Tests of Enhanced Leading Order QCD in W Boson plus Jet Production in 1.96-TeV Proton-Antiproton Collisions

Soushi TSUNO

A dissertation submitted to the Doctoral Program
in Physics, the University of Tsukuba
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Science

January 2004

Abstract

We have studied the W + $\geq$ n jets process in Tevatron Run II experiment. The data used correspond to a total integrated luminosity of 72 pb^{-1} taken from March 2002 through January 2003. The lowest order QCD predictions have been tested with a new prescription of the parton-jet matching, which allows to construct the enhanced LO phase space. According to this procedure, one gets unique results which do not depend on unphysical bias of kinematical cuts to avoide the collinear/infrared divergence in calculation. Namely, one can get the meaningful results in the lowest order prediction. The controllable event samples of the W boson plus jets events by the enhanced lowest order prediction will lead smaller systematic uncertainty than the naive prediction without any cares of the collinear/infrared divergence. We expect our method will be also useful to make systematically small samples as the background estimates in the top quark analysis.

We found a good agreement between data and theory in typical kinematics distributions. The number of events for each inclusive sample up to 3 jets are compared with Monte Carlo calculations. A comparison with Run I results is also presented. This is the first result for the CDF Run II experiment.

Acknowledgements

At first, I would like to thank my thesis adviser Prof. Fumihiko Ukegawa, Prof. Shinhong Kim, Prof. Koji Takikawa, and Prof. Kazuhiko Hara in the institute of physics, University of Tsukuba for useful guidance and financial supports staying at Fermilab.

I learned lots of things from the number of people : Prof. Henry J. Frisch gave lots of advises and supports. His words are always inspired me with many physics interests. Prof. Joey W. Huston is a good person for me to share the idea and discuss about our interests of physics. Prof. Michelangelo Mangano and Prof. Stephen Mrenna showed me the direction for my thesis. I respect their thoughtful idea and works. Dr. Jay R Dittmann and Dr. Mario Martinez-Perez, who are the convenors of QCD physics group, suggested many fruitful comments. Without their helps, I did not complete my thesis. Also, I consulted Anwar A. Bhatti about a jet physics, many times. I also thank Mr. Andrea Messina for helping cross-checking of our results.

On the other hand, my code development in CDF offline software was supported by the CDF Simulation group. I would especially like to thank Prof. Alan S. Thompson, Prof. Marjorie D. Shapiro, Prof. Manfred Paulini, and Dr. Elena A. Gerchtein.

Early in my staying at Fermilab, I had a great opportunity to install the silicon port-card into the CDF detector in the collision hall. I had a precious experience to feel the enormous CDF detector. I would like to thank all the members of Silicon group. Dr. Douglas A. Glenzinski, Dr. Gino Bolla, Dr. Steve Nahn, Prof. Jean Slaughter, and Dr. Lester H. Miller gave me an exciting time of this detector installation.

Finally, I would like to greatly thank the members of Minami-Tateya numerical calculation group lead by Prof. Yoshimitsu Shimizu and the ATLAS-Japan group formed by Japanese members of the ATLAS experiment. Thanks to many teachings and suggestions from Dr. Shigeru Odaka, Dr. Junpei Fujimoto, Dr. Yoshimasa Kurihara, Dr. Tadashi Ishikawa, Dr. Toshiaki Kaneko, and Prof. Kiyoshi Kato, I could learned a lot of physics basis through the development of the event generator.

I would like to thank all members not a few related to me. And to my family – I could not spend so much with you together during this study – I love to say thank you.

Contents

1 Introduction **1**

1.1 Overview 1

1.2 The Standard Model 2

1.3 Quantum Chromodynamics (QCD) 4

1.3.1 Lagrangian 4

1.3.2 Feynman rule 5

1.3.3 Asymptotic freedom 6

1.3.4 Running coupling 7

1.3.5 Parton distribution function 8

1.3.6 Nonperturbative QCD 9

1.3.7 Infrared divergences 9

1.3.8 Total cross section 11

1.3.9 Factorization and renormalization scales 12

1.4 Dynamics of multi-particle final state 14

1.5 Test of QCD using W + jets events 16

2 The CDF Experiment at the Tevatron Collider **19**

2.1 Introduction 19

2.2 Characteristics of Fermilab accelerators 19

2.2.1 Accelerator chain 19

2.2.2 Production of antiproton 20

2.2.3 The Tevatron . . . 22
2.2.4 The Tevatron luminosity . . . 23
2.3 The CDF II detector . . . 25
2.3.1 Overview . . . 25
2.3.2 Tracking systems . . . 26
2.3.3 Calorimeter systems . . . 32
2.3.4 Muon systems . . . 35
2.3.5 Trigger and data acquisition systems . . . 39

3 $\mathbf{W^{\pm} \to e^{\pm}\nu}$ Event Selection 43
3.1 Introduction . . . 43
3.2 Electrons . . . 43
3.2.1 Electron triggers . . . 43
3.2.2 Electron clustering . . . 45
3.2.3 Corrections . . . 46
3.2.4 Electron geometric and kinematic cuts . . . 47
3.2.5 Electron identification cuts . . . 48
3.3 W selection . . . 51
3.3.1 Missing transverse energy ($\not{E}_T \geq 30$ GeV) . . . 51
3.3.2 $Z \to e^+e^-$ removal . . . 51
3.4 Jets . . . 52
3.4.1 Jet identification algorithm, JetClu . . . 52
3.4.2 Variables in JetClu . . . 53
3.4.3 Jet energy corrections . . . 54
3.4.4 Systematic uncertainties . . . 58
3.5 Luminosity measurement . . . 58
3.6 The $W \to e\nu$ event sample . . . 59
3.6.1 Datasets . . . 59
3.6.2 The $W \to e\nu$ event sample . . . 60

3.7 Cross check of inclusive W production cross section 62
3.7.1 Cross section formula 62
3.7.2 Trigger efficiency 63
3.7.3 Electron identification efficiency 64
3.7.4 Z vertex efficiency 67
3.7.5 Acceptance calculation 68
3.7.6 Background estimation 68
3.7.7 Inclusive $W \to e\nu$ production cross section 72
3.8 Summary 73

4 Jet Separation Study **75**
4.1 Introduction 75
4.2 Cone jet algorithms 77
4.2.1 Jet definitions 77
4.2.2 The infrared/collinear problems 77
4.3 Fake jets event generator 80
4.4 The separation power of JetClu algorithm 81
4.4.1 Template data sets 81
4.4.2 E_T distributions 83
4.4.3 Reconstruction ratio 83
4.4.4 Parton-jet matching 86
4.4.5 Jet separation 88
4.5 Comparison with different cone algorithms 92
4.5.1 Comparison with MidPoint and SeedLess cone algorithms 92
4.5.2 Splitting and merging 93
4.5.3 The separation power of MidPoint and SeedLess algorithms 94
4.6 Summary 95

5 Jet Separation Method **101**

5.1 Introduction . . . 101

5.2 Jet separation method . . . 103

5.2.1 Procedure of the parton-jet matching . . . 103

5.2.2 Parton-jet matching efficiency . . . 105

5.2.3 Kinematical dependence . . . 106

5.2.4 Normalization . . . 109

5.3 Summary . . . 110

6 Tests of Enhanced Leading Order QCD in W + jet Events **113**

6.1 Introduction . . . 113

6.2 $W + \geq$ n jets data samples . . . 113

6.2.1 Data samples . . . 113

6.2.2 Jet energy scale uncertainty . . . 114

6.2.3 Jet counting . . . 115

6.3 Comparisons of theory to data . . . 116

6.3.1 Jet E_T distribution . . . 116

6.3.2 Angular and mass distributions . . . 119

6.3.3 Jet multiplicity . . . 119

6.3.4 Ratio of the jet multiplicity . . . 121

6.3.5 Comparison with Run I measurement . . . 124

6.4 Summary . . . 124

7 Predictions of Quantum Chromodynamics **129**

7.1 Introduction . . . 129

7.2 Characteristic of W + n partons process . . . 130

7.2.1 W + n partons process in GR@PPA . . . 130

7.2.2 Input parameters . . . 131

7.2.3 Kinematical cuts . . . 135

7.3 Cross section results . . . 137
7.4 Scale dependence . . . 140
7.5 Uncertainty of PDF . . . 141
7.6 Heavy flavor fraction . . . 143
7.7 Initial state radiation . . . 144
7.8 Summary . . . 148

8 Results and Conclusions 151

A GR@PPA Event Generator 155
A.1 Introduction . . . 155
A.2 Extension of GRACE to $pp/p\bar{p}$ collisions . . . 156
A.2.1 Overview . . . 156
A.2.2 GRACE system . . . 157
A.2.3 Kinematics . . . 159
A.2.4 Flavor assignment . . . 160
A.2.5 Interface to Les Houches Accord . . . 161
A.3 GR@PPA_All . . . 163
A.3.1 Subprocesses . . . 163
A.3.2 Parameters . . . 165
A.3.3 User-defined energy scales and kinematical cuts . . . 169
A.4 Run the code . . . 173
A.4.1 Distribution package . . . 173
A.4.2 Mode selection . . . 174
A.4.3 Dependencies on PYTHIA . . . 175
A.4.4 Sample programs . . . 176
A.4.5 Output of results . . . 176
A.5 Physics processes in GR@PPA . . . 179
A.5.1 4 b-quark processes (ISUB : 160 - 164) . . . 179

A.5.2 W + jets process (ISUB : 100 - 103) 181
A.5.3 Z + jets process (ISUB : 110 - 112) 182
A.5.4 Top pair productions (ISUB : 170) 182
A.5.5 ZZ, WW, ZW processes (ISUB : 550 - 552) 183
A.6 Summary 184

List of Figures

1.1 Example for the soft/collinear singularity. (a) is an example for spacelike branch, and (b) is an example for timelike branch. 10

2.1 A schematic view of the Run II Tevatron accelerator complex at Fermilab. The new Main Injector is shown in the left side of the Tevatron Ring. The CDF detector is located at one of the collision points at B0 on the Tevatron Ring. 21

2.2 Total integrated luminosity delivered by the Tevatron and recorded by CDF in the period from July 2001 to July 2003. In this analysis, we use the data until January 2003 corresponding to 72.0 pb^{-1}. 25

2.3 A schematic view of the CDF II detector. 26

2.4 Longitudinal view of the CDF II Tracking System. 27

2.5 An end view of the CDF II silicon system including the SVX II cooling bulkheads and ISL support structure. 29

2.6 Nominal cell layout for SL2. 30

2.7 East endplate slots. Sense and field planes are at the clock-wise edge of each slot. 30

2.8 Tracks of a particle with positive/negative charge and positive/negative impact parameter. 32

2.9 A quarter view of the CDF II calorimeter systems and that segmentation. . 33

2.10 A sketch of one wedge of the CEM. 35

2.11 A sketch of the Central Electromagnetic Strip chamber (CES). 35

2.12 Cross section of upper part of new end plug calorimeter. 36
2.13 Geometry of the PES in a 45° sector. 36
2.14 Location of the central muon upgrade components in azimuth ϕ and pseudorapidity η for Run II. 38
2.15 The Run II readout functional block diagram of the three level pipelined and buffered trigger system. 40
2.16 Block diagram for the Run II trigger system. 42

3.1 Projection of the electron isolation to the missing E_T after all the selection criteria except the isolation and missing E_T cuts. 62
3.2 Level-1 (top) and Level-3 (bottom) tracking efficiencies as a function of η measured by CES. 65
3.3 E_T distribution of electron candidates from data passing the L2 pre-scaled trigger and some quarity cuts (top) and L2_CEM16 trigger efficiency as a function of E_T (bottom). 66
3.4 Isolation vs. $\not{E}_T$ 70

4.1 An illustration of infrared sensitivity in cone jet clustering. The length of the arrow represents the size of the momentum. In addition to the two partons emission (solid line) in (a), soft radiation (solid wave) is emitted between two particles in (b). 78
4.2 An illustration of collinear sensitivity in jet reconstruction. The jet reconstruction fails while jet in (a) is reconstructed in (b). 79
4.3 Another example of the collinear sensitivity in jet reconstruction. If the algorithm is sensitive to the E_T ordering of particles, this problem will happen. In (a), a jet is successfully reconstructed with the seed tower of the largest E_T. In (b), two jets are reconstructed with the seeds of the left-most particle and right-most particle, when the largest E_T particle splits into two lower E_T particles. 79

4.4 Demonstration of the two parton generation with the particular directions in ϕ-η plane. The first parton is forced to be the fixed direction by the input direction, $(\phi,\eta)=(\pi/3,-0.5)$. The second parton surrounds the first parton with the particular separation angle. The half-circle is taken to distribute in the central region of the calorimeters. From the top-left side, the size of the separation is varied with 0.1 step. 82

4.5 Leading jet E_T distribution reconstructed by JetClu with the cone size of 0.4. The energy correction is not applied. From top-left side, the separation angle, ΔR, between two generated-partons is changed with 0.1 step up to 0.9 of the separation angle. In this sample, both partons p_T are 30 GeV. . 84

4.6 Secondary jet E_T distribution reconstructed by JetClu with the cone size of 0.4. The energy correction is not applied. From top-left side, the separation angle, ΔR, between two generated-partons is changed with 0.1 step up to 0.9 of the separation angle. In this sample, both parton p_T are 30 GeV. . . 85

4.7 Probability of jets acceptance as a function of the separation angle between two partons. Each opened- and closed-circle is the probabilities of the single jet and two jets reconstruction, respectively. The squared represents the no-jet reconstruction probability. The minimum threshold energy of jet clustering is 3 GeV. Furthermore, the jet is required that the either primary partons exists inside the jet, which has the cone size of JetClu. In this sample, both primary partons p_T are 30 GeV. 87

4.8 Sample plot of the fitting result for two jets acceptance. Both primary partons p_T are 30 GeV. 88

4.9 The deviation of the corrected jet energy from the primary parton. The mean value and that RMS are taken. The closed-circle is for the leading jet, and the opened-square for the secondary jet. 90

4.10 The deviation of the observed jet separation from the primary parton separation. The central value and that sigma fitted by the Gaussian distribution with the range of the 25% height on the maximum value are taken. 90

4.11 The phase space of the jet separation as a function of the jet separation angle. The two-jets reconstruction region is the region where the reconstructed jets point to the primary partons with the probability of 95%. The error bar on this plot is taken as the resolution width at one sigma level. 91

4.12 Probability of jets acceptance of different cone algorithms as a function of the separation angle between two partons. Each opened- and closed-circle is the probabilities of the single jet and two jets reconstruction of JetClu, respectively. The squared and triangle marks represent each reconstruction probability of the MidPoint and SeedLess, too. The minimum threshold energy of jet clustering is 3 GeV. Both primary parton p_Ts are 30 GeV. . . 96

4.13 Two-jet separation border of each jet algorithm with the 95%, 50%, and 5% probabilities as a function of the secondary parton p_T, where the jet energy corrections are not applied for the jet E_T. 97

5.1 An illustration of collinear sensitivity in cone jet clustering. If a parton splitted from the primary parton of the ME calculation goes to between R_{cone} and $R_{sep}^{95\%}$, the mixed situation with one- or two-jets reconstruction happens. 105

5.2 Parton-jet matching efficiency after the jet clustering by JetClu algorithm at the detector level as a function of a jet p_T. The closed and opened circle marks are the matching efficiency by the $R_{sep}^{95\%}$ and $R_{cone}(0.4)$ basis, respectively. The hatched region is the systematics of the $R_{sep}^{95\%}$ basis. In addition to the circle plots, the three dashed lines are overlayed with the different separation angle at the generator level, $R_{sep}^{95\%}$, 0.4, and 0.2 from top to bottom, respectively. 107

5.3 Kinematical dependence for the leading jet p_T with the various ΔR (= 0.2, 0.4, $R_{sep}^{95\%}$) cuts at the generator level. 108
5.4 Leading jet p_T distribution after the correction of the matching efficiency. The inclusive leading jet p_T distribution without the parton-jet matching requirement is also superimposed. 108
5.5 An optimization of the matching efficiency to various ΔR cuts in the 3-Dimention. 110
5.6 An optimization of the matching efficiency to various ΔR cuts in the W + 2 jets and 3 jets processes, respectively. 110

6.1 Deviation of the observed jets p_T from the matched parton p_Ts in the current jet energy correction. The dependences of the different matching conditions of $\Delta R = 0.4$ and $R_{sep}^{95\%}$ are also shown. 115
6.2 Jet transverse energy. From the up-most side, the distributions are the highest E_T in $W + \geq 1$ jets events, the second highest E_T in $W + \geq 2$ jets events, and so forth. 118
6.3 Dijet mass distribution and jet separation angle between the highest E_T jet and the second highest E_T jet in $W + \geq 2$ jets events and $W + \geq 3$ jets events, respectively. 120
6.4 Jet multiplicity distribution. The errors on the data points are the sum of the statistical and systematic uncertainty by the jet E_T scale. The lower and upper band on the LO QCD predictions correspond to the energy scale of M_W^2 and $\langle\ p_T\ \rangle^2$, respectively. 122
6.5 Fraction of the jet multiplicity. The number of events in each n jet bin is normalized by the number of events in the $W + \geq 0$ jets events. 123
6.6 Ratio of jet multiplicity. From the top, the ratio of theory to data, the ratio of n jets events to n-1 jets events, and the ratio to the ratio of n jets events to n-1 jets events are shown. 125
6.7 Fraction of jets with a comparison to Run I measurement. 126

7.1 Schematic picture of the jet separation method for the kinematical cuts on the analysis and generator level. From the top-left, the phase space for the parton kinematics by p_T, η, and $\Delta R_{p,p}$ is shown, respectively. 137

7.2 Variation of the cross section for the W + n partons processes as a function of various energy scales μ_R and μ_F. CTEQ6M is used for the estimations. . 142

7.3 Comparison of the largest p_T distribution in the ME(W+1p)+PS and ME(W+0p)+PS processes in PYTHIA. 147

7.4 Comparison of the largest p_T distribution in the ME(W+1p)+PS and ME(W+0p)+PS processes in HERWIG. 147

A.1 Picture showing the structure of GR@PPA. The data flow is indicated schematically. The main building blocks of the GRACE-based event generator are BASES/SPRING and "GRACE output codes". The interface between them has been extended for pp and $p\bar{p}$ collisions. A set of random numbers $\{u_i\}$ given by BASES/SPRING includes two numbers for defining the initial state (x_1 and x_2), in addition to those for defining the final state. The cross section is calculated from the matrix element returned from the "GRACE output codes", by referring to a PDF using x_1 and x_2. Some phase-space cuts are applied by limiting the range of kinematical variables, or by setting the cross section to zero after the kinematical variables are determined. In some cases, several subprocesses are combined to a single subprocess by adding one more random number for defining the quark flavor. This system is interfaced to a generic user common of Les Houches Accord through `UPINIT` and `UPEVNT`. 158

A.2 The invariant mass distribution of leading two b quarks in the four-b events generated by GR@PPA. The distribution is shown separately for each subprocess. 181

A.3 Spin correlation for each charged lepton from $W^{\pm}$ bosons in the $q\bar{q}$ initiative process of top pair production, $q\bar{q} \to t\bar{t} \to b\ell^+\nu\ \bar{b}\ell^-\bar{\nu}$. 183

A.4 Spin correlation for each charged lepton from $W^{\pm}$ bosons in the gg initiative process of top pair production, $q\bar{q} \to t\bar{t} \to b\ell^{+}\nu\ \bar{b}\ell^{-}\bar{\nu}$. 183

List of Tables

1.1 Feynman rules for QCD in a covariant gauge for gluons (curly lines), fermions (solid lines) and ghosts (dashed lines). 6

2.1 The antiproton intensities given are merely examples. Higher antiproton intensities yield proportionally higher luminosities. The initial Run II upgrades are expected to have the ultimate potential to achieve luminosities of 2×10^{32} with 36 antiproton bunch operation. The typical luminosity at the beginning of a store has traditionally translated to integrated luminosity with a 33% duty factor. Operation with antiproton recycling may be somewhat different. 24

2.2 Characteristics of the CDF II calorimeter. 36

2.3 Design parameters of the CDF II Muon Detectors. Pion interaction lengths and multiple scattering are computed at a reference angle of $\theta = 90°$ in CMU and CMP/CSP, at an angle of $\theta = 55°$ in CMX/CSX, and show the range of values for the IMU. 38

3.1 Systematic uncertainties in the luminosity calculation using the CLC and the CDF measurement of the inelastic cross section. 60

3.2 Number of events in each event selection for the W $\rightarrow e\nu$. 61

3.3 The requirements for the high p_T electron trigger and those backup triggers. 65

3.4 Electron identification efficiency for each identification criterion. The number in the bracket denotes the number for the same signed electrons. 67

3.5 QCD background estimates in the W sample. First three columns show the number of events in each region A, B, and C, respectively. The number of QCD events are shown in the last column. 70

4.1 The fitting results of two jets acceptance. The transverse momentum of the first parton is always larger than one of the second parton. The values in third to fifth column show the separation angle at 5%, 50%, and 95% level with respect to the upper asymptotic limit of the fitting function. . . 89

4.2 The weighted mean and deviation for the corrected jet energy and the observed separation angle. 92

4.3 Some properties of the JetClu algorithm. Note that the jet energy correction is not applied. 98

4.4 Some properties of the MidPoint algorithm. Note that the jet energy correction is not applied. 99

4.5 Some properties of the SeedLess algorithm. Note that the jet energy correction is not applied. 99

6.1 Number of events in each jet bin. The ±10% fluctuation of the jet energy correction is taken as the systematic uncertainty. The upper- and lower-subscripts presents the systematics of $R^{95\%}_{sep}$-basis merging/splitting method. 116

7.1 All the subprocesses in the W + n partons (n $\leq$ 3) processes of the GR@PPA. The first column is the number of partons associated with the W boson. The second is the coupling order, where α_{em} and α_s are the electroweak and strong coupling constants, respectively. The third is the subprocesses considered in the ME calculations. ${q_u}^i({q_d}^i)$ is an up(down)-type quark in the i-th generation, where the i is integer number 1,2 for up-type quarks and 1,2,3 for down-type quarks, and g is a gluon. The numbers of Feynman diagrams in the ME calculation to each subprocess under unitary gauge can be seen in the forth column. Because the diagram structure is symmetrical for the z-direction (beam axis), actual number of diagrams should be multiplied by 2. In the fifth column, the weighting fraction for the parton flavor associated with the W boson is related to the quark mixing of CKM parameter, and the subprocess number to run the GR@PPA is appeared in the last column. 132

7.2 Production cross sections of W + n partons processes. The cross sections are estimated with two different energy scales with respect to two different ΔR cuts. The energy scales used are M_W^2 and $\langle\ p_T\ \rangle^2$ (GeV^2), where the renormalization and factorization scales are equivalent. The $\langle\ p_T\ \rangle^2$ is a squared average of parton p_T's. The first column is the number of partons, the second is the subprocess identification number used in GR@PPA listed in Table 7.1, and the others are the estimated cross sections in unit of pb. Although our kinematical cut of the $\Delta R_{p,p}$ at generator level is the $R_{sep}^{95\%}$, it is practical to show the cross sections in the case of the $\Delta R_{p,p} > 0.4$. The other kinematical cuts are the same as those of Eq. (7.8). 139

7.3 The total cross sections used in this analysis. All processes are produced by GR@PPA except W + 0 parton process of HERWIG built-in process. In the W + 0 parton process, the energy scale was taken as $\hat{s}$ in colliding partons. 140

7.4 except W + 0 parton process. 143

7.5 Fractions of c-quark and b-quark in W + n partons processes. Two energy scales of $\mu_R = \mu_F = M_W^2$ and $\langle\ p_T\ \rangle^2$ are considered. The fraction by the PDF uncertainty is also presented as the maximum/minimum deviation founded in a certain PDF set. The kinematical cuts at generator level are the same as Eq. (7.11). For the calculation, CTEQ6M was used as well as the uncertainty study of PDF. 144

A.1 All subprocesses produced by GR@PPA_All. The $p(\bar{p})$ represents a proton(antiproton) composed of u, d, c, s, b-quarks, and gluon. The j is a jet composed of u, d, c, s, b-quarks, and gluon as well as the p. The $f(\bar{f})$ is fermions (quarks and leptons) allowed to decay from vector bosons. The subprocesses are classified by a unique identification number `ISUB` with different coupling orders. The diagram numbers in brackets of the diboson processes are including the cascade decay type from the decaying fermions f. 164

A.2 Input parameters used in GR@PPA_All. All parameters are set in the subroutine `GRCPAR` or `UPINIT`. 170

A.3 Masses and widths set in the subroutine `SETMAS`. 171

A.4 Fundamental constant set in the subroutine `SETMAS`. 171

Chapter 1

Introduction

1.1 Overview

Understanding jet production mechanism is an interesting topic, because most of the searches for new particles like the Higgs boson or SUSY particles involve final states containing one or more than two jets in that signature of the final state. In the top quark physics, which is the main physics program at Tevatron Run II, an analysis of jets is the most crucial part to measure the top mass or the top production cross section. The W boson plus jets process thus has been an important analysis process in hadron colliders not only for the dominant background of the most of precision measurements but also for the probing sample for new physics, due to an easy triggering of a high transverse momentum of the leptons from their bosons.

In this study, we present a comparison with data and theory calculated by the lowest order perturbative calculation. To avoid a theoretical ambiguity of the collinear/infrared enhancement at the lowest order calculation, we apply a parton-jet matching procedure by requiring the clear definition of the parton separation, where each parton distributes within the particular cone size of a jet with an assumption that the doubly counted phase space will happen presumably in the collinear region, as well as the merging/splitting procedure of the cone jet algorithm. According to this procedure, one gets unique results which do not depend on unphysical bias of kinematical cuts to avoide the collinear/infrared divergence in calculation. Namely, one can get the meaningful results in the lowest order prediction. Besides, this prescription arises a sensitivity to the higher order calculation by comparing the lowest order prediction and the real data, since the phase space dominated by the lowest order perturbative calculation is constructed by our parton-jet matching procedure. The controllable event samples of the W boson plus jets events by the enhanced lowest order prediction will lead smaller systematic uncertainty than the naive prediction without any cares of the collinear/infrared divergence. We expect our method will be also useful to make systematically small samples as the background estimates in the top quark analysis. This is the first result for the CDF Run II experiment.

The CDF is successfully taking the collision data since 2002. The data used in this analysis correspond to a total integrated luminosity of 72.0 pb^{-1} taken from March 2002 through January 2003. High-p_T electron triggered samples are used. After good identification cuts on an isolated high-p_T electron and a requirement of an imbalance of calorimeter energy due to the undetected neutrino (missing E_T), a fixed cone jet algorithm is used to define jets. Our purpose of this study is to test multi-body QCD physics by comparing the data with the theory.

The following sections describe the physics backborn of the Standard Model and QCD aspects briefly in Sections 2 and 3. The current problematic feature of QCD dynamics is discussed in Sections 4 and 5.

1.2 The Standard Model

The Standard Model is a quantum field theory which postulates that all matters are composed of quarks and leptons and that the fundamental forces of nature are mediated by gauge bosons. To unify the electromagnetic and weak forces and account for all observed fermion and gauge boson masses, while still keeping the theory renormalizable, the Standard Model effectively introduces a new hypothetical particle, Higgs boson. During the last decades, despite the remarkable success of the Standard Model in high energy physics, nothing is known about the source of its fundamental theoretical basis, the Higgs mechanism. Most of high energy physics programs in current and future colliders are aiming at this discovery.

Leptons and quarks interact with each other via the exchange of gauge bosons. There are four fundamental forces among them: the strong, electromagnetic, weak and gravitational forces. The strong force is attractive force but in the very short distance, this power becomes rather weak. It is responsible for the binding of protons and neutrons to form atomic nuclei. The electromagnetic force works on two charged particles. It is attractive (repulsive) force if the particles have the opposite (same) signs of electric charge. It is responsible for the binding of electrons to nuclei to form atoms. The electromagnetic force is weaker than the strong force by approximately two orders of magnitude, which explains why protons bind to form nuclei despite the electromagnetic forces of repulsion between them. The weak force works in particle decays for such processes as neutron decay. The gravitational force is purely attractive between two massive particles. Since the gravitational force is much weaker than any other forces in the high energy elementally particles, this effects are usually ignored in the Standard Model framework in high energy collider experiments, while it becomes significant for bodies with astronomical mass like our earth or galaxy.

Leptons are one of the three classes of particles in the Standard Model. They are spin-1/2 fermions without strong interactions. There are six known leptons with three generations [1] like,

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix} \quad , \quad \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix} \quad , \quad \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix} \quad . \tag{1.1}$$

They are produced in pair in the same doublets. Note that the charge conjugate one also exists. Like the leptons, there are six types(flavors) of quarks denoted as

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

Each generation consists of a quark with charge $+\frac{2}{3}$ (u, c, t) together with a quark of charge $-\frac{1}{3}$ (d, s, b), in unit of an electric charge e. Besides of the fractional charges, quarks has color charges denoted as R, G, and B. Due to this color charges, free quarks do not appear in the final state which will be described in the latter section. The fundamental forces between fermions are carried by the bosons,

$$g\ (gluon) \quad , \quad \gamma\ (photon) \quad , \quad \begin{pmatrix} W^+ \\ Z^0 \\ W^- \end{pmatrix} \quad . \tag{1.3}$$

Each corresponds to the strong, electromagnetic, and weak force, respectively.

According to the Standard Model, quarks and leptons obtain their masses by the Higgs mechanism, – the spontaneously broken SU(2) $\times$ U(1) electroweak sector. As the result of the symmetry breaking, the Higgs boson, which is the only remained undiscovered piece of the Standard Model, appears. On the other hand, The mass eigen state can be redefined by the complex mass matrix by introducing the Cabibbo-Kobayashi-Maskawa (CKM) matrix [2] which leads to the CP-violating interactions. In the framework of the minimal Standard Model with massless neutrinos, it happens only in the quark sector. Many of experiments have confirmed this evidence. However, even in the lepton sector, the recent neutrino experiment from K2K [3] observes the neutrino oscillation, which is well explained by two-flavor ν_μ - ν_τ mixing. This may be an indication beyond the Standard Model.

1.3 Quantum Chromodynamics (QCD)

1.3.1 Lagrangian

Quantum Chromodynamics (QCD) is the gauge field theory for the strong interactions of colored quarks and gluons, which is a part of the SU(3)$\times$SU(2)$\times$U(1) Standard Model. A quark has three color charges and a gluon has eight color charges. Hadrons however are color-singlet state of quarks, anti-quarks, and gluons which will be described in the latter section. The QCD Lagrangian for the strong interactions of quarks and gluons is

$$L_{QCD} \;=\; -\frac{1}{4}F^{(a)}_{\mu\nu}F^{\mu\nu}_{(a)} \;+\; i\sum_q \bar{\psi}^i_q \gamma^\mu (D_\mu)_{ij}\psi^j_q \;-\; \sum_q m_q \bar{\psi}^i_q \psi_{qi} \;+\; L_{gauge-fixing} \;+\; L_{ghost} \quad , \tag{1.4}$$

$$F_{\mu\nu}^{(a)} = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a - g_s f_{abc} A_\mu^b A_\nu^c \quad , \tag{1.5}$$

$$(D_\mu)_{ij} = \partial_\mu \delta_{ij} + i g_s A_\mu^a (T_{i,j}^a) \quad , \tag{1.6}$$

where $F_{\mu\nu}^{(a)}$ is the field strength tensor for spin-1 gluon field A_μ^a, g_s is the QCD coupling constant, and the f_{abc} (a,b,c = 1,...,N_c^2-1) are the structure constants of the SU(N_c) algebra, where N_c is the number of colors, $N_c = 3$. Capital indices a,b,c run over 8 color degrees of freedom of the gluon field. The third term in Eq.(1.5) gives rise to triplet and quartic gluon self-interactions which distinguishes QCD from QED. The idea of the asymptotic freedom comes from this term. ψ_q^i (i = 1,2,3) are quark fields in triplet color representation. $(D_\mu)_{ij}$ is covariant derivative with the commutation relations of $N_c \times N_c$ matrices $T_{i,j}^a$,

$$[T^a \, , \, T^b] = i f_{abc} T^c \quad . \tag{1.7}$$

The gauge-fixing term $L_{gauge-fixing}$ is introduced to define the gluon propagator. The major choice of gauge is

$$L_{gauge-fixing} = -\frac{\lambda}{2} \sum_a (\partial_\mu A^{\mu a})^2 \quad . \tag{1.8}$$

This is the set of "covariant gauges". The $\lambda = 1$ corresponds to the "Feynman gauge (unitary gauge)". In the case of $\lambda \neq 1$, the ghost field L_{ghost} appears

$$L_{ghost} = (\partial_\mu \bar{c}_a)(\partial^\mu \delta_{ad} - g_s f_{abd} A_b^\mu) c_d \quad , \tag{1.9}$$

where c_a and $\bar{c}_a$ are scalar ghost and antighost fields. The ghost fields ensure that the gauge fixing does not spoil the unitarity of the "physical" S-matrix that governs the scattering of quarks and gluons in perturbation theory.

1.3.2 Feynman rule

Using the each term of QCD Lagrangian, we can obtain inverse quark and gluon propagators, and vertex factors in Feynman rules for perturbative QCD. We summarize Feynman rules in Table 1.1. Although for convenience, Feynman gauge ($\lambda = 1$) is usually used

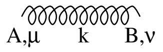
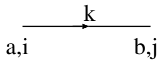
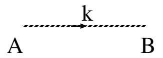
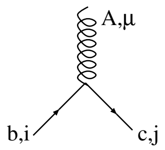
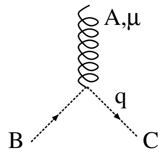
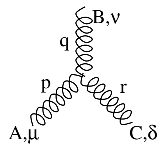
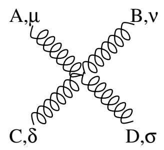
Propagator :	
	$i\frac{\delta_{AB}}{k^2+i\varepsilon}[-g^{\mu\nu} + (1-\lambda)\frac{k^\mu k^\nu}{k^2+i\varepsilon}]$
	$i\frac{\delta_{ab}}{(\not{k}-m+i\varepsilon)_{ji}}$
	$i\frac{\delta_{AB}}{k^2+i\varepsilon}$
Vertex :	
	
$-ig_s(T^A)_{cb}(\gamma^\mu)_{ji}$	$g_s f^{ABC} q^\mu$
	
$-g_s f^{ABC}[(p-q)^\delta g^{\mu\nu} + (q-r)^\mu g^{\nu\delta} + (r-p)^\nu g^{\delta\mu}]$	$-g_s^2 f^{XAC} f^{XBD}[g^{\mu\nu} g^{\delta\sigma} - g^{\mu\sigma} g^{\nu\delta}]$
all momenta incoming, $p+q+r=0$)	$-g_s^2 f^{XAD} f^{XBC}[g^{\mu\nu} g^{\delta\sigma} - g^{\mu\delta} g^{\nu\sigma}]$
	$-g_s^2 f^{XAB} f^{XCD}[g^{\mu\delta} g^{\nu\sigma} - g^{\mu\sigma} g^{\nu\delta}]$

Table 1.1: Feynman rules for QCD in a covariant gauge for gluons (curly lines), fermions (solid lines) and ghosts (dashed lines).

in the perturbative calculation, the covariant gauge is also practical to check the gauge invariance. Note that the ghost fields shown by the dashed lines in Table 1.1 finally cancel out in perturbative calculation at the end.

1.3.3 Asymptotic freedom

The basic idea of the asymptotic freedom is that the emission of virtual gluons by static color sources causes their color charges to “leak out” into the surrounding vacuum. The interaction between distributed charges thus becomes to be weaker at short distances. This is in contrast to QED of which the idea is that the observed charge of electron is smaller at long distances because of screening of its electric charge by vacuum polarization.

And at the end of the long distance in the gluon field, the force becomes much stronger (confinement). At some distance, it becomes easier to make new quarks and antiquarks, which combine to form hadrons, than to keep pulling against the ever-increasing force. That indicates isolated quarks are never observed in our experimental level. Indeed, there is no evidence to find a single quark in scattering experiments. Note that the asymptotic freedom is described by the perturbative QCD, while the confinement is not a result of the perturbation.

1.3.4 Running coupling

According to the idea of the asymptotic freedom and confinement, a perturbation series in QCD is regularized into a dimensionless physical quantity $\alpha_s(Q^2)$, where the $\alpha_s(Q^2)$ decreases with increasing of Q^2. A cut-off mass scale Λ^2 is required to remove ultraviolet divergences at low Q^2 region which is not able to apply a perturbation theory (confinement). We thus can safely use the perturbation theory at high Q^2 (asymptotic freedom). The renormalized coupling α_s depends on Q^2 and Λ^2. When the bare coupling at $Q^2 = \Lambda^2$ is hold by defining a point to subtract the divergences (momentum subtraction scheme)[1], the renormalized coupling $\alpha_s(Q^2)$ is expressed by the renormalization group equation (RGE):

$$Q^2 \frac{\partial \alpha_s}{\partial Q^2} = \beta(\alpha_s) = -b_0 \alpha_s^2 - b_1 \alpha_s^3 + \cdots \quad , \tag{1.10}$$

where the first two perturbative coefficients in the β-function [4] are

$$b_0 = \frac{33 - 2N_f}{12\pi} \quad , \quad b_1 = \frac{153 - 19N_f}{24\pi^2} \quad . \tag{1.11}$$

The N_f is the number of active quark flavors of which mass is much smaller than the scale Q. The third and fourth coefficients b_2 and b_3 in the β-function are also known. The exact analytical solution can be only solved at the lowest order term as

$$\alpha_s(Q^2) = \frac{1}{b_0 \ln(Q^2/\Lambda^2_{QCD})} \tag{1.12}$$

[1] When all the relevant one(and more)-loop diagram in quark-line are summed up, the momentum transfer of the external quark-line is vanished out, and the $\Lambda = \Lambda_{QCD} = M(p_0)$ with some fixed set of external momenta p_0.

where as already mentioned, Λ_{QCD} represents the break-down scale of the perturbative theory which rescues the divergence from logarithmic behavior. That scale typically is a few MeV order decided by experiments.

There is no analytical solution for the higher order RGE. However, the strong coupling $\alpha_s(Q^2)$ can be precisely defined for the first time since the scale dependence is suppressed over the wide range of Q^2. Approximate solutions can be calculated by depending on the different renormalization schemes. The most popular choice of that schemes is to use the $\overline{MS}$-scheme [5]. The approximate solution of the two loop evolution to define Λ_{QCD} could be

$$\alpha_s(Q^2) = \frac{1}{b_0 \ln(Q^2/\Lambda^2_{\overline{MS}})}\Big[1 - \frac{b_1 \ln(Q^2/\Lambda^2_{\overline{MS}})}{b_0 \ln(Q^2/\Lambda^2_{\overline{MS}})} + \mathcal{O}\Big(\frac{\ln^2[\ln(Q^2/\Lambda^2_{\overline{MS}})]}{\ln^2(Q^2/\Lambda^2_{\overline{MS}})}\Big)\Big] \quad . \tag{1.13}$$

The constant $\Lambda_{\overline{MS}}$ is also one fundamental constant of QCD that must be determined from experiments.

Current best fit value [6] of α_s at mass of Z is

$$\alpha_s(M_Z) \;=\; 0.1172 \;\pm\; 0.0020 \quad . \tag{1.14}$$

1.3.5 Parton distribution function

Using similar perturbative expansion with the β-function (1.10), the parton distribution functions $f_a^h(x, Q^2)$ can be brought out by DGLAP (Dokshitzer-Gribov-Lipatov-Altarelli-Parisi) evolution equation [7],

$$Q^2 \frac{df_a^h(x, Q^2)}{dQ^2} = \sum_b \int_x^1 \frac{dz}{z} P_{ab}(\alpha_s(Q^2), z) f_a^h(x, Q^2) \quad , \tag{1.15}$$

at any fixed scale Q. The functions $P_{ab}(\alpha_s(Q^2), z)$ are the Altarelli-Parisi (AP) splitting functions which outputs the probabilities of finding a parton a in a parton b with a fraction z of the longitudinal momentum of the parent parton. Having determined $f_a^h(x, Q^2)$ at a given input scale $Q^2 = Q_0^2$, the evolution equation can be used to compute the PDF's at different perturbative scales Q^2 and larger value of x. They do not depend on the colliding hadron h and thus they are process independent. The AP splitting functions are also expanded by a power series of α_s :

$$P_{ab}(\alpha_s, z) = \alpha_s P_{ab}^{(LO)}(z) + \alpha_s^2 P_{ab}^{(NLO)}(z) + \mathcal{O}(\alpha_s^3) \quad . \tag{1.16}$$

That behavior of the longitudinal momentum z is also logarithmic. The solution of LO and NLO can be found in Ref. [8].

Most representative measurement of PDF is a deep-inelastic lepton-hadron scattering (DIS) since the Q^2 evolution is independent of the unmeasured gluon distribution. Most of the PDF sets are available up to the next-to-leading order (NLL) evolutions of the AP equations with the DIS or $\overline{MS}$ schemes.

1.3.6 Nonperturbative QCD

Corresponding to asymptotic freedom at high momentum scales (short distances), we have infrared slavery: $\alpha_s(Q^2)$ becomes large at low momenta (long distances). Perturbative theory does not reliable for large α_s, so that nonperturbative method must be used. The important low momentum-scale phenomena are the confinement and hadronization. Confinement is the result of (long-distance) property of QCD, which can be treated by lattice techniques. Partons found only in color-singlet bound states (hadrons), size ~ 1 fm, are forming hadrons. If we try to separate those partons in color-singlet state, it becomes energetically favorable to create extra partons, forming additional hadrons. Hadronization is that partons produced in short-distance interactions reorganize themselves (and multiply) to make observed hadrons. Note that the hadronization is a dynamical (long timescale) phenomenon of which nonperturbative models are only available at present.

1.3.7 Infrared divergences

Even in high-energy of the short-distance regime, long-distance aspects of QCD cannot be ignored. Soft or collinear gluon emission gives infrared divergences in perturbative QCD. Light quarks ($m_q \ll \Lambda_{QCD}$) also lead to divergences in the limit $m_q \to 0$ (mass singularities). An example for the soft/collinear singularity is illustrated in Figure 1.1.

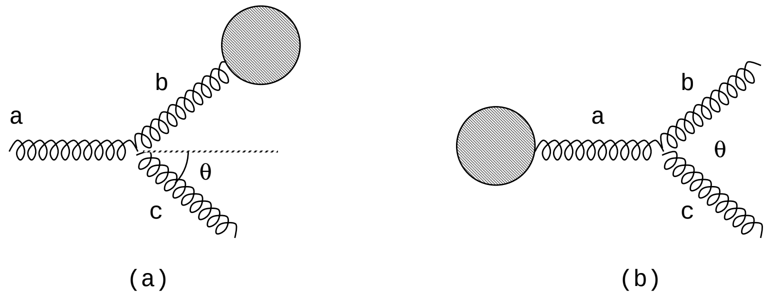


Figure 1.1: Example for the soft/collinear singularity. (a) is an example for spacelike branch, and (b) is an example for timelike branch.

When a gluon splits from the incoming line (a) (spacelike branching), the propagator factor $1/p_b^2$ is expressed as

$$p_b^2 = -E_a E_c (1 - \cos\theta) \leq 0 \quad . \tag{1.17}$$

The propagator factor $1/p_b^2$ diverges as $E_c \to 0$ (soft singularity) or $\theta \to 0$ (collinear or mass singularity). In the case of incoming quark, the quark mass suppresses some divergences, but in the high-energy region which quark masses is almost negligible, those divergence still remain.

When a gluon splits from outgoing line (b) (timelike branching), the propagator factor $1/p_a^2$ is expressed as

$$p_a^2 = E_b E_c (1 - \cos\theta) \geq 0 \quad , \tag{1.18}$$

where the divergences appear when either emitted gluon is soft (E_b or $E_c \to 0$) or when opening angle $\theta \to 0$.

Infrared divergences also appear in the integration of loop diagrams of virtual partons. The infrared divergences indicate dependence on long-distance aspect of QCD not correctly described by pertubative theory. The divergent propagators imply propagation of partons over long distances. When the distance becomes comparable with hadron size ~ 1 fm, quasi-free partons of perturbative calculation are confined/hadronized non-perturbatively, and then apparent divergences disappear. Therefore, we put a limit to use the perturbative theory to perform calculations into two classes of observable.

Infrared safe : The infrared safe quantities should be insensitive to soft/collinear singularities. The infrared divergences in perturbative calculation either cancel between real and virtual contributions or are removed by kinematical factors.

Factorizable : The infrared sensitivity can be absorbed into an overall non-perturbative factor to be determined experimentally.

In either case, infrared divergences must be regularized during perturbative calculation, even though they cancel or factorize in the end.

1.3.8 Total cross section

Consider the physics process

$$h_1(p_1) \ + \ h_2(p_2) \ \rightarrow \ H(Q^2) \ + \ X \quad , \tag{1.19}$$

where $p_{1,2}$ is 4-momenta of the incoming hadrons, $h_{1,2}$, H denotes the observed hard-process triggered by the particle productions such as W/Z or heavy/light quark jets productions as well as Higgs bosons or SUSY particle productions, and the X stands for any unobserved particles produced by the collision. The Q^2 is a typical energy scale which specifies the property of the hard-interaction. The total cross section for this process (1.19) is expressed as

$$\sigma(p_1, p_2; Q^2) = \sum_{i,j,F} \int dx_1 dx_2 d\hat{\Phi}_F f_i^{h1}(x_1, Q^2) f_j^{h2}(x_2, Q^2) \frac{d\hat{\sigma}_{ij\rightarrow F}(\hat{s}; Q^2; \alpha_s(Q^2))}{d\hat{\Phi}_F} + \mathcal{O}((\Lambda_{QCD}^2/Q^2)^p) \quad . \tag{1.20}$$

where $f_i^{h1}(x_1, Q^2)$ is a PDF of the hadron h_1 (p or $\bar{p}$), which gives the probability to find the parton i ($= g, u, \bar{u}, d, \bar{d}, ...$) with an energy fraction x_1 at a probing virtuality of Q^2. The differential cross section $d\hat{\sigma}_{ij\rightarrow F}(\hat{s}; Q^2; \alpha_s(Q^2))/d\hat{\Phi}_F$ represents the parton-level hard-interaction producing the final-state F from a collision of partons, i and j, where $\hat{s}$ is the square of the total initial 4-momentum. The sum is taken over all relevant combinations of i, j and F. Note that in hadron interactions a certain "process" of interest may contain some incoherent subprocesses having different final states, as well as those having different combinations of the initial-state partons. For example, the "two-jet" production process includes all $q\bar{q}'$, $qg(\bar{q}g)$ and gg production processes. The term $\mathcal{O}((\Lambda^2_{QCD}/Q^2)^p)$ generically denotes non-perturbative contributions such as hadronization effects, multiple interactions, contributions of the soft underlying event and so forth.

The cross section of the hard-interaction $\hat{\sigma}_{ij\to F}(\hat{s};Q^2;\alpha_s(Q^2))$ is computable as a power series expansion in the strong coupling $\alpha_s(Q^2)$:

$$\hat{\sigma}_{ij\to F}(\hat{s};Q^2;\alpha_s(Q^2)) = \alpha_s^k(Q^2)\{\hat{\sigma}^{(LO)}_{ij\to F} + \alpha_s(Q^2)\hat{\sigma}^{(NLO)}_{ij\to F} + \alpha_s^2(Q^2)\hat{\sigma}^{(NNLO)}_{ij\to F} + \cdots\} \quad (1.21)$$

where the index k is a number of the order of the $\alpha_s(Q^2)$ in the lowest order (LO) term $\hat{\sigma}^{(LO)}_{ij\to F}$. The LO cross section only gives a crude estimation because of the presence of a magnitude of $\alpha_s(Q^2) \sim 0.1$. Most of the next-to-leading order calculation (NLO) are available. We should note that the factorization formula (1.20) is only applied when the measured cross section is "infrared safe". Namely, even if one high energy particle splits into the two particles moving in the same direction or emits a light particle carrying very small momentum, the measured cross section does not change. For instance, one measures a jet which is a bunch of particles collimated in the almost same area with a given total energy, while one does not measure a single high-p_T hadron as the results of the cross section formula (1.20). If a single high-p_T hadron is measured, the factorization formula has to include an additional convolution with the corresponding parton fragmentation function $d_i^{hadron}(z,Q^2)$. We assume that the hadronization of partons can be separated from the hard-interaction. This is an important consequence. The "jets" must be defined with an insensitive algorithm away from an infrared problem.

1.3.9 Factorization and renormalization scales

The factorization on the right-hand side of Eq.(1.20) involves some degree of arbitrariness that are known as factorization-scheme dependence. This dependence lead to some uncertainties since the perturbative corrections beyond a given order of the factorization scheme may still remain as the size of missing higher-order corrections. These uncertainties could be addressed for the uncertainty of PDF. Of course, physical quantities does not depend on any factorization/renormalization schemes. These schemes are just a calculation method in which the divergence such as collinear singularities are factored out as the PDF or coupling constants at a factorization or renormalization scale Q. One has to care the choice of those scheme when one calculates the cross section.

In addition, the scheme dependence leads to the additional uncertainties to a dependence on the renormalization and factorization scales. The partonic cross section from the hard-interaction depends on its perturbative computable process, while the PDF is the process-independent non-perturbative theory. The renormalization scale μ_R is the scale at which the strong coupling α_s is evaluated, while the factorization scale μ_F is the scale to separate the perturbative domain region of the hard-interaction from the non-perturbative region of PDF. As already mentioned, a "true" physical state should be laid on the same scale $Q = \mu_R = \mu_F$. However, their values cannot be unambiguously fixed. Therefore, the Eq.(1.20) is modified by explicitly introducing the dependence of μ_R and μ_F.

$$\sigma(p_1, p_2; Q^2) = \sum_{i,j,F} \int dx_1 dx_2 d\hat{\Phi}_F f_i^{h1}(x_1, \mu_F^2) f_j^{h2}(x_2, \mu_F^2) \frac{d\hat{\sigma}_{ij\to F}(\hat{s}; Q^2; \mu_F^2; \mu_R^2; \alpha_s(\mu_R))}{d\hat{\Phi}_F} + \mathcal{O}((\Lambda_{QCD}^2/\{Q^2; \mu_F^2; \mu_R^2\})^p) \quad . \tag{1.22}$$

The physical cross section $\sigma(p_1, p_2; Q^2)$ does not depend on the arbitrary scales μ_R, μ_F, but all the other terms in the right-hand side separately depend on these scales. The partonic cross sections from the hard-interaction have the μ_R, μ_F-dependence based on their perturbative expansions. The higher-order perturbative calculation may suppress the μ_R-dependence [9]. The μ_F-dependence of PDF's is also from DGLAP evolution equation. Namely, these μ_R, μ_F-dependences can be understood by a lack of truncated higher-order term. If all-order perturbative calculations are possible, the μ_R, μ_F-dependences disappear. The size of the μ_R, μ_F-dependences are often estimated as a measure of the size of at least some of the uncalculated higher-order terms by setting a lower limit on the size of missing higher-order corrections. Theoretical error caused by truncating the perturbative expansion thus rises up.

The μ_R, μ_F-dependences of the last term $\mathcal{O}((\Lambda^2_{QCD}/\{Q^2;\mu_F^2;\mu_R^2\})^p)$ in the right-hand side of Eq.(1.22) are a remnant from the results of the partonic cross sections and the evolutions of PDF. All invisible effects are included in this term and thus perturbative effects as the observable objects may be more or less disturbed.

1.4 Dynamics of multi-particle final state

Due to an event topology with the multi-particle final state after the QCD evolution and then confinement/hadronization to non-perturbative region, one have to rely on Monte Carlo method, not analytical method. There are two traditional approaches to model the perturbative interactions. One is the matrix-element (ME) method, in which Feynman diagrams are calculated order by order. Exact kinematics, the full interference and helicity structure are taken into account in this approach. However, from the evolution equation Eq.(1.15), the emission of multiple soft gluons plays a significant role in building up the event structure and the ME calculations become increasingly difficult in higher orders, in particular for the loop graphs. This sets a limit to the applicability of MEs. Since the phase space available for gluon emission increases with the available energy, the ME approach becomes less relevant for the full structure of events at higher energies. Thus, the ME approach yields more reliable results for the well-separated jets at high energy region since the perturbative expansion of the ME approach behaves better at higher energy scales based upon the running of α_s.

Another approach is the parton-shower (PS) one, in which the number of partons involved has no explicit upper limit. This method does not use the ME expressions, but uses the approximate expressions with the simplified kinematics, interference and helicity structure. The PS approach is thus expected to give a good approximation for the substructure of jets, but yield less reliable results for the well-separated jets described by power series of α_s in perturbative expansion.

Thus both approaches have been complementary used in many aspects in collider experiments. It is practical to describe both approaches by the term of perturbative expansion of only strong interaction,

$$
\begin{array}{cccccccccccc}
 & & & LO & NLO & & NNLO & & NNNLO & & \cdots & & \\
\sigma & \sim & \mathcal{O}_{tree} & \times \{ \, 1 + & \alpha_s L & + & \alpha_s^2 L^2 & + & \alpha_s^3 L^3 & + & \cdots & & LL \\
 & & & + & \alpha_s & + & \alpha_s^2 L^1 & + & \alpha_s^3 L^2 & + & \cdots & & NLL \\
 & & & & & + & \alpha_s^2 & + & \alpha_s^3 L^1 & + & \cdots & & NNLL \\
 & & & & & & & & \alpha_s^3 & & \cdots & \} & \cdots \quad ,
\end{array}
\tag{1.23}
$$

where the total cross section σ (at Q^2) is logarithmically expanded by L $\equiv \log(Q^2/\Lambda^2)$. Each column in the right hand side of Eq.(1.23) is the order of the strong coupling α_s by the ME calculations and each line is the logarithmic expansion described by the PS calculations. Current most of the ME approach are available at Next-to-leading-order level (NLO), where the tree level calculation is assumed to be the simple $2 \to 2$ process. For the Monte Carlo event generation in hadron colliders, the LO event generators are only used. For the PS approach, the LL ($\equiv \sum(\alpha_s L)^n$) PS is only available, where the logarithmic expansion is summed over all order terms. As already mentioned, the PS approach does not describe the full phase space for the well-separated jets. Thus the ME-based event generators based on the diagonal elements ordered by α_s^n in Eq.(1.23) have been developed instead of the full calculations of NLO level. Those ME calculations only describe the real emission part ordered by α_s^n. Here one theoretical problem is arisen when the LL PS is adopted into the α_s^n-LO ME calculations to make up the realistic event structure. From Eq.(1.23), the α_s^n-LO ME calculations already account for a part of the LL expansion. The overlapped phase space between the α_s^n-LO ME calculations and the LL PS appear. This is known as "double counting" problem. Indeed, the ME-PS matching corrections is discussed in [10] only for the LO level. Currently, so many studies are on-going. It is clear that the kinematical boundary is needed in the α_s^n-LO ME calculations to be safe for the soft/collinear divergences.

1.5 Test of QCD using W + jets events

The W bosons in the hadron-hadron collisions are presumably produced at high energy scale Q^2 around the mass of W boson. Then, the partons associated with the W boson will evolve from the fundamental scale Q_0^2 $(= \Lambda_{QCD}^2)$ on PDF to the virtuality scale at Q^2 along the initial parton line in PDF. When the observable objects in experimental analysis are the well-separated jets associated with W boson, the α_s^n-LO ME calculations embedded into LL PS are used to satisfy this requirement. This is the typical example to test the multi-body dynamics in QCD described in the previous section.

Our purpose is to compare the theory given by the α_s^n-LO ME + LL-PS with the real data. Since only the ME calculation gives a prediction of the total production cross section and the PS does not give any information of cross sections, we measure jets inclusively. Among small number of the coupling orders of α_s^n, the α_s^n-LO ME and LL PS themselves are used with the assumption that the overlapped phase space is presumably small. This is the first order approximation. We then try to construct the phase space with no overlapping phase space in the experimental analysis level. The strategy of this analysis thus is so far: we make each α_s^n-LO ME + LL PS event sample (n = 0,1,2,...), and collect the jets samples inclusively, that is, group the $W + \geq$ n jets event samples, where, for instance, an event which has 2 jets is a member of the $W + \geq 2$ jets event sample but at the same time it can be a member of the $W + \geq 1$ jets event sample,

$$\begin{array}{lcl} \sigma(W + \geq 0\, jets) & : & ME(\alpha_s^0) + LL\,PS \\ \sigma(W + \geq 1\, jets) & : & ME(\alpha_s^1) + LL\,PS \\ \sigma(W + \geq 2\, jets) & : & ME(\alpha_s^2) + LL\,PS \\ & & \cdots \qquad . \end{array}$$

The cross section formula is then expressed again,

$$\sigma(W + \geq n\, jets) \equiv F(parton : jet) \times \sigma(W + n\, partons) \qquad , \tag{1.24}$$

where $\sigma(W + \geq n\ jets)$ is the observed cross section, $\sigma(W + n\ partons)$ is the theoretical cross section supplied by the ME calculation, and $F(parton : jet)$ is a probability function of the transition from the partonic final state to the observable jets. This function includes all the effects of the fragmentation, hadronization, and the underlying activity in the hadron-hadron collisions as well as the parton showering. The definition of jets plays a crucial role to avoid an ambiguity of the $F(parton : jet)$. In the LO-ME level, the $F(parton : jet)$ is drastically affected by the collinear/infrared divergence in calculations rather than kinematic/geometric acceptance of jets.

The above is the same analysis topology used in Run I [11]. However, in Run I, we did not give a quantitative study for the treatment of the collinear/infrared divergence in the α_s^n-LO ME calculations. This study is aiming at giving a clear logic to treat them. That is, we consider a prescription to maximize the function $F(parton : jet)$ by looking at the parton-jet matching which will lead to minimize the systematic uncertainty and at the same time, to minimize the overlapped phase space between ME and PS. As the results, we can construct the phase space dominated by the α_s^n-LO ME. The comparison with the real data will arise a sensitivity to the higher order calculation. We expect our method will be also useful to make systematically small samples as the background estimates in the top quark analysis. This is the first result for the CDF Run II experiment.

Chapter 2

The CDF Experiment at the Tevatron Collider

2.1 Introduction

The CDF (Collier Detector at Fermilab) experiment is a general purpose experiment for the study of $p\bar{p}$ collisions at the Tevatron Collider located at the Fermi National Accelerator Laboratory (Fermilab), in Batavia, Illinois, U.S.. The Tevatron accelerator is the highest-energy proton-antiproton accelerator machine in the world. Using Tevatron accelerator, CDF discovered top-quark events in 1995. The upgrades to Run II experiment is expected to give an opportunity not only to discover a new physics but also to perform precision measurements in electroweak sector. In this chapter, various complex and functions of the Tevatron accelerator and CDF detector are described in brief. For a complete review see [12][13].

2.2 Characteristics of Fermilab accelerators

2.2.1 Accelerator chain

The Tevatron is a circular accelerator of about 1 km of radius. Each proton and antiproton beam is accelerated up to 980 GeV, respectively, then colliding in opposite directions with the centre-of-mass energy of 1.96 TeV. Before reaching the final colliding energy of 980

GeV, there are five acceleration steps: Cockcroft-Walton, Linac, Booster, Main Injector, and Tevatron. The schematic view of the accelerator complex at Fermilab is illustrated in Figure 2.1.

The protons begin as H^- ions which is produced in a bottle of hydrogen gas. The ions are then injected into a Cockcroft-Walton preaccelerator with an initial acceleration of 25 keV. The Cockcroft-Walton then accelerates them up to 750 keV. In the second step, these ions enter a linear accelerator (Linac) with 145 m long, then the ions reach the energy of 400 MeV. At the end of the Linac, the electrons from the H^- ions are stripped off by carbon foil, and the resulting protons are passed into the Booster ring. The Booster ring is a proton synchrotron ring with 75 m radius. The protons are circulated until acquiring 8 GeV, before being collected in bunches of 6×10^{10} particles each for use in the Tevatron (or 5×10^{12} for protons used for the production of antiprotons). The next step is a further acceleration in Main Injector, where protons are accelerated from 8 GeV to 150 GeV. The Main Injector is a proton synchrotron ring with 1 km radius. For the Tevatron Run I operation, the 8 GeV proton bunches were injected into the Main Ring synchrotron, where they were accelerated to 150 GeV. In Run II phase, the existing Main Ring has been replaced by a new accelerator, Main Injector, for the purpose of reducing the inefficiency of the antiproton production. As the result of this, more than factor $\sim$ 2 increases in luminosity are expected. The 150 GeV proton bunches are finally injected into the evacuated beam pipe of the Tevatron ring and brought to an energy of 980 GeV.

2.2.2 Production of antiproton

The antiprotons are created at a target station. The 120 GeV protons extracted from the Main Injector hit the nickel target and produce antiprotons. Antiprotons are produced over a large spread angles on the forward direction with energies of approximately 8 GeV. A cylindrical lithium lens is used to focus the antiprotons into a parallel beam. These parallel beam of antiprotons are then directed to the Debuncher. The Debuncher is a rounded triangular-shaped synchrotron with a mean radius of 90 m. Its primary

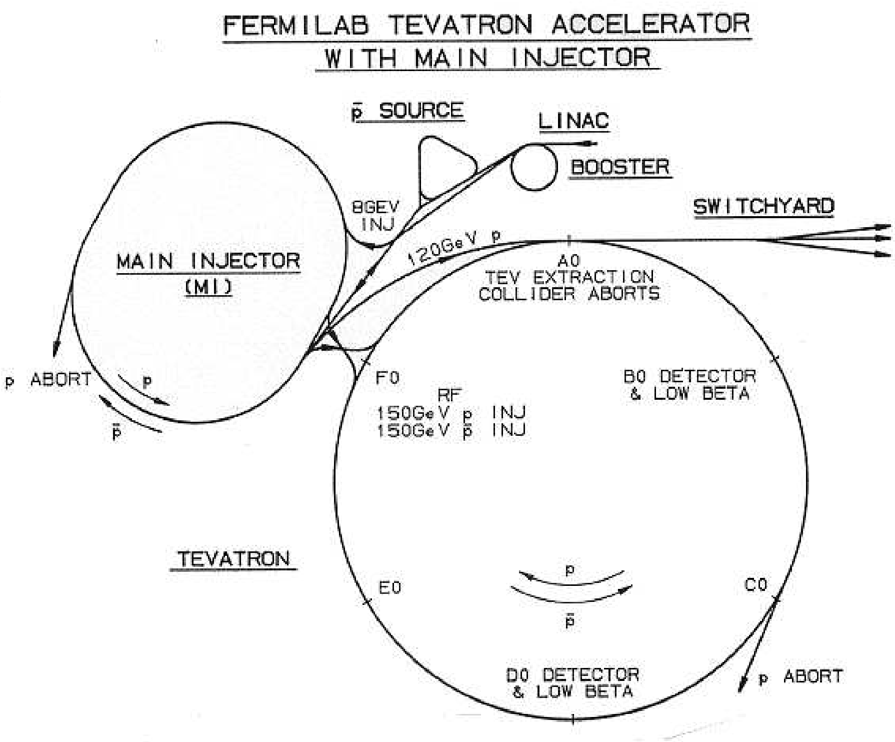


Figure 2.1: A schematic view of the Run II Tevatron accelerator complex at Fermilab. The new Main Injector is shown in the left side of the Tevatron Ring. The CDF detector is located at one of the collision points at B0 on the Tevatron Ring.

purpose is to reduce the momentum spread of the antiprotons by rotating the bunches. There are also beam cooling systems that act to reduce the oscillations in the plane perpendicular to the orbit (transverse plane) as well as reducing the momentum spread of the antiprotons. Then, the antiprotons are transferred to the Accumulator which is the storage ring for the antiprotons. All of the antiprotons are stored here at 8 GeV and cooled until needed. When the stack is large enough, bunches of antiprotons are transferred into the Tevatron ring and accelerated to 980 GeV. The total number of antiprotons in the collider is determined by the product of the antiproton production rate, the typical store duration, and the transmission efficiency from Accumulator to storage in the Tevatron.

In the process of the Tevaron Run II upgrade, a new Antiproton Recycle Ring has been proposed. The role of the Recycler Ring is to provide more antiprotons for the Tevatron, which proportionally doubly increases the luminosity. If the Recycler Ring is available, the antiprotons in the Tevatron at the end of the store will not be dumped but collected into the Recycler Ring, and then sent to the Main Injector and to the Tevatron again.

2.2.3 The Tevatron

The Tevatron is the only cryogenically cooled accelerator at Fermilab. It is a circular synchrotron of radius 1 km that has 8 accelerating cavities in the RF section of the accelerator. The Tevatron receives protons and antiprotons from Main Injector and accelerates them from 150 GeV to 980 GeV. In the Tevatron Collider mode[1], the beam is stored in the Tevatron for several hours. Tevatron can accelerate beam every $120 \sim 200$ seconds for Collider mode. Once 36 bunches of protons and 36 of antiprotons are obtained, the two beams are focused using quadrupole magnets. There are two collision points around the Tevatron ring, B0 and D0. The former is the collision point for the CDF experiment, and the later is for the D0 experiment, respectively.

[1]There also is the Fixed Target mode. In Fixed Target mode, beam is sent to Switchyard over a 20 second period. Then, Tevatron can accelerate beam every 60 seconds for Fixed Target mode.

2.2.4 The Tevatron luminosity

The Tevatron luminosity in the absence of a crossing angle or position offset is given by

$$\mathcal{L} \;=\; \frac{fBN_pN_{\bar{p}}}{2\pi(\sigma_p^2+\sigma_{\bar{p}}^2)}\,F(\sigma_l/\beta^*) \qquad (\mathrm{cm}^{-2}\mathrm{s}^{-1}) \quad , \tag{2.1}$$

where f is the revolution frequency, B is the number of bunches in each beam, $N_p(N_{\bar{p}})$ is the number of protons(antiprotons) in a bunch, $\sigma_p(\sigma_{\bar{p}})$ is the RMS proton(antiproton) beam size at the interaction point, and F is a from factor that depends on the ratio of the bunch length σ_l, to the beta function at the interaction point β^*. The luminosity can be rewritten in a form that more directly displays its dependences on the limiting factors within the Tevatron complex,

$$\mathcal{L} \;=\; \frac{3\gamma f}{\beta^*}(BN_{\bar{p}})(\frac{N_p}{\varepsilon_p})\,\frac{F(\sigma_l/\beta^*)}{(1+\frac{\varepsilon_{N_{\bar{p}}}}{\varepsilon_{N_p}})} \quad , \tag{2.2}$$

where $\varepsilon_{N_p}(\varepsilon_{N_{\bar{p}}})$ is the normalized transverse emittance containing 95% of the proton(antiproton) beam. The major luminosity limitations are dominated by the number of antiprotons $BN_{\bar{p}}$, and the proton beam brightness N_p/ε_p. Note that for a given total number of antiprotons, the luminosity does not depend explicitly on the number of bunches. The luminosity in the Tevatron is proportional to the total antiproton intensity. The collision frequency used for this analysis is every 396 ns for 36 $\times$ 36 bunch crossing rate. The 132 ns mode (140 $\times$ 105) is currently under development. The beam collisions continue typically for 8 hours. At the collision point at CDF detector, the typical beam size is 120 cm for the direction of the beam axis, and 30 μm for the perpendicular plane to the beam axis. The Tevatron Run II parameters are listed in Table.2.1 together with parameters on Run Ib experiment. In Figure 2.2, we show the total integrated luminosity delivered by the Tevatron and recorded by CDF in the period from July 2001 to July 2003. In this analysis, we use the data until January 2003 corresponding to 72.0 pb^{-1}.

RUN	Ib(1993-95) (6×6)	Run IIa (36×36)	Run IIa (140×103)	
Protons/bunch	2.3×10^{11}	2.7×10^{11}	2.7×10^{11}	
Antiprotons/bunch	5.5×10^{10}	3.0×10^{10}	4.0×10^{10}	
Total Antiprotons	3.3×10^{11}	1.1×10^{12}	4.2×10^{12}	
Pbar Production Rate	6.0×10^{10}	1.0×10^{11}	2.1×10^{11}	hr^{-1}
Proton emittance	23π	20π	20π	mm-mrad
Antiproton emittance	13π	15π	15π	mm-mrad
β^*	35	35	35	cm
Energy	900	980	980	GeV
Antiproton Bunches	6	36	103	
Bunch length (rms)	0.60	0.37	0.37	m
Crossing Angle	0	0	136	μrad
Typical Luminosity	0.16×10^{31}	$0.1\text{-}0.8 \times 10^{32}$	2.1×10^{32}	$\mathrm{cm}^{-2}\mathrm{sec}^{-1}$
Integrated Luminosity	3.2	8-15	42	pb^{-1}/week
Bunch Spacing	$\sim$3500	396	132	nsec
Interactions/crossing	2.5	2.3	1.9	

Table 2.1: The antiproton intensities given are merely examples. Higher antiproton intensities yield proportionally higher luminosities. The initial Run II upgrades are expected to have the ultimate potential to achieve luminosities of 2×10^{32} with 36 antiproton bunch operation. The typical luminosity at the beginning of a store has traditionally translated to integrated luminosity with a 33% duty factor. Operation with antiproton recycling may be somewhat different.

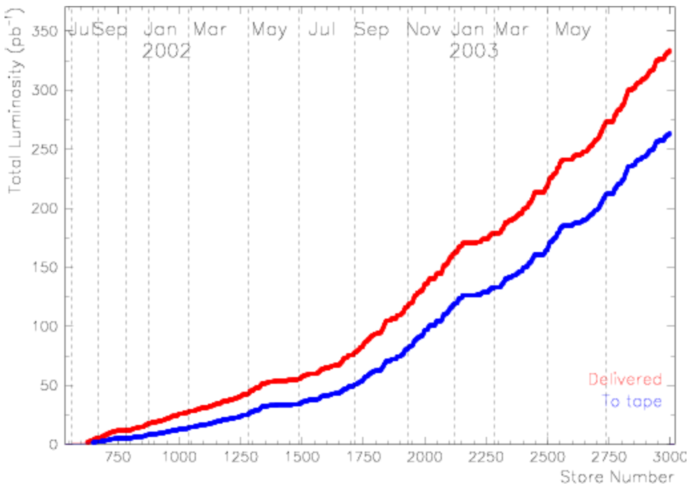


Figure 2.2: Total integrated luminosity delivered by the Tevatron and recorded by CDF in the period from July 2001 to July 2003. In this analysis, we use the data until January 2003 corresponding to 72.0 pb^{-1}.

2.3 The CDF II detector

2.3.1 Overview

CDF II detector is a general purpose solenoidal detector with a goal of the studies for a broad range of high energy physics program. The detector is constructed with the charged particle tracking with fast projective calorimetry and fine grained muon detection.

A schematic view of the CDF II detector is shown in Figure 2.3. Tracking systems are contained in a superconducting solenoid, 1.5 m in radius and 4.8 m in length, which generates a 1.4 T magnetic field parallel to the beam axis. Calorimetry and muon systems are all outside the solenoid. We use a coordinate system where the polar-angle θ is measured from the proton direction, the azimuthal angle ϕ is measured around the beam direction, and the pseudo-rapidity is defined as $\eta = -\ln(\tan(\theta/2))$.

There are some detector upgrades from Run I experiment. The tracking systems have been totally replaced with new detectors. The calorimetry systems are now exclusively scintillator-based. The electronics and trigger systems are fully compliant with the new pipelined configuration, and all the software has been re-written using C++ and an Object Oriented architecture.

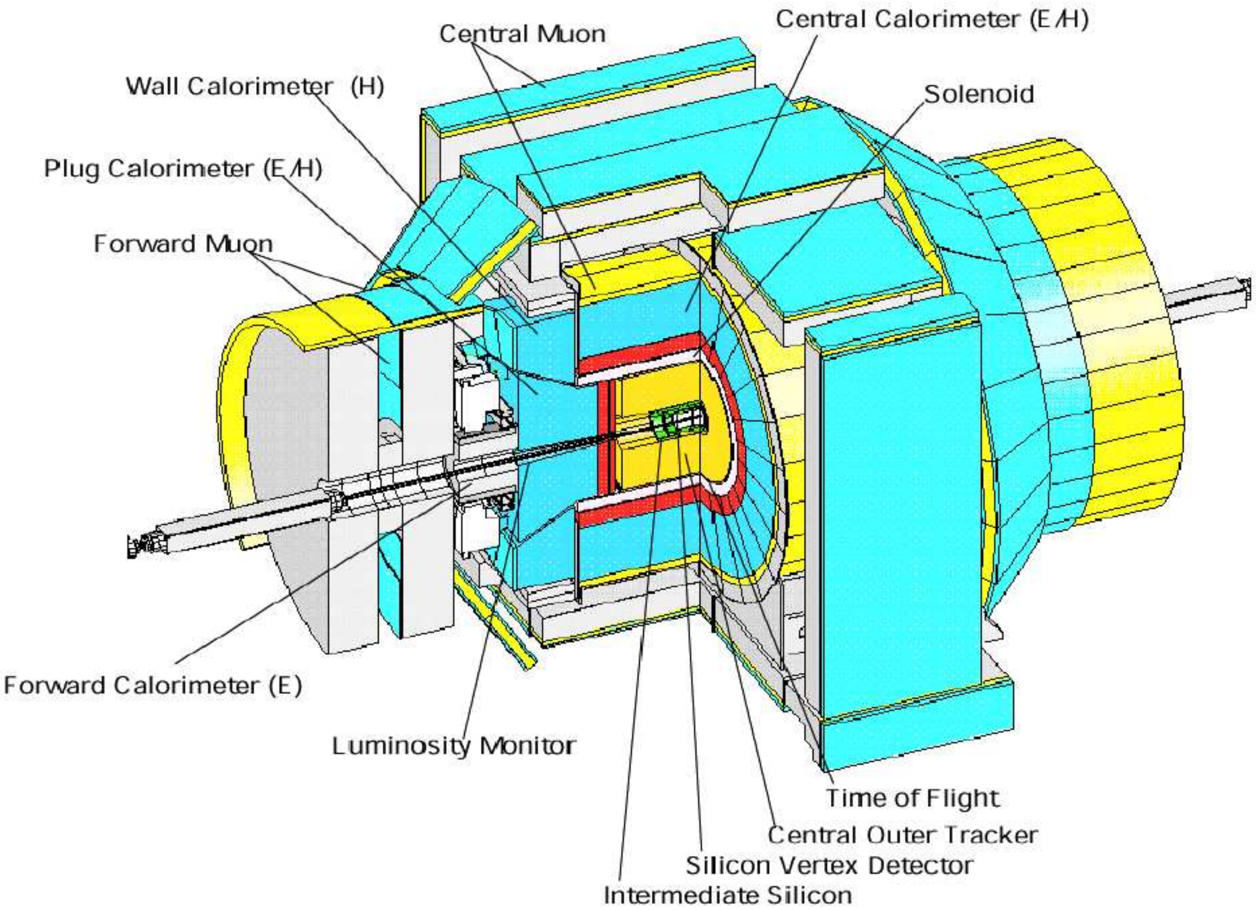


Figure 2.3: A schematic view of the CDF II detector.

2.3.2 Tracking systems

There are two primary tracking detector systems in CDF II: the inner tracking system of the silicon vertex detectors (L00, SVX II, and ISL), and the central outer tracking system of the wire-based drift chamber (COT). The schematic view of the tracking systems is illustrated in Figure 2.4. From the innermost side, Layer00 (L00), Silicon Vertex Detector (SVX II), and Intermediate Silicon Layers (ISL) detectors construct with the inner tracking system, and Central Outer Tracker (COT) constructs with the outer tracking system. The COT covers the central region in the range of $|\eta| \lesssim 1.0$ with the purpose of the high tracking resolution and reconstruction efficiency. The silicon tracker is used not only to precisely reconstruct the track impact parameter or interaction points, but also to be able to perform the silicon stand-alone tracking for the region $(1.0 < |\eta| < 2.0)$ which is not covered by COT.

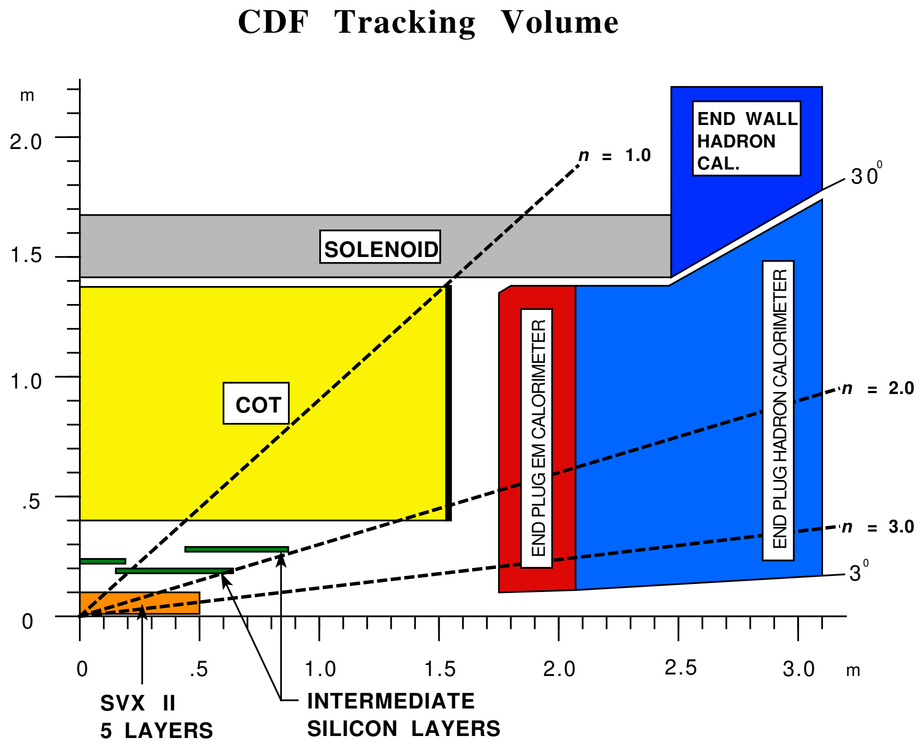


Figure 2.4: Longitudinal view of the CDF II Tracking System.

Inner Tracker: L00 + SVX II + ISL :

The inner tracking system consists of three silicon vertex detectors. We show a detailed end view of the CDF II silicon system in Figure 2.5. Total 7-8 silicon layers allow to archive a good impact parameter resolution and the silicon stand-alone tracking.

The Silicon Vertex Detector (SVX II) is built in three cylindrical barrels with a total length of 96 cm outside L00. Each barrel consists of five radial layers of double sided silicon microstrip detectors between radii of 2.4 and 10.7 cm, and each layer is divided in azimuth into 30° wedges. The first, second, and forth layers from the inner side combine an $r - \phi$ measurement on one side with 90° stereo measurement on the other, and the remaining two layers (third and fifth layers) combine $r - \phi$ with small angle stereo at 1.2°. This is designed to permit good resolution in locating the z-position of secondary vertices and to enhance the 3-D pattern recognition capability of the silicon tracker. The readout chips are mounted on an electrical hybrid on the surface of the silicon detectors. Each readout chip set (SVX3) has 128 channels, each with a charge-sensitive amplifier, 42-cell dual-ported pipeline with four additional cells for buffers, and an ADC. A highly parallel fiber-based data acquisition system reads out the entire detector in approximately 10 μs. This high speed and dual porting of the readout allows the SVX II information to be used for impact parameter discrimination in the SVT processor of the Level-2 trigger.

The Intermediate Silicon Layer (ISL) consists of three silicon layers. In the central region, a single ISL layer is placed at a radius of 22 cm. In the plug region, $1.0 < |\eta| < 2.0$, two layers of silicon are placed at radii of 20 cm and 28 cm. The ISL incorporates many features of the SVX II design. The crystals are double-sided with axial strips on one side and small angle stereo strips of a 1.2° on the other. The readout electronics are identical to the SVX II. In the plug region, the silicon stand-alone tracking is performed, where the COT tracking is not efficient. The ISL thus extends tracking, lepton identification, and b-tagging capabilities over the full region $|\eta| < 2.0$.

The innermost side Layer 00 (L00) [14] has been proposed a part of the upgrade of the CDF II detector. The L00 is a single-sided, radiation-hard silicon layer, placed immediately outside the beam pipe at 1.35-1.62 cm radius. All the readout electronics is identical to those used in SVX II. Being so close to the interaction point, the L00 achieves a good impact parameter resolution for high momentum tracks while that resolution for low momentum tracks is reduced in some regions of the detector because of the increase of the mass of the active volume.

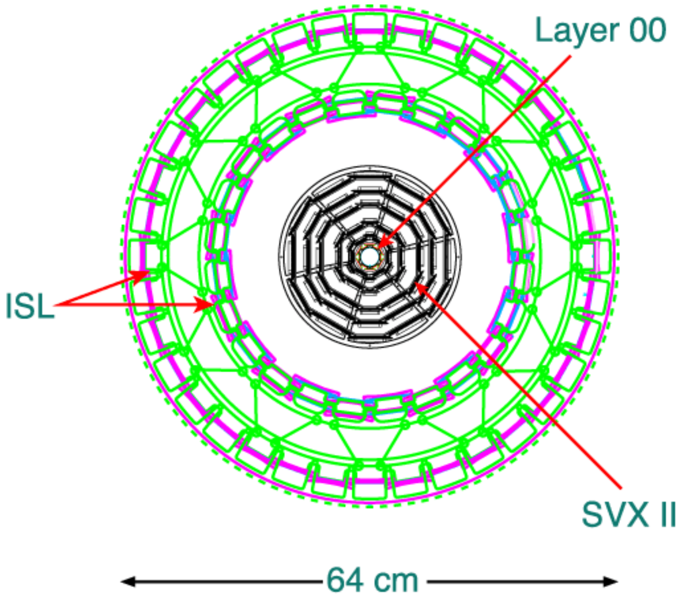


Figure 2.5: An end view of the CDF II silicon system including the SVX II cooling bulkheads and ISL support structure.

Central Outer Tracker (COT) :

The anchor of the Run II CDF tracking system is a large open cell drift chamber for charged particle reconstruction in the central region $|\eta| < 1.0$. The Central Outer Tracker COT is placed between the radii of 40 and 132 cm from the beam line. The design goal of the COT is to reproduce the functionality of the CTC, which was used in Run 0 and I, with drift times less than 100 ns by using smaller drift cells and a fast gas. A gas mixture is Ar-Ethane-CF_4 50:35:15 with a drift velocity of $\sim$100 μm/ns. The basic drift cell has a line of 12 sense wires alternating with shaper wires every 3.8 mm, running down the middle of two gold-on-mylar cathode planes which are separated by ~2 cm. The cell geometry on the superlayer 2 (SL2) is illustrated in Figure 2.6 as an example. Four superlayers are the axial side, and the other four superlayers are the stereo side with a small stereo angle of ±3°. Each superlayer is alternated starting with a stereo superlayer. The wires and cathode planes are strung between two precision milled endplates as shown in Figure 2.7. Both of the planes are tilted at an angle of 35° with respect to the radial direction to compensate for the Lorentz angle of drift electrons caused by the crossing of the solenoidal magnetic field with the drift electric fields of the COT. The tracking information is available for the Level-1 trigger.

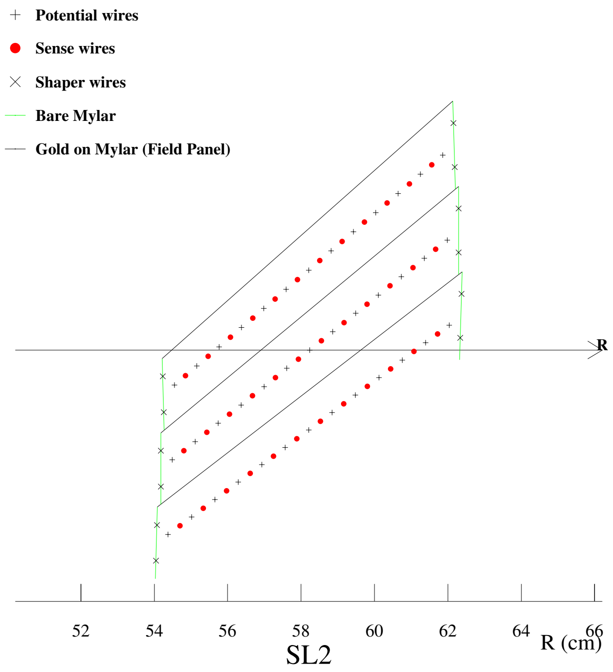


Figure 2.6: Nominal cell layout for SL2.

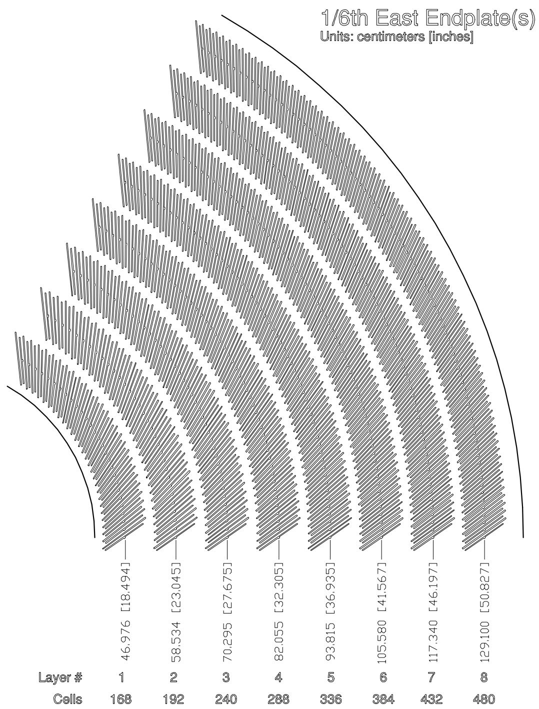


Figure 2.7: East endplate slots. Sense and field planes are at the clock-wise edge of each slot.

The CDF tracking systems are enclosed in a superconducting solenoid with a cylindrical volume 3.5 m long and 2.8 m in a diameter. The tracks of charged particles moving in a uniform magnetic field of 1.4 T give a helicoidal trajectory. The magnetic force in the tracking volume at CDF has the opposite direction to the beam axis (-z). The helix parameterisation at CDF is defined as

$$\vec{\alpha} = (\kappa,\, \phi_0,\, d_0,\, \cot\theta,\, z_0) \quad , \tag{2.3}$$

where,

κ : half curvature (same sign as the charge of the particle),
ϕ_0 : direction of track at point of minimum approach,
d_0 : signed impact parameter: distance between helix and origin at minimum approach,
$\cot\theta$: cotangent of the polar angle at minimum approach,
z_0 : z position at point of minimum approach to origin of helix.

The point of minimum approach is the closest point between helix and the line parallel to z-axis passing a reference point. The reference point is defined as z-axis at CDF. The sign of the curvature is the same as track charge, and the radius of helix is represent as $\frac{1}{2|\kappa|}$. The sign of the impact parameter is the follows: if a track has a positive(negative) charge and the reference point is outside(inside) the circle of the track, then the impact parameter has a positive sign. A drawing for a sign of the track impact parameter in the CDF coordinate systems is given in Figure 2.8. Using this parameterisation, the track momentum is represented as

$$p_T = \frac{c}{|\kappa|} \quad , \tag{2.4}$$

where $c = 1.49895 \times 10^{-3} \cdot B$ [Tesla]. With a uniform magnetic field of $B = 1.4116$ T at CDF, the constant term will be 2.116×10^{-3}. Note that the dE/dx correction *etc* for the track momentum is also applied by another applications.

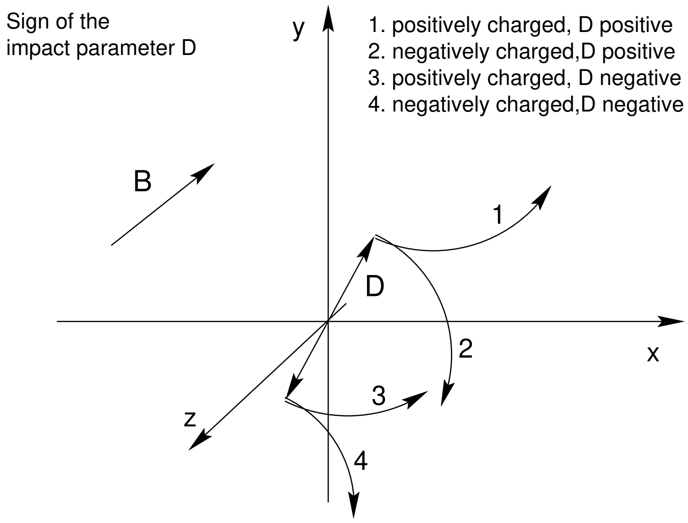
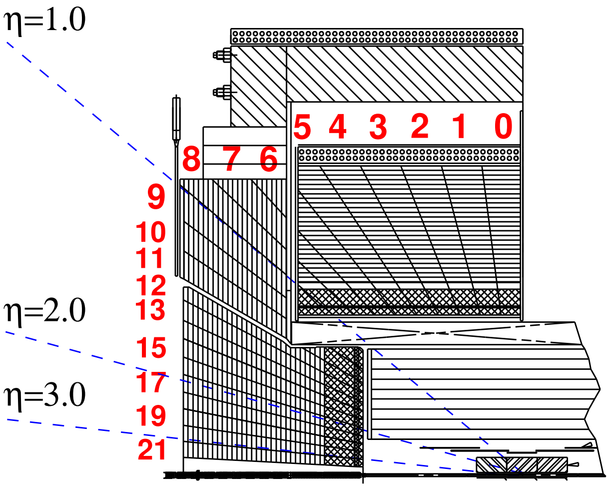


Figure 2.8: Tracks of a particle with positive/negative charge and positive/negative impact parameter.

2.3.3 Calorimeter systems

The CDF II calorimeter systems are located outside the solenoid with two separated devices of the electromagnetic (EM) and hadronic (HA) calorimeters. The hadronic section is placed behind the electromagnetic section. In Run II, the gas calorimeters in the plug region ($|\eta| > 1.1$) used in Run I were all replaced with a new scintillating tile plug calorimeter. The CDF II calorimeters are now all scintillator-based sampling calorimeters, and have a uniform pattern of matched projective towers of EM and hadron calorimeters. The central calorimeters (CEM and CHA) consist of 24 wedges ($15°$) in ϕ, and 10 towers ($\sim$0.1) in half of η region, containing a crack region around $\eta = 0$. In plug region, the segmentation of the plug calorimeters (PEM and PHA) consists of 48 wedges ($7.5°$) and 8 towers ($\sim$0.1) in $1.1 < |\eta| < 1.8$, and 24 wedges ($15°$) and 4 towers ($0.2 \sim 0.6$) in $2.1 < |\eta| < 3.64$. The endwall hadronic calorimeter (WHA) covers a gap between the central and plug hadronic sections. The detecting elements are arranged in a tower geometry pointing back toward to the interaction region. A quarter view of the CDF II calorimeter systems and that segmentation are shown/listed in Figure 2.9.



$\lvert\eta\rvert$ Range	$\Delta\phi$	$\Delta\eta$
0. - 1.1 (1.2h)	15°	$\sim$ 0.1
1.1(1.2h) - 1.8	7.5°	$\sim$ 0.1
1.8 - 2.1	7.5°	$\sim$ 0.16
2.1 - 3.64	15°	0.2 - 0.6

Figure 2.9: A quarter view of the CDF II calorimeter systems and that segmentation.

Electromagnetic Calorimeter (EM) :

The EM calorimeters are a lead/scintillator sampling device made of the sandwiched structure of 31 layers for CEM[2] and 23 layers for PEM, respectively, of which a unit layer is composed of 3.175 and 4.5 mm lead absorber and 5 and 4 mm plastic scintillator for CEM and PEM. When electrons or photons pass through the EM calorimeter, they produce the EM showers, and evolve within the EM section (18 and 21 X_0 (radiation lengths) for CEM and PEM, respectively). On the other hand, hadrons will start evolving later in the HA section due to the small interaction lengths of $\sim$1 λ of the EM section. The scintillation lights are transferred to the photomultipliers (PMT) via wavelength shifters and the number of scintillation lights are counted. A sketch of one wedge of the CEM and the cross section view of upper part of the upgraded plug calorimeter are shown in Figure 2.10 and Figure 2.12, respectively. The EM shower position is measured by a shower max detector (CES, PES) and pre-shower detector (CPR). In the central region (see Figure 2.11), two wire chambers (CES, CPR) are embedded into CEM. The CES is a multi-wire

[2] 31 layers of the plastic scintillator + 30 layers of lead sheets.

proportional chamber embedded near the shower maximum (approximately $6X_0$) into CEM calorimeters. The anode wires are strung parallel to the beam axis and the cathode strips to that perpendicular direction ($r - \phi$). Another wire chamber is placed immediately placed in front of CEM to act as a pre-shower detector (CPR). In the plug region (see Figure 2.13), two scintillator layers act as the shower maximum (PES) and pre-shower detectors. The PES consists of the scintillator-based strip readout divided into eight 45° sector. Each sector contains two layers (called U and V) of 5 mm pitch scintillator strips. The U and V layers are aligned at +22.5° and -22.5° with respect to the radial dimension to provide two-dimensional position measurement, and are inserted into the fifth sampling slot ($\sim 6X_0$) within PEM calorimeter. The first layer with 10 mm thick scintillator in PEM is used as a pre-shower detector. Schematic views of the CES and PES are shown in Figure 2.11 and 2.13, respectively. The electron track matching and the lateral EM shower profile in CES and PES are very powerful to separate the electron, photon, and the photons from the neutral pions. Characteristics of the CDF II EM calorimeter are summarized in Table 2.2.

Hadronic Calorimeter (HA) :

The HA calorimeters are a iron/scintillator sampling device made of the sandwiched structure of 32 layers for CHA and 23 layers for PHA, respectively, of which a unit layer is composed of 2.5 cm and 2 inch iron absorber and 1.0 cm and 6 mm plastic scintillator for CHA and PHA. The WHA is constructed of 15 layers of 5.1 cm iron absorber alternating with 1.0 cm plastic scintillator. The existing iron of the CDF endplugs is used in the new hadron calorimeter: stainless steel disks are attached to the inner 10° cone to extend the coverage to 3°. Two additional stainless steel disks are added behind the electromagnetic section to increase the thickness of the hadron calorimeter. In this way, the magnetic field in the tracking volume and magnetic forces on the end plugs are unchanged. The hadron section has the same tower segmentation as the EM section. Characteristics of the CDF II HA calorimeter are

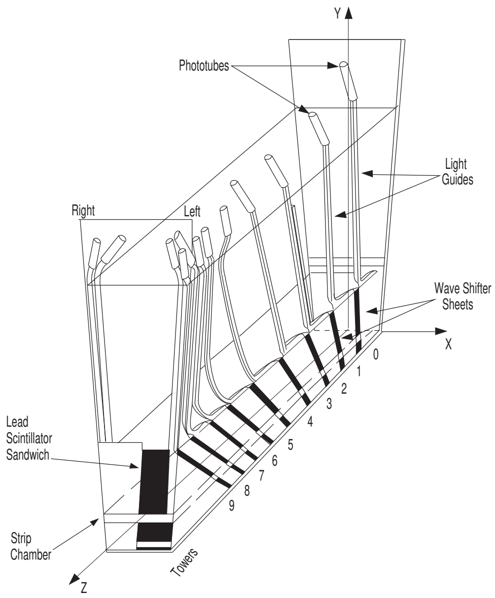


Figure 2.10: A sketch of one wedge of the CEM.

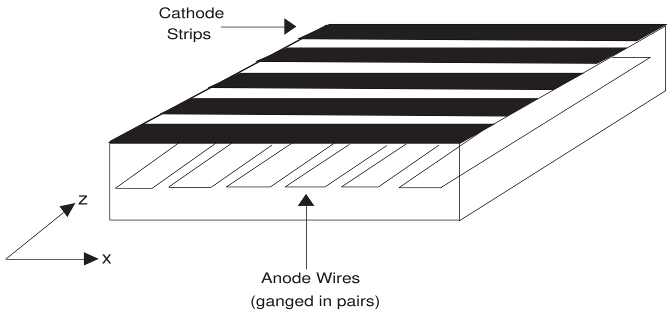


Figure 2.11: A sketch of the Central Electromagnetic Strip chamber (CES).

summarized in Table 2.2.

2.3.4 Muon systems

CDF II uses four systems of scintillators and proportional chambers in the detection of muons over the region $|\eta| < 2.0$. 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. New chambers are added to the CMP and CMX systems to close gaps in the azimuthal coverage in Run II, while the central chambers (CMU) have the almost same configuration without major changes from Run I. And the forward muon systems used in Run I was replaced with a completely new Intermediate Muon System (IMU).

Central muon detectors ($|\eta| < \sim$1.0) :

The Central Muon Detector (CMU) is a set of single-wire drift tubes consisted of 144 modules with 16 rectangular cells per module, located behind $\sim 5.5\lambda$ of absorber (the

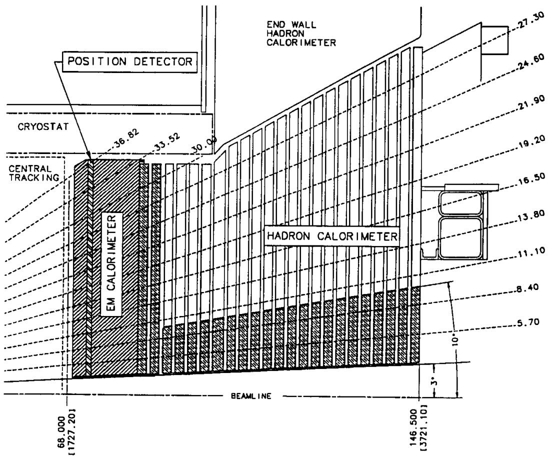


Figure 2.12: Cross section of upper part of new end plug calorimeter.

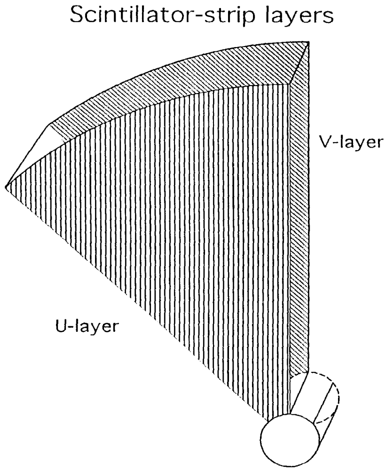


Figure 2.13: Geometry of the PES in a 45° sector.

	Central and End-wall	Plug
Electromagnetic:		
Thickness	$19X_0$, 1λ	$21X_0$, 1λ
Sample (Pb)	$0.6X_0$	$0.8X_0$
Sample (scint.)	5 mm	4.5 mm
Light yield	160 p.e./GeV	300 p.e./GeV
Sampling resolution	$11.6\%/\sqrt{E_T}$	$14\%/\sqrt{E_T}$
Stochastic resolution	$14\%/\sqrt{E_T}$	$16\%/\sqrt{E_T}$
Hadronic:		
Thickness	4.5λ	7λ
Sample (Fe)	1 inch.(central)	2 inch.
	2 inch.(end-wall)	
Sample (scint.)	10 mm	6 mm
Light yield	40 p.e./GeV	39 p.e./GeV
Energy Resolution	$75\%/\sqrt{E_T}$+3%	$80\%/\sqrt{E_T}$+5%

Table 2.2: Characteristics of the CDF II calorimeter.

Central Hadron Calorimeter). Each module consists of four layers. Since the CMU operates proportional mode rather than limited streamer mode, a new pre-amplifier is upgraded to recover that gain used in the limited streamer mode in Run I. The Central Muon Upgrade (CMP) consists of a second set of muon chambers behind an additional 60 cm of steel in the region $55^\circ \leq \theta \leq 90^\circ$. The pseudo-rapidity coverage of the CMP varies with azimuth angle due to the detector geometry that the chambers form a box around the CMU (See Figure 2.14). A layer of scintillation counters (CSP) is also installed on outside the surface of the CMP. The central extension consists of conical sections of drift tubes (CMX) and scintillation counters (CSX) located at each end of the central detector and extending in polar angle from 42° to 55°. This upgrade had been begun in Run I. The gap of the azimuthal coverage of CMX/CSX is now only a 30° region at the top of the detector for Tevatron Main Ring used in Run I and the solenoid refrigerator. The detailed design parameters of the central muon detectors are given in Table 2.3, and a schematic view of the coverage of each detector is shown in Figure 2.14.

Intermediate muon detectors ($\sim$1.0 < $|\eta|$ < $\sim$1.5) :

The Intermediate muon detector (IMU) is designed to trigger on muons with $\eta \leq 1.5$ and to identify off-line muons with $\eta \leq 2.0$. The IMU is almost the same configuration with the CMX/CSX, a set of drift tubes with four layer, located behind the plug calorimeter (6.2-20λ of steel). The forward muon toroidal magnet used in Run I[3] is re-used as the absorber without energizing. Additionally, a ring of steel $\sim$60 cm^2 is welded onto the inner face of the toroids. The muon momentum is measured using the solenoidal magnet field, where the silicon detector gives good momentum resolution. The detailed design parameters of the IMU are given in Table 2.3, and a schematic view of the coverage of each detector is also shown in Figure 2.14.

[3] The position of the toroidal magnet is moved 5.5 m closer to the interaction point.

	CMU	CMP/CSP	CMX/CSX	IMU
Pseudo-rapidity coverage	$\lvert\eta\rvert \leq \sim 0.6$	$\lvert\eta\rvert \leq \sim 0.6$	$\sim 0.6 \leq \lvert\eta\rvert \leq \sim 1.0$	$\sim 1.0 \leq \lvert\eta\rvert \leq \sim 1.5$
Drift tubes:				
thickness	2.68 cm	2.5 cm	2.5 cm	2.5 cm
width	6.35 cm	15 cm	15 cm	8.4 cm
length	226 cm	640 cm	180 cm	363 cm
Max drift time	800 ns	1.4 μs	1.4 μs	800 ns
Total drift tubes	2304	1076	2208	1728
Scintillators:				
thickness	—	2.5 cm	1.5 cm	2.5 cm
width	—	30 cm	30 - 40 cm	17 cm
length	—	320 cm	180 cm	180 cm
Total counters	—	269	324	864
Pion interaction lengths	5.5λ	7.8λ	6.2λ	6.2 - 20λ
Min. muon p_T	1.4 GeV	2.2 GeV 1.4 GeV	1.4 - 2.0 GeV	
Multiple scattering res.	12 cm/p (GeV/p)	15 cm/p (GeV/p)	13 cm/p (GeV/p)	13 - 25 cm/p (GeV/p)

Table 2.3: Design parameters of the CDF II Muon Detectors. Pion interaction lengths and multiple scattering are computed at a reference angle of $\theta = 90°$ in CMU and CMP/CSP, at an angle of $\theta = 55°$ in CMX/CSX, and show the range of values for the IMU.

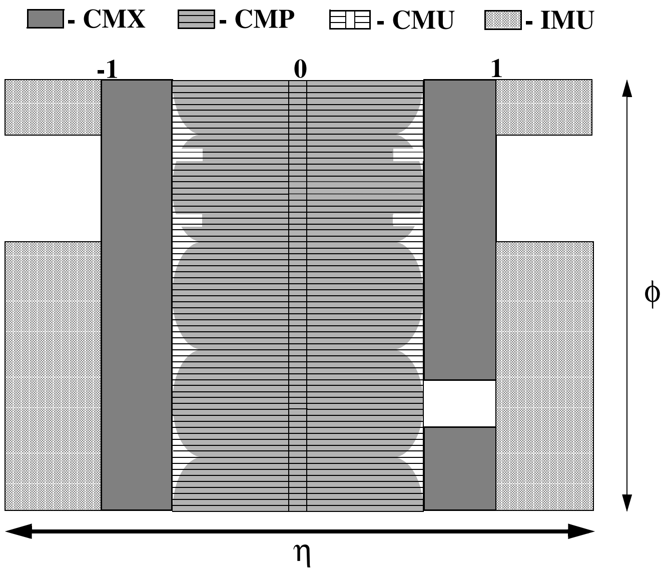


Figure 2.14: Location of the central muon upgrade components in azimuth ϕ and pseudo-rapidity η for Run II.

2.3.5 Trigger and data acquisition systems

There are three trigger paths at CDF trigger system so that a data acquisition system (DAQ) efficiently consumes the collision events within a 132[4] ns bunch-crossing rate. Since all the events cannot be stored, only the interesting events are selected by triggers. In each trigger step, the data size is reduced according to that triggering ability: 40 kHz acceptable rate at Level-1, 300 Hz for Level-2, and 30-50 Hz at Level-3 trigger stage.

Figure 2.15 shows the functional block diagram of the readout electronics. In Run II, the DAQ and trigger systems are largely replaced the systems used in Run I. To accommodate a 132 ns bunch-crossing time and a 4 μs decision time for the first trigger level, all front-end electronics are fully pipelined, with on-board buffering for 42 beam crossings. Data from the calorimeters, the central tracking chamber, and the muon detectors are sent to the Level-1 trigger system, which determines whether a $p\bar{p}$ collision is sufficiently interesting to hold the data for the Level-2 trigger hardware. The Level-1 trigger is a synchronous system with a decision reaching each front-end card at the end of 42-crossing pipeline. Upon a Level-1 trigger accept, the data on each front-end card are transferred to one of four local Level-2 buffers. The second trigger level is an asynchronous system with an average decision time of 20 μs. A Level-2 trigger accept flags an event for readout. Data are collected in DAQ buffers and then transferred via a network switch to a Level-3 CPU node, where the complete event is assembled, analyzed, and, if accepted, written out to permanent storage. These events can also be viewed by online monitoring programs running on other workstations.

Level-1 :

The Level-1 hardware consists of three parallel synchronous processing streams which feed inputs of the single Global Level-1 decision unit. One stream finds calorimeter-based objects, another finds muons, and the third finds tracks in the

[4] In the period of data taking considered in this analysis, the accelerator was operating in 36 bunches mode (396 ns) and the trigger was clocked every 132 ns with the two intermediate clock cycles automatically rejected.

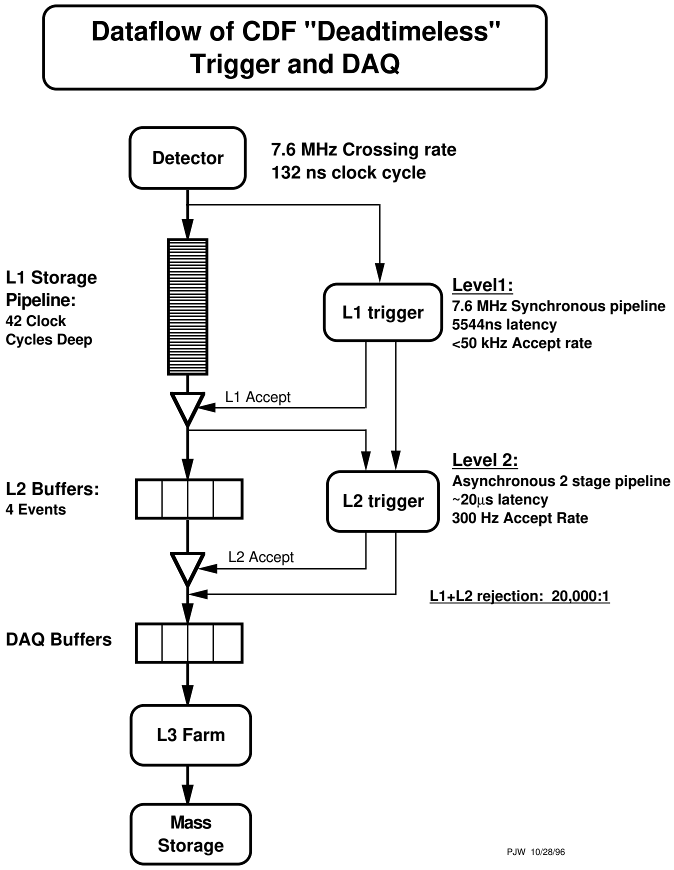


Figure 2.15: The Run II readout functional block diagram of the three level pipelined and buffered trigger system.

central region. The remarkable change for Level-1 in Run II upgrade is the addition of track finding, which was previously available only at Level-2. The 2-D track reconstruction (XFT, the eXtremely Fast Tracker) allows to identify an electron or a muon object at Level-1, or to use the kinematic properties of observed track pairs. Since the muon and electron triggers require the presence of a track pointing at the corresponding outer detector element (XTRP), the tracks must be sent to the calorimeter and muon streams as well as the track only stream. All elements of the Level-1 trigger are synchronized to the same 132 ns clock with a decision made every 132 ns by Global Level-1. The block diagram for the Run II trigger system is presented in Figure 2.16.

Level-2 :

Events accepted by the Level-1 system are processed by the Level-2 hardware. The Level-2 trigger consists of several asynchronous subsystems which provide input data to programmable Level 2 Processors in the Global Level-2 decision crate. Processing for a Level-2 trigger decision starts after the event is written into one of the four Level-2 buffers on all front-end and trigger modules by a Level-1 accept (See Figure 2.16).

The most significant addition to the Level-2 trigger is the Silicon Vertex Tracker (SVT) which provides the track impact parameter with high precision. So that, the secondary vertex information is available at Level-2 triggering. The data from the shower max detector (CES) are also combined at Level-2, which allows to separate electrons and photons. Jet reconstruction is also provided by the Level-2 cluster finder.

Level-3 :

All events accepted by Level-2 trigger are collected in the Event Builder (EVB), and then the EVB assembles those event fragments into one data block and delivers it to the Level-3 trigger system. The Level-3 trigger system is a farm of parallel processors which operate on a Linux PC, where a full event reconstruction is implemented in software. After passing through the Level-3 trigger, the Data Logger system delivers events to the tape device or online monitoring processes. The Level-3 reconstruction program is written in C++ with object-oriented techniques. The same reconstruction program is used in the offline event analysis.

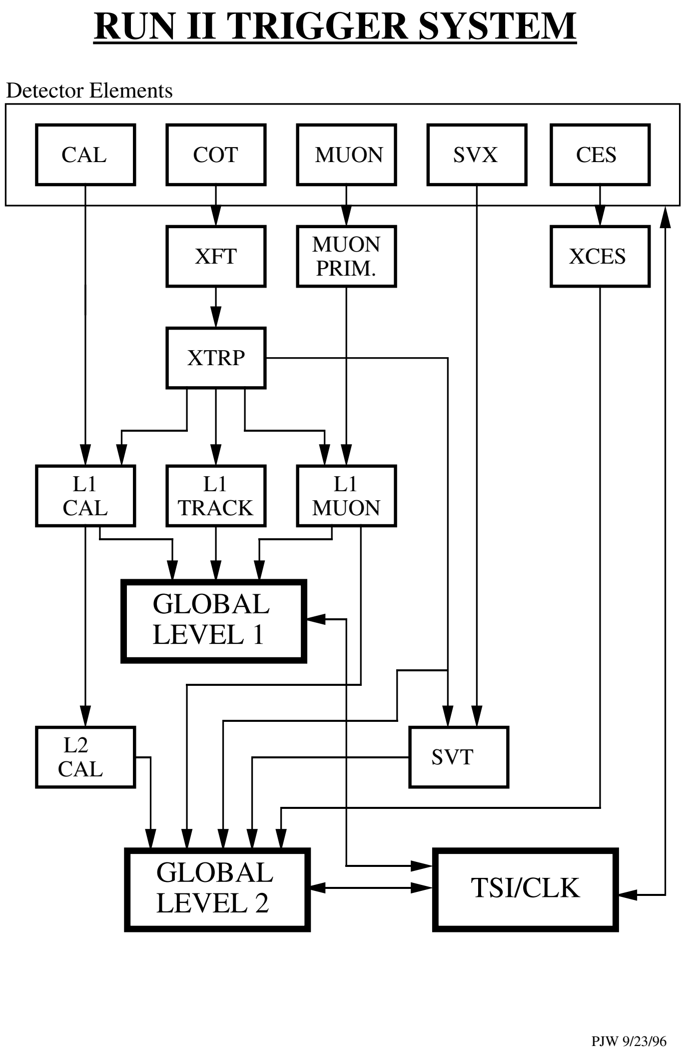


Figure 2.16: Block diagram for the Run II trigger system.

Chapter 3

$\mathbf{W^{\pm} \to e^{\pm}\nu}$ Event Selection

3.1 Introduction

In this Chapter, the details of the selection criteria of W boson are described. The W decays are detected in the leptonic decay mode, via $W \to e\nu$. Unlike the hadronic decays of the W, which are exposed among the huge QCD multijet backgrounds, the $W \to e\nu$ decay mode is easy to recognize the presence by requiring an isolated high-p_T electron and an imbalance of calorimeter energy due to the undetected neutrino. Using that clean signature, we extract about 3×10^5 $W \to e\nu$ events. After obtaining the W event samples by the electron selection and the missing transverse energy cuts, the jets identification cuts are required. To collect the $W \to e\nu$ events is the first step to measure the jet signatures.

3.2 Electrons

3.2.1 Electron triggers

Using information from several detector subsystems, the trajectories of electrons from $p\bar{p}$ collisions can be traced from the interaction region with high precision, though the tracking chambers, and into the electromagnetic calorimeters, as already described in Chapter 2. The identification of electrons begins with the online trigger system. The data passed through that online system, which classifies the electron characteristics, are stored into the tape permanently. Then, the offline reconstructions are performed for these events. All events are reprocessed with the re-calculated electron track parameters with the most up-to-date calibration constants. The best electron candidates are specified with several selection cuts. The results are a sample of tight central electron events that contain $W \to e\nu$ events as a subset. The trigger system is the first step to identify the $W \to e\nu$ events.

Some trigger paths to identify high-p_T electron events are combined as HIGH_PT_ELECTRON trigger. And some of those paths are so designed to estimate the trigger efficiencies on each trigger level. However, the following descriptions are principal requirements for $W \to e\nu$ candidate events to identify a high-p_T electron event in the central region ($|\eta| < 1.1$). The ELECTRON_CENTRAL_18 trigger path is used in this analysis. The W_NOTRACK trigger path is also used to estimate this trigger efficiencies which is a dominant source contributing to an electron trigger efficiency. The ELECTRON_CENTRAL_18 requires the electron objects with an energy in the central electromagnetic calorimeter and a track in the online track reconstruction. The W_NOTRACK is to make only calorimeter requirements for the electrons from W, but no track requirement. In stead of this, a large missing transverse energy is required at Level 1 and Level 3 with higher calorimeter energy thresholds than ELECTRON_CENTRAL_18 trigger. This is so-called a backup trigger to estimate the tracking efficiency of the electron. A detailed description of the trigger efficiencies is provided in Section 3.7.

The online triggers are constructed in three stage of the trigger sets:

Level 1 : The Level 1 triggers are used to extract hard-scattering events from $p\bar{p}$ collisions. Most of electrons from W passed the L1_CEM8_PT8 and L1_EM8_MET15 trigger for ELECTRON_CENTRAL_18 and W_NOTRACK, respectively. The L1_CEM8_PT8 turns on if the energy deposited in any calorimeter towers and the momenta of XFT tracks fitted as a track trajectory exceed their preset thresholds. The energy threshold for the central electromagnetic calorimeter

(CEM) is 8 GeV and a fraction of the hadronic energy, Had/EM, is lower than 0.125. The L1_EM8_&_MET15 requires the missing transverse energy greater than 15 GeV instead of the XFT track of L1_CEM8_PT8. The Level 1 trigger efficiency for $W \to e\nu$ sample is essentially 100 %.

Level 2 : After passing the Level 1 trigger, Level 2 trigger decision is applied with more tighter cuts than Level 1 trigger requirements. For ELECTRON_CENTRAL_18, the L2_CEM16_PT8 requires an energy cluster in the CEM calorimeter with $E_T \geq 16$ GeV and Had/EM ≤ 0.125 as well as an XFT track with $p_T \geq 8$ GeV/c. On the other hand, a slightly higher E_T threshold is required in the W_NOTRACK. The L2_CEM16_L1_MET15 requires the E_T threshold of 16 GeV without track requirements.

Level 3 : The most sophisticated online electron identification takes place at this stage. For ELECTRON_CENTRAL_18, the L3_CEM18_PT9 requires the CEM tower with $E_T \geq 18$ GeV, Had/EM ≤ 0.125, and the Level 3 track with $p_T \geq 9$ GeV, while for W_NOTRACK, the L3_EM25_MET25 requires the missing transverse energy with $\not{E}_T \geq 25$ GeV instead of the requirement for the electron track.

When the trigger requirements of all three levels are combined, the efficiency for identifying a central electron is about 96 %.

3.2.2 Electron clustering

A bunch of the electromagnetic calorimeter towers which satisfy the trigger selection criteria are formed as the electromagnetic clusters, which is a collection of towers distributing around a seed tower with largest E_T (> 3 GeV), where the E_T of a calorimeter tower is defined by $E_T = E \cdot \sin\theta$ as the θ is the polar angle measured from the event vertex to the centroid of the tower. The neighbor tower around the seed tower is only two[1] adjacent towers laid on the $|\eta|$ direction (called "shoulder" tower), not on ϕ. The tower threshold

[1] If only one neighboring tower is found, then, only that tower is included.

in that cluster is $E_T > 100$ MeV. In addition, the ratio of the hadronic to electromagnetic energy associated with the cluster is required to be less than 0.125 in order to reject hadrons.

3.2.3 Corrections

Various corrections are applied in data as well as the MC simulation after triggering the high p_T electron sample. In this section we describe those corrections in brief.

Vertex Correction :

In the trigger level, the electron transverse energy is calculated assuming the interaction point is located at $z = 0.0$ cm. Using obtained z position, the $\not{E}_T$ is corrected. The z position of the interaction point is defined as the z_0 position of the electron track. This requirement are already demonstrated in Run I [15] that using the track z_0 in the $W \to e\nu$ or $Z \to e^+e^-$ events is better than the primary vertex measured by jet vertex algorithm. Additionally, we use the angle of the electron track to calculate the E_T.

CEM Energy Corrections :

Corrections of the CEM response as a function of the position in local coordinates are implemented. We also correct the CEM energies for tower-to-tower gain variations. Additionally, the absolute energy was scaled in the data and in the simulation such that the $Z \to e^+e^-$ mass peak is around 91 GeV. This resulted in $+1.2$ % correction in the data and -0.6 % correction in the simulation. The energy in the simulation was also smeared by 2 % in order to match the resolution observed in the data. Energy dependent variables were re-calculated to reflect these corrections.

Beam Constrained Tracking :

The tracks of the electron candidate are refitted by a beam constrained tracking algorithm only using the COT information. The beam constrained track induces a bias in the track curvature since the track fitting is forced to the beam spot. The change of the track curvature has been studied by looking at the energy-momentum ratio E/p [16]. For data, we apply a correction,

$$(signed)\ \frac{1}{p_T} \ = \ \frac{1}{p_T} \ - \ 0.00042 \ - \ 0.00116 \times \sin(\phi \ + \ 0.3) \tag{3.1}$$

Note that we do not correct the angle of the track.

Leakage Correction :

The electron cluster is basically formed only using one tower in CEM. Only one adjacent tower along the η direction is allowed to be merged into the electron cluster in CEM if that tower has an energy deposition larger than the seed tower E_T. If the electron hits near the edge of the tower in ϕ but not in η, the leakage energy thus increases significantly. The isolation requirement on the electron energy to discriminate between jets and electrons or photon will be disturbed by this energy leakage. Thus, we correct the isolation energy using the parameterized formula [17] as a function of the E_T and x_{CES}:

$$Iso(0.4)_{corr} \ = \ Iso(0.4) \ - \ E_T \times p_0 \cdot \exp\{p_1 \cdot (|x_{CES}| - 21)\} \qquad , \tag{3.2}$$

where $p_0 = 0.0511 \pm 0.0075$ and $p_1 = 0.333 \pm 0.061$. The details about the variable $Iso(0.4)$ is described in latter section.

3.2.4 Electron geometric and kinematic cuts

Electrons from $W \rightarrow e\nu$ decays typically have a large transverse energy. The following geometric and kinematic cuts are applied to identify electrons from W.

Fiducial requirements :

The fiducial volume are taken place to ensure the stable and good response from the electromagnetic calorimeter for the electron candidates. The boundary of the calorimeters in $|\eta_d| < 0.05$ and two chimney towers, which are located for the cryogenic penetration of the magnet, are excluded. In addition to the restriction due to the mechanical design, the fiducial requirements to eliminate regions of poor response are also imposed on the azimuthal boundaries in the local position coordinates of the electron at the calorimeter wedge, $|x_{local}| \geq 21$ cm, and the 90° boundary of the CEM, $z_{local} \leq 9$ cm and $z_{local} \geq 230$ cm, The local positions on the electromagnetic cluster are extrapolated using the maximum p_T track to the plane of the CES in the wedge.

Kinematic requirements ($E_T \geq 20$ GeV, $p_T \geq 10$ GeV, $|\eta_d| < 1.1$) :

The electron candidates are required within the central region, $|\eta_d| < 1.1$, which is covered by CEM. That guarantees precise energy measurements and electron track quality. This requirement reduces up to 50 % of the total number of $W \rightarrow e\nu$ events. The transverse energy of every electron candidate is required to be larger than 20 GeV ($E_T \geq 20$ GeV) within the central region ($|\eta_d| < 1.1$). Also, the track transverse momentum have to be larger than 18 GeV ($p_T \geq 18$ GeV). That allows us to distinguish electrons from photons or neutral pions (π^0). The track transverse momentum is measured from the fully-reconstructed track curvature.

Interaction Vertex ($|Z_{vtx}| \leq 60$ cm) :

To keep the $W \rightarrow e\nu$ interaction within the fiducial volume of the detector, and to maintain the calorimeter projectile tower geometry, we require the interaction vertex to lie within 60 cm around the center of the detector in z. We identify the interaction vertex as the z_0 position of the electron track.

3.2.5 Electron identification cuts

Several electron identification variables are used to reject backgrounds and enhance the fraction of true electrons.

Energy-Momentum Ratio ($E/p \leq 2.0$ or $E_T \geq 50$ GeV) :

The electron passing though the COT material in the magnetic field makes bremsstrahlung. Due to the worse momentum resolution for the high-p_T tracks, the energy-momentum ratio are required to be in the range of $E/p \leq 2.0$ if the electron transverse energy is less than 50 GeV.

Hadronic Energy Fraction (Had/EM $\leq$ 0.055 + 0.00045 $\times$ E) :

Electromagnetic showers are mostly contained within the EM calorimeter. To reject the contamination of hadronic jets that deposit a large fraction of energy in the hadronic calorimeter, small amount of the fraction of hadronic energy are imposed. Since the leakage of electron energy into the hadronic calorimeter increases with energy, the sliding cut

$$Had/EM \ \leq \ 0.055 \ + \ 0.00045 \times E \tag{3.3}$$

is applied to maintain a high efficiency for the electron, where E is the cluster energy in GeV.

Isolation (Iso(0.4) $\leq$ 0.1) :

The isolation is defined as

$$Iso(0.4) \ \equiv \ \frac{E_T^{0.4} - E_T}{E_T} \quad , \tag{3.4}$$

where $E_T^{0.4}$ is the total transverse energy contained within a cone of radius R $= 0.4$ (in $r - \phi$ space) with respect to the center of the electromagnetic cluster. This cut is also imposed to reject the contamination of hadronic jets by Iso(0.4) ≤ 0.1. A small isolation cut leads to a well-separated cluster.

Lateral Shower Profile ($L_{shr} \leq$ 0.2) :

The lateral sharing of energy between the calorimeter towers gives a criterion to identify electrons. The lateral shower profile, L_{shr}, is defined as

$$L_{shr} \ = \ 0.14 \times \sum_i \frac{E_i^{adj} \ - \ E_i^{exp}}{\sqrt{(0.14)^2 E \ + \ (\Delta E_i^{exp})^2}} \quad , \tag{3.5}$$

where the summation is over the two towers adjacent to the seed tower in the same azimuthal wedge, E_i^{adj} is the measured energy in a tower adjacent to the seed tower,

E_i^{exp} is the expected energy in that tower, ΔE_i^{exp} is the uncertainty on E_i^{exp}, and $0.14 \cdot \sqrt{E}$ is the uncertainty on the cluster energy. The expected energy, E_i^{exp}, is a function of the seed tower energy, the impact point, the event vertex, and a shower profile from the test-beam results. L_{shr} is required to be less than 0.2.

Strip Chamber Profile ($\chi^2_{str} \leq$ 10) :

The pulse heights on the CES strips in the electromagnetic shower are compared with channel-by-channel between the observed shower and the expected shower profiles based on the test-beam results. The quantity from the fitted results, χ^2_{str}, is required to be less than 10.

Track Matching ($-$3.0 cm $\leq Q_e \cdot dx \leq$ 1.5 cm and $dz \leq$ 3.0 cm) :

We require a track matching that the extrapolated track of the electron candidate points out into the electromagnetic shower location measured by CES. The differences to each direction, dx and dz, denote the separation between electron track and electron cluster in CEM, where Q_e is a track charge.

Track Segment :

The only well-reconstructed electron tracks are used. We require that the stub of the electron track has more than 3 superlayers with at least 7 hits in each superlayer in COT.

Conversion Removal :

When high energy photons convert to electron-positron pair, those signature will be mimics of the high E_T electron signatures. We remove those fake signals to require an electron track pair with opposite sign points the same origin. At the conversion point, both tracks are almost parallel and follow the direction of the parent photon. We thus require $|\Delta XY| \leq 0.2$ cm and $|\Delta \cot\theta| \leq 0.04$ for an opposite-signed track, where $|\Delta XY|$ is the $R\Delta\phi$ separation at the point of conversion. In order to avoid the over-filtering to the additional conversion of the photon radiated from the real electron, we require no existence of an additional conversion partner which is presumably coming from the real electron.

3.3 W selection

3.3.1 Missing transverse energy ($\not\!\!E_T \geq$ 30 GeV)

Once we get a pure inclusive electron sample, the main remaining task is to make use of the imbalance of the transverse energy to select W boson candidates. If we have more than two tight electron candidates, which satisfied in the selection criteria of the tight electrons described in the previous section, in an event, we define the maximum-E_T electron as the W electron.

To better reject the backgrounds in the multijet samples, we used an higher missing E_T cut than the electron E_T cut. This choice is justified by the fact that the corrections to the missing E_T are not yet fully understood and we preferred to sacrifice some selection efficiency keeping an high purity. The missing E_T is re-calculated using the event z vertex defined by the z_0 of the maximum-E_T tight lepton, and corrected for any tight muon in the event: the muon p_x and p_y components are subtracted and the energy deposits left by the muon in the calorimeter towers are added on a tower-by-tower basis.

3.3.2 $Z \to e^+e^-$ removal

If one of the electrons from Z bosons decaying to the electron-positron pair $Z \to e^+e^-$ goes into the crack region in the calorimeter system, their signature will result in one tight electron plus large missing energy which will fake the $W \to e\nu$ signature. We thus search for the second leg of the tight electrons in the inclusive high-p_T electron sample. If the second largest E_T electron has the opposite sign to the first electron, and the invariant mass of these electron-positron pair satisfies the following criteria,

$$75 \text{ GeV} \leq M_{e^+e^-} \leq 105 \text{ GeV} \quad , \tag{3.6}$$

those events are discarded as fake events of $W \to e\nu$.

3.4 Jets

3.4.1 Jet identification algorithm, JetClu

The partons produced by the hard interaction processes of the $p\bar{p}$ collisions experience the strong force of QCD, and then fragment to the bunch of hadrons. This parton fragmentation process with the subsequent deposition of single particle energies into the tower of the calorimeter creates the typical jet structures observed in collisions producing high transverse energy. We therefore use the energy information on calorimeter towers to reconstruct a jet, but the assignment of towers to the jet is not unique.

The definition of a jet is obviously arbitrary and found by a suitable algorithm. If one tries to extrapolate a parton information using the suitable algorithm, one has to pay attention to the parton configurations in the perturbative calculation. Higher-order QCD processes give more elaborate configurations of partons. For instance, the gluon radiation produced by the association with the hard-scattering partons may not be observed under the loose size of cone, while the dominant configuration however are the presence of the high-pt scattering from the outgoing hard-scattering partons. If the experimental definition of a jet is sufficiently loose, all configurations will fall into the single jet.

Various cone algorithms has been developed with many experiments, where a cone have a circular cross-section in the η-ϕ plane to account for the distortion expected to form the Lorentz transformation. We employ a fixed cone algorithm, JetClu, which is widely used in many analysis at CDF. One merit of using the fixed cone algorithm is a definition of the tracks associated with the cone. This is supported by the algorithms of a heavy flavor tagging like a bottom quark. In the next section, we describe more details about the JetClu algorithm.

The JetClu starts by making a list of all towers with an E_T greater than some threshold, called E_T^{seed}. A second list containing candidates for clustering is made of all towers above a second threshold E_T^{cand}. At the present the default values of E_T^{seed} and E_T^{cand} are 1 GeV and 0.1 GeV, respectively. In the plug and forward calorimeter regions, towers are grouped together in sets of three in ϕ, spanning 15° to correspond to the segmentation of the central calorimeter. Precluster is formed by combining all touching seed towers, which are required to have continuously decreasing tower E_T. The clustering is performed using the tower segmentation without ganging towers used in preclustering. A fixed cone in η-ϕ space of radius R is formed around the centroid determined from the E_T weighted centroid of the precluster. The candidate towers in this cone are merged into the cluster and the centroid is recomputed. Again, all candidates inside the cone around the new centroid are merged in. This process is iterated on until the tower list remains unchanged. The actual cone size used in the analysis is 0.4. The iterative cone algorithm also provides a scheme for treating overlapping clusters. If the towers of one cluster are completely contained within another, the smaller one (lesser E_T) is dropped. If the towers of different clusters partially overlap, an overlap fraction is computed as the sum of the E_T of the common towers divided by the E_T of the smaller cluster. If the fraction is above a cutoff (default is 0.75), then the two clusters are combined. If the fraction is less than the cut, the clusters remain separate, and each tower in the overlap region is assigned to the cluster with the nearest center in η-ϕ space. After the towers are uniquely assigned to clusters, the centroids are recomputed. This tower reshuffling process is iterative, and ends when the tower lists remain fixed.

3.4.2 Variables in JetClu

During the clustering process, the centroid associated with each cluster is calculated by assigning massless four-vectors to each of the electromagnetic and hadronic towers. The four-vectors have a magnitude equal to the energy deposited in the tower, and a direction defined by a unit vector pointing from the event vertex to the center of the face of the calorimeter tower (calculated at the depth that corresponds to shower maximum). A cluster four-vector (p_x, p_y, p_z, E) is then defined by summing over the towers in the cluster:

$$p_x = \sum_i p_x^i \quad , \quad p_y = \sum_i p_y^i \quad , \quad p_z = \sum_i p_z^i \quad , \quad E = \sum_i E^i \quad . \tag{3.7}$$

Given the cluster four-vector, several jet quantities are readily calculated:

$$E_T = E\sin\theta \quad , \quad \eta = -\ln\tan(\frac{\theta}{2}) \quad , \quad \phi = \arctan(\frac{p_y}{p_x}) \quad , \tag{3.8}$$

where

$$\theta = \arcsin\left[\frac{\sqrt{p_x^2+p_y^2}}{\sqrt{p_x^2+p_y^2+p_z^2}}\right] \quad . \tag{3.9}$$

Because the z vertex distribution is spread out along the beam line, forming a Gaussian with a σ of approximately 30 cm, it is necessary to correct the pseudorapidity of all jets from η_d to η (see Section 3.2.3). This shift implies a small energy correction to account for the incidence angle of the jets on the face of the calorimeter.

3.4.3 Jet energy corrections

The measured four-vector of jets generally differs from the energies of the initial partons. This is the result from both instrumental and physical effects such as low energy non-linearities, η crack energy losses, underlying events and clustering. Some of the corrections are decided by the measurable quantities independent of the theory, while some of them rely on the theory prediction. Thus the raw jet energies measured in the calorimeter must be corrected for detector effects at first before they can be compared to physics predictions/models. The correction strategy [18, 19] is the followings:

$$p_T(R) = (p_T^{raw}(R)\times f_{rel} - UEM(R))\times f_{abs}(R) - UE(R) + OC(R) + SPLO \ , \tag{3.10}$$

where R denotes the clustering cone size,

- f_{rel}, the relative jet energy corrections to make the calorimeter response uniform in η (determined from data),
- $UEM(R)$, the multiple interactions corrections (determined from data),
- $f_{abs}(R)$, the absolute jet energy corrections to convert calorimeter cluster p_T to the $\sum(p_T)$ of the particles in the cone (determined from calibration data and simulation),

- $UE(R)$, the underlying event subtraction to obtain particle level "jet" p_T (determined from data),

- $OC(R)$, the out-of-cone corrections to get to the full "parent parton" p_T (determined from simulation), and

- $SPLO$, the splash-out correction to account for the infinite range outside cone. Since $OC(R)$ already refers the energy outside cone, this correction is used only to estimate the systematic uncertainty.

We describe more details about each correction procedure below.

Relative Corrections :

The first step in jet energy corrections is to correct the jets for any variation in response with detector η. For this correction, dijet event samples are used. Since the transverse energy of the two jets in a $2 \to 2$ process should be equal, the energies of jets in the plug and forward calorimeters are scaled to give the energy of an equivalent jet in the central calorimeter. One well-measured central jet ($0.2 < |\eta| < 0.6$) are required and a scale factor is derived from the dijet balance to a second jet. The central calorimeters CEM/CHA are the best understood calorimeters in CDF [20] and the selected region is far away from the cracks. The gain variation depending on the time (run range) in the plug calorimeters is also taken into account [21]. The corrections for the Monte Carlo data [22] are determined separately since some discrepancy between data and simulation can be seen due to a lack of the materials in the detector simulation [23].

CHA and WHA Energy Corrections :

The absolute energy scale corrections had been performed using test beam data by the radiative source calibration run since Run I(0) experiment. The time dependence of the calorimeter response is calibrated by using the minimum ionizing peak obtained with muons. The J/ψ and high-p_T muon samples are used for this calibration and the MIP peak is compare with Run I [24]:

$$E_{cal}^{MIP-RunII}/E_{cal}^{MIP-RunI} \;=\; 0.960 \;\pm\; 0.005 \quad . \tag{3.11}$$

This factor is used as the correction parameter. To check the non-linearity response for hadron sources in CHA, the raw jet energy scale can be checked using photon-jet balancing [25] since the photon energy measured by the electromagnetic calorimeter is very precisely measured [2]. The results differ about 5 % lower from the Run I results. This effect is not understood yet. That leads a large systematic uncertaintly of 100 %.

Multiple Interaction Correction :

In the designed luminosity with 36 bunches, although the average one interaction per beam crossing is expected, the multiple interaction affects the measured jet energy when the energy from these minimum bias events may fall into the jet clustering cone. The transverse energy in a random cone was measured in minimum bias data and parameterized as a function of the number of vertices in the event. This transverse energy is subtracted from each jet to account for multiple interaction in the same bunch crossing as a function of the number of vertices in the event. This correction factor is a linear function of the number of reconstructed vertices in the event [26]. Only vertices associated with at least ≥ 2 COT tracks (addressed as quality 12) in minimum bias events are used to decide this correction factor. Efficiency and fake rate of the vertex reconstruction algorithm are different from in Run I and thus the multiple interaction energy per vertex is slightly larger than those in Run I.

Absolute Corrections :

The jet energy measured by the calorimeters must be corrected for any non-linearity and energy loss in the un-instrumented regions of each calorimeter. The absolute jet corrections account for the response to particle-level energy in the central calorimter.

[2] The CEM corrections are already described in Section 3.2.3.

This correction depends on the jet fragmentation properties. The calibration point is derived using a 50 GeV pion from test beam data. For the non-linearity response, the tuned Monte Carlo events generated by ISAJET [27] with Feynman-Field fragmentation [28] are used for the charged and neutral particles. After fragmentation, the events are processed with a full CDF detector simulation. Each simulated event is compared to the total p_T of all generated particles lying in a cone centered about the measured jet axis. A quadratic spline fit is used to parameterize the mean jet response as a function of E_T for the cone sizes 0.4, 0.7, and 1.0. The same corrections can be used in Run II as long as non-linearity and relative response to π^0 and $\pi^{\pm}$ have not changed since Run I, except for a scale factor.

Underlying Event Corrections :

The underlying event contains all the soft interactions except the hard one. The underlying event energies must be subtracted from the measured jet energy when these particles fall into the clustering cone. The correction procedure is the same as the multiple interaction correction. Events with only one vertex are used to determine the underlying event correction.

Out-of-Cone Corrections :

The jet clustering may not include all the energy from the initiating parton. Some of the partons generated during fragmentation may fall outside the cone chosen for clustering algorithm. Out-of-cone corrections are applied in order to correct the particle-level jet energy to the parton energy (as much as theoretically allowed). These corrections are completely independent of detector/calorimeter performance and depend on the parton fragmentation functions. The correction factor is parameterize as a function of jet p_T. Jets tend to become "narrower" at large energies, and the fractional energy deposited outside the cone decreases. For 0.4 jet cones, the correction decreases from $\sim +25\%$ at $p_T = 15$ GeV to $\sim +15\%$ at $p_T = 100$ GeV. The corrections derived for Run I are used for Run II data.

3.4.4 Systematic uncertainties

We have not determined the absolute and out-of-cone corrections for Run II. In order to be able to use those evaluated for Run I, we compared the raw jet energies of Run II with those of Run I as a calibration point. The γ-jet balance technique was used at this purpose. It resulted in a disagreement between Run I and Run II, for which an additional correction has been applied to the raw jets in Run II [19]. We note the following contributions to the systematic uncertainties to reflect the fact that simulation and corrected data do not agree. Details of the systematic uncertainties on the current jet corrections can be found in [29].

- The raw energy scale outside $0.2 < |\eta| < 0.6$ region has $\sim$2% uncertainty relative to the scale in the central region.
- The raw jet energy scale of the central CDF II calorimeter has $\pm$5% uncertainty (from γ-jet balancing).
- We have reevaluated the underlying event and multiple interaction corrections, but we assign to them the same systematic errors as in Run I. For the underlying event correction, the systematic uncertainty is $\pm$ 30%, and for the multiple interaction correction, 100 MeV/vertex.
- We have not reevaluated the absolute and out-of-cone corrections for Run II. We use the same corrections and uncertainties as those used in Run I [30]. Those uncertainties are $2.5 \sim 3\%$ for the absolute corrections, and $1.4 \sim 7\%$ for the out-of-cone corrections.

3.5 Luminosity measurement

The luminosity for hadron collider experiments can be determined from the counting measurement of inelastic $p\bar{p}$ interactions which has a large cross section measured with an uncertainty of approximately 3% [31]. This process allows precisely to measure the total integrated luminosity. The luminosity monitor has been done at the Cerenkov Luminosity Counters (CLC) by measuring the rate of the inelastic $p\bar{p}$ events. The Cerenkov Counters has a merit that the Cerenkov light(ring) is not insensitive to an accidental events such as the beam halo.

The total integrated luminosity is expressed as

$$\mathcal{L} = \frac{R_{p\bar{p}}}{\varepsilon_{CLC} \cdot \sigma_{in}} \quad , \tag{3.12}$$

where $R_{p\bar{p}}$ is the rate of the inelastic $p\bar{p}$ events measured by CLC, ε_{CLC} is the CLC acceptance determined from data and simulation, and σ_{in} is the inelastic cross section. The inelastic cross section 61.7 mb is obtained by scaling the CDF measurement $\sigma_{in} = 60.4 \pm 2.3$ mb at $\sqrt{s} = 1.8$ TeV to 1.96 TeV [32]. A +1.9 % correction to the luminosity which comes from extrapolating the inelastic cross section from $\sqrt{s} = 1.8$ TeV to 1.96 TeV is also applied. Using these numbers, and requiring the run to belong to the "Good Run List" described in Section 3.6, the total luminosity is estimated to be

$$\mathcal{L} = 72.0 \pm 4.3 \text{ pb}^{-1} \quad . \tag{3.13}$$

The consistency is also cross-checked using other physics processes of W production or jets events. The estimated uncertainty in the CLC measurement is 4.4% and is mostly due to the uncertainty in the absolute normalization of the CLC acceptance. If we use the CDF measurement of the inelastic cross section with the uncertainty of a 2.5%, the total systematic uncertainty of the integrated luminosity would be 6%. The details of the luminosity uncertainty are presented in Table 3.1.

3.6 The $W \to e\nu$ event sample

3.6.1 Datasets

The data used in this analysis were taken with the **HIGH_PT_ELECTRON18** trigger from March 2002 to January 2003 (Run# 141544 $\sim$ 156487) corresponding to an integrated luminosity of 72.0 pb^{-1}. To ensure a good quality data we required each run to

Effect	Uncertainty Estimate
Inelastic cross section (CDF measurement)	2.5 %
CLC acceptance	4.4 %
Detector instability	< 2 %
Detector calibration	< 1.5 %
On-line to off-line transfer	< 1.0 %
Non-linearity at high luminosity ($> 10^{32}$)	N/A
Total Uncertainty	~5.7 %

Table 3.1: Systematic uncertainties in the luminosity calculation using the CLC and the CDF measurement of the inelastic cross section.

satisfy the set of minimal conditions classified as "Good Run". The Good Run conditions in this period are maintained to have at least 10 nb^{-1} in each run when any detector components other than the silicon vertex detector were properly functioning.

The data sets contain not only the $W \to e\nu$ events but also the other physics processes, $Z/\gamma^* \to e^+e^-$, $Z/\gamma^* \to \tau^+\tau^-$, or $W \to \tau\nu$, and QCD processes, where the QCD processes are the semileptonic decay from heavy flavor quarks or the mimics of the electron signature which is mostly π. Among the 72.0 pb^{-1} of data, we collect 2,027,335 events of the inclusive electron sample.

3.6.2 The $W \to e\nu$ event sample

We start to collect the tight electron data sample described in Section 3.2. After obtaining the clean electron signature, the missing transverse energy cut and the physics background rejection cut from $Z \to e^-e^+$ are applied to get W candidates. We extract about 3×10^5 $W \to e\nu$ events with this integrated luminosity. In Table 3.2, we present the number of selected events on each selection criterion. The tight electron data sample contains 87646 events. After the W selection criteria, we find 31726 W candidates.

For the event selection, we use slightly different selection criteria used in the inclusive W production cross section measurement in Run II [33]. To better reject the background in the multijet samples we used an higher missing E_T cut, and, to keep a reasonable

Electron Cuts:	Events	fraction
Central Region ($\lvert\eta_D\rvert < 1.1$)	352,420	1.00
$\lvert Z_{vtx}\rvert \leq 60$ cm	336,250	0.95
Fiducial Region	309,376	0.88
$E_T \geq 20$ GeV	298,461	0.85
$P_T \geq 10$ GeV	298,000	0.85
Had/EM $\leq 0.055+0.00045 \times$ E	297,824	0.85
E/P ≤ 2.0 or $E_T \geq 50$ GeV	296,029	0.84
-3.0 cm $\leq Q_e\ dx \leq$ 1.5 cm	296,010	0.84
$\lvert dz\rvert \leq 3.0$ cm	296,004	0.84
Strip $\chi^2 \leq 10$	295,964	0.84
$L_{shr} \leq 0.2$	295,945	0.84
Track segment	279,344	0.79
Isolation(0.4)	156,266	0.44
Conversion removal	87,646	0.25
W selection Cuts:		
Missing $E_T \geq 30$ GeV	31,729	0.09
Z removal (75 GeV $\leq M_{ee} \leq$ 105 GeV)	31,726	0.09
W candidates:	31,726	0.09

Table 3.2: Number of events in each event selection for the W $\to e\nu$.

amount of events, we lowered the electron E_T threshold which does not play a crucial role in the electron background rejection,

$$E_T \;\geq\; 20\,\text{GeV} \qquad , \qquad \not{E}_T \;\geq\; 30\,\text{GeV} \quad . \tag{3.14}$$

We show the projection of the electron isolation to the missing E_T after all the selection criteria except the isolation and missing E_T cuts in Figure 3.1. Around 30 GeV to 40 GeV in $\not{E}_T$, we can expect the minimum and flat electron isolation. This choice is justified by the fact that the corrections to the missing E_T are not yet fully understood and we preferred to sacrifice some selection efficiency keeping an high purity.

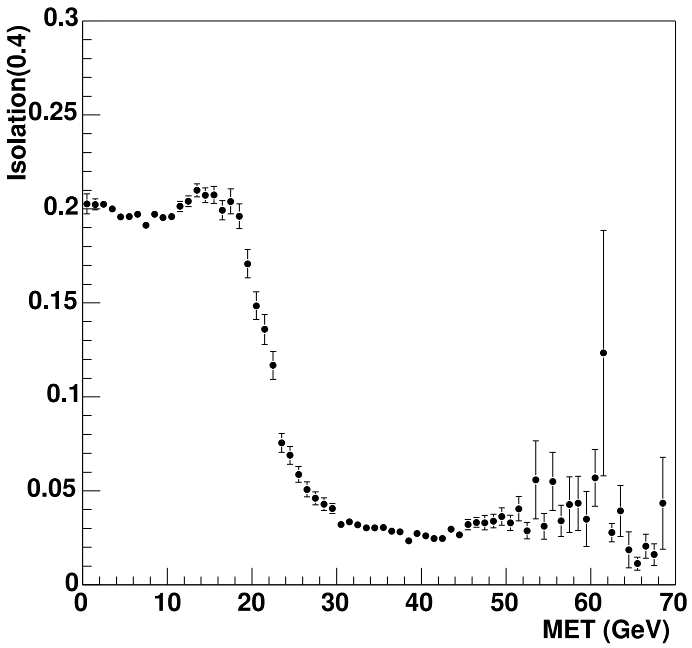


Figure 3.1: Projection of the electron isolation to the missing E_T after all the selection criteria except the isolation and missing E_T cuts.

3.7 Cross check of inclusive W production cross section

3.7.1 Cross section formula

It is practical to check the consistency with the inclusive W production cross section analysis. By using the number in Table 3.2, we can estimate the inclusive W production cross section. The cross section is expressed as

$$\sigma \cdot Br(W \rightarrow e\nu) = \frac{N_W - N_{BG}}{A_W \cdot \varepsilon_{Zvtx} \cdot \varepsilon_c \cdot \varepsilon_T \cdot \int \mathcal{L}dt} \quad , \tag{3.15}$$

where N_W is the number of $W \rightarrow e\nu$ candidate events and N_{BG} is the estimated number of background events. A_W is the kinematic and geometric acceptance taken from Monte Carlo simulation. ε_{Zvtx} is the efficiency of a Z vertex cut ($|Z_{vtx}| \leq 60$ cm). ε_c is the central electron identification efficiency, ε_T is the trigger efficiency and $\int \mathcal{L}dt$ is the integrated luminosity.

3.7.2 Trigger efficiency

In order to estimate the trigger efficiency used in this analysis, we use the so-called "backup" triggers. The backup triggers consist of the almost same trigger requirements except the quantity that should be estimated. Each trigger path has been tested separately and the correspondent efficiency has been calculated at each step [34].

The requirements for the high p_T electron trigger and those backup triggers are listed in Table 3.3. The XFT and CT tracking efficiencies[3] have been calculated using the W_NOTRACK trigger, which demands the same calorimeter requirements used for the high p_T electron trigger, but does not require tracks associated with the EM clusters. The track trigger efficiencies thus can be measured with events coming from this trigger. In order to select the $W \to e\nu$ events, we apply the baseline W selection cuts [34] to both triggers. The L1_XFT_PT8 trigger efficiency is measured by counting the W candidate events passing the L1_CEM8_PT8 trigger:

$$\varepsilon(L1\_XFT\_PT8) = \frac{\#\ of\ W\ candidates\ passing\ L1\_CEM8\_PT8}{\#\ of\ W\ candidates} \quad . \tag{3.16}$$

The efficiency curve is fitted as a function of p_T, and then the resulting trigger efficiency is almost 97%. The Level-3 tracking efficiency is measured by requiring the L1_CEM8_PT8 and L2_CEM16_PT8 bits set and counting the W events with the Level-3 Electron 18 bit set:

$$\varepsilon(L3\_PT9) = \frac{\#\ of\ W's\ passing\ L1\_CEM8\_PT8\ \&\ L2\_CEM16\_PT8\ \&\ L3\_CEM18\_PT9}{\#\ of\ W's\ passing\ L1\_CEM8\_PT8\ \&\ L2\_CEM16\_PT8} \ . \tag{3.17}$$

The tracking efficiency is almost 99 %. However the track trigger efficiency is independent of p_T in the region of $p_T \geq 10$ GeV, but dependent on η for both the Level-1 and Level-3

[3] Since Level-1 track trigger definitions have been changed twice over the period of data taking between July and September 2002, and between September 2002 and January 2003, the tracking efficiencies have been calculated separately.

triggers as shown in Figure 3.2.

The calorimeter trigger efficiency at Level-1 has been calculated using the high p_T inclusive muon sample which requires one tight CMU or CMX muon without an isolation cut. We select events which have CEM activity compatible with the L1_EM8 trigger by combining the calorimeter tower into the trigger tower geometry. The trigger efficiency is estimated by counting the events if the selected events fire the L1_EM8 trigger bit. The L2 calorimeter cluster trigger efficiency has been estimated using the W_NOTRACK_NOL2 trigger, which is the same as the W_NOTRACK trigger, but without any requirement at Level-2 (Level-2 Auto Accept). In order to get high purity electron sample, some quality cuts described in Section 3.2.5 are applied, but E_T cut lowered to 18 GeV and an isolated E_T is less than 4 GeV. The E_T distribution of electron candidates and the resulting L2_CEM16 trigger efficiency are shown in Figure 3.3. The efficiency curve is fitted as a function of E_T. The L2_CEM16 trigger fires about 96% of electrons around $E_T \sim 20$ GeV and the efficiency reaches up to 100% at larger than 25 GeV. At Level-3, since the electron clustering algorithm is almost identical to the one used in offline software, any inefficiency seen is due to differences between offline and trigger energy determinations. The overall scale factor for offline tower energy differs from Level 3 by a few percent depending on data-taking period. In addition, the tower-to-tower gain corrections and the mapping corrections within a tower are applied to offline EM clusters. The total corrections can be as large as 20%. The Level-3 calorimeter trigger efficiency, L3_CEM18, is also fitted, and it reaches 100% at $E_T \sim 23$ GeV.

Since the calorimeter trigger efficiency is essentially 100% at $E_T \sim 25$ GeV, the electron trigger efficiency is almost dominated by the tracking trigger efficiency. Finally, the estimated electron trigger efficiency is $\varepsilon = 96.6 \pm 0.1\%$ for the W candidate events.

3.7.3 Electron identification efficiency

The electron identification efficiency is estimated using the Z-sample where the invariant mass of at least one tight and one loose electron distributes within the mass range of 75

Trigger Name	Level 1	Level 2	Level 3
ELECTRON_CENTRAL_18	L1_CEM8_PT8 (CEM $E_T \geq 8$ GeV, & XFT $p_T \geq 8$ GeV, *etc.*)	L2_CEM16_PT8 (CEM $E_T \geq 16$ GeV, & XFT $p_T \geq 8$ GeV, *etc.*)	L3_ELECTRON_CENTRAL_18 (CEM $E_T \geq 18$ GeV, & COT $p_T \geq 9$ GeV, *etc.*)
W_NOTRACK	L1_EM8_&_MET15 (EM $E_T \geq 8$ GeV, & $\not\!\!E_T \geq 15$ GeV, *etc.*)	L2_CEM16_L1_MET15 (EM $E_T \geq 20$ GeV, *etc.*)	L3_W_NOTRACK_MET25 (CEM $E_T \geq 25$ GeV, & $\not\!\!E_T \geq 15$ GeV, *etc.*)
W_NOTRACK_NO_L2	L1_EM8_&_MET15 (EM $E_T \geq 8$ GeV, & $\not\!\!E_T \geq 15$ GeV, *etc.*)	Auto Accept Prescale 50	L3_W_NOTRACK_MET25 (CEM $E_T \geq 25$ GeV, & $\not\!\!E_T \geq 15$ GeV, *etc.*)
ELECTRON_CENTRAL_8	L1_CEM8_PT8 (CEM $E_T \geq 8$ GeV, & XFT $p_T \geq 8$ GeV, *etc.*)	L2_PS50_L1_CEM8_PT8_CES2_DPS Maximum prescale = 10 (CES $E_T \geq 2$ GeV)	L3_ELECTRON_CENTRAL_8_PT8 (CEM $E_T \geq 8$ GeV, & COT $p_T \geq 8$ GeV, *etc.*)

Table 3.3: The requirements for the high p_T electron trigger and those backup triggers.

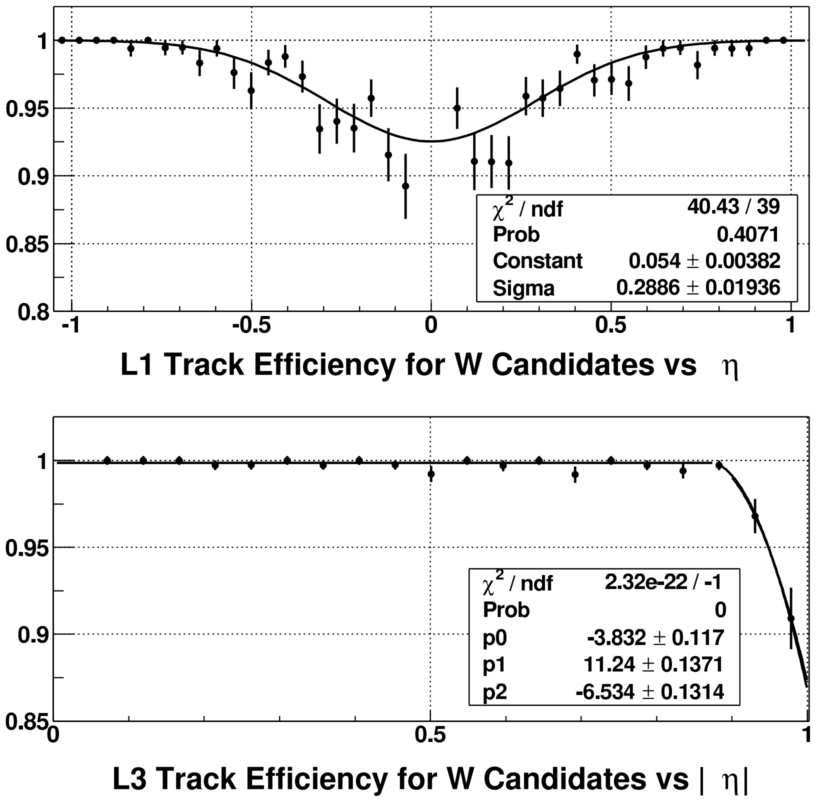


Figure 3.2: Level-1 (top) and Level-3 (bottom) tracking efficiencies as a function of η measured by CES.

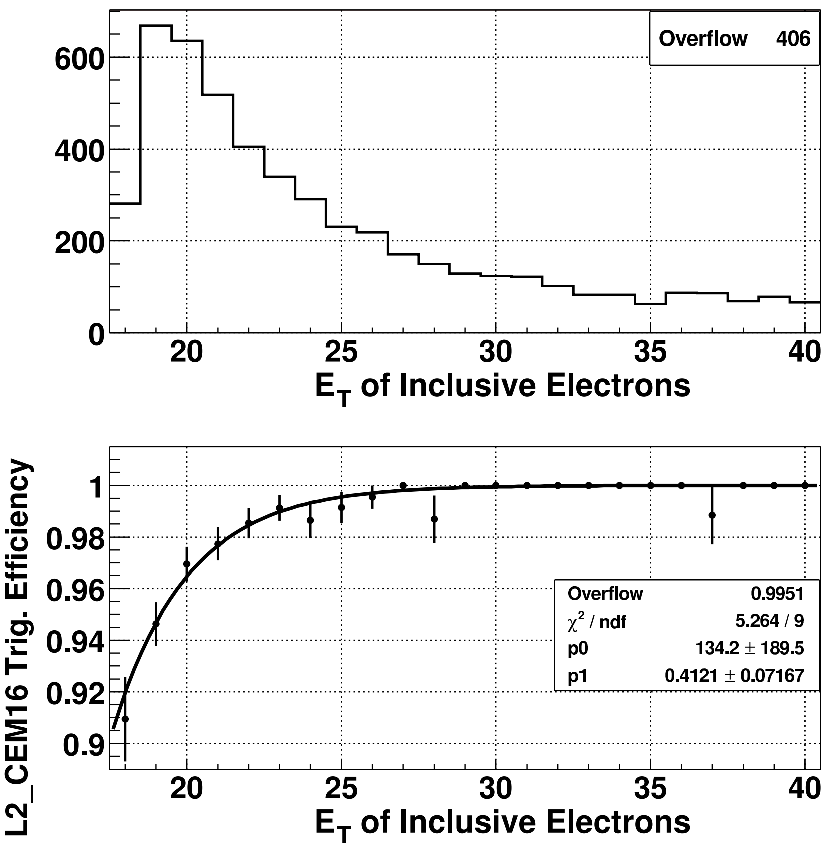


Figure 3.3: E_T distribution of electron candidates from data passing the L2 pre-scaled trigger and some quarity cuts (top) and L2_CEM16 trigger efficiency as a function of E_T (bottom).

GeV $\leq M_{ee} \leq$ 105 GeV, in the same procedure as described in [35], where the "loose" electron is a probe electron which measures the identification efficiency of each cut criteria. The loose electron is defined by the following cuts:

$$E_T \geq 20 \text{ GeV} \quad , \quad P_T \geq 10 \text{ GeV} \quad , and \quad |Z_{vtx}| \leq 60 \text{ cm} \quad , \tag{3.18}$$

where those electrons have an opposite charge to the first (tight) electron and point to the fiducial volume in the detector based on the electron tracks. Each identification efficiency

	Events (SS)	Efficiency $\varepsilon_c^{(i)}$ (%)
Number of Z	1880 (25)	—
Had/EM	1841 (15)	98.8 (70.6)
E/P	1657 (20)	93.1 (85.3)
$Q_e\ dx$	1837 (23)	98.7 (94.1)
$\|dz\|$	1857 (22)	99.3 (91.2)
Strip χ^2	1777 (19)	96.8 (82.4)
L_{shr}	1838 (17)	98.7 (76.5)
Track segment	1787 (23)	97.1 (94.1)
Isolation(0.4)	1777 (15)	96.8 (70.6)
Tight Electron	1329 (9)	82.8 (52.9)

Table 3.4: Electron identification efficiency for each identification criterion. The number in the bracket denotes the number for the same signed electrons.

for the second (loose) electron to the first (tight) one is defined as

$$\varepsilon_c^i = \frac{N_{Ti} + N_{TT}}{N_{CC} + N_{TT}} \quad , \tag{3.19}$$

where N_{CC} is the number of Z events within the invariant mass range, N_{TT} is the number of events where both legs pass the tight central cuts and N_{Ti} is the number of events where one leg passes the tight cuts and the other passes the i-th electron identification cut. For the total electron identification efficiency, the above equation becomes

$$\varepsilon_c = \frac{2\ N_{TT}}{N_{CC} + N_{TT}} \quad , \tag{3.20}$$

when both electron legs pass the tight electron cuts. The results are shown in Table 3.4. The number in the bracket denotes the number for the same signed electrons. Finally, the tight electron identification efficiency is 82.8 ± 0.9%.

3.7.4 Z vertex efficiency

The position of the interaction vertex is required to be $|Z_{vtx}| \leq 60$ cm. Some of events are lost outside Z vertex position due to the Tevatron beam size. The luminosity correction for the $|Z_{vtx}| \leq 60$ cm for the W candidate events is estimated via the $p\bar{p}$ beam luminosity function [36]

$$\frac{d\mathcal{L}}{dz} = N_0 \frac{\exp(-z^2/2\sigma_z^2)}{1 + (\frac{z - z_{min}}{\beta^*})^2} \quad , \tag{3.21}$$

where z is the Z_{vtx}, N_0 is a normalization and the rest are the beam parameters. From the minimum bias data with a wide range of the luminosity in this analysis, we obtain

$$\varepsilon_{Zvtx} = 95.1 \pm 0.1(stat.) \pm 0.5(syst.)\,\% \quad . \tag{3.22}$$

The details of this measurement are described in [37].

3.7.5 Acceptance calculation

The efficiency for geometric and kinematic restrictions on the leptons is referred to as the acceptance. The geometric and kinematic acceptances are calculated separately using a Monte Carlo simulation. The HERWIG $W \to e\nu$ event generator is used to estimate both geometric and kinematic acceptances. The geometric acceptance is the fraction of electrons that deposit energy in a fiducial region of the central electromagnetic calorimeter. The kinematic acceptance is the fraction of electrons and neutrinos to pass the E_T (≥ 20 GeV) and $\not{E}_T$ (≥ 30 GeV) requirements respectively. We obtain these fractions from the simulation. The calculated acceptance is

$$A_W = \varepsilon_{geo.} \cdot \varepsilon_{kin.} = 21.9 \pm 0.14\,\% \quad . \tag{3.23}$$

3.7.6 Background estimation

QCD background :

The most significant $W \to e\nu$ background is the direct QCD production of multijet events. In some QCD multijet events, a jet mimics the signature of an electron, and mismeasured transverse energy results in a large $\not{E}_T$. The semileptonic decay into the electrons from heavy flavor quarks is also considered as the QCD background. The QCD hadronic jets can lead to background to the W signal if the hadron jet containing the electron spreads out so that the electron is isolated in the calorimeters, and if the other jet falls into an uninstrumented region of the detector and then creates a large $\not\!\!E_T$.

These backgrounds have been estimated by using the relation of the electron isolation variable versus the missing transverse energy with the assumption that the electron isolation is independent of the $\not\!\!E_T$ [38]. First of all, we need to obtain the enriched QCD sample. The QCD jets presumably disturbs the electron isolation variable because hadrons in a jet distribute around the semileptonic decayed electron. And a large $\not\!\!E_T$ rejects a large amount of QCD background. Therefore to obtain a QCD background sample, we remove these requirements from our W selection criteria but keep the other selection criteria to W boson. Then, in order to estimate the amount of QCD background in the W sample, we divide this sample into four regions:

Region A : Isolation ≤ 0.1; $\not\!\!E_T \leq 10$ GeV,
Region B : Isolation ≥ 0.3; $\not\!\!E_T \leq 10$ GeV,
Region C : Isolation ≥ 0.3; $\not\!\!E_T \geq 30$ GeV,
Region D : Isolation ≤ 0.1; $\not\!\!E_T \geq 30$ GeV.

We show the isolation-$\not\!\!E_T$ plane with the four regions divided by the above definitions in Figure 3.4. In Figure 3.4, our W events contribute in Region D. From the definitions of the regions, one sees that we have excluded intermediate regions from consideration. This exclusion is to insure that regions A, B, and C are pure multijet and do not include W events. This requirement rejects $W \rightarrow e\nu$ leakage as well as $W \rightarrow \tau\nu$ events which have an average $\not\!\!E_T$ less than $W \rightarrow e\nu$ events but generally larger than 10 GeV. From the assumption of the linear relation of the electron isolation and $\not\!\!E_T$, we directoly count the number of multijet events in W events (Region D):

$$N_{QCD} = \frac{N_A}{N_B} \cdot N_C \quad , \tag{3.24}$$

where N_{QCD} is the number of QCD events in the W sample, and N_A, N_B, and N_C correspond to the number of events in Region A, B, and C, respectively. If the ratio (N_A/N_B) of well-isolated to poorly isolated QCD events is known in the low $\not{E}_T$ region then it is known in the high $\not{E}_T$ region. We present the number of events in each region and the number of the estimated QCD background events in the W sample in Table 3.5. The number in each column corresponds to the number of observed events in each region, respectively. We estimate 385 events as the QCD background events.

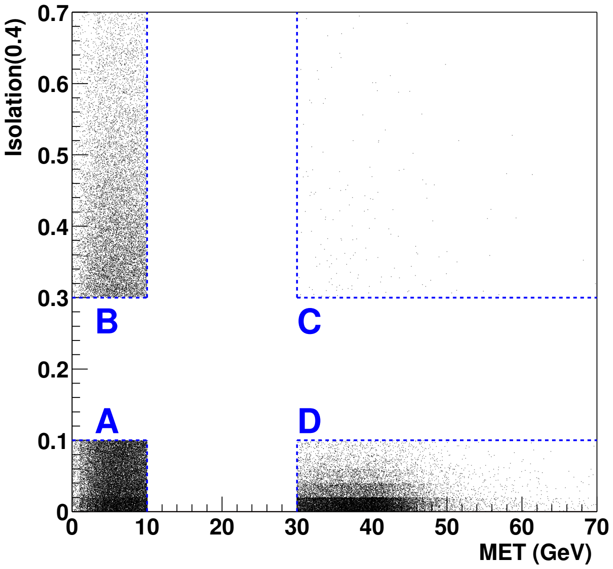


Figure 3.4: Isolation vs. $\not{E}_T$

Region A (N_A)	Region B (N_B)	Region C (N_C)	QCD bkg. (N_{QCD})
29237	15566	205	385

Table 3.5: QCD background estimates in the W sample. First three columns show the number of events in each region A, B, and C, respectively. The number of QCD events are shown in the last column.

Single boson background :

Other physics processes that contribute to our W event sample are $W \to \tau\nu$, and $Z \to e^+e^-$. The production cross section for $W \to \tau\nu$ is identical to that of $W \to e\nu$, and the τ lepton decays to an electron with a branching fraction of 18%. Fortunately, this potentially serious background is substantially reduced by the large $\not{E}_T$ and the electron E_T thresholds. In $Z \to e^+e^-$ events, a large $\not{E}_T$ can be observed if an electron is mismeasured or escapes through an uninstrumented part of the detector. We neglect the contribution of a $Z \to \tau^+\tau^-$ events because the production cross section multiplied by the branching ratio of a $\tau \to e\nu\nu$ is reduced down to 1 % of the production cross section of a $W \to e\nu$.

For the $W \to \tau\nu$ background, we extract the ratio

$$R_{W\to\tau\nu} = \frac{\sigma(W \to \tau\nu)\varepsilon(W \to \tau\nu)}{\sigma(W \to e\nu)\varepsilon(W \to e\nu)} \quad , \tag{3.25}$$

where the ε is the efficiency for finding a W boson which is dependent on the decay mode. In fact, since the production diagrams is equivalent to that of an electron decay mode, we prefer to use this ratio rather than extracting an absolute prediction of $W \to \tau\nu$ in order to remove the renormalization scale dependence inherent in LO QCD predictions. The same procedure is applied for the $Z \to e^+e^-$ background,

$$R_{Z\to e^+e^-} = \frac{\sigma(Z \to e^+e^-)\varepsilon(Z \to e^+e^-)}{\sigma(W \to e\nu)\varepsilon(W \to e\nu)} \quad . \tag{3.26}$$

To estimate these numbers, we use the HERWIG Monte Carlo event generator. We show the numbers from the result of the simulations for these processes:

$$R_{W\to\tau\nu} = 0.0167 \pm 0.0009 \quad , \quad R_{Z\to e^+e^-} = 0.0107 \pm 0.0002 \quad . \tag{3.27}$$

From these ratios, the number of the single boson background events is

$$N_{W\to\tau\nu} = 306 \pm 16(stat.) \quad , \quad N_{Z\to e^+e^-} = 196 \pm 5(stat.) \quad . \tag{3.28}$$

Top background :

The $t\bar{t}$ production is also the other source of background, in which each top quark decays almost exclusively to a W boson and b-quark. Although our W samples are not required to contain b-quarks, the fraction of top events is expected to be significant in the subsamples with a high number of jets. Since our W data selection requires an electron and neutrino, one of the W's from top pair decay is constrained to this decay mode. The other W can decay in any mode but it is the hadronic decay ($W \rightarrow q\bar{q}' \rightarrow$ hadrons) that introduces the largest contamination of our direct single W candidate sample. The top background estimate is derived from a top Monte Carlo sample made by using the HERWIG top event generator with all decay modes allowed and a top mass of 175 GeV followed by a full detector simulation. The W bosons from the top decays are allowed to decay to any final state to obtain every possible background event. In order to extract top background events in our W sample, we use the theoretical top cross section [39] rather than the measured top cross section [40], because we have not established the top cross measurement in Run II experiment yet and we are trying to compare the experimental results to QCD calculations. The top mass is chosen to be 175 GeV [41]. The number of the top contribution in our W event sample is thus calculated as

$$N_{top} \; = \; \frac{N_{top}^{acc}}{N_{top}^{gen}} \, \sigma_{t\bar{t}} \; \times \; 72.0\text{pb}^{-1} \; = \; 33.0 \, \pm \, 0.6(stat.) \quad , \tag{3.29}$$

where $\sigma_{t\bar{t}} = 6.70^{+0.71}_{-0.88}$ pb, N_{top}^{gen} (40,000 events) is the generated events by HERWIG event generator, and N_{top}^{acc} (2,740 events), of these, is the number of events satisfying our W selection.

3.7.7 Inclusive $W \rightarrow e\nu$ production cross section

After rough estimation of the electron identification efficiency, trigger efficiency, acceptance, and background events, we obtain the inclusive W production cross section. The cross section formula is given in Eq.(3.15):

$$\sigma \cdot Br(W \to e\nu) \;=\; \frac{N_W \;-\; N_{BG}}{A_W \cdot \varepsilon_{Zvtx} \cdot \varepsilon_c \cdot \varepsilon_T \cdot \int \mathcal{L}dt} \;=\; 2.57 \pm 0.04 \text{ (nb)} \quad ,$$

where,

$$\begin{aligned} N_W &= 31,726 \pm 178 \quad , \\ N_{BG} &= 920 \pm 30 \quad , \\ A_W &= 21.9 \pm 0.14\,\% \quad , \\ \varepsilon_{Zvtx} &= 95.1 \pm 0.1 \quad , \\ \varepsilon_c &= 82.8 \pm 0.9\% \quad , \\ \varepsilon_T &= 96.6 \pm 0.1\,\% \quad , \\ \text{and } \int \mathcal{L}dt &= 72.0 \pm 4.3 \text{ pb}^{-1} \quad . \end{aligned}$$

Errors are all statistics only except the integrated luminosity. The result is consistent with the inclusive W cross section analysis of 2.67 ± 0.30 nb [33].

3.8 Summary

In this chapter, we have described the selection criteria for the $W(\to e\nu)$ candidate events, and measurement of the inclusive $W(\to e\nu)$ production cross section.

After the selection cuts of W boson, we got 31,726 $W(\to e\nu)$ candidate events. The JetClu algorithm is then used to collect/count jets. The jet energy correction is applied up to the partonic level correction which estimates a parton energy based on a certain simulation, although this energy correction causes large systematics. Since we use a slightly higher $\not\!\!E_T$ cut and lower electron E_T cut to reject a multijet background, the inclusive $W(\to e\nu)$ production cross section was compared with the result of the W production cross section measurement in CDF Run II. We got a good agreement with both results.

Chapter 4

Jet Separation Study

4.1 Introduction

The purpose of a jet identification algorithm is to identify the original[1] parton from the hard scattering process. In our theoretical picture, the partons produced by the hard scattering process evolve approximately within a narrow cone based on the parton showering and hadronization models. A bunch of hadrons are detected at the detector afterwards. The cone algorithm, which is widely used in hadron collider experiments, will be a good approximation in that the hadron bunch distributes within a certain cone as the observable object. In a sense, JetClu algorithm has been the representative algorithm [51]. The recent development of cone algorithms, however, suggests some improvements in order to solve undesirable feature of QCD physics known as the infrared/collinear problem.

One may want the primary parton information from the observed jet. For instance, the top-quark mass analysis requires to measure the kinematical distribution of top-quarks using observed jets. In that case, it is crucial to have a good matching between jets and the primary partons. However, due to incompleteness of the modeling of physical events, the naive treatments can lead to many ambiguities, mainly known as the infrared/collinear

[1] The "original" does not make sense for a physical meaning, here. For example, we would like to mention the outgoing partons from the Matrix Element perturbatively calculated as the core part of the event generator.

singularities. If we require a multi-jet final state, we have a trouble of the "double counting" problem. To avoid this problem, some prescriptions [43] [44] [45] are proposed as a possible solution. The easiest prescription is to look at the phase space dominated by the perturbative calculation of the ME. The pre-clustering of jets at the generator level may work to achieve it [46].

Besides testing those prescriptions, we have to know the performance of the jet clustering algorithms including the detector simulation, not the generator level. Since the jet multiplicity strongly depends on the requirement of the infrared/collinear safety of jet clustering algorithms, it is very important to estimate how well the jet separation works in the cone algorithms. The well-separated jets allow us to see the phase space truncated by the ME dominant region, that is, n partons(cones) are reconstructed as n jets ($n =$ 2,3,4,...). That would enable us to understand the inclusive/exclusive prediction of the jet production processes. The merging of jets should be studied to understand higher order perturbative calculations and the limited ability of the cone jet algorithm.

We have studied the dependence to these infrared/collinear effects of the cone jet algorithms by using a toy Monte Carlo (MC) event generation program. The toy MC program makes two partons with particular momenta and directions. By changing each momentum and separation angle between two partons, the situation of the infrared/collinear region is reproduced. After a hadronization by HERWIG [47] or PYTHIA [48], those events pass through the detector simulation, and then the jet reconstructions are performed. We estimated the separation angle between two reconstructed jets by the cone jet algorithms, where the two partons are reconstructed as the two jets. We are not going to discuss the contribution of the infrared/collinear partons produced by the parton showering in the reconstructed jets, because we only realize the physical meanings when the partons form a bunch of hadrons. We thus keep this study to the lowest order perturbative calculations, while the higher order contribution has the attractive feature of the jet shape. In this paper, we present the well-separated region reconstructed as two jets depending on the jet transverse energies. Note that although the k_T-based clustering [49] would be theoretically better description as long as the parton evolution is based on their k_t generation, we focused on the cone jet algorithms.

This paper is organized as follows: the brief review about the infrared/collinear problems in the cone jet algorithms is described in Section 2. The details of the toy MC are described in Section 3. The results of the separation power of JetClu algorithm are presented in Section 4. More discussions about the algorithm dependency are given in Section 5. A summary is given in Section 6.

4.2 Cone jet algorithms

4.2.1 Jet definitions

A jet coordinate geometry[2] of the JetClu algorithm is already described in Section 3.

In a perturvative calculation of the Matrix Element (ME), the outgoing partons from ME must be observed as jets [50]. For example, look at NLO calculations. The virtual(loop) correction can lead to the cancellation with the soft/collinear enhancement in a certain phase space. In the particle picture, the collinearly emitted parton is presumably absorbed into that parent parton within a certain cone. In order to provide the appropriate jet cross section from the NLO calculation, this invisible parton should not be isolated as the jet. As the same argument, the real emitted parton, in the external line of the Feynman diagrams, should be detected as a jet. Hence we require the additional definition,

$$MC \in J \quad : \quad \sqrt{(y^{MC} - y^{J})^2 + (\phi^{MC} - \phi^{J})^2} \;\leq\; R \quad . \tag{4.1}$$

for the visible outgoing partons.

4.2.2 The infrared/collinear problems

In this section, we review the problematic issues for the cone jet clustering in brief. All details are described in [51]. Theoretical requirement for jet algorithms is that algorithms

[2] In CDF RunII, *E-Scheme* [51] is taken for the other jet clustering algorithms. We however use Run I-based coordinate system in this analysis since the JetClu has been used in Run I.

should not be affected by the infrared and collinear singularities. Algorithms should be insensitive rather than sensitive to the size of the cluster energy. No one looks at a single pion as the jet. Jet should be defined as a bunch of energetic clusters on the calorimeters.

The infrared problems are caused by the presence of soft radiation in the event. The schematic view is illustrated in Figure 4.1. The length of the arrow represents the size of the momentum. In the picture (b) in the right-hand side, soft radiation (solid wave) is emitted between two particles (solid line). These illustrations show that the presence of soft radiation disturbs two jet clusterings in (b), while in the case of no radiation of (a), two jets are reconstructed. In (b), the seed tower appears in this soft radiation regardless of the small momentum.

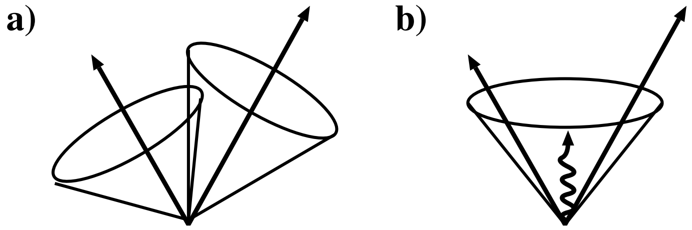


Figure 4.1: An illustration of infrared sensitivity in cone jet clustering. The length of the arrow represents the size of the momentum. In addition to the two partons emission (solid line) in (a), soft radiation (solid wave) is emitted between two particles in (b).

The schematic view of the collinear singularity is illustrated in Figure 4.2. Seed-based algorithms have the collinear problem in general until jets have so large E_T that the sharing of the seed energy between towers does not affect the jet finding. The seed tower threshold is tuned so that the jet reconstruction is 100 % above a certain jet E_T. Otherwise the algorithm without a seed tower is needed for avoiding the collinear singularity.

Another possible collinear problem can arise if the algorithm is sensitive to the E_T ordering of particles. The schematic picture is also illustrated in Figure 4.3. The difference between the two situations is that the central (hardest) particle splits into two almost collinear particles. The separation between the two most distant particles is more than R but less than $2R$. Thus all of the particles can fall within a single cone of radius R around

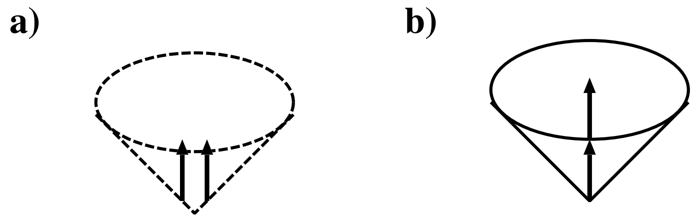


Figure 4.2: An illustration of collinear sensitivity in jet reconstruction. The jet reconstruction fails while jet in (a) is reconstructed in (b).

the central particle as shown in Figure 4.3 (a). However, if the seed tower drifts according to the magnitude of E_T, in the case that the splitting happens in the largest E_T particle, a jet may be found around the second largest particle as shown in Figure 4.3 (b). A jet is reconstructed with the seed of the right-most particle. The left-most particle is a jet by itself. From these situations, the number of jets depends on whether the collinear splitting exists. Clearly, this problem should be avoided by making the selection or ordering of seeds and jet cones independent of the E_T of individual particles.

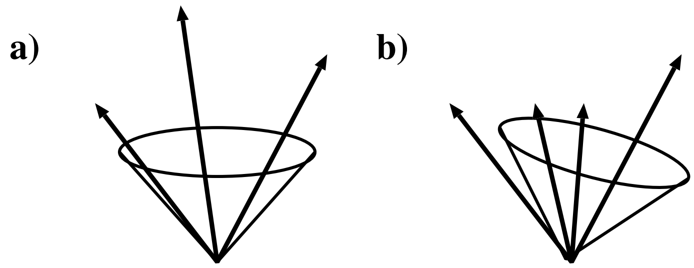


Figure 4.3: Another example of the collinear sensitivity in jet reconstruction. If the algorithm is sensitive to the E_T ordering of particles, this problem will happen. In (a), a jet is successfully reconstructed with the seed tower of the largest E_T. In (b), two jets are reconstructed with the seeds of the left-most particle and right-most particle, when the largest E_T particle splits into two lower E_T particles.

4.3 Fake jets event generator

To see how well the jet separation works in the cone algorithms, we have developed the fake jet event generator. That generator makes two partons with particular momenta and directions. Then, those partons are applied to the parton showering and hadronization of HERWIG or PYTHIA. The development was done in the framework of user interface of `LesHouches` package [52] in CDF offline code. In this section, we describe the details about the generation of two partons. The generation procedures are the followings:

step 1) Decide the flavor combinations of two outgoing partons with uniform weight. The possible assignment of the parton flavor is *gluon*, $u(\bar{u})$-, and $d(\bar{d})$-*quark*.

2) Assign the initial partons according to the outgoing partons decided in the step 1). If both outgoing partons are same flavor, the $u(\bar{u})$- or $d(\bar{d})$-*quark* pair is assigned as the initial partons. The (anti-)quark always comes from (anti-)proton.

3) Set the kinematics of the first parton given in the input transverse momentum, and direction (p_T, ϕ, η). Then, the second parton are set to distribute along with the given transverse momentum and separation angle, ΔR $(=\sqrt{(\Delta\phi_{12})^2 + (\Delta\eta_{12})^2})$.

4) To keep the momentum conservation rule, the third parton is set with their opposite direction. The flavor of this parton is neutrino or graviton[3] depending on the electric charge of the outgoing partons. The third parton is never detected at the detector.

5) Decide color configurations based on the generic QCD 2-2 process. If the color configurations are considered as two different combinations, the configurations are taken account of the weight of s- or t-channel kinematics of two outgoing partons.

[3]Current HERWIG 6.4 is not capable to treat graviton. We assign the third parton as neutrino in spite of the violation of the electric charge. But it does not affect any generation scheme as far as the third parton is stable and undetected particle.

All the kinematics of two outgoing partons are decided by the input, (p_{t1}, ϕ_1, η_1) for the first parton, and (p_{t2}, ΔR_{12}) for the second parton, respectively, where the ΔR_{12} is defined as the squared summation of each difference of ϕ and η for both partons ($=\sqrt{(\Delta\phi_{12})^2 + (\Delta\eta_{12})^2}$). We demonstrate the parton separation in this generator in Figure 4.4. The first parton is produced in the fixed direction by the input direction, (ϕ,η)=(π/3,-0.5) in this figure. The second parton surrounds the first parton with a particular separation angle along a half-circle so as to distribute in the central region of the calorimeters. From the top-left side, the size of the separation is varied with 0.1 step. The range of the separation angle is taken from 0.1 to 0.9. After a hadronization by HERWIG/PYTHIA, those events pass through the detector simulation, and then the jet reconstructions are performed. By changing each momentum and separation angle between two partons, the situation of the infrared/collinear region is reproduced.

4.4 The separation power of JetClu algorithm

4.4.1 Template data sets

We made data sets with various separation angle and momentum of each parton. Due to the different segmentation and the different response between the central and forward calorimeters, both partons were only produced in the central region at first. The first parton was always produced in the same direction, (ϕ,η)=(π/3,-0.5). The second parton was distributed with a given separation angle to the first parton. To avoid the crack region of the calorimeters, the second parton was actually generated within $0.1 < |\eta| < 0.7$. HERWIG was used as the parton showering and hadronization model. Hadronization scale is 1.26 GeV. All other parameters of detector simulation were set to the recommended default parameters. Also, the same manner estimated in the central region is exactly applied for the plug region. We present only the results of the central region to avoid the same discussion. The clusering cone size of 0.4 is used.

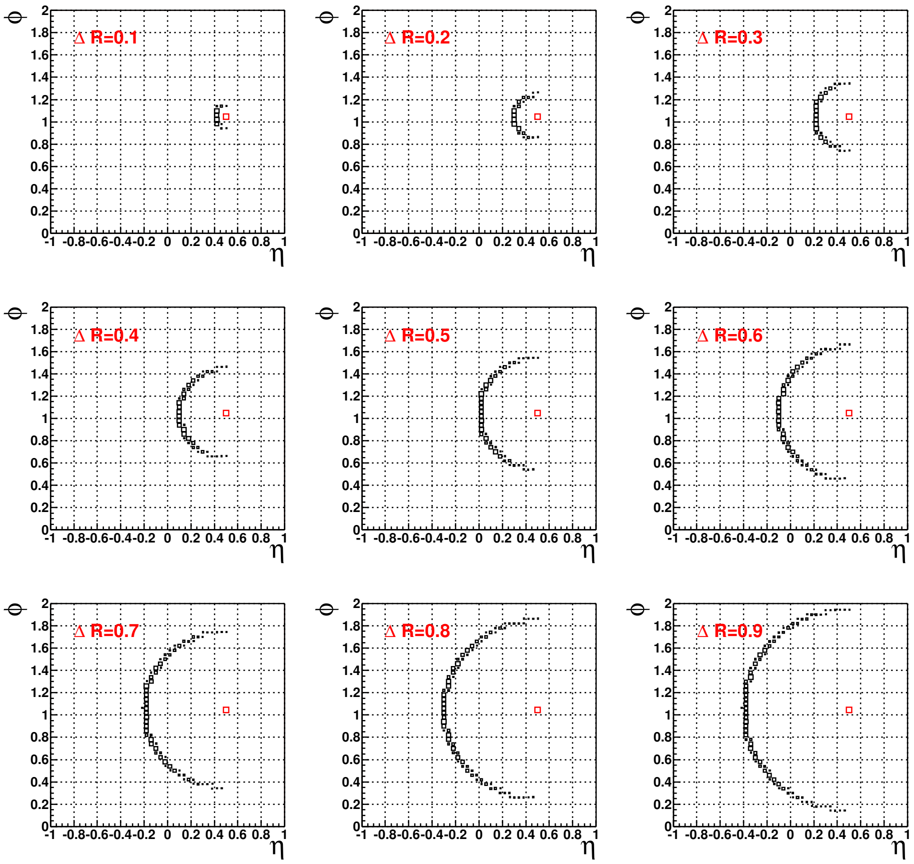


Figure 4.4: Demonstration of the two parton generation with the particular directions in ϕ-η plane. The first parton is forced to be the fixed direction by the input direction, $(\phi,\eta)=(\pi/3,-0.5)$. The second parton surrounds the first parton with the particular separation angle. The half-circle is taken to distribute in the central region of the calorimeters. From the top-left side, the size of the separation is varied with 0.1 step.

4.4.2 E_T distributions

From the template data sets, we show the reconstructed E_T distribution without jet energy corrections in Figure 4.5 for the leading jet, and in Figure 4.6 for the secondary jet. From top-left side, the separation angle, ΔR, between two generated-partons is changed with 0.1 step up to 0.9 of the separation angle. Both transverse momenta of the partons are 30 GeV in this sample. Those figures show the transition behavior of the jet E_T depending on the separation angle.

While the two partons are almost merged as the single jet in the small ΔR region, which results in the larger E_T than the generated p_T of each parton as seen in Figure 4.5, two partons are reconstructed as two jets with increasing of the separation angle in the large ΔR region. We can see two peaks around $\Delta R \sim 0.5$ in Figure 4.5.

Those effects reveal the evidence of the infrared/collinear problems. When the two partons are generated in the collinear region, the algorithm should not be sensitive to the collinear partons. Hence, this algorithm successfully works for the collinear partons since jets should be merged as the single jet. However, in the transition region arround $\Delta R \sim 0.5$, some events make two jets and the others are one jet. Unfortunately, as we cannot specify that these effects come from the soft radiation or E_T ordering of the seed tower, we conclude those are the infrared/collinear problems of the algorithm.

4.4.3 Reconstruction ratio

It is practical to see the ratio of the number of the reconstructed jets. We present the probability of jets acceptance as a function of the separation angle between two partons in Figure 4.7. Each opened- and closed-circle is the probabilities of the single jet and two jets reconstruction, respectively. The squared represents the no-jet reconstruction probability. The minimum threshold energy of jet clustering is 3 GeV. Furthermore, the jet is required to have one of the primary partons inside the jet, cone size of JetClu. Thus, the null-jet probability is caused by these both effects. In this sample, both primary parton p_Ts are 30 GeV.

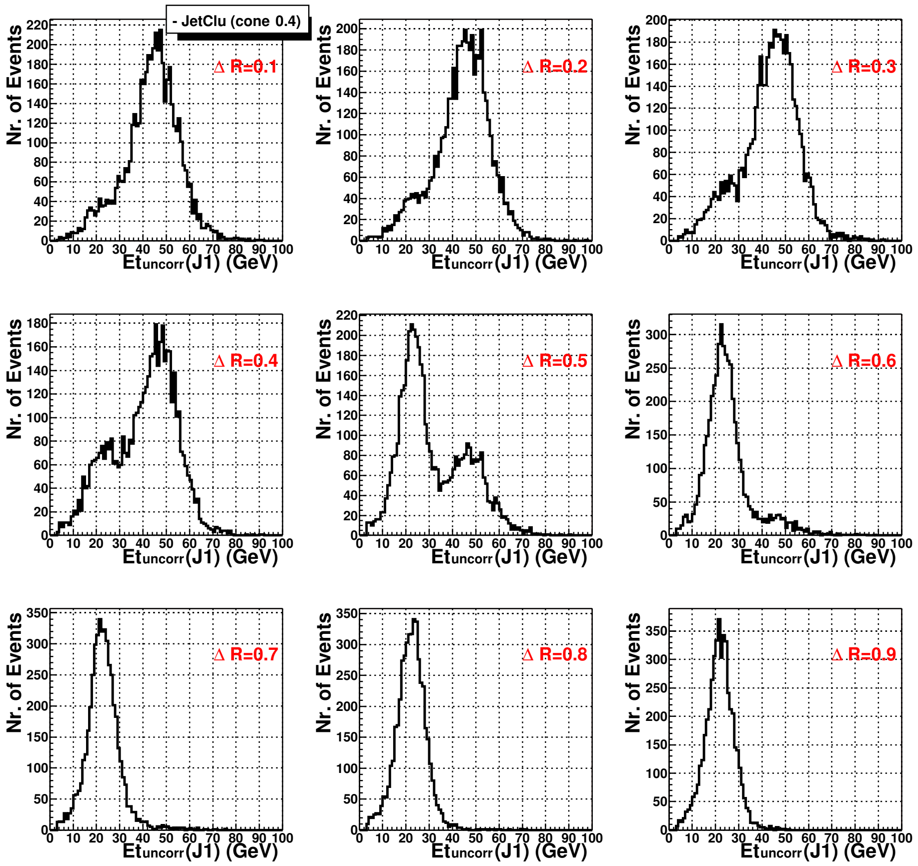


Figure 4.5: Leading jet E_T distribution reconstructed by JetClu with the cone size of 0.4. The energy correction is not applied. From top-left side, the separation angle, ΔR, between two generated-partons is changed with 0.1 step up to 0.9 of the separation angle. In this sample, both partons p_T are 30 GeV.

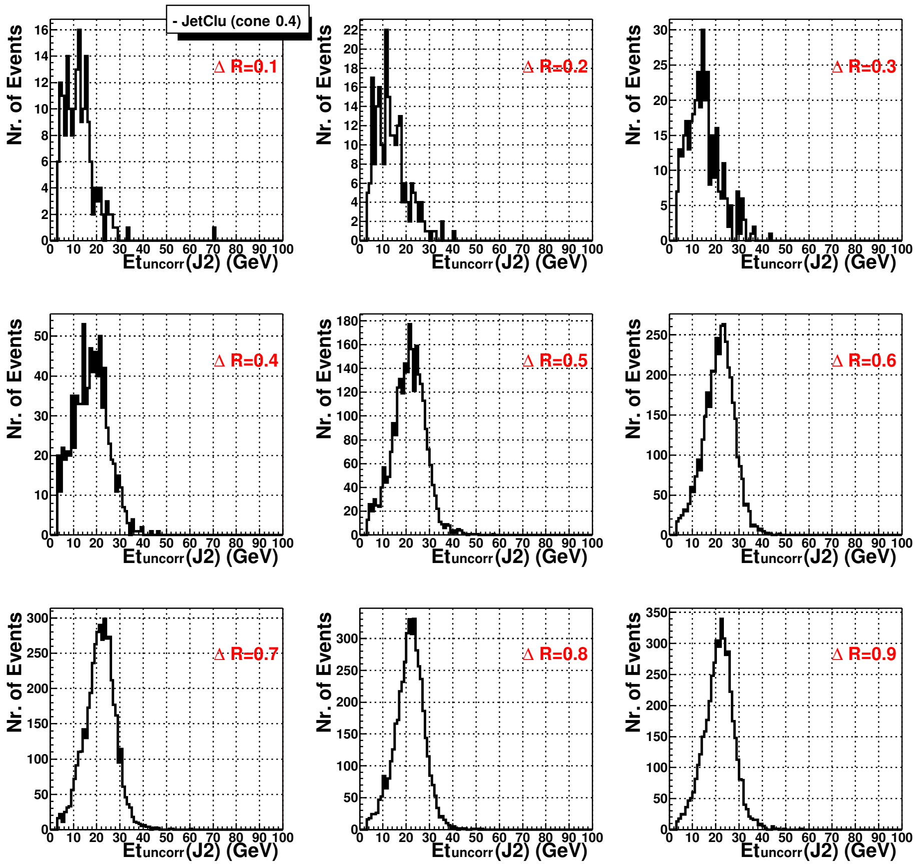


Figure 4.6: Secondary jet E_T distribution reconstructed by JetClu with the cone size of 0.4. The energy correction is not applied. From top-left side, the separation angle, ΔR, between two generated-partons is changed with 0.1 step up to 0.9 of the separation angle. In this sample, both parton p_T are 30 GeV.

As we expect, the transition behavior of JetClu depends on the separation angle between two partons. The well-separated jets become dominant in $\Delta R \gtrsim 0.8$. And the JetClu is free from the collinear problem in $\Delta R \lesssim 0.2$, although to see the collinear region is experimentally less meaningful. In the transition region, there are the mixed states in which events have one or two reconstructed jets. The soft radiation or E_T ordering of the seed tower degrades the jet reconstruction. JetClu is not capable of treating the infrared/collinear problems in this region.

To determine the quantity of the jet separation, we fit the shape of the transition behavior of 2 jets acceptance with the function of

$$f(x) \ = \ p_0 \cdot \tanh(p_1(x - p_2)) \ + \ p_3 \qquad , \tag{4.2}$$

where p_0, p_1, p_2, and p_3 are constants. The fitting function was adopted not because of the fitting χ^2, presumably, but because of simple formula and stability of plateau curve in the large ΔR region. When the hadron fraction from a primary parton results in the additional jets at the minimum threshold (3 GeV) level, the cone in the jet reconstruction may also cover the direction of the another primary parton. In that case, by the above requirement of the jet, we can reconstruct them as two jets. The constant term p_3 was introduced for this persistent fraction of jets. A sample plot of the fitting result is shown in Figure 4.8. The fittings are applied to all data sets. We present the separation angle at 95%, 50%, and 5% two jet acceptance level in Table 4.1.

4.4.4 Parton-jet matching

At the parton level, the jet is a cluster of partons with non-zero color charge. At the hadron level, the cluster is composed of color-singlet hadrons. The transition between the two levels necessarily involves the addition (or subtraction) of at least one colored parton carrying some amount of energy. Thus, the geometry-based matching criteria described in Section 2 would be a better description than the energy-based one. We show the deviations of the corrected jet energy and the observed jet separation from those of the primary parton as a function of the separation angle between two partons in Figure 4.9

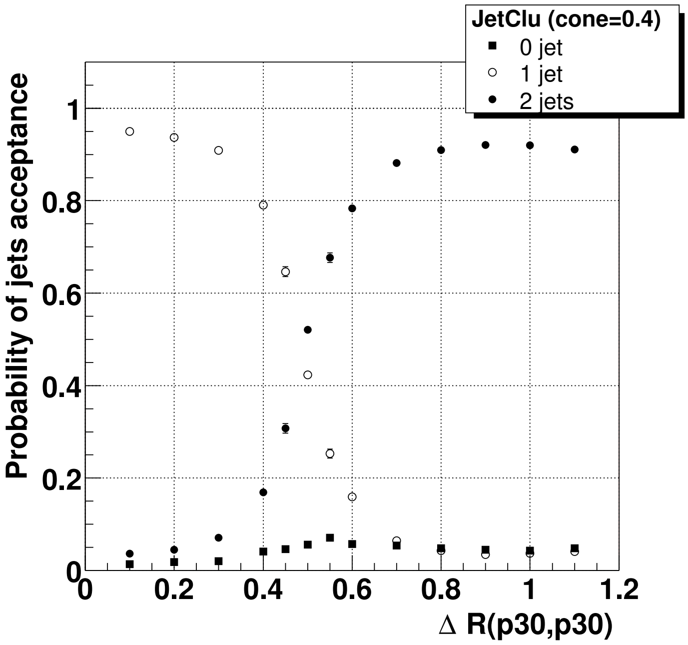


Figure 4.7: Probability of jets acceptance as a function of the separation angle between two partons. Each opened- and closed-circle is the probabilities of the single jet and two jets reconstruction, respectively. The squared represents the no-jet reconstruction probability. The minimum threshold energy of jet clustering is 3 GeV. Furthermore, the jet is required that the either primary partons exists inside the jet, which has the cone size of JetClu. In this sample, both primary partons p_T are 30 GeV.

and Figure 4.10, respectively. The mean value and that RMS are taken as the deviation of the corrected jet energy from the primary parton energy in this case of Figure 4.9. On the other hand, the central value and that sigma fitted by the Gaussian distribution with the range of the 25% height on the maximum value are taken as the deviation of the observed jet direction. In the well-separated region of the large ΔR, the reconstructed jet again describes the primary partons well, while we can see the big discrepancy in small ΔR region.

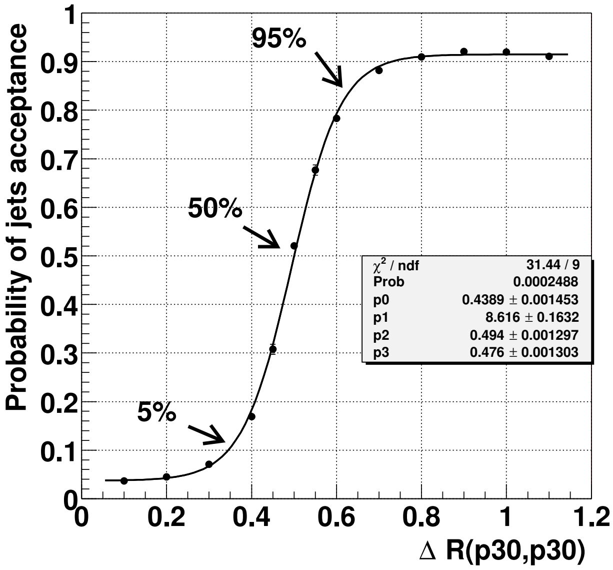


Figure 4.8: Sample plot of the fitting result for two jets acceptance. Both primary partons p_T are 30 GeV.

4.4.5 Jet separation

Using the well-separated region at more than 95% level in Table 4.1, we can get the weighted mean and deviation for the corrected jet energy and the observed separation angle as listed in Table 4.2, where the Gaussian fitting is also applied to the corrected jet energy instead of the RMS value. Those well-defined values clearly point the primary partons. In order to redefine the jet separation as an observable object of JetClu, we assign these obtained values as the resolution of the parton-jet matching. Finally, we present the phase space of the jet separation as a function of the jet separation angle in Figure 4.11. The two-jets reconstruction region is the region where the reconstructed jets point to the primary partons with the probability of 95%. The border line with the 95% probability of the two-jets reconstruction is obtained by the fitting functions of

$$R_{sep}^{95\%}(JETCLU; R_{cone}(0.4); E_{t1}, E_{t2}; CC) = \frac{23.2(1)^{+4.9}_{-3.2}}{min(E_{t1}, E_{t2}) + 56.5(6)^{-9.2}_{+28.9}} + R_{cone}, \tag{4.3}$$

parton p_T(GeV)		2 jets separation point, R_{sep}		
p_{t1}	p_{t2}	5 %	50 %	95 %
20	10	0.093	0.514	0.758
	15	0.120	0.502	0.722
	20	0.141	0.497	0.703
30	10	0.094	0.526	0.777
	15	0.160	0.506	0.728
	20	0.184	0.497	0.712
	30	0.225	0.489	0.662
40	10	0.139	0.527	0.794
	15	0.140	0.511	0.738
	20	0.212	0.501	0.707
	30	0.234	0.489	0.666
	40	0.236	0.483	0.643
50	40	0.249	0.479	0.640
	50	0.270	0.478	0.630

Table 4.1: The fitting results of two jets acceptance. The transverse momentum of the first parton is always larger than one of the second parton. The values in third to fifth column show the separation angle at 5%, 50%, and 95% level with respect to the upper asymptotic limit of the fitting function.

$$R_{sep}^{95\%}(JETCLU; R_{cone}(0.4); E_{t1}, E_{t2}; CP) \;=\; \frac{30.9(4)^{+3.7}_{-7.1}}{min(E_{t1}, E_{t2}) + 87(1)^{-20.8}_{+108.6}} + R_{cone}, \quad (4.4)$$

for the central-central combination and the central-plug combination, respectively, where the boundary condition at the lower limit of the clustering cone size (R_{cone}) is required. In addition to the border of the 95% two-jets reconstruction, the upper/lower bounds of this fitting parameters are also presented according to the parton-jet matching resolution. We should note that the fluctuation of the jet separation is dominated by the size of the resolution of the observed jet direction denoted in the upper and lower case in the Eqs.(4.3) and (4.4). That is insensitive to the fluctuation of the jet energy resolution. The right-most region with more than the 95% two-jets reconstruction probability describes the LO dominant region truncated by the lowest order perturbation calculation. One may apply the minimum separation angle for the jet partons to predict the cross section

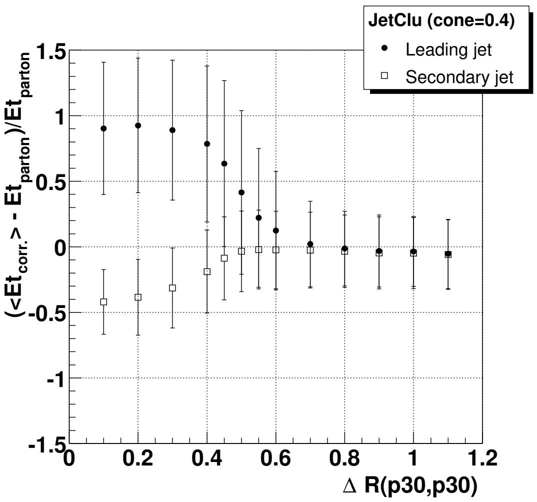


Figure 4.9: The deviation of the corrected jet energy from the primary parton. The mean value and that RMS are taken. The closed-circle is for the leading jet, and the opened-square for the secondary jet.

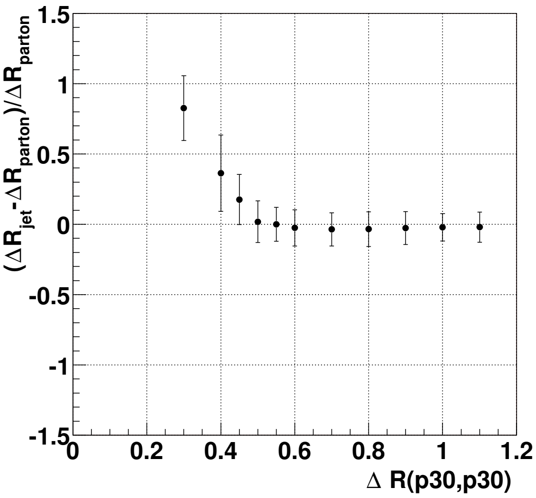


Figure 4.10: The deviation of the observed jet separation from the primary parton separation. The central value and that sigma fitted by the Gaussian distribution with the range of the 25% height on the maximum value are taken.

without kinematical cut dependence of the separation angle.

We also get the border of 50% and 5% two-jets reconstruction probability using the separation angle of the primary partons although those partons are not distincted as the observable object. Clearly, there are the one-jet reconstruction region which the two partons is merged into one jet and the transition region between 5% and 95% reconstruction probability. The treatment for this transition behavior will be addressed as the splitting and merging criteria. It is also interesting that the $\Delta R \approx 0.5$ on the 50% reconstruction border corresponds to the RunI separation requirement (≈ 0.52).

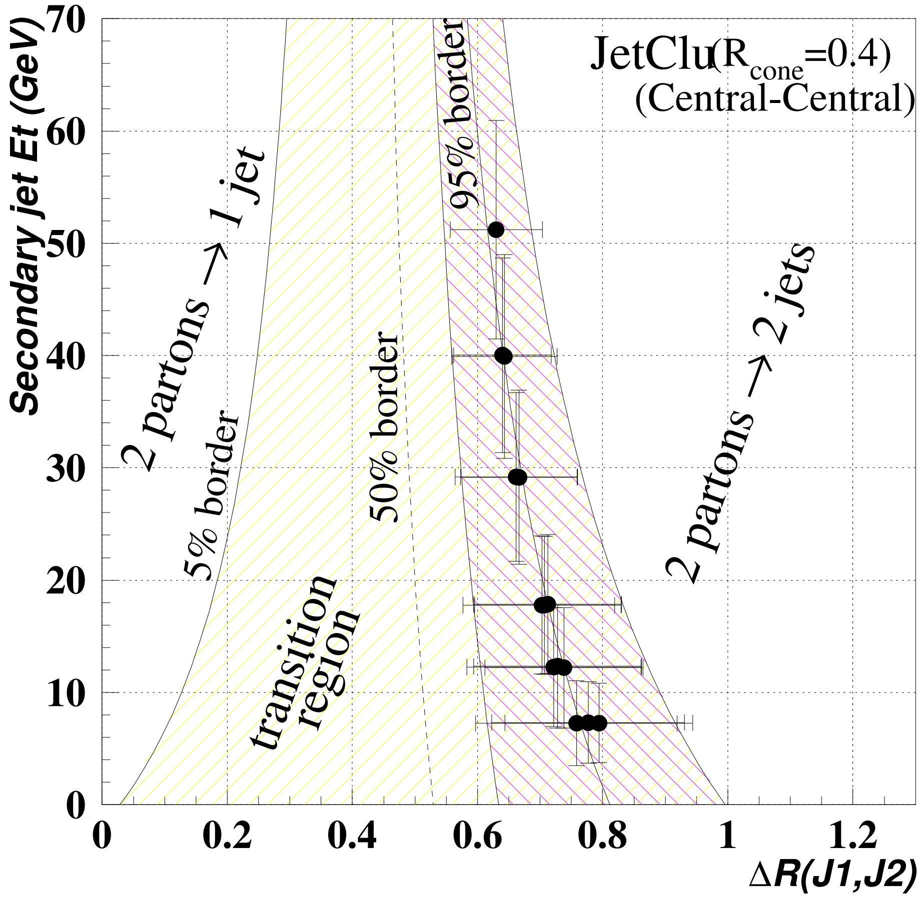


Figure 4.11: The phase space of the jet separation as a function of the jet separation angle. The two-jets reconstruction region is the region where the reconstructed jets point to the primary partons with the probability of 95%. The error bar on this plot is taken as the resolution width at one sigma level.

first parton (GeV)			second parton (GeV)			separation
p_{t1}	$\langle E_T^{corr.} \rangle$	$\langle \sigma_{E_T^{corr.}} \rangle$	p_{t2}	$\langle E_T^{corr.} \rangle$	$\langle \sigma_{E_T^{corr.}} \rangle$	$\langle \sigma_{R_{sep}^{95\%}} \rangle$
20	18.38(6)	5.96(7)	10	7.24(6)	3.77(7)	0.161(2)
	18.37(9)	6.0(1)	15	12.25(9)	5.3(1)	0.139(3)
	18.24(5)	5.91(6)	20	17.77(6)	6.14(7)	0.126(1)
30	29.60(8)	7.08(9)	10	7.30(5)	3.63(6)	0.154(2)
	29.6(1)	7.0(1)	15	12.32(9)	5.5(1)	0.134(2)
	29.45(6)	7.08(8)	20	17.85(6)	6.18(7)	0.118(1)
	29.32(6)	7.12(7)	30	29.19(6)	7.51(8)	0.098(1)
40	40.7(1)	8.4(1)	10	7.27(5)	3.53(6)	0.150(2)
	40.3(1)	8.3(2)	15	12.18(8)	5.3(1)	0.126(2)
	40.43(8)	8.2(1)	20	17.74(6)	6.15(7)	0.112(1)
	40.41(7)	8.42(8)	30	29.15(6)	7.76(8)	0.0933(9)
	40.17(7)	8.36(8)	40	39.91(8)	9.0(1)	0.0847(8)
50	50.9(1)	9.4(1)	40	40.0(1)	8.6(1)	0.078(1)
	50.7(1)	9.8(2)	50	51.2(1)	9.7(1)	0.074(1)

Table 4.2: The weighted mean and deviation for the corrected jet energy and the observed separation angle.

4.5 Comparison with different cone algorithms

4.5.1 Comparison with MidPoint and SeedLess cone algorithms

In this section, we look at the dependency of the jet separation power by the different cone jet algorithms. Midpoint and SeedLess algorithms are compared with the JetClu algorithm. Both algorithms have been developed under a motivation to treat the difficulties of QCD aspects, and will be future standard algorithms.

SeedLess algorithm :

There were some problems induced by the seed-based cone algorithm outlined in Section 4.2. A seedless cone algorithm, which does not have a starting seed tower for iteration, has a ability to remove those problems. It searches the entire calorimeter tower in the detector and finds all atable cones, even if these cones do not have a seed tower at their center. That is, the algorithm is infrared insensitive. However, since this algorithm consumes much CPU time until finding stable cones because of scanning over the whole calorimeter towers, it is difficult to use it at a trigger level, and the studies are on-going.

MidPoint algorithm :

The above seedless algorithm is not efficient to find a jet even in the high E_T jets. Alternative solution is to put a seed tower. This greatly reduces the number of cones that need to be evaluated in the initial stage. As already mentioned, simple seed-based algorithms lead to the infrared/collinear problems. The idea to add a "midpoint" seed tower between two seed towers is that the performance without infrared/collinear problems expected by the seedless algorithm is approximately satisfied. The sensitivity of the algorithm to soft radiation as illustrated in Figure 4.1 is essentially removed. The main change from the JetClu algorithm is the inclusion of midpoints of seeds and centers of larger numbers of seeds as additional seed locations for trial cones.

Note that even if those above algorithms have a possibility to remove the infrared/collinear problems, the treatment of the overlapping cones is another problem which also induces their problems. Explicit definition for splitting or marging the overlapping cones must be needed as far as we use cone algorithms.

4.5.2 Splitting and merging

The lowest order pertubative calculation does not have the sensitivity of the jet shape. The already-tuned non-perturbative hadronization model will cover this problem. However, the splitting and merging procedure causes the fluctuation of the distributions or the cross section. The differences between the lowest order pertubative calculation and the real data indicate the higher order terms of the perturbative calculation. If the merging procedure is applied in ordering of the larger E_T jet, the E_T distribution and jet shape of the largest E_T jet is affected by the contamination of the higher order terms. The splitting and merging procedures are related to the overlapping clusters [53]. The treatment of those overlapping cluster depends on each experiment. If we assume the parton information precisely translates to the hadron level, the invisible content from the higher order calculation will affect the jet shape. This procedure however is only an approximation (due to the color conservation issue) and can lead to undesirable dependence on the details of the showering and hadronization processes. Besides, since in the NLO event generation, the double counting and hadronization problems are not negligible at the detector level, the lowest order perturvative calculation is only available for the event generation to make events reliable in our current technology.

Thus, we explicitly require the obtained jet separation angle from the lowest order parton model rather than using the treatment of the overlapping clusters. In that point, the splitting/merging procedure included in the jet clustering of MidPoint and SeedLess cone algorithms were turned off. Note that since the splitting and merging procedure does not begin until all stable cones have been found, the behavior of the algorithm itself is not affected.

4.5.3 The separation power of MidPoint and SeedLess algorithms

The same analysis described in Section 4 was performed for MidPoint and SeedLess cone algorithms. The minimum E_T of the jet clustering was 3 GeV which is the same as the previous study of the JetClu. Again, the probability of jets acceptance of different cone algorithms are presented as a function of the separation angle between two partons in Figure 4.12. Both primary parton p_Ts are 30 GeV. Each opened- and closed-circle are the probabilities of the single jet and two jets reconstruction of JetClu, respectively. The squared and triangle marks represent each reconstruction probability of the MidPoint and SeedLess, too. The SeedLess algorithm shows almost the same behavior as the JetClu algorithm. The MidPoint algorithm shows some differences from these algorithms. While the MidPoint algorithm reaches the plateau region of two-jets reconstruction in the large ΔR faster than the other algorithms, in the small ΔR, it has lower probability for the one-jet reconstruction and higher for the two-jet reconstruction comparing with the other algorithms. That may result in that the MidPoint algorithm effectively try to separate the two jets putting on the small cone between two-jets clusters. In this case, the MidPoint algorithm is more free from the collinear/infrared problem in the small ΔR region than JetClu and SeedLess algorithms.

We also present the two-jet separation border of each jet algorithm with the 95%, 50%, and 5% probabilities as a function of the secondary parton p_T in Figure 4.13, where the jet energy corrections are not applied. Clearly, the MidPoint algorithm shows smaller two-jets separation angle. The SeedLess cone has sightly larger separation angle. Those results tend to have almost the same behavior on the lowest order perturbation calculation level. Although the comparison of the jet clustering energies of each algorithm is also interesting subjects, we only listed all details of separation properties in Table 4.3, 4.4, and 4.5 for JetClu, MidPoint, and SeedLess cone algorithms, respectively.

4.6 Summary

We have measured the jet separation power on the JetClu algorithm by using the fake-jet event generation program. This program makes two partons with the particular directions and momenta, which allows us to look at the collinear/infrared safeties of the cone jet algorithms at the lowest order perturbation calculation level.

We defined the separation angle with the two-jets reconstruction probability of 95% using the measured direction and energy of the generated particles. The region larger than the separation angle ($> R_{sep}^{95\%}$) corresponds to the truncated phase space with the well-defined jet signature pointing to the primary parton. As the quantitative results, we present the fitting formula in Eqs.(4.3) and (4.4) in the case of the cone size of 0.4. The requirement of the well-separated jets could be a good benchmark to compare the lowest order perturbative calculation with real data. The well-separated jets become dominant in $\Delta R \gtrsim 0.8$. And the JetClu is free from the collinear problem in $\Delta R \lesssim 0.2$, although to see the collinear region is experimentally less meaningful. In the transition region, there are the mixed states in which events have one or two reconstructed jets. The soft

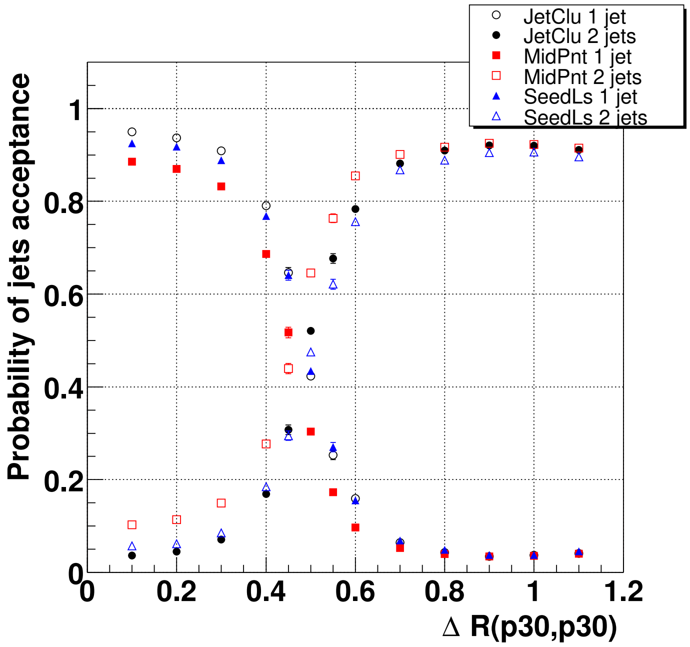


Figure 4.12: Probability of jets acceptance of different cone algorithms as a function of the separation angle between two partons. Each opened- and closed-circle is the probabilities of the single jet and two jets reconstruction of JetClu, respectively. The squared and triangle marks represent each reconstruction probability of the MidPoint and SeedLess, too. The minimum threshold energy of jet clustering is 3 GeV. Both primary parton p_Ts are 30 GeV.

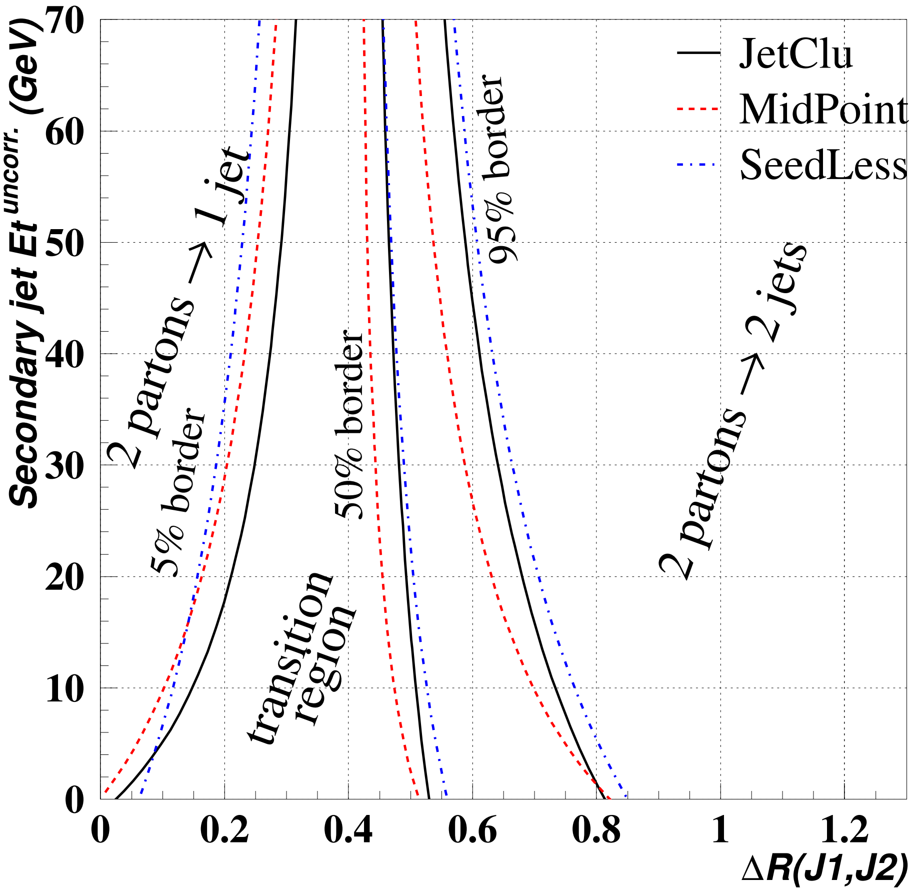


Figure 4.13: Two-jet separation border of each jet algorithm with the 95%, 50%, and 5% probabilities as a function of the secondary parton p_T, where the jet energy corrections are not applied for the jet E_T.

first parton (GeV)			second parton (GeV)			$\Delta R_{j1,j2}$	R_{sep}		
p_{t1}	$\langle E_T^{uncr.} \rangle$	$\langle \sigma_{E_T^{uncr.}} \rangle$	p_{t2}	$\langle E_T^{uncr.} \rangle$	$\langle \sigma_{E_T^{uncr.}} \rangle$	$\langle \sigma_{R_{sep}^{95\%}} \rangle$	5 %	50 %	95 %
20	13.97(5)	4.36(5)	10	5.58(6)	2.95(6)	0.161(2)	0.093	0.514	0.758
	13.98(7)	4.44(9)	15	9.39(6)	3.73(7)	0.139(3)	0.120	0.502	0.722
	13.87(4)	4.30(5)	20	13.43(4)	4.50(5)	0.126(1)	0.141	0.497	0.703
30	22.27(6)	5.38(7)	10	5.64(5)	2.80(6)	0.154(2)	0.094	0.526	0.777
	22.39(8)	5.2(1)	15	9.46(6)	3.89(7)	0.134(2)	0.160	0.506	0.728
	22.20(5)	5.25(6)	20	13.46(4)	4.55(5)	0.118(1)	0.184	0.497	0.712
	22.05(4)	5.28(5)	30	21.80(5)	5.58(6)	0.098(1)	0.225	0.489	0.662
40	30.73(7)	6.31(8)	10	5.68(5)	2.72(5)	0.150(2)	0.139	0.527	0.794
	30.48(9)	6.1(1)	15	9.38(6)	3.81(7)	0.126(2)	0.140	0.511	0.738
	30.50(6)	6.24(7)	20	13.40(4)	4.54(5)	0.112(1)	0.212	0.501	0.707
	30.52(5)	6.30(6)	30	21.77(5)	5.88(6)	0.0933(9)	0.234	0.489	0.666
	30.34(5)	6.29(6)	40	29.93(6)	7.12(8)	0.0847(8)	0.236	0.483	0.643
50	38.6(1)	7.3(1)	40	29.96(9)	6.8(1)	0.078(1)	0.249	0.479	0.640
	38.7(1)	7.3(1)	50	38.5(1)	7.9(1)	0.074(1)	0.270	0.478	0.630

Table 4.3: Some properties of the JetClu algorithm. Note that the jet energy correction is not applied.

radiation or E_T ordering of the seed tower prevent the jet reconstruction. JetClu is not capable of treating the infrared/collinear problems in this region.

The different cone jet algorithms, MidPoint and SeedLess cone algorithm, were also evaluated in the same situation. We looked at the dependence on the different algorithms. While the JetClu and SeedLess algorithms showed almost the same behavior, the MidPoint had the smaller two-jets separation angle than the other algorithms, where the merging/splitting procedure was not used in this analysis.

This study indicates a direction toward more detailed study of the jet shape addressing the splitting and merging issues of jets.

first parton (GeV)			second parton (GeV)			$\Delta R_{j1,j2}$	R_{sep}		
p_{t1}	$\langle E_T^{uncr.}\rangle$	$\langle\sigma_{E_T^{uncr.}}\rangle$	p_{t2}	$\langle E_T^{uncr.}\rangle$	$\langle\sigma_{E_T^{uncr.}}\rangle$	$\langle\sigma_{R_{sep}^{95\%}}\rangle$	5 %	50 %	95 %
20	13.57(4)	4.87(5)	10	5.33(6)	3.04(6)	0.155(2)	0.073	0.475	0.725
	13.40(7)	5.06(9)	15	8.90(6)	4.03(7)	0.140(2)	0.132	0.465	0.672
	13.15(4)	4.86(5)	20	12.65(4)	5.17(5)	0.125(1)	0.146	0.460	0.655
30	21.96(5)	5.47(6)	10	5.43(5)	2.89(6)	0.152(2)	0.062	0.492	0.758
	22.01(8)	5.4(1)	15	8.89(7)	4.22(9)	0.137(2)	0.106	0.469	0.692
	21.65(5)	5.82(6)	20	12.69(4)	5.15(5)	0.118(1)	0.118	0.458	0.665
	21.24(5)	5.93(6)	30	20.77(5)	6.54(7)	0.102(1)	0.182	0.453	0.618
40	30.30(7)	6.38(8)	10	5.41(6)	2.82(6)	0.152(2)	0.039	0.504	0.785
	29.9(1)	6.2(1)	15	8.82(7)	4.14(9)	0.129(2)	0.081	0.480	0.720
	30.03(5)	6.49(7)	20	12.76(4)	5.20(5)	0.112(1)	0.106	0.464	0.680
	29.79(6)	6.63(7)	30	20.86(6)	6.66(7)	0.096(1)	0.170	0.451	0.620
	29.37(6)	6.96(7)	40	28.94(7)	7.85(8)	0.0864(8)	0.204	0.447	0.595
50	37.8(1)	7.5(1)	40	29.0(1)	7.3(1)	0.084(1)	0.178	0.441	0.600
	37.7(1)	8.1(1)	50	37.1(1)	9.2(1)	0.080(1)	0.220	0.443	0.578

Table 4.4: Some properties of the MidPoint algorithm. Note that the jet energy correction is not applied.

first parton (GeV)			second parton (GeV)			$\Delta R_{j1,j2}$	R_{sep}		
p_{t1}	$\langle E_T^{uncr.}\rangle$	$\langle\sigma_{E_T^{uncr.}}\rangle$	p_{t2}	$\langle E_T^{uncr.}\rangle$	$\langle\sigma_{E_T^{uncr.}}\rangle$	$\langle\sigma_{R_{sep}^{95\%}}\rangle$	5 %	50 %	95 %
20	13.49(5)	4.62(6)	10	4.9(1)	3.2(1)	0.164(2)	0.084	0.528	0.792
	13.51(7)	4.59(9)	15	8.93(6)	3.95(8)	0.148(3)	0.138	0.513	0.736
	13.30(4)	4.59(5)	20	12.96(5)	4.77(5)	0.135(1)	0.141	0.504	0.718
30	21.70(6)	5.47(7)	10	5.12(8)	3.04(7)	0.157(2)	0.092	0.550	0.821
	21.8(1)	5.3(1)	15	8.93(8)	4.0(1)	0.140(3)	0.114	0.528	0.770
	21.59(5)	5.52(6)	20	12.95(4)	4.66(5)	0.122(1)	0.120	0.510	0.738
	21.43(4)	5.55(5)	30	21.02(5)	5.95(6)	0.103(1)	0.180	0.495	0.679
40	30.18(7)	6.28(8)	10	5.16(8)	2.93(8)	0.154(2)	0.089	0.556	0.829
	29.7(1)	6.3(1)	15	8.84(8)	4.0(1)	0.135(2)	0.088	0.539	0.801
	29.81(6)	6.55(7)	20	12.84(4)	4.83(5)	0.115(1)	0.125	0.521	0.751
	29.73(6)	6.45(7)	30	20.97(6)	6.13(7)	0.104(1)	0.160	0.500	0.698
	29.49(6)	6.75(7)	40	29.14(7)	7.49(8)	0.0892(9)	0.187	0.491	0.668
50	37.7(1)	7.6(1)	40	29.0(1)	7.3(1)	0.085(1)	0.171	0.491	0.676
	37.8(1)	7.5(1)	50	37.4(1)	8.4(1)	0.082(1)	0.194	0.488	0.659

Table 4.5: Some properties of the SeedLess algorithm. Note that the jet energy correction is not applied.

Chapter 5

Jet Separation Method

5.1 Introduction

For many years, it is known that Quantum Chromodynamics (QCD) — the theory of the genesis developed with many idea and experimental results — still remains the difficult portion on the experimental aspects. All the physics phenomena observed at the hadron colliders originate from the partons in the colliding hadron particles. Thus an optimal exploitation for the precision tests of the Standard Model or new physics searches is unimaginable without the well understanding of many aspects of QCD and their implementation in accurate Monte Carlo programs.

The jet production associated with the vector bosons contains many physics motivated topics. Due to easy triggering of the high transverse momentum of the leptons from their bosons, this process is important not only as the dominant background process of the most of precision measurements but also as the probing sample for the new physics.

Theoretical basis to describe the multi-parton final state in the hadron collisions, however, is still unsatisfied. The parton showering (PS) algorithm has been developed to make multi-particle final state at first. That describes all outgoing partons with the all-order summation at Leading-Log (LL) level in the perturbative calculation. Unfortunately, it was not enough to describe the high p_T region of outgoing partons, especially in the multi-partonic processes. In order to compensate that high p_T region, thus, many

ME-based event generators, known as VecBos [54], Alpgen [55], CompHep [56], Madgraph [57], GR@PPA[58], MCFM [59], and so forth, have been developed with increasing the physics importance of the multi-particle final state in the last decade. Because of its order-dependence of ME calculations, ME-based event generators cannot help using a powerful tool of the parton showering in order to make reliable events in the analysis level. However, once one may naively consider to connect with PS and ME, one encounters a merging problem notoriously known as "double counting problem", which results in the overlapped region described by the PS and ME in the certain phase space. This problem has been the big theoretical issue in the next-to-leading order (NLO) calculation. As a possible solution, various ideas are proposed. Mainly, the PS corrections are advocated in [43]. The LL-subtraction from ME and PDF is discussed in [60], [45], and [44], respectively. Even in the lowest order (LO) level, the "double counting problem" happens when the PS adopts the Q^2 running from higher scale of the hard process of the ME calculation. There obviously seems no established principle for the merging of PS and ME.

Aside from the above attractive ideas, a new prescription [46] was proposed from a point of view of the experimental analysis level using cone jet algorithms. That proposes to require the parton-jet matching, which the primary parton from ME calculation points out to the observed jet, then, only the matched-jets are used in the analysis. This results in the following benefits :

- a large suppression of the fluctuation of the total cross section depending on the kinematical cuts at the generator level,
- a small probability that the PS from the outgoing partons emits over the particular cone size,
- a treatment of the "invisible" partons in the NLO calculation.

At first, the real physics should not depend on the unphysical kinematical cuts on the ME calculations. The soft/collinear enhancement such as the requirements of very small angle cuts between massless partons or p_T cuts may cause the infinite cross section. If the outgoing partons from ME are always observed as the jets, the dependence on the kinematical cuts at the ME calculation will be drastically controlled. Second, by specifying that the primary parton from the ME distributes within the particular cone size, the ME-dominant phase space is constructed, and then the activity of the PS which makes the additional jets is suppressed with the small probability. Since the partons from PS are absorbed into the jet cone, those partons are not detected as the jets. Hence, this may partly avoid the doubly-counted phase space with PS and ME. Third, as far as considering the Feynman picture, the collinearly-emitted partons are absorbed in the primary parton to the cancellation with the loop correction. In order to get the consistent picture with the NLO calculation, these invisible parton should not be isolated as the jet. As the same argument, the real emitted parton, in the external line of the Feynman diagrams, should be detected as a jet. Note that this prescription is only valid in the cone jet algorithms.

In this study, we report an extended algorithm for the parton-jet matching and a clear logic for the jet merging/splitting criteria. We also propose a new procedure for the normalization problem. Finally, we present the test results for this prescription by comparing with data.

5.2 Jet separation method

5.2.1 Procedure of the parton-jet matching

How do we require the parton-jet matching? One may consider to take the matching criteria between the observed jet p_T[1] or direction (separation angle, $R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$) and the primary partons from ME calculation. Since a jet energy resolution is much worse than the one of a jet direction, the separation angle will be a better variable. Also, one may imagine the requirement that the primary parton distributes within the cone (R_{cone}) used in the jet clustering. However, for instance, this requirement is too tight requirement in the smaller cone size such as 0.4 mainly used in the top-quark analysis.

[1] In order to avoid a confused discussion, we refer a transverse momentum p_T for all jet clustering algorithm, although the JetClu algorithm is using a transverse energy E_T.

The small matching efficiency definitely makes systematics worse. And, in this case, there is no solution for the treatment of the overlapping cones. We have to recall that the reason why the parton-jet matching is required is to construct the ME dominant phase space. The ideal matching criteria has a 100% efficiency even in the detector level.

In order to describe the ME dominant phase space, we have studied the jet separation using a toy MC program described in the previous section. In that study, we have estimated the jet separation angle with 95% reconstruction probability as the 2 jets, where 2 partons are always reconstructed as 2 jets. The estimated jet separation angle, $R_{sep}^{95\%}$, is almost twice larger than the cone size used in the jet clustering. We found the mixed-state with one- or two-jets if the separation angle between the observed jets is smaller than $R_{sep}^{95\%}$. This explicitly indicates large ambiguity of the jet reconstruction in the small angle region. Therefore, we propose the extended matching criteria to take $R_{sep}^{95\%}$ instead of R_{cone}, where our definition of the parton-jet matching is that only one parton present in the single jet cone. We also propose that the merging/splitting procedure should be defined in terms of this jet separation, $R_{sep}^{95\%}$. The parton-jet matching and jet merging/splitting criteria are the followings,

$$\begin{array}{ll} if\ \Delta R(parton, jet) \ \le\ R_{sep}^{95\%}, & \text{match}, \\ else, & \text{not match}, \end{array} \tag{5.1}$$

$$\begin{array}{ll} if\ \Delta R(jet, jet) \ \le\ R_{sep}^{95\%}, & \text{merge}, \\ \text{else}, & \text{split}. \end{array} \tag{5.2}$$

If the separation angle between the observed jets is smaller than $R_{sep}^{95\%}$, those jets are merged, otherwise splitted. This merging/splitting criteria correspond to seeing the truncated phase space dominated by the lowest order ME calculation. The schematic view of the jet separation and the merging/splitting criteria is illustrated in Figure 5.1. When a parton splitted from the primary parton of the ME calculation goes out of the cone used in the jet clustering, this parton may be reconstructed as an additional jet with a certain probability. However, if this parton is in the direction between R_{cone} and $R_{sep}^{95\%}$, it can fall within a single cone of R_{cone} around the central particle. In this region, there is a mixed situation with one- or two-jets reconstruction. The cone jet algorithm is not capable of discriminating this parton as a jet or not. This problem is addressed as a collinear problem of the cone jet algorithms. This is the limited ability of the cone jet algorithms. On the other hand, if the parton goes outside $R_{sep}^{95\%}$, it is reconstructed as a jet more than 95% jet separation probability. Hence, to require $R_{sep}^{95\%}$ for the merging/splitting procedure is able to avoid the ambiguity of the fluctuation of a jet clustering in the cone jet algorithms.

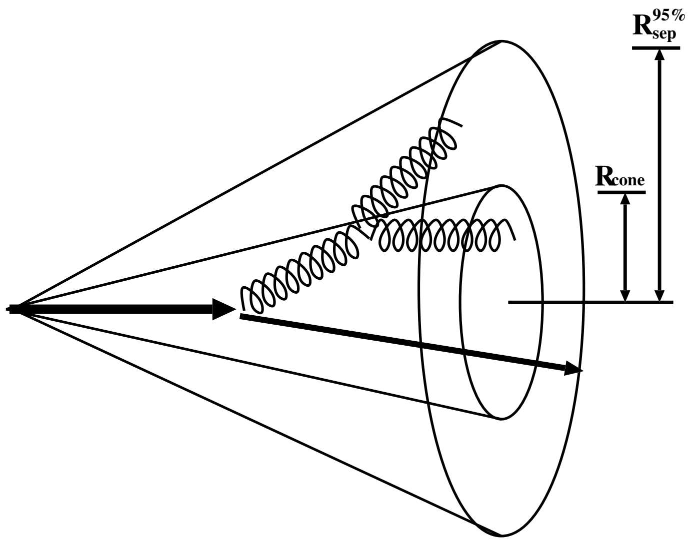


Figure 5.1: An illustration of collinear sensitivity in cone jet clustering. If a parton splitted from the primary parton of the ME calculation goes to between R_{cone} and $R_{sep}^{95\%}$, the mixed situation with one- or two-jets reconstruction happens.

5.2.2 Parton-jet matching efficiency

The parton-jet matching efficiency after the jet clustering by JetClu algorithm at the detector level is estimated as a function of a jet p_T according to the matching procedure discussed in the previous section. A W + 2 jets ME produced by GR@PPA event generator is chosen for discussion. The kinematical cuts at the generator level are $p_T > 8$ GeV, and $|\eta_{det}| < 3.0$, respectively. To detect the primary parton as an isolated jet, the separation angle is set as $\Delta R(\text{p1,p2}) > R_{sep}^{95\%}$. The parton showering and hadronization are performed by HERWIG MC program [47]. The minimum E_T threshold of the jet clustering of JetClu is 3 GeV. It is practical to compare this with the naive matching procedure within the R_{cone}. The plot is shown in Figure 5.2. The closed and opened circle marks are the matching efficiency by the $R_{sep}^{95\%}$ and $R_{cone}(0.4)$ basis, respectively. The hatched region is the systematics of the $R_{sep}^{95\%}$ basis already estimated in the previous section. The matching efficiency is defined as the ratio of the number of matched-jets to the number of most-neighboring jets on the parton :

$$\varepsilon_{parton-jet} = \frac{N(p_T;\ \Delta R(parton, jet) < R_c)}{N(p_T;\ min.\{\Delta R(parton, jet)\})} \quad , \tag{5.3}$$

where R_c is the matching condition required by each procedure, that is, $R_{sep}^{95\%}$ or R_{cone}. From this plot, we can see an improved efficiency of the $R_{sep}^{95\%}$ basis. In the low p_T region, the $R_{sep}^{95\%}$ basis clearly rescues the low p_T jets which is very sensitive to the spread of a jet. We should note that there is no effect on the absolute jet clustering because the denominator in Eq.(5.3) is the relative efficiency to the reconstructed jets.

The separation angle at the generator level also affects the overall matching efficiency. The smaller separation angle in the event generation makes the matching efficiency worse. In order to reduce the systematics of the matching criteria, it would be better to maximize the matching efficiency by setting the larger separation angle at the generator level. In addition to the circle plots in this figure, the three dashed lines are overlayed with the different separation angle at the generator level, $R_{sep}^{95\%}$, 0.4, and 0.2 from top to bottom, respectively. In the case of ΔR=0.2, the overall matching efficiency gets worse.

5.2.3 Kinematical dependence

One of the main reason of the parton-jet matching is a large suppression of the kinematical cut dependence at the generator level. If two partons from the ME calculation are collinearly emitted within the $R_{sep}^{95\%}$, the overlapped phase space is rejected by the parton-jet matching. As this result, no dependence on the kinematical cuts at the generator level is expected. In Figure 5.3, the kinematical dependence for the leading jet p_T is shown with the various ΔR (= 0.2, 0.4, $R_{sep}^{95\%}$) cuts at the generator level. The distribution inclusively contains the highest p_T jet not only for the matched-jets but also the

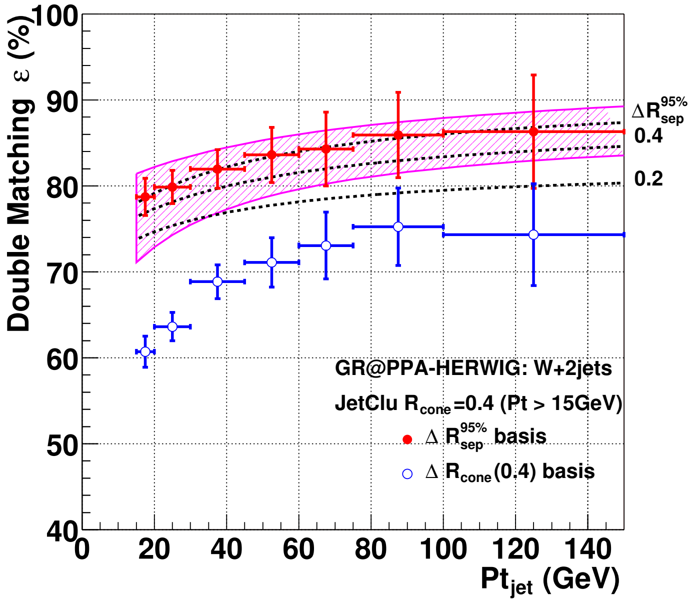


Figure 5.2: Parton-jet matching efficiency after the jet clustering by JetClu algorithm at the detector level as a function of a jet p_T. The closed and opened circle marks are the matching efficiency by the $R_{sep}^{95\%}$ and $R_{cone}(0.4)$ basis, respectively. The hatched region is the systematics of the $R_{sep}^{95\%}$ basis. In addition to the circle plots, the three dashed lines are overlayed with the different separation angle at the generator level, $R_{sep}^{95\%}$, 0.4, and 0.2 from top to bottom, respectively.

no-matched-jets. As expected, all three distributions are almost same regardless as the large fluctuation of the cross section strongly depending on the ΔR cuts. Although there are slight dependence in the low p_T region, this is covered by the systematic uncertainty of the parton-jet matching.

While the parton-jet matching suppresses the kinematical cut dependence, some events are rejected by this requirement in the MC simulation. The requirement of the parton-jet matching however is not applied for the real data. Also, we cannot identify which jets come from the perturbative calculation in the data. When we compare the parton-matched MC sample with the real data, the MC sample is biased by the rejection factor of the parton-jet matching. If this rejection has a flat behavior, we can simply compare this distribution with the data. Unfortunately, in our prescription, this rejection has a p_T dependence as estimated in Figure 5.2. Thus, we have to correct the MC sample by the term of the parton-jet matching efficiency. The leading jet p_T distribution after the correction of the matching efficiency is shown in Figure 5.4. In addition to the histogram of the corrected p_T distribution, the hatched region represents the systematic fluctuation by applying the matching efficiency. The inclusive leading jet p_T distribution without the parton-jet matching requirement is also superimposed as the circle marks together. After applying the efficiency correction, the shape of the distribution in the low p_T region is more corrected than that in the high p_T region. And the corrected distribution reproduces almost the same distribution as that of the no-matching requirement. This correction is necessary to avoid the bias of the parton-jet matching.

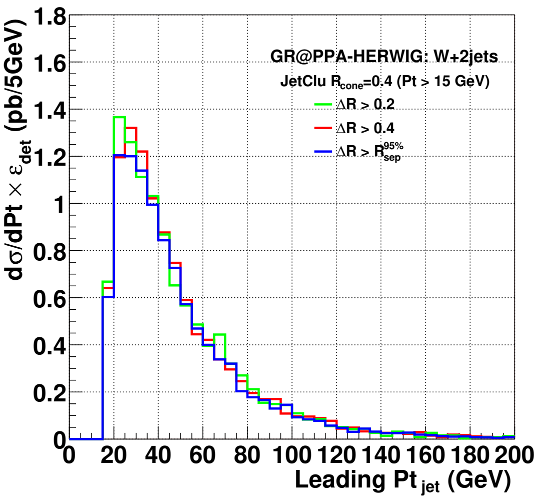


Figure 5.3: Kinematical dependence for the leading jet p_T with the various ΔR ($=$ 0.2, 0.4, $R_{sep}^{95\%}$) cuts at the generator level.

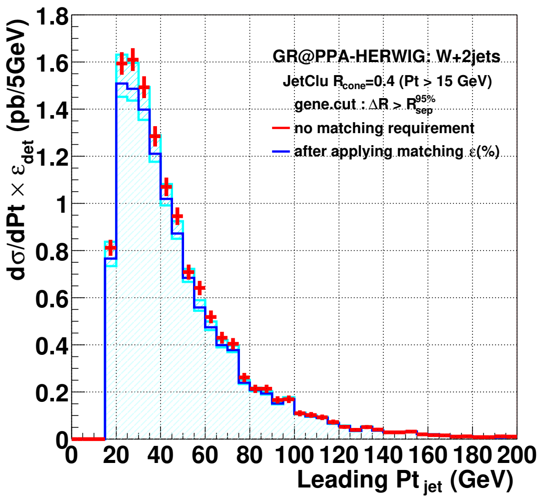


Figure 5.4: Leading jet p_T distribution after the correction of the matching efficiency. The inclusive leading jet p_T distribution without the parton-jet matching requirement is also superimposed.

5.2.4 Normalization

Most of available event generators are based upon the lowest order (LO) calculation. In the LO calculation, nothing tells us about the normalization. Besides, the estimated cross section itself is less meaningful since the LO calculation is not able to avoid the scale dependence. We only realize they can be solved at the higher order calculation. Also, we cannot estimate the absolute efficiency of the cone jet clustering.

Nevertheless, we need a normalization using the output cross section of the LO calculation. That could be a bench mark point toward the NLO calculation. If the NLO event generators up to high jet multiplicity region are theoretically available, we will get a relative result with the LO calculation. Also, the difference is denoted as the k-factor which will be useful for many background analysises although this k-factor is an analysis-dependent. In our prescription, we have described the ME dominant phase space by requiring the jet separation of the $R_{sep}^{95\%}$ limit discussed in the previous section. By requiring the parton-jet matching, we could avoid the overlapped phase space within the smaller jet separation angle. That rejection rate is measured in Figure 5.2. We hence interpret that the extent of the overlapped phase space is replaced with the rejection rate, that is, minimizing the rejection rate corresponds systematically to avoiding the overlapped phase space. We thus propose to set

$$\Delta R(parton, parton) \ > \ R_{sep}^{95\%} \quad , \tag{5.4}$$

at the generator level. This separation cut is the limit point to avoid the ΔR cut dependence at the generator level. The estimated cross section requiring the above kinematical cut will have most systematically small fluctuation. Indeed, the larger separation angle suppresses the fluctuation of the scale dependence because the strong coupling constant is enhanced in the collinear region. We show plots of the matching efficiency versus ΔR cuts in Figure 5.5 and 5.6, accompanied with the averaged $R_{sep}^{95\%}$ point. Clearly, we can see that the $R_{sep}^{95\%}$-based matching criteria points to the maximum matching efficiency.

The cross section would be better to be normalized inclusively, because the PS does not change the cross section calculated by the ME. The requirement of the jet merge/splitting procedure based on the jet separation effectively absorbs the additional PS contribution into the single cone around the primary partons. The emission of the initial state radiation however is independent of the primary parton direction. The PS still have a phase space to fly to the another region which is not described by the jet separation. The isolated PS can make additional jet contents.

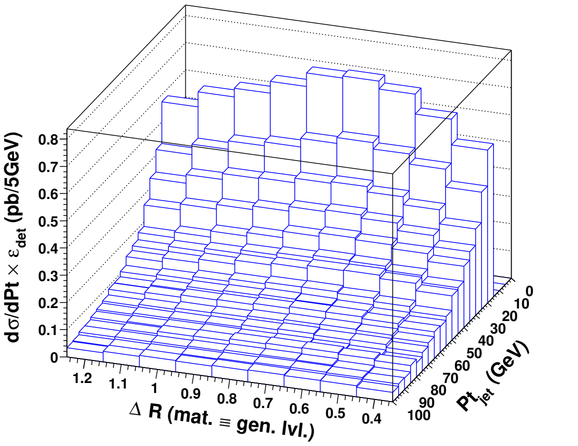


Figure 5.5: An optimization of the matching efficiency to various ΔR cuts in the 3-Dimention.

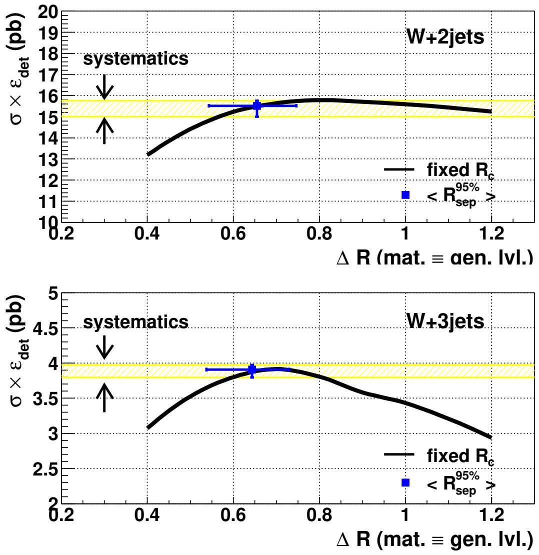


Figure 5.6: An optimization of the matching efficiency to various ΔR cuts in the W + 2 jets and 3 jets processes, respectively.

5.3 Summary

We have proposed a new prescription based upon the jet separation criteria. The jet separation is defined as the separation angle, $R_{sep}^{95\%}$, that the two partons are reconstructed as two jets with the probability of 95%. The parton-jet matching and the jet merging/splitting procedure in the overlapping cones are also required in terms of $R_{sep}^{95\%}$. By requiring the jet separation method, the ME dominant phase space is constructed. The primary parton from the ME calculation goes into this phase space. The other contributions from the parton showering within $R_{sep}^{95\%}$ are absorbed into the single cone around the primary parton direction. That means the double counting between ME and PS are partly avoided.

Chapter 6

Tests of Enhanced Leading Order QCD in W + jet Events

6.1 Introduction

In this chapter, we apply the jet separation method to the real data. Then, we compare it with the lowest order QCD prediction. For the theory prediction, we use GR@PPA Monte Carlo event generator which is described in Chapter 7. The four-vectors of the final state partons from the ME calculation by GR@PPA pass through HERWIG showering Monte Carlo simulation, and pass through the CDF detector simulation. The same selection criteria as applied to the data are applied in the MC samples. The jet E_T distributions in multijets events, jet multiplicity, and its ratio are compared with the data and MC.

6.2 $W + \geq$ n jets data samples

6.2.1 Data samples

After obtaining the W event samples described in Section 3, the jet identification cuts are applied. The JetClu algorithm, a fixed cone clustering algorithm, is used in this analysis. The requirements of the transverse energy and η coverage are :

$$E_T \ \geq \ 15 \text{ GeV} \qquad , \qquad |\eta_{det}| \ \leq \ 2.4 \ . \tag{6.1}$$

The clustering cone size is 0.4. The merging/splitting criteria is applied according to the Jet Separation Method described in the previous section. We collect the jets samples inclusively, that is, group the $W + \geq$ n jets event samples. For instance, an event which has 2 jets is a member of the $W + \geq 2$ jets event sample but at the same time it can be a member of the $W + \geq 1$ jets event sample. The final number of events is listed in Table 6.1 with some systematic uncertainty described in the following subsections.

6.2.2 Jet energy scale uncertainty

Although the systematic uncertainty to the jet energy correction is assigned as described in Section 3.4, we reestimate this uncertainty because our interests are the parton level kinematics and the jet energy correction generally depends on the parton flavor and hard process based upon the typical energy scale [62]. It is practical to see if the jet energy correction reproduces the primary parton p_T from the matched jet p_T. We show the deviation of the observed jets p_T from the matched partons p_T in the current jet energy correction[1] in Figure 6.1. The cone size used in the jet clustering is 0.4. The mean value of energies in each p_T bin is taken as a deviation between the parton and the jet. The dependences on two matching conditions of the ΔR(parton,jet) $= 0.4$ and $R_{sep}^{95\%}$ are also shown as a circle and a opened square in this figure, respectively.

All samples show an good agreement within 5% deviation in high p_T region (≥ 20 GeV). In low p_T region less than 20 GeV, the jet energy correction completely screws up to describe the parton information. There is no dependence on the different matching conditions. More improved jet energy correction such as the p_T dependent correction or the process specific correction recovers the low p_T region. But we assign the uncertainty on the jet energy of $\pm$ 10 % instead of the process specific corrections because some correction factors are quoted from the Run I measurements. Also, since we do not apply any background subtraction, the magnitude of systematics also includes the background contribution.

[1] Offline software: 4.9.1hpt3.

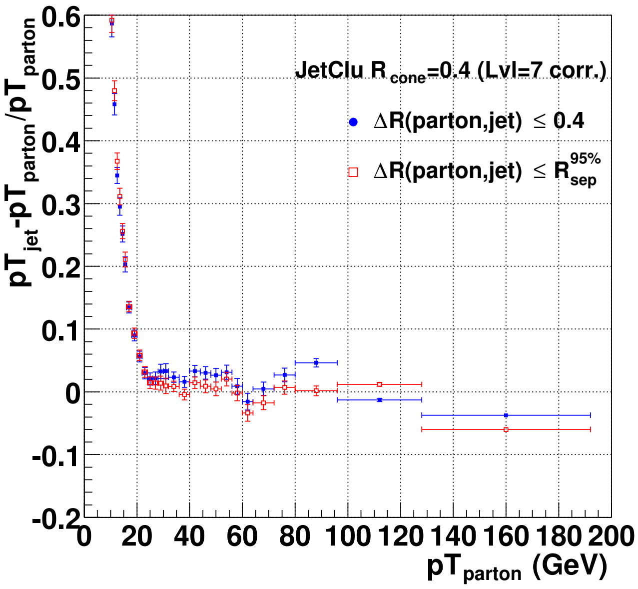


Figure 6.1: Deviation of the observed jets p_T from the matched parton p_Ts in the current jet energy correction. The dependences of the different matching conditions of $\Delta R = 0.4$ and $R^{95\%}_{sep}$ are also shown.

6.2.3 Jet counting

The merging/splitting criteria to the observed jets is taken as the Jet Separation basis $R^{95\%}_{sep}$ described in Section 5. The merging/splitting procedure is applied if two neighboring jets have the transverse energy more than 12 GeV which is the same as Run I criteria. If our jet energy correction is available in the low jet p_T region, this threshold will be reduced. The 12 GeV threshold is a conservative choice to avoid the systematics ambiguity of the jet energy scale. In Table 6.1, we list the number of events in terms of each jet multiplicity with jet energy and merging/splitting systematics effects. It is also practical to show the Run I-based merging/splitting criteria (R_{sep}=0.52) and the case of no-merging/splitting criteria to see how much difference between all of them. Only for the $R^{95\%}_{sep}$-basis, the systematics of this merging/splitting are presented in upper- and lower-subscripts on this

Jet multiplicity	≥ 0	≥ 1	≥ 2	≥ 3	≥ 4
$R_{sep}^{95\%}$-basis	31726	6860^{+12}_{-9}	1524^{-33}_{+55}	341^{-20}_{+28}	59^{-10}_{+12}
syst. +10%	—	+1017	+346	+106	+30
syst. -10%	—	-962	-331	-98	-21
Run I-basis(R=0.52)	31726	6847	1598	373	77
syst. +10%	—	+983	+354	+123	+35
syst. -10%	—	-965	-345	-102	-27
no-merging/splitting	31726	6835	1622	381	84
syst. +10%	—	+995	+329	+118	+39
syst. -10%	—	-957	-350	-107	-28

Table 6.1: Number of events in each jet bin. The $\pm 10\%$ fluctuation of the jet energy correction is taken as the systematic uncertainty. The upper- and lower-subscripts presents the systematics of $R_{sep}^{95\%}$-basis merging/splitting method.

table.

The merging/splitting procedure will change the number of the jet clustering. The merging procedure tends to decrease the jet multiplicity in an event while the splitting procedure increases it. As we expect, the jet merging/splitting procedure especially affects in the large jet multiplicity region. The acceptance of the multijet processes such as the top pair production may be affected by this merging procedure. Our merging/splitting procedure however has less impact on the jet multiplicity in this analysis because the jet energy scale uncertainty contributes much more significantly to the fluctuation of the jet multiplicity.

6.3 Comparisons of theory to data

6.3.1 Jet E_T distribution

The jet transverse energy E_T distribution is presented in Figure 6.2 for each jet process. From the upper-most side, the distributions are the highest E_T in $W + \geq 1$ jets events, the second highest E_T in W + ≥ 2 jet events, and so forth. The data points are presented as a circle dot. The statistical error is only included in this data point. The shade band among the data point is estimated by the fluctuation of the 10% jet energy scale uncertainty described in the previous section. The solid and dashed lines are the LO QCD predictions, except in $W \geq 4$ jets events, produced by GR@PPA event generator with the energy scale of the squared mass of a W boson (M_W^2 (GeV^2)) and the square of the average value of the parton p_T ($\langle\ p_T\ \rangle^2$ (GeV^2)), respectively, where the renormalization and factorization scales are equivalent denoted as the energy scale. The LO QCD prediction in $W \geq 4$ jets events is produced by Alpgen event generator with the energy scale of $M_W^2 + p_{TW}^2$. Those Matrix Element-based event generators are embedded into HERWIG showering Monte Carlo simulation, and then the generated events pass through the full detector simulation. The MC predictions are normalized by the total number of events in each $W + \geq$ n jets data sample. For the detail descriptions about the Monte Carlo event generator, see Section 7. Also, we would emphasize that a new systematics of the parton-jet matching described in the previous section is included in the MC predictions, but that contribution is negligibly small.

For the MC prediction, the energy scale of $\langle\ p_T\ \rangle^2$ varies with the parton p_T's event by event. The lower energy scale is enhanced by the larger strong coupling α_s since the size of a strong coupling constant increases with the lower energy scale. The shape of the jet E_T distribution thus depends on an order of magnitude of the α_s by event basis. At least, the requirement of the parton-jet matching, one parton in one cone in the LO calculation, gives a clear logic that the size of the α_s directly translates the jet E_T shape. Hence, we can expect that the jet E_T distribution has a sensitivity to the choice of the energy scale. We can see the steeper E_T distribution in the case of $\langle\ p_T\ \rangle^2$ than that of M_W^2. The choice of the energy scale M_W^2 is useful as a good bench mark point to compare not only to the different energy scale but also to the higher order calculation because the running strong coupling constant (scale running) is less meaningful in the NLO calculation. Both MC predictions show a good agreement with the data. The choice of $\langle\ p_T\ \rangle^2$ seems better to describe the data well, but is not clear due to the large jet energy uncertainty.

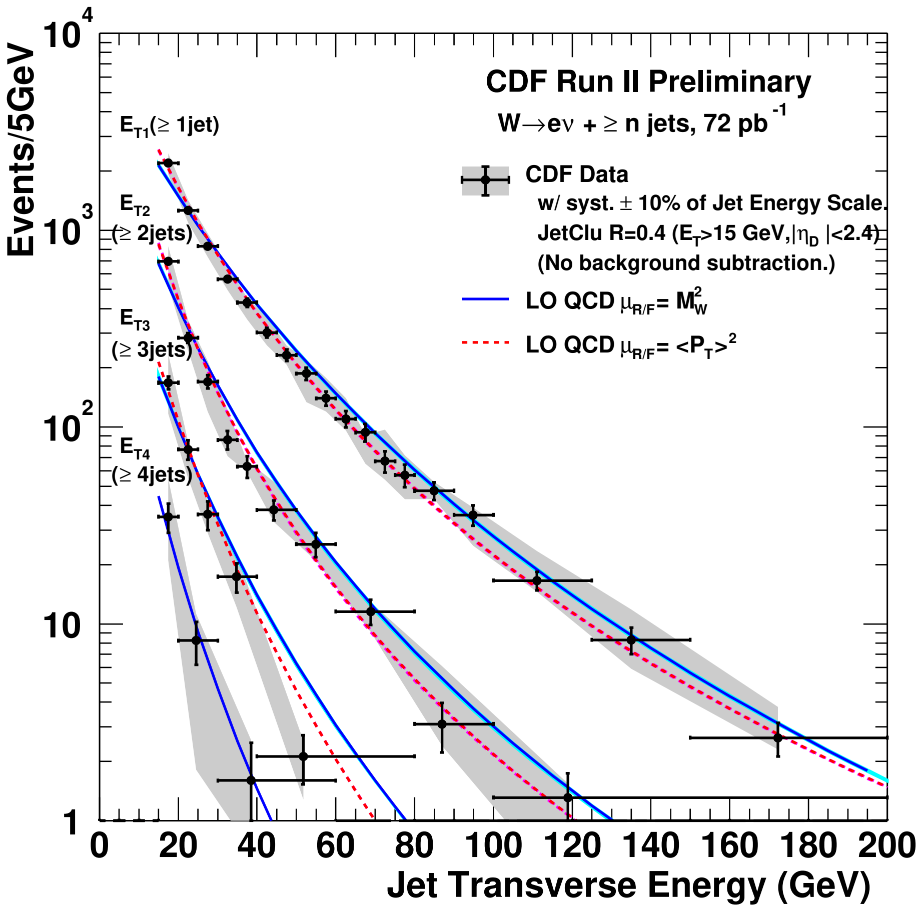


Figure 6.2: Jet transverse energy. From the up-most side, the distributions are the highest E_T in $W + \geq 1$ jets events, the second highest E_T in $W + \geq 2$ jets events, and so forth.

6.3.2 Angular and mass distributions

The invariant mass and angular distribution (ΔR_{jj}) between two jets is a sensitive variable to the collinear/infrared singularity. Some differences may be an indicator to the higher order perturbative calculation. In Figure 6.3, we present the dijet mass distribution and angular distribution between the highest E_T jet and the second highest E_T jet in the $W + \geq 2$ jets events and the $W + \geq 3$ jets events, respectively.

A discrepancy in both mass distributions of $W + \geq 2$ jets and 3 jets events in the data and MC predictions can be seen in this plot. The mass distributions of MC predictions are harder than those of the data. The distribution is better reproduced by the energy scale of $\langle\ p_T\ \rangle^2$. On the other hand, the ΔR_{jj} distributions are insensitive to the energy scale. These features could be seen in Run I measurement [11]. We see that the theory predictions for the ΔR_{jj} distribution remain valid to the resolution limit of jet-jet separation for our analysis.

6.3.3 Jet multiplicity

Using the cross section of the MC, we can compare the number of jets distribution with the data. We present the jet multiplicity distribution in Figure 6.4. The errors on the data points are the sum of the statistical and systematic uncertainty by the jet E_T scale. The lower and upper band on the LO QCD predictions correspond to the energy scale of M_W^2 and $\langle\ p_T\ \rangle^2$, respectively. All the acceptance or detection efficiencies are already included into the number of detected events because the MC events also pass through the detector simulation. The lower energy scale of $\langle\ p_T\ \rangle^2$ yields higher cross sections since it correlates with a larger value of α_s. We have also plotted the leading order theory prediction for the inclusive W production cross section ($W + \geq 0$ jets) by HERWIG built-in process. On this plot, we did not consider any background contributions. However, those background contaminations are almost negligible in the W + 0,1,2,3 jets events. Indeed, those fractions are ~2.8%, ~4.4%, ~4.7%, and ~10.1% in the W + 0,1,2,3 jets events, respectively.

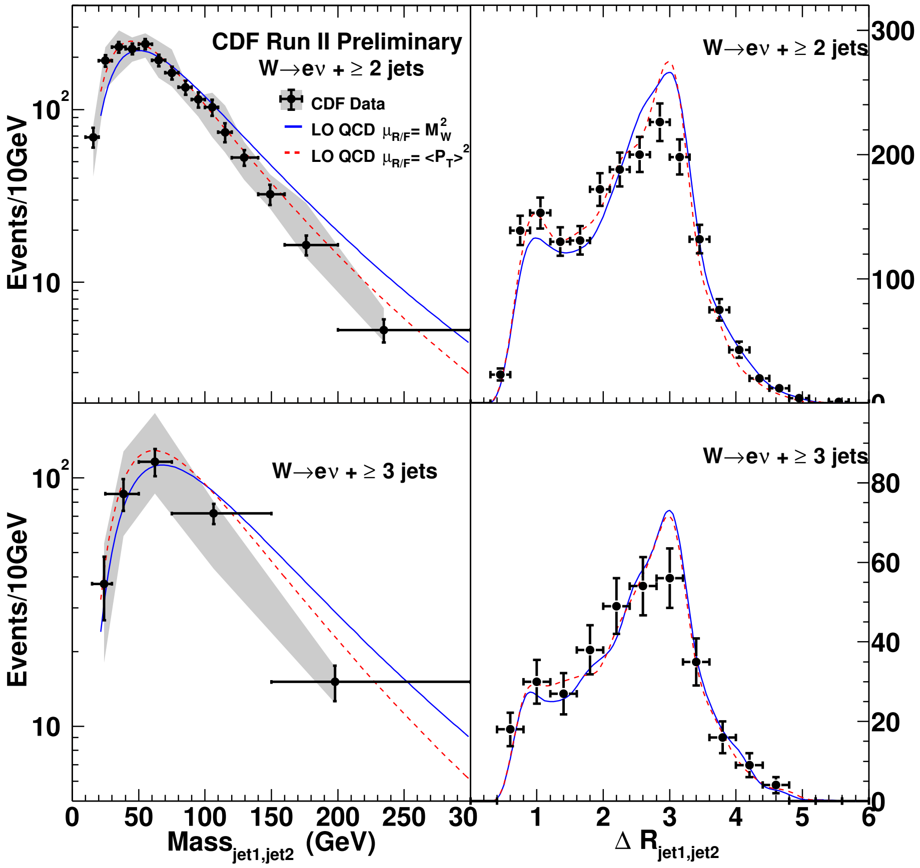


Figure 6.3: Dijet mass distribution and jet separation angle between the highest E_T jet and the second highest E_T jet in $W + \geq 2$ jets events and $W + \geq 3$ jets events, respectively.

The ambiguity for the unphysical parameter like the kinematical cuts on the generator level has been already rejected by the requirement of the parton-jet matching. Since there is only one parton from the ME calculation in the jet cone, the number of jets is proportional to the number of partons, that is, an order of the strong coupling constant. We can see almost linear relation of the jet multiplicity in both the data and MC's. This shows our analysis method well describes the enhance lowest order phase space. The difference of the absolute cross section will be addressed as a lack of the higher order calculations.

It is useful to see the fraction of the jet multiplicity presented in Figure 6.5. The number of events in each n jet bin is normalized by the number of events in $W + \geq 0$ jets events. The relative size of the jet multiplicity on the data is well reproduced by the LO calculations.

6.3.4 Ratio of the jet multiplicity

We show various ratio plots to each jet bin in $W + \geq$ n jets events in Figure 6.6. From the top, the ratio of theory to data, the ratio of n jets events to n-1 jets events, and the ratio to the ratio of n jets events to n-1 jets events,

$$R_{n/(n-1)} = \frac{\sigma_n}{\sigma_{n-1}} , \tag{6.2}$$

are presented. Taking the ratio of the physics variable is to cancel out the uncertainties from the absolute source like the luminosity. The identification efficiency or acceptance etc. may also cancel somehow out.

We see that the absolute cross section predictions agree with the data within a factor of less than 2. Those factors are ~1.2 for the energy scale of $\langle p_T \rangle^2$ and ~1.5 for M_W^2 over the range to the $W + \geq 3$ jets events, respectively. Remarkable feature is that the MC predictions show almost constant behavior in this ratio plot. That means that our analysis method and MC prediction well describe the data. It is interesting to see the ratio $R_{n/(n-1)}$ (middle). The jet counting uncertainties are reduced except for $R_{1/0}$. The $R_{n/(n-1)}$ comparison is valid if higher order QCD corrections to the LO cross sections are

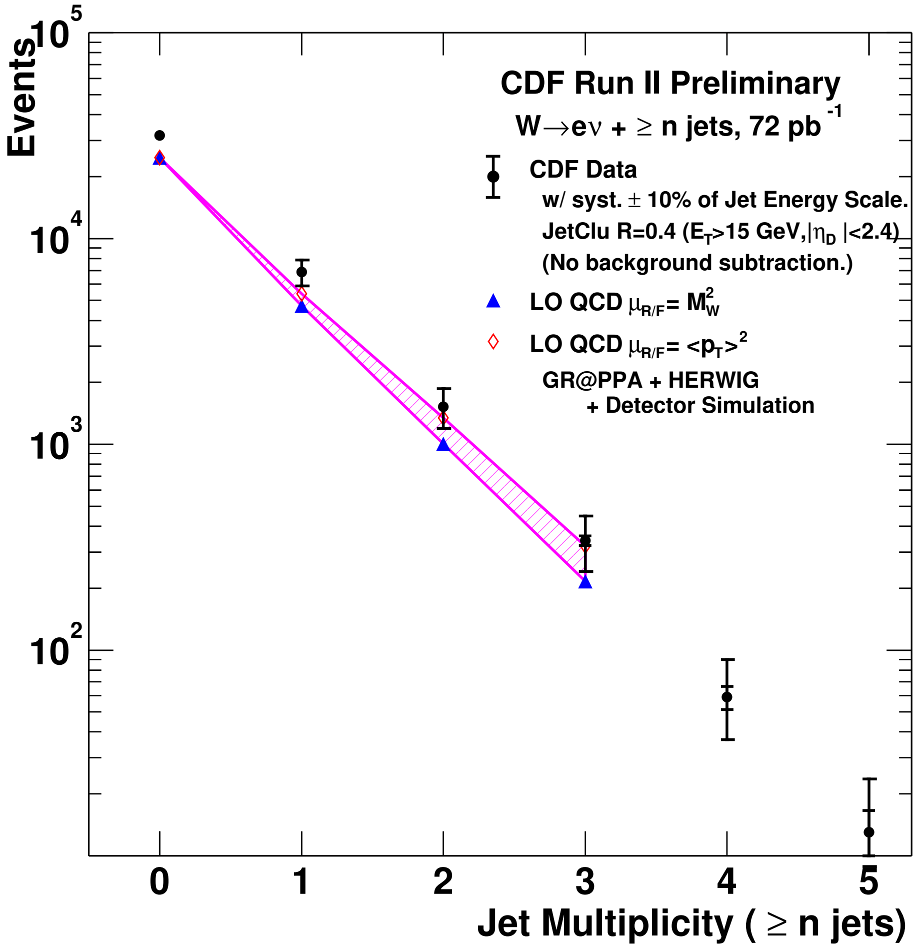


Figure 6.4: Jet multiplicity distribution. The errors on the data points are the sum of the statistical and systematic uncertainty by the jet E_T scale. The lower and upper band on the LO QCD predictions correspond to the energy scale of M_W^2 and $\langle\, p_T \,\rangle^2$, respectively.

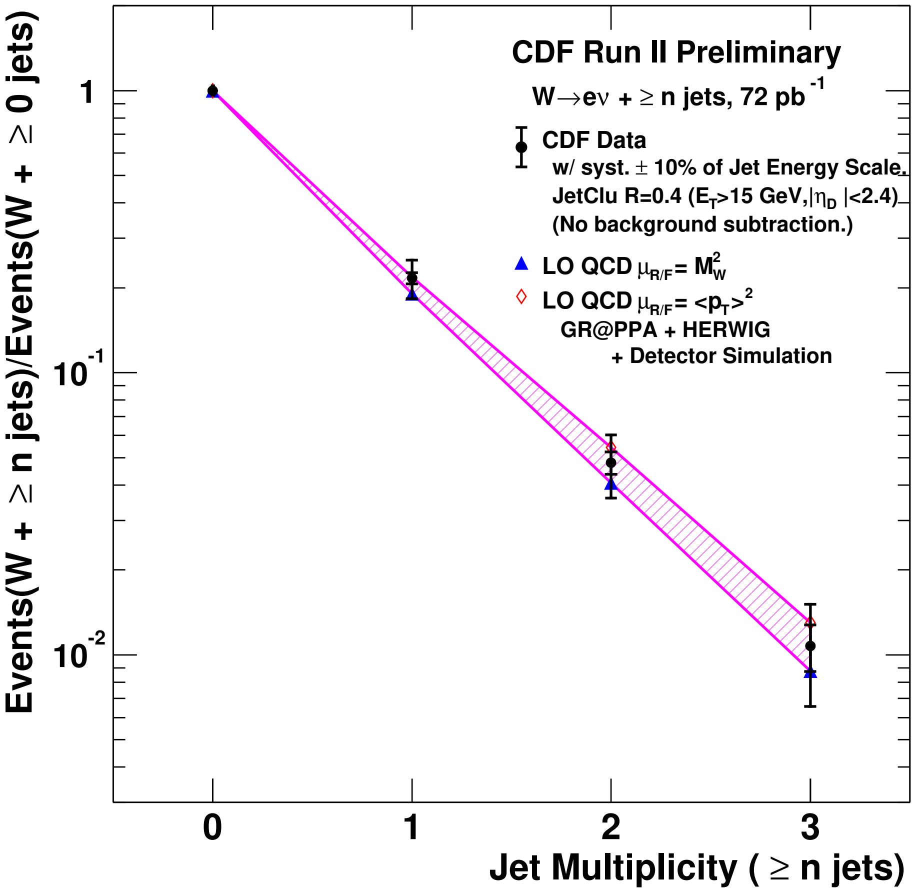


Figure 6.5: Fraction of the jet multiplicity. The number of events in each n jet bin is normalized by the number of events in the $W+\geq 0$ jets events.

not strongly dependent on the number of final state partons. The ratio $R_{n/(n-1)}$ measures the decrease in cross section with the addition of 1 jet. The value of $R_{n/(n-1)}$ thus is clearly dictated by the magnitude of the strong coupling constant since adding an extra jet adds a factor of α_s. We can see the energy scale $\langle\ p_T\ \rangle^2$ gives a better agreement than the M_W^2. In the $R_{n/(n-1)}$ plot, the particular value of $R_{n/(n-1)}$ will vary as a function of the specific jet E_T requirement that define a jet. To remove this dependence to some degree, the ratio (bottom) of data and theory for $R_{n/(n-1)}$ give a sensitivity to an independent comparison of the jet definition and its systematics. With accurate theory predictions and accurate data measurements the value of this ratio is 1.0. If the QCD predictions reproduce the jet kinamtics accurately, the ratio of data to theory is independent of the choice of the jet E_T requirement so that the quantity may be of more general interest.

6.3.5 Comparison with Run I measurement

It is practical to compare this Run II measurement with the Run I measurement. We present the fraction of jets in Figure 6.7. The fraction of the Run II data is slightly larger than the Run I results. This is not obvious feature. The upgraded collision energy of 1.96 TeV and higher instantaneous luminosity will make larger size of the production cross section and fake backgrounds. The fraction of the jet multiplicity normalized by $W + \geq 0$ jets events however should be independent of the collision energy. One of the doubtful source is the different jet definition and the magnitude of backgrounds. In our analysis, no background subtraction is applied. The detailed study is needed for more accurate measument.

6.4 Summary

Data have been compared to the theory predictions at the lowest order perturbative calculation level. Jet separation procedure based on the parton-jet matching requirement is used for the data and theoretical predictions. This requirement is to construct the enhanced LO phase space. For the theory prediction, two choices of the energy scale,

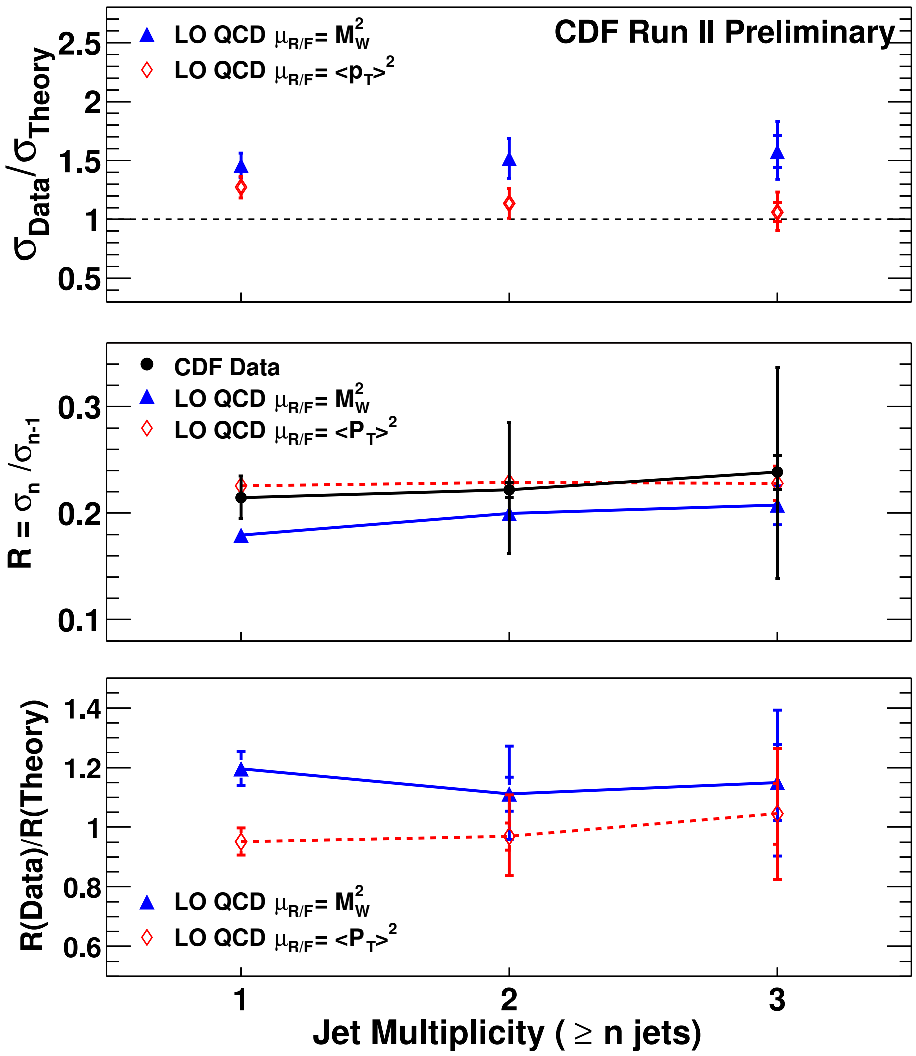


Figure 6.6: Ratio of jet multiplicity. From the top, the ratio of theory to data, the ratio of n jets events to n-1 jets events, and the ratio to the ratio of n jets events to n-1 jets events are shown.

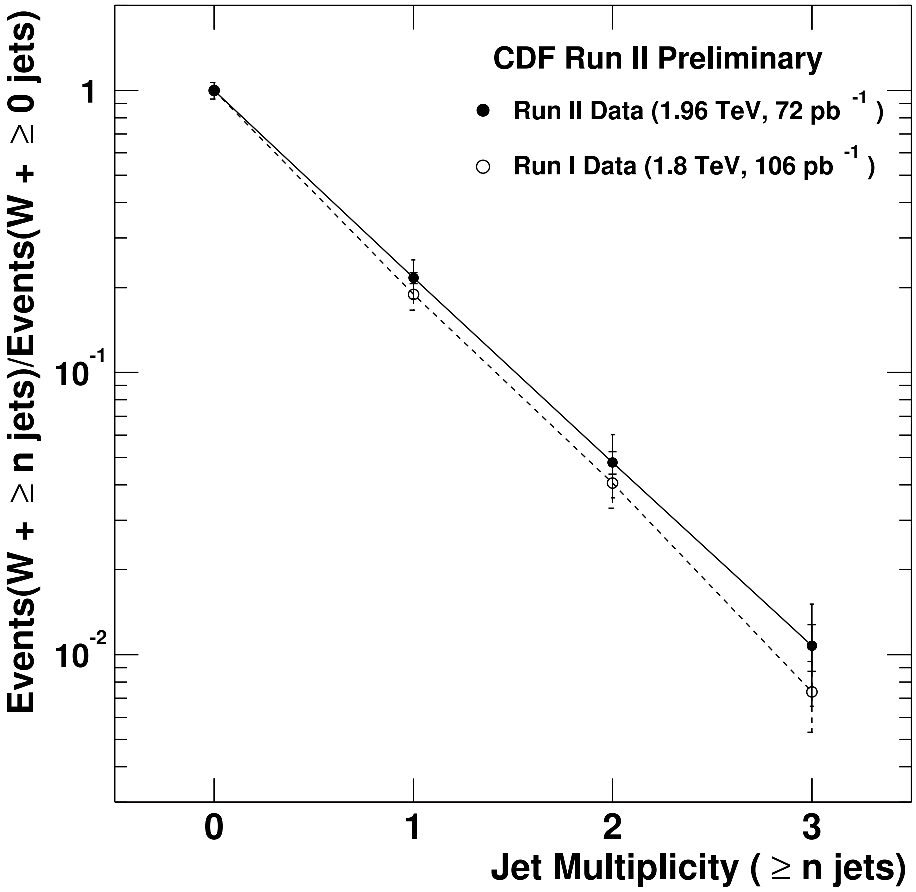


Figure 6.7: Fraction of jets with a comparison to Run I measurement.

$\langle\ p_T\ \rangle^2$ and M_W^2, where the renormalization and factorization scales are equivalent, has been tested. The jet transverse energy, mass and jet-jet separation distribution are compared between data and theory predictions. All distributions show good agreements. The choice of the energy scale of $\langle\ p_T\ \rangle^2$ is preferred to describe data well. Jet multiplicity distribution is also compared up to $W + \geq 3$ jets events. Non-background condition is assumed. But those background effects are almost negligible in less than 3 jets events. The constant (flat) behavior can be seen in the various ratio plots. This is very important feature to certify our rightness of the MC generation and analysis scheme, which is crucial for the measurement of the strong coupling constant. In the comparison with Run I results, Run II results were slightly larger than Run I results.

Chapter 7

Predictions of Quantum Chromodynamics

7.1 Introduction

In this chapter, we describe a theory prediction of the W boson plus n partons processes using a newly developed GR@PPA Monte Carlo event generator. The GR@PPA event generator makes the set of a four-vector of the final state partons as well as the final state parton flavor based on the lowest order matrix element calculation. To make reliable events at the observable level, those information are embedded into the HERWIG showering program, which is described in the later section, then hadronized, and passes through the detector simulation. We especially concentrate here on the parton kinematics, the variation of the total cross section with respect to the choice of the energy scale on the W + n partons MC predictions. We also discuss the uncertainty of PDF and a fraction of a b-quark in the W + n partons. All the technical details of GR@PPA event generator can be seen in Appendix. However, we introduce some details specified on the W + n partons processes for keeping an easy discussion to the later sections.

7.2 Characteristic of W + n partons process

7.2.1 W + n partons process in GR@PPA

GR@PPA [58] is an event generator based on the lowest order matrix element(ME) calculation. The matrix elements are supplied by the GRACE [63] system which is an automatic amplitude calculation program with the helicity amplitude method. There is no approximation for the calculation of the MEs: all fermion masses are nonzero, the helicity information can be traced down to the final state particles. Especially, we would like to mention that the quark mixing effect with CKM parameters (including 3×3 matrix) is also included in the diagram calculation. Hence, the exact treatment of the massive flavor productions of c- or b-quarks, and the process such as a single b-quark production which will be useful for the measurement of the single top production process are possible.

Since the "jet" is any kind of parton flavors at the parton level, the Feynman diagrams for the jet production depend on the processes with each parton flavor combination. The jet production in the GR@PPA is expressed as a summation over all the MEs with different parton combinations in the final state. All the subprocesses in the n jets(partons) process of GR@PPA are thus classified by the initial/final partons combinations and the electroweak/strong coupling order. In Table 7.1, we list up all the subprocesses in the W + n partons (n $\leq$ 3) processes of the GR@PPA. The first column is the number of partons associated with the W boson. The second is the coupling order, where α_{em} and α_s are the electroweak and strong coupling constants, respectively. The third is the subprocesses considered in the ME calculations. ${q_u}^i({q_d}^i)$ is an up(down)-type quark in the i-th generation, where the i is integer number 1,2 for up-type quarks and 1,2,3 for down-type quarks, and g is a gluon. W boson can decay to any kind of fermions, but no interference between the decay products from W and the other final state partons is assumed in the ME calculation. Note that the ME is considered up to the decay particles of the W boson (see the coupling order of each subprocess). The numbers of Feynman diagrams in the ME calculation to each subprocess under unitary gauge can be seen in the forth column. Because the diagram structure is symmetrical for the z-direction (beam axis), actual number of diagrams should be multiplied by 2. In the fifth column, it is useful to show that the weighting fraction for the parton flavor associated with the W boson is related to the quark mixing of CKM (Cabibbo-Kobayashi-Maskawa Matrix) parameter, and the subprocess number to run the GR@PPA appears in the last column. In total, 144, 2,376, and 18,632 diagrams and their interference terms are dealt exactly in the ME calculations of W + 1, 2, and 3 partons processes.

7.2.2 Input parameters

Electroweak parameters :

The electroweak interaction is characterized by only three (plus Higgs mass) parameters[1]. The other parameters related to the electroweak interaction are obtained by those three parameters. If one treats the other electroweak parameters as free parameters, one breaks a gauge invariance, then the calculation results become meaningless. Because the choice of three parameters however is completely arbitrary, the parameters precisely measured by experiments are preferred.

The G_μ scheme is traditionally taken for this purpose.

$$(G_F,\ M_W,\ M_Z)\ =\ (1.16639\times 10^{-5},\ 80.419,\ 91.188)\quad , \tag{7.1}$$

where G_F is the Fermi constant, M_W [64] and M_Z [64] are masses of W and Z boson, respectively. We obtain the other parameters related to the electroweak interaction:

$$\alpha_{em}\ =\ 1/132.51\quad ,\quad \sin^2\theta_W\ =\ 0.2222\quad .$$

Width scheme :

The presence of virtual unstable massive gauge bosons gives rise to singularities in the phase space due to the massive gauge boson propagators, which display poles at $q^2 = M^2$, where q^2 becomes the invariant mass of the decay products of the unstable bosons. These singularities are cured by the introduction of the finite

[1] Additionally, a gauge invariance also demands fermion masses in Yukawa sector.

W + n partons	Coupling order	Subprocess	Diagrams (× 2)	V_{CKM}^{jj}	ISUB
		$q_u{}^i(\bar{q}_u{}^i) + g \rightarrow W^{\pm}(f\bar{f}) + q_d{}^j(\bar{q}_d{}^j)$	24	(i, j)	422
1	$\alpha_{em}^2\alpha_s$	$\bar{q}_d{}^i(q_d{}^i) + g \rightarrow W^{\pm}(f\bar{f}) + \bar{q}_u{}^j(q_u{}^j)$	24	(j, i)	
		$q_u{}^i(\bar{q}_u{}^i) + \bar{q}_d{}^j(q_d{}^j) \rightarrow W^{\pm}(ff) + g$	24	(i, j)	423
		$q_u{}^i(\bar{q}_u{}^i) + g \rightarrow W^{\pm}(ff) + q_d{}^j(\bar{q}_d{}^j) + g$	96	(i, j)	424
		$\bar{q}_d{}^i(q_d{}^i) + g \rightarrow W^{\pm}(f\bar{f}) + \bar{q}_u{}^j(q_u{}^j) + g$	96	(j, i)	
		$q_u{}^i(\bar{q}_u{}^i) + \bar{q}_d{}^j(q_d{}^j) \rightarrow W^{\pm}(ff) + q_u{}^i(\bar{q}_u{}^i) + \bar{q}_u{}^k(q_u{}^k)$	72	(k, j)	
		$q_u{}^i(\bar{q}_u{}^i) + \bar{q}_d{}^j(q_d{}^j) \rightarrow W^{\pm}(f\bar{f}) + q_d{}^j(\bar{q}_d{}^j) + \bar{q}_d{}^k(q_d{}^k)$	96	(i, k)	425
		$q_u{}^i(\bar{q}_u{}^i) + \bar{q}_d{}^j(q_d{}^j) \rightarrow W^{\pm}(f\bar{f}) + q_u{}^k(q_d{}^k) + \bar{q}_u{}^k(\bar{q}_d{}^k)$ (k ≠ i(j))	72	(i, j)	
		$q_u{}^i(\bar{q}_u{}^i) + \bar{q}_d{}^j(q_d{}^j) \rightarrow W^{\pm}(ff) + g + g$	60	(i, j)	426
2	$\alpha_{em}^2\alpha_s^2$	$g + g \rightarrow W^{\pm}(ff) + \bar{q}_u{}^i(q_u{}^i) + q_d{}^j(\bar{q}_d{}^j)$	96	(i, j)	427
		$q_u{}^i(\bar{q}_u{}^i) + \bar{q}_u{}^j(q_u{}^j) \rightarrow W^{\pm}(ff) + \bar{q}_u{}^j(q_u{}^j) + q_d{}^k(\bar{q}_d{}^k)$	72	(i, k)	
		$q_d{}^i(\bar{q}_d{}^i) + \bar{q}_d{}^j(q_d{}^j) \rightarrow W^{\pm}(f\bar{f}) + \bar{q}_u{}^k(q_u{}^k) + q_d{}^i(\bar{q}_d{}^i)$	96	(k, j)	428
		$q_u{}^i(\bar{q}_d{}^i) + \bar{q}_u{}^i(q_d{}^i) \rightarrow W^{\pm}(f\bar{f}) + \bar{q}_u{}^j(q_u{}^j) + q_d{}^k(\bar{q}_d{}^k)$ (i ≠ j(k))	72	(j, k)	
		$q_u{}^i(\bar{q}_u{}^i) + q_u{}^j(\bar{q}_u{}^j) \rightarrow W^{\pm}(ff) + q_u{}^{i,j}(\bar{q}_u{}^{i,j}) + q_d{}^k(\bar{q}_d{}^k)$	96	((i,j), k)	429
		$\bar{q}_d{}^i(q_d{}^i) + \bar{q}_d{}^j(q_d{}^j) \rightarrow W^{\pm}(f\bar{f}) + \bar{q}_d{}^{i,j}(q_d{}^{i,j}) + \bar{q}_u{}^k(q_u{}^k)$	144	(k, (i,j))	
		$q_u{}^i(\bar{q}_u{}^i) + q_d{}^j(\bar{q}_d{}^j) \rightarrow W^{\pm}(ff) + q_d{}^j(\bar{q}_d{}^j) + q_d{}^k(\bar{q}_d{}^k)$	72	(i, k)	430
		$\bar{q}_u{}^i(q_u{}^i) + \bar{q}_d{}^j(q_d{}^j) \rightarrow W^{\pm}(f\bar{f}) + \bar{q}_u{}^i(q_u{}^i) + \bar{q}_u{}^k(q_u{}^k)$	48	(k, j)	
		$q_u{}^i(\bar{q}_u{}^i) + g \rightarrow W^{\pm}(ff) + q_d{}^j(\bar{q}_d{}^j) + g + g$	336	(i, j)	431
		$\bar{q}_d{}^i(q_d{}^i) + g \rightarrow W^{\pm}(f\bar{f}) + \bar{q}_u{}^j(q_u{}^j) + g + g$	336	(j, i)	
		$q_u{}^i(\bar{q}_u{}^i) + g \rightarrow W^{\pm}(ff) + q_d{}^j(\bar{q}_d{}^j) + q_{u,d}{}^k(q_{\bar{u},d}{}^k) + q_{\bar{u},d}{}^k(q_{u,d}{}^k)$	864	(i, j)	
		$q_u{}^i(\bar{q}_u{}^i) + g \rightarrow W^{\pm}(f\bar{f}) + q_d{}^j(\bar{q}_d{}^j) + q_u{}^i(\bar{q}_u{}^i) + \bar{q}_u{}^k(q_u{}^k)$ (k ≠ i)	144	(k, j)	
		$\bar{q}_d{}^i(q_d{}^i) + g \rightarrow W^{\pm}(f\bar{f}) + \bar{q}_u{}^j(q_u{}^j) + q_{\bar{d},u}{}^k(q_{d,u}{}^k) + q_{d,u}{}^k(q_{\bar{d},u}{}^k)$	864	(i, j)	432
		$\bar{q}_d{}^i(q_d{}^i) + g \rightarrow W^{\pm}(f\bar{f}) + \bar{q}_u{}^j(q_u{}^j) + \bar{q}_d{}^i(q_d{}^i) + q_d{}^k(\bar{q}_d{}^k)$ (k ≠ i)	288	(j, k)	
		$q_d{}^i(\bar{q}_d{}^i) + g \rightarrow W^{\pm}(f\bar{f}) + q_d{}^i(\bar{q}_d{}^i) + q_d{}^j(\bar{q}_d{}^j) + \bar{q}_u{}^k(q_u{}^k)$	436	(k, j)	
		$\bar{q}_u{}^i(q_u{}^i) + g \rightarrow W^{\pm}(f\bar{f}) + \bar{q}_u{}^i(q_u{}^i) + \bar{q}_u{}^j(q_u{}^j) + q_d{}^k(\bar{q}_d{}^k)$	288	(j, k)	
		$q_u{}^i(\bar{q}_u{}^i) + \bar{q}_d{}^j(q_d{}^j) \rightarrow W^{\pm}(ff) + q_u{}^i(\bar{q}_u{}^i) + \bar{q}_u{}^k(q_u{}^k) + g$	432	(k, j)	
3	$\alpha_{em}^2\alpha_s^3$	$q_u{}^i(\bar{q}_u{}^i) + \bar{q}_d{}^j(q_d{}^j) \rightarrow W^{\pm}(f\bar{f}) + q_d{}^j(\bar{q}_d{}^j) + \bar{q}_d{}^k(q_d{}^k) + g$	576	(i, k)	433
		$q_u{}^i(\bar{q}_u{}^i) + \bar{q}_d{}^j(q_d{}^j) \rightarrow W^{\pm}(f\bar{f}) + q_u{}^k(q_d{}^k) + \bar{q}_u{}^k(\bar{q}_d{}^k) + g$ (k ≠ i(j))	432	(i, j)	
		$q_u{}^i(\bar{q}_u{}^i) + \bar{q}_d{}^j(q_d{}^j) \rightarrow W^{\pm}(ff) + g + g + g$	144	(i, j)	434
		$g + g \rightarrow W^{\pm}(ff) + \bar{q}_u{}^i(q_u{}^i) + q_d{}^j(\bar{q}_d{}^j) + g$	576	(i, j)	435
		$q_u{}^i(\bar{q}_u{}^i) + \bar{q}_u{}^j(q_u{}^j) \rightarrow W^{\pm}(ff) + \bar{q}_u{}^j(q_u{}^j) + q_d{}^k(\bar{q}_d{}^k) + g$	432	(i, k)	
		$q_d{}^i(\bar{q}_d{}^i) + \bar{q}_d{}^j(q_d{}^j) \rightarrow W^{\pm}(f\bar{f}) + \bar{q}_u{}^k(q_u{}^k) + q_d{}^i(\bar{q}_d{}^i) + g$	576	(k, j)	436
		$q_u{}^i(\bar{q}_d{}^i) + \bar{q}_u{}^i(q_d{}^i) \rightarrow W^{\pm}(f\bar{f}) + \bar{q}_u{}^j(q_u{}^j) + q_d{}^k(\bar{q}_d{}^k) + g$ (i ≠ j(k))	432	(j, k)	
		$q_u{}^i(\bar{q}_u{}^i) + q_u{}^j(\bar{q}_u{}^j) \rightarrow W^{\pm}(ff) + q_u{}^{i,j}(\bar{q}_u{}^{i,j}) + q_d{}^k(\bar{q}_d{}^k) + g$	576	((i,j), k)	437
		$\bar{q}_d{}^i(q_d{}^i) + \bar{q}_d{}^j(q_d{}^j) \rightarrow W^{\pm}(f\bar{f}) + \bar{q}_d{}^{i,j}(q_d{}^{i,j}) + \bar{q}_u{}^k(q_u{}^k) + g$	864	(k, (i,j))	
		$q_u{}^i(\bar{q}_u{}^i) + q_d{}^j(\bar{q}_d{}^j) \rightarrow W^{\pm}(ff) + q_d{}^j(\bar{q}_d{}^j) + q_d{}^k(\bar{q}_d{}^k) + g$	432	(i, k)	438
		$\bar{q}_u{}^i(q_u{}^i) + \bar{q}_d{}^j(q_d{}^j) \rightarrow W^{\pm}(f\bar{f}) + \bar{q}_u{}^i(q_u{}^i) + \bar{q}_u{}^k(q_u{}^k) + g$	288	(k, j)	

Table 7.1: All the subprocesses in the W + n partons (n $\leq$ 3) processes of the GR@PPA. The first column is the number of partons associated with the W boson. The second is the coupling order, where α_{em} and α_s are the electroweak and strong coupling constants, respectively. The third is the subprocesses considered in the ME calculations. $q_u{}^i(q_d{}^i)$ is an up(down)-type quark in the i-th generation, where the i is integer number 1,2 for up-type quarks and 1,2,3 for down-type quarks, and g is a gluon. The numbers of Feynman diagrams in the ME calculation to each subprocess under unitary gauge can be seen in the forth column. Because the diagram structure is symmetrical for the z-direction (beam axis), actual number of diagrams should be multiplied by 2. In the fifth column, the weighting fraction for the parton flavor associated with the W boson is related to the quark mixing of CKM parameter, and the subprocess number to run the GR@PPA is appeared in the last column.

widths of the gauge bosons, which shift the poles away from the real axis. However, in the field theory the widths arise from the imaginary parts of higher-order diagrams describing the gauge boson self-energies resummed to all orders. So the tree level amplitude is supplemented with only a subset of higher-order contributions and this can destroy the gauge invariance of the calculation, which becomes important when very different scales or high energies enter the process under consideration. In hadronic collisions, potential problems arise in multiboson production and in weak boson fusion. Theoretical solutions have been proposed in [65], [66], [67], and [68] since LEP experiment. It seems not trivial in the multi-partonic processes. This is opened issues for the future high energy hadron collider. For the moment, we use a fixed width scheme,

$$\frac{1}{q^2 - M^2} \quad \Rightarrow \quad \frac{1}{q^2 - M^2 + i\Gamma M} \quad , \tag{7.2}$$

to treat a boson propagater, which is more practical way for the single boson production and is already studied at LEP II.

Partial decay width and Branching ratio :

Although the total decay width of bosons can be calculated at the tree level, we use the experimental value of the total decay width as an input parameter in order to reproduce the realistic shape of the mass distribution. The partial decay width however is uniquely decided in the theoretical calculation. Hence, the total cross section is shifted from the correct theoretical prediction at the lowest order calculation to an unphysical result. We thus reweight the total cross section by the branching ratio measured by the experiments.

$$\sigma_{theo.}(X \to f\bar{f}) \;=\; \sigma_{calc.} \times \frac{\Gamma^{tot.}_{exp.}}{\Gamma^{part.}_{theo.}} \cdot Br_{exp.}(X \to f\bar{f}) \quad , \tag{7.3}$$

where $\sigma_{calc.}$ is the calculated cross section using the input total decay width from the experimental value $\Gamma^{tot.}_{exp.}$ when the boson X decay to fermions $f\bar{f}$, $\Gamma^{part.}_{theo.}$ is the partial decay width calculated by theory, and $Br_{exp.}(X \to f\bar{f})$ is the branching ratio of $X \to f\bar{f}$ from the experimental value. The $Br_{exp.}(X \to f\bar{f})$ is used as the input parameter. In the W + n partons process, the production cross section can be expressed as

$$\sigma_{W+n\,partons} \cdot Br(W \to e\nu) \;=\; \sigma_{theo.}(W_{\to e\nu} \,+\, n\,partons) \quad , \tag{7.4}$$

where the W decays to electron(positron) and neutrino. We use the experimental value [64] of the total decay width and the branching ratio of W boson,

$$\Gamma^{tot.}_{exp.} \;=\; 2.12 \text{ GeV} \quad , \quad Br(W \to e\nu) \;=\; 10.66\,\% \quad . \tag{7.5}$$

CKM parameters :

Since we do not identify the b-quark jet using b-tagging algorithms, the partons in the W + n partons should contain b-quark flavors. We use the following squared CKM parameters:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} |V_{ud}| = 0.9752 & |V_{us}| = 0.2210 & |V_{ub}| = 0.0054 \\ |V_{cd}| = 0.2210 & |V_{cs}| = 0.9743 & |V_{cb}| = 0.0419 \\ |V_{td}| = 0.0054 & |V_{ts}| = 0.0419 & |V_{tb}| = 0.9991 \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad , \tag{7.6}$$

where V_{xy} is the mixing angle coupling with an up-type x quark and a down-type y quark. Note that we also consider the b-quark content in PDF in the ME calculations because of the finite size of the mixing with a b-quark V_{ub} and V_{cb}.

Fermion masses :

Because there is no reason to neglect the quark masses, we assign the quark masses,

$$\begin{aligned} m_d = 5 \text{ MeV} \quad , \quad m_s = 500 \text{ MeV} \quad , \quad m_b = 4.7 \text{ GeV} \\ m_u = 2 \text{ MeV} \quad , \quad m_c = 1.5 \text{ GeV} \quad . \end{aligned} \tag{7.7}$$

Proton Distribution Function :

The Les Houches Accord Parton Density Function (LHAPDF) interface [69] was conceived at the Les Houches 2001 workshop [70] in the PDF working group to enable the usage of Parton Density Functions with uncertainties in a uniform manner. When PDFs with uncertainties are considered, a “fit” to the data no longer is described by a single PDF. Instead in its most flexible implementation, a fit is represented by a PDF set consisting of many individual PDF members. Calculating the observable for all the PDF members enables one to reconstruct the uncertainty on the observable. The LHAPDF interface was made with this in mind and manipulates PDF sets.

We use CTEQ6 (**cteq6.LHpdf**) averaging over 40 PDFs [71] with the strong coupling $\alpha_s(M_Z^2) = 0.1180$.

7.2.3 Kinematical cuts

We use the JetClu algorithm, which is a fixed cone algorithm, for a jet clustering. The cone algorithm has been widely used in hadron collider experiments, because a bunch of hadrons from the parton in the hard-collision distributes within a certain cone as the observable object, where a cone is defined as $\Delta R = \sqrt{\phi^2 + \eta^2}$. However, once one may trace back to the original partonic process in the hard-collision, one encounters a notoriously known as an infrared/collinear problem discussed in Section 5. Our interests are a partonic process from the perturbative calculation.

On the other hand, an infrared/collinear problem is also the inherent problem in the lowest order ME calculation. For instance, how do we decide the kinematical cuts at the generator level? The total cross section increases infinitely in the smaller separation angle ΔR between two partons. In the lowest order perturbative calculation, it is not clear where is the border on the calculation limit. We thus have made one assumption that only one parton exists in a jet, so-called, “parton-jet matching”. Two partons in the single jet are not allowed. Note that this assumption is only valid at the lowest order. As one example, let us consider why this assumption is reasonable. When we generate an electron in the LO MC, it is also possible to calculate a photon along with the electron. If this photon goes to closer with the electron, the cross section will be much larger than one of the process without the photon. And this collinearly emitted photon is never detected as an isolated photon. Then, in our experiment, we measure as one electron signature. However, no one uses this calculation result (LO ME) to measure the single electron. We tacitly assume one LO ME electron in one observable electron. That is, if we have two particles in the one observable object in the LO calculation, the LO ME calculation is somewhere broken. (Of course, if one wants to do that, one uses NLO calculation.)

In order to apply the parton-jet matching, we follow a prescription based on the jet separation method described in Section 5. The kinematical cuts of the event generators in hadron collisions generally contain three cuts, (p_T, $|\eta|$, $\Delta R_{p,p}$), for all final state particles. The $\Delta R_{p,p}$ cut is the separation angle cut between two partons if the ME has more than two partons. In the jet separation method, the $\Delta R_{p,p}$ is constrained in order to satisfy the parton-jet matching. We illustrate the schematic picture of the jet separation method for the kinematical cuts on the analysis and generator level in Figure 7.1. From the top-left, the phase space for the parton kinematics by p_T, η, and $\Delta R_{p,p}$ is shown, respectively. The p_T and $|\eta|$ cuts at the generator level are usually decided to be the lower values which cover the whole phase space of the experimental analysis. That is, the parton p_T threshold at the generator level will be lower than the jet $p_T(E_T)$ threshold at the experimental analysis level. Since the events with lower jet E_T than the jet E_T threshold are discarded, the phase space by the theory calculation finally corresponds to the phase space by the experimental analysis. However, we need to pay attention to the $\Delta R_{p,p}$ cut while the p_T and $|\eta|$ cuts are trivial to decide. Events are not discarded even in the lower $\Delta R_{p,p}$ cut if we do not require the parton-jet matching. A naive $\Delta R_{p,p}$ cut results in the overestimated phase space, which does not corresponds to the phase space by the experimental analysis. If we use a simple parton-jet matching scheme that events are rejected if the parton does not match with the jet within a certain $\Delta R_{p,j}$, the theoretical cross section is biased by that parton-jet matching scheme. Instead of a rejection of the un-matched events, the jet separation method gives an exact point of the $\Delta R_{p,j}$ cut, then replaces that rejection with a new systematic uncertainty of the parton-jet matching. Due to the requirement from the jet separation method, we set the kinematical cuts:

$$p_T \;>\; 8 \text{ GeV} \;, \quad |\eta| \;<\; 3.0 \;, \quad \Delta R_{p,p} \;>\; R_{sep}^{95\%} \quad , \tag{7.8}$$

for the final state partons and no cut for leptons.

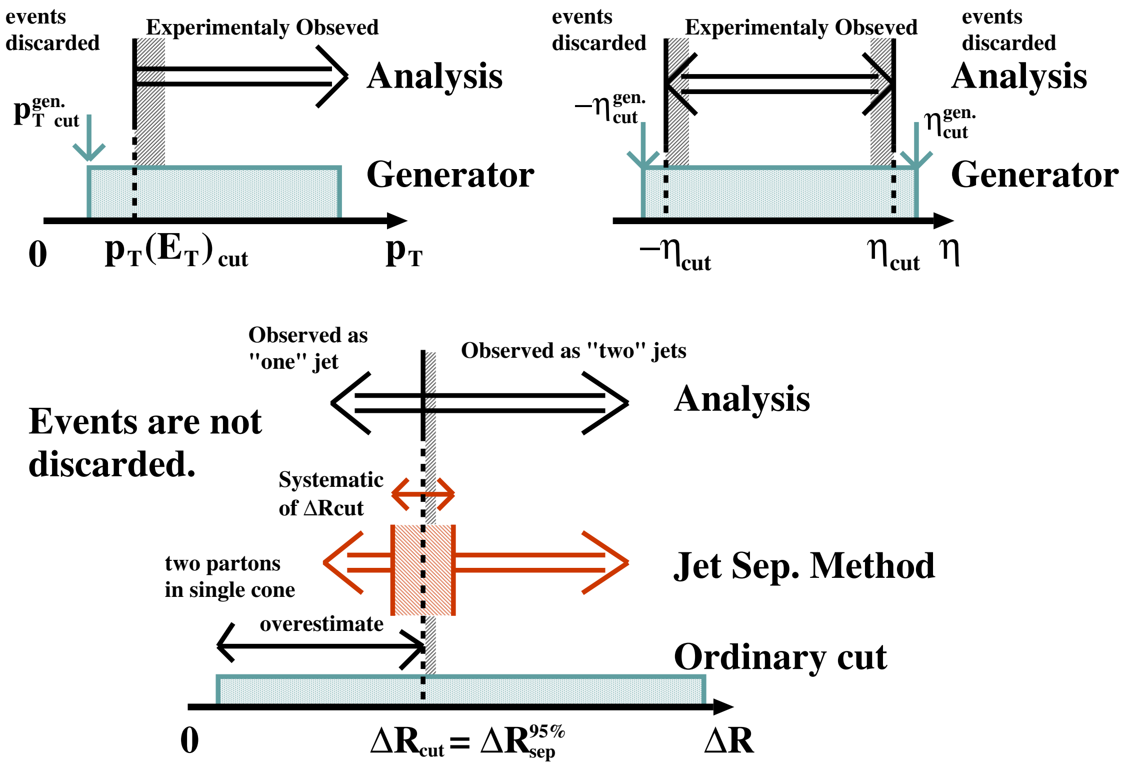


Figure 7.1: Schematic picture of the jet separation method for the kinematical cuts on the analysis and generator level. From the top-left, the phase space for the parton kinematics by p_T, η, and $\Delta R_{p,p}$ is shown, respectively.

7.3 Cross section results

We summarize the production cross sections of W + n partons processes in Table 7.2. The cross sections are estimated with two different energy scales with respect to two different ΔR cuts. The energy scales used are M_W^2 and $\langle\, p_T \,\rangle^2$ (GeV^2), where the renormalization and factorization scales are equivalent. The $\langle\, p_T \,\rangle^2$ is a squared average of parton p_T's defined as

$$\langle\, p_T \,\rangle^2 \;=\; \left(\frac{\sum^{n} p_T^i}{n} \right)^2 \quad , \tag{7.9}$$

where n is the number of partons. Although our kinematical cut of the $\Delta R_{p,p}$ at generator level is the $R_{sep}^{95\%}$, it is practical to show the cross sections in the case of the $\Delta R_{p,p} > 0.4$. The other kinematical cuts are the same as those of Eq. (7.8). In Table 7.2, the first column presents the number of partons, the second is the subprocess identification number used in GR@PPA listed in Table 7.1, and the others are the estimated cross sections in unit of pb.

The results of the case of $\Delta R_{p,p} = R_{sep}^{95\%}$ is smaller than those of $\Delta R_{p,p} = 0.4$. The size of differences for each subprocess depends on the number of calculated partons. The fluctuations by the energy scale M_W^2 to $\langle\, p_T \,\rangle^2$ is also smaller in the case of $\Delta R_{p,p} = R_{sep}^{95\%}$ than that of $\Delta R_{p,p} = 0.4$, and the subproceses with gluons in the final state tend to have larger fluctuations than those with quarks. The dominant contributors in each n parton process is the gluon initiated process, `424`, `431` in 2, and 3 partons processes, except the 1 parton process. We thus have a relation that the jet multiplicity is approximately proportional to a magnitude of the strong coupling constant if the n parton processes are dominated by one particular subprocess

$$f((\geq)\, n\, jets) \;\approx\; C_{exp.}\, (\alpha_s)^n \quad , \tag{7.10}$$

where $C_{exp.}$ is the jet-definition parameters decided by an experiment such as the cone size used in the clustering of jets, the transverse momentum, rapidity and separation cuts for jets and so forth. Note that the ratio to each Eq. (7.10) for n partons processes approximately leads to a cancellation of the term $C_{exp.}$.

Finally, we present our results used in this analysis in Table 7.3. Note that this numbers are affected by the experimental condition after the detector simulation. Almost 6 - 20 % of events survived due to the jet E_T cut efficiency or overall reconstruction efficiency of the jet clustering.

n partons	ISUB	Cross Section (pb)			
		$\Delta R_{p,p} = R_{sep}^{95\%}$		$\Delta R_{p,p} = 0.4$	
		M_W^2	$\langle p_T \rangle^2$	M_W^2	$\langle p_T \rangle^2$
1	422	239.6(4)	340.8(5)	239.6(4)	340.8(5)
	423	426.5(6)	576.3(9)	426.5(6)	576.3(9)
2	424	95.5(1)	182.7(3)	102.2(1)	196.0(3)
	425	12.56(2)	21.47(5)	14.35(3)	24.69(7)
	426	63.1(1)	114.9(2)	75.6(1)	137.7(3)
	427	3.377(5)	7.33(1)	3.457(5)	7.50(1)
	428	6.89(1)	11.22(3)	6.96(1)	11.35(3)
	429	2.844(4)	4.809(8)	2.871(4)	4.846(8)
	430	3.222(5)	5.40(1)	3.256(5)	5.45(1)
3	431	24.92(3)	60.9(1)	31.66(4)	77.2(1)
	432	1.449(4)	3.46(1)	1.924(5)	4.59(1)
	433	5.32(1)	11.28(7)	6.88(2)	14.74(6)
	434	8.48(1)	19.63(5)	13.08(2)	30.06(7)
	435	1.537(2)	4.37(1)	1.760(2)	5.02(1)
	436	4.11(2)	8.13(4)	4.79(2)	9.50(4)
	437	1.181(3)	2.542(9)	1.384(3)	2.955(8)
	438	1.363(3)	2.88(1)	1.591(3)	3.39(1)

Table 7.2: Production cross sections of W + n partons processes. The cross sections are estimated with two different energy scales with respect to two different ΔR cuts. The energy scales used are M_W^2 and $\langle\ p_T\ \rangle^2$ (GeV^2), where the renormalization and factorization scales are equivalent. The $\langle\ p_T\ \rangle^2$ is a squared average of parton p_T's. The first column is the number of partons, the second is the subprocess identification number used in GR@PPA listed in Table 7.1, and the others are the estimated cross sections in unit of pb. Although our kinematical cut of the $\Delta R_{p,p}$ at generator level is the $R_{sep}^{95\%}$, it is practical to show the cross sections in the case of the $\Delta R_{p,p} > 0.4$. The other kinematical cuts are the same as those of Eq. (7.8).

	Cross Section (pb)			
	0 p	1 p	2 p	3 p
$Q^2 = M_W^2$	1939(0)	665.5(8)	187.8(3)	48.36(7)
$Q^2 = \langle\ p_T\ \rangle^2$	1939(0)	917(1)	348.2(5)	113.4(2)

Table 7.3: The total cross sections used in this analysis. All processes are produced by GR@PPA except W + 0 parton process of HERWIG built-in process. In the W + 0 parton process, the energy scale was taken as $\hat{s}$ in colliding partons.

7.4 Scale dependence

Physical quantities should be independent of a particular scheme used to renormalize the theory. This scheme is just a calculation method how the infinite divergence such as a collinear singularity is factored out into PDF or coupling constants at a particular input energy scale. The renormalization scale μ_R is the scale at which the coupling constants are evaluated, and the factorization scale μ_F is the scale to separate the perturbative domain region of a hard-interaction from the non-perturbative region of PDF. These schemes however lead to uncertainties to a dependence on the factorization/renormalization scales, since the perturbative corrections beyond a given order of the scheme may still remain as the size of missing higher-order corrections. The results of perturbative QCD calculations, in general, are a scale dependent. In fact, in the lowest order, these scales are the most ambiguity part having an effect to physical quantities. There seems to be no way to decide the best choice of the energy scale, especially for multi-partonic processes, although an effective choice can be seen in [9]. According to the past analysis at CDF [11], we have taken $\mu_F = \mu_R = \langle\ p_T\ \rangle^2$ and M_W^2 instead of $M_W^2 + p_{TW}^2$ [2].

It is practical to see a dependence of the energy scale with the different choice of μ_F and μ_R. To keep the same condition with our experimental analysis at the theoretical calculation, we take the following kinematical cuts:

$$p_T \ > \ 15 \text{ GeV} \quad , \quad |\eta| \ < \ 2.4 \quad , \quad \Delta R_{p,p} \ > \ R_{sep}^{95\%} \qquad \text{for parton}$$

[2] The total cross section with the energy scale of M_W^2 is not so much different from that of $M_W^2 + p_{TW}^2$.

$$
\begin{aligned}
p_T &> 20 \text{ GeV} \quad , \quad |\eta| < 1.1 && \text{for electron} \\
p_T &> 20 \text{ GeV} && \text{for neutrino.}
\end{aligned} \tag{7.11}
$$

In Figure 7.2, we present the variation of the cross section for the W + n partons processes as a function of various energy scales μ_R and μ_F. The circle, square, and open-circle marks represent the cross sections for the W + 1, 2, and 3 partons processes, respectively. The energy scales are separately taken as the fixed μ_F (= M_W in GeV) with a scale variation of μ_R (dashed line), and the fixed μ_R (= M_W in GeV) with a scale variation of μ_F (dash-dotted line) together with $\mu_R = \mu_F$ (solid line). The PDF used is CTEQ6M with NLO running α_s.

We can see less dependence on μ_F (dash-dotted line) in every W + n partons process, of which order is $\sim 1/\log(\mu_F^2)$ with increasing of μ_F. The fluctuation of the cross section of $\mu_F \sim 10$ GeV to 200 GeV is around ∼30% independent of the number of partons. On the other hand, the dependence on μ_R (dashed line) depends on the number of partons, that is, the order of α_s. It is easy to imagine this behavior if we recall the formula of the α_s. As these results, the fluctuation of the cross section with the same variation of μ_F and μ_R (solid line) is strongly correlated with $\sim(\alpha_s(\mu_R^2))^n/\log(\mu_F^2)$, where n is number of partons. The energy scale of $\langle\ p_T\ \rangle^2$ distributes around ∼30 GeV. At last, we would like to emphasize that the scale dependece may disappear in the higher-order calculation. The NLO calculation plays a key role for this analysis.

7.5 Uncertainty of PDF

A convenient way of evaluating PDF uncertainties is to express the uncertainty in a set of individual PDFs. By evaluating the observable for each of the members of the PDF set, the PDF uncertainty can be determined. It is extracted by scanning the 40 sets of the CTEQ6 compilation. These sets are supposed to cover the systematics of 1 σ level involved in the extraction of PDF from the existing experimental data. The range of variation obtained by changing the PDF sets in given in Table 7.4. To keep the same

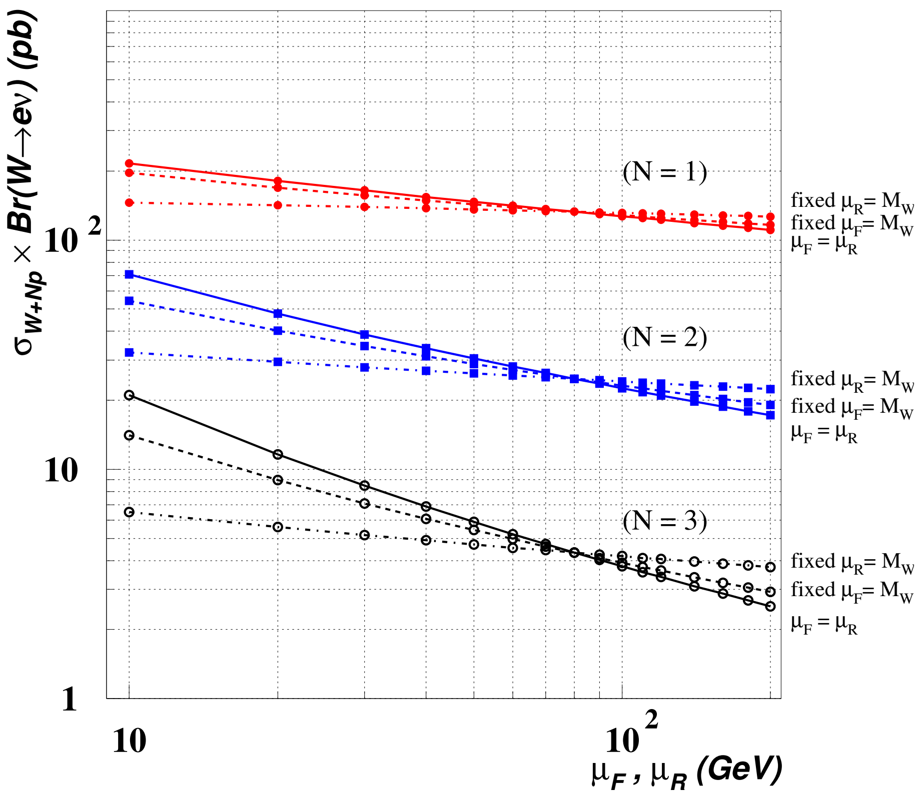


Figure 7.2: Variation of the cross section for the W + n partons processes as a function of various energy scales μ_R and μ_F. CTEQ6M is used for the estimations.

condition with our experimental analysis, the kinematical cuts at the generator level are the same as Eq. (7.11).

The table shows the maximum and minimum deviation scanned by 40 sets together with the nominal cross sections using CTEQ6M. The bracket on the next of values in maximum and minimum deviation indicates the label of the PDF set for which those values were found. The systematics in this table is defined as

$$syst. = \frac{(max.\ value) - (min.\ value)}{(max.\ value) + (min.\ value)} \quad . \tag{7.12}$$

The size of the systematics is around $1.4 \sim 3.4\%$ depending on the parton multiplicity. There is not so large difference between the different energy scales of M_W^2 and $\langle\ p_T\ \rangle^2$.

Cross Section (pb)	$W(\rightarrow e\nu)$ + 1p		$W(\rightarrow e\nu)$ + 2p		$W(\rightarrow e\nu)$ + 3p	
	M_W^2	$\langle\ p_T\ \rangle^2$	M_W^2	$\langle\ p_T\ \rangle^2$	M_W^2	$\langle\ p_T\ \rangle^2$
CTEQ6M	132.6(1)	166.5(1)	24.76(4)	37.0(1)	4.311(7)	7.27(1)
max.(PDF set)	+2.0 (8)	+2.6 (19)	+0.37 (8)	+0.5 (8)	+0.206 (30)	+0.21 (30)
min.(PDF set)	−2.1 (37)	−3.2 (37)	−0.41 (7)	−0.4 (10)	−0.091 (29)	−0.14 (10)
Systematics(%)	1.5	1.7	1.6	1.4	3.4	2.4

Table 7.4: except W + 0 parton process.

Since the overall fluctuation by PDF uncertainty is much smaller than one by the scale variation, we do not consider this contribution in this analysis.

7.6 Heavy flavor fraction

In our analysis, we do not discriminate a heavy flavor fraction in our jet samples. In the top-quark physics, those heavy flavor contents are very important information to reject the huge backgrounds containing non-heavy flavor jets, where b-quark is the heavy flavor in consideration. Since our sample is one of those backgrounds, it is useful to see a heavy flavor fraction in the W + n parton processes. This is useful for the analysis [72] using the b-quark fraction in W + jets sample. We present the c-quark and b-quark fractions in W + n partons processes in Table 7.5. Two energy scales of $\mu_R = \mu_F = M_W^2$ and $\langle\ p_T\ \rangle^2$ are considered. The fraction by the PDF uncertainty is also presented as the maximum/minimum deviation founded in a certain PDF set. The kinematical cuts at generator level are the same as Eq. (7.11). For the calculation, CTEQ6M was used as well as the uncertainty study of PDF described in the previous section. Note that any non-massless quarks are considered in calculation.

The fraction is basically less than ∼1.5% and the size of the PDF uncertainty is around ∼7 - 10% with increasing the number of partons. There is no dependence on the different energy scales. The dependence may be canceled out by taking the ratio of the heavy flavor fractions. The odd number of c-(b-)quarks are significant source of the uncertainty of PDF, because the “sea” quarks are extracted inside PDF, then those partons appear in

Process	c-quark fraction(%)				b-quark fraction(%)			
(Q^2 (GeV^2))	0c	1c	2c	3c	0b	1b	2b	3b
W+1p (M_W^2)	96.0(3)	3.90(6)			99.9(3)	0.005(2)		
max.(PDF set)	+0.1 (26)	+0.14 (17)	N/A	N/A	+<0.1 (3)	+0.004 (37)	N/A	N/A
min.(PDF set)	−0.1 (17)	−0.13 (26)			−<0.1 (37)	−0.004 (3)		
W+2p (M_W^2)	92.7(3)	6.13(7)	1.11(3)		98.8(3)	0.16(1)	0.97(3)	
max.(PDF set)	+0.2 (29)	+0.20 (30)	+0.07 (14)	N/A	+<0.1 (7)	+0.02 (35)	+0.06 (1)	N/A
min.(PDF set)	−0.1 (37)	−0.26 (10)	−0.07 (26)		−<0.1 (1)	−0.04 (18)	−0.05 (35)	
W+3p (M_W^2)	90.8(3)	7.30(8)	1.80(4)	0.030(5)	98.0(3)	0.20(1)	1.75(4)	0
max.(PDF set)	+0.4 (29)	+0.54 (30)	+0.11 (33)	+0.015 (25)	+0.1 (40)	+0.03 (7)	+0.01 (2)	+0.001 (16)
min.(PDF set)	−0.5 (30)	−0.46 (29)	−0.04 (30)	−0.009 (36)	−<0.1 (2)	−0.03 (26)	−0.14 (18)	−<0.001
W+1p ($\langle p_T \rangle^2$)	96.0(3)	3.99(6)			99.9(3)	0.006(2)		
max.(PDF set)	+0.1 (18)	+0.09 (17)	N/A	N/A	+<0.1 (29)	+0.003(8,37)	N/A	N/A
min.(PDF set)	−<0.1 (17)	−0.18 (18)			−<0.1 (8)	−0.003 (29)		
W+2p ($\langle p_T \rangle^2$)	92.5(3)	6.40(8)	1.01(3)		98.9(3)	0.14(1)	0.93(3)	
max.(PDF set)	+0.3 (18)	+0.21 (12)	+0.11 (32)	N/A	+<0.1 (6)	+0.01 (30)	+0.04 (35)	N/A
min.(PDF set)	−0.2 (12)	−0.38 (18)	−0.03 (7)		−<0.1 (17)	−0.04 (36)	−0.07 (6)	
W+3p ($\langle p_T \rangle^2$)	90.5(3)	7.57(8)	1.81(4)	0.038(6)	98.2(3)	0.16(1)	1.62(4)	0
max.(PDF set)	+0.2 (21)	+0.35 (30)	+0.11 (19)	+0.012 (38)	+<0.1 (23)	+0.02 (36)	+0.15 (31)	+<0.001
min.(PDF set)	−0.3 (17)	−0.23 (19)	−0.08 (12)	−0.013 (26)	−0.1 (31)	−0.03 (23)	−0.03 (33)	−<0.001

Table 7.5: Fractions of c-quark and b-quark in W + n partons processes. Two energy scales of $\mu_R = \mu_F = M_W^2$ and $\langle p_T \rangle^2$ are considered. The fraction by the PDF uncertainty is also presented as the maximum/minimum deviation founded in a certain PDF set. The kinematical cuts at generator level are the same as Eq. (7.11). For the calculation, CTEQ6M was used as well as the uncertainty study of PDF.

the final state particles without flavor changing by the weak current interaction with a W boson. Or another possibility is a flavor changing to the c-(b-)quarks by the weak current interaction. In the odd number of b-quarks, the CKM mixing with the third generation also contributes.

7.7 Initial state radiation

After a perturbative calculation by the ME to the hard interaction in a $p\bar{p}$ collision, one have to apply a parton showering (PS), and then hadronization, fragmentation and so forth after the parton showering in order to make the event a realistic one expanded to non-perturbative area. While the parton showering is very powerful method to describe all order summation of the (next-to-)leading log term in a perturbative expansion, it is not easy to combine the measured cross section in an experimental analysis with the theoretical cross section from ME's, because the parton showering does not change the theoretical cross section. The additional jets from the parton showering increase/decrease the jet production rate. One needs to be careful if the observable object corresponds to the theoretical one. Unfortunately, the cone jet algorithm is not suitable for this purpose. We thus measure a jet production rate ($\geq$ n jets) inclusively, as the results are already shown in the previous section. Since the parton-jet matching is the way only to predict stable theoretical results, the inclusive jet measurement is the better suitable analysis scheme [73], when one discusses the overall jet production rate by using the cone jet algorithm. The style of k_T clustering algorithm [49] might be more useful for future analysis.

Using the jet separation method, a probability that the final state radiation (FSR) makes additional jets is largely suppressed. However, there is no correlation with the kinematics between partons produced by a ME and initial state radiation (ISR). The ISR is thus able to make the additional isolated jets. The emission probability of the parton showering is characterized by an input energy scale which decides a starting point of the evolution. In the case of a simple $2 \to 2$ process, that scale of ISR is usually taken as the factorization scale (μ_F). In a multi-partonic process, a typical scale however is not clear as well as no typical energy scale (μ_F, μ_R) in the ME calculation. Therefore, we have three independent free parameters in the ME+PS calculation. We have forced the two scales of μ_F and μ_R to be the same value in calculation. It would be better to constrain the scale of PS as the first order approximation. We propose an assumption that : if we compare the parton p_T of

- the largest p_T distribution (max.{ p_T(ME(W+1jet)) , p_T(PS) }) in ME(W+1p) + PS process, and
- the largest p_T distribution (max.{ p_T(PS) }) in ME(W+0p) + PS process,

then, both p_T distributions should be the same in the low p_T region, since the leading-log (LL) terms described by PS is a good approximation in the very low p_T region ($\sim$

a few GeV). In the case of ME(W+0p) + PS process, there is no overlapping phase space between ME and ISR, because of no final state parton in ME(W+0p). Thus, the ISR describes the LL-term correctly. On the other hand, the ME(W+1p) + PS process potentially has a source of the double counting between ME and PS. The reliable PS algorithm should be the same distribution with the case of the max.$\{p_T(\text{PS})\}$ in ME(W+0p) + PS process. For this purpose, two PS algorithms of PYTHIA version 6.203 and HERWIG version 6.5[3] are tested. We present the comparison plots for PYTHIA and HERWIG in Figure 7.3, and 7.4, respectively. The solid and dot-dashed line represent the largest p_T distribution in ME(W+1p) + PS events with $p_T > 1$ GeV and 500 MeV for the parton in ME(W+1p), respectively. The dashed line represents the largest p_T distribution in ME(W+0p) + PS events, which is purely produced by the PS. The same ME's are fed into PYTHIA and HERWIG, that is, the ME's is not built-in ME's in PYTHIA or HERWIG. The cut-off scale which stops a PS evolution is 1 GeV and 2.5 GeV for PYTHIA and HERWIG, respectively. The kink structure in those plots corresponds to this scale. In the top-right hand side on those plots, each ratio of ME(W+1p) + PS to ME(W+0p) + PS is also shown.

For the HERWIG case, all three distributions show good agreement with ech other in the range of $4 < p_T < 10$ GeV. Since the parton from PS does not reach above the region of $p_T > 10$ GeV, the parton from ME(W+1p) is dominant in the largest p_T region, so that it is natural that both distributions of ME(W+1p) + PS and ME(W+0p) + PS processes are different. The region below the cut-off scale of PS is out of consideration. On the other hand, the behavior of PS in PYTHIA is completely different with one of HERWIG's. In the region of $p_T > 10$ GeV, that behavior in ME(W+1p) + PS process is almost the same as that of HERWIG, because the parton from ME(W+1p) is dominant in the largest p_T region. The behavior in ME(W+0p) + PS of PYTHIA is also slightly smaller than that of HERWIG. In the range of $4 < p_T < 10$ GeV in our interest, both p_T distributions in ME(W+1p) + PS are larger than one in ME(W+0p) + PS. The dependence of the p_T

[3]There is no significant difference between version 6.4 and 6.5.

cut (500 MeV, 1 GeV) in ME(W+1p) can be seen. Those feature is due to the different PS algorithms of PYTHIA and HERWIG. While the PS algorithm of PYTHIA ver.6.203 is a primitive PS algorithm which is controlled by the input evolution scale of PS, the PS algorithm of HERWIG ver.6.5 adopts a veto algorithm [74] which uses resummation techniques. In this example, the partons with a larger p_T are emitted by the primitive PS of PYTHIA with the evolution scale of M_W^2, so that the p_T distribution in large p_T region in ME(W+1p) + PS does not match with the case of ME(W+0p) + PS of which treatment of PS is presumably theoretically correct in a low p_T region. Since we do not prefer to use a new additional parameter in the theoretical prediction, we use HERWIG showering Monte Carlo simulation.

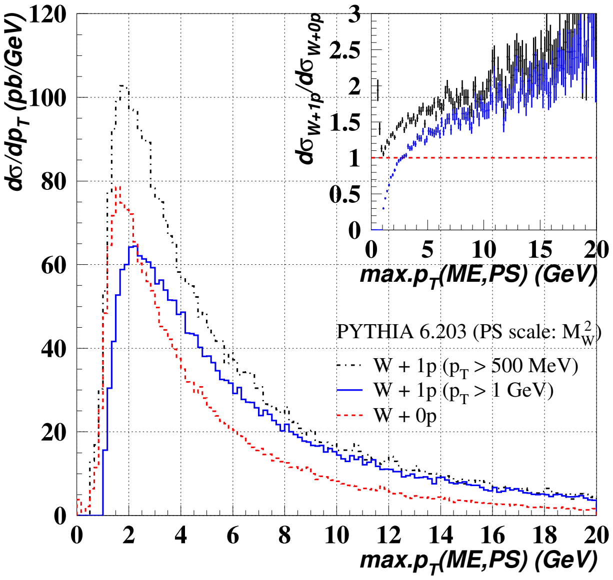


Figure 7.3: Comparison of the largest p_T distribution in the ME(W+1p)+PS and ME(W+0p)+PS processes in PYTHIA.

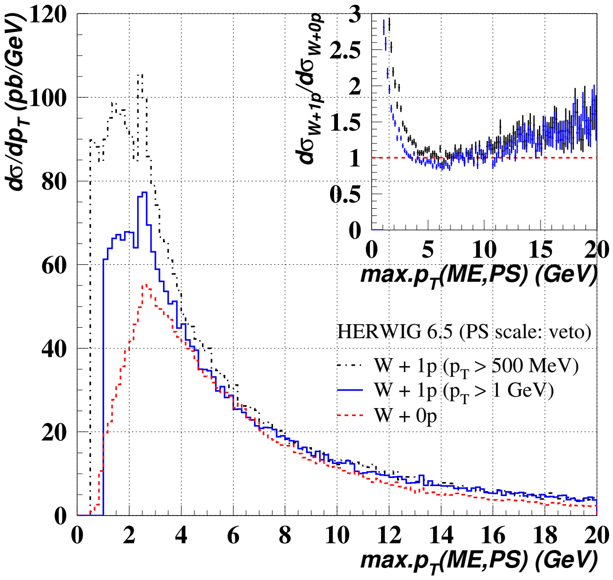


Figure 7.4: Comparison of the largest p_T distribution in the ME(W+1p)+PS and ME(W+0p)+PS processes in HERWIG.

7.8 Summary

As the lowest order prediction of the W + n partons processes, we have studied some aspects: input parameters, kinematical cuts at generator level, scale dependence, PDF uncertainty, heavy flavor fraction, and parton showering of the initial state radiation.

The GR@PPA event generator was used in the Matrix Element(ME) calculation and event generation. The noticeable feature of this event generator is the exact treatment of the lowest order matrix elements with large number of Feynman diagrams. We described the input parameters and width correction scheme used in this analysis in detail, and explained the kinematical cuts at generator level based upon the jet separation method by the simple illustration about the difference of the phase space between analysis level and theoretical level. The detailed numbers of the cross sections used in this analysis are also presented in this chapter. The cross section was found to have the μ_F dependence of $\sim 1/\log(\mu_F^2)$, and the μ_R dependence of $\sim \alpha_s^n$. Then, we found that our choice of $\mu_R = \mu_F = M_W^2$, $\langle\, p_T \,\rangle^2$ tends to have the dependence of the convolution of $\sigma(\mu_R \otimes \mu_F) \sim \alpha_s^n/\log(\mu_F^2)$. For the uncertainty of PDF, we found that magnitude of the uncertainty is negligibly small than the fluctuation of the energy scale. The NLO calculation plays a key role to solve the scale dependence.

By the way, the heavy flavor fraction in the W + n partons processes has also been studied. All finite quark mass and 3×3 CKM mixing allow to estimate the heavy flavor fraction with the PDF uncertainty. Since the "sea" quark content in PDF is a large systematic source, the fraction is affected by the PDF uncertainty. On the other hand, the fraction has less dependence on the energy scale uncertainty. This means the ratio like a heavy flavor fraction approximately cancels the scale dependence. In the lowest order prediction, it would be useful to use some sort of the ratio rather than the absolute value.

To make up a reliable events to use at the detector level, we apply the parton showering (PS) method (+ hadronization) for the events produced by ME calculation. The parton-jet matching(Jet Sep. Method) suppresses a probability that the parton showering makes additional jets. However, the initial state radiation unfortunately is detected as the isolated jets at detector level, because there is no kinematic correlation between ME and PS. Therefore, there is no way to measure exclusively n jet events as long as using cone jet algorithms if one discusses the absolute prediction. And one additional energy scale exists to decide the evolution of the PS. We prefer to constrain this scale rather than using as a free parameter. Two PS algorithms by PYTHIA version 6.203 and HERWIG 6.4(5) have been tested with an assumption that an appropriate treatment reasonably describes the low energy behavior. As a result, we chose HERWIG Monte Carlo simulation.

As the summary, we use GR@PPA + HERWIG + (detector simulation) as the lowest order prediction for this analysis.

Chapter 8

Results and Conclusions

We have studied the W + $\geq$ n jets process in Tevatron Run II experiment. This is the first result for the CDF Run II experiment. The CDF is successfully taking the collision data since 2002. The data used in this analysis correspond to a total integrated luminosity of 72.0 pb^{-1} taken from March 2002 through January 2003. The high-p_T electron triggered samples are used.

In Chapter 3, we described the selection criteria for the $W(\rightarrow e\nu)$ candidate events, and measurement of the inclusive $W(\rightarrow e\nu)$ production cross section. After the selection of W boson, we had 31,726 $W(\rightarrow e\nu)$ candidate events. Since we use a slightly higher $\not\!\!E_T$ cut and lower electron E_T cut to reject a multijet background, the inclusive $W(\rightarrow e\nu)$ production cross section is compared with the result of the W production cross section measurement in CDF Run II. We had a good agreement between both results.

After good qualitative cuts on an isolated high-p_T electron and a requirement of an imbalance of calorimeter energy due to the undetected neutrino (missing E_T), the JetClu algorithm is used to collect/count a jet. The transverse energy and pseudo-rapidity (η) coverage of jets are required as

$$E_T \geq 15 \text{ GeV} \quad , \quad |\eta| \leq 2.4 \, . \tag{8.1}$$

The clustering cone size is 0.4. The jet energy correction is applied up to the partonic level correction which estimates a parton energy based on a certain simulation, although this energy correction leads to large systematics. The merging/splitting criteria is followed by the Jet Separation Method which requires the iterative separation cone between two jets with 95% separation probability estimated by the two partons at the lowest order calculation. The detailed descriptions for the jet separation of the cone jet algorithms can be seen in Chapter 4. We collect the jets samples inclusively, that is, group the $W + \geq$ n jets event samples, where, for instance, an event with 2 jets is a member of the $W + \geq$ 2 jets event sample but at the same time it can be a member of the $W + \geq 1$ jets event sample.

After obtaining the $W + \geq$ n jets event samples, the data have been compared to the theory predictions at the lowest order perturbative calculation level. Jet Separation procedure based on the parton-jet matching requirement is used for the data and theoretical predictions. This requirement is to construct the enhanced LO phase space. With a requirement of the parton-jet matching in the lowest order perturbative calculation, we can suppress undesirable features of the multi-partonic processes in pQCD. As the results, the theory prediction would be meaningful. Chapter 5 dedicates to the description of this parton-jet matching procedure. For the theory prediction, two choices of the energy scale, $\langle\ p_T\ \rangle^2$ and M_W^2, where the renormalization and factorization scales are equivalent, has been tested.

In Chapter 6, the jet transverse energy, mass and jet-jet separation distribution are compared with data and theory predictions. All the distributions show good agreements. The choice of the energy scale of $\langle\ p_T\ \rangle^2$ is preferred to describe data well. Jet multiplicity distribution is also compared up to $W + \geq 3$ jets events. Non-background condition is assumed. But those background effects are almost negligible in the less than 3 jets events. The constant (flat) behavior can be seen in the various ratio plots. This means that the MC generation and our analysis scheme are right, which is crucial for the measurement of the strong coupling constant. As a comparison result, Run II results are slightly larger than Run I results.

It is practical to see a characteristic feature of $W +$ n parton process in the lowest order perturbative calculation. We have studied some aspects: input parameters, kinematic cuts at generator level, scale dependence, PDF uncertainty, heavy flavor fraction, and parton showering of the initial state radiation. Details are described in Chapter 7. A newly developed GR@PPA event generator (Appendix A) was used in the Matrix Element calculation and event generation. To make up a reliable events to use at the detector level, we apply the HERWIG parton showering model (+ hadronization) for the events produced by GR@PPA event generator.

In conclusion, we proposed a clear logic to be able to get a stable prediction even in the lowest order perturbative calculation by requiring our method (Jet Separation Method) of a parton-jet matching. By applying this to real data, we found the current scheme for the theory prediction (LO + LL) is not so out-of-tune. Although the lowest order perturbation has a big ambiguity of the scale dependence, the comparison of the CDF data with the lowest order prediction would be a good benchmark toward to the higher order calculation. Also, taking some ratios of the measured physics quantities was proved to be a valid way to avoid a scale dependence of the lowest order perturbation. Needless to say, the NLO calculation plays a key role to proceed this analysis in future.

Appendix A

GR@PPA Event Generator

A.1 Introduction

Understanding of the electroweak symmetry breaking on TeV scale is the most important mission for the high energy physics program in the present and future colliders of Tevaron, LHC, and Linear Collider. Among the searches for Higgs boson(s) and supersymmetric particles, most of them have the multi-particle final state in their signature. Besides, the nominal Standard Model processes also contribute as background processes. The optimal prediction thus is an essential part of the experimental analysis.

In collider experiments, a theoretical calculation using Monte Carlo method has been widely used in many aspects for new physics searches or a precision measurements. Recent development of the Monte Carlo event generator is represented as [55] and [57]. In our previous work [58], we have developed a complete set of four bottom quarks in the final state at the lowest order perturbative calculation within the Standard Model framework in $pp/p\bar{p}$ collisions, known as GR@PPA_4b event generator. We have succeeded to escape the large number of diagrams in calculation by adopting one additional integration variable in the initial parton flavors coming from the proton density function (PDF). This paper is for the extended work of GR@PPA framework to treat a flavor of the final state parton, where the same technique with the treatment of the initial parton flavor is used. And new processes are added into the Matrix Element calculations as well as $b\bar{b}b\bar{b}$ processes.

We named this extended version GR@PPA_All.

The most noticeable feature of this event generator is the exact treatment of the lowest order matrix elements with large number of all possible Feynman diagrams within the Standard Model. All fermion masses are nonzero, the helicity information can be traced down to the final state particles, and γ/Z mediated processes can be also taken into account exactly. Non-standard processes are also potentially possible [75] in the future GR@PPA framework. In addition, 3×3 CKM matrix are embedded, for instance, as to produce the odd number of b-quarks in the final state. Hence, GR@PPA will be useful to study heavy flavor productions or polarization processes as well as multibody processes. Since GR@PPA outputs the helicity state in the final state particles, it is also preferable to connect the decay package like TAUOLA [76].

This paper is organized as follows: the overview of GR@PPA framework is described in Section A.2. The description of parameters in GR@PPA_All is given in Section A.3. All details about running this program are given in Section A.4. Some physical results are presented in Section A.5. A summary is given in Section A.6.

A.2 Extension of GRACE to $pp/p\bar{p}$ collisions

A.2.1 Overview

GR@PPA is an event generator based on the lowest order matrix element calculation. The matrix elements are supplied by the GRACE system [63] which is an automatic amplitude calculation program with the helicity amplitude method. The core part of the program thus is to calculate a matrix element of a parton-level hard interaction and then, to provide the unweighted events.

Since the GRACE system originally has been developed aiming at applications to lepton-lepton collisions, the generated codes were not directly applicable to hadron-hadron collision interactions. We have developed an extended framework, called GR@PPA (GRACE at PP/Anti-p), to implement those features specific to hadron collisions [77]. A similar extension of GRACE has been realized in a previous work, GRAPE [78], for *ep* collisions. In the present work we adopt a different method (an embedding method), expecting an improvement in the usability of the program.

The primary function of GR@PPA is to determine the initial/final-state partons, *i.e.* their flavor and momenta. The functional block diagram of the GR@PPA system is drawn in Figure A.1. Based on the GRACE output codes, GR@PPA calculates the cross section and generates unweighted parton-level events using BASES/SPRING [79] included in the GRACE system. An extension has been done between BASES/SPRING and the GRACE output code. Since the GR@PPA framework is just an interface to determine the parton species and momenta, it has a possibility that automatically handles the GRACE output code in future.

The output unweighted events are finally stored in Les Houches Accord (LHA) common block in a common user interface, which was proposed at Physics at TeV Colliders II Workshop, Les Houches, France, 2001 [80]. Using this interface, the GR@PPA programs can be totally embedded in any kind of the parton showering programs having the Les Houches Accord, for instance, which is available in PYTHIA version 6.2 [48] and HERWIG version 6.5 [47]. We also keep an old style used in PYTHIA version 6.1 only for the $b\bar{b}b\bar{b}$ processes in the new version of GR@PPA.

A.2.2 GRACE system

As shown in Figure A.1, the basic elements of the system, which are the same as the original GRACE system, are the “GRACE output code” and BASES/SPRING. The “GRACE output code” is a set of FORTRAN codes for calculating the matrix element of a specified process, according to a set of kinematical variables specifying a phase-space point. BASES/SPRING is a multi-dimensional general-purpose Monte Carlo integration and event-generation program set. It generates a set of random numbers to give them to an external function. Using the returned answer, BASES performs an integration and SPRING generates “events” by means of a hit/miss method. The most remarkable feature of the BASES/SPRING system is the utilization of a multi-dimensional grid method

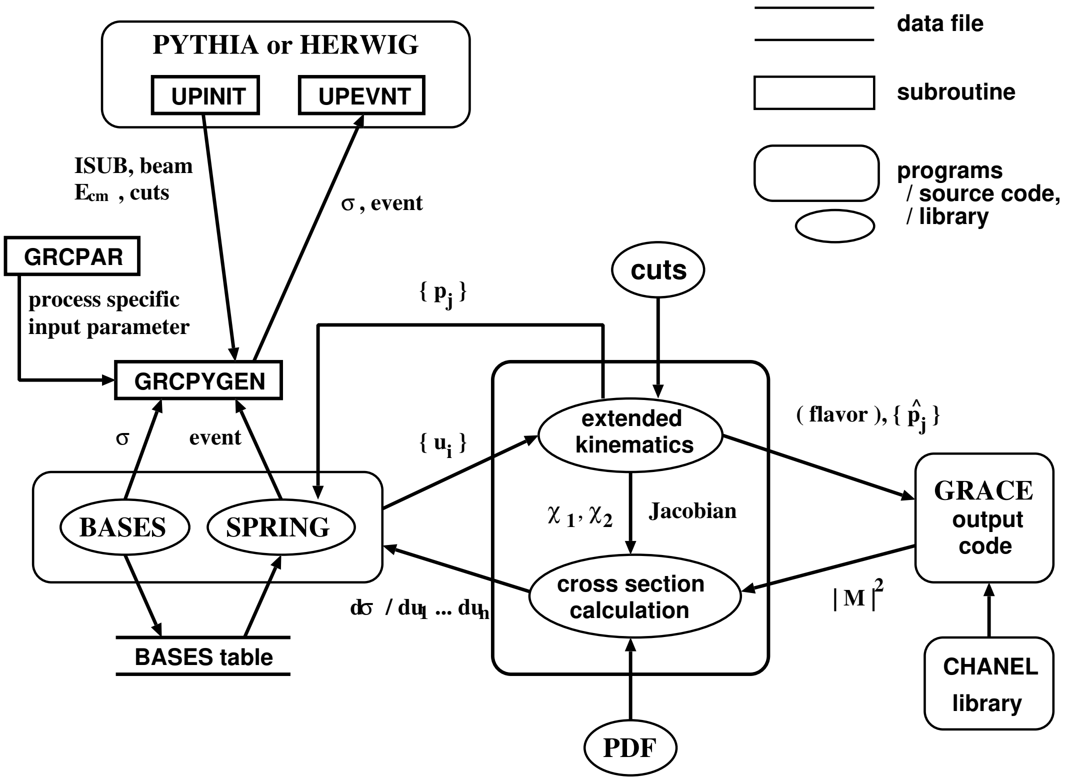


Figure A.1: Picture showing the structure of GR@PPA. The data flow is indicated schematically. The main building blocks of the GRACE-based event generator are BASES/SPRING and "GRACE output codes". The interface between them has been extended for pp and $p\bar{p}$ collisions. A set of random numbers $\{u_i\}$ given by BASES/SPRING includes two numbers for defining the initial state (x_1 and x_2), in addition to those for defining the final state. The cross section is calculated from the matrix element returned from the "GRACE output codes", by referring to a PDF using x_1 and x_2. Some phase-space cuts are applied by limiting the range of kinematical variables, or by setting the cross section to zero after the kinematical variables are determined. In some cases, several subprocesses are combined to a single subprocess by adding one more random number for defining the quark flavor. This system is interfaced to a generic user common of Les Houches Accord through `UPINIT` and `UPEVNT`.

for the random number generation. BASES optimizes the grid setting by an iteration to maximize the efficiency of the integration and the event generation. The optimized setting is stored in an external file (BASES table) to be used later in event generation by SPRING. The codes can be automatically generated using a utility included in the GRACE system. The codes for the processes in Table **??** have been generated by the authors and included in the GR@PPA distribution.

A.2.3 Kinematics

Cross sections with a hard interaction in $pp/p\bar{p}$ collisions can be described as

$$\sigma = \sum_{i,j,F} \int dx_1 \int dx_2 \int d\hat{\Phi}_F f_i^1(x_1, Q^2) f_j^2(x_2, Q^2) \frac{d\hat{\sigma}_{ij\to F}(\hat{s})}{d\hat{\Phi}_F}, \tag{A.1}$$

where $f_i^a(x_a, Q^2)$ is a PDF of the hadron a (p or $\bar{p}$), which gives the probability to find the parton i with an energy fraction x_a at a probing virtuality of Q^2. The differential cross section $d\hat{\sigma}_{ij\to F}(\hat{s})/d\hat{\Phi}_F$ describes the parton-level hard interaction producing the final-state F from a collision of partons, i and j, where $\hat{s}$ is the square of the total initial 4-momentum. The sum is taken over all relevant combinations of i, j and F. Note that in hadron interactions a certain "process" of interest may contain some incoherent subprocesses having different final states, as well as those having different combinations of the initial-state partons. For example, the "two-jet" production process includes all $q\bar{q}'$, $qg(\bar{q}g)$ and gg production processes.

The original GRACE system assumes that both the initial and final states are well-defined. Hence, it can be applied to evaluating $d\hat{\sigma}_{ij\to F}(\hat{s})/d\hat{\Phi}_F$ and its integration over the final-state phase space $\hat{\Phi}_F$ only. An adequate extension is necessary to take into account the variation of the initial and final state both in parton species and their momenta, in order to make the GRACE system applicable to hadron collisions.

The interface has to convert the random numbers given by BASES/SPRING to a set of kinematical variables necessary for the matrix element calculation ("kinematics"), and to convert the returned matrix element to the differential cross section. Singular structures such as the $1/k$ singularity of the photon/gluon radiation and Breit-Wigner resonance structures, has to be taken into account in the conversion to the kinematical variables, using their well-known asymptotic forms. Although the grid method of BASES/SPRING is very flexible and practically very powerful, itself is not capable of dealing with these singularities without any care. In order to determine the initial-state variables, x_1 and x_2, we require BASES/SPRING to provide two additional random numbers in addition to those for the kinematics at $\hat{s}$. Due to a $1/x$ asymptotic behavior of the structure functions, it is convenient for Monte Carlo integration and event generation to rewrite Eq.(A.1) as

$$\sigma = \sum_{i,j,F} \int \frac{d\tau}{\tau} \int dy \int d\hat{\Phi}_F x_1 f_i^1(x_1, Q^2) x_2 f_j^2(x_2, Q^2) \frac{d\hat{\sigma}_{ij\to F}(\hat{s})}{d\hat{\Phi}_F}, \tag{A.2}$$

where,

$$\tau = x_1 x_2 \ , \qquad y = \frac{1}{2} \ln \frac{x_1}{x_2}. \tag{A.3}$$

In GR@PPA, the added two random numbers are converted to τ and y, while taking into account the asymptotic $1/\tau$ behavior for τ, and assuming a flat probability distribution for y. The variable τ determines the center-of-mass (cm) energy of the hard interaction, since $\hat{s} = \tau s$. The variables x_1 and x_2 are derived using Eq.(A.3) in order to refer to PDFs in the conversion of the returned matrix element to the differential cross section, as shown in Figure A.1. The interface finally returns the calculated differential cross section to BASES/SPRING, and at the same time converts the kinematical variables in the cm frame to ones in the laboratory frame, by applying a Lorentz boost determined by y.

A.2.4 Flavor assignment

As already mentioned, a "process" of interest is usually composed of several incoherent subprocesses in hadron interactions. However, the present version of BASES/SPRING can treat only one subprocess at the same time. This does not matter in BASES. It is sufficient to do the integration and the grid optimization sequentially for these subprocesses, one after the other. On the other hand, this is a serious limitation in event generation by

SPRING, because we frequently want to generate events of different subprocesses in a random order.

We applied a slight modification to SPRING to overcome this difficulty. The "BASES table" is prepared for every subprocess by running BASES sequentially over the subprocesses. The modified SPRING works as follows: when SPRING is called at the first time, all relevant "BASES tables" are read into a tentative memory area. The main "BASES table" to be used for random-number generation is replaced in each event, by copying an appropriate one from the tentative memory. This method works well because entire information specific to subprocesses, such as the optimized grid information and the cross section information, is recorded in the "BASES table".

Although we successfully extended the BASES/SPRING to multiple subprocesses, the number of subprocesses is desired to be as small as possible because we have to prepare not only the "BASES table", but also the "GRACE output codes" for every subprocess. In many cases, the difference between the subprocesses is the difference in the quark combination in the initial and/or final states only. The matrix element of these subprocesses is frequently identical, or the difference is only in a few coupling parameters and/or masses. In such cases, it is convenient to add one more integration/differentiation variable to replace the summation in Eqs.(A.1) and (A.2) with an integration. As a result, these subprocesses can share an identical "GRACE output code" and can be treated as a single subprocess by BASES/SPRING. This extension is implemented in GR@PPA for the subprocesses that has the quarks in the initial and/or final state.

A.2.5 Interface to Les Houches Accord

As shown in Figure A.1, GR@PPA includes an interface to a generic process interface LHA format. Using this facility, we can add the effect of initial- and final-state parton showers to the generated events. This effect emerges as a finite overall p_T of the hard interaction system and finite underlying activities. Furthermore, if we activate the hadronization and decay, we can obtain realistic events which can be passed to detector simulators.

The LHA interface routines are supplied by `UPINIT` and `UPEVNT`. The `UPINIT` is used for the initialization of the parameter set commonly used in the showering Monte Carlo and the external generator. The colliding beam and centre-of-mass energy *etc.* are set in this routine. The `UPEVNT` is used for the event generation of the showering Monte Carlo. The output events from the external generator must be stored in the LHA common block in the `UPEVNT`. According to this procedure, GR@PPA is also initialized in `UPINIT` and produced events in `UPEVNT`.

The initialization of GR@PPA parameters are handled in `GRCPAR`. The interface to access to GR@PPA then is the subroutine `GRCPYGEN`, which controls BASES/ SPRING and, as a result, controls all GR@PPA routines. The subroutine `GRCPYGEN` calls BASES or SPRING according to the mode selection determined by an input argument. The total cross section is evaluated in the initialization stage (`UPINIT`) with the mode selection for calling BASES. In the event generation cycle (`UPEVNT`), `GRCPYGEN` is called again with the mode selection for calling SPRING. Since the event generation is totally controlled by SPRING in GR@PPA, the rejection mode by means of a hit/miss method in LHA is not activated any more. The returned argument of `GRCPYGEN` is always set equal to the total cross section evaluated by BASES.

The calling sequence of `GRCPYGEN` is as follows:

```
CALL GRCPYGEN(BEAMS, ISUB, MODE, SIGMA)    ,
```

where the input arguments are

`BEAMS (CHARACTER)` : 'PP' for pp collisions and 'PAP' for $p\bar{p}$ collisions
`ISUB (INTEGER)` : subprocess number
`MODE (INTEGER)` : = 1 for calling BASES, and 0 for calling SPRING,

and the output is

`SIGMA (REAL*8)` : integrated cross section.

The argument `BEAMS` is dummy when `MODE` = 0. PYTHIA requires users to assign a unique subprocess number `ISUB` to every user-defined subprocess. The output `SIGMA` is always equal to the integrated cross section of the subprocess specified by `ISUB`.

The most important task of `GRCPYGEN` in the event-generation cycle is to pass the event information determined in GR@PPA to LHA common blocks. The interfacing rules are all specified by LHA. The information concerning the parton species and momenta, which has been determined in the "kinematics" routines and passed through the user interface routine of SPRING, is copied to the arrays in the common of LHA. The color flow information, which is necessary to perform hadronization, is also recorded, based on the information from SPRING [81].

A.3 GR@PPA_All

A.3.1 Subprocesses

The subprocesses supplied by GR@PPA is limited at the moment since an automatic amplitude generation system of GR@PPA is under development. We however provide these processes with the best optimized kinematics removed the singular behavior in the kinematics. All the subprocesses produced by GR@PPA_All are listed in Table A.1. The $p(\bar{p})$ represents a proton(antiproton) composed of u, d, c, s, b-quarks, and gluon. The j is a jet composed of u, d, c, s, b-quarks, and gluon as well as the p. The $f(\bar{f})$ is fermions (quarks and leptons) allowed to decay from vector bosons. The subprocesses are assigned by `ISUB` number. The total number of Feynman diagrams computed under unitary gauge in each subprocess are also listed in this table. The combinations of decaying fermions f are not account in this number. Because the diagram structure is symmetrical for the z-direction (beam axis), actual number of diagrams should be multiplied by 2. The arguments in the forth column, α_s, α_{em}, and y_x, are the coupling constants for strong and electroweak interactions, and Yukawa interaction with the x-quark, respectively. All subprocesses are classified by those coupling orders. The Higgs boson is assumed to be the minimal SM one. If users intend to be an extended Higgs boson beyond SM, users must put on the proper mass and width, and then normalize to the appropriate cross section. We present the total cross section of each subprocess in the last two columns. The Tevatron and LHC condition are considered as a bench mark point. The input parameters used in this calculation are the default parameters listed in Table A.2 except that the invariant mass of opposite charged lepton pair is larger than 20 GeV, $m_{\ell^+\ell^-} > 20$ GeV, and the choice of the energy scale and PDF. The energy scale ($\mu_F = \mu_R$) is fixed as M_Z^2 (GeV^2), and CTEQ6 (LHA PDF) is used ($\alpha_s(M_Z^2) = 0.1180$).

The integration accuracy achieved in BASES is fairly better than 1% for all subprocesses. Note that a large number of diagrams which are hard to manage manually are included in calculation. The detailed description for each subprocess is appeared in later section.

ISUB	Process	Diagrams ($\times$ 2)	Coupling order	Tevatron σ(2TeV)(pb)	LHC σ(14TeV)(pb)
100	$p + p(\bar{p}) \to W^{\pm}(f\bar{f}) \quad (+X)$	12	(α_{em}^2)	2194(3)	18656(24)
101	$p + p(\bar{p}) \to W^{\pm}(f\bar{f}) + j \quad (+X)$	144	$(\alpha_{em}^2\ \alpha_s)$	225.7(3)	3592(7)
102	$p + p(\bar{p}) \to W^{\pm}(f\bar{f}) + 2\ j \quad (+X)$	2,376	$(\alpha_{em}^2\ \alpha_s^2)$	38.69(7)	1364(3)
103	$p + p(\bar{p}) \to W^{\pm}(f\bar{f}) + 3\ j \quad (+X)$	18,632	$(\alpha_{em}^2\ \alpha_s^3)$	6.47(1)	542(1)
110	$p + p(\bar{p}) \to Z^0/\gamma^*(f\bar{f}) \quad (+X)$	5	(α_{em}^2)	409.5(3)	3154(3)
111	$p + p(\bar{p}) \to Z^0/\gamma^*(f\bar{f}) + j \quad (+X)$	60	$(\alpha_{em}^2\ \alpha_s)$	29.84(2)	465.9(5)
112	$p + p(\bar{p}) \to Z^0/\gamma^*(f\bar{f}) + 2\ j \quad (+X)$	690	$(\alpha_{em}^2\ \alpha_s^2)$	5.513(9)	187.5(3)
160	$p + p(\bar{p}) \to h^0(b\bar{b}) + b + \bar{b} \quad (+X)$	18	$(\alpha_s^2\ y_b^2)$	0.1627(3)(fb)	25.41(7)(fb)
161	$p + p(\bar{p}) \to Z^0/\gamma^*(b\bar{b}) + b + \bar{b} \quad (+X)$	84	$(\alpha_s^2\ \alpha_{em}^2)$	0.2142(1)	8.192(8)
162	$p + p(\bar{p}) \to b + \bar{b} + b + \bar{b} \quad (+X)$	38	(α_s^4)	5.14(1)	569(1)
163	$p + p(\bar{p}) \to h^0(b\bar{b}) + Z^0/\gamma^*(b\bar{b}) \quad (+X)$	25	$(\alpha_{em}^2\ y_b^2)$	6.35(2)(fb)	45.7(1)(fb)
164	$p + p(\bar{p}) \to Z^0/\gamma^*(b\bar{b}) + Z^0/\gamma^*(b\bar{b}) \quad (+X)$	60	(α_{em}^4)	0.01457(3)	0.1065(7)
170	$p + p(\bar{p}) \to t\bar{t} \to b\bar{b} + 4\ f \quad (+X)$	8	$(\alpha_s^2\ \alpha_{em}^4)$	0.07235(8)	6.635(9)
550	$p + p(\bar{p}) \to Z^0/\gamma^*(f\bar{f}) + Z^0/\gamma^*(f'\bar{f}') \quad (+X)$	40(*120)	(α_{em}^4)	0.07423(4)	0.7934(8)
551	$p + p(\bar{p}) \to W^+(f\bar{f}) + W^-(f'\bar{f}') \quad (+X)$	39(*79)	(α_{em}^4)	0.1053(2)	1.419(4)
552	$p + p(\bar{p}) \to Z^0/\gamma^*(f\bar{f}) + W^{\pm}(f'\bar{f}') \quad (+X)$	72(*120)	(α_{em}^4)	0.01159(1)	0.1263(2)

Table A.1: All subprocesses produced by GR@PPA_All. The $p(\bar{p})$ represents a proton(antiproton) composed of u, d, c, s, b-quarks, and gluon. The j is a jet composed of u, d, c, s, b-quarks, and gluon as well as the p. The $f(\bar{f})$ is fermions (quarks and leptons) allowed to decay from vector bosons. The subprocesses are classified by a unique identification number ISUB with different coupling orders. The diagram numbers in brackets of the diboson processes are including the cascade decay type from the decaying fermions f.

A.3.2 Parameters

Although the execution of GR@PPA is controlled by the subroutine GRCPYGEN, the detailed behavior depends on some parameters in common blocks and conditions defined in some subprograms. Users can change those details by changing appropriate parameters and subprograms. All the input parameters except the masses or widths and the other fundamental parameters are set in the subroutine GRCPAR. If the parameters are commonly used in any subprocesses, it is also possible to set in the subroutine UPINIT. The masses and widths, and the other fundamental parameters are set in the subroutine SETMAS. All details of those parameters are described in the following and listed in Table A.2.

Kinematic parameters :

The parameters that is necessary to be given by users are GRCECM, which specifies the cm energy of the beam collision in GeV with the colliding beam species by PP or PAP in the subroutine GRCPYGEN. Optionally, users can define some phase-space cuts in the laboratory frame: GRCPTCUT(I), GRCETACUT(I) and GRCRCONCUT(I). These parameters define the minimum p_T in GeV, the largest pseudorapidity in the absolute value and the minimum separation in $\Delta R(i, other)$ for the i-th particle, respectively. The separation ($\Delta R(i, other)$) is defined for the pair of the i-th particle (i) and every final-state particles ($other$) except gauge bosons as

$$\Delta R(i, other) \;=\; \sqrt{(\Delta\phi(i, other))^2 \,+\, (\Delta\eta(i, other))^2}, \tag{A.4}$$

where $\Delta\phi(i, other)$ and $\Delta\eta(i, other)$ are the separation in the azimuthal angle and the pseudorapidity, respectively. For ease usage, GPTCUT, GETACUT, and GRCONCUT are given for the partons composed of jets (j). The definition is the same as GRCPTCUT(I), GRCETACUT(I) and GRCRCONCUT(I), but these cuts are only applied for jets (j) (See Table A.1). In addition, users can define their own cuts by editing the subroutine GRCUSRCUT in which 4-momenta of all partons are provided through a common block. The instruction is described in the later section. An example is also attached to the file `grcpar.F`. Note that these kinematical cut parameters are applied after the kinematical variables of an event are determined.

`IGJFLV(I)` is an activation switch to decide jet flavors produced by the ME calculation. For instance,

$$\texttt{IGJFLV(I)} \;=\; 1,\,1,\,1,\,1,\,0,\,0,\,1 \quad (I \,=\, 1,7) \tag{A.5}$$

means a jet is composed of d, u, s, c-quarks, and $gluon$, but does not contain b and t-quarks. Currently, the t-quark contents in jets are not supported in any subprocesses of GR@PPA_ALL.

Electroweak parameters :

The electroweak interaction is characterized by only three (plus Higgs mass) parameters[1]. The other parameters related to the electroweak interaction are obtained by those three parameters. If one treat the other electroweak parameters as free parameters, one breaks a gauge invariance, then the calculation results will be meaningless. Because the choice of three parameters however is completely arbitrary, the parameters precisely measured by experiments are preferred. `IGAUGE` provides a choice of gauge invariant sets. For instance, `IGAUGE` = 1 (G_μ scheme) is

$$(G_F,\; M_W,\; M_Z) \;=\; (1.16639 \times 10^{-5},\; 80.419,\; 91.188) \quad , \tag{A.6}$$

where G_F is the Fermi constant, M_W and M_Z are masses of W and Z boson, respectively. We then obtain the other parameters related to the electroweak interaction:

$$\alpha_{em} \;=\; 1/132.51 \quad , \quad \sin^2\theta_W \;=\; 0.2222 \quad .$$

The fundamental parameters of masses or coupling constant are manually given in the subroutine `SETMAS`, which is described in later.

Although the total decay width of bosons can be calculated at the tree level, one may want to use the experimental value of the total decay width as an input parameter

[1] Additionally, a gauge invariance also demands fermion masses in Yukawa sector.

in order to reproduce the realistic shape of the mass distribution. The partial decay width however is uniquely decided in the theoretical calculation. Hence, the total cross section is shifted from the correct theoretical prediction at the lowest order calculation to an unphysical result. We thus provide one optional parameter `IWIDCOR` to reweight the total cross section by the branching ratio inputted by users.

$$\sigma_{theo.}(X \to f\bar{f}) \;=\; \sigma_{calc.} \times \frac{\Gamma^{tot.}_{exp.}}{\Gamma^{part.}_{theo.}} \cdot Br_{exp.}(X \to f\bar{f}) \quad , \tag{A.7}$$

where $\sigma_{calc.}$ is the calculated cross section using the input total decay width $\Gamma^{tot.}_{exp.}$ when the boson X decay to fermions $f\bar{f}$, $\Gamma^{part.}_{theo.}$ is the partial decay width calculated by theory, and $Br_{exp.}(X \to f\bar{f})$ is the input branching ratio of $X \to f\bar{f}$.

The mode selection of vector bosons decays are controlled by `IGWMOD(I)` and `IGZMOD(I)` with a flag of turning on/off:1/0. The branching ratio is set by `GRCWBR` and `GRCZBR`. Note that this input branching ratio is only activated when `IWIDCOR` = 2. The numbering for the decay mode follows the PYTHIA convention.

The Cabibbo-Kobayashi-Maskawa Matrix (CKM) parameters are given in

$$\mathcal{M}_{i,j} \;=\; \begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} \quad , \quad \left\{ \begin{array}{l} i \;=\; u, c, t \\ j \;=\; d, s, b \end{array} \right. \quad , \tag{A.8}$$

where V_{xy} is the mixing angle coupling with an up-type x quark and a down-type y quark. Note that we also consider the b-quark content in PDF in the ME calculations because of the finite size of the mixing with a b-quark $|V_{ub}|$ and $|V_{cb}|$.

BASES/SPRING parameters :

`IBSWRT` controls whether BASES should be called in the initialization or not. The task of BASES is to optimize the integration grids and, after that, store the optimized results in a "BASES table". The execution of BASES consumes much CPU time because a precise evaluation is necessary for an efficient event generation by SPRING. It is not necessary to repeat the execution for identical conditions. A previously optimized result ("BASES table") is reused if `IBSWRT = 1`. It should be noted that, once users change the parameters affect the estimated cross section such as kinematic or electroweak parameters, the "condition" is no longer identical and BASES has to be re-executed by setting `IBSWRT = 0`.

`NCALL` specifies the number of sampling points in each step of the iterative grid optimization in BASES. The larger this number is, the better the conversion would be. However, it takes longer in the CPU time. The optimized values are preset as the default values in each subprocess. The character variable `GRCFILE` gives the "BASES table" file name. A new file must be specified if `IBSWRT = 0`, while an existing file must be specified if `IBSWRT = 1`.

`IRSEED` is a seed of a random number used in SPRING event generator. Users can change this seed in the event generation cycle.

Process specific parameters :

`ICOUP` determines the energy scale (Q^2) for calculating the coupling strengths, α_{em} and α_s, in the matrix element calculation (renormalization scale, μ_R). Some choices are prepared in `ICOUP`(See Table A.2). If `ICOUP` $= 5$, the constant scale given by `GRCQ` in GeV is used in calculation. Users have to set an appropriate value of `GRCQ` in unit of GeV. `IFACT` determines Q^2 for PDF (factorization scale, μ_R). The definition is the same as `ICOUP`. The same choice as `ICOUP` is taken if `IFACT` is not explicitly given. If `IFACT` $= 5$, the constant scale given by `GRCFAQ` in GeV is used in calculation. Those energy scales also allow to change by hand. The subroutine `GRCUSRSETQ` provides 4-momenta of all partons in the ME calculation. Users can define their own energy scale in this subroutine by setting `ICOUP = 6` and/or `IFACT = 6`. The instruction is described in the later section. An example is also attached to the file `grcpar.F`.

`INPFL` is the number of flavors considered in PDF. For instance, the "`INPFL` $=$ 5" means "$u(\bar{u})$", "$d(\bar{d})$", "$s(\bar{s})$", "$c(\bar{c})$" and "$b(\bar{b})$"-quarks as well as *gluons* are considered inside PDF.

`IGRCGEF` is a flag switch (on/off:1/0) to include photon interference in Z/γ^* propagator.

Constant parameters :

The particle masses and decay widths and the other fundamental constants are set in the subroutine `SETMAS`. We show those parameters in Table A.3 and A.4. Note that for the fundamental constant such as G_F or $\sin^2\theta_W$, some of them are set as a "dummy" because the argument `IGAUGE` reevaluates those constants according to the choice of gauge invariant sets. Since GR@PPA does not give any constraint to these parameters, users have to care the mass and total decay width of the Higgs boson which have to be manually controlled.

A.3.3 User-defined energy scales and kinematical cuts

The choice of "`ICOUP(IFACT)` = 6" allows users to take any kind of scales by editing the subroutine `GRCUSRSETQ`. One example can be seen in this subroutine. That example shows the way to take the average transverse mass of two top-quark invariant system as the renormalization/factorization scale in the top pair productions (`ISUB` = `170`). The 4-momenta of all final state particles are given as the variable, `PGRC(i,j) (i=1,4)`, where the argument "`j`" is "j"-th particle from initial state particle with (`i=1,4`) = (p_x,p_y,p_z,E). The 4-momenta of intermediated particles are also give as `PQGRC(i,j) (i=1,4)`, where the definition is quite same as `PGRC(i,j)`. The numbering of j-th particles depends on each subprocess. In the include file, `grchad.inc`, detailed definition of kinematics is described. In the case of top pair production, the numbering convention of the initial/final state particles ($p1$, $p2$, $p3$, $p4$, $p5$, $p6$, $p7$, and $p8$) and their resonance particles ($Q1$, $Q2$, $Q3$, and $Q4$) is

$$\begin{array}{llllll} q(g)\,(p1) \;+\; \bar{q}'(g)\,(p2) & \rightarrow & t\,(Q1) \quad + & \bar{t}\,(Q2) & \rightarrow & W^-\,(Q4) \;+\; \bar{b}\,(p6) \\ & & \hookrightarrow W^+\,(Q3) \;+\; b\,(p3) & & & \hookrightarrow f''\,(p7) \;+\; \bar{f'''}\,(p8) \\ & & \quad \hookrightarrow f\,(p4) \;+\; \bar{f}'\,(p5) & & . & \end{array} \tag{A.9}$$

The example in `GRCUSRSETQ` is the follows.

Parameters	Description
`GRCECM`	CM energy in GeV. (D. = 14000.)
`GRCPTCUT(I)`	p_T cut for i-th particle. (D. = 0. GeV)
`GRCETACUT(I)`	Absolute of pseudo-rapidity $\lvert\eta\rvert$ cut for i-th particle. (D. = 10.)
`GRCRCONCUT(I)`	Separation ($\Delta R(i, other)$ ($= \sqrt{\Delta\phi^2 + \Delta\eta^2}$)) cut between i-th particle and the other particles. (D. = 0.)
`GPTCUT`	p_T cut only for jets (j). (D. = 20. GeV)
`GETACUT`	Absolute of pseudo-rapidity $\lvert\eta\rvert$ cut only for jets (j). (D. = 3.)
`GRCONCUT`	Separation ($\Delta R(j, j)$ cut between two jets (j). (D. = 0.4)
`IGJFLV(I)`	Flag switch (on/off:1/0) for the selection of the jet flavor (I(=1,8):d/u/s/c/b/t/g). (D. = 1 (I=1,2,3,4,5,7))
`IGAUGE`	Gauge invariant sets for the electroweak interaction. (D. = 1) $= 1$: (G_F, M_W, M_Z) $= 2$: ($\alpha_{em}(M_Z)$, M_W, M_Z) $= 3$: (α_0, G_F, M_W) $= 4$: ($\sin^2\theta_W$, G_F, M_W)
`IGWMOD(I)`	Flag switch (on/off:1/0) for the selection of W decay products. (I(=1,20):(du)/(dc)/(su)/(sc)/(bu)/(bc)/ ($e\nu_e$)/($\mu\nu_\mu$)/($\tau\nu_\tau$)) (D. = 1 (I=17))
`GRCWBR(I)`	Branching ratio of W boson. The numbering follows PYTHIA convention.
`IGZMOD(I)`	Flag switch (on/off:1/0) for the selection of Z decay products. (I(=1,32):($d\bar{d}$)/($u\bar{u}$)/($s\bar{s}$)/ ($c\bar{c}$)/($b\bar{b}$)/(e^+e^-)/ ($\nu_e\bar{\nu_e}$)/($\mu^+\mu^-$)/ ($\nu_\mu\bar{\nu_\mu}$)/($\tau^+\tau^-$)/ ($\nu_\tau\bar{\nu_\tau}$) (D. = 1 (I=9))
`GRCZBR(I)`	Branching ratio of Z boson. The numbering follows PYTHIA convention.
`IWIDCOR`	Width correction. If `IWIDCOR` = 2, the total cross section is re-normalized by the input argument of `GRCWBR(I)`, `GRCZBR(I)`. This is useful to estimate the cross section keeping the realistic shape of the mass distribution from the experimental value. If `IWIDCOR` = 1, theoretical calculation is outputed. (D. = 1)
`GRCCKM(I,J)`	Squared of Cabibbo-Kobayashi-Maskawa Matrix (CKM) parameters: $\begin{pmatrix} \lvert V_{ud}\rvert^2 = 0.95113 & \lvert V_{us}\rvert^2 = 0.04884 & \lvert V_{ub}\rvert^2 = 0.00003 \\ \lvert V_{cd}\rvert^2 = 0.04884 & \lvert V_{cs}\rvert^2 = 0.94940 & \lvert V_{cb}\rvert^2 = 0.00176 \\ \lvert V_{td}\rvert^2 = 0.00003 & \lvert V_{ts}\rvert^2 = 0.00176 & \lvert V_{tb}\rvert^2 = 0.99821 \end{pmatrix} \quad (I = u, c, t, J = d, s, b)$
`IBSWRT`	Switch on the implementation of the integration. If users have already made the integration file, it should be `IBSWRT` = 1, which leads to save time consumption. If users have not made the integration file yet, it should be `IBSWRT` = 0. The integration file is automatically made if users do not specify the output file name. Of course, when users change some parameters, you must make this file again with `IBSWRT` = 0. (D. = 0)
`GRCFILE`	Output file name for the BASES integration.
`NCALL`	Number of mapping points for the BASES integration.
`IRSEED`	Seed of random number. (D. = 12345)
`ICOUP`	Switch parameter of the renormalization scale in GeV. $= 1$: $\sqrt{\hat{s}}$ in the hard interaction. $= 2$: average value of transverse squared mass ($< m_T^2 >$). $= 3$: total transverse squared mass ($\sum m_T^2$). $= 4$: maximum transverse squared mass ($max\ m_T^2$). $= 5$: constant value. (Set `GRCQ` in GeV order.) $= 6$: user defined scale. (Edit the subroutine `GRCUSRSETQ`.) If users want to choose the fixed energy scale, it should be set `ICOUP` = 5. Then, users must set "`GRCQ`" which is the finite value in GeV order.
`IFACT`	Switch parameter of the factorization scale. The definitions are quite same as `ICOUP`. If `IFACT` = 5, constant value must be given. (Set `GRCFAQ` in GeV.) If users don't intended special usage, it should be always set to "0", which is taken as the same as the renormalization scale (`ICOUP`). (D. = 0)
`GRCQ`	The value of renormalization scale in GeV order if `ICOUP` = 5.
`GRCFAQ`	The value of factorization scale in GeV order if `IFACT` = 5.
`INPFL`	Number of flavors considered in PDF. For example, the "`INPFL` = 5" means "$u(\bar{u})$", "$d(\bar{d})$", "$s(\bar{s})$", "$c(\bar{c})$" and "$b(\bar{b})$" quarks are considered inside PDF. (D. = 5)
`IGRCGEF`	Flag switch (on/off:1/0) to include photon interference in Z/γ^* propagator. (D. = 1)

Table A.2: Input parameters used in GR@PPA_All. All parameters are set in the subroutine `GRCPAR` or `UPINIT`.

	Mass	Width
$u(\bar{u})$	`AMUQ(1)` (D. = 2.0 (MeV))	`AGUQ(1)` (D. = 0.0)
$c(\bar{c})$	`AMUQ(2)` (D. = 1.5 (GeV))	`AGUQ(2)` (D. = 0.0)
$t(\bar{t})$	`AMUQ(3)` (D. = 175.0 (GeV))	`AGUQ(3)` (D.=1.58(GeV))
$d(\bar{d})$	`AMDQ(1)` (D. = 5.0 (MeV))	`AGDQ(1)` (D. = 0.0)
$s(\bar{s})$	`AMDQ(2)` (D. = 0.5 (GeV))	`AGDQ(2)` (D. = 0.0)
$b(\bar{b})$	`AMDQ(3)` (D. = 4.75 (GeV))	`AGDQ(3)` (D. = 0.0)
$e^{\pm}$	`AMLP(1)` (D. = 0.511 (MeV))	——
$\mu^{\pm}$	`AMLP(2)` (D. = 105.66 (MeV))	——
$\tau^{\pm}$	`AMLP(3)` (D. = 1.7771 (GeV))	——
$\nu(\bar{\nu})_{e,\mu,\tau}$	`AMNU(I)` (`I=1,3`) (D. = 0.0)	——
Z	`AMZ` (D. = 91.188 (GeV))	`AGZ` (D. = 2.446 (GeV))
$W^{\pm}$	`AMW` (D. = 80.419 (GeV))	`AGW` (D. = 2.048 (GeV))
H^0	`AMH` (D. = 120.0 (GeV))	`AGH` (D. = 6.537 MeV)

Table A.3: Masses and widths set in the subroutine `SETMAS`.

Parameters	Description
`ALPHA`	Electromagnetic coupling constant at $Q^2 = M_Z^2$: $\alpha_{em}(M_Z^2)$ (D. = 1.0d0/128.07)
`ALPHAS`	Strong coupling constant : α_s (D. = 0.130)
`ALPHA0`	Electromagnetic coupling constant at $Q^2 = 0$: $\alpha_{em}(0)$ (D. = 1.0d0/137.036)
`GF`	Fermi constant : G_F (D. = 1.16639 $\times$ 10^{-5} (GeV^{-2}))
`STH2`	Square of Weinberg angle : $\sin^2\theta_W$ (D. = 0.231)

Table A.4: Fundamental constant set in the subroutine `SETMAS`.

```
      SUBROUTINE GRCUSRSETQ
      IMPLICIT REAL*8(A-H,O-Z)
      INCLUDE 'grchad.inc'

C...User-defined energy scale.
C
C   The definition is effective only when "ICOUP" or "IFACT" = 6.
C   Users must set their-defined energy scale into "GRCQ" or "GRCFAQ".
C
C...4-momentum in the CM frame of the hard interaction
C   PGRC(i,j)  (i=1,4:(px,py,pz,E)): 4-momentum of j-th particle.
C   PQGRC(i,j) (i=1,4:(px,py,pz,E)): 4-momentum of j-th resonance particle.
C
C   This example shows the way to take the avarage transeverse mass of two
C   top-quark invariant system as the renormalization/factrization scale in
C   the top pair productions (ISUB=170).
C
C   For 6-body kinematics, 4-momenta of (anti-)top quark are given in
C   a parameters : PQGRC(i,1) and PQGRC(i,2).

C...Top pT^2.
      PT2 = PQGRC(1,1)**2 + PQGRC(2,1)**2

C...Top p.
      TP1 = SQRT(PQGRC(1,1)**2+PQGRC(2,1)**2+PQGRC(3,1)**2)
      TP2 = SQRT(PQGRC(1,2)**2+PQGRC(2,2)**2+PQGRC(3,2)**2)

C...Top and Anti-top mass.
      TMAS  = SQRT(PQGRC(4,1)**2-TP1**2)
      TAMAS = SQRT(PQGRC(4,2)**2-TP2**2)

C...Avarage transeverse mass.
```

```
      GRCQ = SQRT((2.0D0*PT2 + TMAS + TAMAS)/2.0D0)
      GRCFAQ = GRCQ

      RETURN
      END
```

Users can also define user original kinematical cuts by editing the subroutine `GRCUSRCUT`. The example is also described in this subroutine. This example shows the cuts of $p_T >$ 20 GeV and $\Delta R > 0.4$ only for light flavors from top decay, but not applied the b-quarks from top. Users should note that the return argument, `IUSRCUT` must be set in anytime with “0” or “1” if events pass their cuts or not.

```
      SUBROUTINE GRCUSRCUT(IUSRCUT)
      IMPLICIT REAL*8(A-H,O-Z)
      INCLUDE 'grchad.inc'
      INTEGER IUSRCUT

C...User-defined kinematical cuts.
C
C   The returned argument is "IUSRCUT".
C   If events pass their cuts, IUSRCUT = 0,
C         not pass             IUSRCUT = 1.
C
C   Users must carefully check the convergence of the BASES iteration
C   if they define a tight cut using this subroutine.
C
C...4-momentum in the CM frame of the hard interaction
C     PGRC(I,J)  (I=1,4:(px,py,pz,E))  : 4-momentum of J-th particle.
C
C...4-momentum in the Laboratory frame.
C     PLGRC(I,J) (I=1,4:(pT,phi,eta,E)): 4-momentum of J-th (J > 2)
C                                        final state particle.
C
C...Useful stuff.
C     FUNCTION GRCKFCD(I)   : Output particle identification code.
C     FUNCTION GRCIVMS(I,J) : Output inv. mass between "I" and "J" particles.
C     FUNCTION GRCDPHI(I,J) : Output Delta Phi between "I" and "J" particles.
C     FUNCTION GRCDRCN(I,J) : Output Delta R between "I" and "J" particles.
C
C   This example shows the cuts of pT > 20 GeV and DeltaR > 0.4 only for
C   light flavors from top decay, but not applied the b-quarks from top.

C...Initialzation.
      IUSRCUT = 0

C...Set cut vatiable.
      PTCUT = 20.0D0
      DRCUT = 0.4D0
```

```
C...Look for light flavor, then check the kinematics.
      DO I = 3,8
         ID = GRCKFCD(I)
         IF (IABS(ID).LT.5) THEN
            IF (PLGRC(1,I).LT.PTCUT) IUSRCUT = 1
            DO J = 3,8
               IF (J.NE.I.AND.GRCDRCN(I,J).LT.DRCUT) IUSRCUT = 1
            ENDDO
         ENDIF
      ENDDO

      RETURN
      END
```

A.4 Run the code

A.4.1 Distribution package

The distribution package is arranged for the use on Unix systems. However, since the structure is rather simple, we expect that the program can be compiled and executed on other platforms without serious difficulties. The package is composed of the following files and directories:

`Makefile`	: the `Makefile` for the setup, `Makefile.aix`, `Makefile.linux` and `Makefile.solaris` are example `Makefiles` for IBM-AIX, Linux and Solaris, respectively,
`README`	: a file describing how to set up the programs,
`VERSION-2.62`	: a note for this version,
`bbbb/`	: Amplitude calculation codes for $b\bar{b}b\bar{b}$ processes,
`wjets/`	: Amplitude calculation codes for W(2f) + jets processes,
`zjets/`	: Amplitude calculation codes for Z(2f) + jets processes,
`top/`	: Amplitude calculation codes for $t\bar{t}$ (6-body) processes,
`diboson/`	: Amplitude calculation codes for WW, ZZ, ZW (4-body) processes,
`basesv5.1/`	: BASES/SPRING (version 5.1) source codes,
`chanel/`	: CHANEL source codes,
`grckinem/`	: source codes of kinematics,

`diagrams/` : typical Feynman diagrams included in the ME calculations,

`example/` : source codes of example programs,

`inc/` : `INCLUDE` files,

`lib/` : the directory to store object libraries; initially empty.

Users have to edit the file `Makefile` to specify an appropriate compiler and associated compile options, as well as the paths to the GR@PPA_All directory, and PYTHIA/HERWIG and CERNLIB/LHAPDF libraries. Those parts to be edited can be found at the top of the `Makefile`. We prepared examples for IBM-AIX, Linux and Solaris systems. All library routines are compiled and combined to object libraries if users execute the command `make` from the GR@PPA_All top directory. The object libraries are then moved to the directory `lib/` if the command `make install` is executed. The `Makefile`s of example programs in `example/` are set up by executing `make example`.

A.4.2 Mode selection

In the default usage, GR@PPA uses LHAPDF (The Les Houches Accord Parton Density Function) [69] which was conceived at the Les Houches 2001 workshop [70] in the PDF working group to enable the usage of Parton Density Functions with uncertainties in a uniform manner. GR@PPA also has an option to select the external PDFs provided by PDF libs [82] and the built-in PDFs inside PYTHIA. Because the subroutine name and calling sequence differ with each PDFs, the mode selection to choose the PDF group is adopted when users compile the program at first. The modification is in the include file `inc/define.h`,

```
#define PYTHIAPDF  1
#define CERNPDFLIB 2
#define LHAPDF     3

#define PDFLIB CERNPDFLIB          .
```

This example shows to choose the CERN PDF library by changing "`PDFLIB`". The same procedure is applied whether to use the LHA common format for the recent showering

MCs[2], or to work in PYTHIA version 6.1. The mode selection is done by setting the argument in "PSMODEL",

```
#define LHACOMMON   0
#define PYTHIA61   61

#define PSMODEL LHACOMMON          .
```

This example shows to choose the LHA common format. The use of PYTHIA version 6.1 is a relic from our previous work, which is only available in $b\bar{b}b\bar{b}$ subprocesses (ISUB $=$ $160 \sim 164$). Note that users must re-make the libraries again (after `make clean`) when users change the mode selections in the include file `inc/define.h`.

A.4.3 Dependencies on PYTHIA

GR@PPA internally uses some utility programs provided by PYTHIA when the mode selection is selected for the use of PYTHIA and the built-in PDF of PYTHIA. The functions `PYALEM` and `PYALPS` are used to determine the Q^2-dependent coupling strengths of QED and QCD in the matrix element calculation. Since the Q^2 is given to these functions as an argument, their behaviors are basically controlled by GR@PPA routines. However, they require additional parameters to define the running. Users are required to set relevant control parameters, such as `PARU(112)` before calling any initialization routines. In addition, GR@PPA uses the PYTHIA function `PYPDFU` for referring to PYTHIA built-in PDFs. Users have to make a choice of PDF by setting the parameter `MSTP(51)`. On the other hand, when users select the external PDFs by CERN PDF library, the method to call PDF lib by setting `MSTP(52) = 2` in PYTHIA is not activated because this method requires a certain manipulation of the PYTHIA library. We support directly to use the PDF library rather than setting the PYTHIA convention to use external PDFs.

[2] The LHA format is available since PYTHIA version 6.2 and HERWIG version 6.5.

A.4.4 Sample programs

Sample programs is provided in each subdirectory in `example/alone`, `example/pythia`, `example/herwig`, respectively. Execution of the command `make example` from the GR@PPA_All top directory sets up the `Makefile` for those programs.

The main code is `alsample.F`, `pysample.F`, and `hwsample.F` for the stand-alone execution, PYTHIA, and HERWIG, respectively. The detailed instructions are described each manual of PYTHIA and HERWIG. In the stand-alone execution, the sample only call the subroutine `UPINIT` and `UPEVNT`. That will be useful to check the function of the program. All the subroutine for the parameter initialization and settings are contained in the file `upinit.F` and `grcpar.F`. Users can select the processes and set their parameters. After that, `GRCPYGEN` is called for the initialization of GR@PPA in `UPINIT`. BASES optimizes the integration grids and evaluates the total cross section here if `IBSWRT = 0`. Note that `GRCPYGEN` has to be called for every subprocess that users want to activate.

An event generation loop follows the initialization part. A call to `UPEVNT` automatically results in a call to `GRCPYGEN` through `UPEVNT`. The source code of `UPEVNT`, dedicated to the use in GR@PPA_All, is `upevnt.F`. The generated event information is returned in the common `HEPEUP`. Users can obtain the event information from this common elsewhere.

A.4.5 Output of results

In the output of the GR@PPA_All execution, GR@PPA_All makes two output files which are related to the BASES integration for one subprocess. One is a binary file which is used in the event generation cycle by SPRING. And the another one is a log file which contains details about the accuracy, the execution time, the cross section, and the weighted shapes of each random number of the BASES integration as well as the phase space cut criteria and parameter settings. From this file, users can get the information if the integration could get enough accuracy. Users should pay appropriate attention to the print out from BASES, especially when they apply tight cuts. Since each subprocess is composed of many coherent diagrams, it is not practicable to take all singularities into account in the "kinematics" definitions. Some very minor ones are ignored in GR@PPA_All. A combination of very tight cuts may enhance the relative contribution of ignored singularities. In such cases, it is likely to happen that, in the BASES iteration, the estimated total cross section jumps (increases) to a value unreasonably different from the previous estimation and, accordingly, the estimated error increases. Users should consider that they must be in such a trouble if they find a jump of, for instance, more than three times the previous error. The results are unreliable in the phase-space region defined by such cuts. The instructive integration accuracy is 0.5 % or better for every iteration. Users should change the parameter `NCALL` to a larger value if this accuracy is not achieved.

As the one example, we show the output results after execution using default settings in the GR@PPA_All distribution package. Users should find two files like `.result` and `.data` as already mentioned. The calculation results such as the total cross section are stored in `bases_w1j.result` which was indicated in `GRCFILE`. Users find the following lines in the file, `bases_w1j.result`:

```
               Convergency Behavior for the Grid Optimization Step
------------------------------------------------------------------------------
<- Result of  each iteration ->  <-     Cumulative Result     -> < CPU  time >
 IT Eff R_Neg   Estimate  Acc %  Estimate(+- Error )order  Acc % ( H: M: Sec )
------------------------------------------------------------------------------
  1  90  0.00  3.630E+03  0.982  3.629635(+-0.035628)E 03  0.982    0: 0: 2.67
  2  87  0.00  3.580E+03  0.482  3.589510(+-0.015526)E 03  0.433    0: 0: 5.30
  3  87  0.00  3.651E+03  0.450  3.618653(+-0.011287)E 03  0.312    0: 0: 7.94
  4  86  0.00  3.624E+03  0.450  3.620391(+-0.009282)E 03  0.256    0: 0:10.56
  5  86  0.00  3.604E+03  0.455  3.616479(+-0.008076)E 03  0.223    0: 0:13.21
------------------------------------------------------------------------------

               Convergency Behavior for the Integration Step
------------------------------------------------------------------------------
<- Result of  each iteration ->  <-     Cumulative Result     -> < CPU  time >
 IT Eff R_Neg   Estimate  Acc %  Estimate(+- Error )order  Acc % ( H: M: Sec )
------------------------------------------------------------------------------
  1  86  0.00  3.600E+03  0.456  3.599852(+-0.016409)E 03  0.456    0: 0:15.82
  2  86  0.00  3.599E+03  0.447  3.599374(+-0.011486)E 03  0.319    0: 0:18.45
  3  86  0.00  3.610E+03  0.447  3.602813(+-0.009355)E 03  0.260    0: 0:21.07
  4  86  0.00  3.621E+03  0.451  3.607206(+-0.008119)E 03  0.225    0: 0:23.70
  5  86  0.00  3.615E+03  0.447  3.608788(+-0.007256)E 03  0.201    0: 0:26.32
------------------------------------------------------------------------------
```

The integration was performed in two integration steps. Each step has five iterations.

The print format consists of the result of each iteration, the cumulative result and the computing time used. In the result of each iteration, the sampling efficiency (the percentage of the points inside of the kinematical boundary), the ratio of the number of the negative valued sampling points to the total number of sampling points in unit of percent, the estimate of integral of the iteration and the estimated accuracy in unit of percent are shown. In the cumulative result, the cumulative estimates of integral and error are listed up in addition to the accuracy in unit of percent. The estimate of integrals are in unit of pb. The computing time in this table is measured from the beginning of the grid optimization step till the end of the current iteration.

In the convergency behavior for the grid optimization step, it should be checked that the accuracy for each iteration does decrease iteration by iteration and converge to a stable value. If not the case, it is recommended to increase the number of sampling points, `INCALL` and try again. When the increment of number of sampling points does not help to improve the behavior, the current choice of the integration variables may not be suitable for the behavior of integrand. In that case, users should give up their parameter choices.

Finally, users find the summary table in the end of this file, after following some histograms which are used for a debug of kinematics.

```
Summary Table
=============
  CM Energy : 14000.000 GeV
  Beam type :       Proton-Proton Collision

  Process : [101] pp(bar)->W(2f) + 1 jet

  3 body final state
      Pt Cut of Particle 1 :  0.000 GeV
      Pt Cut of Particle 2 :  0.000 GeV
      Pt Cut of Particle 3 : 20.000 GeV
     Eta Cut of Particle 1 : 10.000
     Eta Cut of Particle 2 : 10.000
     Eta Cut of Particle 3 :  3.000
   Rcone Cut of Particle 1 :  0.000
   Rcone Cut of Particle 2 :  0.000
   Rcone Cut of Particle 3 :  0.400
        Missing Pt(Et) Cut :  0.000 GeV
```

```
 Option of Renormalization Scale :  5  80.419   GeV
 Option of   Factorization Scale :  0

   W Forced decay :    1
                  :   17    0   0.108

 Num. of Mapping for Integration :       491520
 Num. of Variable Rejection      :            0
 Num. of Cut Rejection           :        62763
 Num. of Kinematical Rejection   :          112
 Num. of Additional Rejection    :            0
===========================================
 Total Cross Section :    0.360879D+04   pb
```

These lines contains all user-selected parameters. The total cross section is also presented in the last line. The value actually is the same as the last cumulative result of the integration. Based on this table, these parameters are checked in order to escape inconsistency with the integration before running SPRING event generation cycle.

A.5 Physics processes in GR@PPA

A.5.1 4 b-quark processes (ISUB : 160 - 164)

The properties of the Higgs bosons depend on the underlying theory. The minimal supersymmetric extension of the Standard Model (MSSM) [83], which is considered to be a promising theory to solve difficulties in the SM, requires the existence of three neutral Higgs bosons. Among them, the CP-odd one and, in many cases, one of the two CP-even ones, have appreciably large couplings to the bottom quark over a wide parameter range (large $\tan\beta$ regions). The production associated with a bottom quark pair is a promising search channel in this case.

These Higgs bosons with large couplings to the bottom quark predominantly decay to a bottom quark pair. Therefore, this process can be experimentally tagged as four bottom-quark events, and the Higgs boson production can be identified by a resonant enhancement in the invariant mass spectrum of two bottom quarks. In spite of such a clear signature, a discovery in this way is not trivial because of the presence of the huge QCD background [84, 85]. Actually, a previous study for LHC [86] showed a discouraging result. Because only crude estimates based on many approximations have been available for the background processes, the prospects are still quite ambiguous.

The subprocesses, assigned in the subprocess number `160` to `164`, include all order-four tree level interactions within the Standard Model. Typical diagrams are shown in Appendix A.7. As listed in Table A.1, each subprocess is divided according to the difference in the order of the couplings. In the SM Higgs production processes, we ignore the Yukawa couplings of light (u, d, s and c) quarks. The subprocess (`160`) including y_b^2 are composed of diagrams including a bottom-quark pair production mediated by the Higgs boson. Namely, they are the "signal" processes. The subprocesses (`161`) classified as $\alpha_s^2\alpha_{em}^2$ includes irreducible "Z^0 background". Those (`162`) classified as α_s^4 is the non-resonant but most serious "QCD backgrounds". The contribution of the subprocesses (`163, 164`) classified as $\alpha_{em}^2 y_b^2$ and α_{em}^4 is expected to be small but included for the completeness. Note that the interference between those diagrams belonging to different subprocesses are ignored in GR@PPA.

The total cross sections estimated by GR@PPA with the default parameters listed in Table A.4 are also presented in Table A.1 for each subprocess. The Higgs boson mass and width are assumed to be 120 GeV/c^2 and 6.54 MeV, respectively. The b quark mass is set to 4.75 GeV. The renormalization and factorization scales (Q^2) are chosen to be identical to be a mass of Z boson. The results for `ISUB` = `160`, and `162` for both Tevatron Run-II and LHC conditions were found to be in good agreement with corresponding results from CompHEP [87].

The invariant mass distributions of two b quarks having the largest and the second largest transverse energy (E_t) with respect to the beam direction are shown in Figure A.2 for the Tevatron Run-II case. The results were obtained by turning off all simulations by PYTHIA. The peaks corresponding to the production of the Higgs boson and the Z boson are clearly seen. We can also see that the contribution of pure QCD subprocesses are quite large. Adequate phase-space cuts and/or an appreciable enhancement are necessary so that the Higgs boson signal become visible. It should be noted that, in the $\alpha_s^2\alpha_{em}^2$

subprocess (`ISUB` = `161`), off-resonance effects are clearly seen below the Z-boson peak. This shows that the electroweak effects (both Z and γ exchanges) are exactly evaluated in this program.

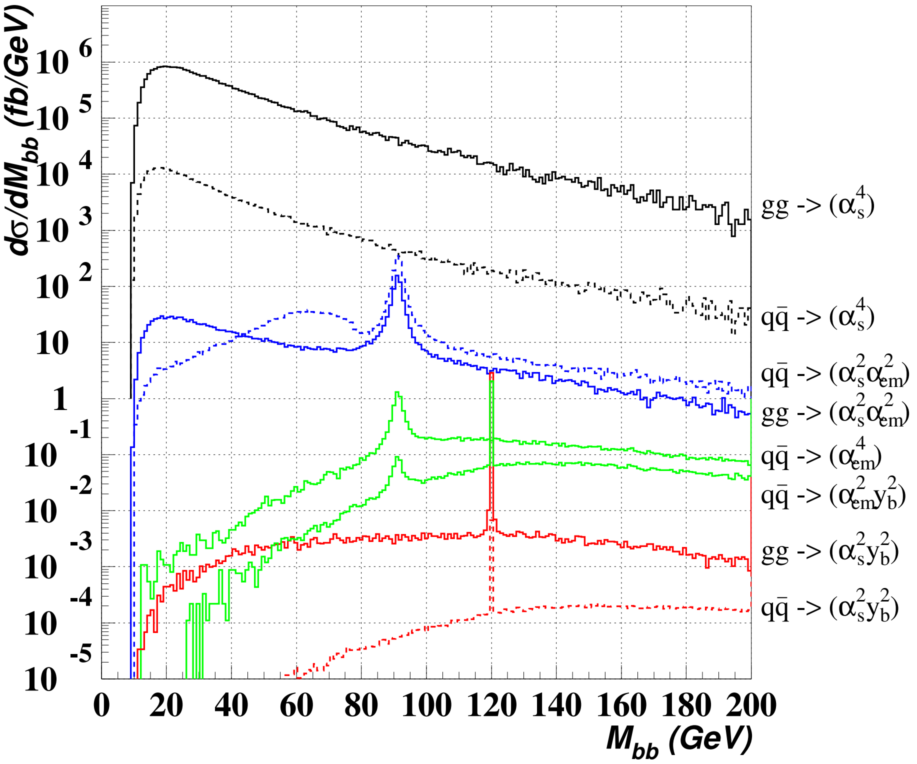


Figure A.2: The invariant mass distribution of leading two b quarks in the four-b events generated by GR@PPA. The distribution is shown separately for each subprocess.

A.5.2 W + jets process (ISUB : 100 - 103)

Due to an ease triggering of a high transverse momentum of the leptons from their bosons, the W + jets process has been a practical analysis process in hadron colliders not only for the dominant background of the most of precision measurements but also for the probing sample for new physics.

The subprocesses `ISUB` = $100 \sim 103$ cover all the possible processes with up to 3 jets associated with W boson, respectively. The decay process from W boson to fermions is also considered in the ME calculations although we simply use the notation W. All spin correlations and finite width effects are thus dealt exactly. The decay mode of the W boson is set by `IGWMOD`. Note that when the W boson is forced to decay to a quark pair, the interference between that quarks and jet partons is not account in the ME calculations. The jets are composed of u, d, c, s, b-quarks, and *gluon*. The flavor selection of jets is controlled by `IGJFLV`. The benchmark cross sections of these processes are listed in Table A.1.

A.5.3 Z + jets process (ISUB : 110 - 112)

The subprocesses `ISUB` = 110 $\sim$ 112 cover all the possible processes with up to 2 jets associated with Z boson, respectively, as well as the W + jets process. The notation Z means the Z/γ^* propagator from which the decay process to fermions is also considered in the ME calculations. All spin correlations and finite width effects are thus dealt exactly. The Z/γ^* interference effects are switched off by `IGRCGEF` option. In that case, only Z boson propagator is accounted in the ME calculations. The decay mode of the Z boson (and γ^*) is set by `IGZMOD`. Note that if the decay products are selected to a quark pair, the interference between that quarks and jet partons is not account in the ME calculations. The jets are composed of u, d, c, s, b-quarks, and *gluon*. The flavor selection of jets is controlled by `IGJFLV`. The benchmark cross sections of these processes are listed in Table A.1.

A.5.4 Top pair productions (ISUB : 170)

The subprocess `170` is top pair production processes. The exact treatment of 6-body ME calculations allows to be related to the spin correlations with final-state particles. We assume that there is no interfere between top quark lines in the Feynman diagrams due to a short distance of the top life time. Hence, the matrix elements which is considered in GR@PPA_All only includes top pair production diagrams.

As one example, the spin correlation between decayed particles from $W^\pm$ bosons can be specified using each angle with respect to the outgoing directions of top quarks. Due to the experimental difficulties, charged leptons are used in this kind of analysis [88]. Defining θ_+ as the angle between one of the charged leptons and the axis of quantization in the rest frame of its parent top quark, and similarly defining θ_- for the other charged lepton, the spin correlation can be seen as Figure A.3 for the $q\bar{q}$ initiative process, and Figure A.4 for the gg initiative process, respectively. Due to the different spin configuration of the top pair productions, there are different spin correlations according to the initial partons.

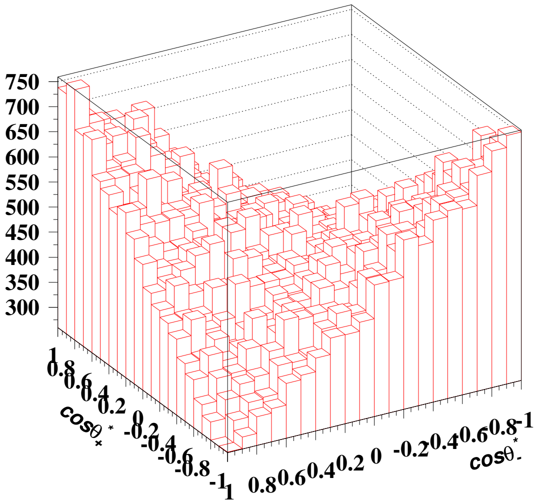


Figure A.3: Spin correlation for each charged lepton from $W^\pm$ bosons in the $q\bar{q}$ initiative process of top pair production, $q\bar{q} \to t\bar{t} \to b\ell^+\nu\ \bar{b}\ell^-\bar{\nu}$.

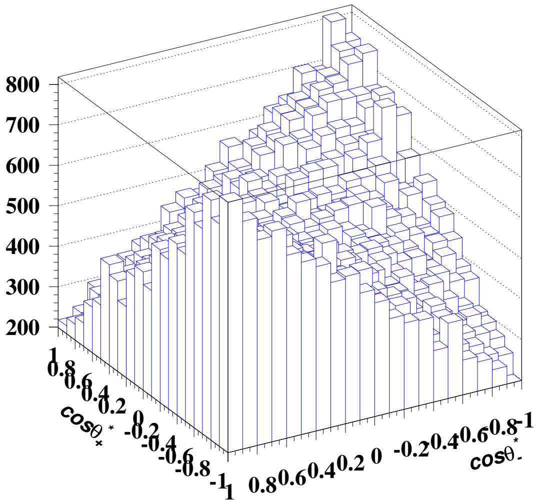


Figure A.4: Spin correlation for each charged lepton from $W^\pm$ bosons in the gg initiative process of top pair production, $q\bar{q} \to t\bar{t} \to b\ell^+\nu\ \bar{b}\ell^-\bar{\nu}$.

A.5.5 ZZ, WW, ZW processes (ISUB : 550 - 552)

A pair creation of the vector bosons is studied in the subprocesses `ISUB` $= 550 \sim 552$. The MEs fully cover with up to the decay fermions from their vector bosons. Thus, the Z/γ^* interference is dealt exactly. In addition, note that the same particle exchange processes interfere each other. For instance, it happens when one Z (and γ^*) decays to e^+e^- and then, another Z or W also decays to e^+e^- or $e^\pm$ and $\nu_e(\bar{\nu}_e)$, which can be seen in the ZZ process (`ISUB` = 550) and ZW process (`ISUB` = 552). Thus the spins of the decay fermions from both Z bosons correlates each other. This may also be useful to study the polarization processes.

A.6 Summary

We have developed an event generator GR@PPA_All for pp and $p\bar{p}$ collisions. The program is based on the newly developed GR@PPA framework, an extension of the GRACE automatic event-generator generation system to hadron collisions. This extension allows to incorporate the variation in the initial/final state into the GRACE system. The program includes an interface to the LHA common format to the showering event generators, together with PYTHIA 6.1 for old style in our previous study. This implementation is to add simulations of the initial- and final-state parton showers, hadronization and decays, as well as the additional underlying activities.

We should note that a large number of diagrams which are hard to manage manually are included in GR@PPA. It should be emphasized that GR@PPA can generate not only the signal events, but also various irreducible backgrounds. Especially, the most dangerous QCD background, which have been evaluated using crude approximations so far, can be evaluated on the basis of an exact tree-level calculation. Since the GR@PPA is not a process-specific event generator, our processes implemented by GR@PPA_All will increase with our physics interests.

Bibliography

[1] The LEP collaborations, Combination procedure for the precise determination of Z boson parameters from results of the LEP experiments, CERN-EP-2000-153.

[2] M. Kobayashi and T. Maskawa, *Prog. Theor. Phys.* **49** (1973) 652.

[3] M. H. Ahn *et al.*, *Phys. Rev. Lett.* **90** (2003) 041801.

[4] O.V. Tarasov, A.A. Vladimirov and A.Y. Zharkov, *Phys. Lett.* **B 93** (1980) 429.

[5] W.A. Bardeen *et al.*, *Phys. Rev.* **D 18** (1978) 3998.

[6] K. Hagiwara *et al.*, *Phys. Rev.* **D 66** (2002) 010001.

[7] V.N. Gribov and L.N. Lipatov, *Sov. J. Nucl. Phys.* **15** (1972) 438;
V.N. Gribov and L.N. Lipatov, *Sov. J. Nucl. Phys.* **15** (1972) 675;
G. Altarelli and G. Parisi, *Nucl. Phys.* **B 126** (1977) 298;
Yu.L. Dokshitzer, *Sov. Phys. JETP* **46** (1977) 641.

[8] R.K. Ellis, W.J. Stirling and B.R. Webber, QCD and Collider Physics, Cambridge University Press.

[9] P.M. Stevenson, *Phys. Lett.* **B 100** (1981) 61; *Phys. Rev.* **D 23** (1981) 2916.

[10] M.H. Seymour, *Comput. Phys. Commun.* **90** (1995) 95.

[11] T. Affolder *et al.*, *Phys. Rev.* **D 63** (2001) 072003.

[12] J. Mariiner, *et al.*, The Tevatron Run II Handbook, See www document, http://www-bd.fnal.gov/lug/runII_handbook/RunII_index.html.

[13] R. Blair, *et al.*, The CDF II Detector Technical Design Report, FERMILAB-Pub-96/390-E, November, 1996.

[14] The CDF II Collaboration, Proposal for Enhancement of the CDF II Detector: An Inner Silicon Layer and A Time Of Flight Detector, Frmilab-Proposal-909, presented to the Fermilab Director and PAC, October, 1998.

[15] M. Brozovic *et al.*, Jet Vertex Resolution in a Run 1 Z Event Sample, CDF/PUB/JET/PUBLIC/5698.

[16] M. Coca *et al.*, A First Look at Run 2 High p_T Electrons, CDF/MEMO/ELECTRON/CDFR/5803.

[17] A. Wyatt and B. Heinemann, Correction for Leakage Energy in the Central and Plug Calorimeters in Run II, CDF/ANAL/ELECTRON/CDFR/6167.

[18] L. Keeble, NEW JET CORRECTION FUNCTION QDJSCO VERSION 2.0, CDF/ANAL/JET/CDFR/1513.

[19] J-F. Arguin *et al.*, Generic Jet Energy Corrections for Run II data used for the Winter Conferences, CDF/ANAL/JET/CDFR/6280.

[20] S. Demers, *et al.*, Response of the Central Calorimeter to low momentum, isolated hadrons in Run2, CDF/ANAL/CALORIMETRY/CDFR/5874.

[21] H.S. Budd and W.K. Sakumoto, Determining Gains of Plug Towers using Plug Laser and Source Data, CDF/DOC/CALORIMETRY/CDFR/6433.

[22] C. Currat, *et al.*, Single Pion Response in the CDF Calorimeters and Tuning GFLASH, CDF/ANAL/CALORIMETRY/CDFR/5886.

[23] C. Currat, *et al.*, Comparison of Jet Momenta in Data and Monte Carlo after Plug time dependent Corrections, CDF/PHYS/JET/CDFR/6420.

[24] H.S. Kim, MIP Energy Scale Comparisons of CHA/WHA between Run I and Run II, CDF/DOC/CALORIMETRY/CDFR/6392.

[25] G. Latino, Jet EM Fraction Studies: RunI-RunII Comparison on Gamma-Jet Data, CDF/ANAL/JET/CDFR/6152.

[26] J.-F. Arguin, and B. Heinemann, Underlying Event Corrections in Run II, CDF/ANAL/JET/CDFR/6239.

[27] F.E. Paige and S.D. Protopopescu, ISAJET 5.30: A Monte Carlo event generator for pp and $\bar{p}p$ interactions, BNL report no. BNL-38034, 1986.

[28] R.D. Field and R.P. Feynman, *Nucl. Phys.* **B 136** (1978) 1.

[29] J-F Arguin *et al.*, Systematic Uncertainties Associated with Jet Corrections for Winter 2003 Conferences, CDF/PHYS/JET/CDFR/6419.

[30] L. Galtieri and J. Lys, How well do we understand jets in Run I? Study of the jet energy scale for raw jet et > 8 GeV, CDF/ANAL/TOP/CDFR/3983.

[31] F. Abe *et al.*, *Phys. Rev.* **D 50** (1994) 5550.

[32] S. Klimenko *et al.*, Averaging of the inelastic cross-section measured by the CDF and the E811 experiments, CDF/DOC/CDF/CDFR/6314.

[33] M. Coca *et al.*, W $\rightarrow e\nu$ Cross Section Analysis with Run II Data, CDF/DOC/ELECTROWEAK/CDFR/6300.

[34] Y.K. Kim *et al.*, Trigger Efficiencies for High P_T Electrons, CDF/DOC/ELECTRON/CDFR/6234.

[35] M. Coca and E. Halkiadakis, Central Electron Identification Efficiencies for Winter 2003 conferences, CDF/DOC/ELECTROWEAK/CDFR/6262.

[36] W. Badgett and P. Derwent, Event z Vertex Cut Efficiency as a Luminosity Correction for Run Ia, CDF/ANAL/ELECTROWEAK/CDFR/2703.

[37] W.K. Sakumoto and A. Hocker, Event $|Z_{vtx}| < 60$ cm Cut Efficiency for Run II, CDF/ANAL/ELECTROWEAK/CDFR/6331.

[38] F. Abe *et al.*, *Phys. Rev. Lett.* **64** (1990) 152.

[39] M. Cacciari *et al.*, The $t\bar{t}$ Cross-Section at 1.8 and 1.96 TeV: A study of the systematics due to parton densities and scale dependence, hep-ph/0303085.

[40] F. Abe *et al.*, *Phys. Rev. Lett.* **80** (2773) 1998.;
For the latest results, M. Kruse, Combination of the Winter 2003 Dilepton and Lepton Plus Jets top cross-sections, CDF/PHYS/TOP/GROUP/6509.

[41] F. Abe *et al.*, *Phys. Rev. Lett.* **82** (271) 1999.;
For the latest results, Y.K. Kim, M. Shochet, Averaging the Run II Top Mass Measurements, CDF/DOC/TOP/CDFR/6625.

[42] F. Abe *et al.*, *Phys. Rev. Lett.* **74** (2626) 1995.

[43] S. Catani *et al.*, *J. High Energy Phys.* **11** (2001) 063.

[44] John C. Collins, and Xiaomin Zu, *J. High Energy Phys.* **6** (2002) 018.

[45] Y. Kurihara *et al.*, *Nucl. Phys.* **B 654** (2003) 301.

[46] M.L. Mangano, Exploring theoretical systematics in the ME-to-shower MC merging for multijet processm in Proceedings of Matrix Element/Monte Carlo Tuning Workshop, Fermilab, Nov. 15, 2002.

[47] G. Marchesini and B. R. Webber, *Nucl. Phys.* **B 310** (1988) 461; G. Marchesini *et al.*, *Comput. Phys. Commun.* **67** (1992) 465.

[48] T. Sjöstrand, *Comput. Phys. Commun.* **82** (1994) 74.

[49] S. Catani *et al.*, *Nucl. Phys.* **B 406** (1993) 187.

[50] M.L. Mangano, private communication.

[51] Gerald C. Blazey *et al.*, Run II Jet Physics; CDF/PUB/JET/PUBLIC/5293.

[52] S. Tsuno *et al.*, LesHouchesModule: A CDF-interface module for Les Houches Accords; CDF/DOC/MONTECARLO/CDFR/6333.

[53] S.D. Ellis *et al.*, *Phys. Rev. Lett.* **69** (1992) 25.

[54] F.A. Berends *et al.*, *Nucl. Phys.* **B 357** (1991) 32.

[55] M.L. Mangano *et al.*, *J. High Energy Phys.* **07** (2003) 001.

[56] A. Pukhov *et al.*, CompHEP - a package for evaluation of Feynman diagrams and integration over multi-particle phase space. User's manual for version 33.; hep-ph/9908288.

[57] F. Maltoni and T. Stelzer, MadEvent: Automatic event generation with MadGraph; hep-ph/0208156.

[58] S. Tsuno *et al.*, *Comput. Phys. Commun.* **151** (2003) 216.

[59] J. Campbell and R.K. Ellis, *Phys. Rev.* **D 65** (2002) 113007.

[60] S. Frixione, and B.R. Webber, *J. High Energy Phys.* **06** (2002) 029.

[61] S. Tsuno, Jet Separation Study, CDF/PHYS/JET/CDFR/6442.

[62] J.-F. Arguin, and P.K. Sinervo, Revisiting the Top-Specific Jet Energy Corrections, CDF/PHYS/TOP/CDFR/6404.

[63] T. Ishikawa *et al.*, GRACE manual, KEK Report 92-19 (1993);
F. Yuasa et al., *Prog. Theor. Phys. Suppl.* **138** (2000) 18;
For the latest progress, see hep-ph/0007053.

[64] D.E. Groom *et al.*, *Eur. Phys. J.* **C 15** (2000) 1.

[65] U. Baur *et al.*, *Nucl. Phys.* **B 375** (1992) 3;
Y. Kurihara *et al.*, *Phys. Lett.* **B 349** (1995) 367.

[66] A. Aeppli *et al.*, *Nucl. Phys.* **B 428** (1994) 126.

[67] U. Baur and D. Zeppenfeld, *Phys. Rev. Lett.* **75** (1995) 1002;
W. Beenakker *et al.*, *Nucl. Phys.* **B 500** (1997) 255.

[68] W. Beenakker *et al.*, *Nucl. Phys.* **B 548** (1999) 3.

[69] See www document, http://vircol.fnal.gov/ .

[70] S. Alekhin *et al.*, Physics at TeV Colliders II Workshop, Les Houches, France, May 2001.

[71] J. Pumplin *et al.*, *J. High Energy Phys.* **07** (2002) 012.

[72] H. Bachacou *et al.*, Preliminary Studies of Heavy Flavor Contents in W+jets Sample, CDF/DOC/TOP/CDFR/6570.

[73] S.D. Ellis and D.E. Soper, *Phys. Rev.* **D 48** (1993) 3160.

[74] M. Cacciari and S. Catani, *Nucl. Phys.* **B 617** (2001) 253.

[75] J. Fujimoto *et al.*, *Comput. Phys. Commun.* **111** (1998) 185.

[76] S. Jadach *et al.*, *Comput. Phys. Commun.* **76** (1993) 361.

[77] K. Sato *et al.*, Proc. VII International Workshop on Advanced Computing and Analysis Techniques in Physics Research (ACAT 2000), P.C. Bhat and M. Kasemann, AIP Conference Proceedings **583** 214 (2001).

[78] T. Abe, *Comput. Phys. Commun.* **136** (2001) 126.

[79] S. Kawabata, *Comput. Phys. Commun.* **41** (1985) 466;
S. Kawabata, *Comput. Phys. Commun.* **88** (1995) 309.

[80] Proposed at Physics at TeV Colliders II Workshop, Les Houches, France, May 2001; See also, E. Boos *et al.*, Generic User Process Interface for Event Generators; hep-hp/0109068

[81] J. Fujimoto *et al.*, *Comput. Phys. Commun.* **100** (1997) 128.

[82] H. Plothow-Besch, *Comput. Phys. Commun.* **75** (1993) 396.

[83] For reviews, H.P. Nilles, *Phys. Rept.* **110** (1984) 1;
H.E. Haber and G.L. Kane, *Phys. Rept.* **117** (1985) 75.

[84] J. Dai, J.F. Gunion, and R. Vega, *Phys. Lett.* **B 345** (1995) 29;
J. Dai, J.F. Gunion, and R. Vega, *Phys. Lett.* **B 387** (1996) 801.

[85] T. Affolder *et al.*, *Phys. Rev. Lett.* **86** (2001) 4472.

[86] E. Richter-Wąs, and D. Froidevaux, *Z. Physik* **C 76** (1997) 665.

[87] V. Ilyin, private communication.

[88] B. Abbott *et al.*, *Phys. Rev. Lett.* **85** (2000) 256.