Measurement of the Transverse Momentum of Dielectron Pairs in Proton-Antiproton Collisions

by

Dylan P. Casey

Submitted in Partial Fulfillment

of the

Requirements for the Degree

Doctor of Philosophy

Supervised by

Professor Thomas Ferbel

Department of Physics and Astronomy

The College

Arts and Sciences

University of Rochester

Rochester, New York

1997

To Brenda and Simon.

Curriculum Vitae

The author was born in Lansing, Michigan on 20 June 1968. He attended Michigan State University from 1986 to 1991, and graduated with a Bachelor of Arts and a Bachelor of Science degree in 1991. He came to the University of Rochester in the fall of 1991 for graduate studies in Physics. He received a graduate student fellowship from the Department of Education from 1991 through 1994. He received a Messersmith Fellowship in 1994 and 1995. He pursued his research in Particle Physics under the direction of Professor Thomas Ferbel and received the Master of Arts degree in 1993.

Acknowledgments

My journey through graduate school involved the assistance, encouragement, and support of a great number of people. General thanks are extended to the Department of Education, the Department of Energy, and the University of Rochester, all of whom enabled me to pursue my graduate education. There are, also, a few individuals that require special thanks.

In the first couple of years, before the great migration to Fermilab, I made many friends at Rochester and fulfilled part of a long-time dream of playing in a band. Thanks to Dave Branning and Bob Henderson, in particular, who shared many hours of talk—physics and otherwise—with me. Dave, Charlie Freeman, and John Bloomer managed to turn the band, Mojo Heroes, into something, even cutting a CD, in the years after I left Rochester. Though my involvement in the band was early and fleeting, I will always cherish the memories of playing our original music.

Many thanks to my fellow Rochester grad students at D∅ over the years—Cathy Cretsinger, Ki Suk Hahn, Kathy Fatyga, Eunil Won, and Hui Zhu. I couldn’t have asked for better people to grill me at weekly meetings and gripe with whenever possible.

Along with Ki Suk, and postdocs Bob Hirosky and Stefan Gruenendahl, I enjoyed a weekly getaway known as the “Rochester Lunch” during most of my time at Fermilab. My own schedule often hinged on the weekly excursion for good conversation, decent eats and drawn-out coffee. All three are constant reminders that doing good physics doesn’t mean giving up good living.

Marc Paterno may take credit for plunging me (willingly and enthusiastically) along the path of Bayesian statistics. The work on unfolding is as much his as it is mine and that particular collaboration has been one of the most satisfying of my graduate career.

Many thanks to the members of the WZ group at D∅, particularly Ian Adam, Eric Flattum, Norman Graf, and Harry Melanson for their constant help, thoughtful evaluation, and friendly collaboration.

Gian DiLoreto consistently reminds me of one of the main reasons I got into physics in the first place—to find out how things work. We spent many hours stuck under our vehicles together trying to figure out how we were going to get to work the next day. Thanks for all the evenings spent drinking a beer on the back porch talking about physics, motorcycles, and our joint passion for Lyle Lovett and twangy guitar. I look forward to attending your defense at MSU and joining you on a “victory ride” atop our 750s.

Jim and Kelly Jaques were the gems of our life in Oak Park, and their friendship is inseparable from my thoughts of that time. Few people would deal well with friends knocking on the door early Sunday morning, 1 year-old son in tow, wondering if going to breakfast sounded good. Thanks, Jim, for your constant encouragement and friendship over these past couple of years. Whether during our daily commute, over a bottle of wine and great food, or our regular morning cup of coffee, I learned from you about physics, good living, and good friendships.

My advisor, Tom Ferbel, was a source of inspiration as a physicist. His guidance, support, and encouragement was a highlight of my education at Rochester.

My parents, Patrick and Rose Casey, have provided me unconditional love and support my whole life and I cannot thank them enough. I hope to raise my family as well as they have raised theirs. I have, hands down, the finest siblings on the face of the planet—Brian, Kim, Josh, and Dave—thanks for everything.

Without question, I would not have made it without the love and support of my wife Brenda and my son Simon. I am the luckiest man alive.

Abstract

We present a measurement of the transverse momentum distribution of dielectron pairs with invariant mass near the mass of the Z boson. The data were obtained using the D0 detector during the 1994-1995 run of the Tevatron Collider at Fermilab. The data used in the measurement corresponds to an integrated luminosity of 108.5 pb^{-1}. The measurement is compared to current phenomenology for vector boson production in proton-antiproton interactions, and the results are found to be consistent with expectation from Quantum Chromodynamics (QCD).

A Time to Measure

There is a time for taking measure as close
As caliper and gauge can touch, as bright
As light detectors catch the topmost
Quark in brief, deep-in-matter flight.

We can measure. We can chart events,
Track their trending presence, draw a line,
In seeming randomness see the sense
Of things, their steady being there through time.

Upon this firm, expectant knowing, all
Of life (almost) is built, piece by measured
Piece in swing of time, in ageless call
To shape the shapeless, from it form our treasured

Useful works -- for shelter, food (or war).
We measure but ourselves and want for more.

Patrick C. Casey
1997

Contents

List of Tables

List of Figures

Forward

The University of Rochester is one of 48 institutions participating in the D∅ experiment at Fermilab. More than 400 physicists and students contribute to the effort. The Rochester group has been active in the monitoring of liquid argon purity, data acquisition and reconstruction, and analysis of data in the areas of QCD, supersymmetry, and top-quark physics. The group has also participated extensively in the development and construction of the scintillating-fiber tracking chamber to be included in the current upgrade of the D∅ detector.

I participated in the Cosmic Ray Test Stand—a 3000 channel test of the principles of operation and construction of the scintillating-fiber tracker. I was responsible for the characterization of the visible light photon counters (VLPCs), which were developed for D∅, used in the test stand, and will serve as the read-out sensors for the upgraded central tracker. I participated extensively in the offline processing of the data, monitoring of the reconstruction processes, and writing the code for transferring data from the processing farm to the main computer-storage areas. I also co-wrote the QCDWZ ntuple-maker—a program that put the data into the standard "ntuple" form used in analysis at D∅, and was used extensively by others studying W and Z phenomena. My analysis project has been the extraction of the transverse momentum distribution of the Z boson.

Chapter 1
Introduction

This is a thesis in experimental elementary particle physics. Elementary particle physics is the study, measurement, and mathematical formulation of the result of physical interaction at short distances. A very successful model of particle phenomena, called the Standard Model, has been developed over the past 20 years. Being a thesis in experimental physics, most of this document will involve the measurement of a specific physical quantity. In particular, we will describe the measurement of the transverse momentum of electron pairs produced in proton-antiproton ($p\bar{p}$) interactions. The results of the measurement will be compared with the prediction of the Standard Model. Before getting into the details of the measurement, we will summarize the basic assertions of the Standard Model, the main ingredients and methods of an experiment in particle physics, and discuss aspects of the formalism for the process under study.

1.1 The Standard Model

The Standard Model is a collection of very similar theories that attempt to describe matter at distances down to $\leq 10^{-18}$ cm. It rests on two tremendous intellectual achievements of modern science—*special relativity* and *quantum mechanics*. Special relativity, developed virtually single-handedly by Albert Einstein[1], revised our understanding of space and time. Einstein showed that the perception of distance and time depends upon the relative motion of observers, and thereby purged physics of the concept of an absolute reference frame. Quantum mechanics was developed by a cadre of physicists in order to explain puzzling experimental results in atomic spectroscopy[2][3]. The fundamental assumption in quantum mechanics is that the world at very short distances is granular rather than continuous. For example, the observed energy levels of electrons in a helium atom are discrete—each energy level is quantitatively distinct with no observed energy values in between the first and second, the second and third, and so on. There are several consequences of a quantum mechanical description of the world. Most strikingly, interactions between objects are no longer deterministic, but probabilistic. The characteristics of all interactions, for example, the exact time when a particle will decay or where it will be observed following a collision, are known only within a quantified degree of uncertainty. Though debates still rage regarding the exact physical interpretation of the theory[4], the quantum mechanical prediction agrees very well with every measurement done to test it.

Extending quantum mechanics and joining it with special relativity, one obtains *relativistic quantum field theory*. Here, not only are the particles themselves quantized, but so are the interactions—each are considered a quantum of a field permeating space. For example, an electron is the discrete and indivisible piece of the electron matter field. The photon is the discrete and indivisible piece of the electromagnetic field, carrying the electromagnetic force from one charged object to another. The quantum field theories comprising the Standard Model are part of a special class of theories called *gauge theories*. In gauge theories, one obtains the structure of particle interactions by demanding that the mathematical relationship, called the *Lagrangian*, describing the energy of the particle be invariant under local phase transformations. If a quantity is invariant under some transformation, it is said to have a *sym-*

metry.[1] A local phase transformation is one in which the field of the particle is allowed to change by a space-time dependent phase. Demanding that the Lagrangian for the particle be symmetric under such a transformation requires the addition of terms to the Lagrangian that correspond to particle interactions. These new field is called a *gauge* field, and the quanta of the field are called gauge bosons. For example, demanding that the kinetic energy of an electron be invariant under a local phase transformation,[2] requires that one introduce the electromagnetic vector potential (A_μ) which mediates electromagnetic interactions between electrons. The quantum of the electromagnetic interaction is the photon.

In sum, the Standard Model has the following, rather amazing property. The presence of each of the forces is a consequence of an underlying symmetry within the theory. This is a deep and complex property of the Standard Model, and a detailed explication of the these structures is beyond the scope of this limited introduction. References [5] and [6] provide detailed discussions of gauge theories and the Standard Model. Reference, [7] has a general (non-mathematical) explanation of the place and structure of gauge theory in the Standard Model.

A difficulty arises when one uses quantum field theory to calculate a physical result. Exact calculations in the theory are generally not possible to perform, and one must develop methods to obtain approximate values of measurable quantities. By far the most successful and ubiquitious approximation method is *perturbation theory*. In perturbation theory, the interactions between particles are written as an infinite sum of terms in increasing powers of the interaction strength (also called the coupling constant). Each of the individual terms in the series is calculable. Initial attempts to develop predictions from quantum field theory using

1 A simple example of a symmetry is the rotational symmetry of an equilateral triangle under rotations of 60^o about the center of the triangle. After such a rotation, the triangle is indistinguishable from the non-rotated version. Symmetries lie at the heart of the Standard Model.

2 As one might guess, there are many different kinds of local phase transformations. The language for describing transformations in general is called *group theory*. In this language, the photon arises as a consequence of requiring the Lagrangian to be symmetric under local U(1) transformations.

the leading terms of the sum were very successful. However, the sub-leading terms (the next terms in the sum) were found to be infinite. This required some re-thinking of perturbation theory. After a great deal of effort, these infinities were brought under control through a process called *renormalization*[6]. Roughly put, in renormalizing a theory, one matches terms which are positively infinite with those that are negatively infinite—the left over part (the *residue*) is the finite result providing the desired prediction.

1.2 Particles and Forces

In the Standard Model, all matter is comprised of *quarks* and *leptons*. They are the fundamental, indivisible particles in the theory. There are six types of each, as shown in Table 1.1. Ordinary atomic matter is made up only of electrons orbiting a nucleus, which is composed of the up and down quarks. The other particles are made only at particle accelerators and in interactions of cosmic rays in the atmosphere. Every particle also has a corresponding anti-particle, which is identical in mass, but whose quantum numbers (charge, for instance) are reversed. Collectively, particles that are made up of quarks are called hadrons. Although leptons can observed individually, quarks have never been observed outside of a bound state of three valence quarks (called baryons) or quark-antiquark pairs (called mesons).[1,2]

The three interactions (forces) included in the Standard Model are the strong, weak, and electromagnetic (see Table 1.2).[3] Gravity is not included because there is at present no quantum field theory for gravity. However, gravity's exclusion is of little consequence in terms of

---

[1] There have been many searches for evidence of hadronic matter that is not of the standard baryon or meson variety. Most recently, an experiment at Brookhaven National Laboratory has claimed to have observed a state consistent with a $q\bar{q}g$ structure. Such so-called "C exotic" states are interesting because they have quantum numbers that are inconsistent with the particle being simple $q\bar{q}$ or three quark systems[9].

[2] It should be noted that the top quark does not form bound states, unlike all of the other quarks, because it decays too quickly.

[3] Formally, each of the forces is a consequence of requiring a specific internal symmetry of the free particle Lagrangian. Specifically, the Standard Model is a U(1)xSU(2)xSU(3) gauge theory. Each of the symmetries corresponds to one of the forces—U(1) for the electromagnetic, SU(2) for the weak, and SU(3) for the strong.

	Particle	Electric Charge	Mass (GeV/c^2)	Interactions
Quarks	*up (u)*	*+2/3*	*several MeV*	*EM,Weak, Strong*
	down (d)	*-!/3*	*several MeV*	*EM,Weak, Strong*
	charm (c)	*+2/3*	*1.0-1.3*	*EM,Weak, Strong*
	strange (s)	*-1/3*	*0.1-0.3*	*EM,Weak, Strong*
	top (t)	*+2/3*	*~175*	*EM,Weak, Strong*
	bottom (b)	*-1/3*	*4.1-4.5*	*EM,Weak, Strong*
Leptons	*electron (e)*	*-1*	$5.11x10^{-4}$	*EM, Weak*
	electron neutrino (ν_e)	*0*	$<5.1x10^{-9}$	*Weak*
	muon (μ)	*-1*	*0.1057*	*EM,Weak*
	muon neutrino (ν_μ)	*0*	$<2.7x10^{-4}$	*Weak*
	tau (τ)	*-1*	*1.771*	*EM,Weak*
	tau neutrino (ν_τ)	*0*	*<0.031*	*Weak*

Table 1.1 Compilation of the particles that are part of the Standard Model. The particle masses are taken from the Particle Data Book[10]. Note that the quark masses are ambiguous. This is due to the fact that all but the top quark are always bound in pairs and triplets (called *mesons* and *baryons*). For each of the above particles, there also exist anti-particles with opposite quantum numbers.

	Gauge Boson	Force	Mass (GeV/c^2)
Force Carriers	*Photon*	*EM*	*0*
	$W^{\pm}$,*Z*	*Weak*	*80.33, 91.187*
	gluons (8)	*Strong*	*0*

Table 1.2 Compilation of the force carriers in the Standard Model.

successfully describing sub-atomic interactions, because the force of gravitational attraction between two particles is much smaller (~10^{40} times smaller than the electromagnetic force) than any of the other forces affecting individual particles.

The strong force affects only quarks, and is responsible for holding the nucleus together despite the extraordinary repulsive forces present between the like-charged protons. The theory of the strong force is called *quantum chromodynamics* (QCD). In the theory, the strong force is ascribed to the exchange of *gluons*, the quanta of the *color* field (hence the name QCD) responsible for the strong force. An electron has electric charge, and its interaction with another charged object is mediated by a photon, while a quark has “color charge”, and its interaction with another “colored object” is mediated by a gluon. The word color is used metaphorically to describe the elements of the strong force. There are 3 kinds of “charges” in the strong force, called red, green and blue. Just as ordinary atomic matter is electrically neutral (the positive charge the nucleus and the negative charge of the orbiting electrons are balanced), the total color charge of a proton is zero because the color charges of the quarks combine to form a “colorless” object. The metaphor with visible color is evident here—red, green and blue, are so-called primary colors that form a colorless (white) combination. All hadrons within the Standard Model are colorless objects.

One of the salient features of QCD is that the interaction strength gets smaller as the energy scale of the interaction increases. The end result of this effect is that at high energies (equivalently, short distances), quarks and gluons within a hadron behave almost independently of one another. This defining quality of QCD is called *asymptotic freedom*. It results from the fact that the color field quanta, the gluons, can interact with each other.[1] Consequently, for high enough interaction energy, the techniques of perturbation theory should work because the expansion of the interaction in powers of the interaction strength can converge. Thus we speak of perturbative and non-perturbative regimes as regions of energy scale in which the interaction strength is small and in which perturbation theory can be used, and regions in which the interaction strength is large and perturbation theory is expected to fail. Asymptotic freedom also suggests why quarks and gluons do not exist in isolation. That is, because the coupling strength of two colored objects increases as the distance between them increases (α_S gets larger for larger distances), it requires an ever greater amount of energy to separate them further. In principle, no amount of energy can isolate one quark from another, effectively, keeping them *confined*. Though not proven rigorously to be a consequence of QCD, confinement is accepted as a general feature of the theory[7][8].

The electromagnetic force is the most familiar interaction. We rely on a detailed understanding of electromagnetism whenever we turn on a TV, drive a car, use a computer, or turn

[1] This is the same thing as saying the field is “Non-Abelian”. Technically, the generators of the field, the gluons, do not commute with each other[11].

on a light.The electromagnetic force is also the most well-understood of the three. Calculations using *quantum electrodynamics*, QED, the portion of the Standard Model applicable to electromagnetic interactions, have been verified to astounding accuracies. Possibly, the most well-known and dramatic result is the agreement between theory and the measured value of the anomalous magnetic moment of the electron to better than one part in one billion[12].

The weak force is rarely observed in everyday experience, but is important on the sub-atomic level. It has history as being one of the most puzzling interactions. The neutrino, which early on was associated with the weak force, was first proposed by Pauli in order to account for the observed electron energy spectrum in β decay[12]. Specifically, the observed energy of the electron was inconsistent with the conservation of energy, implying the existence of an unobserved particle in the interaction.[1] As the process was studied more closely, it became evident that more was amiss. The experiments showed two important things. First, that the interaction was *universal*, *i.e.* the interaction was of the same form and strength for a plethora of different nuclear as well as charged pion decays. This indicated that there was something unique about the force—in particular, it wasn't just the electromagnetic or strong force in disguise. Second, that the existence of both $O^{14} \rightarrow N^{14*} + e^+ + \nu$ and $He^6 \rightarrow Li^6 + e^- + \nu$ transitions, the first of which has no net change in nuclear angular momentum ($0^+ \rightarrow 0^+$) and the second of which does ($0^+ \rightarrow 1^+$), implied that there were two distinct possibilities for describing the interaction, assuming that parity were conserved[12]. (Parity is a space reflection symmetry.) If the underlying physics of a process does not change if one reflects all of the spacial coordinates, *i.e.*, let $\vec{x} \rightarrow -\vec{x}$, then one says that parity is conserved. However, it was found that parity was violated in weak interactions. Rather than merely clinging to the dearly held prejudice that parity conservation is a general quality of all interactions, T.D. Lee

1 The other alternative would require the violation of the principle of energy conservation. This alternative, flying in the face of one of the most dearly held assumptions in all of physics, was, in fact, proposed by Neils Bohr to account for the experimental results[12].

and C.N. Yang searched the literature, revealing a lack of evidence for or against parity conservation[13]. Soon, C.S. Wu, E. Ambler, *et al.*performed the experiment showing that, in fact, one of the great prejudices of physics was a fraud—parity was violated[12]. The effect, observed in Co^{60} decays, was quickly verified in $\pi \rightarrow \mu\nu \rightarrow e\nu\nu\nu$ decays[12]. While not commonly experienced, the study of the weak force was re-shaping both our understanding of fundamental phenomena and our view of the world.

Once it was clear that parity was violated in such interactions, all nuclear decays were brought together and described in the famous *V-A* form, referring to the vector and axial-vector components of the interaction[14][15]. However, there was one last problem. The production rate for neutrino interactions was observed to grow linearly with neutrino energy, yet the form of theory of weak interactions (called the *Fermi* theory) required an upper bound on the production rate. The observation and prediction suggested a unitarity violation at sufficiently high energy (~300 GeV in the center of mass). The problem could be alleviated by modifying the theory to include a massive vector boson, called the W, as the mediator of the weak force. However, the theory was not renormalizable—the infinities present in the non-leading-order terms of the perturbative expansion could not be purged. A new theory of weak interactions was needed. That resulting theory revealed not only a complete and consistent theory for weak interactions, but also unified the weak and electromagnetic interactions under a single, electroweak, force. This unification is the crowning achievement of the Standard Model.

Consistent with the earlier discussion of the underlying symmetries of the Standard Model, the weak and electromagnetic forces are united by demanding gauge invariance under a local phase transformation.[1] Four fields are postulated, W^+,W^-,W^o, and B^o. The charged bosons correspond to the charged carriers of the weak force known as the $W^\pm$ bosons. The neutral components of the field mix together yielding the photon and a new neutral particle,

[1] The invariance is under U(1)xSU(2) local phase transformations.

the Z. Yet, one is left with an important problem—neither the matter particles of the theory (the electrons and quarks) nor the force carrying particles mediating the interaction have mass. This is a problem at two levels. First, we know that some of the matter particles, the electron and the quarks, do have mass. Second, if the gauge bosons that transmit the weak force had no mass, then the weak force would have infinite range and we would be able to observe it readily at large distances. On the contrary, the range of the weak force appears to be only about 10^{-15} cm, and is seldom encountered outside the region of particle collisions or the decay of certain nuclei. The final piece of the electroweak theory that resolves both of these issues is called the *Higgs mechanism*. One postulates the existence of a new field (called the Higgs field) which has the special property that the lowest energy state (the ground state) of the field is not zero. Such a field is said to have a non-zero vacuum expectation value. In the presence of this field, the symmetries present in the Lagrangian are broken and the particles thereby obtain mass. This effect, in which the Lagrangian has a set of symmetries that are not shared by the states themselves, is called a *spontaneously broken symmetry*. In the simplest version of the Higgs mechanism, there are four degrees of freedom (each corresponding to an individual gauge boson) for describing the Higgs field. Three of them are entangled with the electroweak fields, and together they are re-interpreted as the massive W and Z bosons that carry the weak force. The photon remains massless. Yet another particle, called the Higgs boson, is predicted by the theory, corresponding to the remaining quantum of the Higgs field.

When the electroweak (EW) theory was first proposed in 1967[16][17], a massive W, though anticipated even in the earlier Fermi theory, had not yet been observed. Also, the EW theory predicted the presence of so-called *neutral current* interactions—weak force interactions that were mediated by the neutral gauge boson, the Z. Neutral currents were first observed in 1973[12]. The W and Z bosons themselves were found in 1983, after a massive effort by the UA1 and UA2 collaborations at CERN in Geneva, Switzerland[18][19][20][21].

The Higgs boson has not yet been observed, but the most recent searches have set a lower limit on its mass as 71 GeV[22].[1] It is generally supposed that it can not be heavier than 1 TeV [10].

There are many sources providing in-depth descriptions of elementary particle physics. References [8][12][23][24] have been useful for the author, and the source of most of the above summary.

1.3 Experiments in Particle Physics

As has been alluded to, the development of particle physics has been a continuing confrontation between the direct observation of phenomena and the mathematical formulation of the physics describing those observations. Prior to the development of quantum mechanics in the late 1920s and also prior to the development of the electroweak theory and QCD in the 1960s, experiment was "ahead" of theory—there were many observations for which there was little understanding of the underlying mechanisms.[2] More recently, theory has been "ahead" of experiment in that many experiments have been undertaken to search for specific particles predicted by the mathematical model. As mentioned above, the W and Z bosons were predicted by the electroweak theory, and observed in the early 1980s. The top quark was predicted once the bottom quark had been observed in 1977, but was not observed until 1995[26][27]. In fact, most of the experimental projects undertaken in particle physics today focus on measuring parameters within the Standard Model and looking for deviations or inconsistencies from the predictions (precision tests), or on searching for new particles predicted by various extensions of the Standard Model (super-symmetry[23], for instance).

---

1 This is the value for a scalar, neutral Higgs, which is the consequence of the simplest form of the Higgs mechanism. However, there are other methods for invoking the mechanism, and these yield different predictions for the Higgs bosons.

2 The advent of high-energy accelerators allowed the production of a veritable "zoo" of hadronic particles for which there was little systematic explanation, until Gell-Mann and others introduced the ideas of the SU(3) symmetry and quarks into the phenomenology[24][25].

All particle physics experiments have several things in common. In fact, the general structure of a particle physics experiment has hardly changed since Rutherford's group first scattered α-particles from gold foils, and showed that atoms have very massive, positively-charged nuclei at their centers. Here, we briefly define some of the important terms and concepts involved in scattering experiments that will be used later in the thesis.

Most particle physics experiments deal with colliding particles. Basically, one throws one set of particles at another and observes the results of the collision. Before the 1970s, this involved firing a beam of particles at a stationary target. Such experiments are called fixed-target experiments. With the advent of the Intersecting Storage Ring (ISR) at CERN, there have also been colliding beam experiments, in which beams of particles, usually electrons or protons, are made to collide into one another. The data to be reported in this dissertation were collected with the D∅ detector at Fermi National Accelerator Laboratory (Fermilab), which is a $p\bar{p}$ collider located near Chicago. The collider is part of a series of particle accelerators (see Figure 1.1). Each individual accelerator brings bunches of protons and antiprotons a little closer to their final momentum, much like shifting though the gears on a car. The final stage, called the Tevatron, consists of a ring of magnets ~4miles in circumference. Currently, the final energy of the particles when they are made to collide is 0.9 TeV, resulting in a center-of-momentum energy of 1.8 TeV. Magnets keep the bunches of protons and antiprotons in orbit as they travel around the ring inside an evacuated pipe. The protons and antiprotons travel through the ring in opposite directions and are made to collide with each other at two locations. The D∅ detector (pronounced "D-zero") is located at the "D∅" position of the ring. For technical reasons, one being in order to fix the time between collisions, the particles travel around the ring in compact bunches. There are about 10^{11} particles in each bunch. In our data, there was about a 70% probability for one interaction at each beam crossing. Details about the accelerator can be found in Ref [28]. 1

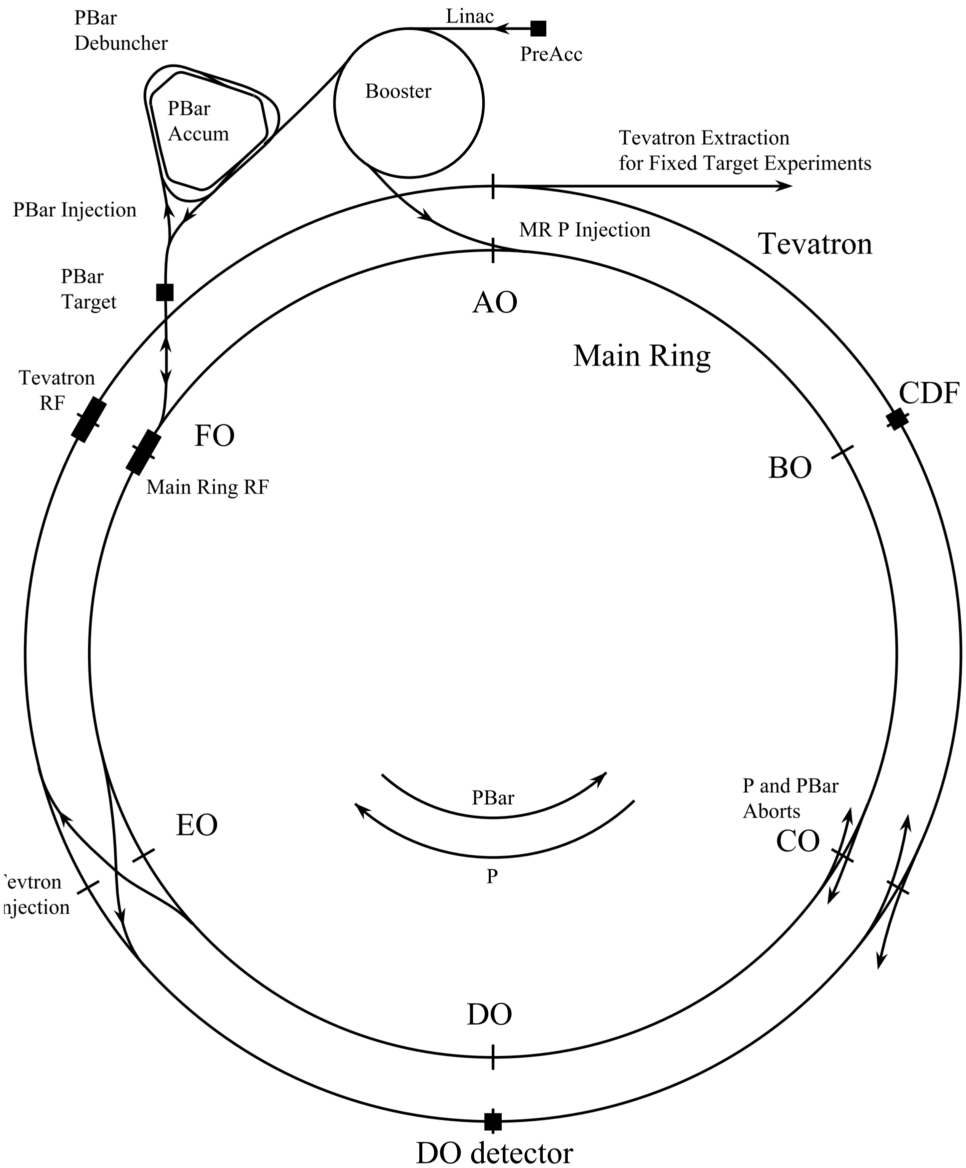

Figure 1.1 Sketch of the Fermilab accelerator complex.

Each collision between two particles is called an event. Each experiment collects and studies millions of events.[1] There are also many types of distinguishable particles that can be created in these collisions—electrons, photons, kaons, rhos, pions, etc. The task in the experiment is to identify the objects in each event, group similar types of events together and study their properties. In that way we can learn how the particles were made and learn something about the underlying physics.

We often observe large groups of particles that are closely associated with on another. At times, these particles are nearly indistinguishable because they overlap in space (for instance, their energy depositions in the calorimeter). Together, we call this stream of particles a *jet*. Jets are the result of the strong interaction. As mentioned in the discussion of the strong force, individual quarks have never been observed, only pairs and triplets appear to exist within hadrons. In the context of QCD, jets are interpreted as the end results of a quark or gluon fragmentation into colorless hadrons. As a quark pulls further and further away from its original confined system, *e.g.*, within the proton, enough energy is built up in the color field to create new quark/antiquark pairs. The produced quarks and gluons are rearranged until all are confined again into individual hadrons. Figure 1.2 shows a picture of a collision observed in the DØ detector in which one can see a single particle, in this case an electron, and a jet as primary features of the event. There is also an imbalance in the transverse momentum in the event, suggesting the production of a neutrino that carried away the transverse momentum, but went undetected because it interacts only via the weak force. This “missing” transverse momentum is usually referred to as missing E_T or $\not{E}_T$. The displayed event corresponds, in fact, to W+*jet* production, with $W \rightarrow e\nu$. When an event is comprised only of jets, we call that

[1] The presence of very large, computer-based data sets is a relatively new phenomenon in particle physics, constrained primarily by technology. A common feature of experiments 30 years ago involved recording information on film of the paths of the particles as they traveled through some active medium. The structure of each of the events was then discerned by carefully analyzing each of the thousands of photographs and stored on tape for subsequent analysis. Today, most of the information about an event can be encoded into a computer database, stored on disk, and analyzed via computer algorithms.

event a *multijet* event. Since the strong force is the dominant force at the collision energies of the Tevatron, multijet events are the most common types of events.

Among the quantities used to describe individual particles, the 4-momentum and the invariant mass are among the most important. The 4-momentum is a four-element vector that is comprised of the three components of the relativistic vector momentum and the total relativistic energy of the particle. The invariant mass is obtained from the square of the 4-momentum, $p^{\mu}p_{\mu} = m^2c^4 = E^2 - |\vec{p}|^2c^2$ .The invariant mass of a particle is a particularly useful parameter because it does not depend upon the relative motion between the particle and the observer (the experiment).

The transverse momentum is the part of the total momentum of a particle produced in a collision that is perpendicular to the direction of the incoming beams. It is determined by $p_T = \sqrt{p_x^2 + p_y^2}$ (see Figure 1.3 for an illustration). The transverse momentum is useful for reasons similar to that of the invariant mass—its value is invariant to boosts along the collision axis.

After measuring the kinematic parameters of produced particles and classifying events into different categories we can obtain the *cross section* (usually written as σ), which is related to the rate of particle production for any particular final state. Another important quantity is the *differential cross section*. This is the rate of particle production as a function of some kinematic parameter. In this thesis, we will present $d\sigma/dp_T$, the cross section for producing *ee* pairs as a function of the transverse momentum of the dielectron system.

1.4 Motivations for Measuring p_T^{ee}

Here, we outline some of the motivation for the measurement of the transverse momentum (p_T) distribution of *ee* pairs (p_T^{ee}) and the current understanding of the underlying physics.

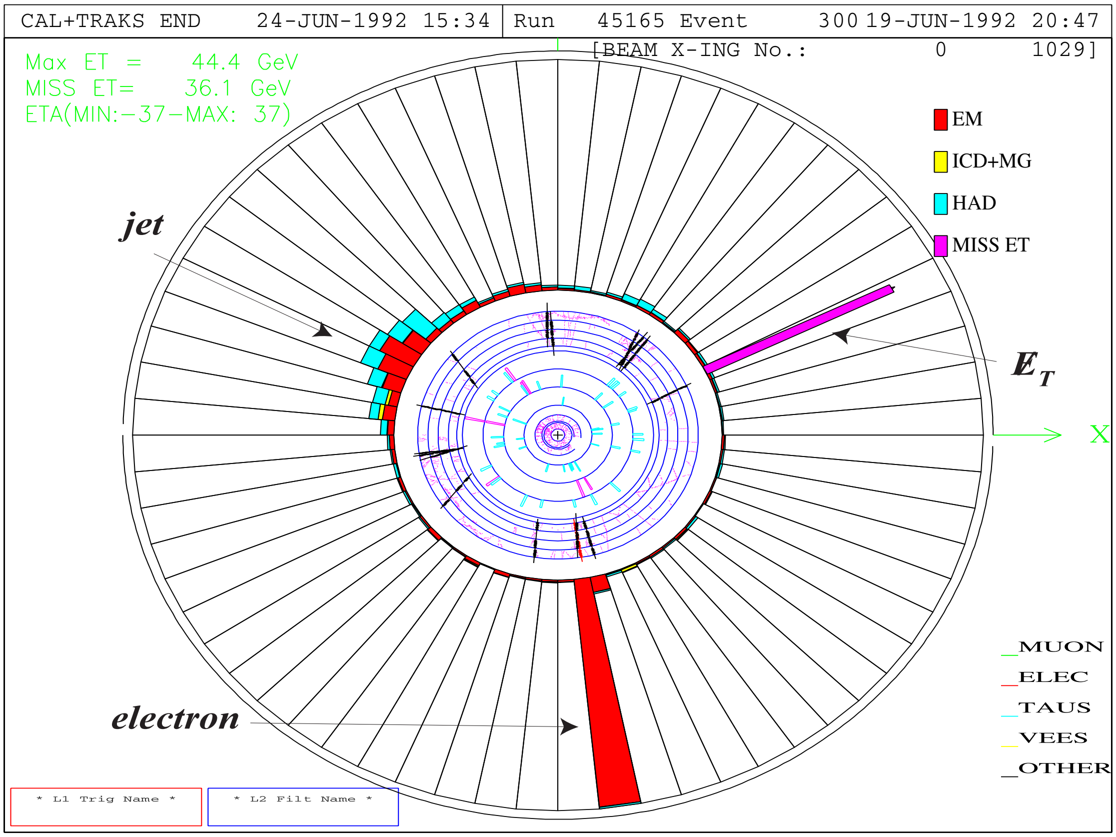

Figure 1.2 An event display for the transverse momenta in an electron + jet event, in this case the electron is from a W→eν decay. Note that the electron (bottom right) is much more narrow than the jet (top left). The spike pointing to the top right corresponds the missing transverse energy in the event, and is identified as the neutrino.

Since we will be focusing on the invariant mass range near the Z pole (M_{ee}~M_Z), the events are dominated by Z production—98% of *ee* pairs in the region 75<M_{ee}<105 GeV at D∅ are due to Z production[29]. The remaining few percent are almost exclusively due to the production of *ee* pairs through a mechanism referred to as *Drell-Yan* production[24]. The two processes cannot be distinguished, and interfere quantum mechanically with each other. In leading order, the *Z* originates from $q\bar{q} \to Z \to e^+e^-$, while Drell-Yan production proceeds throughout the coupling of $q\bar{q}$ to virtual photons: $q\bar{q} \to \gamma^* \to e^+e^-$. Other processes also contribute to inclusive dielectron production in $p\bar{p}$ collisions. One example is $t\bar{t}$ production, where the top quarks decay to W+b and both Ws decay into an electron and a neutrino. However, the

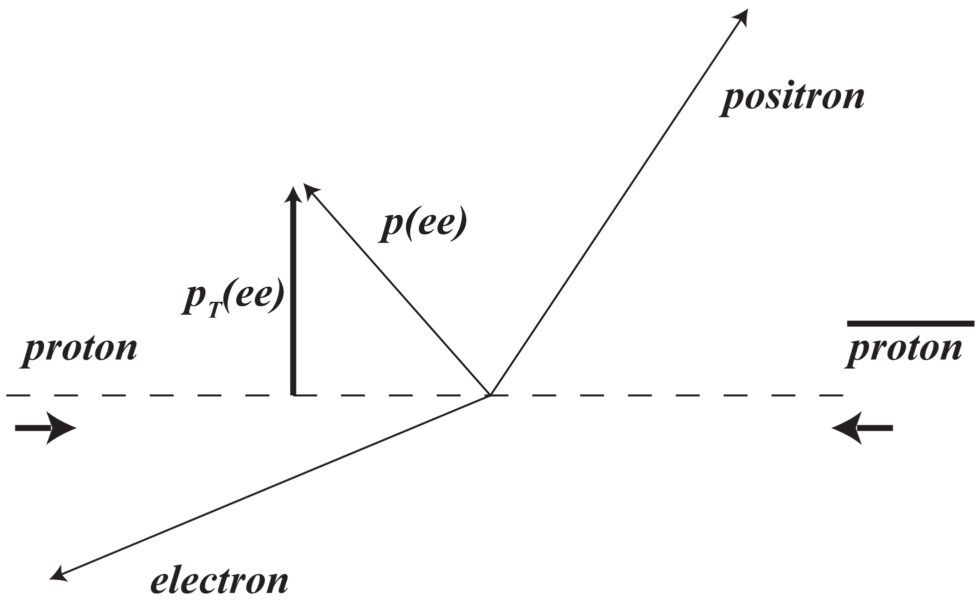

Figure 1.3 Illustration of how the transverse momentum of the di-electron pair is determined and how the transverse momentum of an object relates to the total momentum of that object.

production rate for these background contributions is very small and can, in general, be ignored. Contributions from this and other rare dielectron processes will be discussed later.

Studying the p_T^{ee} for distribution M_{ee}~M_Z amounts to studying the underlying production process, $q\bar{q} \rightarrow Z \rightarrow ee$. This process is sensitive to the dynamics of $q\bar{q}$ annihilation in hadron collisions. In particular, the measurement of $d\sigma/dp_T$ for *ee* pairs provides a sensitive test of the formalism used to describe vector boson production at low p_T. This formalism, called resummation, will be outlined briefly in the next section. This formalism is applicable to *W*, *Z* and γ^* mediated lepton/anti-lepton production. This begs the question of why not use the W for such investigations because there are 10 times more *W*s produced than *Z*s.[1] The *Z* is chosen over the W because the p_T measurement does not suffer from the same level of experimental imprecision as the measurement of the *W*. The *Z* properties from its decay into two electrons

[1] For Drell-Yan production with $M_{ee} < M_Z$, the E_T of the electrons is significantly lower than those in dielectron production with $M_{ee} \sim M_Z$. Since the ability to reject jets which mimic electrons decreases with decreasing electron E_T, these events are difficult to identify efficiently. For Drell-Yan production with $M_{ee} > M_Z$, there are so few events that measurement is still not competitive with the *Z* measurement.

can be measured much more precisely than those of the *W*, even in its cleanest channel of an electron and neutrino.[1] Figure 1.4 compares the resolution in p_T for the *W* and the *Z* at the DØ detector. Because of the better resolution, the *Z* measurement is expected to provide a much more satisfactory evaluation of the phenomenology, despite the fact that there are ten times fewer events.[2]

Besides being of interest on its own merits, the p_T^{ee} measurement, especially in the low p_T region, where the cross section is highest, will ultimately help yield a more precise *W* mass measurement. (This is discussed below.) The *W* mass, in combination with the top quark mass, provide direct constraints on the mass of the Higgs (see Figure 1.5). Because of large multijet backgrounds, the *W* can be reliably detected using only the decay $W \to e\nu$ rather than in the more copious $W \to jj$ channel. However, since the neutrino escapes detection, its momentum must be inferred via momentum conservation in the collision. This can provide a good measure of the transverse momentum of the neutrino, but the total momentum is impossible to reconstruct because of the unknown longitudinal momentum of the annihilating quarks that originate from the colliding hadrons. (The constituents that do not participate in the collision are emitted at small angles, and cannot be detected.)

There are three methods for measuring the mass of the *W* (M_W) in the $W \to e\nu$ channel. Each exploits the dependence of some measurable kinematic quantity on M_W: the transverse mass, $m_T = \sqrt{2E_T^e E_T^\nu(1 - \cos\theta_{e\nu})}$ [32]; the transverse momentum of the electron (p_T^e), and the transverse momentum of the neutrino (p_T^ν). Distributions in these quantities are closely related to M_W. A fit to the data can single out the best value of the mass. However, the pre-

1 In general, the Z is often used as the calibrating signal in an experiment. The mass is known to very high precision from e^+e^- collisions[10], and since all of the decay products (both electrons) can be well measured, many experimental quantities (for instance, the energy scale of the calorimeter) are set through a comparison of the observed and standard mass of the *Z*.

2 In fact, an analysis on the first 12 pb^{-1} of W data shows that the W p_T distribution is insensitive to significant changes in the theory.

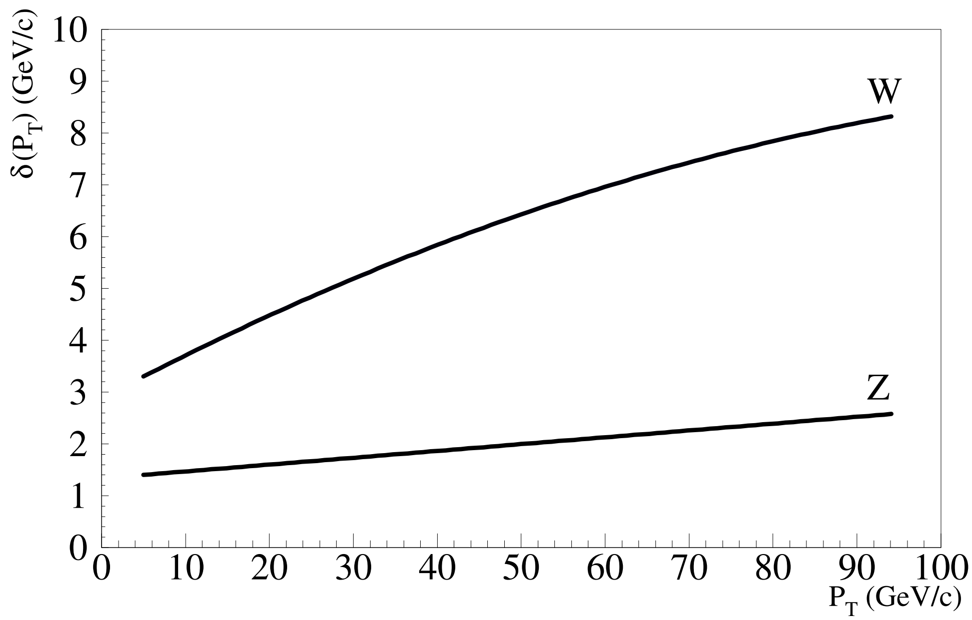

Figure 1.4 Comparison of the expected transverse-momentum resolution for *W* and *Z* bosons as a function of transverse momentum. The result was determined using a simulation of vector boson production, accounting for the D∅ detector resolution.

dicted distributions in these kinematic parameters also depend on the specific production properties of the *W* boson. Currently, the method with highest precision is the one based on the transverse mass, in part because it is fairly insensitive to the transverse momentum of the W. However, it is expected that the next attempt (in Run II at Fermilab) to measure M_W in $p\bar{p}$ collisions, the m_T method will be systematically limited by the missing E_T (E_T^{υ}) measurement. The expected large data set (~2×10^6 $W\rightarrow e\nu$ events) means that the statistical errors will be very small. As can be seen in Table 1.3, the uncertainty in detector resolution contributes more uncertainty to M_W when using the transverse mass method than for the p_T^e method. In the high statistics environment of the next run, we anticipate that the theoretical uncertainties in the p_T^e measurement will be easier to minimize than the systematic uncertainties in the transverse mass measurement. However, the transverse momentum of the *W* contributes directly to the electron p_T, and any uncertainty in the production model for the W feeds directly into an uncertainty in the *W* mass determination[33]. By using the closely-related *Z* production processes to constrain the *W* production model, this uncertainty can be minimized thereby yielding a more precise W mass.

Uncertainty	m_T method (MeV)	p_T^e method (MeV)
p_T model for the W	*5*	*28*
Total Theoretical	*34*	*62*
Detector	*60*	*46*
Total Systematic	*70*	*79*
Total Statistical	*95*	*108*
Total Uncertainty	*118*	*134*

Table 1.3 Comparison of some of the sources of uncertainty in measuring the W boson mass. The uncertainty in modeling p_T^W is included in the total theoretical uncertainty, which is part of the total systematic uncertainty. Note the large difference in the contribution of the p_T model to the transverse mass measurement versus the p_T^e measurement.

While the low p_T region is of genuine practical (M_W) and theoretical (QCD resummation) interest, the investigation of *Z* production at high p_T can be used to search for departures from perturbative QCD. An excess of events in this region could signal the presence of new physics.

1.5 *Z* Production Formalism

We briefly summarize the understanding of Z production within the context of QCD.

The production of the Z in $p\bar{p}$ collisions is expected to be described by the QCD portion of the Standard Model. In the parton model[24][25], *Z*s are produced in simple longitundinal collisions of $q\bar{q}$ constituents of the proton and antiproton, and consequently cannot have transverse momentum (< 1 GeV/c). Consequently, the fact that produced *Z* bosons have p_T must be attributed to QCD gluon radiation prior to quark-antiquark annihilation into the Z. Quarks radiate gluons as they travel through the color field within the proton (antiproton) just as accelerated electrons radiate photons. Gluon radiation increases in proportion to the time available for such annihilation, which goes as 1/Q[34]. (Q is the energy scale for the process, *e.g.*, Q ~ M_Z in *Z* production.) These radiated gluons carry away transverse momentum from

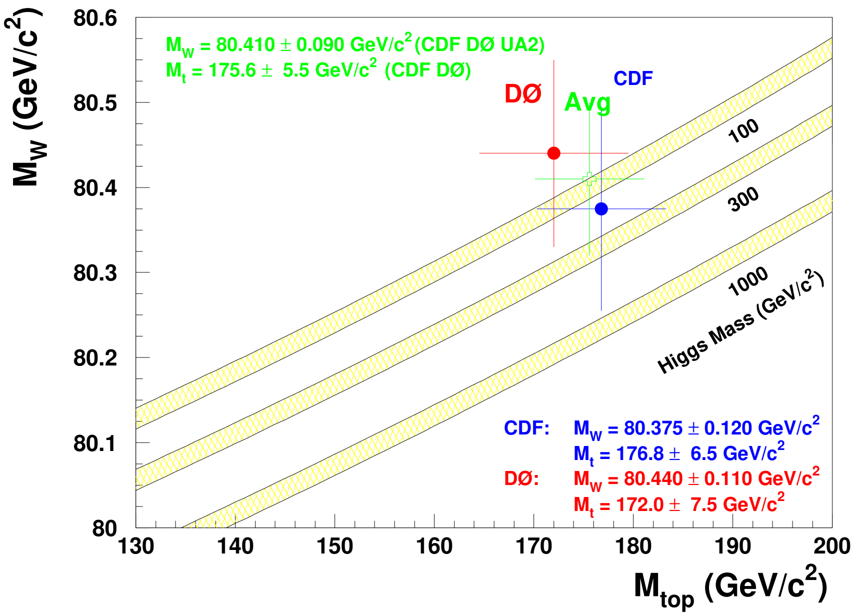

Figure 1.5 Comparison of the measured *W* mass and *t* quark mass values to the possible values of the Higgs mass[26][27].

the annihilating quarks. Momentum conservation requires that any observed p_T of the Z must originate in the p_T of the incoming quarks in the collision. Thus, one expects that the observed transverse momentum distribution of any dielectron system (of scale Q ~ M_{ee}) will broaden as a function of Q. This is, indeed, the effect observed. At M_{ee}~10 GeV, the typical p_T for Drell-Yan pairs is about 1 GeV/c[34]. For W-production (M = 80 GeV), the average p_T is about 5 GeV/c. Finally, for Z-production (M = 92 GeV), the average p_T is about 6 GeV/c.

The differential cross section is given by:

$$\frac{d^2\sigma_V}{dp_T dy} = \sum_{i,j}\int dx_1 dx_2 f(x_1) f(x_2) \frac{d^2\sigma(ij \to V)}{dp_T dy} \qquad 1.1$$

where y is the rapidity of V; x_1 and x_2 are the momentum fractions of the colliding partons; $f(x_1)$ and $f(x_2)$ are the parton distribution functions (PDFs) for the incoming partons, essentially the probability of the parton to have the momentum fraction x of the hadron; and

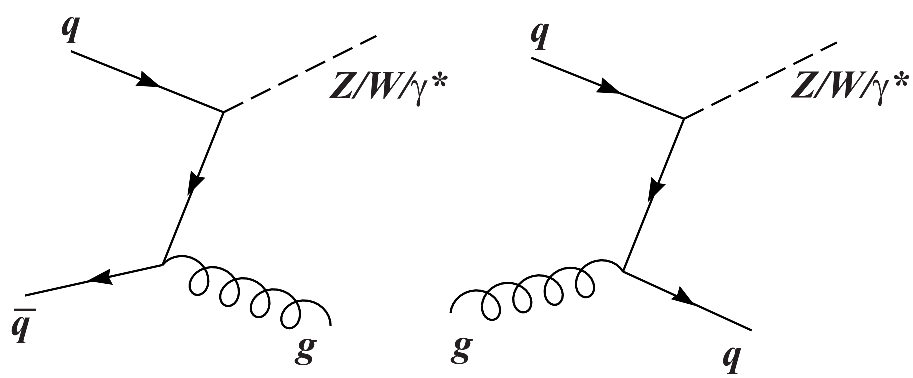

Figure 1.6 Leading order (α_s) Feynman diagrams for vector boson production with $p_T \neq 0$. (The diagram that goes as α_s^0 has $p_T = 0$.)

$\sigma(ij \to V)$ is the partonic cross section for production of the state *V*. The sum is over the contributing parton flavors (*i.e.*, up, down, etc.).

For moderate to high values of p_T of the *Z* ($p_T \sim M_Z$), the production cross section can be calculated using standard perturbative methods. The leading order diagrams where the vector boson $p_T \neq 0$ are shown in Figure 1.6. The result has the standard form

$$\frac{d^2\sigma}{dp_T^2 dy} \sim \alpha_W \alpha_S (a_1 + a_2 \alpha_S + a_3 \alpha_S^2 + \ldots) \qquad 1.2$$

where α_W is the weak interaction strength and α_s is the strong interaction strength.
We will call this the fixed-order perturbative result. The result up to $O(\alpha_S^2)$ has been calculated by Arnold and Reno[35].

This calculation is appropriate only for $p_T \sim M$, and diverges for $p_T \to 0$. The dominant contributions to the perturbative result are of the form

$$\frac{d^2\sigma}{dp_T^2 dy} \sim \frac{\alpha_W \alpha_S}{p_T^2} \ln\left(\frac{Q^2}{p_T^2}\right)\left[v_1 + v_2 \alpha_S \ln^2\left(\frac{Q^2}{p_T^2}\right) + v_3 \alpha_S^2 \ln^4\left(\frac{Q^2}{p_T^2}\right) + \ldots\right] \qquad 1.3$$

where Q^2 is the square of the mass of the dilepton system, in our case, M_Z^2. Note that there is the singularity as $p_T \to 0$ due to the $(1/p_T)^2$ term. Since $\alpha_s \approx 1 / \left(a + \ln\left(\frac{Q^2}{\Lambda^2}\right)\right)$ is fixed by the

Q^2 scale, the logarithmic terms also diverge as $p_T \to 0$. The perturbative expansion in α_s, which depends on the Q^2 scale, in not sufficient to regulate the behavior of the calculated cross section due to these logarithmic terms. Effectively, there are two scales in the problem; Q^2 and $p_T{}^2$. Because the cross section peaks at $p_T \ll M_Z$, the standard perturbative expansion given in Eq.1.3 will be insufficient to describe the regions of largest cross section (*i.e.*, small p_T).

The above problem is addressed using a technique called resummation. The salient features, along with the appropriate references, are given below. Since the problem is fundamentally mathematical, there is a great deal of formalism.

The v_i terms in Eq.1.3 are not independent. Considering all terms in the perturbative series that are at least as singular as $(1/p_T)^2$, one obtains a series of the following form, as $p_T \to 0$

$$\frac{d^2\sigma}{dp_T^2 dy} \sim \frac{\alpha_W}{p_T^2} \sum_{n=1}^{\infty} \sum_{m=0}^{2n-1} {}_n a_m \alpha_S^n \ln^m\left(\frac{Q^2}{p_T^2}\right) \qquad 1.4$$

which can be written, suppressing the coefficients ${}_n a_m$, as:

$$\frac{\alpha_W}{p_T^2}[\alpha_S(L+1) + \alpha_S^2(L^3 + L^2 + L + 1) + \alpha_S^3(L^5 + \ldots + 1) + \ldots] \qquad 1.5$$

where $L=\ln(Q^2/p_T{}^2)$. Note that each term of the expansion goes as $\alpha_S L^2$ rather than α_S. Equation 1.5 can be re-organized into a series of infinite sums whose leading behavior increases in powers of α_S, as follows:

$$\frac{d\sigma}{dp_T^2 dy} \sim \frac{\alpha_W}{p_T^2}[\alpha_S Z_1 + \alpha_S^2 Z_2 + \ldots]. \qquad 1.6$$

where each term involves an infinite sum of terms going as $\alpha_S{}^n L^m$,

$$\begin{aligned}
\alpha_S Z_1 &\sim \alpha_S(L+1) + \alpha_S^2(L^3 + L^2) + \alpha_S^3(L^5 + L^4) + \ldots \\
\alpha_S^2 Z_2 &\sim \alpha_S^2(L+1) + \alpha_S^3(L^3 + L^2) + \alpha_S^4(L^5 + L^4) + \ldots \\
\alpha_S^3 Z_3 &\sim \alpha_S^3(L+1) + \alpha_S^4(L^3 + L^2) + \ldots
\end{aligned} \qquad 1.7$$

Note that every higher term in the re-organization decreases by a factor of α_S rather than $\alpha_S L^2$. (We have continued to suppress the coefficients ${}_n a_m$ in Eq.1.7.) Each term in the new expansion depends on only a few calculable coefficients. This is the principle of the resummation technique — re-organizing the perturbative expansion to provide a result that depends only upon a few calculable coefficients and a series that is an effective expansion in α_S.

The above re-organization and resummation was first performed by Collins, Soper, and Sterman[34], and is often called the CSS formalism.

The resummed form for the differential cross section is expressed as a fourier transform to impact parameter space:

$$\frac{d\sigma}{dp_T^2} \sim \int_0^{\infty} d^2 b e^{i\vec{p}_T \cdot \vec{b}} W(b, Q) + Y . \qquad 1.8$$

W(b,Q) sums to all orders the terms that are at least as singular as $(1/p_T)^2$ (as $p_T \to 0$). The *Y* term is a correction based upon the fixed-order perturbative result. It is determined by computing the fixed-order perturbative result and subtracting the part of the series that is as singular as $(1/p_T)^2$ (called the "asymptotic piece" in Ref. [35]). Both quantities diverge as $(1/p_T)^2$ as $p_T \to 0$, however, the difference is well-defined and can be shown to correspond to the terms left out of the resummation procedure.

The first term has the form

$$W(b, Q) \sim e^{-S(b, Q)} . \qquad 1.9$$

The exponent, *S(b,Q)*, is called the Sudakov factor for historical reasons, and has the general renormalization group form,

$$S(b, Q) = \int_{(b_o/b)^2}^{Q^2} \frac{d\mu^2}{\mu^2}\left[\ln\left(\frac{Q^2}{\mu^2}\right)A(\alpha_s(\mu)) + B(\alpha_s(\mu))\right] \qquad 1.10$$

where

$$A(\alpha_s(\mu)) = \sum_j \alpha_S^j A^{(j)} \qquad B(\alpha_s(\mu)) = \sum_j \alpha_S^j B^{(j)}. \tag{1.11}$$

The $A^{(j)}$ and $B^{(j)}$ values are determined by comparing the above, order by order in α_s, to the fixed-order perturbative result. Considering the first-order approximation for $A(\alpha_s)$ and $B(\alpha_s)$, *i.e.*, keeping just $A^{(1)}$ and $B^{(1)}$, the exponentiation of the Sudakov factor corresponds to the first series in Eq.1.7. Knowing $A^{(1)}$ and $B^{(1)}$ yields the infinite series corresponding to the $O(\alpha_s)$ resummation. Similarly, by adding $A^{(2)}$ and $B^{(2)}$, one obtains the second series in Eq.1.7, bringing the approximation to $O(\alpha^2{}_s)$. The evaluation of the coefficients required for the sums was performed to $O(\alpha_S{}^2)$ by Davies, Weber, and Stirling. An enumeration of these coefficients and more detailed discussion of the technical details can be found in [36][37][38][39] [40].

There is an important caveat in the procedure. While the transformation from p_T-space to impact-parameter space facilitates the resummation calculation, the Fourier integral becomes undefined for $b \geq 1/\Lambda_{QCD}$, which corresponds to p_T becoming too small. Essentially, this marks the region where one crosses from a (reasonably) well-defined perturbative region into the non-perturbative region, in which perturbative techniques fail, and one must invoke a parameterization. This is accomplished by replacing $W(b,Q)$ with

$$W(b, Q) \to W(b_*, Q)e^{-S_{NP}(b, Q)} \tag{1.12}$$

where

$$b_* = \frac{b}{\sqrt{1 + \left(\frac{b}{b_{max}}\right)^2}} \tag{1.13}$$

smoothly cuts off the variation of $W(b, Q)$ for $b > b_{max}$, and $S_{NP}(b, Q)$ is a non-perturbative function that controls $W(b, Q)$ for $b \gg b_{max}$, *i.e.*, at low p_T.

$S_{NP}(b)$ must satisfy the following requirements. At small b, corresponding to high p_T, $b^* \to b + O(b^2)$. The non-perturbative effects must decrease with increasing p_T (decreasing b), hence $S_{NP}(b)\rightarrow 0$, as $b\rightarrow 0$. Also, infinite values of b cannot contribute to the cross section, and therefore we require $S_{NP}(b)\rightarrow\infty$, as $b\rightarrow\infty$. The value of b_{max} is arbitrary, but is generally taken to be ~ 0.5 GeV^{-1}.

The non-perturbative function $S_{NP}(b,Q)$ can be shown to have the following universal form[34]:

$$S_{NP}(b, Q) = h_1(b, x_A) + h_1(b, x_B) + h_2(b)\ln\left(\frac{Q}{2Q_o}\right) \tag{1.14}$$

where x_A and x_B are the momentum fractions of the annihilating quarks; Q_o is an arbitrary momentum scale; and $h_1(b,x_A)$, $h_1(b,x_B)$, and $h_2(b)$ are functions to be determined from experiment. The lack of dependence of $h_2(b)$ on the momentum fractions of the incoming quarks has led to speculation that it may contain some deeper relevance to the gluonic structure of the proton.[41] Discussion of the form of Eq.1.14 and how the formalism compares with the data will be undertaken in Chapter 4.

1.6 The Plan

The remaining chapters briefly discuss the D∅ detector, data selection criteria, and the corrections applied to the data in determining the final distributions. The final chapter presents the results and conclusions of our study. We present the measured differential cross section for p_T^{ee} and compare this to the current phenomenology taking into account the known detector resolution. Considerable effort was invested in a new method to unfold the measured distribution in order to extract a detector-independent measurement, but, unfortunately, significant technical difficulties were encountered in execution. The Appendix details the solution to the problem, and gives an example proving the viability of the method in a simpler situation. There we also discuss some of the difficulties encountered when we attempted to solve the problem at hand. This matter is being pursued at present, and we anticipate an approximate result for the unfolded p_T^{ee} distribution in the near future.

Finally, we present the upper limits for contributions to the cross section for p_T^{ee} above 50 GeV/c from sources beyond the Standard Model. We compare the observed cross section for dielectrons with $p_T > 50$ GeV/c to that expected from Z and Drell-Yan production and use the principle of *statistical inference* to obtain the probability distribution for contributions not included in the Standard Model.

Chapter 2
The Detector and the Data

To collect data in particle physics usually requires the construction of large complex detectors and sophisticated data acquisition systems. One must understand the detailed characteristics of the detectors and the software tools used to transform the individual electronic signals into the objects that we want to study. Hundreds of individuals have contributed to the construction and characterization of the DØ detector and copious documentation is available regarding all technical issues. For completeness, in this thesis, we will define only the most important terms and provide a brief description of some of the relevant parts of the detector. We will provide references below to more extensive information.

2.1 Definitions and Conventions

Two coordinate systems are used in DØ (Figure 2.1): The *detector frame*, where the origin of the coordinate system corresponds to the geometrical center of the detector, and the *physics frame*, where the origin of the coordinate system corresponds to the location of each $p\bar{p}$ interaction (or *event vertex*). In general, the position of the event vertex varies mainly along the direction of the beam; this is because the transverse dimensions of the beam are small, and there is very little transverse dispersion in the beams. Both coordinate frames are right-handed systems in which the positive *z*-axis coincides with the direction of the proton beam. Also, in both coordinate systems, the *y*-axis is chosen to point up (relative to ground). When applying fiducial or detector-motivated selection criteria to the data, the geometric variables are generally defined in the detector frame, whereas, physical quantities (the four-momentum of an electron, for instance) are calculated using the physics frame. Figure 2.2 shows the distribution in the *z*-coordinate (the vertex position along the beam direction) of the event relative to the center of the detector.

The angle θ is the standard coordinate that defines the polar angle relative to the beam direction, and ϕ defines the azimuth about the beam direction, with $\phi = 0$ given by the *x*-axis.

E_T (*transverse energy*) is the total energy of a particle or system of particles “transverse” to the direction of the beam. $E_T = E\sin\theta \approx p_T$ is an important quantity because it is essentially invariant under Lorentz boosts (space-time transformations) along the direction of the beam. Because of this feature, the E_T of an electron in the center-of-mass frame of the collision is nearly the same as the E_T we observe in our detector. We use E_T because it is the energy (*E)* that is measured in the calorimeter, and not the momentum (*p)*.

Τηε η (*pseudo-rapidity*) of a produced object is defined as $\eta = -\ln\left(\tan\frac{\theta}{2}\right)$. The η of a particle corresponds to the *rapidity* ($y = \frac{1}{2}\ln\left(\frac{E - p_z}{E + p_z}\right)$) when the mass of the particle is small relative to its p_T. For particles such as pions or kaons, one can usually set $\eta \sim y$ in high-p_T collisions. The rapidity is a useful variable because it is also invariant under Lorentz boosts along the beam direction.

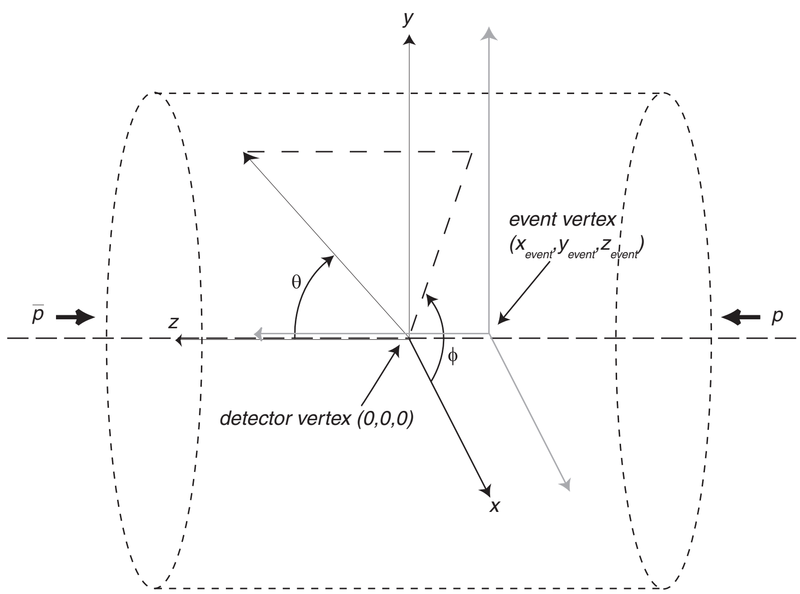

Figure 2.1 Illustration of the coordinate systems at D∅.

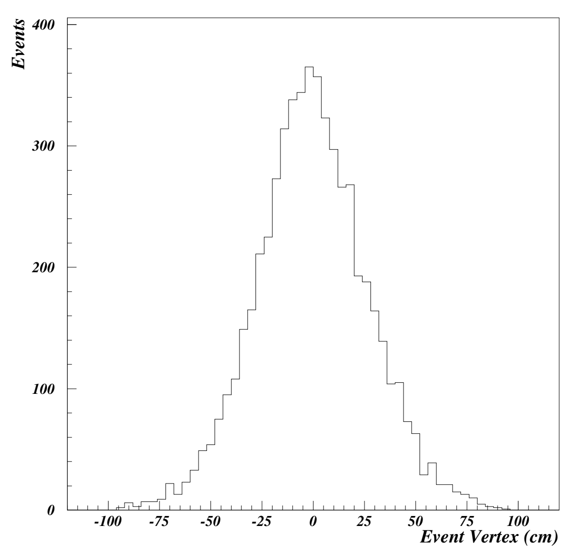

Figure 2.2 Distribution of the z-coordinate of the vertex (for Z events) relative to the center of the detector.

2.2 The Detector

D∅ is a general purpose particle detector designed to measure electrons, muons, and jets produced in $p\bar{p}$ collisions[42]. Before any detailed physics analysis can occur, we must determine the energies and trajectories of the individual particles produced in the interaction, and with that information, we can attempt to build a picture of the underlying processes. Each of the detector systems at D∅ provides a piece of the needed information. In the following, we will describe the components of D∅ that are primary to our analysis. Because the D∅ detector does not have a central magnetic field, it is impossible to distinguish the electric charge of the electron from that of the positron in Z decays. Thus, the “dielectron” events under study will be mainly e^+e^- events, but will have a small background from same-charge electrons.

2.2.1 Calorimetery

D∅ has three *calorimeters* that are used to measure the energy and positions of electrons and jets (see Figures 2.3 and 2.4).Each calorimeter is composed primarily of plates of uranium absorber, separated by gaps of liquid argon that function as the sensitive ionization medium. The *central calorimeter* covers the region $|\eta| \le 1.0$. The two *endcap calorimeters* cover the regions $1.4 < |\eta| < 4.2$. Throughout this thesis, electrons that were detected in the central calorimeter will be referred to as “CC” electrons. Similarly, electrons detected in one of the endcap calorimeters will be referred to as “EC” electrons. Since the different calorimeters correspond to different kinematic regions and have different detection efficiencies, each dielectron event falls into one of 3 categories, based upon the value of the pseudo-rapidity of the two decay electrons — CCCC, CCEC, and ECEC. Because of resolution considerations, only the CCCC and CCEC type events are used in the final analysis.

As can be seen in Figure 2.4, the calorimeters are constructed to be projective in pseudo-rapidity. Over most of the detector, the standard segmentation is 0.1x0.1 in η–ϕ. Each 0.2x0.2

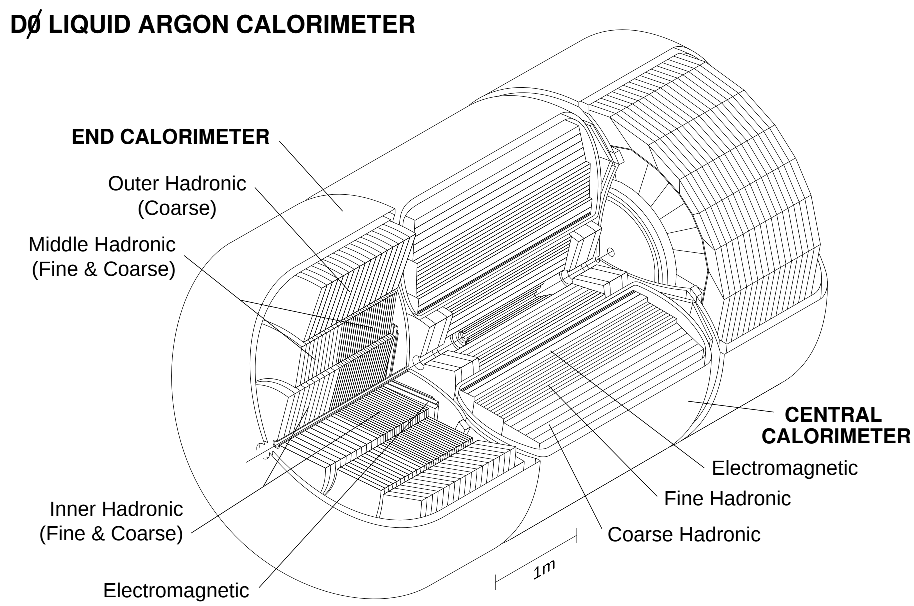

Figure 2.3 Diagram of the D∅ detector, highlighting the calorimeters.

η–ϕ section is called a *trigger tower*. The energy in such towers is used during data acquisition for selecting events that are considered worth writing to tape (see Section 2.3).

The calorimeters are constructed in layers[42]. The inner-most layer is the *electromagnetic* (EM) layer. Its third section is the most finely segmented (0.05x0.05 in η–ϕ). Outside of the EM layer is the *hadronic* layer, which has the standard η-ϕ segmentation, but is much thicker. The names of the calorimeter sections are motivated by their function. In general, highly energetic electrons travel through very little material (one radiation length is ~3.2 mm of uranium) before interacting. The electromagnetic shower of particles produced by an electron traversing material is highly collimated. Consequently, the EM section of the calorimeter is relatively thin and finely segmented. The third layer is the most finely segmented because the shower maximum for high p_T electrons (> 25 GeV) occurs within this layer. In contrast, jets are made up of many hadronic particles that interact primarily with the nuclei via the strong force (interaction lengths in uranium are about 10 cm). Consequently, jets are generally

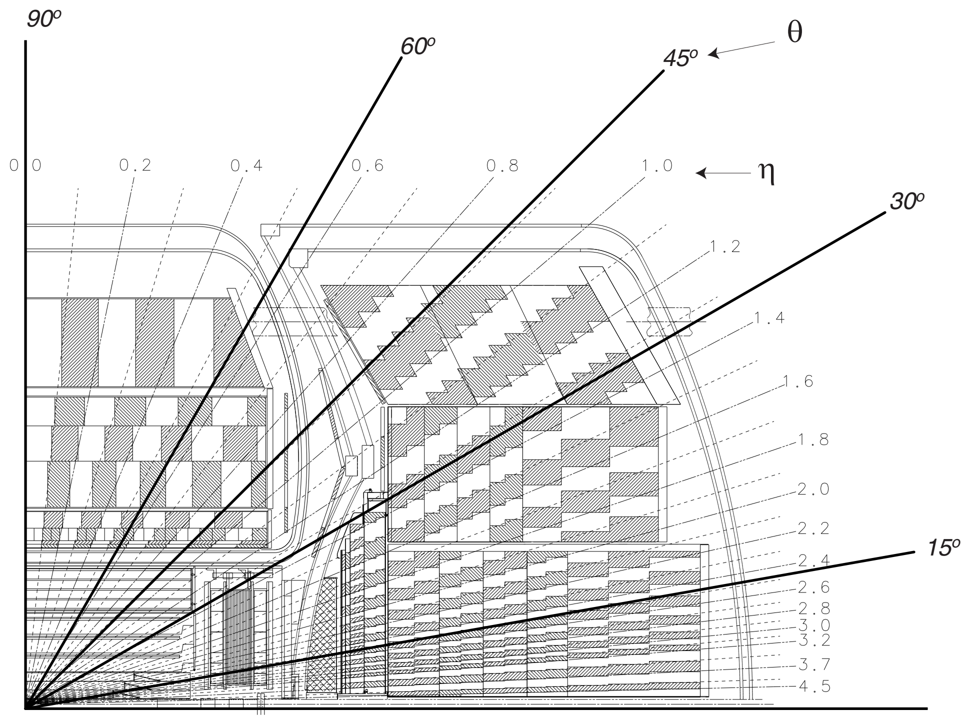

Figure 2.4 Illustration of a quarter section of the DØ detector. The alternately shaded and un-shaded regions denote individual calorimeter towers (trigger towers are $0.2x0.2$).

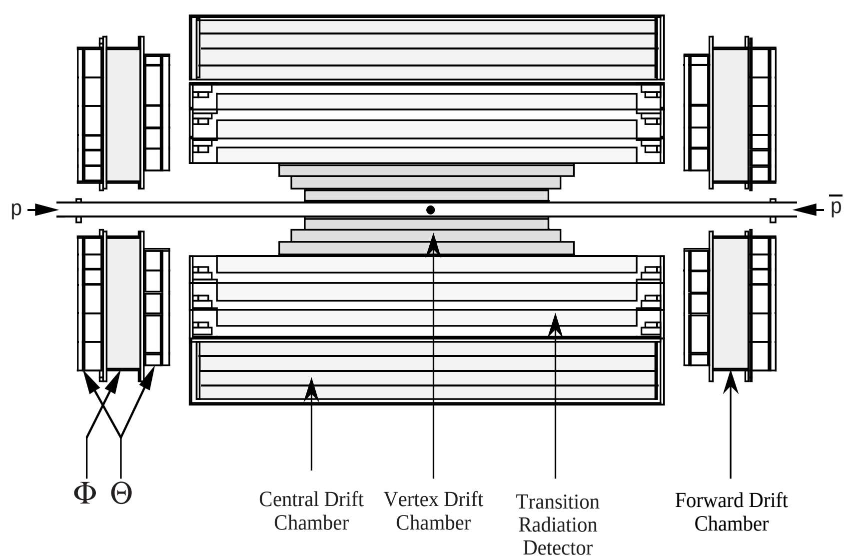

Figure 2.5 Diagram of the tracking systems of the DØ detector. Only the Central Drift Chamber (CDC) and the Forward Drift Chambers (FDCs) are used in this analysis.

more diffuse objects and have more extended energy deposition than electrons. The hadronic sections of the calorimeter are therefore relatively thick and can be more coarsely segmented.

2.2.2 Tracking

There are also three devices (see Fig. 2.5) used to determine the trajectories of electrons and charged particles in the jets — one *central drift chamber* (CDC) and two *forward drift chambers* (FDC). The charged particle trajectories are called *tracks*. Each complete track is built up from the individual hits in the chambers, where the charged particles ionized the gaseous medium. The general principles of design and operation of calorimeters and drift chambers are discussed in [43].

2.2.3 Other Systems

There are other detector systems at D∅ (*e.g.*, a large muon detection system), but they are not relevant for our analysis. The documentation for all of the detector elements at D∅ can found in Refs. [42] and [44].

2.3 Data Acquisition and Reconstruction

The steps from $p\bar{p}$ collisions to obtaining quantities suitable for analysis can be broken up into two parts. First, there is *data acquisition* — deciding which events to write to tape and which to ignore. Second, there is *data reconstruction* — transforming the individual electrical pulses from the thousands of detector channels read out in each event into a database containing electrons and jets. Below is a brief summary of each of these parts. For more detailed accounts see Ref. [42].

2.3.1 Trigger System

Acquiring data at D∅ proceeds in a hierarchical series of choices. The main problem is that it is not possible, nor desirable, to write out and examine every produced $p\bar{p}$ collision. The rate of interaction is approximately 500 kHz at a luminosity of 10^{31} events/cm^2-sec, but only a few events per second can be written to tape for later analysis. Also, most of the events are not particularly interesting. Even if we could write every interaction to tape, most of the data would be ignore in later analyses. The process of choosing the desired events is called “triggering”, and each level has its own constraints. At each stage there is a limited amount of information available for making any choice, and a limited amount of time in which to choose. At D∅ we have three primary levels in the triggering system — Level-0 (L0), Level-1 (L1), and Level-2 (L2).

The L0 trigger is a simple hardware trigger that requires some evidence for a “hard” interaction (large momentum transfer) between the proton and antiproton as they collide inside the detector.

The L1 trigger is a more complex hardware trigger that uses coarse information from the calorimeters in order to pick out interesting events. Rather than looking at the distribution of energy in every cell in the calorimeter, large groups of cells are considered all at once. As mentioned previously, the large groups of cells in $\eta\phi$ (0.2x0.2) are called *towers*. The L1 trigger used in this analysis required the presence of at least two towers with energy greater than 7 GeV in the EM sections of the calorimetry.

The L2 trigger is a software trigger implemented on a set of computers, where essentially all of the detector information is available for analysis. The final decisions made at this level use fast reconstruction algorithms to identify electron and jet showers. Only 2-3 μsec are available for the decision whether to keep any event. The L2 trigger for this analysis required one well-isolated electron-like object[1] with E_T > 20 GeV, and a second electron-like object with E_T > 16 GeV.

2.3.2 Reconstruction

Once events are selected by the L2 system, all of the detector information is written to tape. This low-level detector information, containing calorimeter cell energies, and pulses from drift chamber wires, has to be reconstructed to form the candidate electrons, jets, etc., for subsequent analysis. The name of the program written to reconstruct the data is RECO and it has several software releases. This analysis uses RECO versions 12.13 and later modifications[42]. Descriptions of algorithms used for reconstructing electron clusters and determining the tracks associated with those clusters are given in Ref. [29]. In general, the reconstruction of each event in this analysis is same as for nearly all other analyses at D∅ [44]. One significant difference involves the selection of the event vertex. The procedure we used is outlined below, and many of the W and Z analyses (including the W mass and inclusive cross section measurements) use this same procedure.

Because the colliding proton and antiproton bunches are dispersed over tens of centimeters along the beam direction, the collisions do not usually occur at the exact center of the detector, and the location of the interaction point must therefore be determined for each collision. The algorithm built into RECO determines the event vertex based upon the multiplicity of tracks found in the CDC. The point along the beamline to which the largest number of tracks extrapolate is assigned as the vertex. This vertex will be referred to as the RECO vertex. The RECO vertex was not used as the event vertex for the *Z* analysis. This is because electrons are very well-measured, and the two electrons from *Z* decay completely determine the vertex for a such events, we therefore use the electron tracks to define the event vertex.

---

[1] A "well-isolated electron object" has 90% of its energy in the EM section of the calorimetry and satisfies very loose clustering requirements, and an isolation requirement that there be very little hadronic energy deposition nearby.

The value of the z_{vertex} for each electron is determined by extrapolating the trajectory from the cluster centroid in the calorimeter, through the track center-of-gravity in the tracking chamber, to the beam axis as sketched in see Figure 2.6. Each accepted event is required to have $|z_{vertex}| < 100\,\text{cm}$, which constrains it to the central region of the detector. In cases where a track cannot be associated with the electron candidate, the RECO vertex is chosen as tentative z_{vertex} for that electron.

The final event vertex is chosen as the z_{vertex} of the central-most electron cluster that has a well-matching associated track. To be accepted, an event must have at least one good electron—*i.e.*, an EM cluster with a good track. Once the event vertex is determined, all electron kinematic quantities for the event are re-calculated using that final event vertex and the positions of the clusters in the calorimeter.

2.3.3 Electron Quality Criteria

Once events are reconstructed, they are examined more closely to confirm the presence of electrons. This proceeds using quantities that reflect the general quality of any electron candidate and that characterize the event. To accept an event, we require that each of the parameters listed below lie within some acceptable range. The criteria for the parameters were determined through analyses of simulated events, and data from the test beam and data from the collider[29]. The most relevant variables are:

f_{iso} (isolation fraction) quantifies how narrow and “isolated” are the electron energy depositions relative to other energy depositions in the calorimeter. The quantity is defined by:

$$f_{iso} = \frac{E_{tot}(R < 0.4) - E_{EM}(R < 0.2)}{E_{EM}(R < 0.2)} \qquad 2.1$$

where $E_{tot}(R < 0.4)$ is the total energy within a cone of radius 0.4 around the electron cluster, and $E_{EM}(R < 0.2)$ is the energy deposited in the EM section of the calorimetery within a cone of radius 0.2 around the electron cluster. The radius, R, is defined in η-ϕ as $\sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

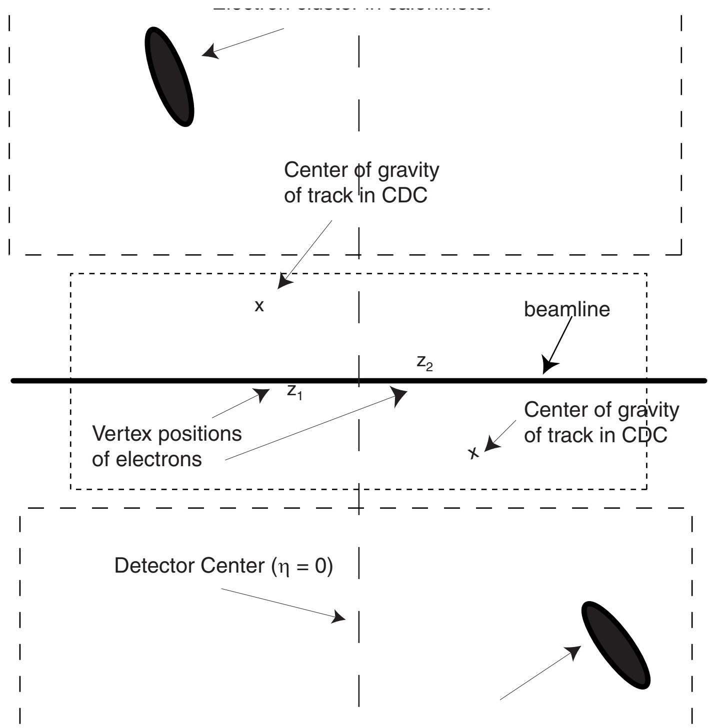

Figure 2.6 Diagram of vertexing method for Z events. In the diagram above, z_1 would be used as the event vertex since it is associated with the central-most electromagnetic cluster in the calorimeter.

High-p_T electrons produced in $p\bar{p}$ collisions, such as those produced in processes involving W or Z decays, are well-isolated in η-ϕ.

f_{em} (electromagnetic fraction) is the fraction of an electron's total energy that is deposited in the electromagnetic (EM) section of the calorimeter.

χ^2 (H-matrix χ^2) quantifies how well the pattern of energy deposition for a candidate electron agrees with that expected[30][31]. It is defined by

$$\chi^2 = \sum_{ii}^{41} (x_i - \bar{x}_i) H_{ij} (x_j - \bar{x}_j) \qquad 2.2$$

where

$$H = C^{-1} \qquad C_{ij} = \frac{1}{N} \sum_{k=1}^{N} (x_i^k - \bar{x}_i)(x_j^k - \bar{x}_j) . \qquad 2.3$$

The 41 observables in the correlation matrix (C) correspond to the fractional energy depositions in the EM1, EM2, and EM4 layers of the EM section of the calorimeter, the fractional energy in each cell of a 6x6 array surrounding the highest energy tower in the EM3 section of the calorimetery, the logarithm of the total cluster energy, and the z-coordinate of the interaction vertex. The elements of the correlation matrix were calculated using N simulated electrons (using GEANT) with a range in energy of 10 GeV to 150 GeV. The elements of the H-matrix are determined from the inverse of the correlation matrix.

σ_{trk} (track-match significance) quantifies how well a track reconstructed in the drift chambers matches the associated energy cluster in the calorimeter.

Figure 2.7 compares the distributions of the above quantities for electrons from $Z \to ee$ decay (see criteria described in the next section) and multijet events (in which one jet mimicked an electron shower). (See Chapter 3.2 for more details on multijet background events.)

2.4 Signal Selection

Acceptable dielectron candidates are obtained by selecting events that have at least two electron candidates located in well-understood regions of the detector. We summarize below the requirements imposed in our analysis.

2.4.1 Fiducial and Kinematic Criteria

To study electrons in well-understood regions of the detector, they must satisfy certain specific kinematic and geometric requirements. Only the two highest-E_T electron candidates in the event are used to reconstruct the Z candidate, and both are required to have $E_T > 25$ GeV. CC electrons are also required to be in the region of $|\eta_{det}|<1.1$. EC electrons are required to be in the region $1.5<|\eta_{det}|<2.5$. In order to avoid cracks between neighboring ϕ modules of the CC calorimeter, the ϕ of the electron is required to be at least 0.005 radians away from the position of the crack. Table 2.1 summarizes the kinematic and fiducial-volume criteria.

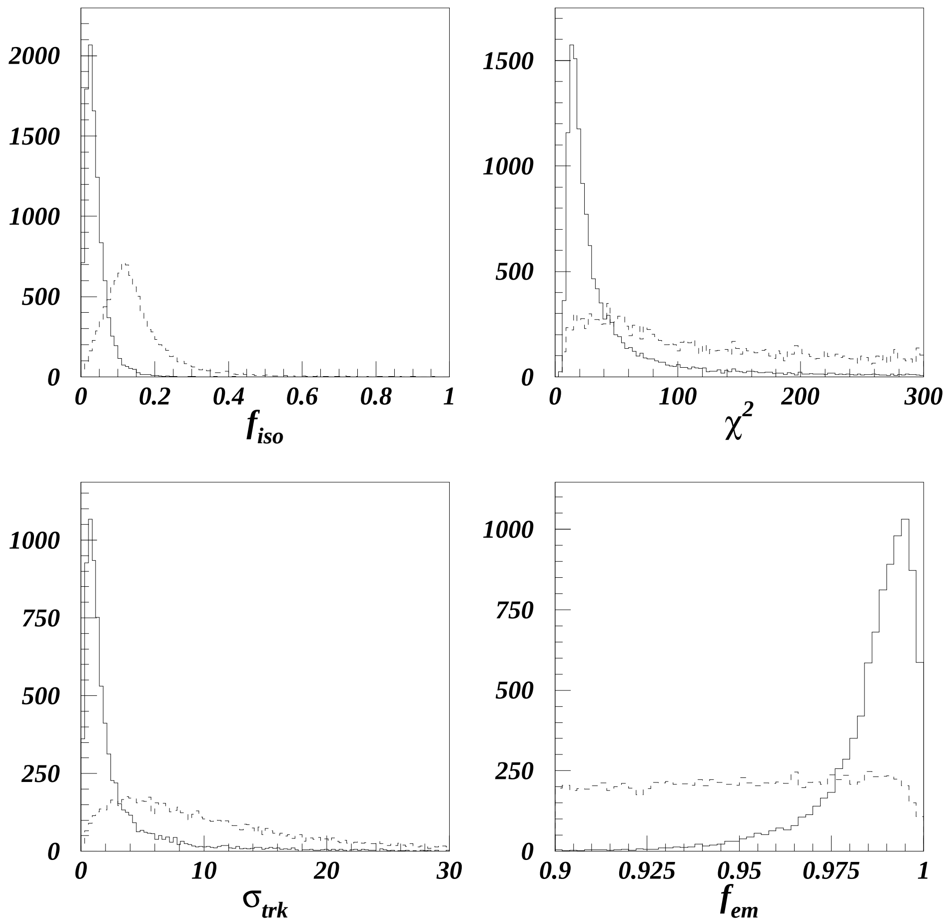

Figure 2.7 A comparison the distributions of the electron quality variables in $Z \rightarrow ee$ data (solid) and multijet events (dashed). The data were obtained by requiring all quality criteria to be satisfied, with the exception of the one being examined. The multijet distributions were obtained by requiring poor isolation for all cases except the isolation plot, in which case a poor χ^2 was required for the jet mimicking the electron

2.4.2 Quality Criteria

To be acceptable, candidates for Z production, the two decay electrons are required to be isolated and to satisfy the cluster shape requirements of f_{iso}< 0.15 and an H-matrix χ^2 < 100. In addition, at least one of the two electrons must have a good associated track, with σ_{trk}<5(10) for CC(EC) electrons. Finally, the events are required to have an invariant mass near the known value of the Z mass, $75 < M_{ee} < 105$ GeV/c^2. Table 2.2 gives a summary of the electron quality criteria used in this analysis. The "tight" criteria provide a cleaner sample of electrons.

Table 2.3 summarizes the total number of events passing all of the cutoffs in each topology. As mentioned previously, only the CCCC and CCEC events are used in the final analysis.

Parameter	Value
η -*CC*	$\lvert\eta_{det}\rvert < 1.1$
η -*EC*	$1.5 < \lvert\eta_{det}\rvert < 2.5$
Mass window	$75 < M_{ee} < 105\ GeV/c^2$
E_T	$> 25\ GeV$
crack ϕ *(CC)*	$\lvert\phi - \phi_o\rvert < 0.05\ rad$

Table 2.1 Fiducial volume requirements for accepted electrons.

e$^-$ Quality	"Tight" Value	"Loose" Value
f_{iso}	< 0.15	< 0.15
f_{em}	> 0.95	> 0.95
χ^2	< 100	< 100
σ_{trk}	< 5/10	none

Table 2.2 Electron quality criteria for signal. Each Z is required to have at least one "tight" and one "loose" electron, *i.e.,* at least one of the electron clusters in the calorimeter must have a well-matching track in the tracking chambers.

Figure 2.8 shows the full mass, and the p_T distribution for *ee* pairs with $75 < M_{ee} < 105$ GeV/c^2 in the final *Z* event sample (CCCC + CCEC). The event with highest p_T is at ~ 280 GeV/c. There are 157 events with p_T > 50 GeV/c. Figure 2.9 shows the p_T distribution sepa-

Topology	Events
CCCC	*3594*
CCEC	*2813*
ECEC	*673*

Table 2.3 Number of events according to event topology, after imposition of analysis requirements.

rately for the CCCC, CCEC and Fig. 2.10 shows the combined event sample. These distributions comprise the primary focus of this thesis.

2.5 Resolutions and Detector Modeling

Both the acceptance and the resolution-smeared theory were calculated using a simulation technique that was originally developed for measuring the W mass at D∅, and which has been used in many of the general W/Z analyses[29]. First, the mass of the Z is selected randomly according to an energy-dependent Breit-Wigner line-shape. The Z is then decayed at rest into two electrons that are boosted into the frame appropriate for the collision. The latter is defined by selecting a p_T and rapidity of the Z, at random from some given theoretical prediction. The positions and energies of the electrons are then smeared by the amounts expected from the resolution of each of the relevant detectors. Figure 2.11 shows the expected resolutions for various parameters of interest, and Fig. 2.12 compares Z data with simulated results for several physical distributions.

In addition, certain specific features of the experiment are also implemented in the above detector simulation. For instance, the electron-E_T requirement at the L2 trigger has a “turn-on”. That is, the trigger is not fully efficient for electrons with E_T = 25 GeV, but becomes fully efficient closer to 30 GeV[29]. The effect is small, as can be seen in Figure 2.13. Also included in the simulation is the tracking efficiency as a function of electron rapidity. The largest effect occurs near the edges of the tracking system, corresponding to rapidities of 1.0-1.1. The effect is measured using distributions observed in Z events, and parameterized for each of the calorimeters. The tracking efficiency and its parameterization are shown in Fig.

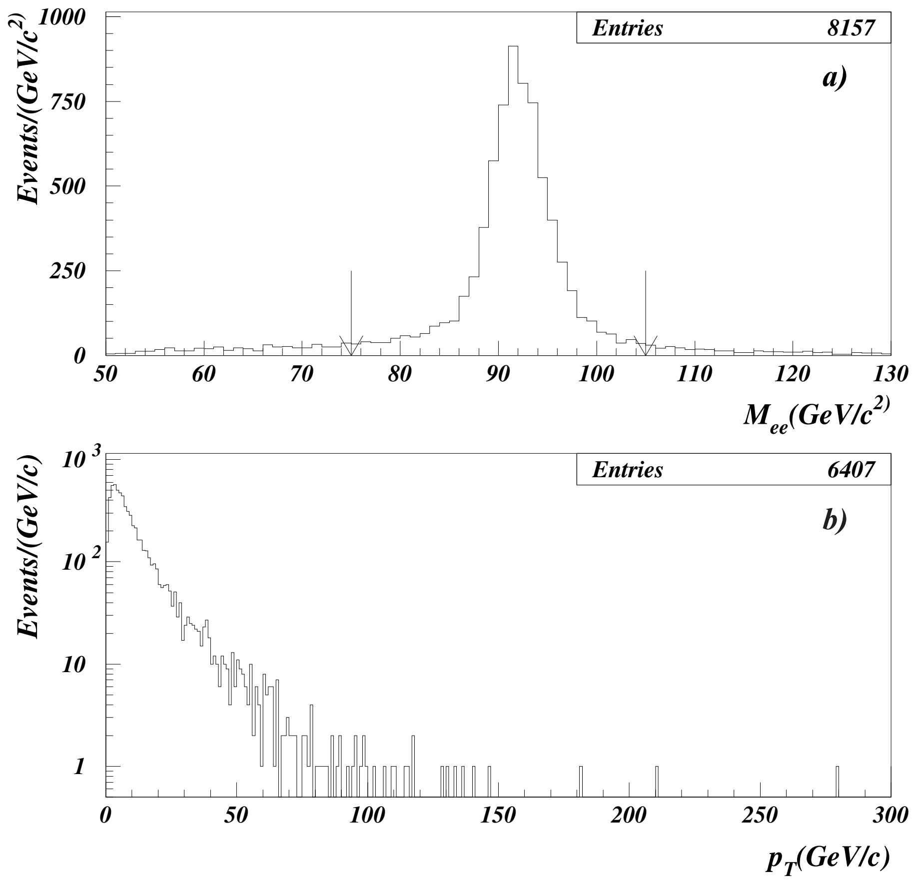

Figure 2.8 Mass distribution for all accepted *ee* pairs (a), and the p_T distribution for those events with $75 > M_{ee} < 105$ GeV/c^2 (b).

2.13. The details of the detector simulation can be found in Refs. [33], [45], and [46].The systematic uncertainties affecting the p_T^{ee}, which can be extracted from the detector simulation, will be discussed more fully in Section 3.1.3.

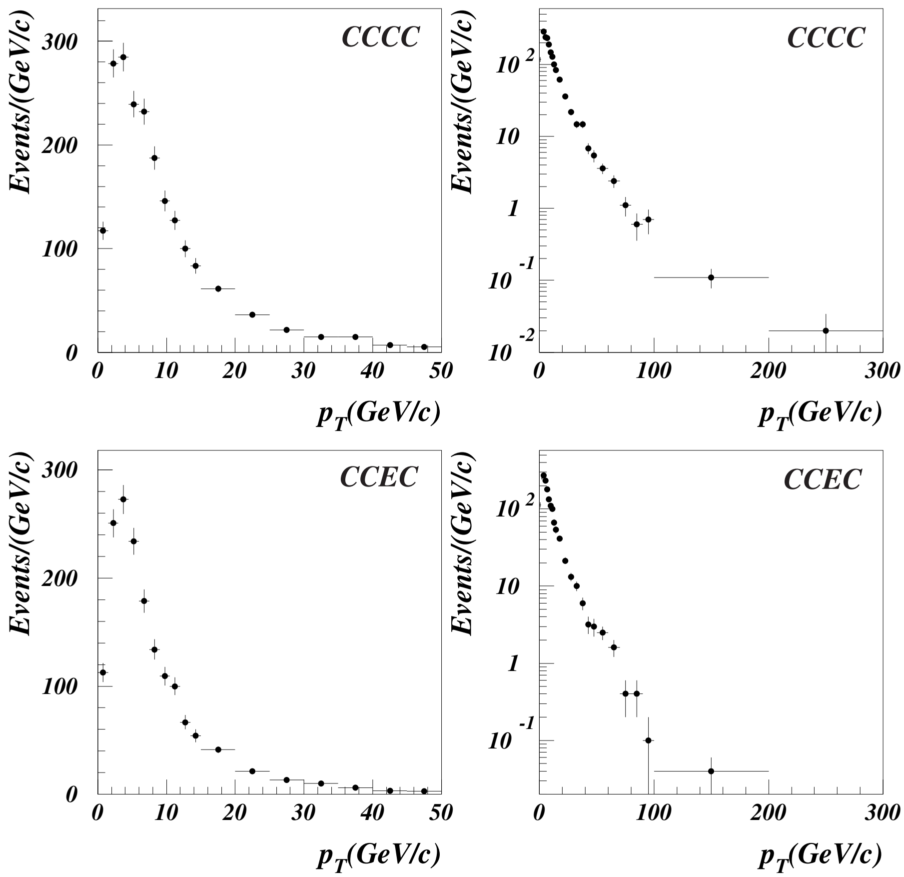

Figure 2.9 The p_T distributions of *ee* pairs with $75 > M_{ee} < 105$ GeV/c^2, separately for the CCCC and CCEC topologies.

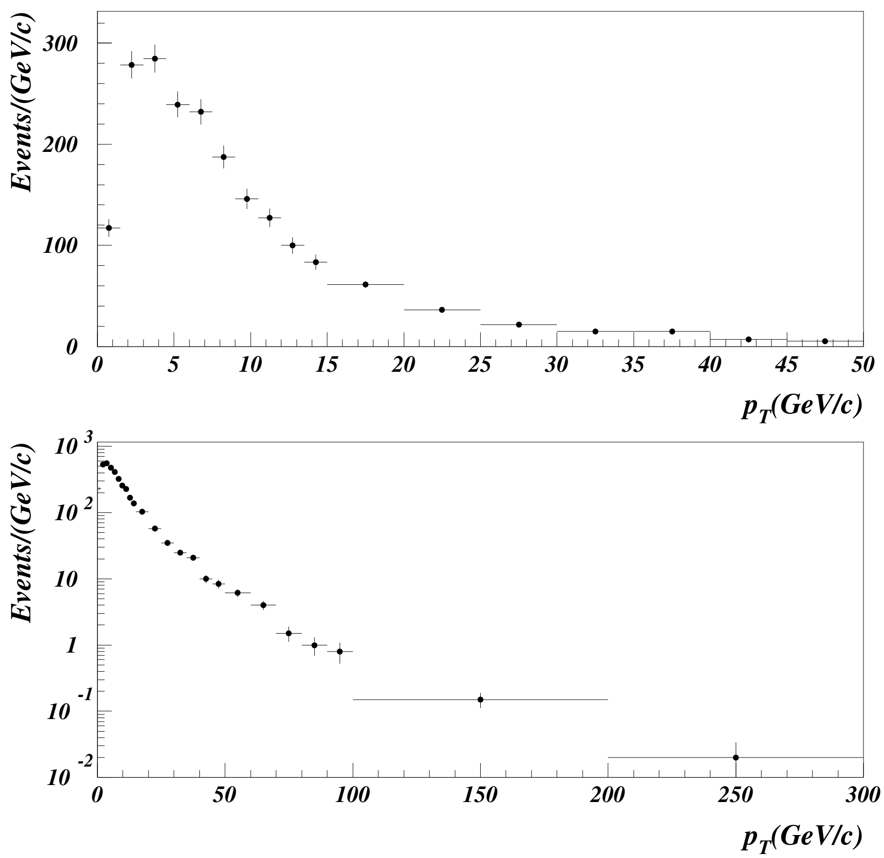

Figure 2.10 The p_T distributions of *ee* pairs with $75 > M_{ee} < 105$ GeV/c^2 for the sum of CCCC and CCEC topologies

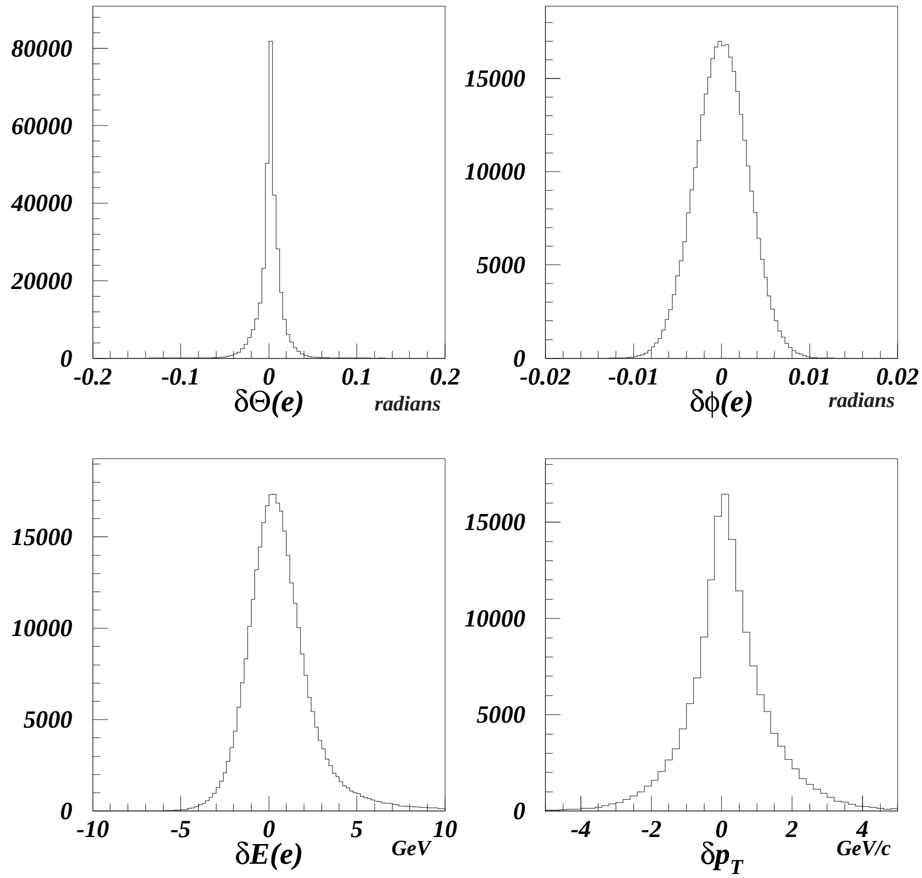

Figure 2.11 Resolutions expected for the θ, ϕ, and E of electrons, and for the p_T of the Z boson in $Z \rightarrow ee$ data.

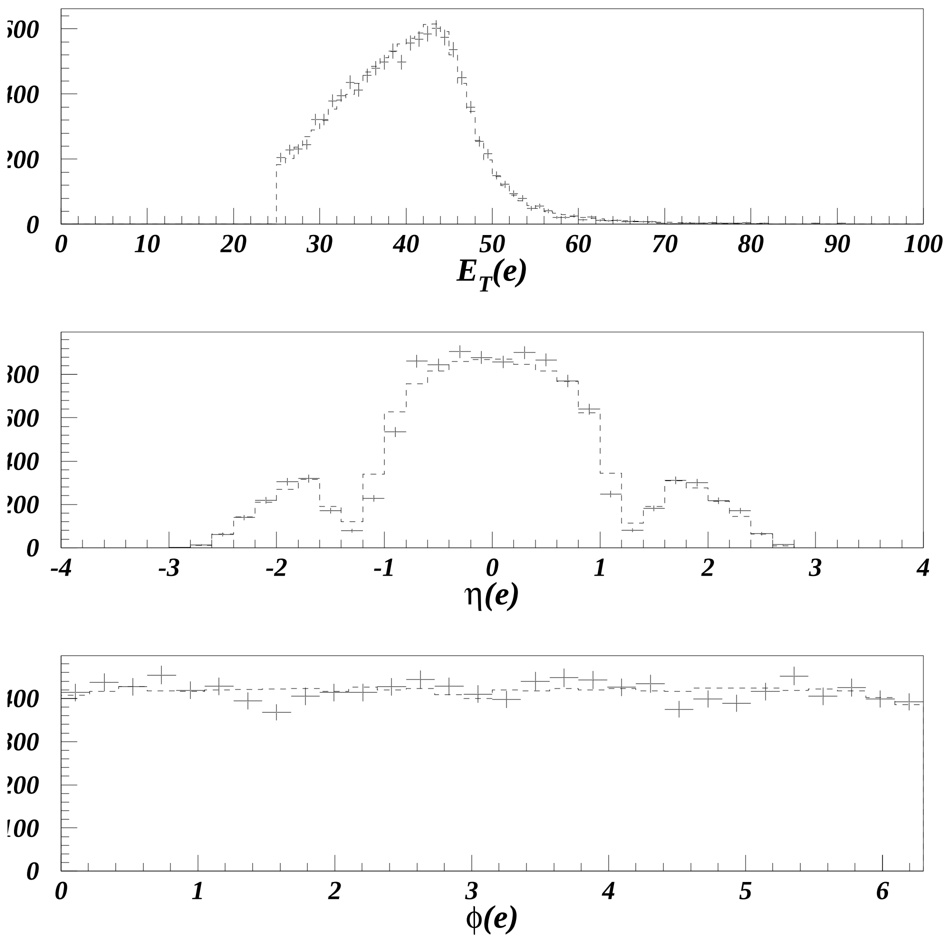

Figure 2.12 Comparison of *Z* data (crosses) to results of simulations (histograms).

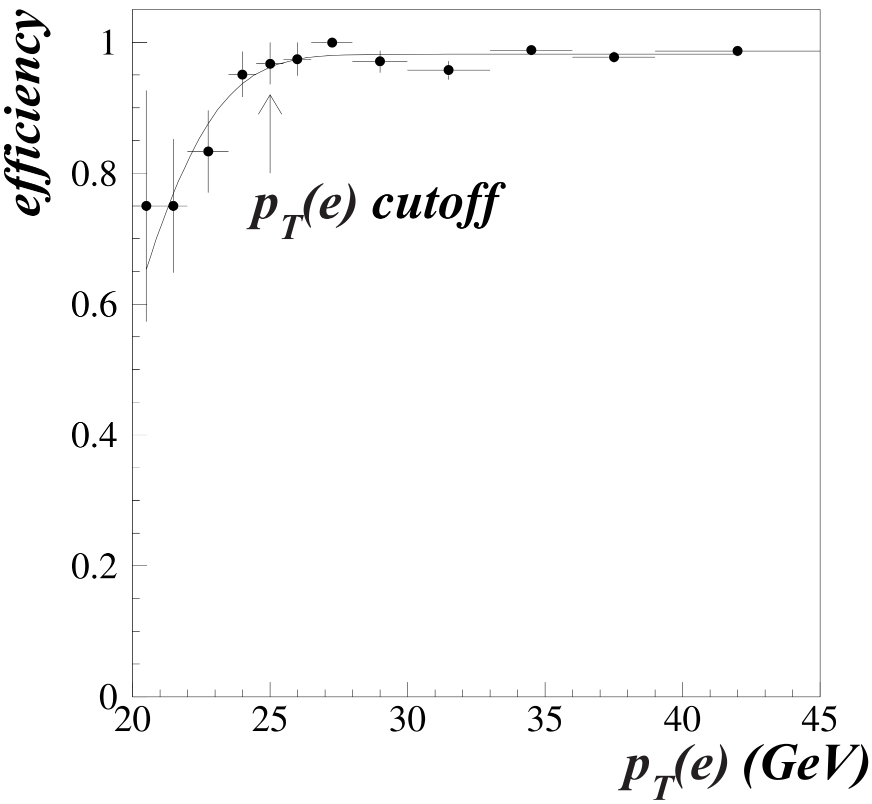

Figure 2.13 Electron detection efficiency as a function of p_T defined at L2.

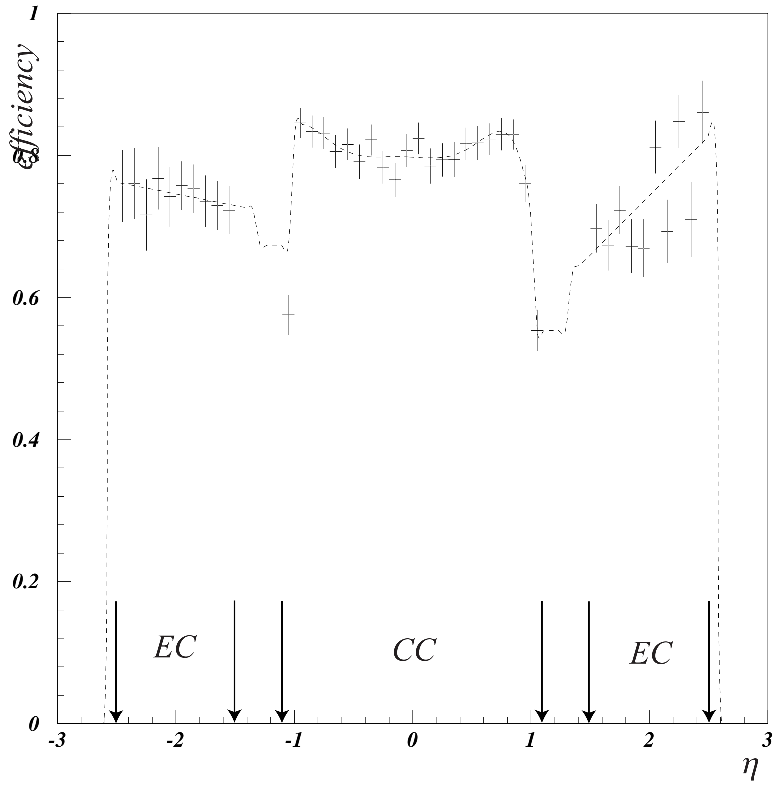

Figure 2.14 Tracking efficiency as a function of detector rapidity (crosses) compared to a fitted *ad hoc* dependence. The arrows indicate regions of acceptance for CC and EC electrons.

Chapter 3
Data Corrections

Once the data are selected, corrections are required in order to determine the differential cross section. The corrections are based upon our understanding of the detector, efficiencies and acceptances of signal, and the background events that get into our data sample (look like dielectron events, but are not).

3.1 Efficiencies and Acceptances

We enhance the number of signal events relative to background by examining each event in successively more detailed ways, rejecting both background and signal events along the way. Consequently, the effects of any choices of selection criteria must be studied and accounted for in the final result.

3.1.1 p_T-Dependent Efficiencies

We will first describe how we determine the efficiency of the electron selection criteria as a function of p_T of the Z. We have ignored the absolute electron identification efficiency because we normalize our result to the known integrated total cross section for Z production[29].

In principle, all the electron selection criteria can impact the observed p_T of the Z, however, it has been found that the p_T of the Z is most sensitive to the electron isolation requirement[46]. Nearby jet activity spoils the isolation of an electron, causing it to fail the selection criteria. The effect depends upon the detailed kinematics of the event, in particular, the location of hadronic activity (*e.g.*, associated jet production), and can depend upon the p_T of the vector boson.

Other analyses have wrestled with this problem[47]. In particular, for the M_W measurement, the effect of jet activity near an electron shower has been parameterized in terms of the projection of the transverse component of the total recoil E_T onto the direction of the p_T of the electron ($u_{||}$). This is illustrated in Fig. 3.1. The parameterization was obtained using a combination of simulated electrons and W data. The showering of the simulated electrons was generated with the DØ GEANT detector simulation program[48]. The simulated electrons were required to satisfy the standard kinematic criteria for W events, and were overlaid onto W events with the observed electron left in place. (It was required that the simulated electron not overlap with the observed electron.) The parameters for the simulated electrons (*e.g.*, E_T, isolation, χ^2) agreed well with the values for electrons observed in W data. As an example, Figure 3.2 compares the value of the average isolation (f_{iso}) versus $u_{||}$ for high-E_T electrons from W decays with results of the detector simulation. The agreement tells us that the effect of the hadronic activity on the electron is well modeled in the simulation. Although the parameterization was developed using W events, it is assumed to be applicable to Z events, because the parameterization summarizes the effect of hadronic activity on high-p_T electrons, regardless of the origin of that activity. That is, to the extent that hadronic activity in an event can be represented by $u_{||}$, the parameterization should be applicable to electrons in any relatively similar event.

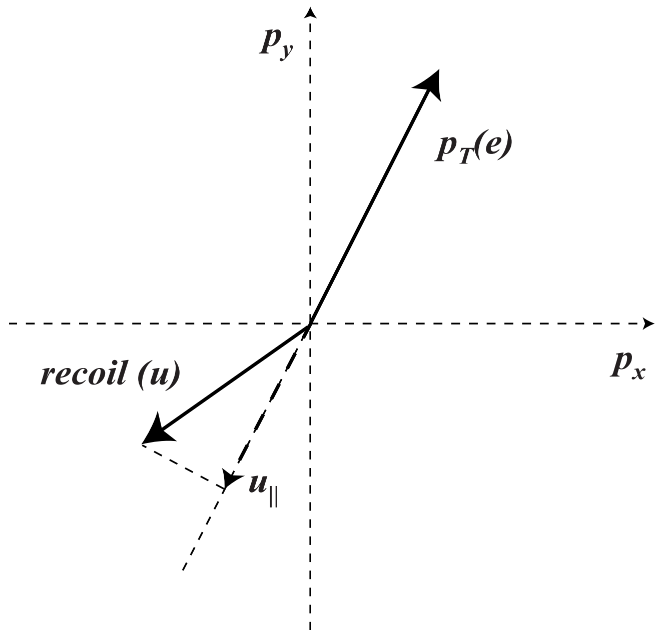

Figure 3.1 Illustration of the relationship between the transverse momentum of the electron, the vector E_T of the hadron recoil in the calorimeter, and $u_{||}$, the projection of the recoil onto the electron direction.

The final parameterization for the efficiency as a function of $u_{||}$ is:

$$\varepsilon(u_{||}) = \alpha \qquad (u_{||} < u_o) \tag{3.1}$$

$$\varepsilon(u_{||}) = \alpha(1 - s(u_{||} - u_o)) \qquad (u_{||} > u_o) \tag{3.2}$$

where u_o is the value of $u_{||}$ at which the efficiency begins to decrease as a function of $u_{||}$, and s is the rate of decrease. The values obtained from the best fit are $u_o = 3.85 \pm 0.55$ GeV, $s = 0.013 \pm 0.001$ GeV^{-1}. α is the overall efficiency, which is ignored in the parameterization. A complete description of the methodology used to obtain the parameterization can be found in Refs. [46] and [47].

Several assumptions used in deriving the $u_{||}$ parameterization make it incomplete for application to the measurement of the p_T of the *Z*. In particular, the analysis required p_T(W) < 30 GeV/c, thereby limiting the applicability of the parameterization to that region. In addition, one would expect that the use of the entire hadronic recoil to describe the jet activity might break down at higher p_T because of the likelihood of having higher accompanying jet multiplicities.

In order to overcome such drawbacks, and to obtain a reasonable parameterization of the electron identification efficiency for Z-p_T > 30 GeV/c, a p_T-dependent electron identification efficiency was obtained by using smeared $Z \rightarrow ee$ HERWIG events, overlaid onto randomly selected $p\bar{p}$ collisions (zero-bias events). The efficiency as a function of Z-p_T is defined by the ratio of the p_T distribution of the smeared HERWIG events, with kinematic, fiducial and electron quality requirements imposed, and the same distribution, but imposing only the kinematic and fiducial requirements. Figure 3.3 shows the resulting efficiency as a function of Z-p_T. The efficiency determined from the $u_{||}$ parameterization is also shown. The distributions have been normalized to each other in the region p_T < 30 GeV/c, where the $u_{||}$ parameterization is reliable. Also shown is the ratio of the two normalized results for p_T < 30 GeV/c. The agreement of the HERWIG analysis with the $u_{||}$ analysis is taken as confirmation of the HERWIG result, because, as mentioned before, the model of the $u_{||}$ analysis has been shown to be reliable for p_T < 30 GeV/c[47]. We will use the HERWIG result to obtain the p_T -dependent event selection efficiency for all p_T values.

Figure 3.4 shows the result of linear parameterization of the efficiency that will be used in the final correction to Z-p_T. In the p_T range of < 20 GeV/c, the fit has a slope of -0.004 (GeV/c)$^{-1}$ and intercept of 1.00. For p_T values greater than 20 GeV/c, the efficiency is taken as a constant of value 0.935, with an uncertainty of ±0.05.

Again, it should be noted that the above values of efficiency are not absolute. (In fact, the efficiency from the HERWIG study was normalized to that from the $u_{||}$ analysis.) The absolute efficiency includes not only effects from electron identification, but has other contributions from the L0 trigger, etc. By normalizing to the known total efficiency of 0.76 (integrated over all p_T values)[29], we deduce that a scale factor of 0.78 has to be applied to the above parameterization. Including this factor, the p_T dependence of the efficiency, in absolute terms, is given by (0.78±0.02)-(0.004±0.0001)p_T, for p_T < 20 GeV/c, and 0.73±0.04 for p_T >20 GeV/c. We assume the efficiency to be the same (constant) above 100 GeV/c as in the region of 20-100 GeV/c. This is the simplest assumption we can make given the statistics of the simulation. The efficiency at large p_T could be no larger than at $p_T = 0$, and would correspond to about a 1.5 standard deviation change in that region, and this will be reflected in the uncertainty on the extracted differential cross section. We do not expect the efficiency to decrease in the region beyond 100 GeV/c. Kinematically, the jets will tend to be in the hemisphere opposite the electrons in the events. Events with high jet multiplicity may have instances in which the large-E_T jets balance most of transverse momentum of the event, but smaller-E_T jets can overlap with one of the electrons. However, because the electrons are very energetic and therefore highly collimated, low energy jets are not likely to affect such electrons to the extent of causing them to fail the isolation criteria. We assign estimated uncertainties of ±3% per bin below 20 GeV/c, and ±5% in the region above 20 GeV/c.

3.1.2 Acceptance

The acceptance for dielectron pairs is defined by the fiducial and kinematic cutoffs used to select the signal. The parameterized detector simulation referred to in Section 2.5 is used to determine the overall acceptance as a function of p_T. The effects of the L2 trigger on E_T, cutoff on rapidity, the ϕ cracks in the calorimetry, rapidity-dependence of the tracking efficiency, and the final E_T requirements, are all included in the calculation. The detector simulation for Z events is run both with and without imposing fiducial and kinematic requirements, and then the two p_T distributions are divided bin by bin to obtain the acceptance as a function of p_T. Figure 3.5 shows how the various selection criteria affect the shape of the acceptance for the signal in p_T. The strongest effect comes from the E_T and rapidity requirements—the other

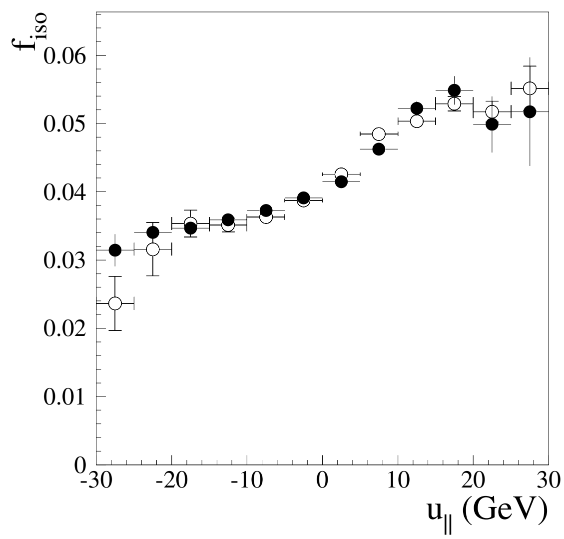

Figure 3.2 The mean electron isolation (f_{iso}) versus $u_{||}$ for a GEANT simulation (open circles) and high-E_T electron data from $W \to e\nu$ events (closed circles).

selections only change the normalization. Figure 3.6 shows the acceptance for the two event topologies, as well as for the combined event sample.

The mass requirement on the dielectron pairs was ignored in the final acceptance calculation. This simplifies comparisons with predictions, which are calculated at a fixed Z mass, and therefore does not take into account any dependence on dielectron mass. Figure 3.7 compares the p_T distribution for dielectron pairs with invariant mass near that of the Z to those with invariant mass above and below the nominal Z mass, and supports the expectation that any p_T dependence on mass (near the Z mass peak) must be very small.

3.1.3 Systematic Uncertainties

There are uncertainties on all the parameters used in the detector simulation. The effect of uncertainties in the energy scale and resolution for electrons, in the tracking[1], etc., on the total

[1] There is a z-dependent bias on the position of the center-of-gravity of a track in the CDC. The bias is linear and the scale refers to the rate of change of the bias as a function of z[46].

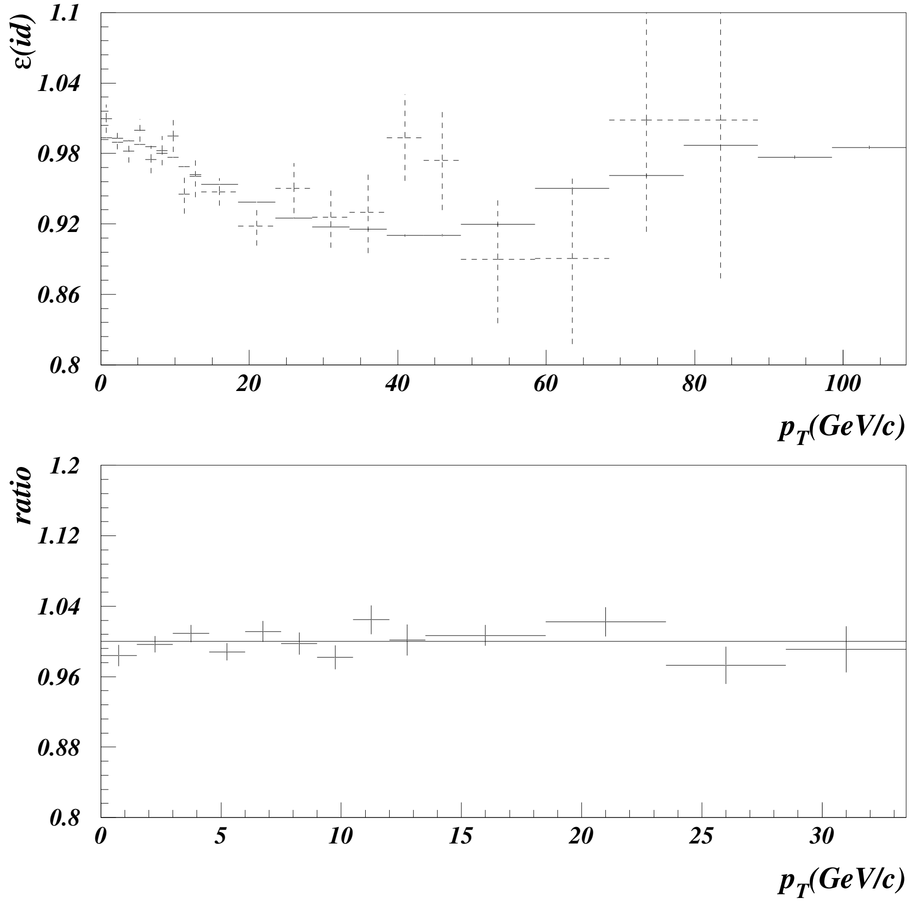

Figure 3.3 Comparison of the efficiency dependence on p_T as determined using HERWIG (dashed) $u_{||}$ (solid). The bottom plot shows the ratio in the range 0-30 GeV/c, where they are expect to agree.

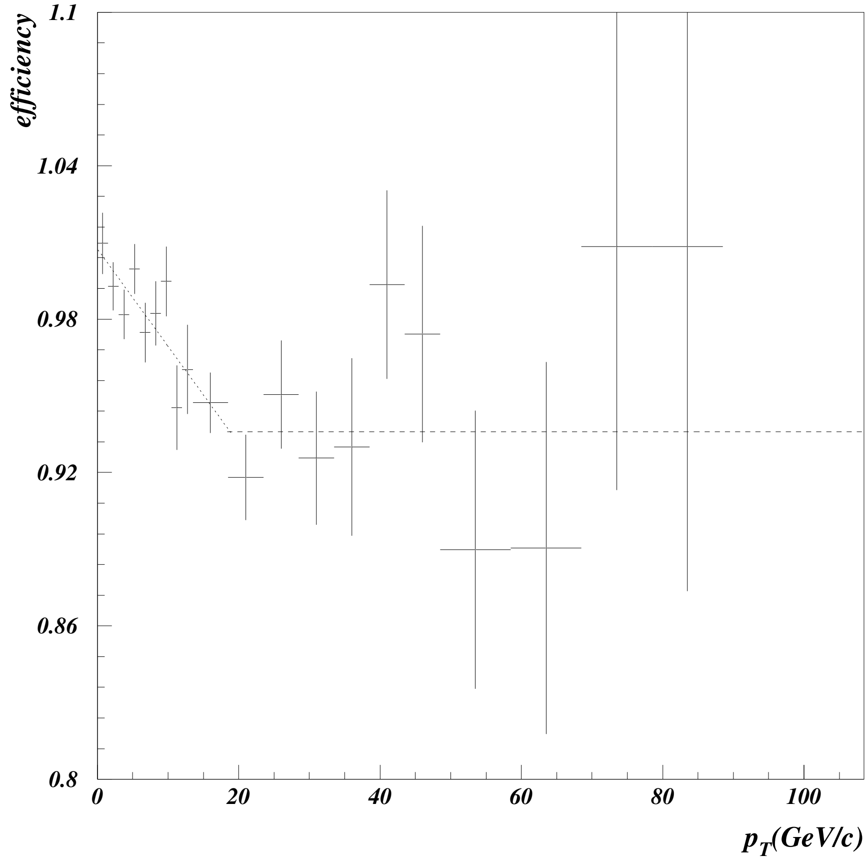

Figure 3.4 Final efficiency as a function of p_T. The top two plots show the individual fits used to obtain the parameterization of the efficiency. The bottom plot shows the final result with the uncertainties included.

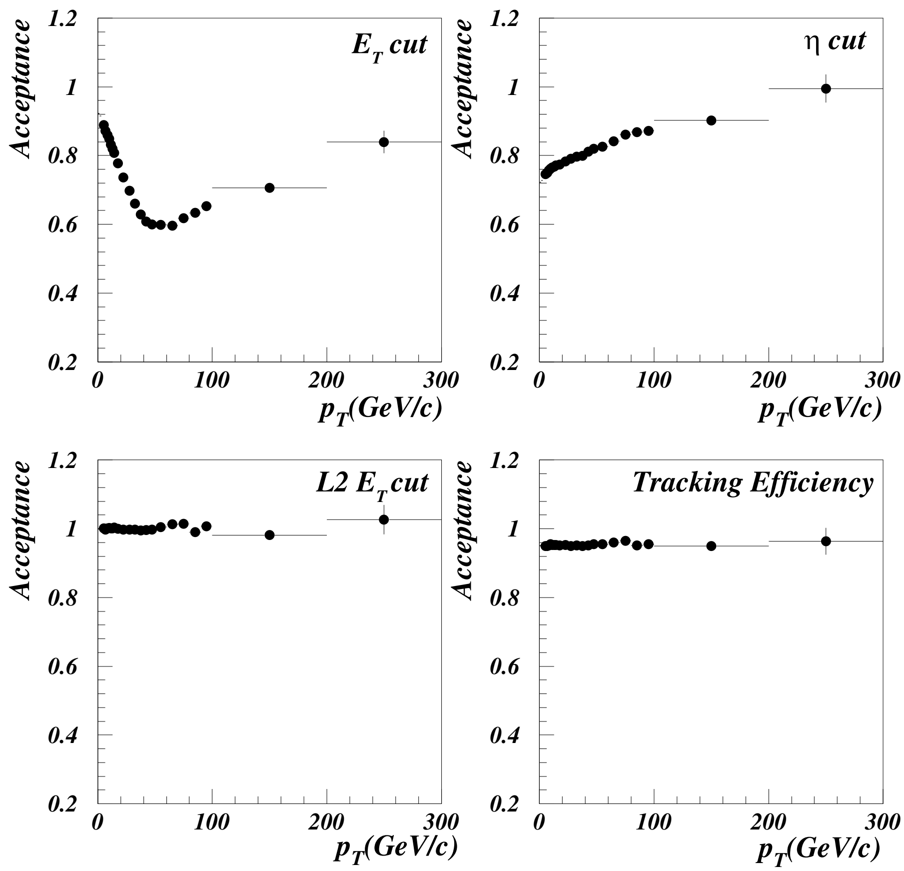

Figure 3.5 Effect of the fiducial and kinematic requirements on the acceptance as a function of transverse momentum of the *Z*.

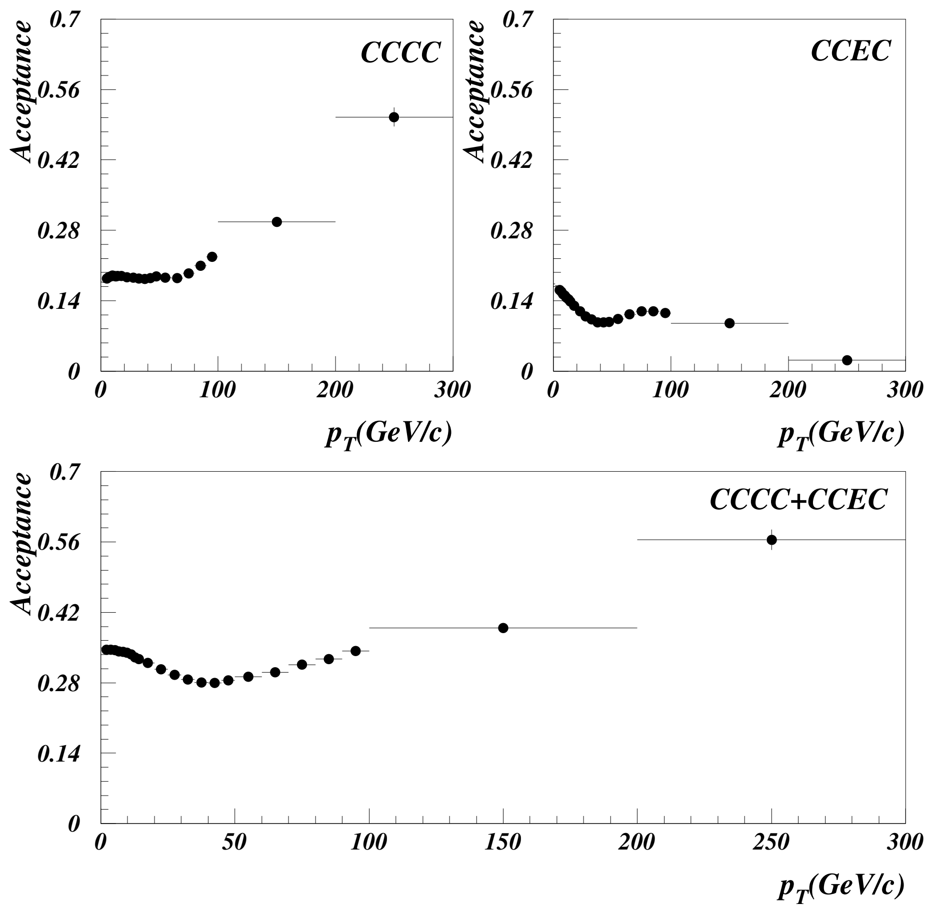

Figure 3.6 Final acceptance as a function of Z-p_T for the two primary topologies and for the complete event sample. In combining the CCCC and CCEC events, the distributions were weighted according to their respective number events.

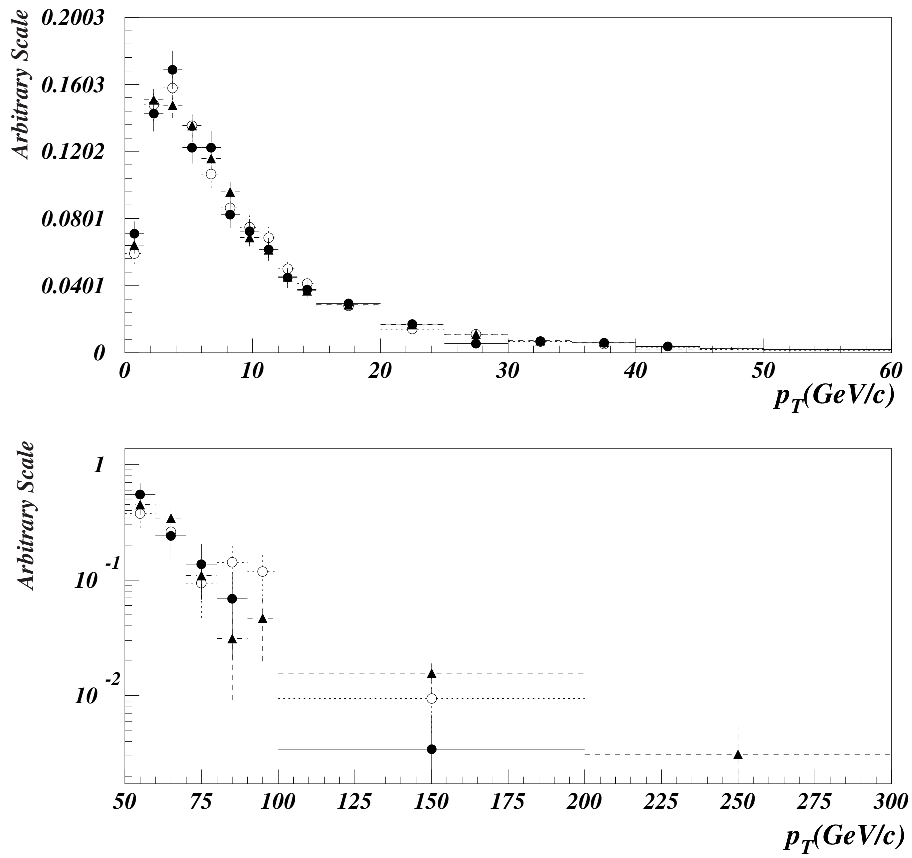

Figure 3.7 Comparison of the transverse momentum distribution of dielectron pairs with mass very close to the nominal Z mass, $90 < M_{ee} < 92$ GeV/c^2 (solid circles), to those in the mass regions $75 < M_{ee} < 90$ GeV (open circles) and $92 < M_{ee} < 105$ GeV/c^2 (closed triangles).

uncertainty in acceptance, was assessed for each bin of p_T, by varying the uncertainties by twice their known values (σ)[46]. (The 2σ value is used for purely historic reasons.) Figure 3.8 shows the nominal acceptance and those obtained by varying the different resolutions. The largest changes are observed at high p_T. If we parameterize this systematic uncertainty as a linear function of p_T, we obtain $\pm 0.01 + 0.0001 p_T$. This resulting band of uncertainty is also shown in Fig. 3.8.

Because we determine the acceptance bin by bin in p_T, we are less sensitive to the underlying model for the p_T spectrum used in the detector simulation. However, we are still sensitive to the assumed rapidity distribution of the *Z* in each bin of p_T. The uncertainty in the predicted rapidity of the *Z* is expected to be dominated by the parton distribution function (PDF) used for modeling Z production. The uncertainty in acceptance due to the choice of PDF was found to be ±0.2% for the inclusive measurement of the *Z* cross section[29]. This constrains the low p_T region, where the cross section is largest, to an uncertainty that is far smaller than the uncertainty from the variations in the detector model. The PDF dependence of the acceptance at high-p_T should also be small because the distribution of x values of the partons should be largely independent of the p_T of the *Z*. (This was verified using the ResBos simulation of vector boson production [40].) We ignore the additional small uncertainty from dependences of different PDFs on p_T.

3.2 Background

The primary background to dielectron production is multijet production, in which the jets have a large electromagnetic component (most of the energy is deposited in the EM section of the calorimeter) and/or they are mismeasured in some way that causes them to pass the electron selection criteria. There are also contributions to the dielectron signal that are not from misidentification of electrons, but correspond to other processes that differ from the one we

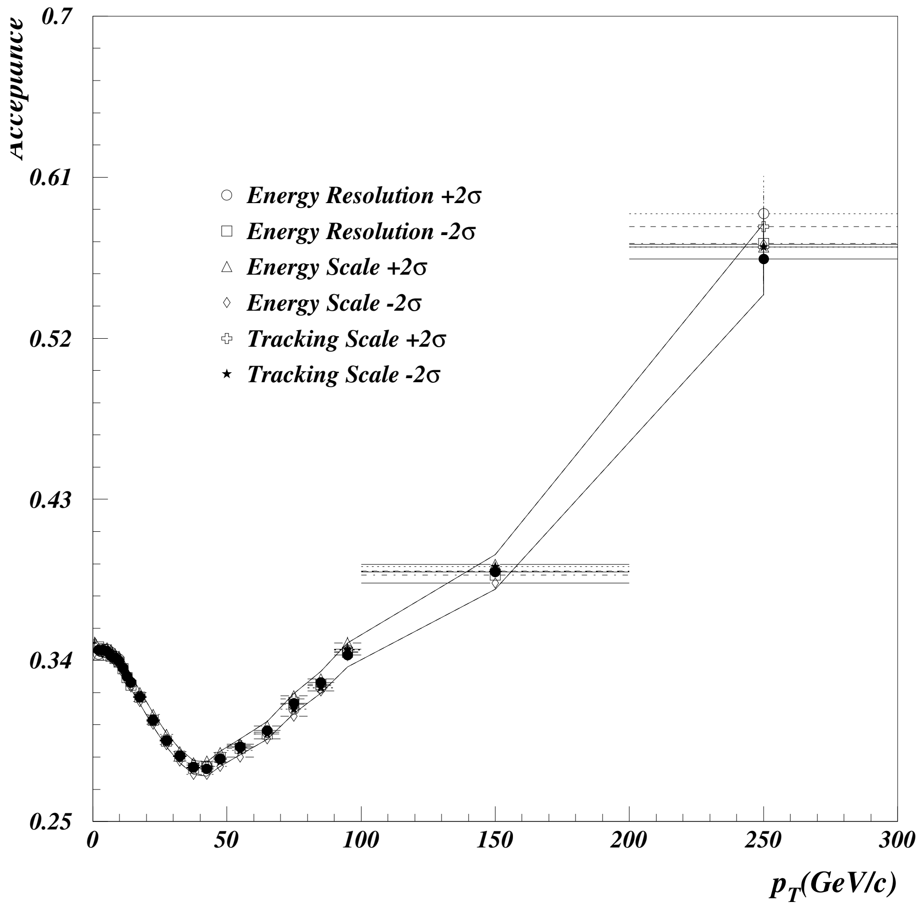

Figure 3.8 Effect of parameter uncertainties on the acceptance in each bin of p_T. The band corresponds to a parameterization of the uncertainty as a function of p_T, as discussed in the text.

are trying to measure, *e.g.*, $Z \rightarrow \tau\tau$ and $t\bar{t}$ production. Such incoherent processes have p_T dependences that are quite different from the Z/γ* mediated production of the *Z* and Drell-Yan pairs, and must be taken into account prior to comparison of data and theoretical prediction. Such backgrounds are discussed below.

3.2.1 Multijet Background

The two primary contributions to the multijet background arise from direct-γ production and multijet events. Figure 3.9 compares the lowest-order (α_s^2) Feynman diagrams for these processes to *Z* production. (Lowest-order direct-photon production is of order $\alpha_s\alpha_{EM}$.) Although the total cross section for these QCD processes is much larger than for *Z* production, the probability that a jet is mistaken for an electron in the calorimeter is only ~10^{-3} /jet[29], and consequently, we expect the total background to $Z \rightarrow ee$ events to be only a few percent. The cross section for inclusive photon production at $p_T > 25$ GeV is about 10^{-3} of inclusive jet production[49]. Because only one jet must be mistaken for an electron, the consequent contribution of direct-photon production to the background equals (approximately) that from all-jets. (For both sources there must also be a charged track associated with the photon or the electromagnetic-like jet in the calorimeter.)

The two QCD mechanisms have similar p_T production distributions, but can differ in their contributions to the background because of the nature of the required misidentifications. Also, since only certain jets can mimic electrons, one might expect that only highly electromagnetic jets (objects that were electron candidates because most of their energy was deposited in the EM section of the calorimeter, but that failed the final quality criteria) would be of interest for describing the shape of the background.

Three types of event samples were studied in order to examine if differences in production mechanism or detector resolution would result in a significant variation of the background characteristics: dijet events, direct-γ events, and dielectron events in which both electrons

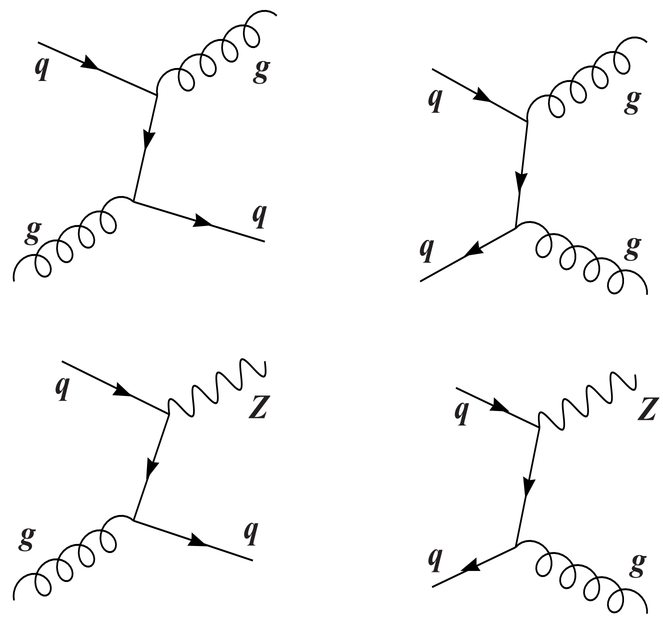

Figure 3.9 Lowest order (α_s^2) Feynman diagrams for QCD background processes compared to diagrams for *Z* signal.

failed the quality criteria. Tables 3.1 - 3.3summarize the selection requirements for each background sample. Detailed definitions of the cutoffs used in the multijet sample are given in Ref. [50]. Similarly, the definitions for cutoffs in the direct-γ sample are given in Ref. [51]. Figures 3.10 - 3.13show the invariant mass and p_T distributions for the samples in both the CCCC and CCEC event topologies. All the distributions appear to agree. It can be concluded that differences in production mechanism and detector resolution are inconsequential for determining the shape of the background to dielectron events.

The direct-γ sample was chosen to characterize the background, for two reasons. First, the sample has the greatest number of events. Second, one expects the direct-γ data sample to provide a good approximation of the combination of backgrounds from dijet and true direct-γ contribution to the background. This is because about half of the direct-γ sample consists of misidentified jets, and therefore has the approximate balance of dijet and direct-γ events expected from QCD background. (As discussed at the beginning of the section, the direct-γ

and multijet events are expected to provide about the same level of contribution to the background because the rate of jet misidentification is comparable to the difference between cross sections for the two processes.)

Parameter	Value
η -*CC*	$\lvert\eta_{det}\rvert < 1.1$
η -*EC*	$1.5 < \lvert\eta_{det}\rvert < 2.5$
Mass window	$75 < M_{jj} < 105$ GeV/c^2
E_T	*> 25 GeV*
$\dfrac{\not{E}_T}{E_T^{leading\ jet}}$	*< 0.7 GeV*
$f_{hotcell}$	*20*
f_{ch}	*< 0.4*
f_{em}	$0.05 < f_{em} < 0.95$
trigger	*jet_min*

Table 3.1 The cutoffs used to generate the *jj* sample. These were chosen to be in accord with the QCD group's "good-jet" requirements. A description of the parameters can be found in Ref. [50].

Parameter	Value
η -*CC*	$\lvert\eta_{det}\rvert < 1.1$
η -*EC*	$1.5 < \lvert\eta_{det}\rvert < 2.5$
Mass window	$75 < M_{j\gamma} < 105$ GeV/c^2
E_T	*> 25 GeV*
ϕ *crack in CC* (γ)	$\lvert\phi\rvert \pm 0.05$
$E_T^{iso}(\gamma)$	*< 2.0*
$f_{em}(\gamma)$	*> 0.95*
$\chi^2(\gamma)$	*< 100*
tracking(γ)	*no track*
trigger	*em1_gis*

Table 3.2 The cutoffs used to generate the γj sample and were chosen to be in accord with the direct-photon group's analysis. Unless noted to the contrary, each cutoff was applied to both the photon and the jet. See Ref. [51] for a description of the parameters.

Parameter	Value
η *-CC*	$\lvert\eta_{det}\rvert < 1.1$
η *-EC*	$1.5 < \lvert\eta_{det}\rvert < 2.5$
Mass window	$75 < M_{ee} < 105$ GeV/c^2
E_T	*> 25 GeV*
χ^2	*> 150*
trigger	*em2_eis_esc*

Table 3.3 The cutoffs used to generate the “bad-electron” sample, which were used to study background to *Z* events.

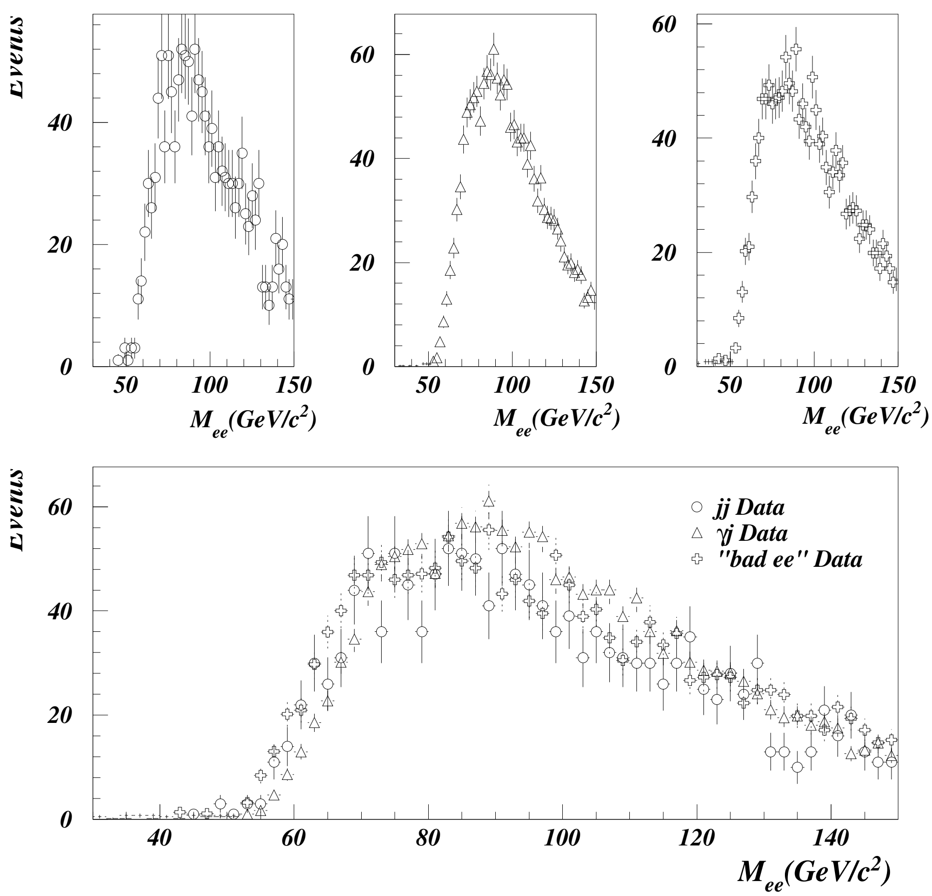

Figure 3.10 Invariant mass distributions for the three types of QCD multijet background samples in the CCEC topology.

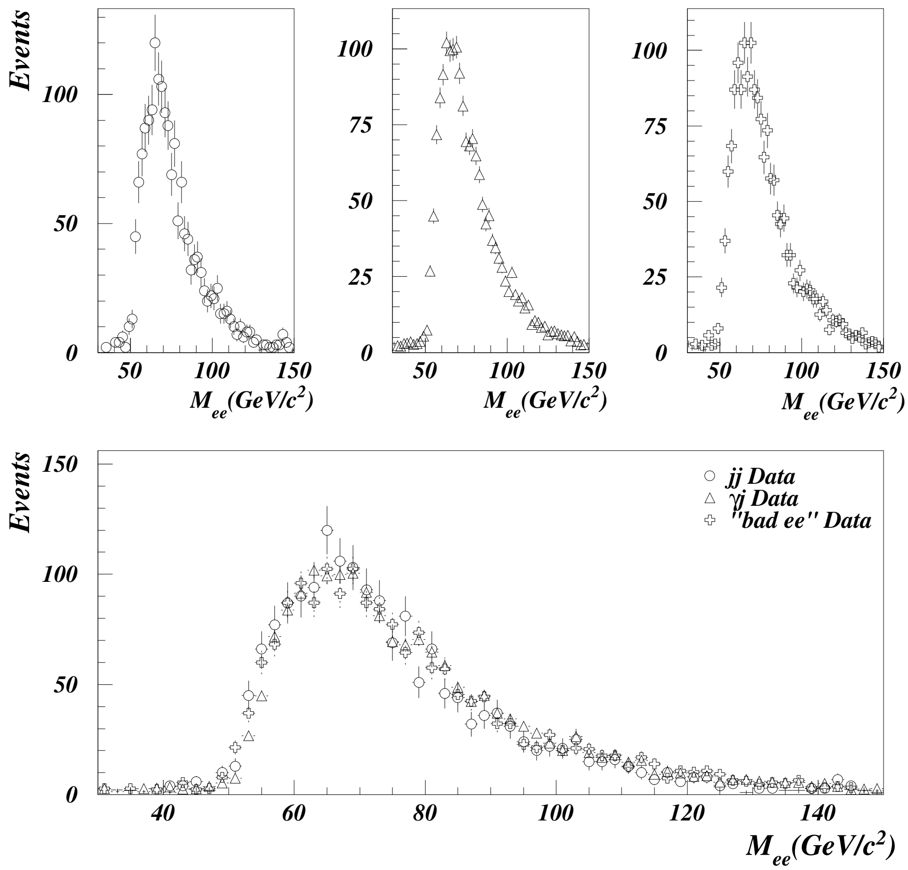

Figure 3.11 Invariant mass distributions for the three types of QCD multijet background samples in the CCCC topology.

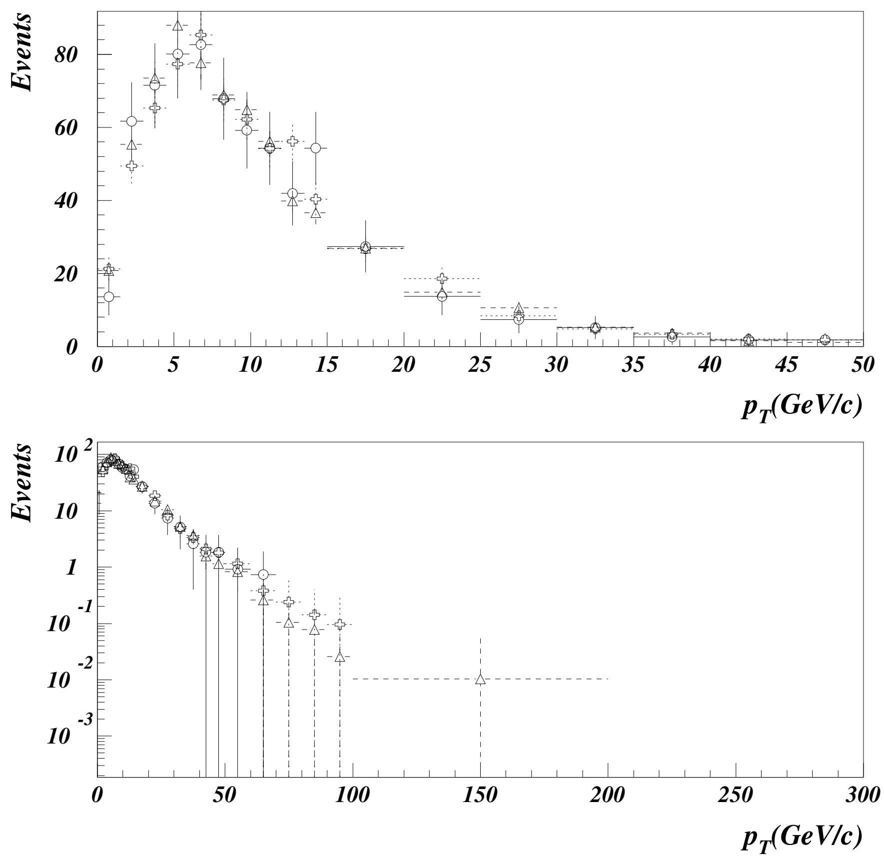

Figure 3.12 Comparison of the shape of the transverse momentum distributions for the three multijet background samples for the CCEC event topology.

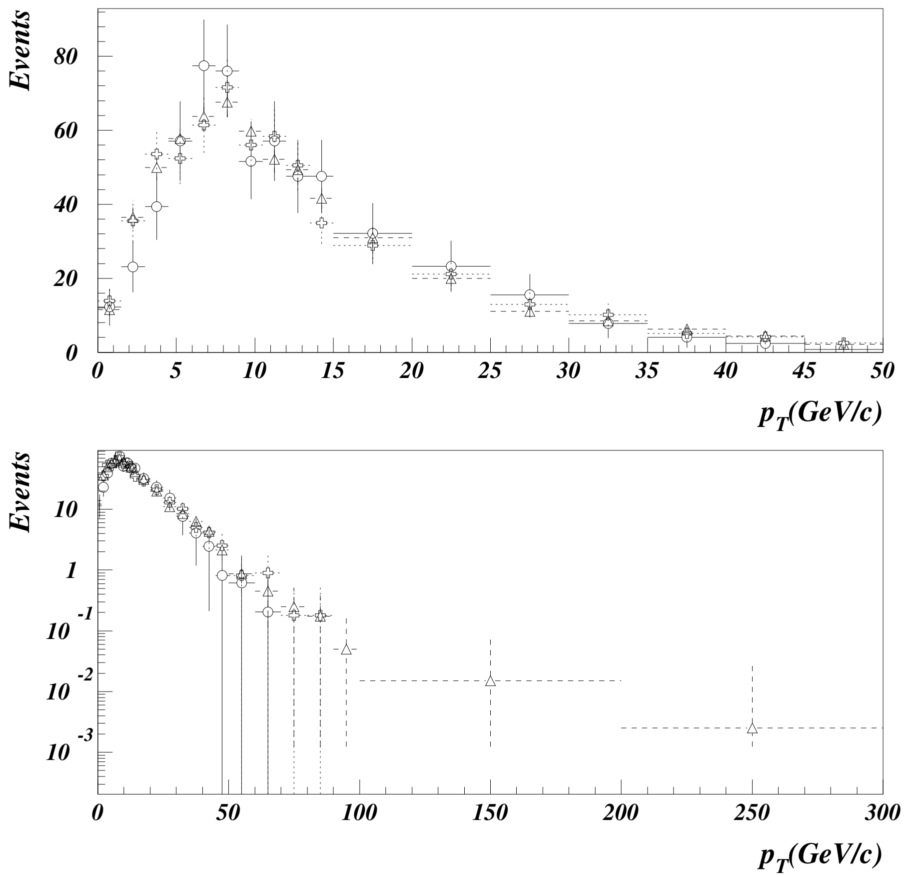

Figure 3.13 Comparison of the shape of the transverse momentum distributions for the three multijet background samples for the CCCC event topology.

3.2.2 p_T-Dependence of QCD Background

Because the shape and amount of QCD background differs in the CCCC and CCEC event topologies, the final background was obtained by combining the CCCC and CCEC results, weighted by their relative contributions to the background. That distribution was then normalized to the appropriate background fraction. The determination of both the relative and total background contributions is discussed below. Figure 3.14 shows the dependence of the background on p_T for the combined events in the CCCC and CCEC regions. The combined result was fitted separately for $p_T < 20$ GeV/c and $p_T > 20$ Gev/c, using the following parameterization:

$$C(p_T + \alpha)^2 e^{\beta p_T} \qquad (p_T < 20\ GeV/c) \tag{3.3}$$

$$e^{\alpha + \beta p_T} \qquad (p_T > 20\ GeV/c) \tag{3.4}$$

Figure 3.15 shows the results of the fit. The values of the extracted parameters are given in Table 3.4.

	$p_T < 20$ GeV/c	$p_T > 20$ GeV/c
α	1.28±0.26	4.37±0.25
β	-0.24±0.01	-0.094±0.008
C	2.84±0.44	—

Table 3.4 Values of the parameters obtained from the fit to the multijet background.

3.2.3 QCD Multijet Background Level

Because the background dependence on mass is different for each event topology, the level of the multijet background was determined separately for each set of data. Using this background and the known resonance of the *Z*, we can obtain the relative contribution of background by performing a maximum-likelihood fit for the amounts of background and signal in the data. (This was the method used in for determining the total background from multijet events in our previous measurement of the total cross section of the *Z*[29].)

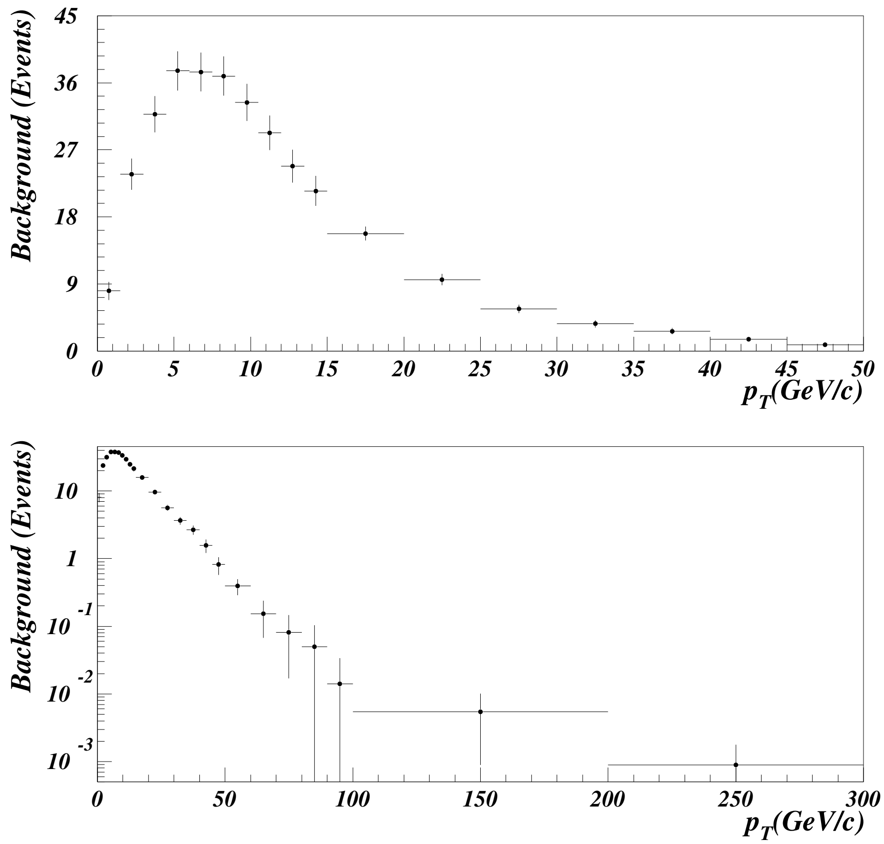

Figure 3.14 Final background distribution as a function of p_T. The distribution for the two event topologies have been combined according to their fraction in the total background.

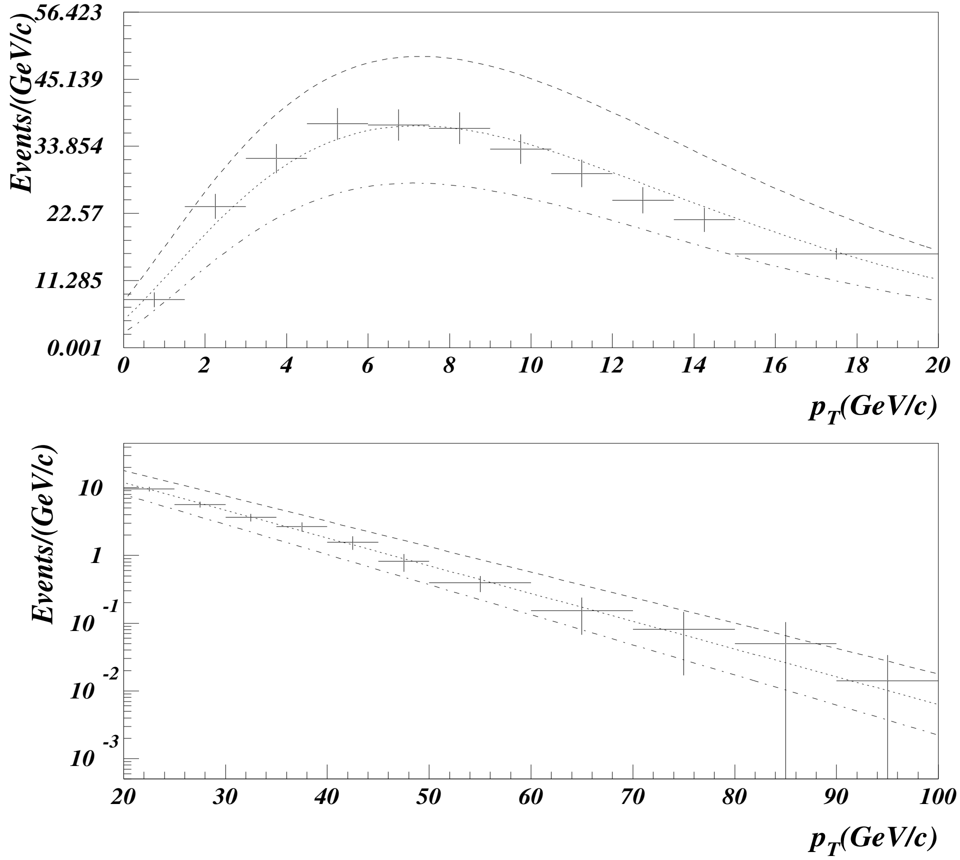

Figure 3.15 Fits to obtain the shape of the multijet background as a function of p_T. The dotted lines correspond to coherent 1σ deviations in the values of the fit parameters. The uncertainties on the points are purely statistical.

The Monte Carlo generator PYTHIA[52] is used to produce the invariant mass spectrum for the signal, which contains contributions from both *Z* and Drell-Yan production (and their quantum-mechanical interference). The generated four-momenta were then smeared using the detector simulation described previously. The direct-γ data sample was used to represent the invariant mass spectrum of the background.

We obtained the amount of multijet background contaminating the data sample by performing a binned maximum likelihood fit to the sum of the signal (PYTHIA) and background (direct-γ data as described in the previous section):

$$N_{data}(m_i) = c_1 \cdot N_{PYTHIA}(m_i) + c_2 \cdot N_{\gamma + jet}(m_i) \qquad 3.5$$

where c_1 and c_2 are the normalization factors for the signal and background contributions respectively, and m_i is the i^{th} mass bin. The fit was performed in the dielectron invariant mass window of 50-130 GeV/c^2. Figure 3.16 shows the best fit to the dielectron invariant mass data for each of the event topologies. The final value for the multijet background in the data is defined by normalizing the fitted parameter c_1 to the number of events observed in the mass window of the *Z* (75-105 GeV/c^2):

$$f_{back} = c_1 \cdot \frac{N_{tot}(data)}{N_{tot}(\gamma+jet)} \cdot \frac{N_{75\text{-}105}(\gamma+jet)}{N_{75\text{-}105}(data)} \qquad 3.6$$

where

$$N_{tot}(sample) = \sum_{all\ \ m_i} N_{sample}(m_i) \qquad 3.7$$

$$N_{75-105}(sample) = \sum_{75 < m_i < 105} N_{sample}(m_i) \qquad 3.8$$

Table 3.5 summarizes the background contributions in each of the topologies, where the specified uncertainties are purely statistical[53].

The systematic uncertainty on the background in each bin is assigned to be 25%, except for the last two bins, which are assigned uncertainties of 50%. The 25% value is dominated by the 20% uncertainty in overall normalization, and the extra 10% (added in quadrature) is taken from the fit.

3.2.4 Multijet Background at High p_T

In order to estimate the amount of background in the data at high p_T, we use the result shown in Section 3.2.2, but consider only p_T values above 50 GeV/c. We use the normalization obtained from the total background contribution, which is dominated by low-p_T production. We derive that 7.8±2 multijet events are expected to contaminate the dielectron sample for p_T > 50 GeV/c.

3.2.5 Other Sources of Dielectron Signal

Although $Z \to ee$ and direct-γ events make up nearly the entire dielectron signal, there are contributions from other sources, such as $Z \to \tau\tau \to eeX$, $t\bar{t} \to eeX$ and diboson (WW, ZZ, WZ, Wγ, Zγ) production in dielectron final states. The expected contributions from these sources are estimated below.

3.2.5.1 $Z \to \tau\tau$ Contribution

The $Z \to \tau\tau$ contribution to the dielectron signal has been examined in a previous analysis [54]. The dielectron event rate in our accepted range of mass from $Z \to \tau\tau$ is calculated as < 2.6×10^{-6} per Z event[54]. For the current sample, this corresponds to < 0.009 events for all values of dielectron p_T. Thus this contribution is ignored.

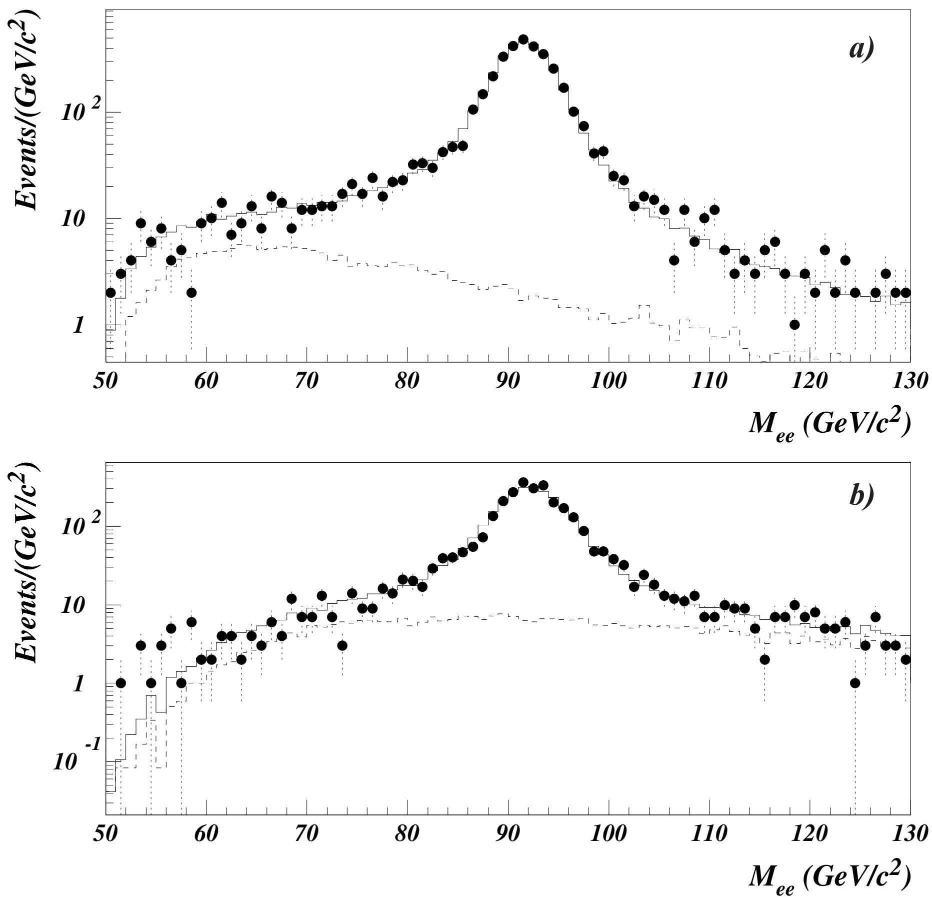

Figure 3.16 Comparison of the dielectron invariant mass distribution (closed circles) to the fit of PYTHIA plus background (histogram). Plot a) is for the CCCC and b) is for the CCEC topologies.

Data	f_{back}
CCCC	0.043 ± 0.01
CCEC	0.057 ± 0.01
Total	0.049 ± 0.01

Table 3.5 Background fractions in the two primary event topologies.The total background is obtained by summing the individual topologies, weighted by their fraction of events. The uncertainties are purely statistical.

3.2.5.2 $\mathbf{t\bar{t}}$ Contribution

The dielectron background contribution from $t\bar{t}$ production is concentrated mainly at high p_T of the $Z \to ee$ signal (Fig. 3.17). The fraction was determined using the HERWIG simulator for $t\bar{t}$ production, smeared with the known D∅ detector resolutions[48]. Electron contributions from both $W \to e\nu$ and $W \to \tau X \to eX$ channels were considered. For a $t\bar{t}$ cross section of 6.4 *pb* [26], and the standard branching ratios for the W, the calculated geometric and kinematic acceptance from HERWIG was 0.01±0.006. Including electron identification efficiency for our dielectron events, we expect ~ 0.36 events in the entire sample and ~0.2 events with dielectron $p_T > 50$GeV/c. Considering the small number of events expected, the $t\bar{t}$ contribution is ignored. Figure 3.17 displays the mass spectrum and p_T distribution expected from $t\bar{t}$ background.

3.2.5.3 Diboson Contribution

Figures 3.18 - 3.20 show the p_T spectra for the two highest-E_T electrons in WW, ZZ, and Wγ events generated with the HERWIG simulator, using the known D∅ detector resolutions. No Wγ events out of ~3000 generated passed the cutoffs. The reason is that very few events have photons with $E_T > 25$ GeV/c and an invariant mass ($M_{e\gamma}$) near the Z mass.

The resulting acceptances for background events with $p_T > 50$ GeV/c are given in Table 3.6. Included in the table are the assumed production cross sections multiplied by branching ratios ($\sigma \cdot B$) for W and Z into electron states. The WW cross section ($10.2^{+6.3}_{-5.1}$ *pb*) and branching ratio to dielectrons (0.011) was obtained from Ref. [55]. The value $\sigma(W\gamma) \cdot B(W \to e\nu) = 4.17 \pm 0.3\ pb$, was obtained from Ref. [56] and assumes $p_T^{\gamma} > 10$ GeV/c. The WZ and ZZ cross sections are each taken from plots in Ref. [57]. Given their small size, these contributions have been ignored in our analysis.

Process	Acceptance for dielectron satisfying *Z* signal criteria	$\sigma \cdot$ *Branching ratio* (pb)	Expected events
$t\bar{t}$	*0.01±0.006*	*0.43±0.01 pb*	*~0.2*
$W\gamma$	*<0.0003*	*4.17±0.3 pb*	*<0.1*
WW	*0.016±0.007*	*0.06±0.03*	*~0.009*
WZ	*0.016±0.007*	*0.03±0.01*	*~0.005*
ZZ	*0.046±0.002*	*0.03±0.01*	*~0.005*

Table 3.6 The expected numbers of events from diboson and top processes. Each case assumes a total luminosity of 108.5 and dielectron identification efficiency of 0.73.

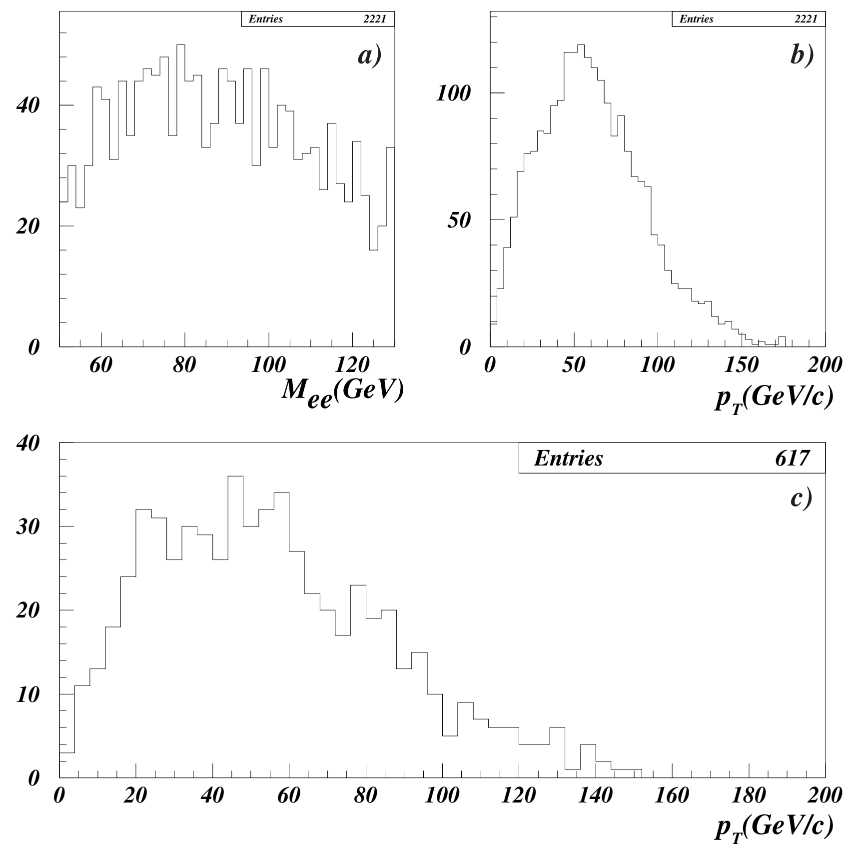

Figure 3.17 Dielectron spectra from $t\bar{t}$ events generated using the HERWIG Monte Carlo. Kinematic and fiducial selection criteria have been applied. (a) shows the mass spectrum of the dielectron pairs near the mass of the Z, (b) shows the p_T spectrum for all events, and (c) the p_T spectrum after requiring $75 < M_{ee} < 105$ GeV/c^2.

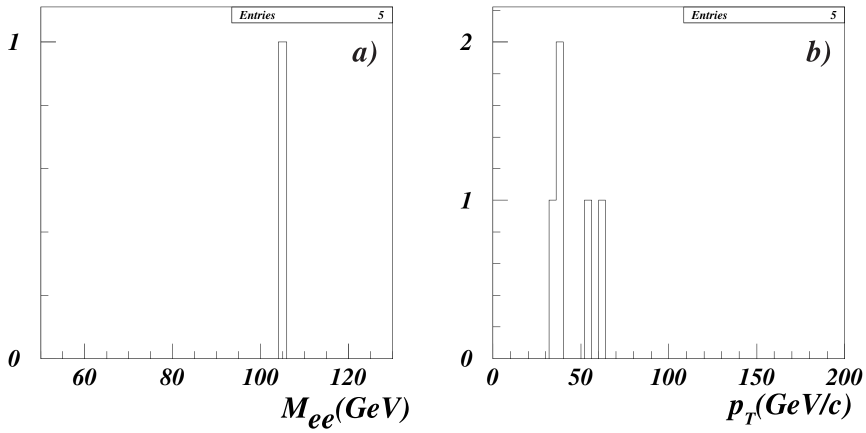

Figure 3.18 Dielectron spectra from $W\gamma$ events generated using the HERWIG Monte Carlo. Kinematic and fiducial selection criteria have been applied. (a) shows the mass spectrum of the dielectron pairs near the mass of the Z, (b) shows the p_T spectrum for all events. Out of ~3000 original events, none remain in the mass range of the Z ($75 < M_{ee} < 105$ GeV/c^2).

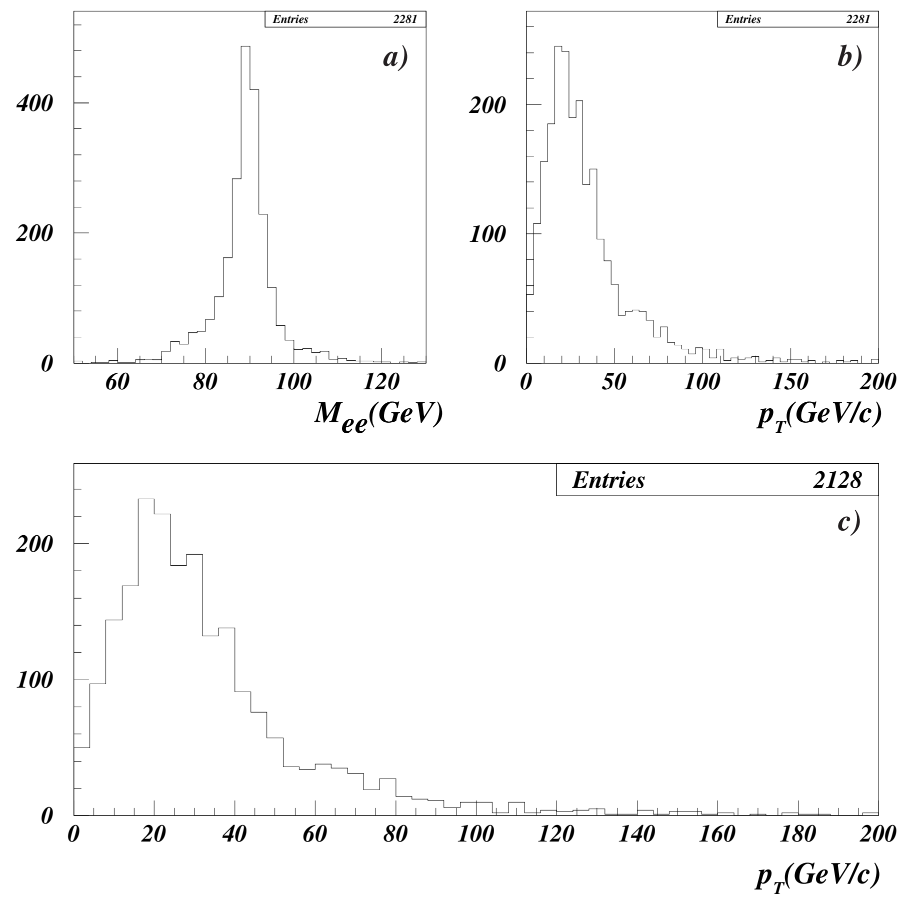

Figure 3.19 Dielectron spectra from $ZZ \to eejj$ events generated using the HERWIG Monte Carlo. Kinematic and fiducial selection criteria have been applied. (a) shows the mass spectrum of the dielectron pairs near the mass of the Z, (b) shows the p_T spectrum for all events, and (c) the p_T spectrum after requiring $75 < M_{ee} < 105$ GeV/c^2.

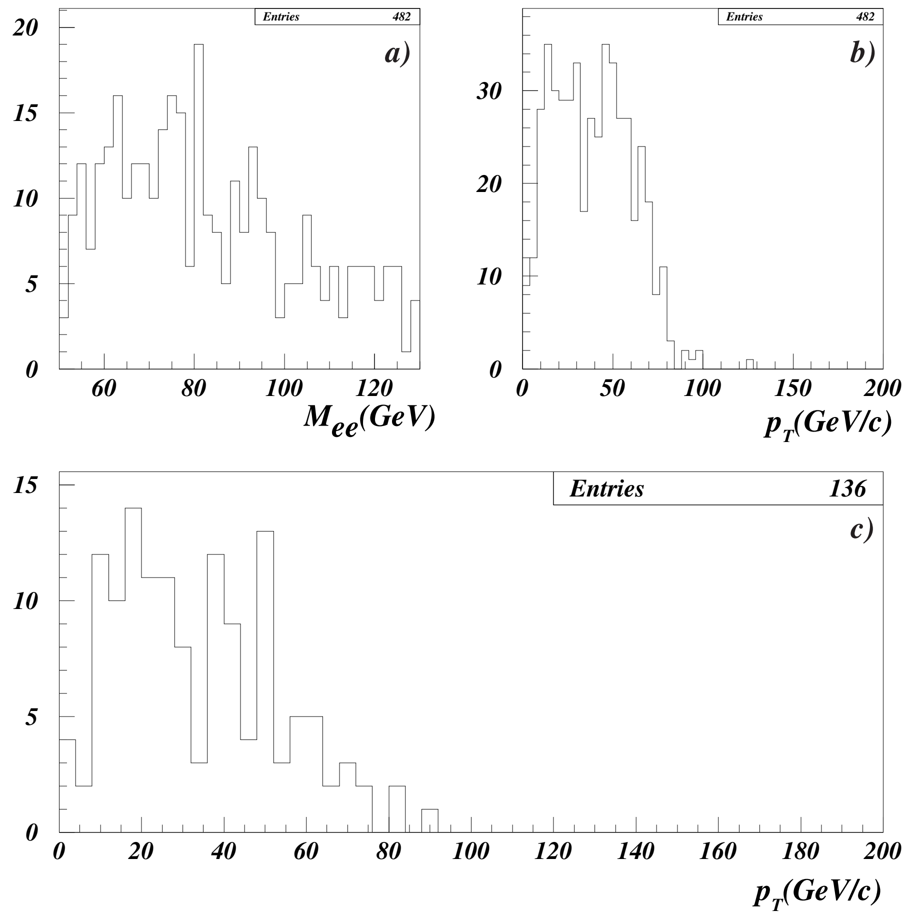

Figure 3.20 Dielectron spectra from $WW \to ee$ events generated using the HERWIG production Monte Carlo with D0 detector smearing. Kinematic and fiducial cuts are also applied. The top left plot shows the mass spectrum of the dielectron pairs near the mass of the Z. The top right plot shows the p_T spectrum for the whole event sample. The bottom plot is the p_T spectrum after requiring $75 < M_{ee} < 105$ GeV/c^2.

Chapter 4
Results and Conclusions

We now combine the ingredients from the previous chapters to extract the inclusive differential cross section for Z production, and provide a limit on contributions from sources beyond the Standard Model for the region of $p_T > 50$ GeV/c.

4.1 Differential Distribution

Table 4.1 gives the number of events observed as a function of p_T. Also included is the dependence of the product of the acceptance and the efficiency (ε_{tot}), and the expected number of background events (b), all as a function of p_T. (The systematic uncertainties are also shown.) The differential cross section, $d\sigma/dp_T$, is obtained by calculating the cross section in each bin of p_T, accounting for the width of each bin. That is,

$$\left(\frac{d\sigma}{dp_T}\right)_i = \frac{\sigma_i}{\Delta_i^{bin}} \qquad 4.1$$

We obtain the cross section and uncertainty in each bin using the principle of *statistical inference*, also known as *Bayesian Statistics*[59][60]. Appendix A discusses some of the features and motivation for using statistical inference to analyze data.

bin	p_T range (GeV/c)	ε_{tot}	$\delta\varepsilon_{tot}$	b	δb	data
1	0-1.5	0.35	0.011	15	4	345
2	1.5-3	0.35	0.011	31	8	794
3	3-4.5	0.346	0.011	44	11	836
4	4.5-6	0.346	0.011	53	13	710
5	6-7.5	0.333	0.011	56	14	616
6	7.5-9	0.333	0.011	55	14	482
7	9-10.5	0.339	0.011	52	13	383
8	10.5-12	0.326	0.011	47	10	341
9	12-13.5	0.316	0.010	41	10	250
10	13.5-15	0.310	0.016	36	9	206
11	15-20	0.300	0.016	80	20	512
12	20-25	0.291	0.015	47	12	287
13	25-30	0.282	0.015	29	7	174
14	30-35	0.272	0.014	18	4.5	124
15	35-40	0.263	0.014	10	2.8	104
16	40-45	0.263	0.014	7.2	1.8	50
17	45-50	0.263	0.014	4.5	1.1	42
18	50-60	0.272	0.014	4.4	1.1	61
19	60-70	0.282	0.015	1.7	0.42	40
20	70-80	0.300	0.016	0.67	0.17	15
21	80-90	0.310	0.017	0.26	0.065	10
22	90-100	0.319	0.018	0.10	0.025	8
23	100-200	0.366	0.021	0.55	0.25	15
24	200-300	0.526	0.037	0.09	0.045	2

Table 4.1 Input contributing to the cross section in each bin: ε_{tot} and $\delta\varepsilon_{tot}$ are the products of the acceptance and identification efficiencies and combined uncertainties; b and δb are the number of background events and the systematic uncertainties; data corresponds to the number of observed events.

4.1.1 Determining the Individual Cross Sections

The expected number of events μ_i in each bin i is related to the underlying cross section as follows:

$$\mu_i = L\varepsilon_i\sigma_i + b_i \qquad 4.1$$

where L is the total luminosity for the data, ε_i is the overall detection efficiency for the process, σ_i is the cross section, and b_i is the number of background events in each bin. We cannot just replace the *expected* number of events, μ_i with the *observed* number of events d_i. Our one observation is not an observation of the mean. However, the observed number of events and the expected number of events are related through a probability distribution, in our case an assumed poisson distribution:

$$P(d|L, \varepsilon, b, \sigma, I) = \frac{e^{-(L\varepsilon\sigma + b)}(L\varepsilon\sigma + b)^d}{d!} \tag{4.2}$$

where the bin index has been suppressed.

Applying Bayes' Theorem to our problem, we invert Eq.4.2 (see Section A.1.1),

$$P(\sigma, L, \varepsilon, b|d, I) = \frac{P(d|\sigma, L, \varepsilon, b, I)P(\sigma, L, \varepsilon, b|I)}{\zeta}. \tag{4.3}$$

where ζ is the normalization, $P(\sigma, L, \varepsilon, b|d, I)$ is the joint posterior probability and $P(\sigma, L, \varepsilon, b|I)$ is the joint prior probability. $P(d|\sigma, L, \varepsilon, b, I)$ is often called the *likelihood function*. Assuming that the individual parameters are logically independent, that is, the cross section does not depend on the luminosity, then Eq.4.3 can be re-written as

$$P(\sigma, L, \varepsilon, b|d, I) = \frac{P(d|\sigma, L, \varepsilon, b, I)P(\sigma|I)P(L|I)P(\varepsilon|I)P(b|I)}{\zeta} \tag{4.4}$$

Marginalizing the nuisance parameters, L, b, ε, on the right side of Eq.4.4, we get:

$$P(\sigma|d, I) = \frac{\int dbdLd\varepsilon P(d|\sigma, L, \varepsilon, b, I)P(\sigma|I)P(L|I)P(\varepsilon|I)P(b|I)}{\zeta}. \tag{4.5}$$

We take the priors for the efficiency (ε) and background (b) to be gaussian distributed with their estimated mean values and standard deviations as the parameters. In the calculation of the binned differential cross section, the uncertainty on the luminosity only changes the overall normalization of the distribution. We can therefore integrate over L using a delta-function for its distribution.

The integration in Eq.4.5 was performed using the numerical integrator MISER[62], and the data given in Table 4.1. The final values of the cross section are taken as the means of the individual probability distributions, and the uncertainties are taken as single standard deviations about the mean. Although the geometrical acceptance and identification efficiency were determined separately, we combine them into a single overall event efficiency (denoted by ε_{tot} in the table) by simply multiplying them together. The uncertainties can be added in quadrature, since the priors for both are take as gaussian. The values of the cross section are taken as the means of the individual probability distributions, and the uncertainties are taken as standard deviations. Once the cross section in each bin is determined, we obtain the differential cross section by accounting for the bin widths, as shown in Eq.4.1. The total differential distribution is normalized to the D0 inclusive cross section ($\sigma_Z \cdot B(Z \to ee) = 210\ pb$). Table 4.2 shows the values of the cross section. The differential cross section is displayed in Figure 4.3. For completeness, we have included in Figs. 4.4-4.7 the probability distributions for cross section for each of the bins.

4.2 Comparison with Predictions

In Section 1.5, we outlined the features of the theory used to describe vector boson production. In general, current descriptions using fixed-order perturbative calculations (LO or NLO) attempt to describe the high-p_T region, and use the resummation method to describe the low-p_T region. Inevitably, there is some cutoff in p_T, above which the fixed order perturbative result can be used, and below which the resummation calculation should be valid[35]. We will use the NLO calculation, implemented with a program written by Arnold and Kaufman[35]. In the low-p_T region, however, there is an ambiguity in the choice of the non-perturbative function that describes the physics at low p_T (see Section 1.5). Thus, we will compare our results to both of the low-p_T parameterizations[35][39].

4.2.1 Low p_T Region

The non-perturbative function S_{NP} has the general form (Section 1.5),

bin	p_T range (GeV/c)	σ (pb)	$\delta\sigma$ (pb)	$\frac{d\sigma}{dp_T}$ (nb)	$\delta\left(\frac{d\sigma}{dp_T}\right)$ (nb)
1	0-1.5	8.74	0.57	7.42	0.48
2	1.5-3	20.16	0.99	17.13	0.84
3	3-4.5	21.17	1.06	17.99	0.90
4	4.5-6	17.56	0.96	14.93	0.82
5	6-7.5	15.56	0.94	13.22	0.80
6	7.5-9	11.87	0.82	10.09	0.70
7	9-10.5	9.32	0.73	7.92	0.62
8	10.5-12	8.36	0.66	7.10	0.56
9	12-13.5	6.14	0.58	5.22	0.49
10	13.5-15	5.15	0.54	4.38	0.46
11	15-20	13.34	1.22	3.40	0.29
12	20-25	7.67	0.77	1.96	0.20
13	25-30	4.80	0.55	1.23	0.14
14	30-35	3.64	0.45	0.93	0.11
15	35-40	3.61	0.41	0.92	0.11
16	40-45	1.54	0.27	0.39	0.069
17	45-50	1.35	0.24	0.35	0.062
18	50-60	1.96	0.29	0.25	0.037
19	60-70	1.29	0.22	0.16	0.028
20	70-80	0.47	0.12	0.060	0.016
21	80-90	0.32	0.10	0.041	0.013
22	90-100	0.26	0.084	0.033	0.011
23	100-200	0.39	0.104	0.0005	0.0013
24	200-300	0.05	0.029	0.00064	0.00037

Table 4.2 Cross section and differential cross section for Z production and its uncertainty for each bin in p_T.

$$S_{NP}(b, Q) = h_1(b, x_A) + h_1(b, x_B) + h_2(b)\ln\left(\frac{Q}{2Q_o}\right) \quad 4.1$$

The two suggested parameterizations of this non-perturbative function are[36][39],

$$S_{NP}^{DWS}(b, Q) = g_1 b^2 + g_2 b^2 \ln\left(\frac{Q}{2Q_o}\right) \quad 4.2$$

$$S_{NP}^{LY}(b, Q) = g_1 b^2 + g_2 b^2 \ln\left(\frac{Q}{2Q_o}\right) + g_3 b \ln(100 x_A x_B) \quad 4.3$$

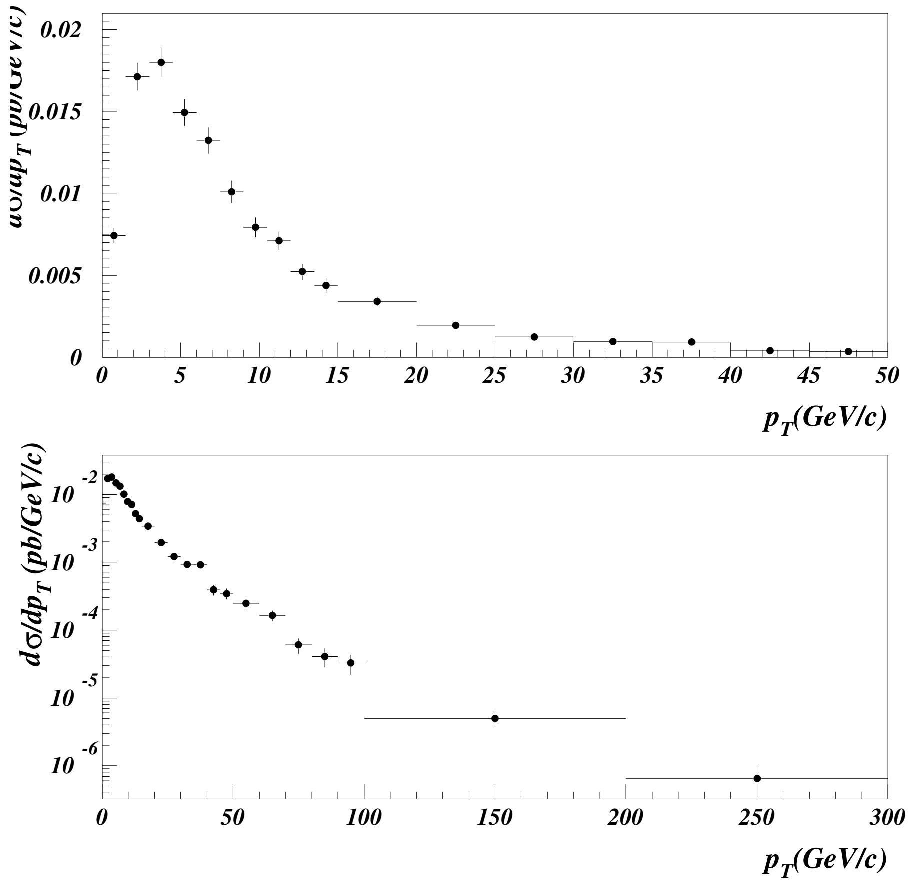

Figure 4.3 Final result for of the measurement of the differential cross section.

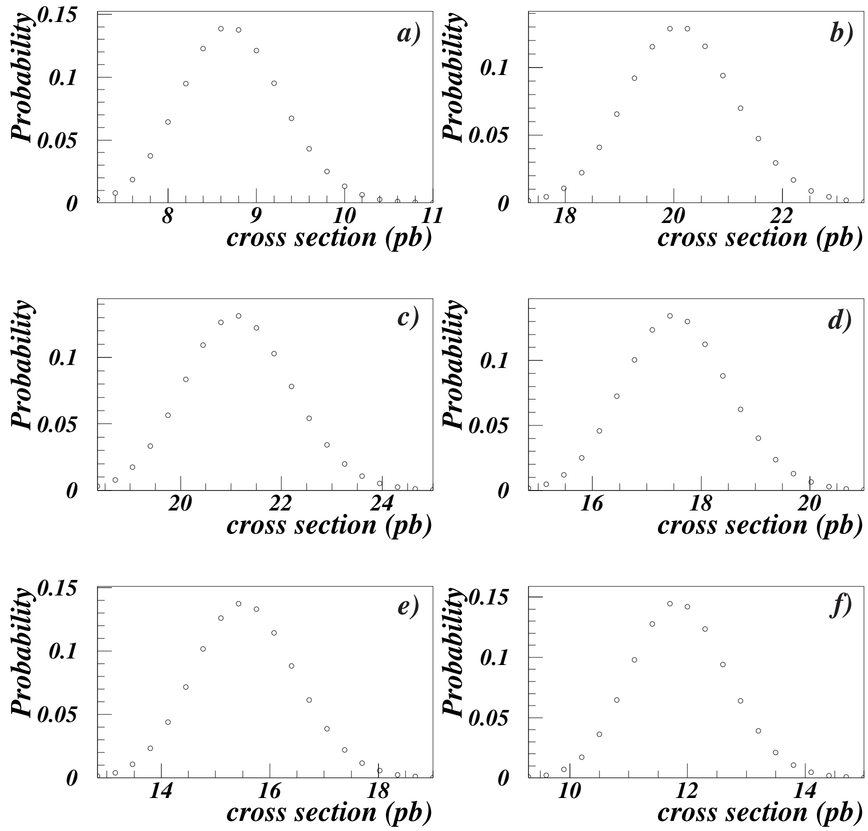

Figure 4.4 Normalized probability distributions for the cross section in individual p_T bins 1-6 given in a)-f), respectively.

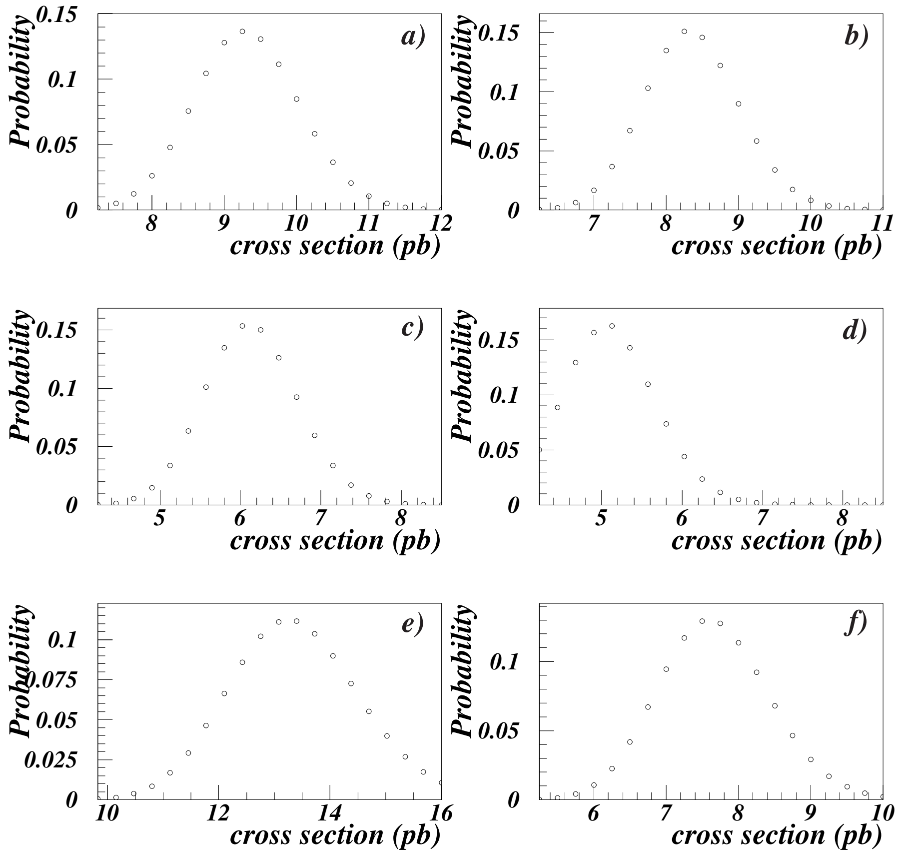

Figure 4.5 Normalized probability distributions for the cross section in individual p_T bins 7-12 given in a)-f), respectively.

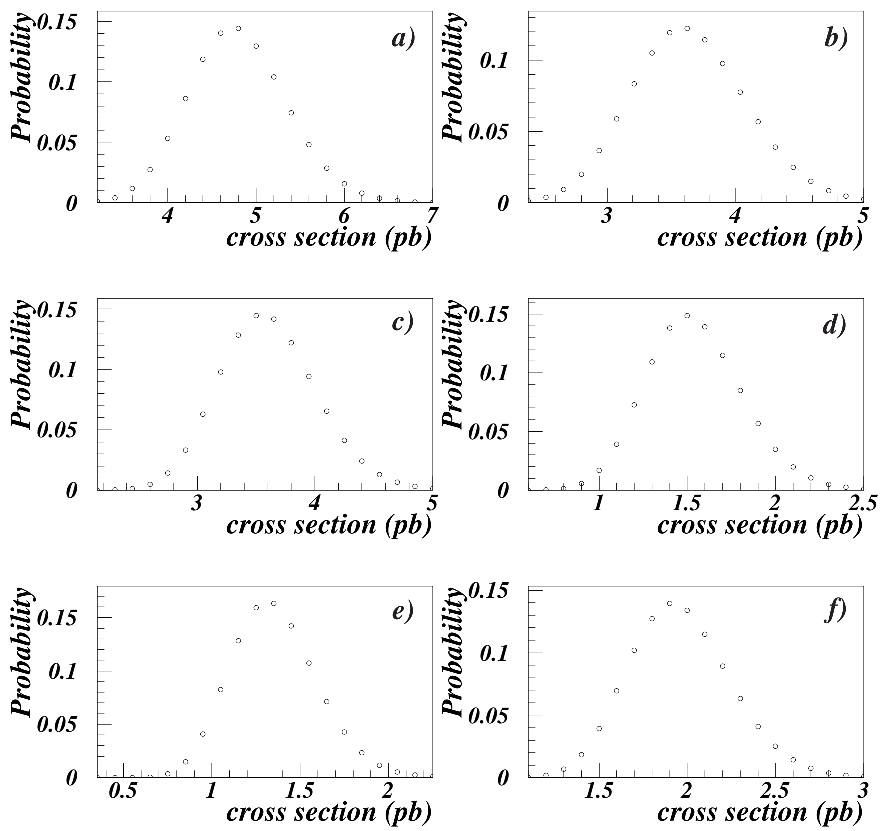

Figure 4.6 Normalized probability distributions for the cross section in individual p_T bins 13-18 given in a)-f), respectively.

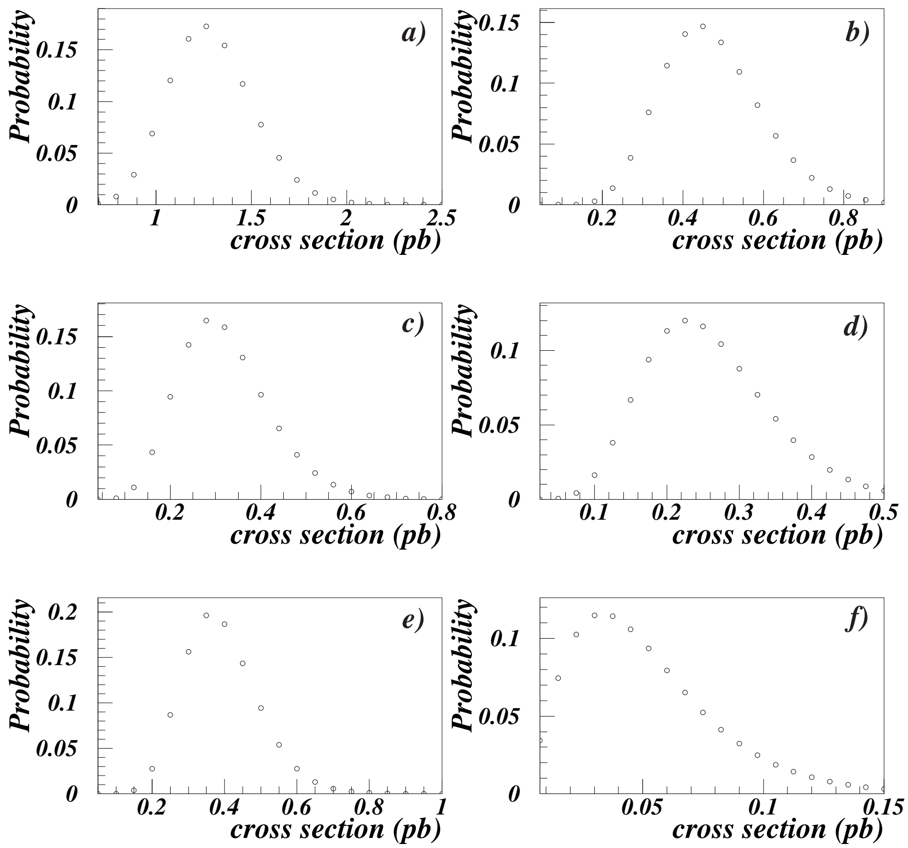

Figure 4.7 Normalized probability distributions for the cross section in individual p_T bins 19-24 given in a)-f), respectively.

where x_A and x_B are the momentum fractions of the incoming partons, b is the fourier-conjugate to the transverse momentum (impact parameter), and Q_o is an arbitrary momentum scale. The parameters, g_1, g_2, and g_3 cannot be predicted by QCD, and must therefore be measured to provide guidance to phenomenology.

The parameterization in Eq.4.2, due to Davies, Weber, and Stirling[37], was appropriated by Arnold and Kaufman in their analysis, which included the NLO fixed-order perturbative calculation[35]. Early fits to determine g_1 and g_2 were based on older data for Drell-Yan production[68]. The more recent form of Ladinsky and Yuan[39] was used in fits to high-statistics Drell-Yan data and the low-statistics Z data from the 1992-93 CDF run at the Tevatron[67]. The fit by Arnold-Kaufman, using Eq.4.2 as the non-perturbative function, yields g_1= 0.15 GeV^2 and g_2=0.4 GeV^2, where b_{max}= 0.5 GeV^{-1}, Q_o= 2 GeV, using the parton distribution functions (PDFs) of Duke and Owens[65]. Ladinsky and Yuan, using Eq.4.3, the CTEQ3M PDFs[66], b_{max}= 0.5 GeV^{-1}, and Q_o= 1.6 GeV determine the parameter values $g_1 = 0.11^{+0.04}_{-0.03}$ GeV^2, $g_2 = 0.58^{+0.1}_{-0.2}$ GeV^2, and $g_3 = -1.5^{+0.1}_{-0.1}$ GeV^{-1}.

The non-perturbative parameterization should be needed only at low p_T, and is therefore that region where we compare the two phenomenologies. Figure 4.8 compares our data to the Ladinsky-Yuan, and to the Davies, Weber, Stirling parameterizations of S_{NP} for $p_T < 50$ GeV/c. The theory has been smeared according the known D∅ detector resolution (described in Section 2.5), and normalized to the data. Figure 4.9 shows the ratio of the two predictions relative to the data. The p_T-dependence in Ladinksy-Yuan agrees better with data. The χ^2 for the comparison of data and the Ladinsky-Yuan parameterization is 18.6/16 d.o.f., while for the Davies-Weber-Stirling parameterization it is 83/16 d.o.f.

4.2.2 High-p_T Region

For the region of $p_T > 50$ GeV/c, we compare our data to the NLO perturbative calculation provided by Arnold-Kaufman. The overall normalization of the data is obtained from the entire spectrum (as before). The NLO result is normalized to the total cross section prediction for $p_T > 50$ GeV/c ($\sigma \cdot Br = 4.79$ pb). The data and the prediction are in reasonable agreement, (χ^2/d.o.f. = 9.3/6), although we have ignored any uncertainty in the theory. The theory is considered reliable to about a 10% level[35], and with the poor statistics, we observe little disagreement.

One way to quantify this comparison is to consider the observed inclusive cross section at high-p_T and the probability for sources of events other than for Z/γ* production. We do this

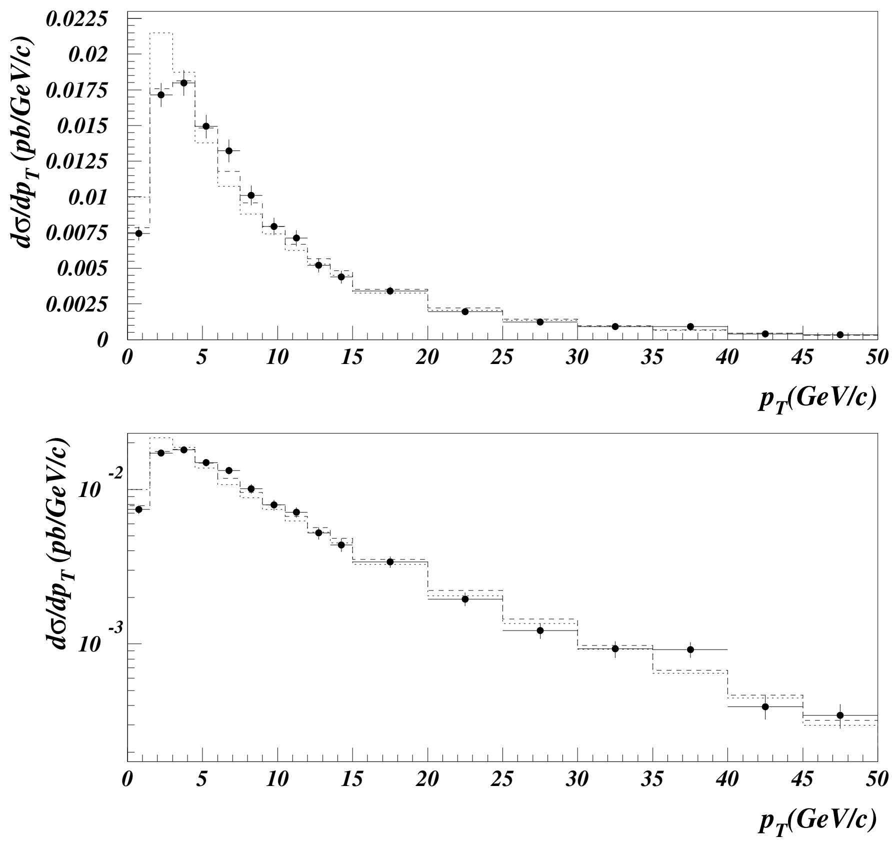

Figure 4.8 Comparison of data (solid circles) to smeared theory; Arnold-Kaufman (dotted histogram) and Ladinsky-Yuan (dashed histogram).

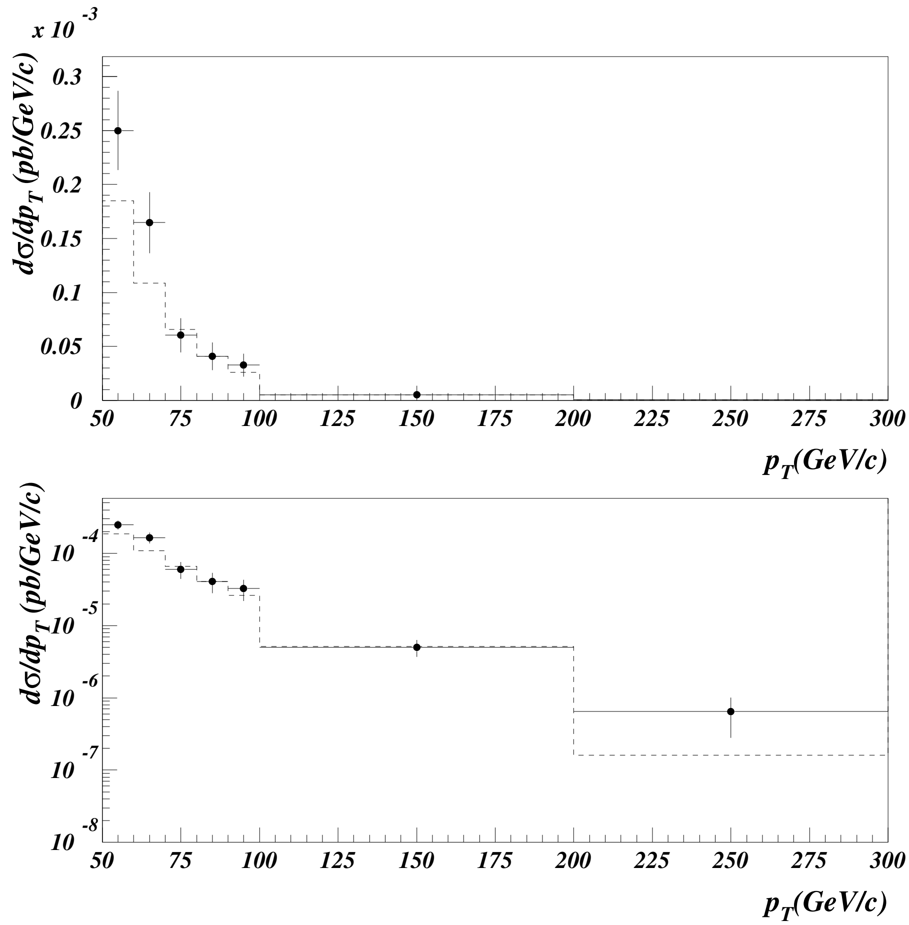

Figure 4.9 Comparison between data and the NLO prediction in the high p_T region.

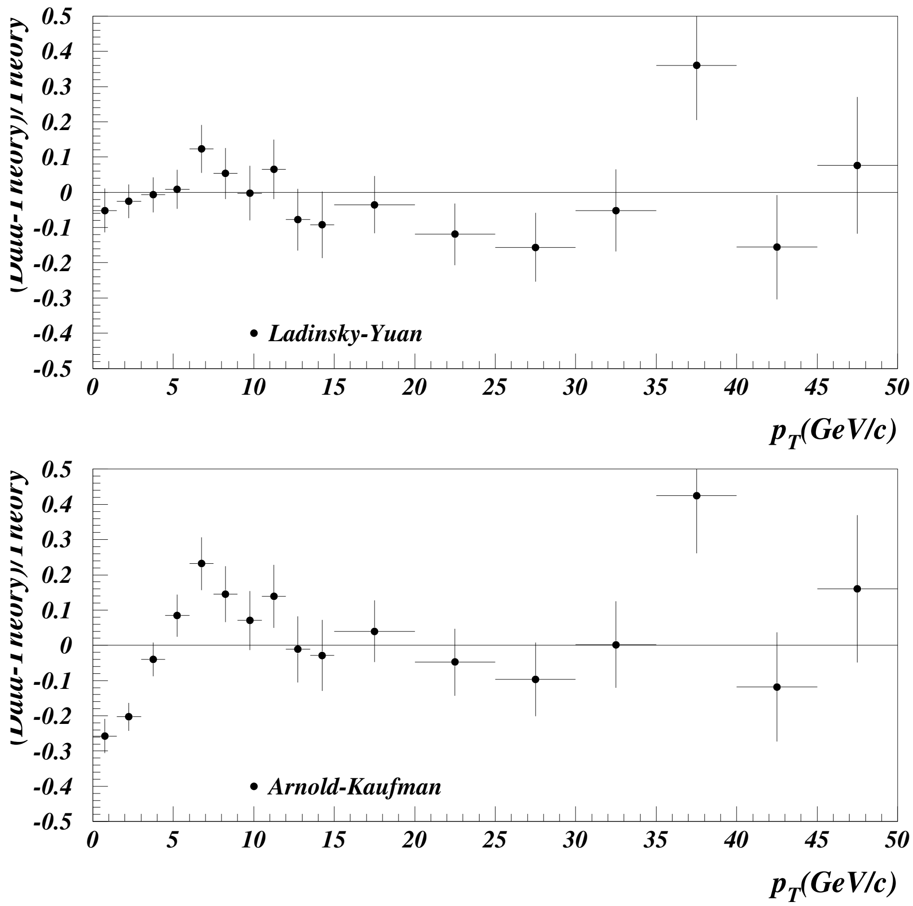

Figure 4.10 The fractional difference between the data and the predictions for two predictions in the region p_T < 50 GeV/c

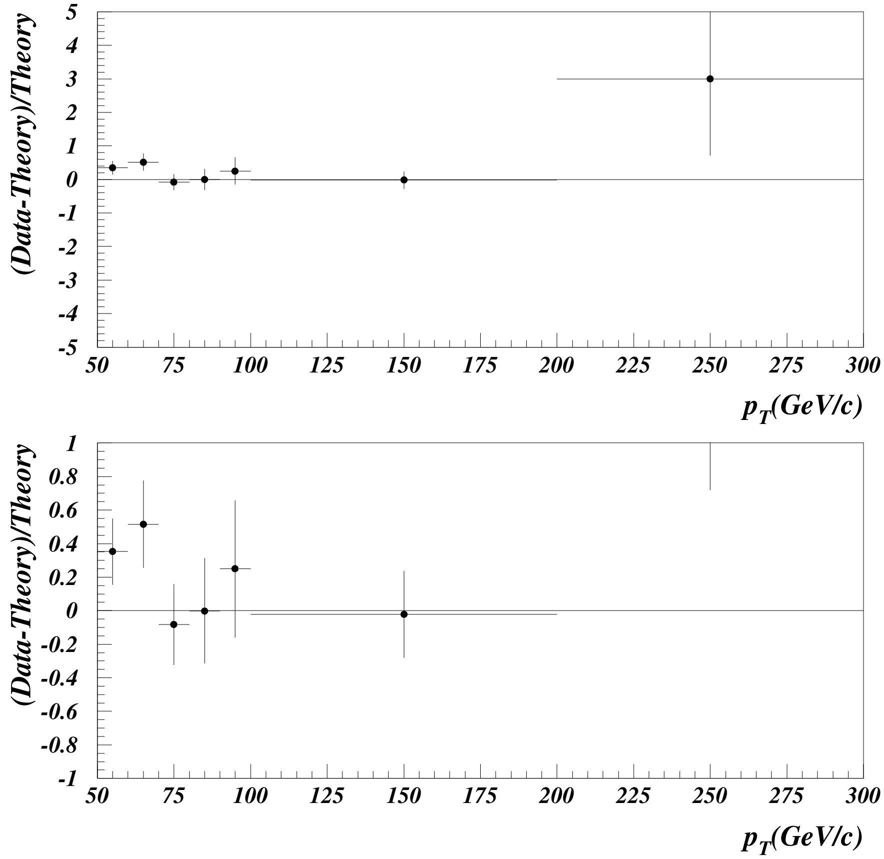

Figure 4.11 The fractional difference between the data and the prediction for the NLO prediction in the region p_T > 50 GeV/c. The only difference between the plots is the vertical scale.

using probability theory, as presented earlier in this Chapter and in the Appendix. For measuring the inclusive cross section, we use Eq.4.1, considering all the events with $p_T > 50$ GeV/c. In order to consider the possibility of other sources, we modify Eq.4.1 as follows:

$$\mu = L\varepsilon(\sigma'_{Z/\gamma*} + \sigma'_{new}) + b \qquad 4.4$$

where L is the total integrated luminosity, ε is the dielectron identification efficiency for events with $p_T > 50$ GeV/c, b is the multijet background, $\sigma'_{Z/\gamma*}$ is the acceptance times cross section for Z/γ* production, and σ'_{new} is the acceptance times cross section for "anything else" that goes to dielectrons. Applying Bayes' Theorem, and following the steps used previously, we obtain the posterior probability for σ'_{new},

$$P(\sigma'_{new}|D, I) = \qquad 4.5$$
$$\int dL\,db\,d\varepsilon\,d\sigma'_{Z/\gamma*}P(D|L, \varepsilon, b, \sigma'_{Z/\gamma*}, \sigma'_{new})P(L|I)P(\varepsilon|I)P(b|I)(\sigma'_{Z/\gamma*}|I)P(\sigma'_{new}|I)$$

In principle, we could include priors for the acceptances explicitly, and therefore account for any systematic uncertainties associated with them. However, we know that $\delta a_{Z/\gamma*}/a_{Z/\gamma*} <$ 1% [29] which is much less than either $\delta\varepsilon/\varepsilon$ or $\delta L/L$ (both are about 5%, as shown earlier), so including the uncertainty on the acceptance for the known cross section does not affect the result. Including an explicit acceptance term for the σ'_{new} term would be useful if we were searching for a specific process, but if we are looking for anything, then setting $a_{new} = 1$ is appropriate. For $\sigma'_{Z/\gamma*}$ we take the absolute NLO prediction for the Z/γ* production cross section from the Arnold and Reno calculation as implemented in the program written by Arnold and Kaufman[35]. Including the acceptance (0.31), we have $\sigma'_{Z/\gamma*} = 1.5\pm0.15\,pb$, where we have taken the uncertainty in the calculation to be 10%.

The background in both cases is taken from the multijet analysis shown earlier. The efficiency, ε, is taken to be the absolutely normalized event identification efficiency of 0.73, as described before (Chapter 3). We assume a 5% uncertainty on the identification efficiency.

The resulting probability distribution of the inclusive cross section for $p_T > 50$ GeV/c is shown in Figure 4.12. No significant excess is evident over the NLO calculation of 4.79 *pb*. The probability distribution for sources other than Z/γ* production is shown in Fig. 4.12, and, as expected, the uncertainty is so large as to preclude a satisfactory lower limit on other physics.

4.3 Conclusion

We have measured the transverse momentum distribution for dielectron production in $p\bar{p}$ collision with the DØ detector at Fermilab. We have compared the measured distribution to resolution-smeared theoretical distribution and found that the Ladinsky-Yuan QCD-inspired parameterization of the low-p_T physics matches the data well. We have also found no evidence for dielectron production beyond expected contributions from processes that can be described by the Standard Model for the region of $p_T > 50$ GeV/c.

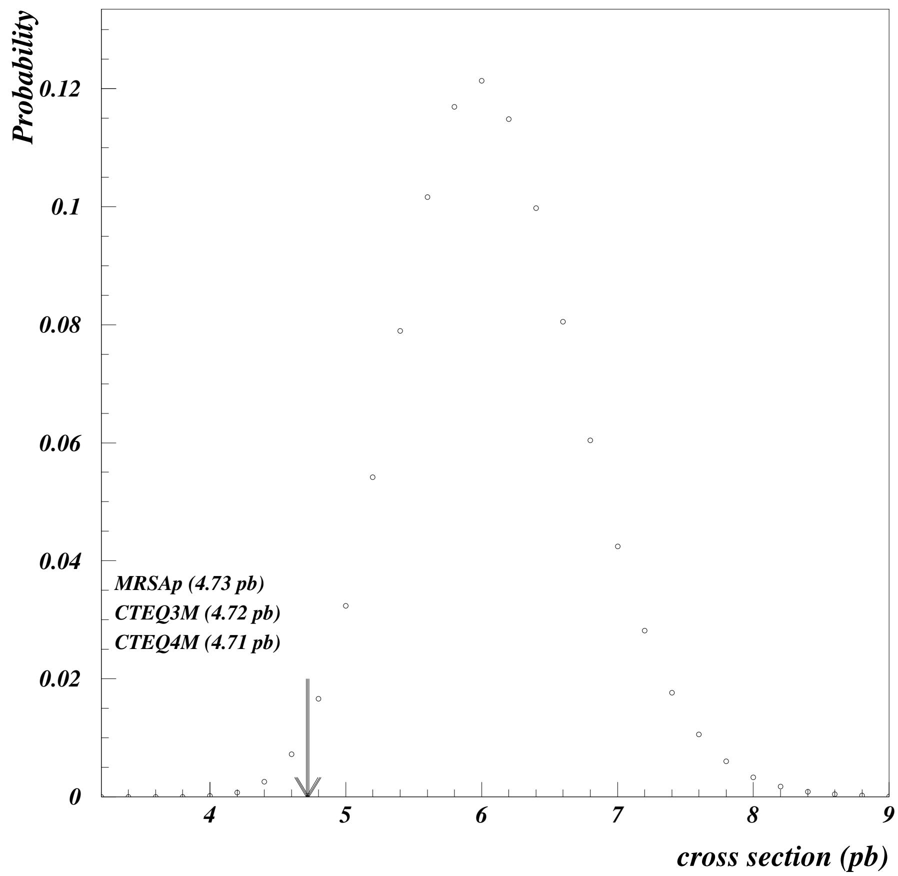

Figure 4.12 Normalized probability distribution of the observed cross section for dielectrons with $p_T > 50$ GeV/c. The arrows indicate the NLO results using the indicated PDFs[35].

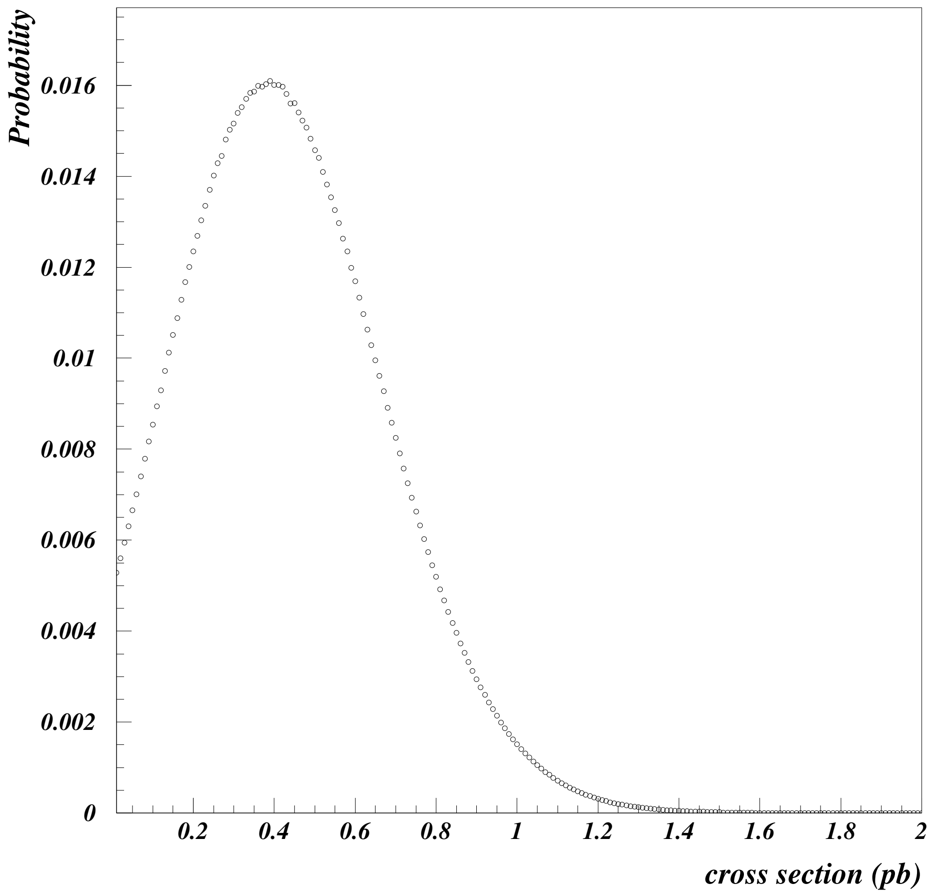

Figure 4.13 Normalized probability distribution of the cross section due to sources other than Z/γ^* for dielectrons with $p_T > 50$ GeV/c. We have assumed 100% geometrical acceptance for such sources/

References

[1] A. Einstein, *The Principle of Relativity,* Dover Publications, 1924.

[2] P.A.M.Dirac, *The Principles of Quantum Mechanics,* 4th ed., Clarendon Press, Oxford, 1958.

[3] J.J.Sakuri, *Modern Quantum Mechanics*, Addison-Wesley Publishing Company, 1985.

[4] M. Jammer, *The Philosophy of Quantum Mechanics*, John-Wiley and Sons, 1974.

[5] I.J.R. Aitchison and A.J.G. Hey, *Gauge Theories in Particle Physics*, Adam Hilger, 1989.

[6] C. Quigg, *Gauge Theories of the Strong, Weak, and Electromagnetic Forces*, Addison-Wesley Publishing Company, 1983.

[7] G. ‘t Hooft, “Gauge Theories of the Forces between Elementary Particles”, *Sci Am*, **104**, (1980).

[8] K. Gottfried and V. Weisskopf, *Concepts of Particle Physics*, vols. I and II, Oxford University Press, New York,1986.

[9] The E852 Collaboration, Phys. Rev. Lett. **79,** 1630 (1997).

[10] Particle Data Group, “Review of Particle Properties,” *Phys. Rev. D***54**, (1996).

[11] F. Halzen and A. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*, John Wiley and Sons, 1984.

[12] R.N. Cahn and G. Goldhaber, *The Experimental Foundations of Particle Physics*, Cambridge University Press, 1989.

[13] T.D. Lee and C.N. Yang, *Phys. Rev.* **104**, 254 (1956).

[14] Feynman and Gell-Mann, *Phys. Rev.* **109**, 193 (1957).

[15] Marshak and Sudarshan, *Phys. Rev.* **109**, 1860 (1957).

[16] S.L. Glashow, *Nucl. Phys.* **22**, 579 (1968).

[17] S. Weinberg, *Phys. Rev. Lett.* **19**, 1264 (1967).

[18] UA1 Collaboration, *Phys. Lett. B***122**, 103 (1982).

[19] UA2 Collaboration, *Phys. Lett. B***122**, 476 (1983).

[20] UA1 Collaboration, *Phys. Lett. B***126**, 398 (1983).

[21] UA2 Collaboration, *Phys. Lett. B***129**, 130 (1983).

[22] LEP Electroweak Working Group, as reported by E.Gross at Rencontres de Moriond, March 1997.

[23] G. Kane, *Modern Elementary Particle Physics*, Addison-Wesley Publishing Company, 1983.

[24] M. Gell-Mann, *Phys. Lett.*, **8** 214 (1964).

[25] G. Zwieg, CERN Report 8182/Th. 401 (1964).

[26] D∅ Collaboration, *Phys. Rev. Lett.* **74**, 2632 (1995).

[27] CDF Collaboration, *Phys. Rev. Lett.* **74**, 2626 (1995).

[28] J. Thompson, "Introduction to Colliding Beams at Fermilab", FERMILAB-TM-1909.

[29] J. Tarazi, J. Thompson, "Measurement of the $W \to e\upsilon$ and $Z \to ee$ Inclusive Cross Sections using Run 1b/1c Data", D∅ Note (1997).

[30] M. Narain, "Electron Identification in the D∅ Detector" in *Proceedings of the American Physical Society Division of Particles and Fields Conference*, Fermilab (1992).

[31] R. Raja *et al.*, "H-matrix analysis for top → lepton + jets", D∅ Note 1006 (1990).

[32] J. Smith, W.L. VanNeerveen, J.A.M. Vermaseren, *Phys. Rev. Lett.***50,** 22 (1983).

[33] I. Adam PhD thesis, *Measurement of the W Boson Mass with the D0 Detector using the Electron E_T Spectrum*, unpublished (1997).

[34] J.C. Collins, D.E. Soper, G. Sterman, *Nucl. Phys. B***250**, 199 (1985).

[35] P.B. Arnold and R.P. Kaufman, *Nucl. Phys. B***349**, 381 (1991).

[36] C.T.H Davies, B.R. Weber, W.J. Stirling, *Nucl. Phys. B***256**, 413 (1985).

[37] C.T.H. Davies, W.J. Stirling, *Nucl. Phys. B***244**, 337 (1984).

[38] G. Altarelli, R.K. Ellis, M. Greco, and G. Martinelli, *Nucl. Phys. B***246**, 12 (1984).

[39] G.A. Ladinsky and C.-P. Yuan, "The Nonperturbative Regime in QCD Resummation for Gauge Boson Production at Hadron Colliders", *Phys. Rev. D***50**, 4239 (1994).

[40] C. Balazs and C.P. Yuan, *Phys. Rev.* D**56**, 5558 (1997).

[41] G. P. Korchemsky, G. Sterman, *Nucl. Phys. B***437**, 415 (1995).

[42] D∅ Collaboration, "The D∅ Detector", *Nucl. Instr. Meth. A***338**, 185 (1994).

[43] T. Ferbel (ed), *Experimental Techniques in High Energy Physics*, Addison-Wesley Publishing Company (1987).

[44] S. Snyder, PhD thesis, SUNY Stony Brook, *Measurement of the Top Quark at D∅*, unpublished (1995).

[45] E. Flattum, PhD thesis, Michigan State University, *A Measurement of the W Boson Mass in $p\bar{p}$ Collisions at $\sqrt{s}$ =1.8 TeV*, unpublished (1997).

[46] I. Adam, E. Flattum, and U. Heintz, "W Boson Mass Measurement using Run 1B Data", D∅Note 3198 (1997).

[47] U. Heintz, D∅Note in preparation.

[48] R. Brun *et al.*,GEANT Program Manual, CERN Program Library (unpublished).

[49] N. Graf, private communication.

[50] R. Hirosky, private communication.,

[51] S. Jerger, PhD thesis, Michigan State University, *Inclusive Direct Photon Production in the Central and Forward Rapidity Regions in Proton-Antiproton Collisions at* $\sqrt{s}$ *= 1800 GeV,* unpublished (1997).

[52] H.U. Bengtsson and T. Sjorstrand, *Comput. Phys. Commun*. **46**, 43 (1987).

[53] MINUIT, Application Software Group, Computing and Networks Division, CERN, Geneva, Switzerland.

[54] Z.Jiang, PhD thesis, SUNY Stony Brook, *Measurement of the Z boson Transverse Momentum Distribution in* $p\bar{p}$ *Collisions at* $\sqrt{s}$ *= 1.8 TeV,* unpublished (1995).

[55] CDF collaboration, *Phys. Rev. Lett.* **78**, 4536 (1997).

[56] D∅ Collaboration, *Phys. Rev. Lett.* **78**, 3634 (1997).

[57] E.Eitchen, I. Hinchliffe, K. Lane, C. Quigg, *Rev Mod Phys* **56**, 579 (1984).

[58] G. Marchesini *et al.*, Comput. Phys. Commun. **67**, 465 (1992).

[59] E.T.Jaynes, *Probability Theory — The Logic of Science*, in preparation. Copies of the manuscript are available from http://bayes.wustl.edu.

[60] T. J. Loredo, “From Laplace to Supernova SN1987A: Bayesian Inference in Astrophysics,” in *Maximum Entropy and Bayesian Methods*, P. Fougere *ed.*, Kluwer Academic Publishers, Dordrecht the Netherlands (1990).

[61] R.T. Cox, *Am. Jour. Phys* **14**, 1 (1946).

[62] W.H. Press *et al.*, *Numerical Recipes in C*, (1996); P.R. Bevington, *Data Reduction and Error Analysis for the Physical Sciences*, McGraw-Hill Book Company, New York (1969).

[63] G.L.Bretthorst, “An Introduction to Parameter Estimation Using Bayesian Probability Theory,” in *Maximum Entropy and Bayesian Methods*, P. Fougere *ed.*, Kluwer Academic Publishers, Dordrecht the Netherlands (1990).

[64] W.T. Eadie, D. Dijard, F.E. James, M. Roos, and B. Sadoulet, *Statistical Methods in Experimental Physics*, North-Holland Publishing Company, Amsterdam (1971).

[65] P.B. Arnold and R.P. Kaufman, *Nucl. Phys. B***349**, 381 (1991).

[66] CTEQ Collaboration, *Phys. Rev. D***55**, 1280 (1997).

[67] CDF Collaboration, *Phys. Rev. Lett.* **67**, 2937 (1991).

[68] D.W. Duke and J.F. Owens, *Phys. Rev.* **D30,** 49 (1984).

[69] W.H. Press *et al.*, *Numerical Recipes in C*, (1996).

[70] MACSYMA 2.0 User’s Guide, MACSYMA, Inc. unpublished (1995).

[71] S. Wolfram, *The Mathematica Book*, *3rd ed*, Cambridge University Press, 1996.

[72] Tchakov and Manankova,” *MILXy Way: How Much Better than VEGAS can One Integrate Many Dimensions?”*, FERMILAB-CONF-95-213-T (1995).

Appendix A Statistical Inference and an Unsmearing Method

This first section of this appendix presents the motivation and method for using the principle of statistical inference to extract the differential cross section from our data that was shown in Chapter 4. In the second section, we present a new method for accounting for effects of detector resolution on the data.

A.1 Statistical Inference

As motivation for the direct use of probability theory in order to extract a measurement from a data set, consider the standard quantity quoted as the uncertainty of a measurement of some quantity, x — the standard deviation $\bar{\sigma} = \sqrt{<x>^2 - <x^2>}$. The first and second moments of the probability distribution can be written explicitly as:

$$<x> = \int xP(x|D, I)dx \tag{A.1}$$

$$<x^2> = \int x^2P(x|D, I)dx \tag{A.2}$$

where $P(x|D, I)$ is the probability of x given the data, D, and the general conditionals, I. The inclusion of the general conditional, I, follows the convention of Jaynes[59], Loredo[60], and Bretthorst[63]. It represents the general assumptions and conditions that are involved in obtaining the data set, and reminds us that *all* probabilities are conditional on prior information.

The fundamental encoding of information regarding the measurement of x is contained in $P(x|D, I)$—the probability distribution for x given the data obtained. That is, $\int_a^b P(x|D, I)dx$ defines the degree of plausibility assigned to the statement "given our data, the true value of x is between a and b". If $P(x|D, I)$ is gaussian, then the mean, $<x>$, describes the most probable value of x and the region $[<x> - \bar{\sigma}, <x> + \bar{\sigma}]$ defines the minimum interval of 68% certainty. Of course, when $P(x|D, I)$ is not symmetric, then $<x>$ may not be the most probable value, and if it is not gaussian, then the region $[x - \bar{\sigma}, x + \bar{\sigma}]$ will usually have a different probability content.

Quantities such as the standard deviation are very useful for characterizing our state of knowledge, but merely reflect the complete measure of information, which is contained in the probability distribution. Let us consider our measurement of the cross section. When we say that the probability distribution for the cross section is $P(\sigma|D, I)$, we mean that we do not know the true value precisely, but, to our best knowledge, we believe that it lies between α and β with probability $\int_\alpha^\beta P(\sigma|D, I)dx$. We assign the probability P=0 only when we know with certainty that some range is not possible, and P=1 when we know with certainty that the value lies in the specified range. It should be recognized that we are not saying that the true value of the cross section is "distributed" in some given way, only that our knowledge of the true value is imperfect, and our best estimate can be described by some probability distribution. This probability distribution represents our state of knowledge, not the state of nature.

A.1.1 Bayes' Theorem

Of course, under normal circumstances, we do not know $P(x|D, I)$—it is what we are trying to measure. However, we can construct a function related to what we do know,

$P(D|x, \phi, I)$, where x represents some quantity we are interested in, *e.g.*, the cross section, and ϕ represents some set of the parameters that we are not interested, *e.g.*, the background and efficiency associated with the process we are studying. That is, $P(D|x, \phi, I)$ embodies a model of how the inherent processes are altered by our detector, *e.g.*, how the cross section is related to its true value.

In order to obtain $P(x|D, I)$, we have to invert the probability $P(D|x, \phi, I)$. The only way to do this is with the help of Bayes' Theorem:

$$P(x, \phi|D, I) = \frac{P(D|x, \phi, I)P(x, \phi|I)}{P(D|I)} \qquad \text{A.3}$$

The probability $P(x, \phi|I)$ is called the *joint prior probability* for the parameters x and ϕ, and is generically called a "prior"—a quantity that depends only upon the assumptions that we bring to the problem. The probability $P(x, \phi|D, I)$ is called the *joint posterior probability* for the parameters x and ϕ. ("Prior" represents our knowledge before, and "posterior" after, we examine the data.) $P(D|I)$ normalizes the posterior probability ($\int\int P(x, \phi|D, I)dxd\phi \equiv 1$) and we will write it as ζ. The prior probabilities for x and ϕ can often be logically independent, and then we can rewrite the joint prior probability as:

$$P(x, \phi|I) = P(x|\phi, I) = P(x|I)P(\phi|I). \qquad \text{A.4}$$

Even if x and ϕ are not logically independent, we can include any correlations by construction of the appropriate joint prior. When there are no correlations,

we can rewrite Eq.A.3 as

$$P(x, \phi|D, I) = \frac{P(D|x, \phi, I)P(x|I)P(\phi|I)}{\zeta}. \qquad \text{A.5}$$

The prior $P(\phi|I)$ contains the prior knowledge of the parameters we are not interested in, *e.g.*, our estimates of the efficiency and background, and their respective uncertainties. Through such priors, we can include the systematic uncertainties on our measurement. For instance, our knowledge of the detection efficiency might be described sufficiently by a gaussian with an average efficiency, $<\varepsilon>$, and a standard deviation, $\bar{\sigma}_{\varepsilon}$. $P(x|I)$ represents our prior knowledge of the parameter we are measuring. Often, we take this to be some *least-informative* prior. For example, when measuring the cross section for *Z* production, we can use a constant prior for the range between zero and the total $p\bar{p}$ cross section. However, at times there may be good reasons for not using a least-informative prior. For instance, we may know, that the quantity *x* has an exponential form or there may be a previous experiment that constrained *x*, and we want to include that information. The priors $P(\phi|I)$ and $P(x|I)$ differ to the extent that our knowledge of the quantities *x* and ϕ differs. We usually know a great deal about the efficiency, but not so much about the cross section. Our level of knowledge is incorporated explicitly into the measurement via these priors. As such, no explicit distinction arises between statistical and systematic uncertainties. In fact, the distinction is somewhat blurred—both are merely considered as constraints upon our knowledge, one reflecting the fact that we have limited data and the other reflecting the fact that we have limited understanding of the apparatus we are using for the measurement.

The posterior probability given in Eq.A.5 is not what we want. It still contains some dependence on parameters we are not interested in—the “nuisance parameters” ϕ. We purge these by integrating over ϕ. This process is called marginalization, and it yields our final target, the posterior probability density for *x*,

$$P(x|D, I) = \int d\phi P(x, \phi|D, I) . \tag{A.6}$$

From the posterior probability, $P(x|D, I)$, we can calculate the means and standard deviations, or most probable values and confidence intervals, or upper and lower limits; in fact, we can calculate whichever descriptive statistic we choose to employ to describe *x*.

The preceding discussion is completely general. The rules of probability and Bayes’ Theorem constitute a quantification for reasoning in the face uncertainty — logical induction. However, considering probability as a quantification of a state of knowledge, is not the common approach taken in past texts on statistical analysis[64]. Probability is usually defined in terms of a frequency distribution for a quantity, when taken over a long time period. This distinction between probability as a “statement of knowledge” and probability as a “frequency distribution” is non-trivial, and it is reflected in a long history of vehement debate. We will not enlarge on those very interesting arguments, but refer the reader to Refs. [59], [60], and [61] regarding the subtleties.

It is important to recognize that the above use of probability theory is completely logical and well-defined. Interpreting the probability $P(A|BC)$ as a statement of our knowledge of A, given that B and C are true, the common rules for addition and multiplication of probabilities follow immediately from three *desiderata*:

(i) All probabilities are represented by real numbers.

(ii) The results of calculations with these probabilities are consistent, *i.e.* one cannot simultaneously obtain that $P(A|B) = 1$ and $P(A|B) = 0$.

(iii) In the limit of certain truth ($P(A|B) = 1$) or certain falsehood ($P(A|B) = 0$), the results of inference reduce to deductive logic.

The addition and multiplication rules that follow are[61]:

$$P(A|B) + P(\bar{A}|B) = 1 \qquad \text{A.7}$$

$$P(AB|C) = P(A|BC)P(B|C) \qquad \text{A.8}$$

where $\bar{A}$ is the compliment of A.

A trivial consequence of Eq.A.8, is Bayes’ Theorem:

$$P(A|BC) = \frac{P(B|AC)P(A|C)}{P(B|C)}. \qquad \text{A.9}$$

Thus, the rules for reasoning in the face uncertainty, for inverting logical relationships, can be derived from merely three, well-motivated, criteria for *how* we wish to reason!

A.1.2 Comments on Priors

In any actual problem, we are most likely to find that we have estimates for all the parameters x and ϕ, and that all have non-negligible uncertainties. As long as the best estimate of each of the parameters is significantly far away from any unphysical region (for example, the efficiencies are not near 0 or 1), and as long as our knowledge of one parameter is not influenced by knowledge of another (that is, they are independent), then the appropriate distribution to represent out state of knowledge of the parameter (the prior distribution) is the gaussian distribution[59]. For the mean of the gaussian, we use our best estimate of the parameter; for the standard deviation, we use our estimate of the uncertainty. If the parameter is close to an unphysical region, then the appropriate prior distribution can be found using the technique of Maximum Entropy[59]. It is, in fact, this technique that defines the gaussian as the appropriate prior distribution when the above conditions are satisfied. If is not claimed that the quantity is distributed as a gaussian, but only that it is the best way to represent out state of knowledge of that quantity.

Determining the appropriate prior for a parameter that we wish to measure is a more delicate matter. In particular, we often wish to use some sort of least-informative prior, but there is ambiguity in defining what is meant by “least-informative”.[1] A reasonable (but not only) choice is, as noted previously, a prior uniformly distributed in some range (a “flat” prior). If one is interested in fitting a given functional form to some set of data points, then one would use that functional form directly in the model, and would specify the appropriate priors for the fit parameters rather than the prior values for cross section. If there were some previous knowledge of the data that we wanted to incorporate into the fit (this would happen if we wished to combine results from two different experiments), and if the previous result was

---

[1] This, in part historical, originates from the attempt to generate a more “objective” criterion for analysis using frequency distributions. These issues, as well as the issue of “objective” probability, are discussed in Ref. [59].

independent of our other (nuisance) parameters, we would merely incorporate this previous knowledge in an appropriate prior. That is, rather than take a uniform prior for the result, we could use the posterior probability from the previous analysis as our prior. (This is the way Bayes' Theorem provides the means for adding two data sets.) In our analysis, we use gaussian distributions for the nuisance parameters and a flat distribution, bounded below by 0, for the prior for the cross section.

A.2 Unsmearing Method

One of the original goals of this analysis was to make the measurement "detector independent", that is, unfold the smearing and resolution effects in the data and present a final differential cross section that can be compared to any (non-smeared) theoretical prediction. This appendix outlines a new method devised to accomplish this task, and shows a simple example confirming that the method is viable. The technical difficulties encountered in implementing our solution, prevented it from being applied in the current analysis. We discuss some of the difficulties in Section A.2.6.

The goals of the unfolding are the following. The solution should:

- account for all statistical and systematic uncertainties in a coherent and quantitative manner.
- be valid for large and small data samples and be insensitive to the presence of any known kinematic boundaries, *e.g.*, $p_T = 0$.
- provide the ability to account for prior information about the detector and the measurement, *e.g.*, all physical cross sections must be positive quantities.
- provide uncertainties that have true probability content.

Our procedure is a straightforward extension of the statistical inference method used to extract the individual cross sections in Chapter 4.

A.2.1 Introduction

When making a measurement of a distribution, one is always confronted with the fact that output of one's detector is not a perfect reflection of the process being measured. Even when offsets and biases are accounted for in corrections to the final measurement, the finite resolution of the detector will cause the measured quantity to be smeared with respect to the true quantity. In a binned measurement of the cross section, this amounts to there being a probability that an event produced in bin i will be observed in bin j. Usually, we have information regarding how the experimental apparatus affects the quality of what we are measuring—we know the smearing matrix for the detector. We then want to use this knowledge to obtain the underlying distribution of the data before it was compromised by the measurement process—we want to extract the "true" distribution from the measured one. This is often called "deconvoluting", "unsmearing", or "unfolding" the data, and is done in analyses of measurements of everything from colliding particle experiments to astronomical observations. Often, the solution is cast as a functional or matrix inversion problem. That is, assuming a linear relationship between the true and measured quantities, we can relate the two as follows:

$$d = St + b \qquad \text{A.10}$$

where d stands for our data, t for the true quantity, and S describes the smearing from bin to bin, and b is the contamination (background) in d. Equation A.10 can be cast in either a discrete or continuous form. The result of an experiment provides d, and understanding of the experimental environment yields S and b. The solution to Eq.A.10, is often obtained by applying the inverse of the smearing matrix:

$$t = S^{-1}(d-b)\,. \qquad \text{A.11}$$

The statistical uncertainties are propagated through using standard methods[64], and usually added in quadrature. The systematic uncertainties are usually represented by a band apart from the measurement, because the standard methods do not allow for systematic and statistical errors to be combined in a consistent manner.

There are two well-known problems with this solution to the unfolding:

- If the smearing matrix is singular, or very nearly singular, S^{-1} does not exist or is poorly determined, and there is no unique result for t. This is easily understood—the smearing matrix blurs the information about the true distribution, which leads to whole class of solutions for t.
- The presence of background (or noise) in the data effectively changes the smearing matrix. This is because the term $St = \varepsilon$, where ε is small compared to b, can be arbitrarily added to the equation. As a result, small changes in data lead to large changes in t, and one observes an instability in the determination of t.

There are additional objections:

- The relationship in Equation A.11 does not account for the fact that our data represent a single and unique measurement of the underlying process. In particular, for a binned differential cross section, the number of events observed in a given bin represent one sampling of a poisson process, and not an observation of the mean. Even if we have a perfect detector ($S \equiv 1$) with no backgrounds, S still connects only the cross section in a given bin to the expected number of events in that bin. Our single observation is only one case in the process. This information is never included when solving the problem as a matrix or functional inverse.
- We often have valuable prior information that we want to include in our measurement, but cannot in standard methods. For instance, in a measurement of a cross section, we know that the true value cannot be negative. We also know many details about the allowed values of the elements of S and b, which we might also want to include in our final result. Essentially, we want to include our systematic uncertainty about the measurement before reaching a final conclusion.
- In general, standard analysis methods are *ad hoc* and lack consistent mathematical foundation, and can lead to non-sensible results and undesirable consequences[59].

While other methods can be utilized to extract t, and circumvent some of the above objections (designed typically to address the singularity and instability issues), they are *ad hoc*, applicable only in specific cases, and do not address all the other general objections[60]. The following method has the advantage of being well-founded in probability theory, and being applicable to all cases involving the unfolding of binned data.

A.2.2 Differential Cross Section

Following the discussion in Section A.1, we can summarize the steps used to obtain an unfolded differential cross section:

- Devise a model for observing the data
- Generate $P(D|x, \phi, I)$
- Invert $P(D|x, \phi, I)$ using Bayes' Theorem to obtain $P(x, \phi|D, I)$
- Marginalize over the uninteresting quantities ϕ
- Generate descriptive statistics such as means and standard deviations using $P(x|D, I)$

The following sections go through this procedure in greater detail.

A.2.3 The Model

Assuming a linear dependence of the expected number of events μ_k, in each bin k, to the cross section σ_j for each bin, we can write:

$$\mu_k = L\varepsilon_k \sum_{j=1}^{N} S_{kj}\sigma_j + b_k \qquad \text{A.12}$$

where L is the total luminosity, ε_k is the efficiency in bin k, σ_j is the cross section in bin j, b_k is the background in bin k, and S_{kj} describes the smearing from bin j to bin k. This model is the same as the one used in Section 4.1.1, except that here the observed number of events in each bin k has contributions from many of the bins that comprise the true distribution.

The probability of observing d_k events in bin k is given by the poisson distribution with mean μ_k:

$$P(d_k|\mu_k, I) = \frac{e^{-\mu_k}\mu_k^{d_k}}{d_k!} \quad , \qquad \text{A.13}$$

and the probability of observing the entire spectrum where $d = \{d_1 \ldots d_N\}$, given $\mu = \{\mu_1 \ldots \mu_N\}$, is

$$P(d|\mu, I) = \prod_{k=1}^{N} P(d_k|\mu_k, I)\,. \qquad \text{A.14}$$

Substituting for μ_k in Eq.A.14 using Eq.A.12, the probability of observing the data set is:

$$P(d|\sigma, S, \varepsilon, L, b) = \prod_{k=1}^{N} \frac{e^{-L\varepsilon_k \sum_{j=1}^{N} S_{kj}\sigma_j + b_k}\left(L\varepsilon_k \sum_{j=1}^{N} S_{kj}\sigma_j + b_k\right)^{d_k}}{d_k!} \qquad \text{A.15}$$

A.2.4 Applying Bayes' Theorem

In order to obtain the final probability distribution for the cross section, we must invert Eq.A.15 using Bayes' Theorem:

$$P(\sigma, S, \varepsilon, L, b|d, I) = \frac{P(d|\sigma, S, \varepsilon, L, b, I)P(L|I)P(S|I)P(\varepsilon|I)P(b|I)P(\sigma|I)}{\zeta} \qquad \text{A.16}$$

where we have assumed that the model parameters are logically independent. Note, that the quantities b, ε, and σ are all N-dimensional vectors, and that S is an NxN matrix. The final probabilities for each bin are then obtained by marginalizing over the nuisance parameters, yielding:

$$P(\sigma_k|d, I) = \int dL \int_{all\ k} db \int_{all\ k} d\varepsilon \int \int_{all\ k} dS \int_{(i \neq k)} d\sigma \;\; P(\sigma, S, \varepsilon, L, b|d, I)\,. \qquad \text{A.17}$$

A.2.5 The Explicit Form

Given any model, Eq.A.17 provides a general form for extracting the cross section. The total luminosity contributes only to the overall normalization of the cross section, and will not affect the distribution, and we can consequently ignore that integration. Also, there are N^2 integrations involving the smearing matrix. For any sizable smearing matrix, this will mean far too many dimensions to consider integrating, so we will ignore those integrations, and use the central values of S for determining of the probability distributions for the cross sections. That is, we ignore the systematic uncertainty on the smearing matrix. Some of this information can be recovered by averaging over a few discrete values of S. That is, for any value of k, generate a family of $P(\sigma_k|d, I)$ values corresponding to different values of the elements of the smearing matrix, and take the weighted average to obtain the final value of $P(\sigma_k|d, I)$. Making such simplifications, Eq.A.17 becomes:

$$P(\sigma_k|d, I) = \int_{all\ k} db \int_{all\ k} d\varepsilon \int_{(i \neq k)} d\sigma \; P(d|\sigma, S, \varepsilon, L, b, I)P(\varepsilon|I)P(b|I)P(\sigma|I) \qquad \text{A.18}$$

where we have substituted for the joint posterior probability, and ignored the normalization.

The priors for the efficiency and the background can be taken as gaussian. We take the prior for the cross section in any bin as a constant with a lower bound of 0 and some upper bound σ_{max}. The combined priors are just the products of priors for the individual bins:

$$P(\varepsilon|I) = \prod_{i=1}^{N} P(\varepsilon_i|I) = \prod_{i=1}^{N} e^{-\frac{1}{2}\left(\frac{\varepsilon_i - \bar{\varepsilon}_i}{\sigma_\varepsilon^i}\right)^2} \qquad \text{A.19}$$

$$P(b|I) = \prod_{i=1}^{N} P(b_i|I) = \prod_{i=1}^{N} e^{-\frac{1}{2}\left(\frac{b - \bar{b}_i}{\sigma_b^i}\right)^2} . \qquad \text{A.20}$$

where, $\bar{\varepsilon}_i$ and $\bar{b}_i$ are the best estimates of the efficiency and background in each bin, respectively. The σ_b^i and σ_ε^i are uncertainties on the background and efficiency in each bin.

Substituting for the priors and for the likelihood function, Equation A.18 then becomes,

$$P(\sigma_k|d, I) = \qquad \text{A.21}$$
$$\int_{all\ i} db_i \int_{all\ i} d\varepsilon_i \int_{(i \neq k)} d\sigma_i \prod_{i=1}^{N} \left(\left(L\varepsilon_i \sum_{j=1}^{N} S_{ij}\sigma_j + b_i \right)^{d_i} e^{-\left(L\varepsilon_i \sum_{j=1}^{N} S_{ij}\sigma_j + b_i \right)} \right) \left(e^{-\frac{1}{2}\left(\frac{b - \bar{b}_i}{\sigma_b^i}\right)^2} \right) \left(e^{-\frac{1}{2}\left(\frac{\varepsilon_i - \bar{\varepsilon}_i}{\sigma_\varepsilon^i}\right)^2} \right)$$

The lower and upper bounds for the integration over ε_i are 0 and 1, while the lower and upper bounds for b_i are some constants b_i^{min} and b_i^{max}. If any efficiency is close to either 1 or 0, then the gaussian distribution may not describe our knowledge, and a different distribution should be used. A distribution that would maximize the range of values allowed could be determined using the method of maximum entropy[59].

A.2.6 Discussion of the Solution

The greatest complication in solving Eq.A.21 is the presence of the smearing matrix. If not for the smearing, the problem would correspond to an integral over a trivial product of independent gaussian and poisson distributions. To investigate the difficulties, we can ignore the background and efficiency terms, and thereby simplify the problem to the form:

$$P(x_k|d, I) = \int_{(i \neq k)} dx_i \prod_{i=1}^{N} \left(\sum_{j=1}^{N} M_{ij} x_j \right)^{d_i} e^{-\left(\sum_{j=1}^{N} M_{ij} x_j \right)} \qquad \text{A.22}$$

The product of polynomials is to the power d_i. If the d_i are small (*i.e.*, we have few events), then the number of terms is manageable and the above method is easy to implement analytically. However, because in general both N and d_i can be large, just the number of terms in the sum is very large. The number of terms in the multinomial alone is $\frac{(N + d_i - 1)!}{(N-1)!d_i!}$. For example, with $N = 5$ bins and $d_i = 100$ events, we have $\sim 10^6$ terms. Including the product over the number of bins, there are $\sim 10^{30}$ terms required in order to calculate the probability distribution—far too many to consider, even for the simplified case.

In most cases, the smearing matrix is be essentially diagonal. (Of course, a perfect detector would be characterized by a diagonal smearing matrix.) Under such circumstances, the number of dimensions of integration is something less than N. If the smearing matrix is diagonal, then the integral is easily separable into a product of single dimensional integrals. However, as the smearing becomes more severe, the number of true bins which were sources for the events observed in bin i becomes larger and larger and therefore the minimum number of integrations becomes large and larger. Nevertheless, when there is sufficient data, and the binning of the measurement is determined mainly by the experimental resolution, one would expect N to be at least of the order of 5. (Otherwise we would not be taking advantage of the resolution of our detector.) Thus, for a large number of events, we still have a very large number of terms, even if we limit the number of bins that we integrate to only those that contribute to the smearing in the bin under consideration.

When d_i are large, the poisson distribution can be approximated by a gaussian, and the problem can, again, be solved analytically, after appropriate substitutions in the exponent. We attempted this using the symbolic calculation engines Macsyma[70] and Mathematica[71]. Even for a 5-dimensional smearing matrix, the complexity of the substitutions proved to be too time-consuming for convergence.

Most of our effort in approaching this problem has relied on using techniques of numerical integration. In the simplest examples, we ignored priors for the nuisance parameters, and considered only the unsmearing of the cross section. We initially used the program VEGAS[69] to accomplish the integration, and achieved success in a 5 dimensional case, which is presented below. However, higher-dimensional matrices proved successively more difficult to handle. VEGAS integrates a function by subdividing the integration space along the axes of integration, attempting to concentrate the evaluation of the integrand to regions of the integration space that have the largest contributions to the integrand. However, if the integral involves many coordinates, then this dividing process is often not very effective. This type of an integrand we have called “diagonal”, and VEGAS does not handle that very well.

Another method of numerical integration, called MILX[72], seemed more promising because it addresses directly the diagonal nature of the integrand. MILX breaks up an integrand in terms of simplices, rather than the hyper-cubes of VEGAS. An example in two dimensions will help illustrate the difference. Integrating a two-dimensional function, VEGAS breaks up the plane into squares, optimizing the size of the squares such that they are small in regions where the integrand has large contributions and is varying rapidly, but large in regions where the integrand is small. The size of the squares reflects the relative number of evaluations of the integral in each region. MILX, on the other hand, breaks up the plane into triangles, which correspond to geometrical forms of a simplex in 2 dimensions. Again, the sizes of the triangles will vary in order to cover the integrand where it makes the largest contributions and changes most rapidly. If the integrand has significant contributions diagonally across the plane, the MILX method will be more efficient in covering the space than VEGAS. The difference is small in the 2-dimensional case, however, the difference can grow significantly with more dimensions.

The greatest problem with MILX is that it requires tuning to any specific problem, and that the method is still under development. After consultation with the authors (particularly with Dr. Galina Manankova), several problems still remained intractable. The most significant difficultly is that MILX fails when the number of data points is high. Effectively, the contribution to the integrand is too localized, and MILX fails to find the region where the contributions are most significant. Work is still in progress to tune MILX, but is not close to solution. In fact, we are currently pursuing other directions for resolving the difficulties. In addition to pressing on with the MILX method, we are also re-examining the analytical solution in an attempt to apply analytic approximation procedures that might lead to convergence.

A.2.7 A Simple Example

We present a simple example to demonstrate that our technique successfully unsmears data to reproduce the input distribution. The model assumes $\mu_k = LS_{kj}\sigma_j$, and does not include background. All efficiencies are taken to be 100%. We use nanobarns (*nb*) for the units of cross section, and assume the total luminosity of the data sample is $L = 10000\, nb^{-1}$. The priors for the luminosity and the smearing matrix elements are taken to be delta functions. Thus, the only numerical integration is done over the values of cross section. The assumed arbitrary smearing matrix is shown in Table A.1. It is taken purposefully as non-diagonal in order to emphasize the impact of smearing, and thereby the effectiveness of the unsmearing. The assumed true and smeared cross section, and the observed distribution of events (for one experiment) are shown in Table A.2. The smeared distribution was obtained by applying the smearing matrix to the true distribution. The number of events in any bin was obtained by a simple sampling of a poisson distribution, with a mean defined by the smeared cross section. The integrations were performed using VEGAS.

Figure A.1 shows the results of the study. The difference between the true distribution (squares) and the smeared one (triangles) is very large. The unsmeared result from two independent experiments (diamonds and circles) agree very well with the unsmeared input. The points on the unsmeared values are the most probable values determined from the probability distribution. The uncertainties represent the 68% certainty intervals of the probability distributions. The agreement clearly demonstrates the veracity of our unsmearing procedure.

0	*0*	*0.3*	*0.5*	*0.1*
0.1	*0*	*0*	*0.5*	*0.2*
0.5	*0*	*0*	*0*	*0.7*
0.4	*0.2*	*0*	*0*	*0*
0	*0.8*	*0.7*	*0*	*0*

Table A.1 Smearing matrix assumed for the 5 bin test.

bin	true (*nb*)	smeared (*nb*)	data (*nb*)
1	*0.1*	*0.68*	*6773*
2	*0.5*	*0.42*	*4306*
3	*1.0*	*0.26*	*2531*
4	*0.7*	*0.14*	*1430*
5	*0.3*	*1.1*	*11088*

Table A.2 Numerical values for the true and smeared cross section (*nb*), and the distribution in the number of events for a single test experiment.

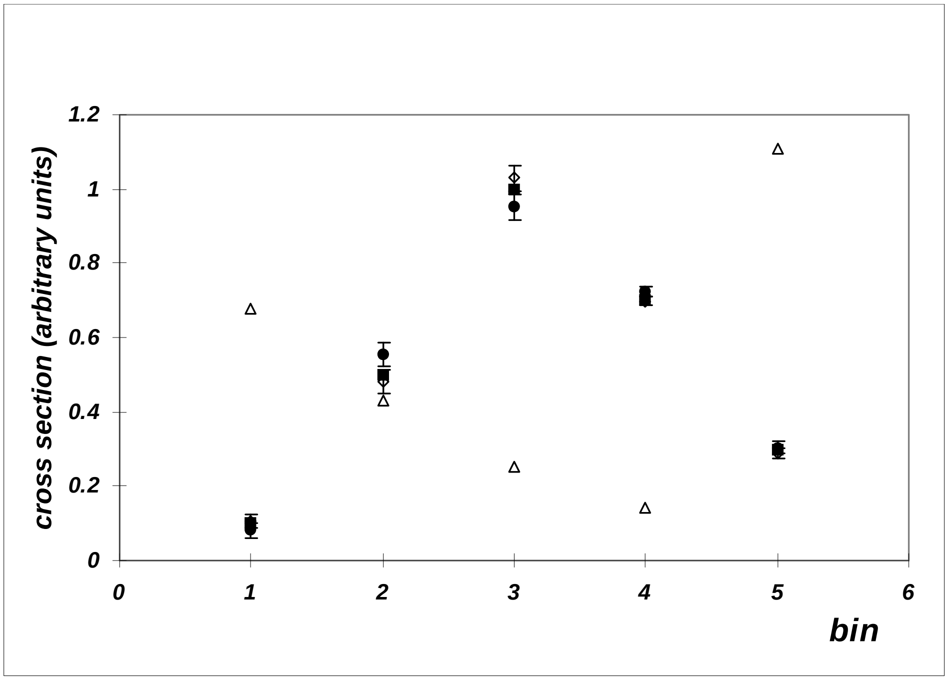

Figure A.1 Results of the 5 bin test. The true distribution corresponds to the squares. The open triangles show the smeared distribution. The points with error bars show the unsmeared results from two independent experiments.