regions of the detector and that they can be well reconstructed.

Electrons

This analysis includes electrons in the central and the plug calorimeter region. The selection criteria are similar, except for small differences, which will be pointed out.

Central electrons, denoted as *CEM* are required to have large transverse energy, $E_T > 20$ GeV. Also the electrons are required to be *fiducial* to the central calorimeter shower maximum detector (CES). This is basically a requirement that the CES cluster position is in a good coverage region, away from the uninstrumented or poorly instrumented regions of the the central calorimeter wedges. This insures that electromagnetic showers are properly reconstructed. The fiducial requirements are described below [70]:

- The z position of the CES cluster, z_{CES} , must be 9 cm away from the $z=0$ cm plane, to avoid the gap between the two central calorimeter halves. About 5% of central electrons will fail this cut. Also there is an upper bound requirement, $z_{CES} < 230$ cm.
- The shower position in CES, in the transverse plane, must be not further than 21.0 cm from the center of the wedge, to avoid the wedge ϕ boundary. A wedge at the CES radius spans $(-24.25, 24.25)$ cm in x_{CES} range. So about 13% of the electrons will be lost due to this requirement.
- In the CDF central calorimeter there is a wedge module, know an *the chimney module*, which is cut away to allow access to the superconducting solenoid. In this wedge there are only 7 complete towers (0-6), tower 7 is not full, while towers 8 and 9 are missing. The seed tower of the EMCluster could not be in Tower 7 of the chimney module (Wedge 05E), which corresponds to $0.77 < \eta < 1.0$, $75^\circ < \phi < 90^\circ$ and $z_{CES} > 193$ cm.

- Also the seed tower cannot be Tower 9, as it is incomplete².

Plug electrons are required to have $E_T > 20$ GeV and to have a reconstructed cluster in the PES, the plug EM shower maximum detector. The pseudorapidity of the PES cluster is constrained to the range $1.2 < |\eta| < 2.0$, as the fake charge rate for Phoenix (PHX) electrons increase with $|\eta|$. Also as part of the fiducial requirement, at least a *PHX track* is required to point to the cluster.

Muons

The muon candidates are required to have a COT-only track, with beam-constrained momentum, $P_T > 20$ GeV. Based on the existence of hits matching the high-momentum track, in case the muon passes through a muon chamber, or the muon is not fiducial to any chamber, the muons are classified in the following categories:

- CMUP: If the muon has stubs in CMU and CMP detectors only, matching the track. There are no extra fiducial requirements.
- CMX: If the muon has only a stub in CMX detector that matches the track. There are no extra fiducial requirements.
- CMU: If the muon has only a CMU stub and it is fiducial only to CMU.
- CMP: If the muon has only a CMP stub and it is fiducial only to CMP.
- CMIO: If the muon has no CMU, CMP, CMX or BMU stubs, and also it is not fiducial to any muon detector. A CMIO muon is also called a *stubless muon*.

The fiducial information is obtained using *MuonFiducialTool* [57], which extrapolates the muon track to the muon chambers and checks if the track intersects a particular muon detector or not. There is a drawback to this procedure because multiple Coulomb scattering is not included.

² The highest η tower, Tower 9, is shaped like a right triangle and near the outer end, such that there can be substantial energy leakage.

6.2 Lepton Identification

We select a lepton (electron or muon) using the baseline cuts proposed by the Electroweak/Top Groups[27], with small differences for categories which are used *only* by the dilepton analysis.

6.2.1 Electron Identification

There are two classes of electrons used in the analysis:

- Central Electrons: CEM
- Plug Electrons: PHX

The variables used to identify a central electron are almost identical (except for leakage corrected isolation) with the ones used in Run I:

- E_T : The transverse electromagnetic energy deposited by the electron in the CEM (central electromagnetic calorimeter) is calculated as the electromagnetic cluster energy multiplied by $\sin \theta$, where θ is the polar angle provided by the best COT track pointing to the EM cluster. An electron cluster is made from a seed EM tower and at most one more shoulder tower, passing some well defined requirements. The maximum cluster size could have two towers in pseudorapidity ($\Delta\eta \approx 0.22$) and one tower in azimuth ($\Delta\phi \approx 0.26 \text{ rad}$)
- P_T : The transverse momentum of the COT beam constrained track, as measured using the COT track curvature in the magnetic field.
- E_{had}/E_{em} : The ratio of the hadronic (CHA/WHA) calorimeter energy to the electromagnetic (CEM) calorimeter energy for the electron cluster.
- E/P: The ratio of the EM cluster transverse energy, E_T , to the COT track transverse momentum, P_T .

- L_{shr} : The lateral shower profile for electrons. This variable compares the energy in CEM towers adjacent to the seed tower between data and test beam electrons [86]. Mathematically, it is defined as:

$$L_{shr} = 0.14 \frac{\sum_i (E_i^{meas} - E_i^{pred})}{\sqrt{(0.14\sqrt{E_{EM}})^2 + \sum_i (\sigma_i^{pred})^2}}, \quad (6.1)$$

where the sums are over the towers in the EM cluster *adjacent to the seed tower* and in the *same wedge as the seed tower*; either a one or two tower sum. E_i^{meas} is the measured energy in an adjacent tower i , E_i^{pred} is the predicted energy deposit, by an electron, in the adjacent tower i , E_{EM} is the total electromagnetic energy in the cluster, and σ_i^{pred} is an estimate of the uncertainty in E_i^{pred} . This is *slightly different* than what is in Run I notes/theses ([66]), but it is what was actually used in the Fortran code and translated into the Run II C++ framework.

- $Q * \Delta x$: The distance in the r - ϕ plane between the extrapolated, beam constrained, COT track and the best matching CES cluster, multiplied by the lepton charge.
- Δz : The distance in the r - z plane between the extrapolated, beam constrained, COT track and the best matching CES cluster.
- χ_{strip}^2 : The χ^2 comparison of the CES shower profile in the r - z view with the same profile extracted from test beam electrons
- z_0 : The z intersection of the track with the beam axis in the r - z plane.
- Track quality cuts: The electron associated track must have passed through at least 3 axial and 3 stereo superlayers (SL), each with at least 7 hits out of 12.
- $\Delta z_{leptons}$: The distance along the beam axis between z_0 of the two leptons. Also the event's primary vertex, which is the z_0 of the highest lepton, is required to be less than 60 cm from $z=0$.

Distributions of these variables for electrons in the data and Pythia 4.9.1 $Z \rightarrow ee$ Monte Carlo are shown in Figures 6.1 and 6.2. The distributions are made using the second leg

of the $Z^0 \rightarrow e^+e^-$ data sample, where the first leg is required to pass the tight central electron requirements (Table 6.1) and the second leg must be fiducial (track fiduciality or FIDELE = 4) and $E_T > 20$ GeV. Also the invariant mass of the two electron system is required to fall inside Z mass range (75, 105) GeV. To further remove any background from the unbiased electron sample obtained, we apply all the ID cuts except the one plotted (These are known as “N-1” distributions).

Variable	Cut
Geometrical and Kinematical Requirements	
E_T (two-tower EM sum) * ($\sin(\theta)$ of BC track) Fiducial	> 20 GeV Fidele = 1 (Ces $ X < 21$ cm, $9 < \text{Ces } Z < 230$ cm, Tower 9 excluded most of tower next to chimney included)
Identification Requirements	
P_T (COT-only BC track) E/P (for $E_T < 50$ GeV only) Track $ z_0 $ E_{had}/E_{em} $Lshr$ (using track-based strip cluster, z0 of track) charge-signed $ \Delta X $ $ \Delta Z $ χ_{strip}^2 Track Type COT track quality	> 10 GeV/c 2.0 < 60.0 cm $< 0.055 + 0.00045 * E$ < 0.2 $-3.0 < Q_{track} * \Delta X < 1.5$ cm < 3.0 cm < 10 PADtrack ≥ 3 axial, ≥ 3 stereo SL with > 6 hits each
Isolation Requirements	
Fractional Calorimeter Isolation E_T (with PJW leakage Correction [78])	< 0.1

Table 6.1: Selection requirements for CEM electrons.

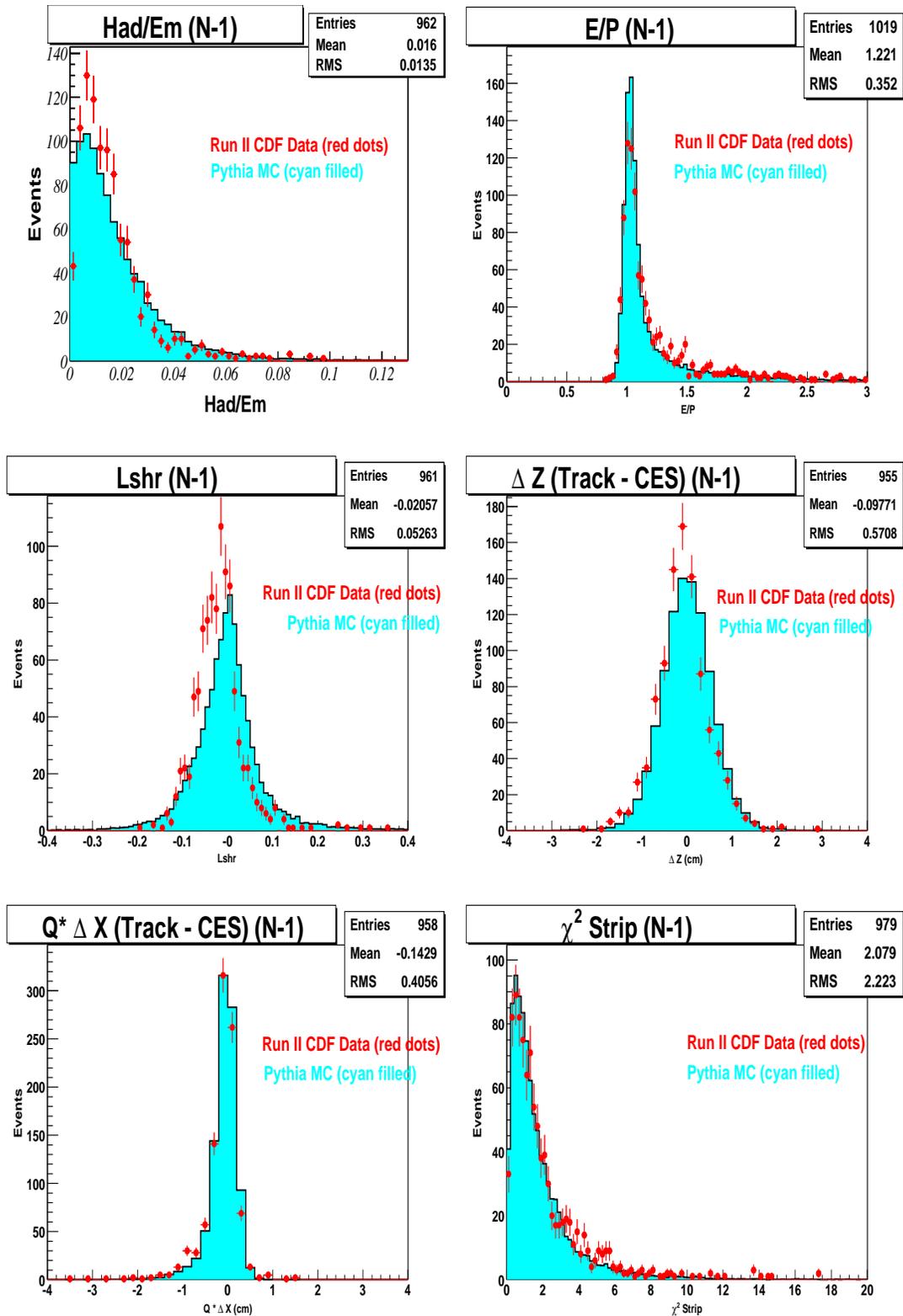


Figure 6.1: Tight central electron ID variables E_{had}/E_{em} , E/P , L_{shr} , ΔZ , χ^2_{strip} and $Q^* \Delta X$. These are 'N-1' plots and are made using $Z \rightarrow e^+e^-$ Run II data and compared with the Pythia MC sample.

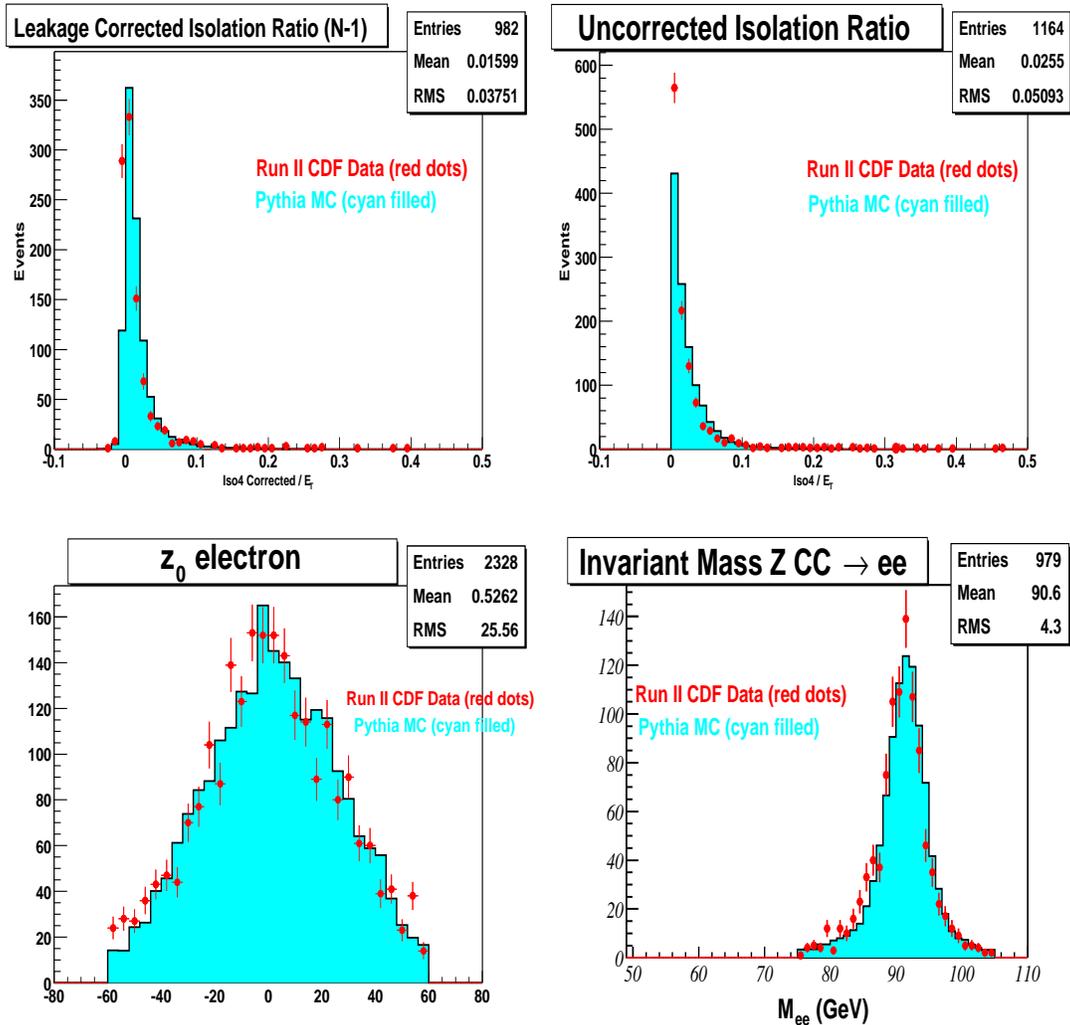


Figure 6.2: Tight central electron ID variables iso_{frac} , z_0 , and the dielectron invariant mass are shown. These are 'N-1' plots and are made using $Z \rightarrow e^+e^-$ Run II data and compared with the Pythia MC sample.

The variables used to identify the Phoenix plug electrons are briefly defined below [23]:

- **Fiducial cuts:** The electron PES cluster used for the analysis are required to have pseudorapidity in the window $1.2 < |\eta_{PES}| < 2.0$. The region beyond 2.0 is excluded because the slight increase in the acceptance is overwhelmed by the increase in the backgrounds.
- E_T : The transverse energy of the electromagnetic cluster in the PEM calorimeter.

- E_{had}/E_{em} : The ratio of the hadronic (PHA) calorimeter energy to the electromagnetic (PEM) calorimeter energy for the electron cluster.
- χ_{PEM3X3}^2 : It compares the electromagnetic shower profile of a given PEM cluster with the shower shape from the test beam data.
- PEM3X3FitTowers: It is the number of towers used by the 3X3 PEM cluster fit algorithm.
- PESProfileRatio5by9: It is defined as:

$$\frac{\text{Sum of the energy in the central 5 strips of a PES cluster}}{\text{Sum of the energy in the all strips (9) of a PES cluster}} \quad (6.2)$$

and it is a measure of how isolated the electron PES cluster is. The variable is defined for both U and V layers of a PES 2-d cluster.

- $\Delta R(PHXtrack - PES)$: The angular separation between the extrapolated PHX track at the PES and the position of the electron cluster in PES ($\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$).

The selection requirements for a Phoenix (PHX) plug electron are shown in Table 6.2³.

³The ID variables are compared between the data and simulation in various papers [29].

Variable	Cut
Geometrical and Kinematical Requirements	
E_T	> 20 GeV
Fiducial	$1.2 < \eta_{PES} < 2.0$
At least a PHX track	Yes
Identification Requirements	
E_{had}/E_{em}	< 0.05
PEM3X3FitTowers	$\neq 0$
χ^2_{PEM3X3}	< 10.0
PESProfileRatio5by9(U and V layers)	> 0.65
$\Delta R(PHXtrack - PES)$	< 3 cm
SiTrack Hits	≥ 3
Phoenix Track $ z_0 $	< 60.0 cm
Isolation Requirements	
Fractional Calorimeter Isolation E_T (with leakage correction [14])	< 0.1

Table 6.2: Selection requirements for PHX electrons.

6.2.2 Electron Efficiencies

As we have seen, the identification of an electron is comprised of many steps, starting from the trigger cuts and ending with the requirements from the Table 6.1. The Monte Carlo simulation does not always model the data very well, and this is corrected by applying scale factors.

Electron Identification Efficiencies

The electron identification efficiency is the probability that a high- E_T electron that passed the trigger cuts will pass the tight identification criteria (See [72; 47] for details). The results for each electron type are given in Table 6.3. ICEM refers to isolated central electron, NCEM refers to non-isolated central electron (iso < 0.1), and PHX to Phoenix plug electron.

Lepton Category	Data	MC	Id Efficiency Scale Factor
ICEM	82.5 ± 0.5	85.47 ± 0.16	0.965 ± 0.006 [72]
NCEM	75.0 ± 8.5	78.1 ± 1.8	0.96 ± 0.11 [72]
PHX	65.1 ± 0.8	74.9 ± 0.2	0.87 ± 0.01 [47]

Table 6.3: Electron identification efficiencies and the electrons scale factors (data/MC) used to correct the acceptance, for ICEM, NCEM and PHX.

6.2.3 Muon Identification

The muon identification variables that are used are detailed below and the cuts on them are listed in Table 6.4.

- P_T : The transverse momentum of the COT beam-constrained track, as measured using the track curvature in the COT in the magnetic field.
- E_{HAD} , E_{EM} : The energy the muon candidate deposits in the hadronic, respectively electromagnetic part of calorimeter.
- $|\Delta x|$: The distance in the r - ϕ plane between the extrapolated track and the muon stub at the chamber radius.
- d_0 : The muon track impact parameter. The cuts on d_0 are tighter or looser, depending on whether the track has silicon hits or not, respectively.
- Track quality cuts: The muon associated track is required to pass through at least 3 axial and 3 stereo superlayers (SL), each with at least 7 hits out of 12.

Variable	Cut	Category the cut only applies to
Geometrical and Kinematical Requirements		
P_T (COT-only BC track) Stubs	$> 20 \text{ GeV}/c$	
Fiducial	CMU and CMP only CMX only CMU only CMP only CMU only CMP only not to CMU, CMP, CMP, CMX	CMUP CMX CMU CMP CMU CMP CMIO
Identification Requirements		
E_{had} E_{em} $E_{em} + E_{had}$ Track $ d_0 $	$< 6 + \max(0, 0.028 * (p - 100)) \text{ GeV}$ $< 2 + \max(0, 0.0115 * (p - 100)) \text{ GeV}$ $> 0.1 \text{ GeV}$ $< 0.02 \text{ cm}$ (if track has Si hits) OR $< 0.2 \text{ cm}$ (if not)	CMIO
$ \Delta x_{CMU} $ $ \Delta x_{CMP} $ $ \Delta x_{CMX} $ Track Type Track $ z_0 $ COT track quality	$< 3.0 \text{ cm}$ $< 5.0 \text{ cm}$ $< 6.0 \text{ cm}$ PADtrack $< 60.0 \text{ cm}$ ≥ 3 axial, ≥ 3 stereo SL with > 6 hits each	CMUP or CMU CMUP or CMP CMX
Isolation Requirements		
Fractional Calorimeter Isolation	< 0.1	

Table 6.4: Selection requirements for muons.

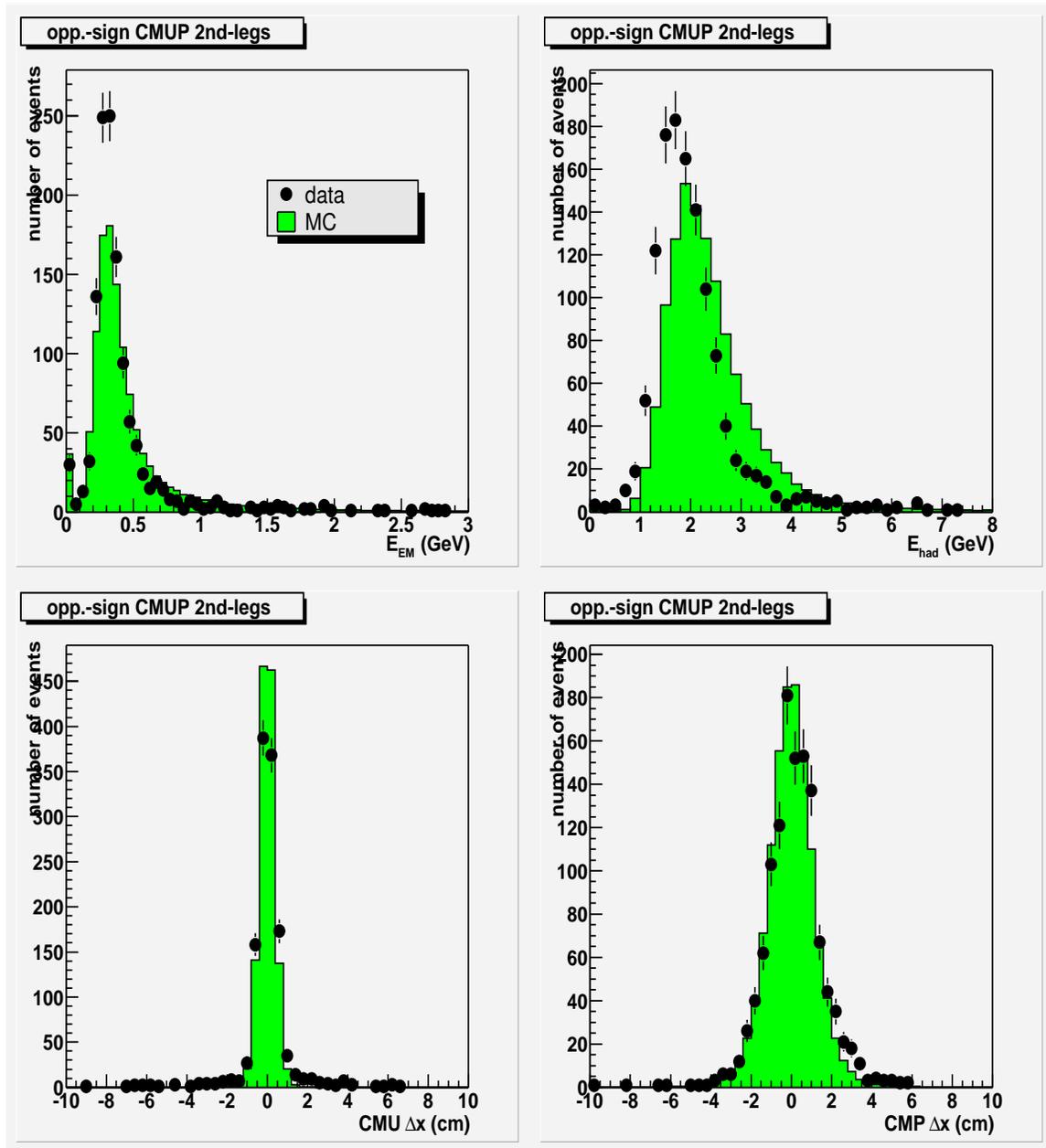


Figure 6.3: CMUP electron ID variables from the data compared with Monte Carlo expectations.

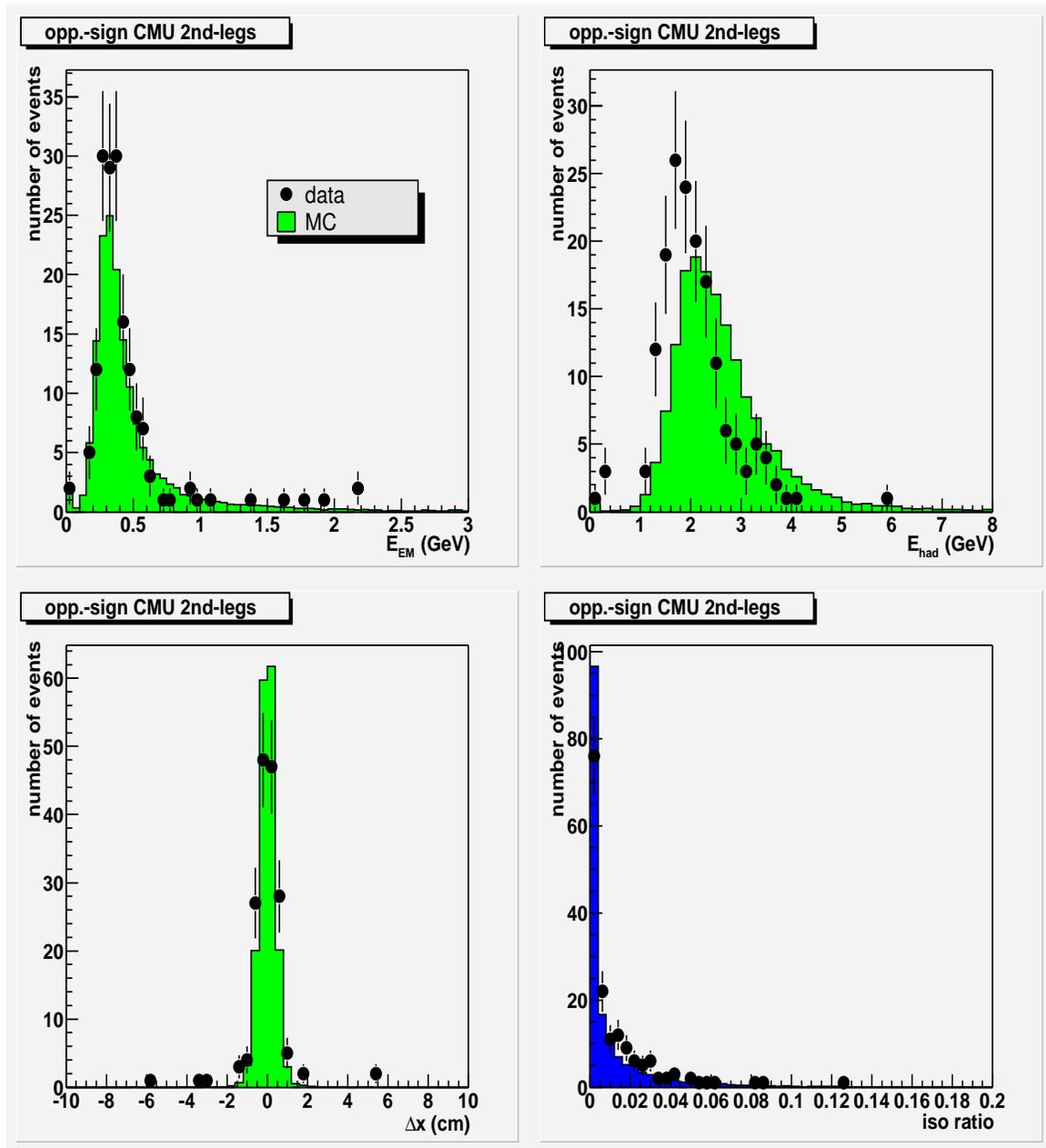


Figure 6.4: CMU electron ID variables from the data compared with Monte Carlo expectations.

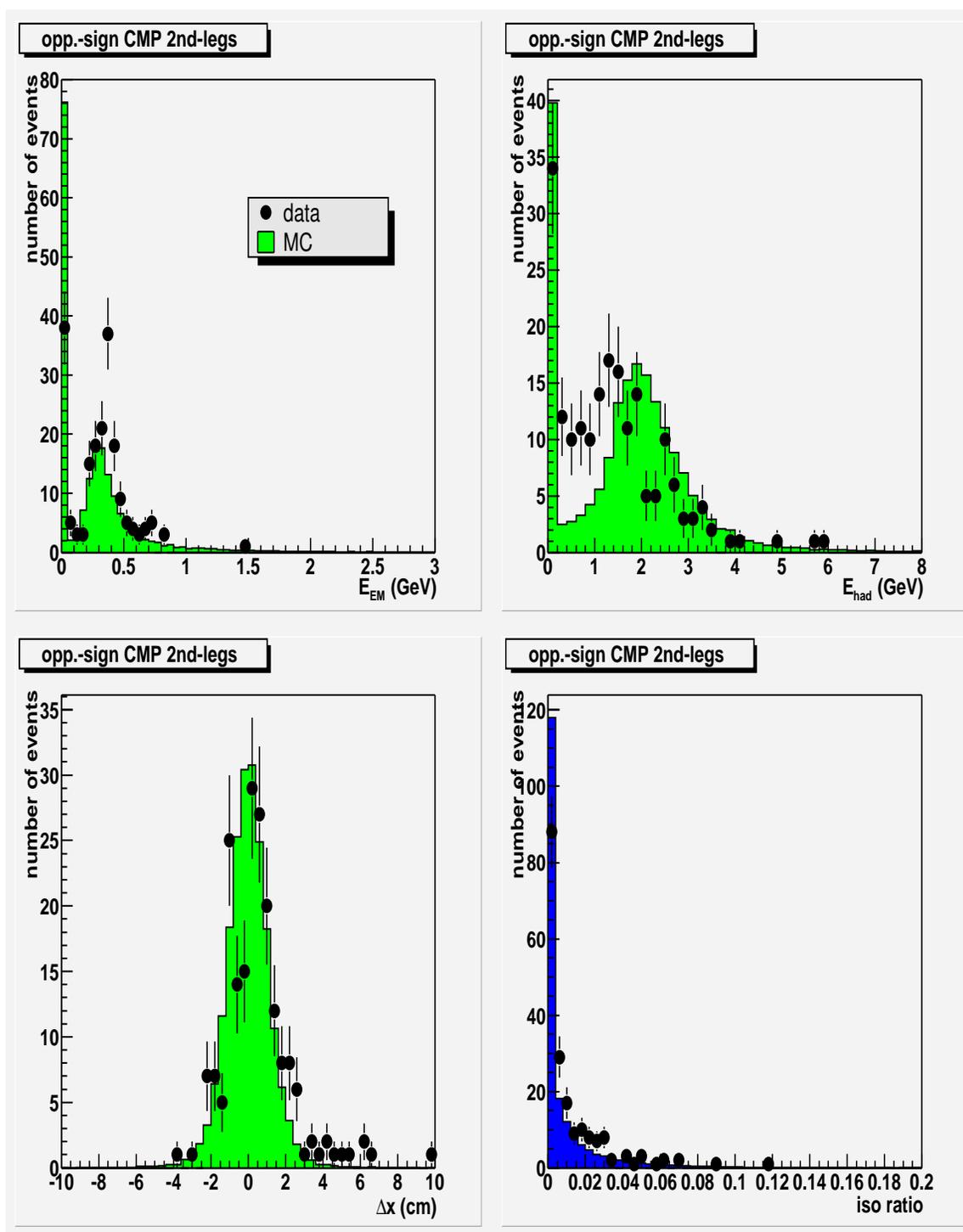


Figure 6.5: CMP ID variables from the data compared with Monte Carlo expectations.

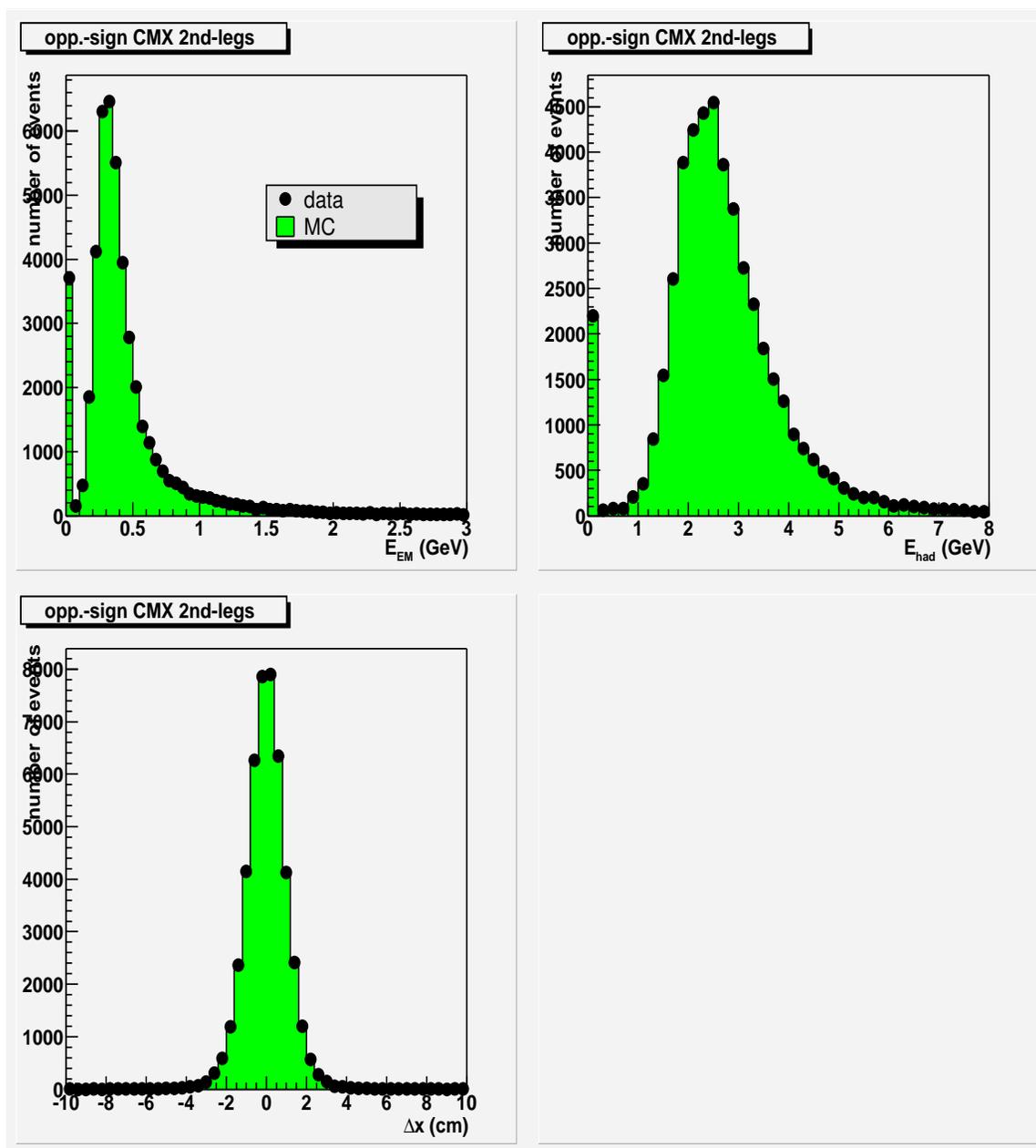


Figure 6.6: CMX ID variables from the data compared with Monte Carlo expectations.

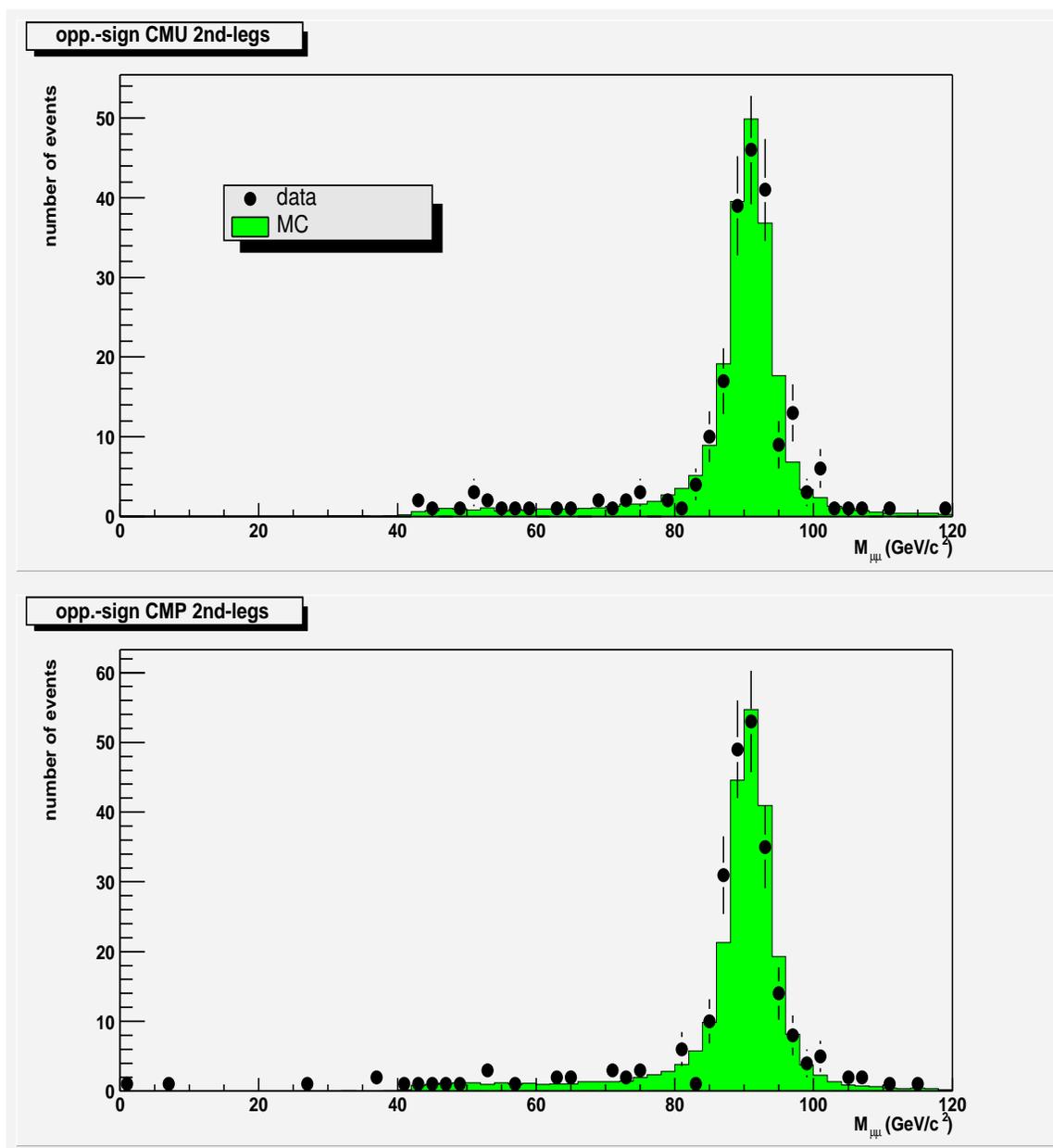


Figure 6.7: Invariant dimuon mass using at least a CMU-only or CMP-only muon.

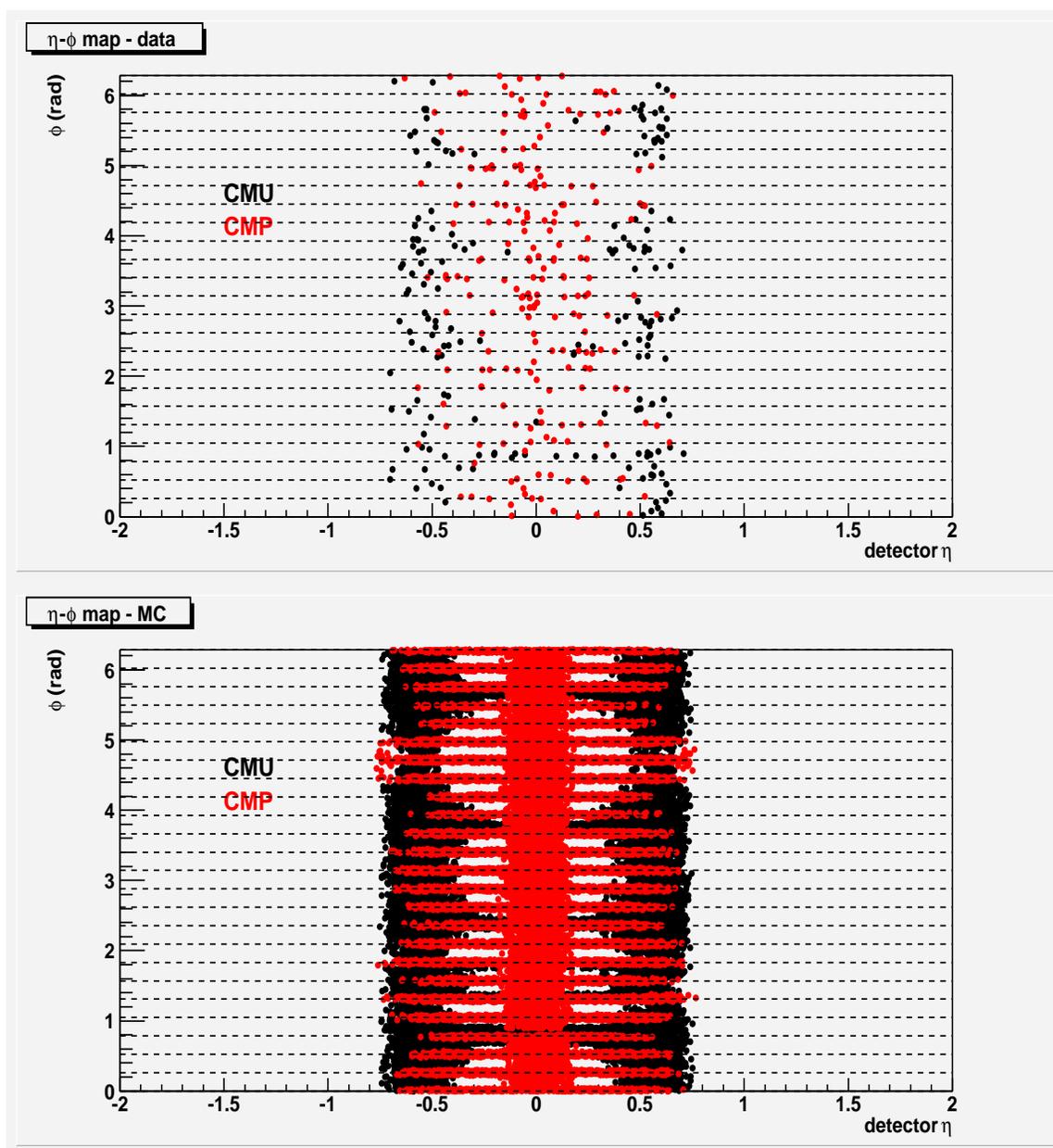


Figure 6.8: The $\eta - \phi$ map of CMU-only and CMP-only muons in the Run II data (upper plot) and Monte Carlo (lower plot).

6.2.4 Muon Efficiencies

Muon Identification Efficiencies

The muon identification efficiencies for each muon type are shown in Table 6.5, while the

details of the determination are presented in reference [98]. NIMUO refers to any nonisolated muon, be it a CMUO, CMX, CMU or CMP. Because of the limited statistics, we do not estimate the identification efficiency separately for various non-isolated muon categories.

Lepton Category	Data	MC	Id Efficiency
CMUP	0.851±0.007	0.907±0.002	0.94±0.01 [98]
CMX	0.901±0.008	0.888±0.003	1.015±0.008 [98]
CMU	0.905±0.011	0.912±0.003	0.993±0.013
CMP	0.919±0.011	0.935±0.002	0.983±0.011
NIMUO(all)	0.842±0.034	0.854±0.008	0.986±0.041

Table 6.5: Muon identification efficiencies and the muons scale factors (data/MC) used to correct the acceptance.

6.3 Conversion and Cosmic removal

After the previously described cuts, we further clean up the sample but removing well identifiable possible sources of background, photon conversions and cosmic rays.

6.4 Isolation cut

For each top dilepton event at least one lepton has to be isolated. Any lepton can be non-isolated, except PHX and CMIO, which are required to be isolated. For this analysis only the calorimeter isolation fraction I_{cal} is used. I_{cal} is corrected for lateral leakage[14] in the calorimeter⁴, and is defined as :

$$I_{cal} = \frac{\Delta E_T^{0.4}}{E_T}, \quad \text{for electrons, and} \quad (6.3)$$

$$I_{cal} = \frac{\Delta E_T^{0.4}}{P_T}, \quad \text{for muons,} \quad (6.4)$$

⁴The lateral leakage is the energy in the outer region of the EM shower that outside the the calorimetry towers associated with the shower.

where ΔE_T is the calorimeter transverse energy in the $\Delta R=0.4$ cone, around the lepton direction axis, excluding the energy deposition associated with the lepton.

The isolation requirement reduces dramatically the possibility that a lepton from a semileptonic b decay, typically surrounded by jet activity, will pass the lepton identification criteria.

6.5 Dilepton Event Classification

The leptons used in this analysis fall in two categories, depending if they can be the trigger lepton or not:

- Tight leptons: CEM, PHX, CMUP and CMX
- Loose leptons: CMU, CMP, CMIO

For each category shown in Table 6.6, at least one tight lepton is required to be isolated. There are 41 subcategories, based on the isolation requirements.

Category	Trigger	Luminosity
ee categories		
CEM-CEM	CEM_18	193 pb ⁻¹
CEM-PHX	CEM_18	162 pb ⁻¹
eμ categories		
CEM-CMUP	CEM_18 CMUP_18	193 pb ⁻¹
CEM-CMIO/U/P	CEM_18	193 pb ⁻¹
CEM-CMX	CEM_18 CMX_18	175 pb ⁻¹
PHX-CMUP	CMUP_18	162 pb ⁻¹
PHX-CMX	CMX_18	150 pb ⁻¹
PHX-CMIO/U/P	MET_PEM	162 pb ⁻¹
μμ categories		
CMUP-CMUP	CMUP_18	193 pb ⁻¹
CMUP-CMIO/U/P	CMUP_18	193 pb ⁻¹
CMX-CMIO/U/P	CMX_18	175 pb ⁻¹
CMX-CMX	CMX_18	175 pb ⁻¹
CMX-CMUP	CMX_18 CMUP_18	175 pb ⁻¹

Table 6.6: The dilepton categories and the triggers used to collect the events. Each dilepton category has at least one tight lepton, **in bold**, which can trigger the event. Also each category has at least one tight isolated lepton, while the partner can be nonisolated.

6.6 Invariant Mass cut

The same lepton-flavor dilepton data sample, selected by requiring two well-identified leptons, is dominated by the events from Z^0 decays. To reduce this background, a special treatment is applied when the invariant mass $M_{\ell\ell}$ is in the range (76, 106) GeV. In the past these events were simply rejected and thereby reducing the signal acceptance in both the ee and $\mu\mu$ channels (Figure 6.9) by 24%. A new variable, known as *jet significance* was shown to be very efficient in rejecting the Drell-Yan (DY) background ([73; 48]) and it is briefly described below.

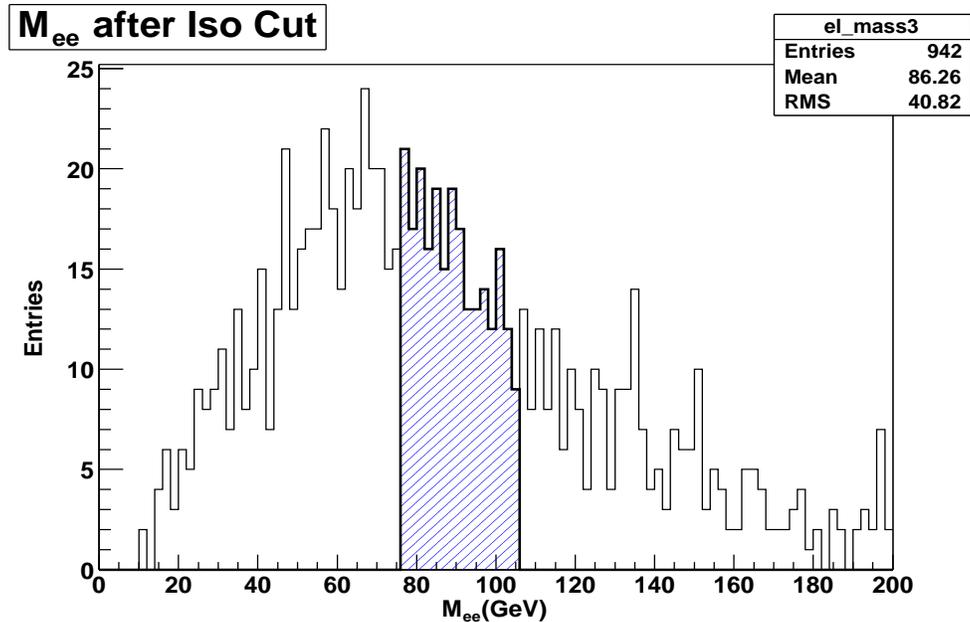


Figure 6.9: Invariant mass for ee and $\mu\mu$ top dilepton events, after lepton ID and isolation cuts. The shaded region corresponds to the Z window and represents about 24% of the events.

Jet significance

A Drell-Yan event has no real missing energy, as the final state objects should leave all their energy in the detector. However the detector is not hermetic, having cracks in ϕ between the central wedges, between the east and west central calorimeter regions ($\eta = 0$), between the central and plug calorimeters ($|\eta| = 1.1$), and in the beam hole region ($|\eta| > 3.64$). Also the calorimeter response is nonlinear and only the particles energy above some threshold are included in the calorimeter clusters. The most common source of missing energy is due to the jet energies being mismeasured, even after the jet corrections are applied, as discussed in section 6.9. If only one object in the event is mismeasured and the fake \cancel{E}_T is below 50 GeV, an L cut, which quantifies how close the MET direction is to a mismeasured object, efficiently rejects these events. The L cut is detailed in Section 6.8. However, a large number of DY events will still pass the L cut primarily because either the false \cancel{E}_T is greater than 50 GeV, or more than an object is mismeasured and the \cancel{E}_T direction is randomly oriented. Let's introduce a new variable, *the jet significance*, defined as

$$\text{jetsig} = \frac{\cancel{E}_T}{\sqrt{\sum_{|\Delta\phi(\cancel{E}_T, \text{jet})| < 90^\circ} (\vec{E}_{T\text{jet}} \cdot (\frac{\vec{\cancel{E}_T}}{\cancel{E}_T}))}} \quad (6.5)$$

where the sum is over all the jets with $|\eta| < 2.5$ and in the same hemisphere (xy semiplane) as the \cancel{E}_T (i.e. $|\phi_{\cancel{E}_T} - \phi_{\text{jet}}| < \pi/2$).

A few variables were explored to recover $t\bar{t}$ dilepton events within the $Z \rightarrow e^+e^-$ window, such as a missing energy significance cut, a tighter invariant mass cut, a tighter \cancel{E}_T cut. The jet significance⁵ was found to be the best discriminant between the events with real \cancel{E}_T , such as $t\bar{t}$ or WW and the Drell-Yan. However even after a jetsig requirement, there are still $Z \rightarrow e^+e^-$ events with large \cancel{E}_T passing this cut. Therefore to further reject them it is required that \cancel{E}_T not be close to a jet, to complement the L cut, which is applied only for $\cancel{E}_T < 50$ GeV.

The selection requirements are summarized in Table 6.7.

Inside	Outside
jetsig < 8	-
$\Delta\phi(\cancel{E}_T, \text{nearest jet}) < 10^\circ$	-

Table 6.7: Selection requirements applied only for $ee, \mu\mu$ events. *Inside* means for $76 < M_{\ell\ell} < 106$ GeV, where $\ell = e, \mu$.

Using these two new cuts instead of a rejection of events inside the Z mass window, 90% of the $t\bar{t}$ acceptance loss is recovered, while 80 % of $Z^0 \rightarrow e^+e^-$ events are rejected (Table 6.8).

⁵In fact the jet significance is nothing but a *missing energy significance*, $\cancel{E}_T / \sigma_{\cancel{E}_T}$, where $\sigma_{\cancel{E}_T}$ is, (up to a multiplicative constant), the resolution on \cancel{E}_T due to the uncertainty on the jet energies.

Cut efficiency (%)		
Variable	$t\bar{t}$	$Z^0 \rightarrow e^+e^-$
$\text{jetsig} < 8$	90.34 ± 2.45	32.74 ± 4.41
$\Delta\phi(\cancel{E}_T, \text{nearestjet}) < 10^\circ$	96.95 ± 1.50	54.05 ± 8.19

Table 6.8: Selection requirements applied only for $ee, \mu\mu$ events. *Inside* means for $76 < M_{\ell\ell} < 106$ GeV, where $\ell = e, \mu$.

6.7 \cancel{E}_T cut

A $t\bar{t}$ dilepton event has two neutrinos in the final state, which cannot be detected and generate a large energy imbalance in the transverse plane known as *missing energy* (\cancel{E}_T). The default or *raw missing energy*, associated with every event is a 2 dimensional vector $(\cancel{E}_{T_x}, \cancel{E}_{T_y})$, in the xy transverse plane, equal to the negative of the vector sum of all the transverse energy in calorimeter. The magnitude of $\vec{\cancel{E}}_T$, is denoted \cancel{E}_T . The raw missing energy undergoes three corrections, which tend to correct for few ways in which the \cancel{E}_T could be mismeasured.

- **Primary vertex correction:** The raw missing energy stored in CdfMet bank assumes that the event primary vertex is $z = 0$; using the full offline event information one reconstruct the event primary vertex (for dilepton events we use the z_0 of the highest transverse energy lepton in the event) and then recalculate $(\cancel{E}_{T_x}, \cancel{E}_{T_y})$.
- **Muon correction:** The muons are minimum ionizing and they deposited very little energy in the calorimeter. Therefore we add back to $(\cancel{E}_{T_x}, \cancel{E}_{T_y})$ the transverse energy deposited in the electromagnetic and hadronic calorimeter and subtract the COT beam constrained muon momentum components, (P_{Tx}, P_{Ty}) . We correct *for all the muons*, passing the dilepton selection cuts, be they isolated or not; they could be CMUP, CMP, CMU, CMX or CMIO.
- **Jet correction:** Every jet passing the dilepton selection is corrected for detector effects such as to go back to parton level energies. The \cancel{E}_T is then corrected for the difference between the corrected and uncorrected jet transverse energies.

From this point on \cancel{E}_T will refer to the corrected missing energy, unless explicitly stated otherwise.

6.8 $\Delta\phi$ or “L” cut

The previous selection requirement gets rid of the background events with small missing energy magnitude (Drell Yan mainly, but also $Z \rightarrow \tau\tau$). The $\Delta\phi$ cut rejects the events with $|\cancel{E}_T| < 50$ GeV, if the \cancel{E}_T is too close to either a lepton or a jet, explicitly $\Delta\phi(\cancel{E}_T, \text{nearest } \ell \text{ or jet}) < 20^\circ$. This requirement consists of two angular constraints: the angle between the \cancel{E}_T and closest lepton, and the angle between the \cancel{E}_T and the nearest jet. The reason both are used is as follows:

- The DY background events have no physical missing energy. The common way fake \cancel{E}_T is “created” is due to a jet being badly measured or lost in the detector cracks⁶. So requiring a good separation between \cancel{E}_T and the nearest jet is a very efficient way to reject the DY.
- The $\Delta\phi(\cancel{E}_T, \text{nearest lepton})$ cut was used in Run I mainly to reject dileptons from di-tau events and it was preserved for historical reasons. However its power is limited. Once the events are required to have at least 2 jets, $\Delta\phi$ is randomly distributed, and it should be dropped in the future.

6.9 At least 2 Jets

The jets are the result of the fragmentation and hadronization of the colored objects, such as quarks and gluons. There is assumed to be a one-to-one correspondence between a parton and its corresponding jet. The necessary steps to reconstruct a jet, and the jet clustering algorithm used in this analysis, are described in Section 4.6.

⁶To understand better why the jet energy could be mismeasured, see Section .

Jet Corrections

There are up seven corrections, consecutively applied to a *raw jet*, such that at the end the new jet 4-momentum reproduces closely the corresponding parton. They are briefly described below[45]:

- Level 1: *Relative energy corrections* - These are extracted using dijet balancing. *They make the response uniform in η , across the calorimeters.* The corrections are dependent on the raw jet energy and η , such that the energy scale is the same everywhere and equal to the response in the well understood region, $0.2 < |\eta| < 0.6$ ⁷.
- Level 2: *Time-dependent corrections* - These consist of calorimeter time-dependent corrections and are applied to each calorimeter tower energy.
- Level 3: *Raw energy scale corrections* - These rely on a comparison of photon-jet balancing of data from Run I and Run II. The results show that Run II energy scale is too low. An overall scale factor is applied such that Run II calorimeter raw scale is same as the one in Run I. The usage of Run I scale as reference is made because at Level 5 the Run I corrections are used.
- Level 4: *Multiple interaction corrections* - It corrects the calorimeter-level jet energy for any additional interaction in the event. Because for the run range used in this analysis the average number of primary vertices per event is about 1.2, this correction is not applied to the jets used in our analysis⁸.
- Level 5: *Absolute energy correction* - It corrects the jet energy observed in the calorimeter to the $\sum P_T$ of the particles within the cone of same size around the parton direction which matched the jet direction within $\delta R < 0.5$. The correction factor is parameterized separately for below and above 100 GeV. These correction is momentarily taken from Run I.
- Level 6: *Underlying event correction (UE)* - It subtracts the underlying energy from the particle-level jet energy. The UE energy was measured from minimum bias data

⁷This range is preferred as it avoids the cracks at $\eta = 0$ or 1.1.

⁸The underlying event energy per extra interaction, in a 0.4 cone, is about 260 MeV [56].

and corrected for calorimeter response by multiplying it by 1.6. However this correction is small.

- Level 7: *Out-of-cone corrections* - It corrects the particle-level energy for the radiation lost outside the clustering cone, taking the "jet energy" back to "parent parton energy".

Jet Requirement

The jets used for the event selection are corrected for level 1, 2, 3 and 5 only. Dilepton events are required to have at least two jets, with $E_T > 15$ GeV.

6.10 H_T requirement

The top quark is much heavier than any other known quark or lepton. Also at the Tevatron the top decays almost at rest, therefore many of its decay products are central and *the total energy flow in the transverse plane* of $p\bar{p}$ collisions is a good discriminant between $t\bar{t}$ and the SM backgrounds. We define the scalar sum of transverse energy of objects in the event, denoted H_T , as below:

$$H_T = \sum_{leptons} E_{T\ell} + \sum_{jets} E_{Tj} + \cancel{E}_T \quad (6.6)$$

where *only the objects passing the selection requirements* are included in the sum, after all the corrections, are applied. The H_T distributions for the main backgrounds are shown in Figure 6.10.

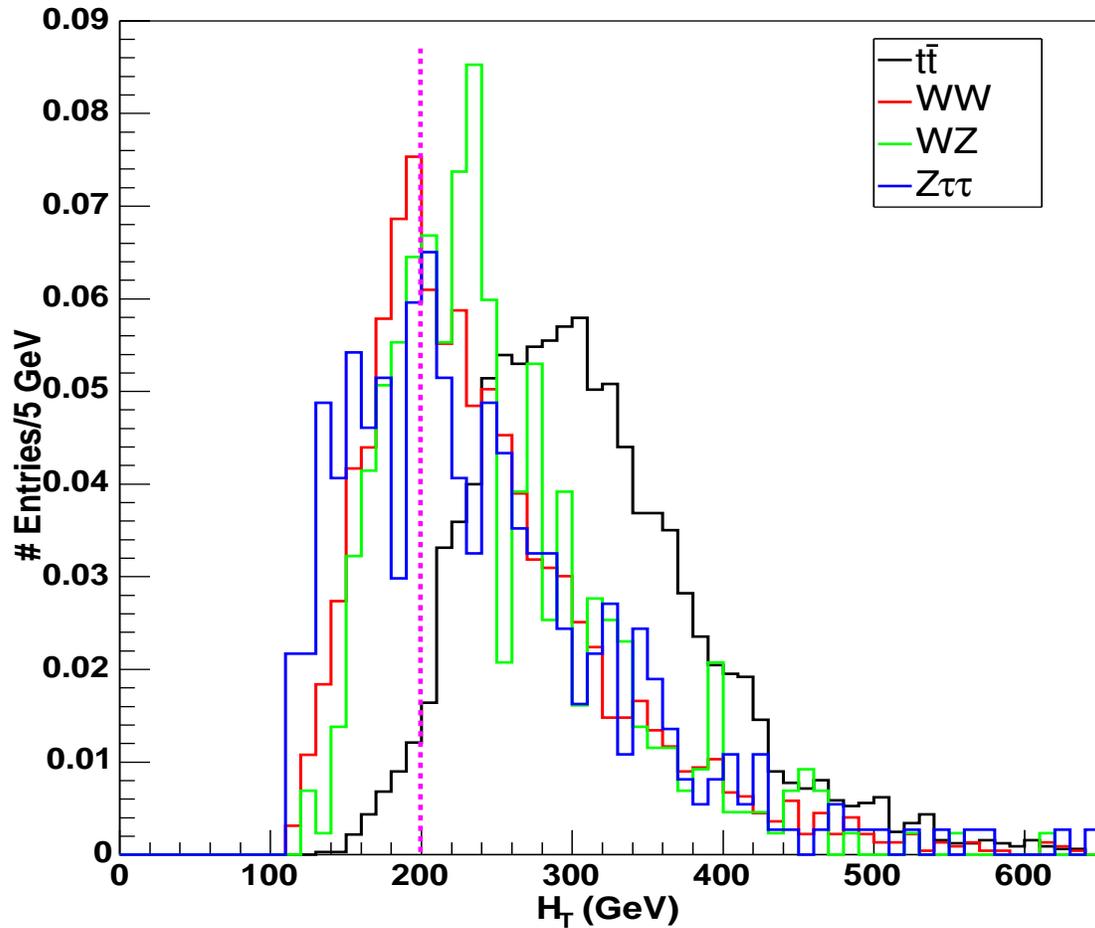


Figure 6.10: H_T distributions for $t\bar{t}$ signal and WW, WZ and $Z \rightarrow \tau\tau$.

Figure 6.11 shows the dilepton signal (S) and backgrounds (B) as a function of H_T , for 126 pb^{-1} . Figure 6.12 shows $S/\sqrt{S+B}$, S/B , and $S/(S+B)$ vs. H_T ([73]). The optimal $S/\sqrt{S+B}$ is obtained when we require $H_T > 195 \text{ GeV}$. Since the optimization is very broad, it was chosen to use a $H_T > 200 \text{ GeV}$ requirement for the event selection.

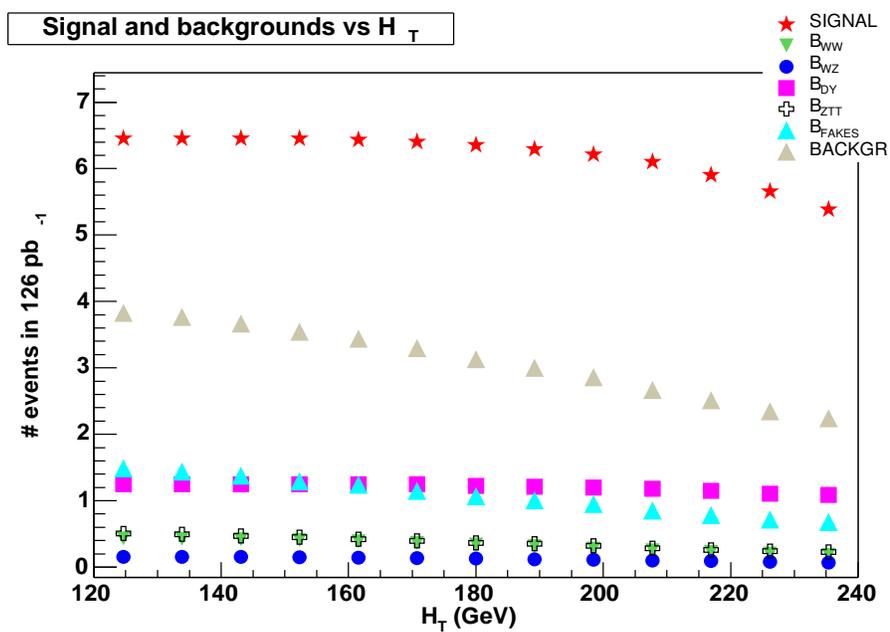


Figure 6.11: Signal and background as a function of the cut placed on H_T .

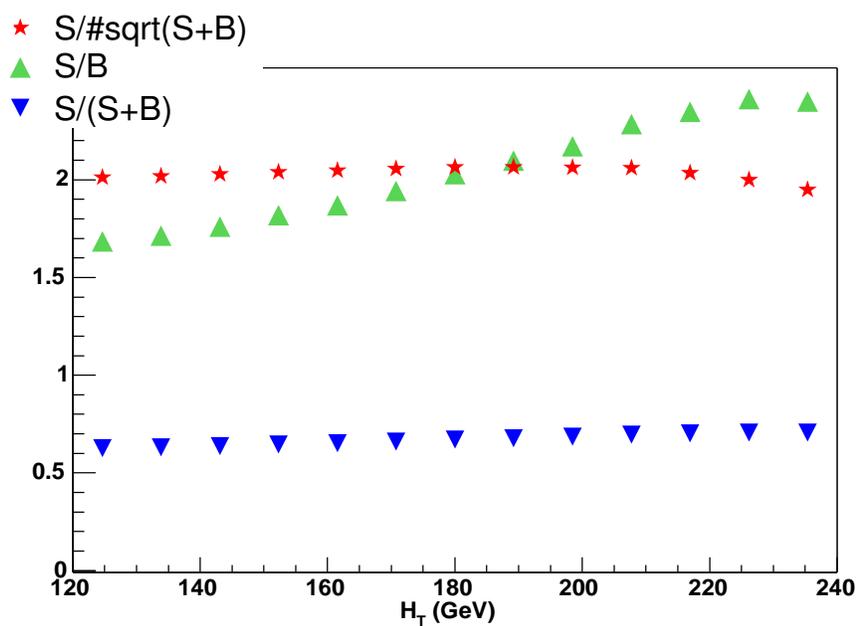


Figure 6.12: Various ratios of signal-to-background as a function of cut placed on H_T .

6.11 Opposite Sign

The final state leptons are required to be oppositely charged, as it is explicitly required that the signal acceptance comes from events with a pair (W^+, W^-) decaying leptonically. This cut is effective in reducing the QCD fake background to about a half. Also the number of same-sign events are a good cross-check of the fake lepton background. Applying this cut later offers the advantage to spot possible excess of same-sign events due to any detector problems or possible new physics.

PHX Charge Fake Rate

PHX electrons have a non-zero charge fake rate (see Figure 6.13), which is determined by calculating the fraction of same-sign CP⁹ Z events in the data, in the mass window (76,106). Based on 3149 Z events (OS+SS)¹⁰, of which 411 are SS, one determines a fake rate $f = 0.130 \pm 0.006$, which it is applied to the CP portion of the DY background (CP accounts for about half, as determined from the ratio of the number of CC/CP¹¹ Z events in the data). Also all the backgrounds which are estimated using MC are corrected for the difference between the fake rate in the data and MC.

⁹CP stands for one lepton in central, one in the plug calorimeter.

¹⁰OS stands for opposite sign, while SS stands for same sign.

¹¹CC stands for two lepton in central calorimeter.

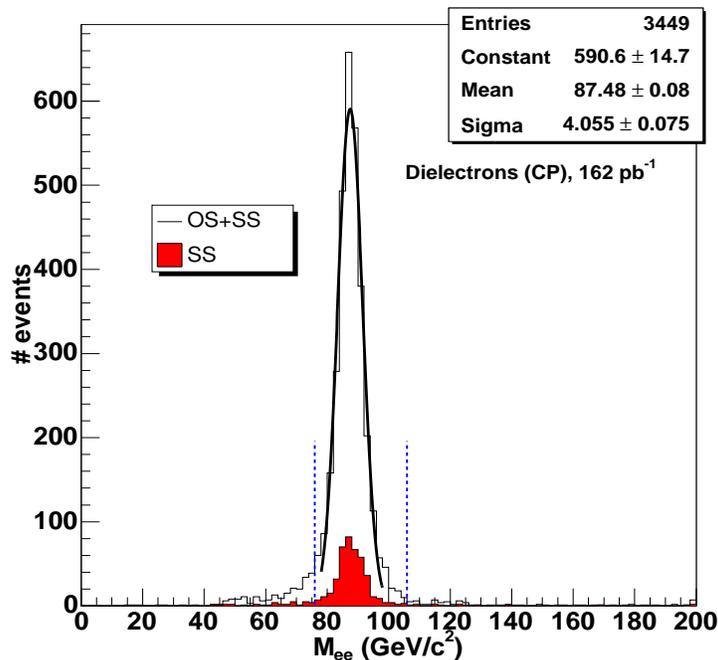


Figure 6.13: Invariant mass distribution for CP ee events in the data. Same-sign events are shown in red.

6.12 Trilepton Events

The events with 3 or more leptons in the final state passing all the dilepton identification requirements, are expected to be very rare based on the Standard Model predictions¹². Therefore, for the cross section measurement, only events with exactly two leptons are considered. However, any possible trilepton event seen in the data will be very interesting and analyzed separately.

¹²The expected number of trileptons is expected to be less than 0.1 events in 193 pb^{-1}

7 Top Dilepton Acceptance

The top dilepton acceptance is needed to measure the $t\bar{t}$ cross-section in the dilepton channel, as it will be seen in Chapter 9. As a result of the $p\bar{p}$ collisions, a number of top dilepton events are produced inside the detector. Of these, only a limited number of top dilepton events can be isolated, due to inefficiencies of the triggers or of the various selection criteria, applied to separate the signal from the overwhelming backgrounds. The top dilepton acceptance represents the fraction of produced top dilepton events, which survive the selection process, or the probability that a $t\bar{t}$ dilepton event will pass the selection requirements. Mathematically the acceptance ϵ_{dil} takes the following expression:

$$\epsilon_{dil} = \epsilon_{geom-kin} \cdot \epsilon_{ID} \cdot \epsilon_{trig} \cdot \epsilon_{conv} \cdot \epsilon_{cosmics} \cdot \epsilon_{iso} \cdot \epsilon_{Zmass} \cdot \epsilon_{\cancel{E}_T} \cdot \epsilon_{\Delta\phi} \cdot \epsilon_{2jet} \cdot \epsilon_{H_T} \cdot \epsilon_{OS} \quad (7.1)$$

The top dilepton acceptance ϵ_{dil} is a product of *efficiencies*, each individual efficiency being determined in respect with the number of events passing the previous cut, therefore the order is *important*. The acceptance is quoted with respect with the inclusive number of $t\bar{t}$ events. However only the *real dilepton events* are considered in the acceptance calculation,

when both W bosons decay leptonically¹. The total acceptance can be expressed as

$$\epsilon_{dil} = br \cdot \epsilon_{real-dil}, \quad (7.2)$$

where br is the branching ratio of dilepton channel relative to the total $t\bar{t}$ cross-section, $br = \frac{4}{81}$, and $\epsilon_{real-dil}$ is the fraction of real dilepton events identified.

The $t\bar{t}$ dilepton acceptance is determined using a Pythia Monte Carlo sample, with a top mass of 175 GeV. However, corrections are applied whenever the data can be used to cross-check how well the simulation models it. The lepton identification and trigger efficiencies are extracted from data.

7.1 Effects of the selection requirements on the acceptance

In this section the effect of the selection criteria on the number of dilepton events will be briefly discussed, with the focus on the cuts which cause the largest signal losses, geometrical and kinematical, and lepton identification requirements.

The effect of the geometrical and kinematical requirements

The first step in selecting a dilepton event is to require 2 lepton candidates, passing very loose cuts, detailed in Section 6.2. The *effect of kinematical cut* is easily understood by looking at the p_T distribution of leptons from W decays, shown in Figure 7.1. By selecting a lepton with E_T or p_T above 20 GeV, 16% of them fails the cut. Out of all the real dilepton events, only $\epsilon_{kin} = 0.84 \times 0.84 = 71\%$ pass the kinematical criterion. The *geometrical cut* basically requires that all the muons be central ($|\eta| < 1.0$), either having stubs in CMUP, CMP or CMX, or be nonfiducial stubless. For electrons the pseudorapidity extends up to

¹The lepton+jets $t\bar{t}$ events could mimic a top dilepton event, in case one of the b quarks decays to a high- p_T electron or muon. But these events are counted also in the fake background and properly subtracted from the number of dilepton candidates observed in the data

$|\eta| = 2.0$, in the plug region, but with one lepton always in the central region of the CDF detector ($|\eta| < 1.1$). As Figure 7.1 shows only 74 % of leptons are in the region $|\eta| < 1.1$, so only about $\epsilon_{geom-cc} = 0.74 \times 0.74 = 55$ % of events have both leptons central. In case of the electrons we are doing slightly better, as events with one plug, one central electron are included. However the central electrons are required to be fiducial, which corresponds to about 18.7% loss for one lepton, as argued in Section 3.3. The net effect on central-central dielectron events is that only $0.81 \times 0.81 = 66$ % of them have both electrons fiducial. This loss is partly compensated by allowing plug electrons. In case of the muons, the inclusion of the stubless muons recovers most of the nonfiducial muons. Now a very rough estimate for the geometrical and kinematical with respect with just real dilepton events, which partly justify the $\epsilon_{geom-kin}$ efficiency, is $0.71 \times 0.55 = 40$ %. Using the generator level information, out of 17979 dilepton events ($WW \rightarrow \ell\ell$), $\ell = e$ or μ , 9473 pass the geometrical and kinematical requirements, for an efficiency of 52.7 %. The higher efficiency is due to the fact one of the leptons passing the geometrical and kinematical requirements could come from a quark. However the majority of these fake leptons will fail the identification or isolation cuts. Also 5310 out of 22247 dilepton events with at least a one tau pass the geometrical and kinematical cuts, for an efficiency of 23.9 %.

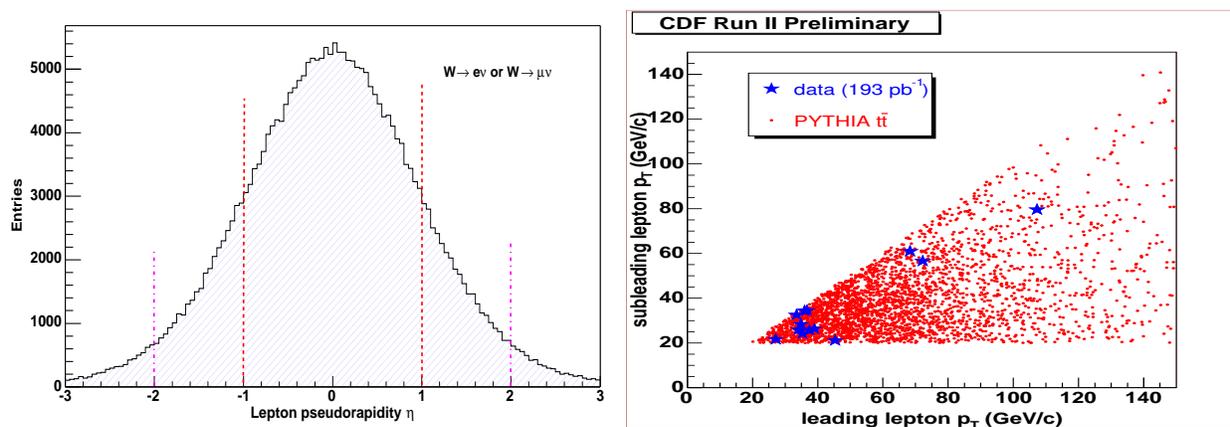


Figure 7.1: The pseudorapidity and the transverse momentum of the leptons from W decays, at the generator level.

The effect of the lepton ID requirements

Again, without going into details, one can estimate the effect of the lepton ID requirements on the dilepton events. The efficiency of the electron identification cuts ϵ_T , for an well isolated electron ($\text{iso} < 0.02$) is about 86 %, as shown in Table 6([72]). This implies that only $0.86 * 0.86 = 74\%$ of the di-electron events pass the identification criteria. This efficiency is an upper bound, as the leptons inside a $t\bar{t}$ event could be close to a jet, thus becoming less isolated than the leptons in a Z decay, so ϵ_T lower.

The rest of the selection criteria have high efficiencies, above 90 %, as can be seen from Figure 7.2 or Table 7.1.

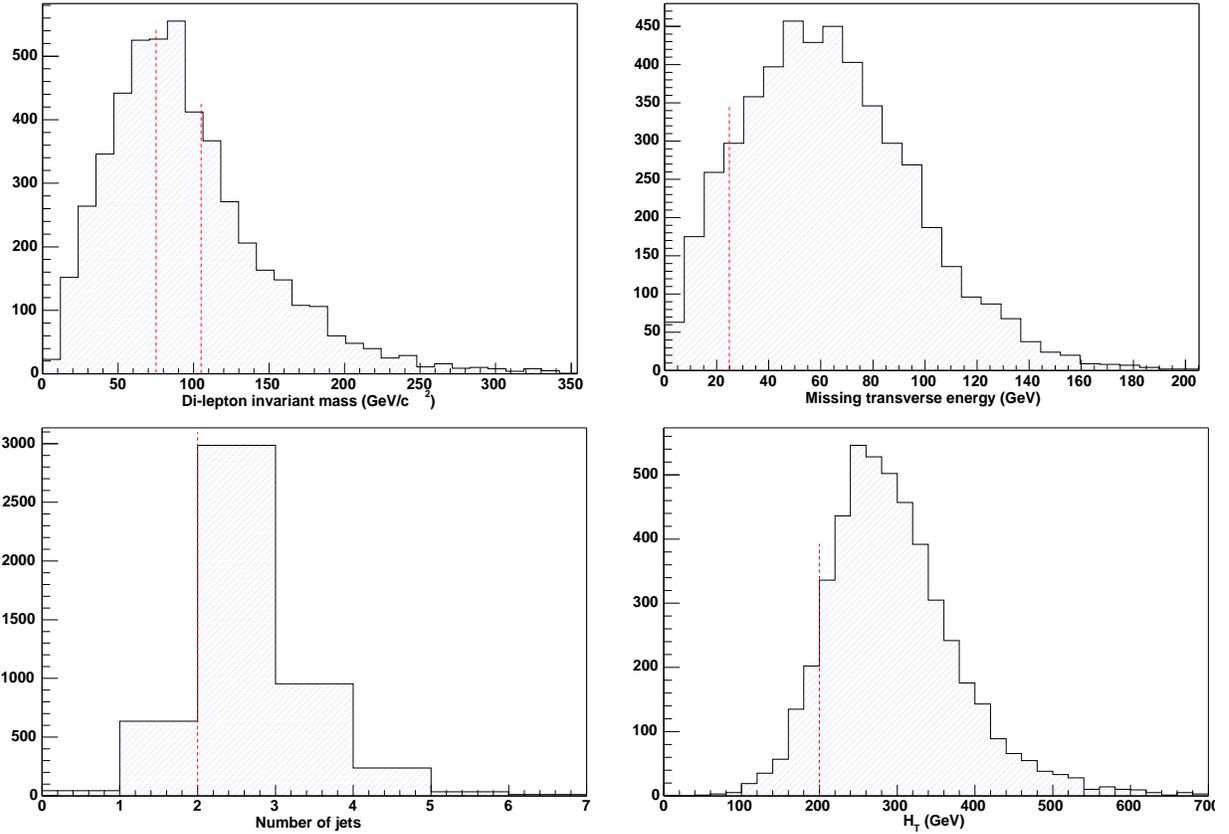


Figure 7.2: Invariant mass $M_{\ell\ell}$, the missing transverse energy, the number of jets and the scalar transverse sum H_T distributions are shown for $t\bar{t}$ events. The dotted line shows where the selection cut is placed.

7.2 Results

The number of events falling in each dilepton category, after applying the selection criteria on the Pythia MC sample top 175, are shown in Tables B.1-B.4. The trilepton events are preserved as a separate category, even if they are not used in the final acceptance or in the cross-section calculation. The last row in the Table B.4 shows the total number of dilepton events passing each cut. Using the numbers from these tables, the individual pieces needed to calculate ϵ_{dil} in formula 7.1 are extracted and shown in Table 7.1. The *raw dilepton acceptance* using the Pythia sample is $\epsilon_{dil}(raw) = 0.79 \pm 0.01$. However these number does not include the trigger efficiencies, identification efficiencies scale factors, reconstruction efficiencies or PHX charge fake rate. Applying all of the above, and also properly waiting the dilepton categories for the data luminosities specified in Table 5.2, the *dilepton acceptance luminosity weighted* becomes:

$$\epsilon_{dil} \times \mathcal{L} = (1.22 \pm 0.01) pb^{-1} \quad (7.3)$$

where there uncertainty is statistical only.

Now let's define the weighted luminosity

$$\mathcal{L}_w = \sum_i \mathcal{L}_i \times w_i \quad (7.4)$$

where the w_i represents the fraction of events falling in the luminosity category i . Using the values from Table 7.2, one gets the weighted luminosity for our data sample:

$$\mathcal{L}_w = 185 pb^{-1} \quad (7.5)$$

Then we can define the effective acceptance $\epsilon_{dil-eff}$, such that $\epsilon_{dil-eff} \cdot \mathcal{L}_w = \epsilon_{dil} \cdot \mathcal{L}$. Using the above values one gets:

$$\epsilon_{dil-eff} = (0.66 \pm 0.01)\% \quad (7.6)$$

Variable	Number of events passing the cut	Efficiencies(%)
$WW \rightarrow \ell\ell$	44226	-
$\epsilon_{geom-kin}$	12145	3.05 ± 0.03
ϵ_{ID}	5319	43.80 ± 0.45
ϵ_{iso}	5070	95.32 ± 0.29
ϵ_{conv}	4898	96.61 ± 0.25
ϵ_{mass}	4831	98.63 ± 0.17
$\epsilon_{\cancel{E}_T}$	4264	88.26 ± 0.46
$\epsilon_{\Delta\Phi}$	3905	91.58 ± 0.43
ϵ_{2-jet}	3327	85.20 ± 0.57
ϵ_{H_T}	3209	96.45 ± 0.32
ϵ_{OS}	3132	97.60 ± 0.27
ϵ_{dil}	3132	0.79 ± 0.014

Table 7.1: Individual cut efficiencies and the total dilepton acceptance, from a Pythia $t\bar{t}$ sample of 398037 events.

Category	Luminosity	Fraction of acceptance (%)
CEM/CMUP	193	68.52
CEM/CMUP and CMX	175	16.51
CEM/CMUP and Si	162	13.76
CEM/CMUP and CMX and Si	150	1.21

Table 7.2: Fraction of total dilepton acceptance corresponding to various luminosities.

Signal Composition

It is useful to quantify the contributions to the total acceptance *by event topology*. In this context the leptons could be either central (C) or plug (P), or could be isolated (I) or nonisolated (NI). As can be seen in Table 7.3, the dominant portion comes from two central isolated lepton categories. However 15 % of events have one PHX plug lepton, while 10 % have one non-isolated lepton.

Topology	Fraction of acceptance (%)
CC-I	75.9
CC-NI	9.1
CP-I	14.0
CP-NI	1.0

Table 7.3: Contribution to the $t\bar{t}$ signal acceptance from various event topologies.

By lepton flavor, the ee , $e\mu$ and $\mu\mu$ contributions to the ϵ_{dil} are shown in Table 7.4. While the pseudorapidity coverage is significantly larger for electron ($\eta_{ele} < 2.0$) than for muons ($\eta_{muo} < 1.0$), the inclusion of stubless muons recovers most of the nonfiducial muons, such that the ee and $\mu\mu$ contribute evenly to the acceptance.

Lepton Flavor	Contribution (%)
ee	22.2
$e\mu$	54.3
$\mu\mu$	23.5

Table 7.4: Contribution to the $t\bar{t}$ signal acceptance from various event topologies.

Table 7.5 shows the contribution to the acceptance by the generator level lepton pairs ($WW \rightarrow \ell\ell$), where $\ell = e, \mu$ or τ . About 10 % of the dilepton acceptance comes from a $\ell+\tau$ pair, where the τ further decays to an electron or muon and neutrinos.

Lepton Pair $WW \rightarrow$	Fraction of Acceptance
ee	20%
$\mu\mu$	22%
$e\mu$	49%
$e\tau$	4%
$\mu\tau$	5%
$\tau\tau$	< 1%

Table 7.5: Contribution to the $t\bar{t}$ signal acceptance from generator level lepton pairs.

Acceptance dependence on the top mass

Top dilepton acceptance increases with the top mass, as shown in Figure 7.3. Over the

mass range of 170 to 190 GeV the acceptance increases by about 10 %, well below the current uncertainty on the $t\bar{t}$ cross-section.

As the top mass increases, so does the energy of W bosons, and the leptons from their decay will pass the kinematical cuts with a greater probability. Also the 2-jet cut efficiency increases, as the jets will become more energetic. The fraction of events which fall in ee , $e\mu$ and $\mu\mu$ remain practically the same as the top mass increases.

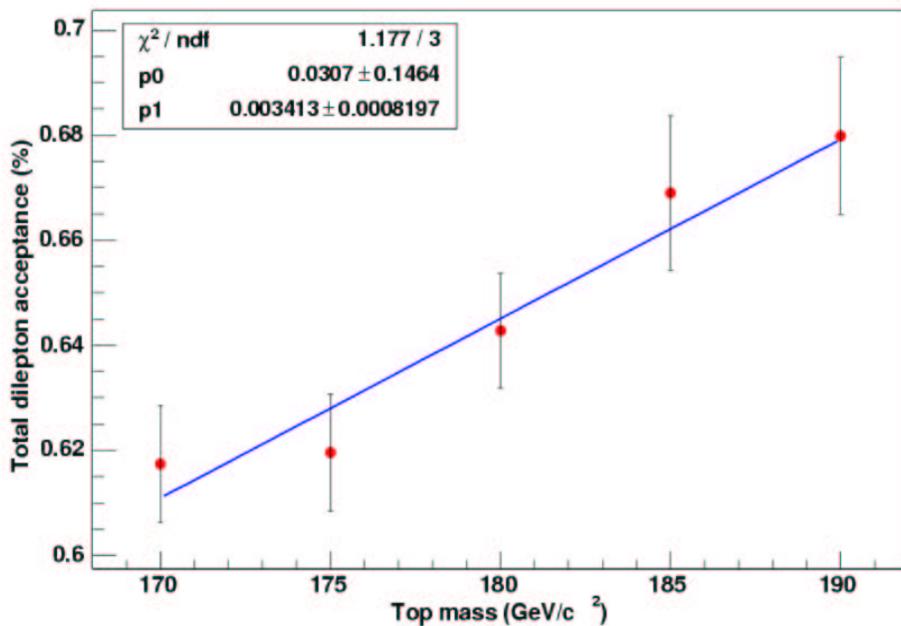


Figure 7.3: Total dilepton acceptance versus top mass

7.3 Systematic uncertainties in the acceptance

The sources of systematic uncertainties on the dilepton acceptance are discussed briefly below([74]):

Monte Carlo generators uncertainty

This systematic uncertainty is calculated as the difference between the dilepton acceptance using Pythia, and a sample generated with Herwig. However there are *two factors* which have to be considered: the branching ratios ² $B_{W\ell\nu} = B(W \rightarrow \ell\nu)$ used in Monte Carlo generators differ between Pythia ($B_{W\ell\nu} = 0.111$) and Herwig ($B_{W\ell\nu} = 0.108$), and also Herwig $t\bar{t}$ samples lack the QED final state radiation (FSR). To overcome the first issue, the Herwig acceptance is reweighted by $(0.108)^2 / (0.111)^2 = 0.947$. By comparing the reweighted acceptance from a Herwig sample ($\epsilon_{dil} = 0.894 \pm 0.016$) with the dilepton acceptance using a Pythia sample with no QED FSR ($\epsilon_{dil} = 0.946 \pm 0.022$), a systematic uncertainty of 5.5% is found. Both samples were generated for a top mass of 175 GeV.

Jet Energy Scale

The raw jets are applied a set of corrections before being used in the event selection, as described in Section 6.9. The correction factors have some uncertainties, parametrized as a function of jet E_T or η , due to various effects, such as fragmentation and QCD radiation modeling, or the detector response to single particles in simulation. The jet correction factors are shifted up or down by 1σ and new top dilepton acceptances are obtained. Taking the semi-difference of the shifted up and down acceptance divided by the default acceptance yields a systematic uncertainty of 4.7%.

Lepton ID Scale Factors

The simulation does not always models data very well. The lepton identification efficiencies are determined using an unbiased sample of leptons from Z boson leptonic decays[72; 98]. The baseline data to Monte Carlo scale factors (SF) had a 2.5% uncertainty associated with them. The SFs for electrons and muons show some dependence on the number of jets in the event, as can be seen in Figure 7.4. The jets which are too close to leptons could alter some identification variables. Due to the limited number of Z events with two or more jets, the uncertainties of scale factors blow up at high jet multiplicities. A second method which weights the SF by ΔR^3 between the high- p_T lepton and the closest jet is used ([46]) as a higher statistics test. The weighted averaged SFs using $t\bar{t}$ events agree with the baseline ones used to estimate the top dilepton acceptance. Nevertheless, the uncertainty was doubled to

²Herwig uses the theoretical branching ratio of $\frac{1}{9}$, while Pythia, the PDG([87]) value.

³ $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$

account for possible effects on the SF due to sources other than nearby jet activity. A 5% uncertainty was assigned on the ID SFs.

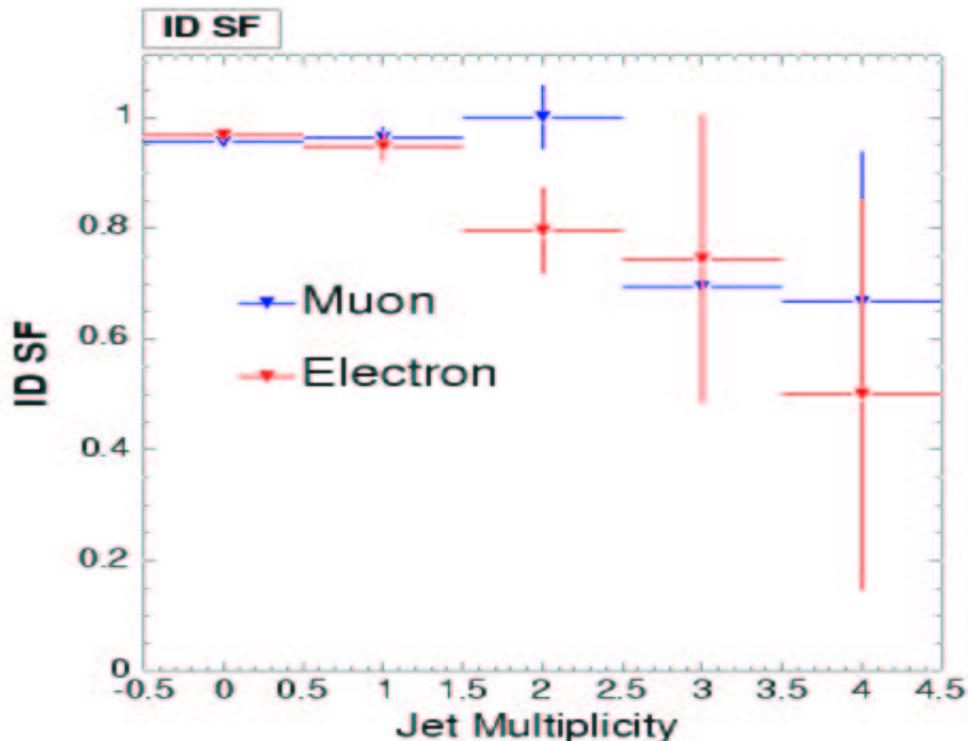


Figure 7.4: Lepton ID data to Monte Carlo scale factor, as a function of jet multiplicities.

Initial State/Final State Radiation

This uncertainty, on how well Pythia Monte Carlo generator models the amount of initial and final state radiation, is determined using two extra Pythia samples: the first has both QED and QCD radiation turned off, while the second uses a different tuning for the underlying event than the default Pythia does. The uncertainty is estimated by taking the sum in quadrature of the relative difference of the two samples with respect to the default Pythia. This results in a systematic error of 1.7 %.

Parton Distribution Function (PDF) Uncertainty

To estimate this uncertainty, the effect on the acceptance is estimated from three contributions: using different PDFs, using a different α_s coupling constant and varying PDFs

within the fit uncertainties. The last contribution was not available at the time the current analysis was sent to publication[4], but it was estimated later and it turns out to be very small and covered by the quoted uncertainty. Each of these contributions are described below.

The relative acceptance difference between Pythia default sample and a Pythia $t\bar{t}$ generated with MRST PDFs, with the same α_s is determined to be 3.03%.

To estimate the effect of varying α_s , three samples with MRST PDFs are used: one is the default sample, used above, and two more with the same PDFs, but lower α_s for the ISR, and FSR respectively. These samples were statistically limited and an upper uncertainty of 2.5 % is used.

A third component to the PDF uncertainty was considered, due to the variation of PDFs around the best fit minimum. Using a Pythia sample, with CTEQ6M PDFs, there are 20 eigenvectors, obtained by diagonalizing the Hessian matrix around the fit minimum. Only three eigenvectors, which give the maximum change in the quark or gluon parton distribution functions are considered. By shifting up and down, by 1σ , the three eigenvectors, one could determine new acceptance, use the semi-differences and sum them in quadrature to get an uncertainty estimate. At the time of writing the thesis the statistics of the available samples was limited and a whole procedure of estimating this uncertainty was being revisited. Also later studies[10] estimated this uncertainty to be around 1%.

Given that between the above three contributions there is some degree of correlation and the sample used were also statistically limited, we felt that it was appropriate to quote 6.0 % as the uncertainty due to PDFs.

Source	Relative Uncertainty (%)
Monte Carlo Generators	5.5
Jet Energy Scale	4.7
Lepton ID SF	2×5.0
ISR/FSR	1.7
PDF	6.0
Total	13.8

Table 7.6: Uncertainties of the $t\bar{t}$ signal acceptance.

7.4 Acceptance summary

The $t\bar{t}$ dilepton acceptance, for a 175 GeV top mass, using a Pythia sample, is

$$\epsilon_{dil} = 0.66 \pm 0.01(stat) \pm 0.09(syst). \quad (7.7)$$

Table 7.7 summarizes the top dilepton acceptances.

Total raw acceptance	$0.78 \pm 0.01 (stat) \pm 0.11 (syst)$
Total weighted acceptance	$0.66 \pm 0.01 (stat) \pm 0.09 (syst)$

Table 7.7: $t\bar{t}$ dilepton acceptance summary.

8 Top Dilepton Backgrounds

In this chapter we discuss the main background sources and estimate their contribution to the observed data candidate dilepton events. Although we are interested in the expected background in the two or more jets bin (the $t\bar{t}$ signal region), we determine the background contributions in the 0 and 1 jet bins as well. This allows to test our predictions in a region where the data sample is comprised of background events, and top contribution is insignificant.

There are various processes which could resemble a top dilepton event in the CDF detector, either because the event has a similar topology or because of object misidentification or false transverse missing energy. Thus the top dilepton backgrounds are classified in two classes:

- Instrumental Backgrounds
 - QCD Fakes
 - Drell-Yan Production
- Physical (Irreducible) Backgrounds
 - Di-tau Production
 - Diboson Production

The physical backgrounds come from events with a top dilepton-like topology, while the instrumental backgrounds are due to the imprecision in measuring \cancel{E}_T or discriminating a lepton from a jet ¹.

8.1 QCD Fake Leptons

The fake dilepton background consists mainly of events where a W -boson is produced in association with jets. One lepton comes from the W -boson decaying leptonically, which also gives rise to missing energy. The other “lepton” is from a jet which *mimics* a lepton. The largest processes which contribute to this background are from the $W + \geq 3$ jets processes². Figure 8.1 shows an example of such a process.

The nature of misidentification for electrons and muons is different. Thus highly energetic jets which contain a photon, a π^0 or an η , could leave most of their energy in the electromagnetic calorimeter. A charged object (like a π^+ or π^-) might produce a track pointing to the electromagnetic cluster, making the jet to look like an electron. There are various ways in which a jet could fake a muon: a punch-through (a hadron which reaches the muon chambers), a decay-in-flight (low P_T kaons that decay just before or in the calorimeter and a high- P_T track, pointing back to the primary vertex is reconstructed) or muons from semileptonic decays. In the later case, usually the muon is inside the jet and fails the isolation requirements.

¹The instrumental backgrounds would not be present if one could build an *ideal detector*, while the physical backgrounds are by no means distinguishable from the signal events, on an event-by-event basis.

²This sample includes small contributions from processes like $Wb\bar{b}$, $Wc\bar{c}$ or Wc , where the *fake* lepton could come from a semileptonic decay. But the dominant contribution is due to a generic jet faking a lepton.

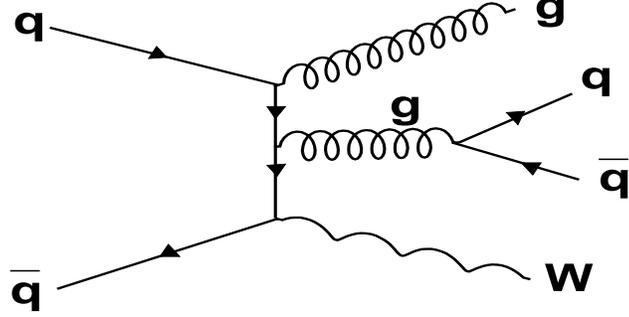


Figure 8.1: Example of a W boson production in association with 3 jets. Such event could fake a top dilepton, if W decays leptonically and one of the three final-state jets fakes a lepton.

A fake lepton charge in a $W+\geq 3$ jets event could be equally positive or negative³, therefore one expects the same level of fake background in same-sign top dilepton events as in opposite-sign dilepton events. With large statistics, one could *completely determine this background by counting the number of same-sign events in the data*. Unfortunately this background is expected to be small for the current dataset, so we could only use the number of same-sign data events as a *cross-check*.

The fake background is estimated using the inclusive jet samples (triggered by a jet with E_T above 20, 50, 70 or 100 GeV, as described in Section 5.6) to calculate the *probability* or the *fake rate* for a *fakeable* jet to “impersonate” a lepton, and apply it to the number of fakeable jets found in the $W+\geq 3$ jet data events ([37]). We measure the electron and muon fake rates for the following categories of electron and muon definitions:

- ICEM: \equiv *isolated CEM electrons*.
- NCEM: \equiv *non-isolated CEM electrons*.
- IMUO: \equiv *isolated muons*, which consist of ICMUP, ICMX, ICMU and ICMP⁴. This category is further divided in tight and loose, as below.

³Any charge-asymmetry is difficult to quantify within the available limited statistics.

⁴Remember that these categories consist of *isolated* CMUP, CMX, CMU and CMP muons.

- TMUO \equiv *tight isolated muons*, which include ICMUP and ICMX.
- LMUO \equiv *loose isolated muons*, which consist of ICMU, ICMP and ICMIO, and
- NMUO \equiv *non-isolated muons*, which consist of NCMUP, NCMX, NCMP and NCMU⁵.

The muons were grouped this way due to limited statistics.

The highest E_T jet in an event is not used, as it is likely to be the trigger jet and could introduce a bias in the fake estimation. To avoid contamination from real leptons from W or Z decays, when calculating the *electron fake rates*, we consider only QCD events with missing transverse energy less than 20 GeV and reject the events with two leptons falling inside the Z mass window ($76 < M_{ll} < 106$ GeV). In the case of the *muon fake rates* calculation, a \cancel{E}_T cut might introduce a serious bias. The $W(\rightarrow e\nu)+0$ jet events are rejected, because we require that the fake lepton not be the highest E_T jet, which is most of the time the trigger jet, i.e. the electron from the W decay. Also only $W(\rightarrow \mu\nu)+\geq 1$ jet events will make it into the jet samples because of the jet trigger requirement. However, a fake muon sometimes induces a large missing energy in an event after the \cancel{E}_T correction. Therefore a \cancel{E}_T cut could bias the fake rates and it is appropriate not to make such a requirement. Also if there is a signal contamination, it will likely decrease the higher the jet E_T threshold is. This is due to the fact the jets in W+multijet samples largely come from ISR/FSR and their E_T spectra falls exponentially and is softer than a b-jet from a top quark decay. A systematic uncertainty is assigned to the background estimate, due to the effect of the missing energy cut on the fake rates.

⁵Remember that these categories consist of *non-isolated* CMUP, CMX, CMU and CMP muons.

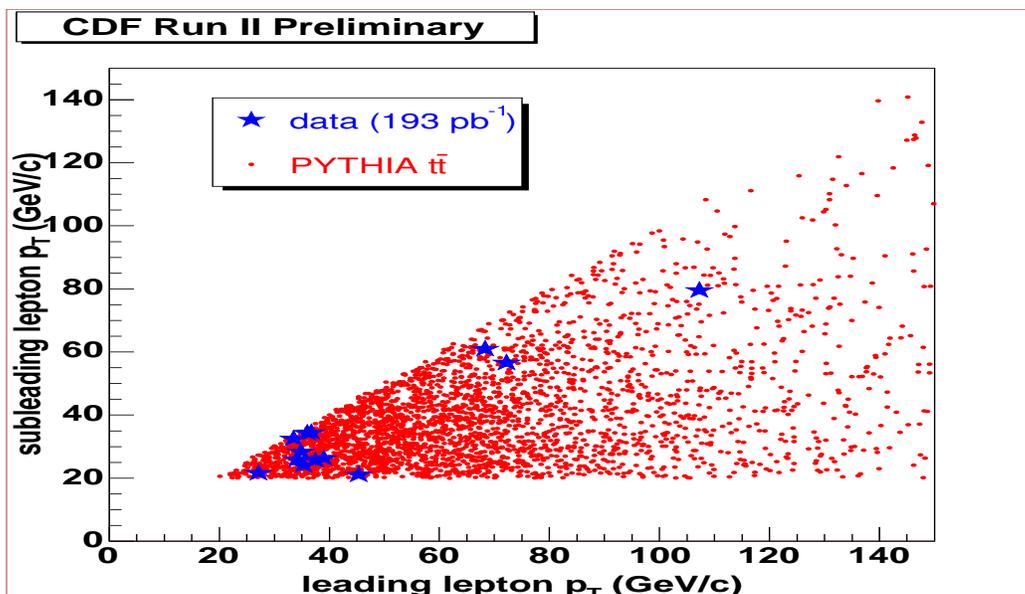


Figure 8.2: Missing energy for events with only one isolated (ICEM) or non-isolated (NCEM) central electron, from the inclusive sample $jet50$, defined in Section 5.6. The bump around 40 GeV shows a significant contribution from W events, with W decaying leptonically.

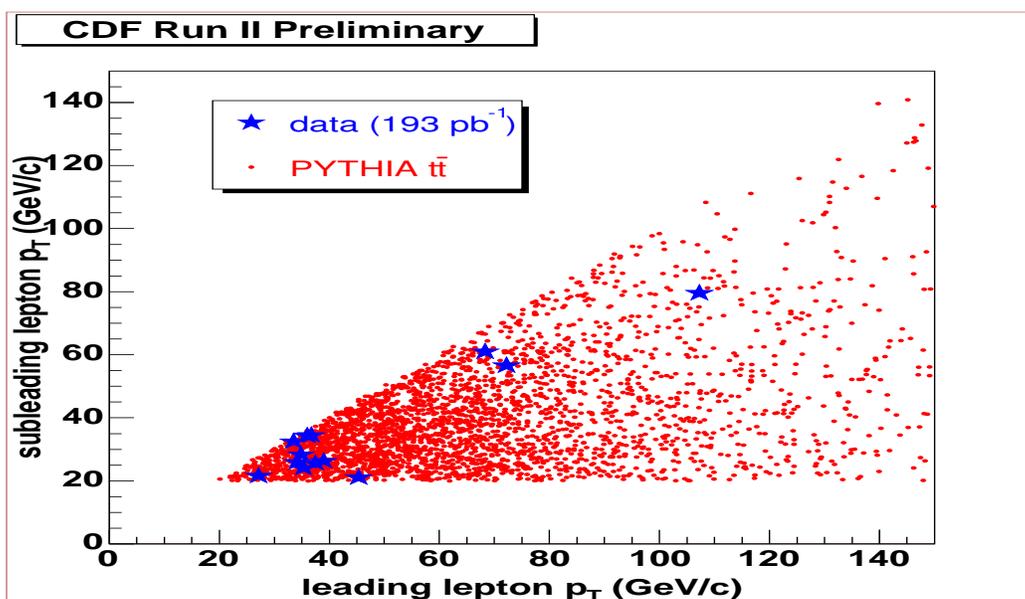


Figure 8.3: Missing energy for events with only one isolated (IMUO) or non-isolated (NMUO) muon, from inclusive sample $jet50$. There is no obvious contribution from W events, with W decaying leptonically.

Figures 8.2 and 8.3 show the \cancel{E}_T distribution for events with a fake lepton, in the *jet50* sample. One could see a clear W signal in electron case, an excess of events with $\cancel{E}_T \approx 40$ GeV, however most of the events are in 0 jet bin. In the muon case only $W(\rightarrow \mu\nu) + \geq 1$ jet events will make it into the jet samples. The number of fake leptons in jet samples is limited and therefore a precise determination of the fraction of real leptons is a challenging task.

The *fake rate* is defined as the number of observed fake leptons divided by the number of fakeable jets. It is highly dependent on the definition of the *denominator*. A *fakeable jet object* is defined (see Table 8.1) in such a way that the kinematics of the fakeable jet are defined in same manner as the lepton it fakes. One advantage of this definition is that it allows one to parametrize the fake rate as a function of the fakeable objects energy or momentum, and the charge of the fakeable jet is well defined (except sometimes for a PHX electron fakeable object). Another possibility is to use the standard jet kinematics definitions for the jet fakeable object; this method ([22]) was shown to give a fake estimate in agreement with the one used in this thesis.

Lepton Type	CDF Object	Fakeable Jet Criteria
ICEM	CdfEmObject	$E_T > 20$ GeV $E_{had}/E_{em} < 0.125 \parallel E_T > 100$ GeV IsoFrac < 0.1 At least a track
NCEM	CdfEmObject	$E_T > 20$ GeV $E_{had}/E_{em} < 0.125 \parallel E_T > 100$ GeV At least a track
PHX	CdfEmObject	$E_T > 20$ GeV $E_{had}/E_{em} < 0.125 \parallel E_T > 100$ GeV IsoFrac < 0.1 $1.0 < \eta_{PES} < 2.0$
TMUO, LMUO	CdfMuonObject	$P_T > 20$ GeV $E/P < 1$ IsoFrac < 0.1
NMUO	CdfMuonObject	$P_T > 20$ GeV $E/P < 1$

Table 8.1: Jet fakeable object definitions.

The fake rates obtained from different samples agrees well within statistics (see Table

8.2) and for the final estimates, the *jet50* sample predictions are used. This sample has the largest statistics and the fakeable jet transverse energy or momentum spectrum agrees well with the one from $W+\geq 3$ jets data.

Category	Data Sample	Observed	Predicted (Average)	Predicted (E_T vs CalFracIso)
PHX	jet20	51	49.54 ± 4.81	49.25 ± 5.82
	jet70	75	56.75 ± 5.51	64.89 ± 8.61
	jet100	69	65.82 ± 6.39	114.16 ± 30.81
ICEM	jet20	15	13.10 ± 2.73	14.86 ± 3.79
	jet70	14	14.46 ± 3.014	13.08 ± 6.42
	jet100	24	20.39 ± 4.25	10.87 ± 33.52
NCEM	jet20	34	31.46 ± 2.55	32.07 ± 3.19
	jet70	63	97.89 ± 7.94	84.78 ± 14.50
	jet100	67	111.85 ± 9.07	77.00 ± 69.84
TMUO	jet20	22	28.93 ± 5.196	33.98 ± 7.39
	jet70	33	31.46 ± 5.65	30.32 ± 5.62
	jet100	69	76.63 ± 13.76	76.29 ± 13.99
LMUO	jet20	18	34.53 ± 5.67	37.33 ± 7.43
	jet70	21	37.55 ± 6.17	36.60 ± 6.20
	jet100	50	91.46 ± 15.04	91.52 ± 15.38
NMUO	jet20	72	58.37 ± 4.54	72.68 ± 8.78
	jet70	88	134.98 ± 10.51	109.98 ± 25.73
	jet100	100	237.81 ± 18.51	157.41 ± 147.16

Table 8.2: Number of observed and predicted fakes for electron categories. We use jet50 to predict jet20, jet70 and jet100. Two predicted numbers are shown: one using the average fake ratio from jet50 and another based on a 2D parametrization of the fake rate: Isolation Ratio and E_T .

First the fake contribution is estimated after all the selection criteria, including \cancel{E}_T , but before any jet requirements. Then the probability $\epsilon(\cancel{E}_T)$ that a fake event will have 0, 1 or more jets is determined as a whole, by counting together the events with a fake electron or muon, to overcome the limited statistics available in each subcategory. Also a similar probability is calculated for the event to pass H_T and have opposite sign leptons. Table 8.3 contains the expected number of fake background events in different top dilepton event jet bins, after the H_T and opposite sign requirements.

	0 jet	1 jet	≥ 2 jets	H_T	OS
Efficiency w.r.t $\cancel{E}_T, \epsilon(\cancel{E}_T)$	0.45 ± 0.01	0.35 ± 0.01	0.20 ± 0.01	0.16 ± 0.01	0.09 ± 0.01
ELE-PHX	0.45 ± 0.23	0.35 ± 0.18	0.20 ± 0.10	0.16 ± 0.08	0.05 ± 0.06
MUO-PHX	0.55 ± 0.24	0.43 ± 0.19	0.24 ± 0.11	0.19 ± 0.08	0.05 ± 0.06
ELE-ICEM	0.33 ± 0.28	0.26 ± 0.22	0.14 ± 0.12	0.11 ± 0.10	0.09 ± 0.09
MUO-ICEM	0.17 ± 0.15	0.13 ± 0.12	0.08 ± 0.07	0.06 ± 0.05	0.04 ± 0.05
TELE-NCEM	0.63 ± 0.38	0.50 ± 0.30	0.28 ± 0.17	0.22 ± 0.13	0.12 ± 0.11
TMUO-NCEM	0.42 ± 0.34	0.33 ± 0.26	0.19 ± 0.15	0.15 ± 0.12	0.08 ± 0.10
ELE-TMUO	1.10 ± 0.35	0.87 ± 0.27	0.49 ± 0.16	0.39 ± 0.12	0.20 ± 0.09
MUO-TMUO	0.33 ± 0.20	0.26 ± 0.16	0.15 ± 0.09	0.12 ± 0.07	0.07 ± 0.06
TELE-LMUO	0.66 ± 0.26	0.52 ± 0.20	0.29 ± 0.11	0.23 ± 0.09	0.11 ± 0.07
TMUO-LMUO	0.18 ± 0.15	0.14 ± 0.11	0.08 ± 0.06	0.06 ± 0.05	0.04 ± 0.05
TELE-NMUO	0.79 ± 0.57	0.62 ± 0.44	0.35 ± 0.25	0.28 ± 0.20	0.18 ± 0.20
TMUO-NMUO	0.34 ± 0.48	0.27 ± 0.37	0.15 ± 0.21	0.12 ± 0.17	0.08 ± 0.16
Total	5.95 ± 1.06	4.68 ± 0.84	2.66 ± 0.48	2.09 ± 0.38	1.10 ± 0.33

Table 8.3: Expected number of fake background events in 193 pb^{-1} . The errors for each top dilepton category are *statistical only*.

The fake estimate is compared with the number of same-sign events. Table 8.4 shows good agreement between the number of SS data events and the expected contribution from two sources: fake background and the effect of the PHX charge misidentification⁶ on physical backgrounds (diboson and Drell-Yan mainly).

Same-Sign contribution	0 jet	1 jet	≥ 2 jet
SS Predicted	2.51 ± 0.48^{stat}	1.94 ± 0.37^{stat}	1.01 ± 0.20^{stat}
SS PHX Charge Fake	0.61 ± 0.25	0.26 ± 0.1	0.08 ± 0.03
SS Observed	3	2	0

Table 8.4: Fake background prediction for same-sign events, in different jet bins. Also the number of same-sign events in the data is shown. There are *two ways* to get a same-sign event: *either* from fake background, when a $W \geq 3$ event fakes a dilepton, *or* an opposite-sign event has a charged misidentified PHX electron.

Systematic Uncertainties

⁶The PHX electrons have the charge determined wrong about 13% of the time, at large pseudorapidity η . The details are given in Section 6.11.

There are four sources of systematic uncertainties in the fake lepton background calculation, summarized in Table 8.5:

- The *Method* uncertainty is the statistical uncertainty on the fake estimate and is mainly due to the lack of statistics when binning of fake rate in E_T and Isolation. If one would use an average fake rate this will be drastically reduced. However the parametrization of the fake rates significantly improves the agreement between the observed fakes in jet20, jet70 and jet100 predicted by jet50 for non-isolated categories..
- Fake rates vary depending on which jet sample one uses; the fake estimates for different jet samples are compared before applying the “2 jet” requirement and we take half the sum of minimum/maximum difference with respect to the central value, calculated from jet50, as a systematic uncertainty.
- To remove the real leptons from W boson decays we apply a \cancel{E}_T cut for events with fake electrons, but not with fake muons for reasons already mentioned. By applying \cancel{E}_T or not simultaneously for both fake e’s and μ ’s events, we see a 20% variation in the fake background. We take this as a systematic uncertainty.
- Because the fake rates derived from the various jet samples predict different shapes for the fake E_T spectra, their predicted H_T shapes, and hence H_T cut efficiencies, will differ. *To quantify* by how much, we build an H_T distribution where each W+jets event gets weighted by the probability for it to yield a fake lepton, where the probability comes from the fake rate determined from jet samples jet20, jet50, jet70 or jet100. This gives us four different H_T distributions shown in Figure 8.4. In fact, each bin corresponds to the fake estimate in that H_T range. The H_T cut efficiency can vary at most by 14 % depending on which jet sample is used to determine the shape. We take this as a systematic uncertainty.

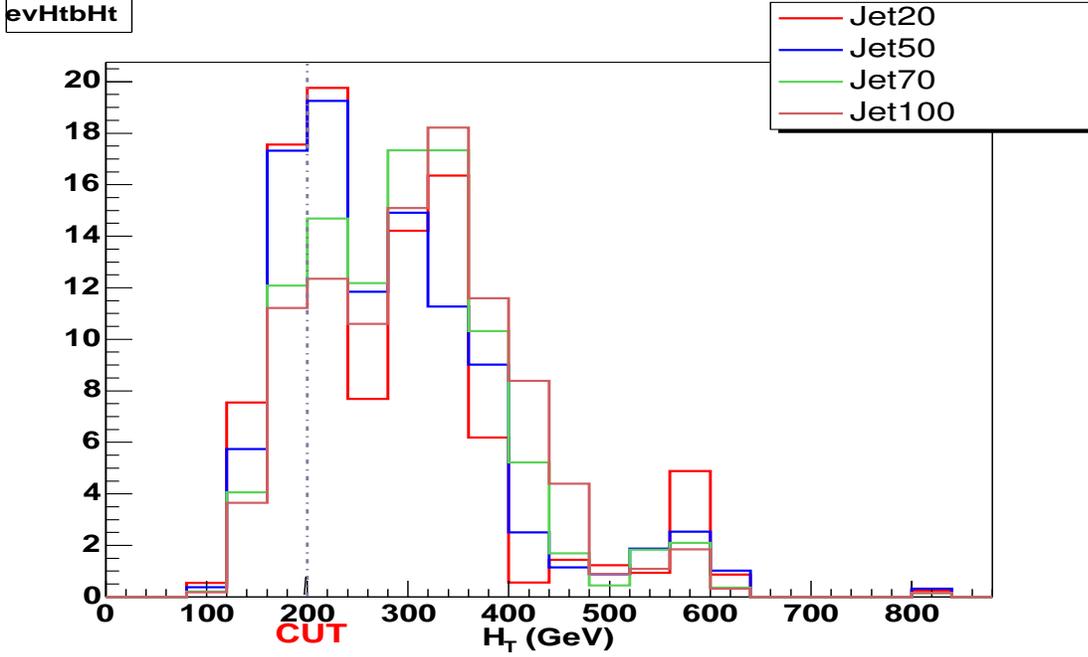


Figure 8.4: For these H_T distributions, each $W+\geq 3$ jets event is weighted by the probability for it to yield a fake. The fake rates are determined from jet20, jet50, jet70 or jet100. All four distributions are normalized to an area equal of 100.

Source	Relative Uncertainty (%)
Method	31
Different Jet Samples	9
Residual real electrons	20
H_T Shape	14
Total	40.5

Table 8.5: Fake background systematic uncertainties.

8.2 Drell-Yan lepton pair production

In hadronic collisions the quarks and anti-quarks could annihilate to produce a virtual photon, or a Z boson, which decays to a pair of electrons or muons. The inclusive hadronic reaction $A+B \rightarrow \ell^+\ell^- + \text{anything}$, where $\ell = e$ or μ , is known as the *Drell-Yan (DY) process* [24]. The DY cross-section is dominated by the lowest order diagram, with a pair of leptons

and no jets in the final state. Only the higher order processes, with at least two jets from QCD radiation (Figure 8.5), could mimic a $t\bar{t}$ dilepton topology. Also the higher mass DY cross-section is dominated by the Z contribution close to the Z^0 pole region. Therefore one uses different strategies to remove DY background *inside the Z^0 mass window* versus *outside the mass window*. *Inside* the mass windows of $76 < M_{\ell\ell} < 106$ GeV (where $\ell = e$ or μ), we impose tighter requirements. *Outside* this region, there is very little DY background. In general, the outside the mass window contribution comes about when the lepton E_T/P_T is mismeasured.

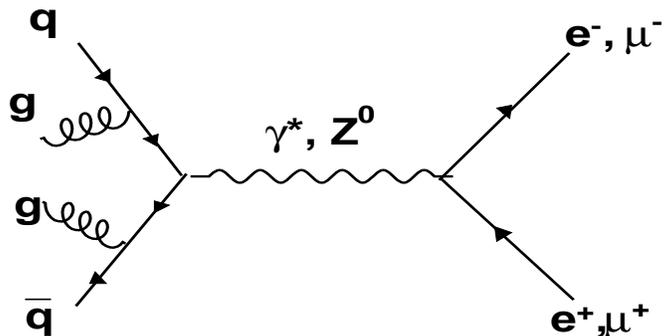


Figure 8.5: Example of a Drell Yan production Feynman diagram, in association with two jets, from ISR.

Drell-Yan events *have no inherent missing transverse energy* (\cancel{E}_T), since there are no neutrinos in the final state. The fake missing energy comes from the mismeasurements of tracks in the calorimeter (MIP) and the mismeasurements of jets due to the detector cracks or other effects. In both cases the \cancel{E}_T will point in the direction of the mismeasured track or jet object. This background is very dependent on the tails of the \cancel{E}_T distribution in DY with multi-jets events. The first step to reduce this background is to correct, using our best information, the raw missing energy, event by event. The list of jet corrections is presented in Section 6.9.

The Drell-Yan background estimate is determined from Z events with large \cancel{E}_T in the data, combined with a few cut efficiencies determined from Monte Carlo. In an earlier Run II iteration of the $t\bar{t}$ cross-section measurement ([29]), when a larger η coverage for

PHX electrons was allowed [26], the Monte Carlo was normalized to the data in each jet bin; however, with the smaller η coverage this method suffers greatly from lack of statistics. Therefore a different approach (similar to [28]) was adopted, where the data is integrated over all jet bins and then the events are distributed in jet bins, based on Monte Carlo expectation. The H_T efficiency is taken from Monte Carlo. For dielectron and dimuon events inside the mass window (76, 106) GeV there is a *jet significance requirement*. The background estimate has two parts, inside and outside the mass window:

$$N_{\text{DY}}^{\ell\ell} = N_{\text{DY}}^{\ell\ell, \text{in}} + N_{\text{DY}}^{\ell\ell, \text{o}}. \quad (8.1)$$

where $\ell = e$ or μ , and $N_{\text{DY}}^{\ell\ell, \text{in}}$ is the background inside the mass window, and $N_{\text{DY}}^{\ell\ell, \text{out}}$ is the background outside the mass window.

Specifically the formula can be further expanded into jet bins as

$$N_{\text{DY}}^j = f_{Z_{\text{veto}}}^j N_{Z_{\text{veto}}} + f_{\cancel{E}_T}^j R_{\text{o}/\text{i}}^j N_{\cancel{E}_T} \quad (8.2)$$

where the superscript j refers to the jet bin ($j = 0, 1$ or 2 , where 2 refers to ≥ 2 jets), the first term in the sum corresponds to $N_{\text{DY}}^{\ell\ell, \text{in}}$, and the second term in the sum corresponds to $N_{\text{DY}}^{\ell\ell, \text{out}}$. The f^j 's gives the relative fraction of each jet bin for the “in” and “out” categories, the subscript “ \cancel{E}_T ” means “after the \cancel{E}_T cut”, the subscript “ Z_{veto} ” means “after the \cancel{E}_T and Z_{veto} cuts,”⁷ $R_{\text{o}/\text{i}}$ means the ratio of number of events outside to inside the Z mass window after the \cancel{E}_T cut, and the N 's refer to the number of Z events observed in the data.

The \cancel{E}_T distributions for events with Z's passing a loose selection criteria are shown in Figure 8.6. Note the very small statistics at large \cancel{E}_T , which makes it impossible to look further at the “2 jet” bin events. These events contain a non-negligible contribution from $t\bar{t}$, WW, WZ, and DY ditau events, which are subtracted from the number of Z loose data events⁸. These contributions are determined in Section 8.2.1 and shown in Table 8.8.

⁷The “ \cancel{E}_T ” cut is $\cancel{E}_T > 25$ GeV & $[\Delta\phi(\cancel{E}_T, \text{nearest } \ell \text{ or } j) > 20^\circ \mid \cancel{E}_T > 50$ GeV (see Section 6.7)]. The “ Z_{veto} ” cut is jetSig > 8 & $\Delta\phi(\cancel{E}_T, \text{nearest } j) > 10^\circ$ (see Section 6.8).

⁸A 30% uncertainty is assumed on these non-Z contributions.

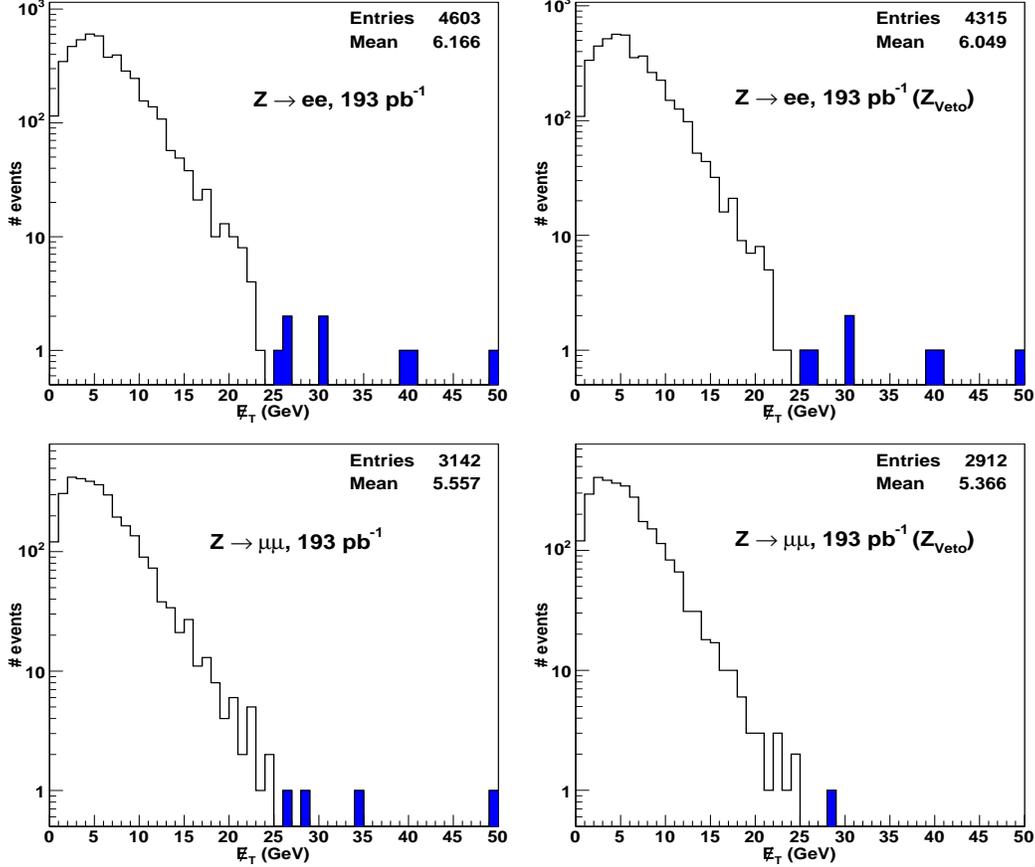


Figure 8.6: \cancel{E}_T distribution for Z loose events in the data. The $[\Delta\phi(\cancel{E}_T, \text{nearest } \ell \text{ or } j) > 20^\circ \mid \cancel{E}_T > 50 \text{ GeV}]$ cut is already applied. Top plots are for electrons; bottom plots are for muons. Left plots are without the Z_{veto} cut; right plots are with the Z_{veto} cut. The entries in blue show the events with $\cancel{E}_T > 25 \text{ GeV}$. Overflows are shown in the last bin.

The Pythia DY MC (datasets `ztop0e` and `ztop0m`) are used to determine the fraction of events in different jet multiplicity bins and the ratios of number of events outside to inside, shown in Figures 8.7 and 8.8. The jet multiplicity distributions are sculpted by the \cancel{E}_T and Z_{veto} cuts, as seen in Figure 8.7. In the events with no jets the \cancel{E}_T is primarily generated by mismeasured leptons. This can be seen by looking at the invariant mass distributions of Figure 8.8. If the leptons are well-measured, there has to be mismeasured jets in the events to create large fake \cancel{E}_T , hence the higher jet multiplicities in events with large \cancel{E}_T . The Z_{veto} cut then biases back toward small jet multiplicities; the probability for a small jetsig increases the more jets there are in the event and jetsig is undefined for events with no jets. These events pass jetsig by definition.

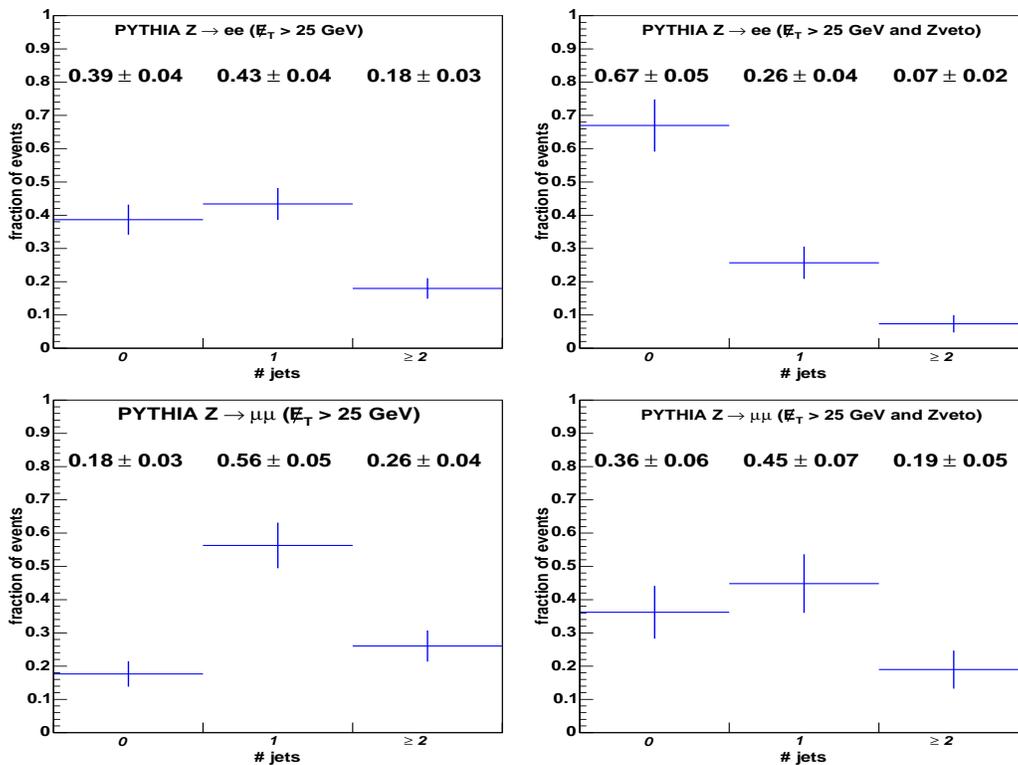


Figure 8.7: Jet multiplicity distributions for Pythia DY events. The E_T cut is applied. Overflows are shown in the last bin. Top plots are for electrons; bottom plots are for muons. Left plots are without the Z_{veto} cut; right plots are with the Z_{veto} cut.

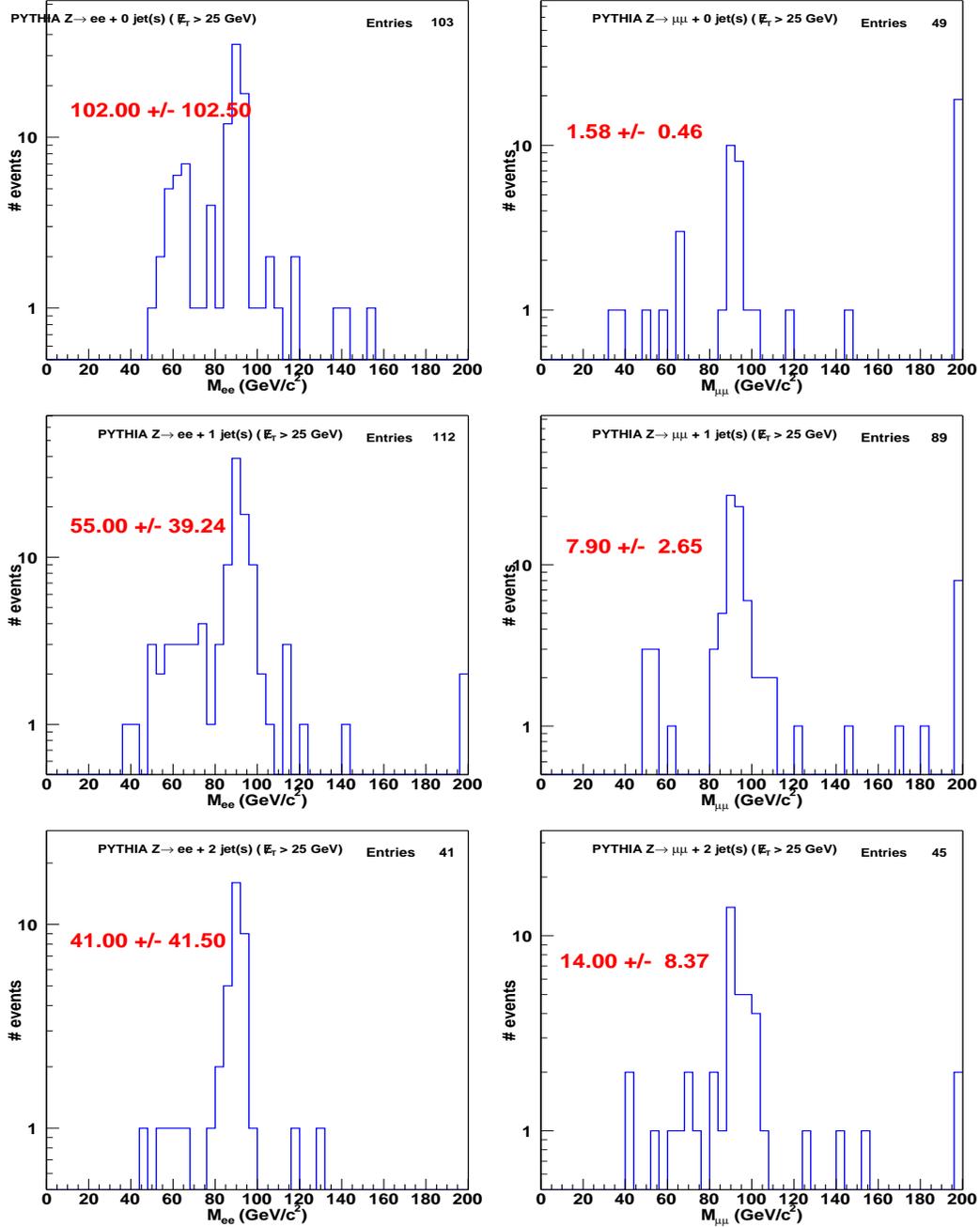


Figure 8.8: Dilepton invariant mass distributions from the Pythia DY Monte Carlo. E_T cut is applied. Left plots are for electrons, right plots are for muons. Top to bottom are for 0, 1, and ≥ 2 jets bins. Also shown is the ratio of events outside/inside the Z window. Overflows are shown in the last bin

DY events with ≥ 2 jets are further subject to the H_T cut, its efficiency being determined from Monte Carlo. Furthermore, the H_T cut depends strongly on the dilepton

invariant mass, so its efficiency has to be determined separately for the two mass regions (outside/inside). For this purpose the AlpGen DY+2p MC samples $\text{atop}\{64,23,61\}$ (electrons) and $\text{atop}\{65,27,52\}$ (muons), described in Table 8.6, were used. Table 8.7 summarizes the H_T cut efficiencies.

Dataset ID	Sample	Cross-section (pb) at LO
atop64	$Z \rightarrow ee + 2p, M_Z \in (10,75)$	50.20 ± 0.09
atop23	$Z \rightarrow ee + 2p, M_Z \in (75,105)$	23.34 ± 0.04
atop61	$Z \rightarrow ee + 2p, M_Z \in (105,800)$	0.631 ± 0.001
atop65	$Z \rightarrow \mu\mu + 2p, M_Z \in (10,75)$	50.20 ± 0.09
atop27	$Z \rightarrow \mu\mu + 2p, M_Z \in (75,105)$	23.34 ± 0.04
atop52	$Z \rightarrow \mu\mu + 2p, M_Z \in (105,800)$	0.631 ± 0.001

Table 8.6: AlpGen DY+2p MC samples used to determine the H_T cut efficiency.

	in	out
ee	0.53 ± 0.13	0.89 ± 0.02
$\mu\mu$	0.82 ± 0.15	0.75 ± 0.02

Table 8.7: H_T cut efficiencies taken from AlpGen DY+2p MC. Uncertainties are statistical only.

Finally we apply the opposite sign cut, which in principle should be 100% efficient for DY events. However, PHX electrons have a non-zero charge misidentification rate (see Section 6.11) and this will affect the CP portion of the DY background. CP accounts for about half, as determined from the ratio of CC/CP Z s in the data.

8.2.1 Correction for the $t\bar{t}$, WW, WZ and $Z \rightarrow \tau\tau$ contributions

The Z mass region is dominated by the DY contribution, but there are also *smaller portions* from $t\bar{t}$, WW/WZ and $Z \rightarrow \tau\tau$, which have to be subtracted from the number of loose Z candidates with large \cancel{E}_T seen in the data. The estimates are performed using Monte Carlo samples and are given in Table 8.8, by dilepton flavor.

Process	ee		$\mu\mu$	
	\cancel{E}_T cut	\cancel{E}_T, Z_{veto}	\cancel{E}_T cut	\cancel{E}_T, Z_{veto}
WW	0.83	0.82	0.62	0.61
WZ	0.73	0.68	0.56	0.53
$t\bar{t}$	0.65	0.56	0.66	0.55
$Z \rightarrow \tau\tau$	0.04	0.04	0.02	0.02

Table 8.8: Expected number of non-DY events inside the Z window. A 30% uncertainty is assumed on these non-DY contributions.

Systematic Uncertainties

There are two main systematic uncertainties:

1. The largest uncertainty is due to the limited number of loose Z events in data with large \cancel{E}_T and it is 100%. It is due to just a very few data events, with $\cancel{E}_T > 25$ GeV.
2. H_T cut can be particularly susceptible to the jet energy scale; a 20% effect was seen on these efficiencies when fluctuating jets up and down in E_T by 1σ of their resolution and is taken as a systematic.

The DY background is summarized in Table 8.9.

	0j	1j	$\geq 2j$	H_T	OS
ee	4.2 ± 1.9	2.2 ± 0.9	0.57 ± 0.27	0.38 ± 0.28	0.36 ± 0.27
$\mu\mu$	0.2 ± 0.7	0.1 ± 0.6	0.11 ± 0.34	0.07 ± 0.34	0.07 ± 0.34
$\ell\ell$	4.4 ± 2.0	2.2 ± 1.1	0.69 ± 0.44	0.45 ± 0.44	0.43 ± 0.43

Table 8.9: Summary of DY background.

8.3 Di-tau production

The di-tau production from the decay of Z bosons could be a source of irreducible dilepton background if both τ 's decay leptonically. However the need for at least two jets from QCD radiation (Figure 8.9) reduces drastically this background. The branching ratio

of the leptonic decay of a τ , $\text{BR}(\tau \rightarrow \ell\nu_\tau\nu_\ell, \ell = e \text{ or } \mu) = 18\%$ ([87]), implies that $\text{BR}(Z^0 \rightarrow \tau\tau \rightarrow \ell\nu_\tau\nu_\ell\ell\nu_\tau\nu_\ell) = (2 \times 18\%)^2 = 13\%$.

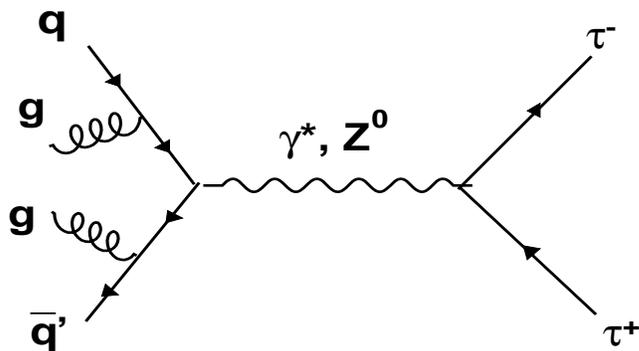


Figure 8.9: Example of a di-tau production Feynman diagram in association with two jets.

For this background estimate we use a *Pythia 6.2* $\gamma^*/Z \rightarrow \tau^-\tau^+$ sample. 48 million events were generated, each γ^*/Z with a CM energy ≥ 30 GeV. Then the events are filtered to keep only those with at least 2 leptons with $E_T > 17$ GeV and $|\eta| < 2.5$. This reduces our sample to about 1.0 % of the initial size (422,363 events). The estimated cross-section for the generated process is 335.3 ± 5.7 pb, with $\sim 2\%$ uncertainty, coming primarily from the uncertainty on the PDF calculation [12]. The effective luminosity of this sample is 143.2 ± 2.4 fb $^{-1}$. The events are processed through the dilepton analysis selection and the detailed breakdown is shown in Tables B.5- B.8. The summary of cut efficiencies is presented in Table 8.10.

Cut	Number of events passing	Efficiency (%)
$\epsilon_{geom.Pt}$	112697	-
ϵ_{ID}	69606	61.76 ± 0.1448
ϵ_{iso}	68278	98.09 ± 0.05
ϵ_{Conv}	67222	98.45 ± 0.05
ϵ_{mass}	67136	99.87 ± 0.01
$\epsilon_{\cancel{E}_T}$	3335	4.93 ± 0.08
$\epsilon_{\Delta\Phi}$	1232	36.94 ± 0.84
ϵ_{2-jet}	371	30.11 ± 1.31
ϵ_{H_T}	232	62.53 ± 2.51
ϵ_{OS}	231	99.57 ± 0.43

Table 8.10: Efficiencies of the selection cuts applied to a Pythia $Z \rightarrow \tau\tau$ sample. Each efficiency is calculated from the number of events that have passed all the preceding cuts. The errors are statistical only.

To get the background estimate, a *correction factor* F_{2j} , equal to the ratio of the fraction of events with ≥ 2 jets, in data ($f_{2j}(data)$) over MC ($f_{2j}(MC)$), is applied. This is necessary as Pythia is a leading order (LO) generator, with the jets coming from inverse parton showering, therefore it *underestimates* the amount the radiation observed in the data. F_{2j} is determined by comparing the fraction of events with ≥ 2 jets, in $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu\mu$ data samples with Monte Carlo. The results are summarized in Table 8.11. Also, a similar scale factor is determined for 1 jet bin.

Dilepton Channel	Njets	Data $f_{2j}(data)$	MC $f_{2j}(MC)$	Ratio (Data/MC) F_{2j}
ee	1 jet	$15.1 \pm 0.8 \%$	$11.8 \pm 0.08 \%$	1.28 ± 0.07
ee	2 jet	$3.0 \pm 0.4 \%$	$1.92 \pm 0.03 \%$	1.56 ± 0.21
$\mu\mu$	1 jet	$15.0 \pm 0.7 \%$	$13.3 \pm 0.11 \%$	1.13 ± 0.05
$\mu\mu$	2 jet	$3.9 \pm 0.4 \%$	$2.25 \pm 0.05 \%$	1.73 ± 0.18
Average	1 jet	$15.0 \pm 0.5 \%$	$12.5 \pm 0.01 \%$	1.13 ± 0.04
Average	2 jet	$3.5 \pm 0.3 \%$	$2.085 \pm 0.03 \%$	1.68 ± 0.15

Table 8.11: Fraction of events with 1 or ≥ 2 jets for data and Pythia Monte Carlo.

Systematic Uncertainties

There are two systematic uncertainties for the $Z \rightarrow \tau\tau$ background estimate, briefly

discussed below:

1. The largest uncertainty comes from the jet energy scale; a 29% variation is seen on the background estimate when shifting jets up and down by 1σ of the energy scale, and is taken as a systematic.
2. A 10 % uncertainty due to the uncertainty on the 2-jet scale factor, F_{2j} , is assigned.

The background estimate in 193 pb^{-1} is shown in Table 8.12.

	0jets	1jet	≥ 2 jets	$H_T + \text{OS}$
ee	0.05 ± 0.01	0.22 ± 0.07	0.15 ± 0.05	0.09 ± 0.03
$e\mu$	0.11 ± 0.03	0.45 ± 0.14	0.36 ± 0.11	0.22 ± 0.07
$\mu\mu$	0.04 ± 0.01	0.19 ± 0.06	0.17 ± 0.04	0.11 ± 0.03
Total	0.19 ± 0.06	0.86 ± 0.26	0.69 ± 0.21	0.42 ± 0.13

Table 8.12: $Z \rightarrow \tau\tau$ background estimate for 193 pb^{-1} , in different jet bins. The systematic and statistical uncertainties are included.

8.4 Vector boson pair production

The hadronic production of a vector boson pair ($VV = WW, WZ$ or ZZ), followed by V decaying leptonically, $p\bar{p} \rightarrow VV \rightarrow \ell\ell + X$, could produce a final state indistinguishable from a top dilepton event, on an event-by-event basis.

The lowest order diagrams are shown in Figures 8.10 and 8.11. The latest calculations of the diboson production cross sections, at next-to-leading order (NLO), predict the following values at $\sqrt{s} = 1.96 \text{ TeV}$ ([59], [58], [101])⁹:

$$\begin{aligned}
 \sigma(p\bar{p} \rightarrow WW) &= 13.25 \text{ pb} \\
 \sigma(p\bar{p} \rightarrow WZ) &= 3.78 \text{ pb} \\
 \sigma(p\bar{p} \rightarrow ZZ) &= 1.43 \text{ pb}
 \end{aligned}
 \tag{8.3}$$

⁹To be precise the ZZ and WZ cross sections are calculated at $\sqrt{s} = 2 \text{ TeV}$ using MCFM program with CTEQ LHAPDF.

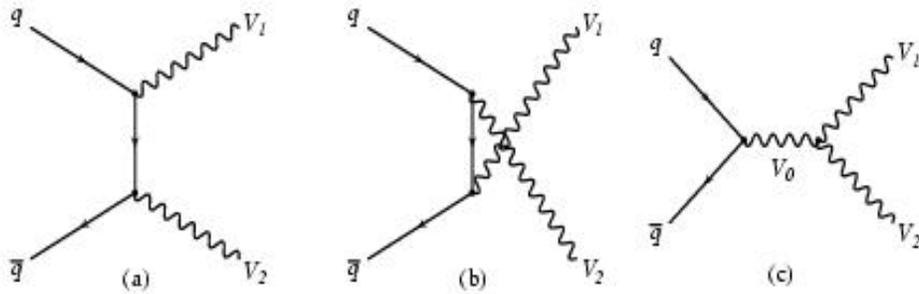


Figure 8.10: Leading order Feynman diagrams for the vector boson pair production. The V_0 , V_1 , and V_2 are assigned as following: $V_0 = \gamma$ or Z , $V_1 = W^+$, and $V_2 = W^-$ for WW production and $V_0 = V_1 = W$ and $V_2 = Z$ for WZ production.

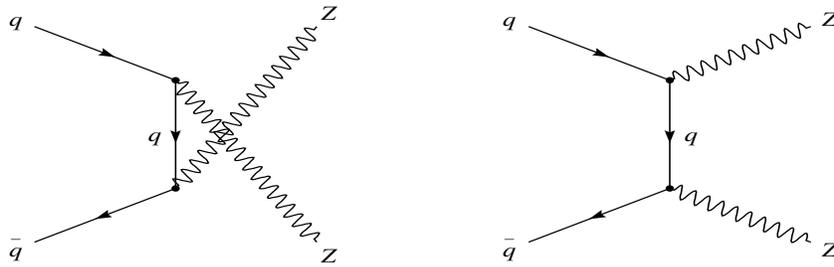


Figure 8.11: Feynman diagrams for the tree level processes contributing to $p\bar{p} \rightarrow ZZ$ in SM.

The largest diboson background in the $t\bar{t}$ dilepton channel comes from WW production. The contribution from ZZ is significantly smaller than the one from WW in the signal region. The estimate of WW background is discussed extensively in the next section, while the WZ and ZZ contributions are briefly mentioned.

8.4.1 WW Production

WW pair production in association with at least two jets from initial or final state radiation may pass the dilepton selection criteria if both W bosons decays leptonically. Such final state will also have real \cancel{E}_T from the two neutrinos. The process is $p\bar{p} \rightarrow WW + \geq 2$ jets $\rightarrow l\nu l\nu + \geq 2$ jets, and is shown in Figure 8.12.

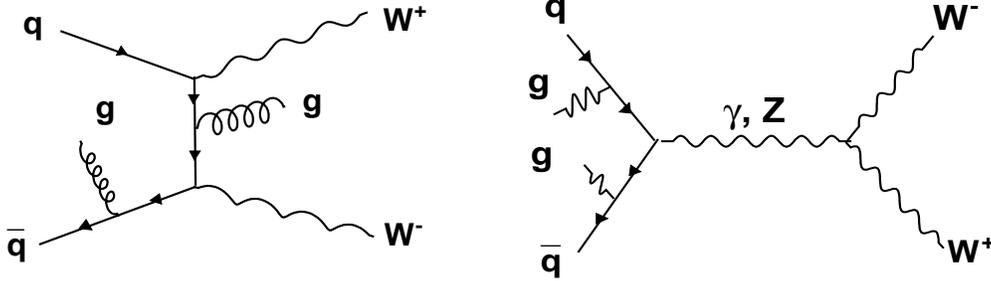


Figure 8.12: Example of a WW production in association with two jets.

The WW background is estimated using a large WW Pythia sample, referred to within CDF as *wtop0f*, generated with Pythia using the 4.9.1 version of the CDF offline software. It consisted of 828,000 WW events where both Ws were forced to decay leptonically: $W \rightarrow \ell$, $\ell = e, \mu$ or τ . The equivalent inclusive WW sample size before decaying the Ws was 7.45 million events. Using the $\sigma_{WW}^{NLO} = 13.25 \text{ pb}$ [59], and taking account of the branching ratio for both Ws to decay leptonically, the effective luminosity of the *wtop0f* sample is $\sim 557 \text{ fb}^{-1}$.

The expected number of WW events passing the $t\bar{t}$ dilepton selection criteria can be written as:

$$N^{WW} = \sigma(p\bar{p} \rightarrow WW) \cdot Br(WW \rightarrow \ell\nu\ell\nu) \times L \times \epsilon_{\cancel{E}_T}^{WW} \times \epsilon_{2j}^{WW} \times \epsilon_{H_T, OS}^{WW} \quad (8.4)$$

where $\epsilon_{\cancel{E}_T}^{WW}$ is the efficiency of all the cuts, except the 2-jet, H_T and OS cuts. ϵ_{2j}^{WW} is the efficiency of 2-jet cut, after all the previous cuts were applied (including \cancel{E}_T and $\Delta\phi$) and $\epsilon_{H_T, OS}^{WW}$ is the efficiency of H_T and OS cuts, with respect to the number of events passing all the other cuts. It is expected that Pythia will not be able to reproduce the fraction of events with at least two jets seen in the data, therefore a correction to the ϵ_{2j}^{WW} efficiency is necessary. This is determined using $Z^0 \rightarrow e^+e^-$ and $Z^0 \rightarrow \mu^+\mu^-$ data, compared with Pythia Monte Carlo samples as was done for $Z \rightarrow \tau\tau$. The conclusions are summarized in Table 8.12. The Pythia WW sample is run through the dilepton analysis and the detailed breakdown of the events surviving the selection requirements is shown in Tables B.9-B.12. Also, for PHX electrons, a correction due to the PHX fake charge rate is applied.

Systematic Uncertainties

There are two main systematic uncertainties to the calculation of the WW background:

1. The largest uncertainty is due to the jet energy scale; a 20% effect was seen on these efficiencies when shifting jets up and down by 1σ , and this is taken as a systematic.
2. By comparing the background estimate using Pythia MC with the estimate from Alpgen WW+2p Monte Carlo, a 20% uncertainty is assigned equal to the difference between estimates.

	0jets	1jet	≥ 2 jets	H_T+OS
ee	2.83 ± 0.80	0.66 ± 0.19	0.25 ± 0.07	0.15 ± 0.04
$e\mu$	5.76 ± 1.63	1.32 ± 0.37	0.50 ± 0.14	0.32 ± 0.09
$\mu\mu$	2.33 ± 0.66	0.58 ± 0.16	0.2 ± 0.06	0.13 ± 0.04
Total	10.92 ± 3.09	2.56 ± 0.72	0.95 ± 0.27	0.61 ± 0.17

Table 8.13: WW background estimate for 193 pb^{-1} , in different jet bins.

8.4.2 WZ/ZZ Production

A WZ event will most likely mimic a top dilepton signature when the Z boson decays leptonically, into two high momentum leptons and the W boson decays hadronically, creating 2 energetic jets. The branching ratio for such a decay channel is 4.6 %, given that the branching ratios $\text{BR}(Z \rightarrow \ell^+ \ell^-) = 6.8\%$, where $\ell = e$ or μ , and $\text{BR}(W \rightarrow q\bar{q}) = 68\%$, where $q\bar{q} = ud$ or cs .

The WZ background is estimated using a WZ Pythia sample of 1.39 million events (wtop0q). A dilepton filter is applied after the generation, demanding two leptons (e , μ or τ), with $P_T > 10 \text{ GeV}$ and $|\eta| < 2.0$. The filter has a $(5.8 \pm 0.01)\%$ efficiency, 80430 events passing the filter. Using the NLO cross-section mentioned in section 8.3, the sample integrated luminosity is 350.18 fb^{-1} . The cut efficiencies for the WZ Pythia Monte Carlo are shown in Table 8.14.

The ZZ background is determined using a Pythia MC sample, ztop4z (1364042 events). Both Zs are generated with the requirement that $M_Z > 30 \text{ GeV}$ and then freely decayed. A

post-generator filter requires at least two leptons (e/μ), with $P_T > 1$ GeV. Only 21.3% of events pass the filter. The sample integrated luminosity is 953.9 fb^{-1} . The cut efficiencies for ZZ are shown in Table 8.15.

Cut	Number of events	Efficiency (%)
$\epsilon_{geom.Pt}$	35365	44.65 ± 0.1766
ϵ_{ID}	21266	60.13 ± 0.2604
ϵ_{iso}	20751	97.58 ± 0.1054
ϵ_{Conv}	18509	89.2 ± 0.2155
ϵ_{mass}	10876	58.76 ± 0.3618
$\epsilon_{\cancel{E}_T}$	7471	68.69 ± 0.4447
$\epsilon_{\Delta\Phi}$	6752	90.38 ± 0.3412
ϵ_{2-jet}	600	8.886 ± 0.3463
ϵ_{H_T}	433	72.17 ± 1.83
ϵ_{OS}	335	77.37 ± 2.011
ϵ_{all}	335	0.423 ± 0.02306

Table 8.14: Efficiencies of the selection cuts applied to a Pythia WZ sample. Each efficiency is calculated from the number of events that have passed all the preceding cuts. The errors are statistical only.

Cut	Number of events	Efficiency (%)
$\epsilon_{geom.Pt}$	78161	26.9 ± 0.08227
ϵ_{ID}	41828	53.52 ± 0.1784
ϵ_{iso}	40518	96.87 ± 0.08516
ϵ_{Conv}	37977	93.73 ± 0.1204
ϵ_{mass}	18087	47.63 ± 0.2563
$\epsilon_{\cancel{E}_T}$	9014	49.84 ± 0.3718
$\epsilon_{\Delta\Phi}$	7989	88.63 ± 0.3344
ϵ_{2-jet}	645	8.074 ± 0.3048
ϵ_{H_T}	533	82.64 ± 1.492
ϵ_{OS}	470	88.18 ± 1.398
ϵ_{all}	470	0.1618 ± 0.007456

Table 8.15: Efficiencies of the selection cuts applied to a Pythia ZZ sample. Each efficiency is calculated from the number of events that have passed all the preceding cuts. The errors are statistical only.

Systematic Uncertainties

There are two main systematic uncertainties for the WZ background and they are similar to those for the WW:

1. The largest uncertainty is due to the jet energy scale; a 20% effect was seen on selection efficiencies when shifting jets up and down by 1σ in jet systematic uncertainty, and is taken as a systematic
2. By comparing the background estimate using Pythia MC with the estimate from Alpgen WW+2p, a 20% uncertainty is assigned equal to the difference between estimates.

For ZZ a 30 % systematic uncertainty, similar with the one on WZ, is assigned.

The background estimates in 193 pb^{-1} are shown in Tables 8.16 and 8.17. As it can be seen, the ZZ is negligible after 2 jet requirement, but it is important in 0 jet bin.

	0jets	1jet	≥ 2 jets	H_T+OS
ee	0.40 ± 0.11	0.46 ± 0.13	0.10 ± 0.03	0.057 ± 0.02
$e\mu$	0.36 ± 0.10	0.26 ± 0.07	0.08 ± 0.02	0.027 ± 0.01
$\mu\mu$	0.45 ± 0.13	0.26 ± 0.07	0.06 ± 0.02	0.045 ± 0.01
Total	1.23 ± 0.3480	0.98 ± 0.28	0.24 ± 0.07	0.13 ± 0.04

Table 8.16: WZ background estimate for 193 pb^{-1} , in different jet bins.

	0jets	1jet	≥ 2 jets	H_T+OS
ee	0.44 ± 0.13	0.12 ± 0.04	0.039 ± 0.013	0.03 ± 0.01
$e\mu$	0.02 ± 0.006	0.05 ± 0.02	0.024 ± 0.008	0.01 ± 0.003
$\mu\mu$	0.39 ± 0.12	0.11 ± 0.03	0.038 ± 0.013	0.03 ± 0.01
Total	0.85 ± 0.26	0.28 ± 0.09	0.101 ± 0.03	0.073 ± 0.021

Table 8.17: ZZ background estimate for 193 pb^{-1} , in different jet bins.

8.5 Other small sources of background

8.5.1 b quark pair production

$b\bar{b}$

Because a second lepton is allowed to be non-isolated, the real leptons from heavy flavor decay contribution to the fake estimate might be underestimated if the fake rate for a b-jet is much higher than for a generic jet. A upper limit of the contribution of $b\bar{b}$ to the fake background yields a 12.3 % increase, which is well within our uncertainty on the fake background.

$Wb\bar{b}$

A $Wb\bar{b}$ event, with one b-jet faking a lepton could mimic a top dilepton event, if the event has an extra jet from QCD radiation. This contribution is included in the fake estimate, except that it might be underestimated if the fake rate for a b-jet is much higher than for a generic jet. Monte Carlo samples were use to determine the order of magnitude for this background, and found to be negligible.

The other possible backgrounds, such as WH or radiative Z^0 bosons are negligible.

8.6 Summary of backgrounds

The background estimates from the previous sections are summarized in Table 8.19. Also the systematic uncertainties assigned to the various backgrounds are shown in Table 8.18.

	Source	Uncertainty (%)
Signal Acceptance		
$t\bar{t}$	Lepton Id SF	10
	Jet Energy Scale	4.7
	ISR/FSR	1.7
	PDFs	6
	Different Generators (Pythia vs Herwig)	5.5
Backgrounds		
Fakes	Method	31
	Jet Samples Composition	9
	H_T	14
	Real leptons in Jet Samples	20
DY($ee, \mu\mu$)	Method	100
	Jet Energy Scale	20
WW/WZ	MC Generator	20
	Jet Energy Scale	36
$Z \rightarrow \tau\tau$	2 jet scale factor	10
	Jet Energy Scale	29

Table 8.18: Grand Table of systematic uncertainties.

	0-jet	1-jet	2-jet	H_T+OS
$e\mu$				
WW/WZ	6.12 ± 1.73	1.58 ± 0.44	0.58 ± 0.16	0.35 ± 0.10
$Z \rightarrow \tau\tau$	0.11 ± 0.03	0.45 ± 0.14	0.36 ± 0.11	0.22 ± 0.07
Fakes	3.69 ± 1.51	2.9 ± 1.19	1.64 ± 0.67	0.69 ± 0.28
ZZ	0.020 ± 0.006	0.05 ± 0.02	0.038 ± 0.013	0.01 ± 0.003
Total Background	9.94 ± 2.30	4.98 ± 1.28	2.62 ± 0.70	1.27 ± 0.31
$t\bar{t}$ 175	0.05 ± 0.01	0.71 ± 0.14	4.7 ± 0.658	4.5 ± 0.6
CDF Run II Data	6	3	10	9
ee				
WW/WZ	3.23 ± 0.91	1.12 ± 0.32	0.35 ± 0.1	0.21 ± 0.06
$Z \rightarrow \tau\tau$	0.05 ± 0.01	0.22 ± 0.07	0.15 ± 0.05	0.09 ± 0.03
Fakes	1.41 ± 0.58	1.11 ± 0.46	0.62 ± 0.25	0.26 ± 0.11
DY	4.2 ± 1.9	2.2 ± 0.9	0.57 ± 0.27	0.36 ± 0.27
ZZ	0.44 ± 0.13	0.12 ± 0.04	0.039 ± 0.013	0.03 ± 0.01
Total Background	9.33 ± 2.19	4.77 ± 1.06	1.73 ± 0.39	0.95 ± 0.30
$t\bar{t}$ 175	0.04 ± 0.01	0.38 ± 0.05	2.00 ± 0.28	1.9 ± 0.3
CDF Run II Data	9	6	1	1
$\mu\mu$				
WW/WZ	2.78 ± 0.79	0.84 ± 0.23	0.26 ± 0.08	0.18 ± 0.05
$Z \rightarrow \tau\tau$	0.04 ± 0.01	0.19 ± 0.06	0.17 ± 0.04	0.11 ± 0.03
Fakes	0.85 ± 0.35	0.67 ± 0.27	0.38 ± 0.16	0.16 ± 0.07
DY	0.2 ± 0.7	0.1 ± 0.6	0.11 ± 0.34	0.07 ± 0.34
ZZ	0.39 ± 0.12	0.11 ± 0.03	0.038 ± 0.013	0.03 ± 0.01
Total Background	4.26 ± 1.12	1.91 ± 0.70	0.96 ± 0.39	0.55 ± 0.35
$t\bar{t}$ 175	0.01 ± 0.001	0.30 ± 0.04	1.95 ± 0.27	1.8 ± 0.3
CDF Run II Data	4	2	3	3
$ee+e\mu+\mu\mu$				
WW/WZ	12.15 ± 3.44	3.54 ± 1.00	1.19 ± 0.34	0.74 ± 0.21
$Z \rightarrow \tau\tau$	0.19 ± 0.06	0.86 ± 0.26	0.69 ± 0.21	0.42 ± 0.13
Fakes	5.95 ± 2.44	4.68 ± 1.92	2.64 ± 1.92	1.1 ± 0.5
DY	4.4 ± 2.0	2.2 ± 1.1	0.69 ± 0.44	0.43 ± 0.43
ZZ	0.85 ± 0.26	0.28 ± 0.09	0.10 ± 0.03	0.073 ± 0.021
Total Background	23.54 ± 4.68	11.56 ± 2.44	5.31 ± 2.01	2.8 ± 0.7
$t\bar{t}$ 175	0.10 ± 0.01	1.39 ± 0.19	8.68 ± 1.21	8.2 ± 1.1
CDF Run II Data	19	11	14	13

Table 8.19: Summary of the background contributions to the $t\bar{t}$ dilepton channels. Also the estimates are shown before the 2-jet cut, in jet multiplicity bins.

9 Unveiling the detector data

In this chapter we *open up the box*, by looking at the CDF data. Also we compare few kinematical distributions with the Standard Model predictions. The $t\bar{t}$ cross-section is calculated and compared with the theoretical predictions.

9.1 The top dilepton candidates

Starting with the 193 pb^{-1} dataset¹, described in Chapter 5, and applying the selection criteria, presented in Chapter 6, a number of *13 events* survive all the requirements. The number of events passing consecutive cuts, for each dilepton channel and categories, are summarized in Tables 9.1-9.4.

¹In fact, the luminosities for various categories range from 150 to 193 pb^{-1} , as seen in Table 6.6, with the largest fraction of events with a 193 pb^{-1} of luminosity. The weighted luminosity is 185 pb^{-1} , as calculated in Chapter 7.

Category	Dilepton Cut					NJets			H_T	OS
	ID	Iso	Z_{veto}	\cancel{E}_T	$\Delta\phi$	0j	1j	2j		
TCE/TCE	4090	4024	3277	17	8	5	2	1	1	1
TCE/NTCE			167	6	2	1	1	0	0	0
TCE/PHX	3731	3640	3145	9	6	3	3	0	0	0
NTCE/IPHX			92	1	0	0	0	0	0	0
ee	7821	7664	6681	33	16	9	6	1	1	1

Table 9.1: Breakdown of the ee top dilepton channel analysis of the 193 pb⁻¹ of data.

Category	Dilepton Cut					NJets			H_T	OS
	ID	Iso	Z_{veto}	\cancel{E}_T	$\Delta\phi$	0j	1j	2j		
CMUP/CMUP	1486	1476	870	5	2	1	1	0	0	0
CMUP/NCMUP			31	0	0	0	0	0	0	0
CMUP/CMU	477	471	342	1	0	0	0	0	0	0
CMUP/NCMU			6	0	0	0	0	0	0	0
CMUP/CMP	745	715	451	3	3	1	1	1	1	1
CMUP/NCMP			13	0	0	0	0	0	0	0
CMUP/CMX	995	992	836	3	2	1	0	1	1	1
CMUP/NCMX			11	0	0	0	0	0	0	0
NCMUP/CMX			20	2	0	0	0	0	0	0
CMUP/CMIO	954	910	757	1	0	0	0	0	0	0
CMX/CMX	331	328	250	1	1	0	0	1	1	1
CMX/NCMX			14	0	0	0	0	0	0	0
CMX/CMU	277	269	226	0	0	0	0	0	0	0
CMX/NCMU			2	0	0	0	0	0	0	0
CMX/CMP	262	256	212	0	0	0	0	0	0	0
CMX/NCMP			11	0	0	0	0	0	0	0
CMX/CMIO	436	410	350	3	1	1	0	0	0	0
$\mu\mu$	5963	5827	4402	19	9	4	2	3	3	3

Table 9.2: Breakdown of the $\mu\mu$ top dilepton channel analysis of the 193 pb⁻¹ of data.

Category	Dilepton Cut					NJets			H_T	OS
	ID	Iso	Z_{veto}	\cancel{E}_T	$\Delta\phi$	0j	1j	2j		
TCE/CMUP	46	31	18	4	3	1	0	2	1	1
TCE/NCMUP			5	0	0	0	0	0	0	0
NTCE/CMUP			1	1	1	0	0	1	1	1
TCE/CMU	8	7	5	1	1	0	0	1	1	1
TCE/NCMU			1	0	0	0	0	0	0	0
TCE/CMP	20	9	4	2	2	1	0	1	1	1
TCE/NCMP			1	0	0	0	0	0	0	0
TCE/CMX	27	21	11	4	4	1	0	3	3	3
TCE/NCMX			2	0	0	0	0	0	0	0
NTCE/CMX			1	0	0	0	0	0	0	0
TCE/CMIO	70	24	20	3	3	2	0	1	1	1
PHX/CMUP	14	10	10	3	3	0	2	1	1	1
PHX/NCMUP			0	0	0	0	0	0	0	0
PHX/CMU	1	1	1	0	0	0	0	0	0	0
PHX/NCMU			0	0	0	0	0	0	0	0
PHX/CMP	6	5	3	1	1	1	0	0	0	0
PHX/NCMP			2	0	0	0	0	0	0	0
PHX/CMX	10	8	6	2	1	0	1	0	0	0
PHX/NCMX			1	0	0	0	0	0	0	0
PHX/CMIO	21	5	5	0	0	0	0	0	0	0
$e\mu$	223	121	121	21	19	6	3	10	9	9

Table 9.3: Breakdown of the $e\mu$ top dilepton channel analysis of the 193 pb^{-1} of data.

Category	Dilepton Cut					NJets			H_T	OS
	ID	Iso	Z_{veto}	\cancel{E}_T	$\Delta\phi$	0j	1j	2j		
3ℓ	8	7	0	0	0	0	0	0	0	0
Total 14243	14014	13618	11203	73	44	19	11	14	13	13

Table 9.4: Multilepton and summary of the analysis of the 193 pb^{-1} of data.

A list of run and event numbers for the 13 candidate events is given in Table 9.5.

Event category	Run/Event number	Trigger leptons
<i>ee</i>		
TCE/TCE	153374/ 2276742	CEM_18
<i>eμ</i>		
CMUP/NITCE	151978/507773	CMUP_18 && CEM_18
TCE/CMUP	167629/180103	CMUP_18 && CEM_18
TCE/CMP	143257/760520	CEM_18
TCE/CMU	165364/592961	CEM_18
TCE/CMX	155114/478702	CMX_18 && CEM_18
TCE/CMX	156484/3099305	CMX_18 && CEM_18
TCE/CMX	167631/2058969	CMX_18 && CEM_18
TCE/CMIO	161633/963604	CEM_18
PHX/CMUP	163064/10576918	CMUP_18 && MET_PEM
<i>μμ</i>		
CMUP/CMP	162820/7050764	CMUP_18
CMUP/CMX	165198/1827962	CMUP_18 && CMX_18
CMX/CMX	153325/599511	CMX_18

Table 9.5: Run and event numbers for the candidate events. Also shown are the triggers the event was collected on.

Kinematical Distributions: Data versus Monte Carlo

In this paragraph a few kinematical distributions are compared with the Standard Model expectations, which consist of the SM expected backgrounds and a $t\bar{t}$ Pythia Monte Carlo signal sample, with a 175 GeV top mass. All the distributions have the backgrounds normalized to the expected contributions from Chapter 8, while for the $t\bar{t}$ contribution the measured cross-section value is used, unless explicitly stated otherwise.

Figure 9.1 shows the lepton p_T distribution. It is intriguing to see that out of the 26 leptons, 19 have a momentum less than 30 GeV. The probability that the p_T distribution of the candidates is consistent with the SM null hypothesis is about 2%([13]).

Two more distributions regarding the leptons are shown in Figure 9.2. The angle between the leptons in transverse plane, $\Delta\Phi_u$, for the dilepton candidates is compared with the $t\bar{t}$ Monte Carlo. The signal distribution is normalized to 10 times the expectation. Second distribution shows the di-lepton invariant mass for the 13 candidates. The dielectron event has an invariant mass of 59.7 GeV, while the dimuon invariant masses 40.0 GeV, 61.2 GeV, and 58.8 GeV respectively. Both data histograms agree well with the expectations.

In Run I the CDF top dilepton sample had a slight excess of large \cancel{E}_T events, as 4 events out of 10 candidates were with \cancel{E}_T above 100 GeV[66]. Figure 9.4 shows the Run II missing transverse distribution, in good agreement with the Standard Model expectation. Only 2 events have \cancel{E}_T larger than 100 GeV. In the same figure, lower plot, the data events passing all the selection requirements, but the L cut², are shown, together with the $t\bar{t}$ Monte Carlo. The lower left region, with low \cancel{E}_T and small $\Delta\phi(\cancel{E}_T, l/j)$ is populated by DY background events.

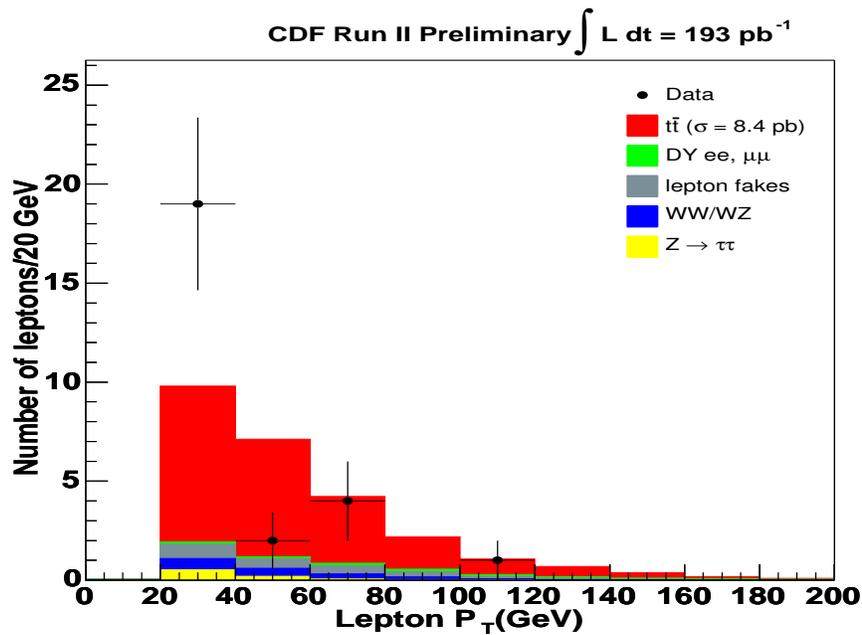


Figure 9.1: Lepton momenta distribution of the 13 candidates, the expected background and $t\bar{t}$ Pythia Monte Carlo signal for a 175 GeV top mass are shown.

²The L cut requires that $\Delta\Phi(\cancel{E}_T, l/j) > 20^\circ$, if $\cancel{E}_T < 50 \text{ GeV}$.

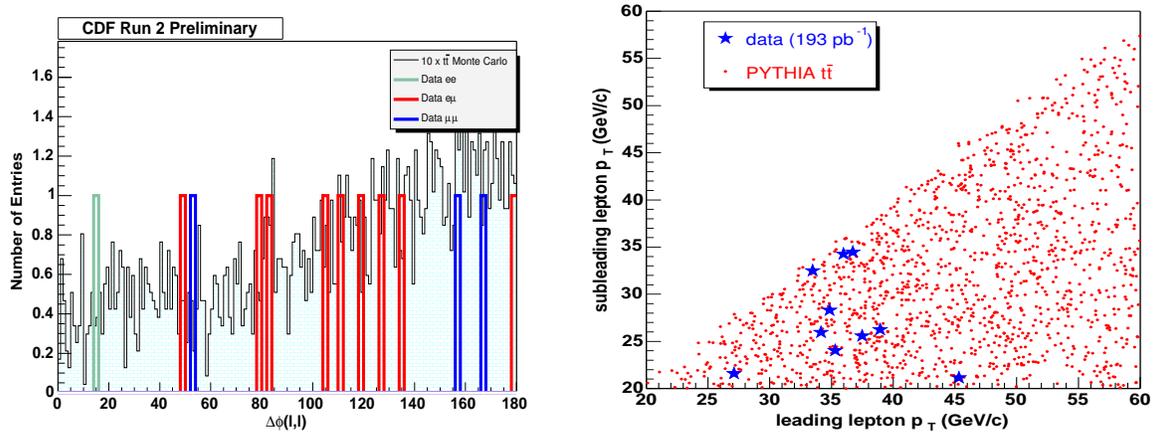


Figure 9.2: The angle, $\Delta\Phi_{ll}$, between the leptons in transverse plane.

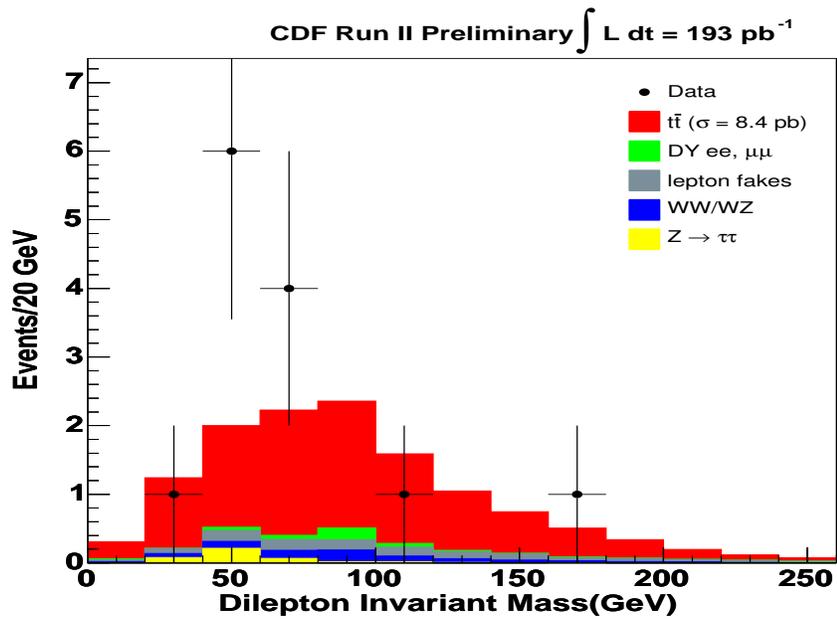


Figure 9.3: The dilepton invariant mass distribution, for the 13 candidates.

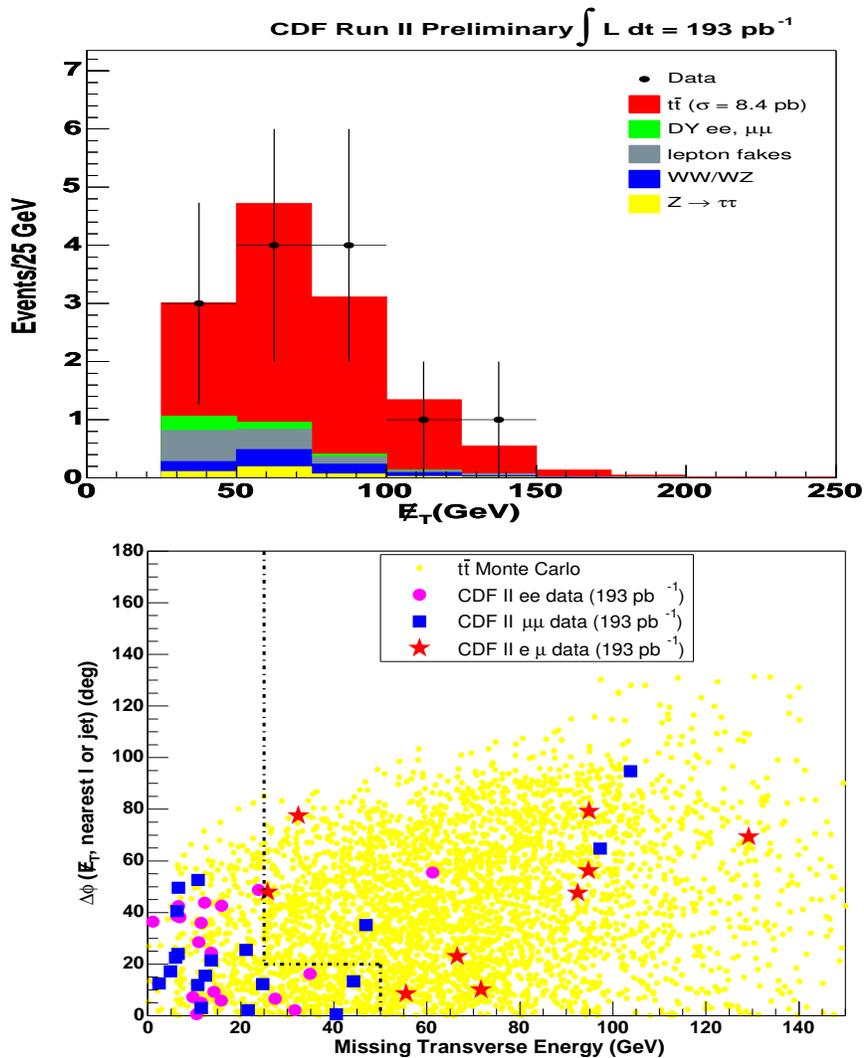


Figure 9.4: Missing transverse energy (upper plot) for the data compared with the Standard Model expectation. The azimuthal angle between the \cancel{E}_T and the nearest lepton or jet versus the \cancel{E}_T (lower plot), for the 13 candidates and $t\bar{t}$ Pythia Monte Carlo with $m_t = 175 \text{ GeV}$. The dashed line shows the L cut.

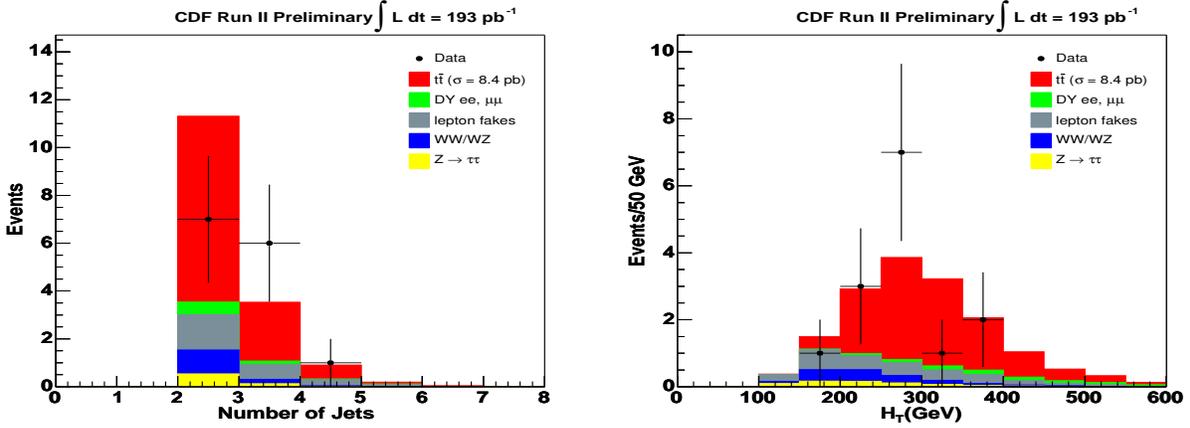


Figure 9.5: The left plot is the jet multiplicity distribution for the 13 candidates, compared with SM expectation, while the right plot shows the scalar transverse energy sum, H_T (GeV).

9.2 b-tags in the dilepton candidates

The identification of the jet as coming from a b quark, after fragmentation and hadronization, is known as *b tagging*. This analysis does not make any b tagging requirements of the jets in the event, as there are better ways to separate the top dilepton signal from backgrounds³. However each top dilepton event has 2 b jets, coming from the top decaying to a bottom quark and a W boson. Moreover the dilepton backgrounds, such as WW, WZ, DY or $Z \rightarrow \tau\tau$ have mainly gluon jets. Therefore there is a strong hint that the top dilepton candidates with b-tags are probably from $t\bar{t}$ decays. There are three major b-tagging algorithms used in CDF to identify a b jet: SECVTX (SECondary VerTeX), SLT(Secondary Lepton Tagging), and Jet Probability.

The SECVTX efficiency to tag at least a b jet per dilepton event is determined from the expression [44]:

$$\epsilon_{btag}^{evt} = F_{1b} \cdot \epsilon_{btag} \cdot S + F_{2b} \cdot 2 \cdot \epsilon_{btag} \cdot S \cdot (1 - \epsilon_{btag} \cdot S) + F_{2b} \cdot \epsilon_{btag}^2 \cdot S^2 \quad (9.1)$$

³The use of b tagging in the dilepton event selection would allow to achieve a S/B ~ 10 , however with a large loss in acceptance of about 50 %

where S is the data to Monte Carlo b-tagging scale factor, F_{1b} , F_{2b} are the fraction of events with 1, respectively 2 taggable b-jets, and ϵ_{btag} is the b-tagging efficiency per jet, determined from Monte Carlo. Using the values from Table 9.6, the following efficiencies, characteristic to a dilepton event, are calculated:

- probability to tag only one b jet is

$$\epsilon_{1-btag}^{evt} = 0.44 \pm 0.13 \quad (9.2)$$

- probability to tag two b jets is

$$\epsilon_{2-btag}^{evt} = 0.12 \pm 0.04 \quad (9.3)$$

- probability to tag at least one b-jet in a $t\bar{t}$ dilepton event is

$$\epsilon_{btag}^{evt} = 0.56 \pm 0.17 \quad (9.4)$$

Variable	Value
F_{1b}	0.543 ± 0.008
F_{2b}	0.364 ± 0.009
S	0.86 ± 0.07
ϵ_{btag}	0.489 ± 0.013

Table 9.6: The variables used to determine the SECVTX b-tagging efficiencies.

Table 9.7 lists the dilepton candidates with the total number of jets and the number of b-tags.

Event category	Run/Event number	# jets	# btags
TCE/TCE	153374/2276742	3	0
CMUP/NITCE	151978/507773	2	1
TCE/CMUP	167629/180103	2	0
TCE/CMP	143257/760520	4	bad SVX run
TCE/CMU	165364/592961	3	1
TCE/CMX	155114/478702	3	1
TCE/CMX	156484/3099305	3	0
TCE/CMX	167631/2058969	2	2
TCE/CMIO	161633/963604	3	0
PHX/CMUP	163064/10576918	2	0
CMUP/CMP	162820/7050764	2	2
CMUP/CMX	165198/1827962	2	2
CMX/CMX	153325/599511	3	1

Table 9.7: The b tagging information for the top dilepton candidates. The event 507773 has the b-tagged jet away from the NICEM, while the event 592961 has the lowest E_T jet b-tagged.

The observed number of b-tagged jets is shown in Table 9.8 and it is in good agreement with the expectations.

B-tags	Observed	Expected
Single Tagged Events	4	4.6 ± 1.3
Double Tagged Events	3	1.3 ± 0.5
Total Tagged Events	7	5.9 ± 1.8

Table 9.8: The number of observed and expected SECVTX single and double-tagged jets.

9.3 $t\bar{t}$ cross section estimation

In this section all the pieces needed to estimate the $t\bar{t}$ dilepton cross-section are combined. The cross-section $\sigma(p\bar{p} \rightarrow t\bar{t} + X)$, in the dilepton channel, is calculated using the following expression:

$$\sigma_{t\bar{t}} = \frac{N^{obs} - N^{bak}}{\epsilon_{dil} \cdot \mathcal{L}} \quad (9.5)$$

where the expression symbols represent

- N^{obs} is the number of observed events, passing the selection cuts,
- N^{bak} is the expected number of Standard Model background events,
- ϵ_{dil} is the top dilepton acceptance, described in Chapter 7, and
- \mathcal{L} is the total integrated luminosity of the dataset, $\mathcal{L} = \int L dt$.

Component	Value	Reference
N^{obs}	13 events	Chapter 9.1
N^{bak}	2.7 ± 0.7 events	Chapter 8
ϵ_{dil}	(0.66 ± 0.09) %	Chapter 7
\mathcal{L}	185	Chapter 5

Table 9.9: The values used to calculate the top dilepton cross-section. The weighted values for the acceptance and luminosity, as defined in Chapter 7, are used.

The $t\bar{t}$ production cross-section can be calculated by maximizing the following likelihood expression[66],[83]:

$$L = P(N^{obs}, \sigma_{t\bar{t}} \cdot \epsilon_{dil} \cdot \mathcal{L} + N^{bak}) \cdot G(L, \mathcal{L}, \sigma_{\mathcal{L}}) \cdot G(b, N^{bak}, \sigma_{N^{bak}}) \cdot G(\epsilon, \epsilon_{dil}, \sigma_{\epsilon_{dil}}) \quad (9.6)$$

where the uncertainties are calculated as the cross-section variation around the extremum, such that the change in the likelihood is $\Delta(\ln L) = \pm \frac{1}{2}$. $P(N^{obs}, \sigma_{t\bar{t}} \cdot \epsilon_{dil} \cdot \mathcal{L} + N^{bak})$ is the Poisson probability of observing N^{obs} events, given the expected number is $(\sigma_{t\bar{t}} \cdot \epsilon_{dil} \cdot \mathcal{L} + N^{bak})$, for a $t\bar{t}$ cross-section $\sigma_{t\bar{t}}$. \mathcal{L} , N^{bak} , ϵ_{dil} are allowed to fluctuate within their uncertainties. $G(p, P, \sigma_p)$ is a Gaussian for parameter p, with mean P and width σ_p . The cross-section is:

$$\sigma_{t\bar{t}} = 8.4_{-2.7}^{+3.2}(stat)_{-1.0}^{+1.8}(syst) \pm 0.5(lum) \text{ pb} \quad (9.7)$$

A alternative way of estimating the systematic uncertainties would be to draw pseudo-trials from Gaussian distributions, for the number of background events, $G(b, N^{bak}, \sigma_{N^{bak}})$ and the top dilepton acceptance, $G(\epsilon, \epsilon_{dil}, \sigma_{\epsilon_{dil}})$. The uncertainty due to the luminosity is

preferred to be separated. For every pseudotrial we recalculate the cross-section, using the expression 9.5. The resulting distribution of the cross-sections is shown in Figure 9.6[20]. The median is used as the central value (50% percentile point) and the 84.14 % and 15.86 % percentile as +1 sigma and -1 sigma respectively. The cross-section calculated this way is

$$\sigma_{t\bar{t}} = 8.4_{-2.7}^{+3.2}(stat)_{-1.1}^{+1.5}(syst) \pm 0.5(lum) \text{ pb} \quad (9.8)$$

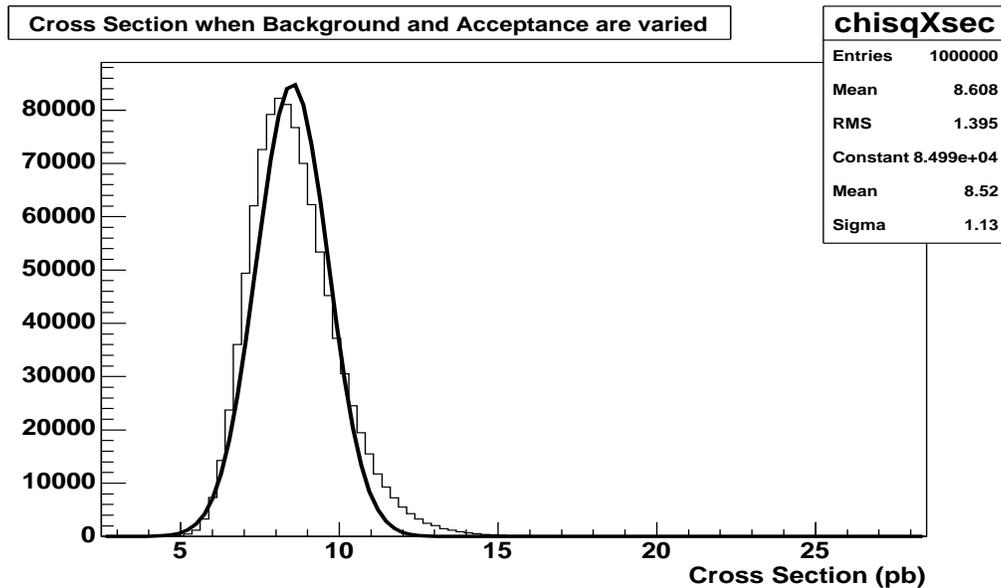


Figure 9.6: Distribution of the cross-section when varying the background and the acceptance with Gaussian constraints.

These two methods give consistent answers and as the measured cross-section we will use the second value, shown in 9.8.

10 Conclusions and Future Prospects

10.1 Conclusions

Using 193 pb^{-1} of Run II data, recorded by the upgraded CDF detector, we selected events produced via the $t\bar{t}$ dilepton decay channel ($t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow \ell^+\nu_\ell b\ell^-\bar{\nu}_\ell\bar{b}$). We observed 13 candidate events with an expected background of 2.8 ± 0.7 events. Out of the 13 candidates, one is in the ee channel, 3 in the $\mu\mu$ channel and 9 in the $e\mu$ channel. With the above values we measured the $t\bar{t}$ production cross-section in the dilepton channel to be:

$$\sigma_{t\bar{t}} = 8.4_{-2.7}^{+3.2}(\text{stat})_{-1.1}^{+1.5}(\text{syst}) \pm 0.5(\text{lum}) \text{ pb} \quad (10.1)$$

This measurement is in good agreement with the Standard Model prediction of $6.70_{-0.88}^{+0.71}$ pb. Also we compared several kinematical distributions with the Standard Model predictions and find good agreement. However the transverse momentum of the leptons is softer in the data compared to the SM expectation. This it is probably a fluctuation, as a complementary measurement, performed with a larger number of top dilepton events, sees good agreement[4]. Using the b-tagging information, we find that 7 of the 13 candidates have at least one jet tagged as coming from a b quark, while 5.9 ± 1.8 events are expected.

10.2 Future Prospects

The $t\bar{t}$ production cross-section measurement in dilepton channel is dominated by the statistical uncertainty, about twice larger than the systematic part. With about 2 fb^{-1} (expected in a few years), the statistical error will be reduced by only 30 % if the same selection is used. Therefore the key to improve the precision of the measurement is to significantly increase acceptance, while keeping the background under control.

The present dilepton sample is a *starting point* for other top analyses, such as the top mass measurement in the dilepton channel or measurement of the W helicity using the transverse momentum spectrum of the lepton. Top is still one of the least well-studied components of the SM. Therefore it is important to use the larger and larger Run II top samples¹, becoming available to study its properties (in particular its mass, the production cross-section and branching ratios), as well as look for new physics which alters the top production or decay mechanisms. There are many models of new physics which modify top's expected SM behavior. Also at the Large Hadron Collider, top will be one of the main background for many new physics searches, so as many of its properties as possible should be precisely measured at the Tevatron.

With 2 fb^{-1} , by combining the $t\bar{t}$ production cross-section measurements from three main channels (dilepton, lepton+jets and all-hadronic) the uncertainty of the measurement might be reduced to about 7-10 %. Table 10.1 shows the precision one expects to achieve with 2 fb^{-1} of Run II data.

Measurement	Estimated Uncertainty	Importance
$\delta\sigma_{t\bar{t}}$	7 %	QCD Couplings
$\delta[\sigma_{\ell\ell}/\sigma_{t\bar{t}}]$	12 %	Non-SM Decays
$\delta[\sigma \times BR(Z' \rightarrow t\bar{t})]$	90 fb	New Physics
$\delta\sigma_{t\bar{b}X+b\bar{t}X}$	24 %	Single Top Observation
$\delta\Gamma(t \rightarrow Wb)$	26 %	Top Width

Table 10.1: Few Top Physics Prospects for 2 fb^{-1}

¹The recorded Run I luminosity was 106 pb^{-1} , so we have already almost twice more data; also many cross-sections increased with the center of mass energy, for example in going from Run I energy of 1.8 TeV to Run II energy of 1.96 TeV, $\sigma_{t\bar{t}}$ increase by 30%

The present weak constraints on the parameters of the top quark leave considerable room for deviations from SM expectations. In the next few years, the improvement of our knowledge of the top quark may possibly help reveal the true mechanism of electroweak symmetry breaking and mass generation, and also help to understand the physics beyond the Standard Model.

Bibliography

- [1] , “Fermilab Accelerator Web Page,” <http://www.fnal.gov/pub/inquiring/physics/accelerators/chainacce>
2004. 181
- [2] , “Tevatron Run II Handbook,” <http://www-bd.fnal.gov/pub/inquiring/physics/accelerators/chainaccel.html>, 2004. 181
- [3] A. Bhatti, M.D.Shapiro, J. Lamoureux, “DSTOWE,” *Internal CDF Note 3952*, 2000.
62
- [4] D. Acosta *et al.*, CDF Collaboration, “Measurement of $t\bar{t}$ Production Cross Section in $p\bar{p}$ Collisions at 1.96 TeV using Dilepton Events,” *Internal CDF Note 6882*, 2004.
127, 169
- [5] Eric Laenen, “Top Quark production dynamics,” *Talk presented at XXXVIIIth Rencontres de Moriond, QCD and Hadronic Interactions*, 2003. 23
- [6] M. Barnett, L. Hall, “Multilepton Signal For Supersymmetric Particles in the Fermilab Tevatron Data ?,” *Phys. Rev. Lett.* 77, 16, 1996. 9
- [7] N. Kidonakis, E. Laenen, S. Moch, R. Vogt, “Sudakov Resummation and Finite Order Expansions of Heavy Quark Hadroproduction Cross Sections,” *Phys. Rev. D* 64, 114001, *hep-ph/0105041*, 2001. 23
- [8] R. Bonciani, S. Catani, M.L. Mangano, P. Nason, “NLL resummation of the heavy-quark hadroproduction cross-section,” *Nucl. Phys. B* 529,424, *hep-ph/9801375*, 1998. 23
- [9] R. P. Feynman, “Photon Hadron Interactions,” *W. A. Benjamin, Reading, Mass*, 1972.

- [10] Stephen Miller, “PDF uncertainties in Top & QCD,” *Talk presented at Joint Physics Group Meeting on March 19, 2004*, 2004. 127
- [11] W. Beenakker, W.L. van Neerven, R. Meng, G.A. Schuler, J. Smith, “QCD corrections to heavy quark production in hadron hadron collisions,” *Nucl. Phys. B351*, 507, 1991. 22
- [12] A.D. Martin, R.G. Roberts, W.J. Stirling, R.S. Thorne, “Uncertainties of predictions from parton distributions. I: experimental errors,” *hep-ph/0211080*, 2002. 146
- [13] Andrew Ivanov; Andy Hocker; Paul Tipton, “Search for Anomalous Kinematics in ttbar Dilepton Events at CDF,” *Internal CDF Note 7114*, 2004. 160
- [14] Beate Heinemann Angela Wyatt, “Correction for Leakage Energy in the Central and Plug Calorimeters in Run II,” *Internal CDF Note 6167*, 2002. 95, 104
- [15] B. W. Harris *et al*, “The Fully Differential Single Top Quark Cross-Section in Next-To-Leading Order QCD.,” *Phys. Rev. D*, 66:054024, 2002. 15
- [16] B.R.Ko, S.H.Oh, C.Wang, “SVXII Standalone Tracking,” *Internal CDF Note 6440*, 2003. 65
- [17] C. Ciobanu *et al*., “Updated Search for Single Top Production using 162 pb^{-1} ,” *Internal CDF Note 6795*, 2004. 16
- [18] The CDF Collaboration C. Ginsburg, “CDF Run 2 Muon System,” *Internal CDF Note 6741*, 2003. 48
- [19] Campagnari C., Franklin M., “The discovery of the top quark,” *Rev. Mod. Phys.* 69:137, 1997. 12
- [20] Charles Plager, “Thanks to Charles Plager for kindly providing me the plot.,” ., 2004. 168
- [21] Chris Hays, Peter Tamburello, Ashutosh Kotwal, Peter Wittich, Rick Snider, “The COT Pattern Recognition Algorithm and Offline Code,” *Internal CDF Note 6992*, 2004. 66

- [22] Chris Hill, Corrinne Mills, David Goldstein, Mircea Coca, “Fake Lepton Backgrounds for the Summer 2003 Top Dilepton Cross Section,” *Internal CDF Note 6592*, 2003. 134
- [23] Cigdem Issever *et al.*, “Plug Electron Baseline Cuts as defined in Summer 2003 and their Efficiencies,” *Internal CDF Note 6789*, 2004. 93
- [24] S. Drell and T.-M. Yan., “Massive Lepton-Pair Production in Hadron-Hadron Collisions at High Energies.,” *Phys. Rev. Lett.*, 25:316, 1970. 138
- [25] E. Thomson *et al.*, “,” *IEEE Trans. Nucl. Sci.*, 49(1063), 2002. 81
- [26] Andy Hocker *et al.*, “Determination of Drell-Yan Backgrounds for the Run II Top Dilepton Cross Section, Summer 2003,” *Internal CDF Note 6591*, 2003. 140
- [27] Cigdem Issever *et al.*, “ETF Web Page,” <http://www-cdf.fnal.gov/internal/physics/top/r2leptons/etf/etfnewhome/BaseLineCuts/baselinecutsall.html>, 2003. 89
- [28] D. Glenzinski *et al.*, “Drell-Yan Rejection and Estimation for the Tight Lepton and Isolated Track Sample,” *Internal CDF Note 6582*, 2004. 140
- [29] D. Goldstein *et al.*, “A measurement of the $t\bar{t}$ cross-section using dileptons in the central and endplug detectors,” *Internal CDF Note 6588*, 2003. 94, 139
- [30] F. Abachi *et al.*, “Observation of the Top Quark.,” *Phys. Rev. Lett.*, 74:2632, 1995. 12
- [31] F. Abe *et al.*, “Evidence for the top quark production in $p\bar{p}$ collisions at $\sqrt{s}=1.8$ TeV.,” *Phys. Rev. D*, 50:2966, 1994. 12
- [32] F. Abe *et al.*, “Observation of Top Quark Production in $p\bar{p}$ collisions with the Collider Detector at Fermilab.,” *Phys. Rev. Lett.*, 74:2626, 1995. 12
- [33] Gerald C. Blazey *et al.*, “Run II Jet Physics,” *Internal CDF Note 5293*, 2000. 75
- [34] M. Coca *et al.*, “Winter 2003 Measurement of the $t\bar{t}$ Cross Section in the Dilepton Decay Channel,” *Internal CDF Note 6319*, 2003.

- [35] M. L. Perl *et al.*, “Evidence for Anomalous Lepton Production in e^+e^- Annihilation,” *Phys. Rev. Lett.* **35**, 1489, 1975. 12
- [36] M. Martinez *et al.*, “DQM Web Page,” <http://www-cdf.fnal.gov/internal/dqm/goodrun/v4/goodv4.html>, 2004. 78
- [37] Mircea Coca *et al.*, “A Second Determination of the Fake Lepton Background for the Summer 2004 Top Dilepton Analysis,” *Internal CDF Note 6742*, 2004. 131
- [38] P. Azzi-Bacchetta *et al.*, “Proposal for a Very Low Mass, Very Small Radius Silicon Layer In the CDF II Upgrade,” *Internal CDF Note 4924*, 1999. 43
- [39] R. Blair *et al.*, “THE CDF-II DETECTOR: Technical Design Report,” *FERMILAB-Pub-96/390-E*, 1996. 26, 44
- [40] S. Herb *et al.*, “Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions,” *Phys. Rev. Lett.*, **39**:252, 1977. 12
- [41] W. Innes *et al.*, “Observation of a Structure in the Υ Region,” *Phys. Rev. Lett.*, **39**:1240, 1977. 12
- [42] D. Acosta *et al.*, “The CDF Cerenkov Luminosity Monitor,” *Internal CDF Note 5559*, 2001. 49
- [43] D. Acosta *et al.*, “A First Look at the CLC Luminosity Measurements,” *Internal CDF Note 6052*, 2002. 78
- [44] D. Amidei *et al.*, “Measurement of the $t\bar{t}$ Production Cross Section in the SECVTX Tagged Sample,” *Internal CDF Note 6585*, 2004. 164
- [45] J-F Arguin *et al.*, “Generic Jet Energy Corrections for Run II data used for the Winter Conferences,” *Internal CDF Note 6280*, 2003. 111
- [46] K. Bloom *et al.*, “High p_T Lepton ID Efficiency Scale Factor Studies in 4.11.1,” *Internal CDF Note 6858*, 2004. 125

- [47] M.Coca *et al.*, “A Measurement of the $t\bar{t}$ cross-section using dileptons,” *Internal CDF Note 6830*, 2003. 95, 96
- [48] Henry Frisch, “Separating top and Z’s in dilepton events inside the Z mass window,” *Internal CDF Note 3387*, 1995. 106
- [49] G. Arnison *et al.*, “Experimental Observation of Lepton Pairs of Invariant Mass Around 95 GeV/c² at the CERN SPS Collider,” *Phys. Lett. B 126*, 398, 1983. 9
- [50] G. Arnison *et al.*, “Further Evidence for Charged Intermediate Vector Bosons at the SPS Collider,” *Phys. Lett. B 129*, 273, 1983. 9
- [51] Hooft, G.’t, *Renormalizable Lagrangians for massive Yang-Mills fields*, Nucl. Phys. B35, 167, 1971. 9
- [52] Hooft, G.’t, *Renormalization of massless Yang-Mills fields*, Nucl. Phys. B33, 173, 1971. 9
- [53] Hooft, G.’t, Veltman, M., *Diagrammatica*, CERN Yellow Reports, 1973. 6
- [54] Kerson Huang, *Quarks, Leptons and Gauge Fields*, World Scientific, 1992.
- [55] I. Bigi *et al.*, “Production and decay properties of ultra-heavy quarks,” *Phys. Lett. B*, 181:157, 1998. 16
- [56] B. Heinemann J.-F. Arguin, “Underlying Event Corrections in Run II,” *Internal CDF Note 6239*, 2003. 111
- [57] Jim Bellinger *et al.*, “A Guide to Muon Reconstruction for Run 2,” *Internal CDF Note 5870*, 2002. 72, 88
- [58] J.M. Campbell *et al.*, “MCFM v3.4.3 Users Guide,” <http://mcfm.fnal.gov>, 2004. 148
- [59] J.M. Campbell, R. K. Ellis, “An update on vector boson pair production at hadron colliders,” *hep-ph/9905386*, 1999. 148, 150

- [60] J.Thom, D.Glenzinski, M.Herndon, C.-J.Lin, A.Yagil, “Determination of the Run II COT Tracking Efficiency using the W-No-Track Sample,” *Internal CDF Note 6866*, 2004. 66
- [61] K. Anikeev *et al*, “EVB and Level3 for Aces,” *Internal CDF Note 5793*, 2001. 57
- [62] K. Bloom and W.-M. Yao, ““Outside-In” Silicon Tracking at CDF,” *Internal CDF Note 5991*, 2002. 65
- [63] K. Nakamura *et al.*, Super-Kamiokande Collaboration, “Atmospheric neutrino results from Super-Kamiokande,” *Nucl. Phys. Proc. Suppl.*, 66:282-289, 1998. 8
- [64] K. Yasuoka *et al*, “,” *Nuclear Instruments and Methods*, A267(315), 1988. 71
- [65] Kirsten Tollefson, “TDWG Report,” *CDF Week - Collaboration Meeting*, July 26, 2004. 58
- [66] Mark Kruse, *Observation of Top Quark Pair Production in the Dilepton Decay Channel from Proton-Antiproton Collisions at $\sqrt{s}=1.8$ TeV*, PhD thesis, Perdue University, 1996. 90, 161, 167
- [67] L. Nodulman, “Central Electron Issues for Summer 2003,” *Internal CDF Note 6558*, 2003. 72
- [68] M. Cacciari, S. Frixione, M.L. Mangano, P. Nason, G. Ridolfi, “The t-ttbar cross-section at 1.8 and 1.96 TeV: a study of the systematics due to parton densities and scale dependence,” *hep-ph/0303085*, 2003. 23, 24
- [69] M. Coca *et al*, “A First Look at Run 2 High p_T Electrons,” *Internal CDF Note 5803*, 2003. 71
- [70] Giulia Manca, *A measurement of the ratio R in p \bar{p} collisions at $\sqrt{s}=1.96$ TeV*, PhD thesis, Oxford University, 2003. 87
- [71] Mehen T., “Summing $O(\alpha^{n+1})$ Corrections to Top Quark Decays,” *Phys. Lett. B*, 417:353, 1998. 16

- [72] Sarah Lockwitz Mircea Coca, Eva Halkiadakis, “Central Electron Identification Efficiencies for the 200 pb^{-1} Run 2 Dataset,” *Internal CDF Note 6580*, 2004. 95, 96, 120, 125
- [73] Mircea Coca, Paul Tipton, “Optimization Studies for the Top Dilepton Cross-Section Measurement,” *Internal CDF Note 6579*, 2003. 106, 113
- [74] Monica Tecchio, “Acceptance and background systematics for the Top Dilepton cross-section measurement,” *Internal CDF Note 6590*, 2003. 124
- [75] T.K. Nelson, “The CDF Layer 00 Detector,” *Internal CDF Note 5780*, 2001. 43
- [76] P. Azzi, G. Busetto, P. Gatti, A. Ribon, “Histogram Tracking in the COT,” *Internal CDF Note 5562*, 2001. 65
- [77] P. Tamburello, “Electron Task Force Web Page,” http://www-cdf.fnal.gov/internal/physics/top/r2leptons/etf/agenda/10_15_01.html, 2003. 72
- [78] P. Wilson, “Calorimeter Isolation and Lateral Shower Leakage for Photons and Electrons,” *Internal CDF Note 4170*, 1998. 91
- [79] Paolo Gatti, “Performance of the new tracking system at CDF II,” *Internal CDF Note 5561*, 2001. 65
- [80] P.Nason, S. Dawson, and R. K. Ellis, “The total cross-section for the production of heavy quarks in hadronic collisions,” *Nucl. Phys. B* 303,607, 1988. 20
- [81] Q. R. Ahmad *et al.*, SNO Collaboration, “Direct evidence for neutrino flavor transformation from neutral-current interactions in the Sudbury Neutrino Observatory,” *Phys. Rev. Lett.*, 89:011301, *nucl-ex/0204008*, 2002. 8
- [82] R. Barate *et al.*, ALEPH Collaboration, “hep-ex/0306033,” *Phys. Lett. B* 565, 2003. 10
- [83] B. WINER R. HUGHES, T. LISS, “Combined SVX-SLT-Dilepton Top Cross Section, Run1A+Run1B,” *Internal CDF Note 3111*, 1995. 167

- [84] Rajan Gupta, “Introduction to Lattice QCD,” *hep-lat/9807028*, 1997. 8
- [85] Robert G. Wagner, “Electron Identification for Run II: Algorithms,” *Internal CDF Note 5466*, 2003. 68
- [86] Robert G. Wagner (ANL), “Understanding and Using Lshr,” *Internal CDF Note 6249*, 2003. 90
- [87] S. Eidelman *et al*, “Review of Particle Physics,” *Phys. Lett. B592*, 1, 2004. 6, 125, 146
- [88] S. L. Glashow, J. Illiopoulos, and L. Maiani, “Weak Interactions with Lepton-Hadron Symmetry,” *Phys. Rev. D 2*, 1285, 1970. 9
- [89] S. Van der Meer, “Stochastic Cooling in the CERN Antiproton Accumulator,” *IEEE Trans. Nucl. Sci. 29(1994)*, 1981. 184
- [90] J.Konigsberg S.Klimenko, “Average cross-section of the inelastic $p\bar{p}$ scattering at 1.8 TeV,” *Internal CDF Note 6314*, 2003. 78
- [91] Steve Kuhlmann, “Preradiator Chamber Aging Tests,” *Internal CDF Note 1290*, 1990. 35
- [92] The CDF Collaboration, “Physics Performance of the CDF II Detector with an inner silicon layer a time of flight detector,” *Internal CDF Note 5264*, 2000. 50
- [93] The CDF Collaboration, the D0 Collaboration, the Tevatron Electroweak Working Group, “Combination of CDF and D0 Results on the Top-Quark Mass,” *hep-ex/0404010*, 2004. 86
- [94] The Top Averaging Collaboration, “Combining the top quark mass results for CDF and D/O/,” *FERMILAB-TM-2084*,, 1999. 86
- [95] Evelyn Thomson, “Description of data samples for Top and Electroweak groups for Winter 2003,” *Internal CDF Note 6265*, 2003. 77

- [96] Tim Nelson, Rick Snider, David Stuart, “Forward Electron Tracking with the Phoenix-Mods Package,” *Internal CDF Note 6278*, 2003. 65
- [97] V. Berger, R. Phillips, *Collider Physics - Updated Edition*, Addison Wesley Publishing Company, 1996. 6
- [98] Lucio Cerrito Victoria Martin, “Muon Cuts and Efficiencies for 4.11 Analyses,” *Internal CDF Note 6825*, 2004. 104, 125
- [99] W. Beenakker *et al*, “Electroweak one loop contributions to top pair production in hadron colliders,” *Nucl. Phys. B*411,343, 1994. 22
- [100] W. Beenakker, H. Kuijf, W.L. van Neerven , J. Smith, “QCD corrections to heavy quark production in p anti-p collisions,” *Phys.Rev. D* 40, 54, 1989. 22
- [101] W. Giele *et al*, “The LHAPDF interface,” <http://vircol.fnal.gov>, 2004. 148
- [102] Yimei Huang, Chris Hays, and Ashutosh Kotwal, “Inside-Out Tracking,” *Internal CDF Note 6706*, 2003. 65

A The Fermilab Accelerator

The CDF detector records the result of the proton-antiproton collisions produced at the Tevatron accelerator, part of the Fermilab accelerator complex[1]. Until the Large Hadron Collider (LHC) begins to operate at CERN, in Europe, the Tevatron[2] is the highest energy collider in the world, and is currently the only facility able to produce and study top quarks.

A.1 A proton-antiproton collider

From optics it is known that the space resolution achieved by scattering of one particle from another is limited by the wavelength λ of their relative motion, $\lambda = 2\pi/k$, where k is their relative momentum. So to *probe small distances* requires large k , which implies very high energy of the center of mass, commonly denoted as \sqrt{s} . Also the second reason to ask for high energy is *new particle production*.

For a fixed target experiment, when a particle A, with energy E_A , meets a particle B, of mass m_B , in a stationary target, their center of mass energy is

$$\sqrt{s} = \sqrt{m_A^2 + m_B^2 + 2 \cdot E_A \cdot m_B} \approx \sqrt{2 \cdot E_A \cdot m_B} \quad (\text{A.1})$$

assuming $E_A \gg m_A, m_B$.

On the other hand, in a collider experiment, where a particle A collides with a particle B, the available energy becomes

$$\sqrt{s} = \sqrt{m_A^2 + m_B^2 + 2 \cdot E_A \cdot E_B - 2 \cdot \mathbf{p}_A \cdot \mathbf{p}_B} \approx \sqrt{4 \cdot E_A \cdot E_B} \quad (\text{A.2})$$

assuming $E_A, E_B \gg m_A, m_B$. Therefore the available energy at a collider increases linearly with the beam energy, while for a fixed target increases only with the square of beam energy. This makes an accelerator with two colliding beams a better choice.

Another major characteristic of a collider experiment, beyond \sqrt{s} , is the luminosity L , the product of incident beam flux with the mean target density. If the cross section of a process is σ , the number of events ΔN produced in a time Δt is $\sigma \cdot \int_{\Delta t} L dt$. One defines the *integrated luminosity* \mathcal{L} , as the $\mathcal{L} = \int_{\Delta t} L dt$. L is known as the *instantaneous luminosity*.

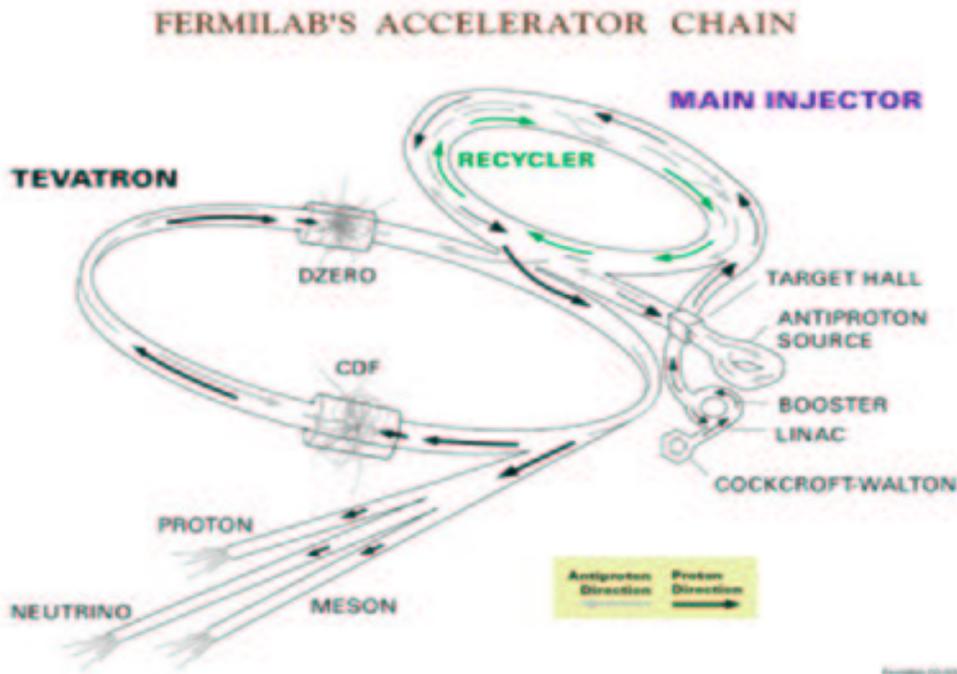


Figure A.1: Diagram of the Tevatron accelerator complex. The Antiproton Source consists of the Debuncher and the Accumulator.

For the Tevatron the instantaneous luminosity, a measure of the collision rates, is ex-

pressed as:

$$\mathcal{L} = \frac{N_p N_{\bar{p}} B f_0}{4\pi\sigma^2} \quad (\text{A.3})$$

where $N_p(N_{\bar{p}})$ represents the number of protons (antiprotons) per bunch, B is the number of bunches of each type, f_0 is the revolution frequency of the bunches and, σ is the transverse area of the bunches.

A.2 The Production of Protons and Antiprotons

After the Run I ended few upgrades were made to the Tevatron accelerator complex, for example the Main Ring was replaced by the Main Injector, and a new antiproton storage ring, the Recycler, was built, all with the goal to increase the instantaneous luminosity.

Protons production

The process of creating protons consists of *three stages*, described below. In the *first stage*, the gaseous hydrogen H_2 is ionized to produce negative ions H^- . Then the hydrogen anions are accelerated by a Cockcroft-Walton electrostatic preaccelerator until they reach an energy of 750 keV. In the *next stage* a 150 m long linear accelerator (LINAC) uses oscillating electric fields to accelerate the H^- ions to 400 MeV. At the end of the LINAC tunnel, the negative ions strike a thin carbon foil which strips the electrons off, giving rise to a proton beam. In the *third stage* the protons are injected to the Booster, a proton synchrotron of 150 m diameter. In the Booster the protons are accelerated to an energy of 8 GeV and bunches of $6 \cdot 10^{10}$ protons are being formed, which are then sent to Main Injector. Here they are accelerated from 8 to 150 GeV, then they are sent into the Tevatron, to collide with the antiprotons.

Main Injector

The Main Injector, completed in 1999, replaces the Main Ring, used in Run I. It is a 3 km in circumference synchrotron, built to increase the antiproton production efficiency. By using the Main Injector allows to increase by a factor of ten the number of antiprotons at the beginning of a store, as compared with the Run I.

Antiprotons production

The antiprotons are more difficult to produce. Protons from the Main Injector, which reached 120 GeV, are sent toward a nickel target to produce secondary particles, few of which are \bar{p} s. The antiprotons are selected from the rest of the particles and focused through lithium lenses. Next they enter the Debuncher, a triangular-shaped synchrotron which uses the *stochastic cooling*¹ to decrease the momentum distribution of the bunches. This process increases the spatial distribution of the particles, producing a continuous beam. Therefore the antiprotons are sent to an Accumulator loop, where the bunch structure is recreated and then they are further cooled down. Then the bunches are accelerated to 150 GeV and injected into the Tevatron.

Recycler

A major limiting factor to attain a high luminosity is the number of antiprotons². Therefore a Recycler Ring was constructed in the same tunnel as the Main Injector. It is supposed to collect the antiprotons left at the end of a store (about 75% of the initial number) and reinject them in the next store. The Recycler is not yet commissioned.

A.3 The Tevatron

The Tevatron, shown in Figure A.1, is a 1 km radius circular accelerator, colliding protons with antiprotons, at a center of mass energy of 1.96 TeV. The protons and antiprotons are injected into the Tevatron, in the opposite direction and on different orbits. They enter at 150 GeV energy and are accelerated at 980 GeV. The two beams are then focused using magnetic quadrupoles in two regions around the ring, where the proton and antiprotons collide every 396 ns³. At each collision region there is a collider experiment, DØ in the region D0, and CDF in B0 region respectively. The corresponding detectors record the particles produced during the p- \bar{p} collisions.

¹The stochastic cooling was developed by Simon van der Meer[89] at CERN and was the most significant development toward high intensity antiproton beams.

²The instantaneous luminosity depends linearly on the number of antiprotons, as seen in Eq. A.3

³Initial plan was that the time between two consecutive collisions would be 132 ns.

B MC Breakdown Tables

We show the breakdown of the backgrounds, which used Monte Carlo generators in the estimation process. These numbers are shown before the scale factors, luminosity weighting or Phoenix fake rates are folded in.

Category	Dilepton Cut							NJets			H_T	OS
	Geo Kin	ID	Iso	Conv Cosm	Z_{veto}	\cancel{E}_T	$\Delta\phi$	0j	1j	2j		
TCE/TCE	1545	845	828	782	627	568	526	8	69	449	438	438
TCE/NTCE					103	94	86	4	24	58	55	49
TCE/PHX	538	365	356	341	303	277	259	2	43	214	207	195
NTCE/IPHX					24	21	20	0	5	15	15	13
ee	2083	1210	1184	1123	1057	960	891	14	141	736	715	695

Table B.1: Breakdown of the e/e final states of the $t\bar{t}$ acceptance using Pythia MC.

Category	Geo	ID	Iso	CC	Z_{veto}	\cancel{E}_T	$\Delta\phi$	0j	1j	2j	H_T	OS
CMUP/CMUP	346	332	330	330	270	241	221	0	36	185	175	175
CMUP/NCMUP					38	35	32	1	1	30	29	28
CMUP/CMU	122	115	98	98	81	72	65	1	12	52	51	51
CMUP/NCMU					13	13	11	0	1	10	10	8
CMUP/CMP	172	170	160	160	139	123	114	1	15	98	94	94
CMUP/NCMP					16	16	14	0	0	14	14	11
CMUP/CMX	259	207	207	207	172	152	140	0	15	125	120	120
CMUP/NCMX					10	10	10	0	0	10	10	10
NCMUP/CMX					11	11	10	0	0	10	9	8
CMUP/CMIO	455	184	148	148	139	128	116	0	23	93	87	87
CMX/CMX	65	45	45	45	36	32	31	0	4	27	27	27
CMX/NCMX					8	8	8	0	0	8	8	6
CMX/CMU	53	41	33	33	28	24	24	0	1	23	23	23
CMX/NCMU					3	3	3	0	0	3	3	3
CMX/CMP	72	63	58	58	47	41	37	0	0	37	34	34
CMX/NCMP					3	2	2	0	1	1	1	1
CMX/CMIO	229	91	74	74	68	61	60	0	10	50	49	49
$\mu\mu$	1773	1248	1153	1153	1082	972	898	3	119	776	744	735

Table B.2: Breakdown of the μ/μ final states of the $t\bar{t}$ acceptance using Pythia MC.

Category	Geo	ID	Iso	CC	Z_{veto}	\cancel{E}_T	$\Delta\phi$	0j	1j	2j	H_T	OS
TCE/CMUP	1647	1078	1072	1039	890	782	705	4	89	612	595	594
TCE/NCMUP					92	83	75	0	5	70	68	64
NTCE/CMUP					57	52	49	3	11	35	34	33
TCE/CMU	300	206	190	184	166	147	132	2	14	116	110	110
TCE/NCMU					18	14	11	0	0	11	8	8
TCE/CMP	432	304	277	270	240	212	197	1	28	168	163	163
TCE/NCMP					30	29	28	0	1	27	25	24
TCE/CMX	636	369	366	356	299	274	251	1	40	210	203	203
TCE/NCMX					29	24	20	0	2	18	18	17
NTCE/CMX					28	25	23	1	6	16	16	16
TCE/CMIO	1483	361	311	301	301	268	249	3	30	216	209	209
PHX/CMUP	290	241	236	233	217	180	163	1	28	134	129	114
PHX/NCMUP					16	12	10	0	0	10	10	8
PHX/CMU	60	45	45	45	39	33	31	1	3	27	25	25
PHX/NCMU					6	6	6	0	1	5	5	5
PHX/CMP	83	68	67	66	58	49	45	0	3	42	39	37
PHX/NCMP					8	5	4	0	1	3	2	2
PHX/CMX	95	66	65	65	63	54	46	2	4	40	37	36
PHX/NCMX					2	2	2	0	0	2	2	2
PHX/CMIO	271	75	63	63	63	52	44	1	6	37	36	32
$e\mu$	5297	2813	2692	2622	2692	2303	2091	20	272	1799	1734	1702

Table B.3: Breakdown of the e/μ final states of the $t\bar{t}$ acceptance using Pythia MC.

Category	Geo	ID	Iso	CC	Z_{veto}	\cancel{E}_T	$\Delta\phi$	0j	1j	2j	H_T	OS
3ℓ	2992	48	41	36	0	29	25	0	9	16	16	0
Total	12786	5319	5070	4898	4831	4264	3905	37	541	3327	3209	3132

Table B.4: Multilepton and summary of the $t\bar{t}$ acceptance using Pythia MC. The trigger efficiencies, scale factors and the rest of the corrections are not applied.

Category	Dilepton Cut							NJets			H_T	OS
	Geom Kin	ID	Iso	Conv Cosm	Z_{veto}	\cancel{E}_T	$\Delta\phi$	0j	1j	2j		
TCE/TCE	11238	10172	10158	9811	9116	473	179	27	101	51	36	36
TCE/NTCE					666	35	11	1	7	3	2	2
TCE/PHX	7199	6768	6668	6486	6216	264	91	13	58	20	8	8
NTCE/IPHX					244	12	5	0	3	2	1	1
ee	18437	16940	16826	16297	16242	784	286	41	169	76	47	47

Table B.5: Breakdown of the e/e final states of the $Z \rightarrow \tau\tau$ acceptance using Pythia MC.

Category	Dilepton Cut							NJets			H_T	OS
	Geo	ID	Iso	CC	Z_{veto}	\cancel{E}_T	$\Delta\phi$	0j	1j	2j		
CMUP/CMUP	4203	3569	3559	3559	3325	157	43	7	23	13	9	9
CMUP/NCMUP					225	17	3	1	2	0	0	0
CMUP/CMU	1608	1410	1368	1368	1321	63	30	4	15	11	8	8
CMUP/NCMU					44	4	2	0	0	2	1	1
CMUP/CMP	2400	2101	2032	2032	1976	87	36	3	20	13	10	10
CMUP/NCMP					53	2	2	0	1	1	1	1
CMUP/CMX	3910	2798	2794	2794	2601	125	50	5	27	18	10	10
CMUP/NCMX					94	9	3	0	1	2	0	0
NCMUP/CMX					95	13	4	0	1	3	3	3
CMUP/CMIO	4754	2624	2439	2439	2435	127	45	9	23	13	5	5
CMX/CMX	1131	748	748	748	698	42	13	2	7	4	3	3
CMX/NCMX					46	2	0	0	0	0	0	0
CMX/CMU	903	667	647	647	617	26	7	0	3	4	2	2
CMX/NCMU					29	5	1	1	0	0	0	0
CMX/CMP	1035	756	726	726	707	45	17	1	10	6	5	5
CMX/NCMP					18	0	0	0	0	0	0	0
CMX/CMIO	2384	1056	985	985	983	63	37	2	27	8	4	4
$\mu\mu$	22328	15729	15298	15298	15267	787	293	35	160	98	61	61

Table B.6: Breakdown of the μ/μ final states of the $Z \rightarrow \tau\tau$ acceptance using Pythia MC.

Category	Dilepton Cut							NJets			H_T	OS
	Geo	ID	Iso	CC	Z_{veto}	\cancel{E}_T	$\Delta\phi$	0j	1j	2j		
TCE/CMUP	14332	12069	12052	11857	11099	588	209	30	116	63	39	39
TCE/NCMUP					365	35	7	0	6	1	1	1
NTCE/CMUP					393	37	13	0	4	9	7	7
TCE/CMU	3061	2510	2432	2395	2300	121	48	13	25	10	6	6
TCE/NCMU					95	5	2	1	1	0	0	0
TCE/CMP	3856	3366	3238	3186	3082	171	58	16	25	17	12	12
TCE/NCMP					104	8	3	0	1	2	0	0
TCE/CMX	7325	4983	4968	4877	4582	226	97	12	49	36	20	20
TCE/NCMX					161	11	5	0	3	2	2	2
NTCE/CMX					134	9	3	0	2	1	1	1
TCE/CMIO	16159	4683	4363	4294	4294	208	82	16	51	15	9	9
PHX/CMUP	3775	3377	3328	3300	3204	110	41	2	24	15	10	9
PHX/NCMUP					96	3	2	0	2	0	0	0
PHX/CMU	923	858	843	839	812	36	8	2	4	2	2	2
PHX/NCMU					27	3	0	0	0	0	0	0
PHX/CMP	971	895	880	875	840	22	6	0	4	2	0	0
PHX/NCMP					35	3	1	0	1	0	0	0
PHX/CMX	2873	2154	2129	2111	2035	85	36	4	23	9	5	5
PHX/NCMX					76	5	2	0	1	1	0	0
PHX/CMIO	6277	2025	1910	1893	1893	76	29	1	16	12	10	10
$e\mu$	59552	36920	36143	35627	35627	1762	652	97	358	197	124	123

Table B.7: Breakdown of the e/μ final states of the $Z \rightarrow \tau\tau$ acceptance using Pythia MC.

Category	Dilepton Cut							NJets			H_T	OS
	Geo	ID	Iso	CC	Z_{veto}	\cancel{E}_T	$\Delta\phi$	0j	1j	2j		
3ℓ	10055	17	11	8	0	2	1	1	0	0	0	0
4ℓ	331	0	0	0	0	0	0	0	0	0	0	0
Total 422363	112697	69606	68278	67222	67136	3335	1232	174	687	371	232	231

Table B.8: Multilepton and summary of the $Z \rightarrow \tau\tau$ acceptance using Pythia MC. The trigger efficiencies, scale factors and the rest of the corrections are not applied.

Category	Dilepton Cut						NJets			H_T	OS
	ID	Iso	CC	Z_{veto}	\cancel{E}_T	$\Delta\phi$	0j	1j	2j		
TCE/TCE	12093	11949	11434	10725	7893	7591	6052	1230	309	200	200
TCE/NTCE				618	466	448	316	95	37	24	24
TCE/PHX	9115	9021	8734	8440	6141	5928	4760	954	214	139	133
NTCE/IPHX				213	149	145	102	30	13	9	9
ee	21208	20970	20168	19996	14649	14112	11230	2309	573	372	366

Table B.9: Breakdown of the e/e final states of the WW acceptance using Pythia MC.

Category	Dilepton Cut						NJets			H_T	OS
	ID	Iso	CC	Z_{veto}	\cancel{E}_T	$\Delta\phi$	0j	1j	2j		
CMUP/CMUP	4103	4076	4076	3795	2755	2633	2100	433	100	71	71
CMUP/NCMUP				255	209	204	107	79	18	12	12
CMUP/CMU	1720	1654	1654	1602	1131	1093	862	196	35	25	25
CMUP/NCMU				43	33	32	16	11	5	1	1
CMUP/CMP	2327	2256	2256	2178	1543	1486	1169	260	57	41	41
CMUP/NCMP				63	51	50	16	27	7	5	5
CMUP/CMX	3355	3353	3353	3146	2264	2178	1723	378	77	52	52
CMUP/NCMX				91	70	67	36	21	10	6	6
NCMUP/CMX				88	71	68	22	29	17	10	9
CMUP/CMIO	2964	2785	2785	2769	2066	1985	1560	337	88	53	53
CMX/CMX	705	704	704	659	448	425	359	51	15	9	9
CMX/NCMX				40	27	27	12	11	4	3	3
CMX/CMU	738	718	718	691	512	483	385	90	8	7	7
CMX/NCMU				20	17	16	5	11	0	0	0
CMX/CMP	941	900	900	866	638	616	522	69	25	15	15
CMX/NCMP				24	18	18	8	7	3	1	1
CMX/CMIO	1346	1256	1256	1243	925	896	755	116	25	18	18
$\mu\mu$	18199	17702	17702	17573	12778	12277	9657	2126	494	329	328

Table B.10: Breakdown of the μ/μ final states of the WW acceptance using Pythia MC.

Category	Dilepton Cut						NJets			H_T	OS
	ID	Iso	CC	Z_{veto}	\cancel{E}_T	$\Delta\phi$	0j	1j	2j		
TCE/CMUP	13946	13845	13523	12756	9183	8822	7113	1371	338	235	235
TCE/NCMUP				426	334	318	160	118	40	25	25
NTCE/CMUP				341	245	235	159	63	13	7	7
TCE/CMU	2885	2793	2726	2629	1931	1851	1503	284	64	42	42
TCE/NCMU				97	83	80	49	22	9	5	5
TCE/CMP	3791	3677	3607	3466	2536	2439	1958	371	110	74	74
TCE/NCMP				141	106	102	62	27	13	9	8
TCE/CMX	5826	5794	5677	5339	3925	3778	3064	569	145	89	89
TCE/NCMX				202	169	164	89	52	23	17	17
NTCE/CMX				136	103	100	66	24	10	6	6
TCE/CMIO	5444	5049	4951	4951	3612	3476	2749	573	154	104	104
PHX/CMUP	5160	5107	5072	4938	3491	3376	2736	518	122	75	64
PHX/NCMUP				134	97	96	40	36	20	13	12
PHX/CMU	1186	1171	1162	1133	820	786	644	117	25	16	14
PHX/NCMU				29	25	24	12	10	2	0	0
PHX/CMP	1363	1343	1331	1296	917	880	714	137	29	20	19
PHX/NCMP				35	28	28	15	10	3	3	3
PHX/CMX	2417	2392	2370	2310	1655	1590	1281	257	52	22	20
PHX/NCMX				60	44	42	15	19	8	6	5
PHX/CMIO	2226	2112	2087	2087	1472	1428	1186	211	31	19	19
$e\mu$	44244	43283	42506	43283	30776	29615	23615	4789	1211	787	768

Table B.11: Breakdown of the e/μ final states of the WW acceptance using Pythia MC.

Category	Dilepton Cut						NJets			H_T	OS	
	ID	Iso	C	Z_{veto}	\cancel{E}_T	$\Delta\phi$	0j	1j	2j			
3ℓ	35	24	9	0	8	8	3	5	0	0	0	
4ℓ	0	0	0	0	0	0	0	0	0	0	0	
Total	146298	83686	81979	80376	80852	58211	56012	44505	9229	2278	1488	1462

Table B.12: Multilepton and summary of the WW acceptance using Pythia MC. The trigger efficiencies, scale factors and the rest of the corrections are not applied.

Category	Dilepton Cut							NJets			H_T	OS
	Geo	ID	Iso	CC	Z_{veto}	\cancel{E}_T	$\Delta\phi$	0j	1j	2j		
TCE/TCE	5023	4873	4866	4673	2138	1143	1043	449	481	113	85	71
TCE/NTCE					260	102	90	44	36	10	7	6
TCE/PHX	4453	4457	4378	4250	2082	1078	974	387	489	98	68	50
NTCE/IPHX					116	40	34	15	13	6	4	2
ee	9476	9330	9244	8923	4596	2363	2141	895	1019	227	164	129

Table B.13: Breakdown of the e/e final states of the WZ acceptance using Pythia MC.

Category	Dilepton Cut							NJets			H_T	OS
	Geo	ID	Iso	CC	Z_{veto}	\cancel{E}_T	$\Delta\phi$	0j	1j	2j		
CMUP/CMUP	1713	1582	1580	1580	821	396	368	219	122	27	20	20
CMUP/NCMUP					80	39	36	15	13	8	7	7
CMUP/CMU	747	700	668	668	363	181	168	92	58	18	15	15
CMUP/NCMU					11	3	3	1	1	1	1	1
CMUP/CMP	963	906	866	866	436	200	178	110	57	11	8	7
CMUP/NCMP					19	8	5	1	4	0	0	0
CMUP/CMX	1883	1481	1477	1477	737	372	342	190	125	27	21	17
CMUP/NCMX					22	11	9	2	7	0	0	0
NCMUP/CMX					38	13	12	4	5	3	3	2
CMUP/CMIO	1987	1312	1192	1192	655	324	292	184	88	20	17	16
CMX/CMX	454	325	323	323	161	79	72	41	28	3	2	2
CMX/NCMX					9	5	5	3	1	1	0	0
CMX/CMU	395	305	285	285	140	52	48	32	14	2	2	2
CMX/NCMU					8	4	3	0	2	1	1	1
CMX/CMP	504	413	391	391	188	85	79	46	26	7	5	5
CMX/NCMP					12	8	7	2	3	2	1	1
CMX/CMIO	1024	585	530	530	306	153	143	85	45	13	7	7
$\mu\mu$	9670	7609	7312	7312	4006	1933	1770	1027	599	144	110	103

Table B.14: Breakdown of the μ/μ final states of the WZ acceptance using Pythia MC.

Category	Dilepton Cut							NJets			H_T	OS
	Geo	ID	Iso	CC	Z_{veto}	\cancel{E}_T	$\Delta\phi$	0j	1j	2j		
TCE/CMUP	846	643	643	623	570	457	414	220	149	45	23	14
TCE/NCMUP					35	25	23	7	9	7	4	1
NTCE/CMUP					18	11	10	3	5	2	1	1
TCE/CMU	186	117	113	111	108	81	77	43	25	9	7	3
TCE/NCMU					3	2	2	0	1	1	1	0
TCE/CMP	241	184	175	164	148	110	101	58	38	5	5	2
TCE/NCMP					16	13	12	4	5	3	2	2
TCE/CMX	452	282	280	274	250	194	178	85	72	21	17	8
TCE/NCMX					14	12	12	3	7	2	0	0
NTCE/CMX					10	8	8	3	4	1	1	1
TCE/CMIO	2441	328	289	283	283	224	209	110	69	30	24	12
PHX/CMUP	433	368	367	363	342	263	231	129	82	20	13	10
PHX/NCMUP					21	18	18	3	8	7	5	2
PHX/CMU	81	74	75	73	71	61	56	31	15	10	6	4
PHX/NCMU					2	2	2	0	1	1	0	0
PHX/CMP	95	86	85	84	82	65	57	31	25	1	0	0
PHX/NCMP					2	2	2	1	1	0	0	0
PHX/CMX	200	156	155	153	149	113	107	53	42	12	5	0
PHX/NCMX					4	3	3	0	1	2	1	1
PHX/CMIO	899	156	149	146	146	103	97	57	32	8	4	2
$e\mu$	5874	2394	2331	2274	2331	1767	1619	841	591	187	119	63

Table B.15: Breakdown of the e/μ final states of the WZ acceptance using Pythia MC.

Category	Dilepton Cut							NJets			H_T	OS	
	Geo	ID	Iso	CC	Z_{veto}	\cancel{E}_T	$\Delta\phi$	0j	1j	2j			
3ℓ	7702	1933	1864	1798	0	1408	1222	966	214	42	40	0	
4ℓ	1050	2	1	0	0	0	0	0	0	0	0	0	
Total	33847	32722	21266	20751	18509	10933	7471	6752	3729	2423	600	433	295

Table B.16: Multilepton and summary of the WZ acceptance using Pythia MC. The trigger efficiencies, scale factors and the rest of the corrections are not applied.