SEARCH FOR LEPTOQUARKS DECAYING TO μ + X WITH THE DØ DETECTOR AT THE FERMILAB TEVATRON COLLIDER

Name: Daniel J. Karmgard
Department: Department of Physics
Major Professor: Vasken Hagopian
Major Professor: Sharon Hagopian
Degree: Doctor of Philosophy
Term Degree Awarded: Spring, 1999

We describe a search for the pair production of second generation leptoquarks that decay to muons plus other particles in 94 pb^{-1} of data taken with the DØ detector at the Fermilab Tevatron (center-of-mass energy $\sqrt{s} = 1.8$ TeV) from 1993–96. The search places limits on the cross sections and mass of second generation leptoquarks for various branching ratios and couplings. For both scalar leptoquarks decaying into a muon and a quark the mass limit is 200 GeV/c^2 while for one scalar leptoquark decaying into a muon and a quark with the other scalar leptoquark decaying into a neutrino and a quark the mass limit is 160 GeV/c^2 at the 95% confidence level.

THE FLORIDA STATE UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

SEARCH FOR LEPTOQUARKS DECAYING TO μ + X WITH THE DØ DETECTOR AT THE FERMILAB TEVATRON COLLIDER

By

DANIEL J. KARMGARD

A Dissertation submitted to the
Department of Physics
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Degree Awarded:
Spring Semester, 1999

The members of the Committee approve the dissertation of Daniel J. Karmgard defended on March 5, 1999.

Vasken Hagopian
Professor Co-Directing Dissertation

Sharon Hagopian
Professor Co-Directing Dissertation

Myra Hurt
Outside Committee Member

Howard A. Baer
Committee Member

Lawrence Dennis
Committee Member

Harrison B. Prosper
Committee Member

This dissertation is dedicated to my family and my teachers. You believed in me until I believed in myself. Thank you.

ACKNOWLEDGEMENTS

On a collaboration such as the D0 experiment, there are a large number of people who contribute to a work such as this. Ultimately, everyone who has worked on the experiment has contributed something, be it building the detector, designing the data acquisition system, writing the software, or taking the data. They are to numerous to name individually but their efforts have made this work possible.

There are, however, a number of people who have contributed directly to this work. First and foremost, to Vasken and Sharon Hagopian for comment, criticism and encouragement. I would also like to acknowledge the Leptoquark Working Group, who instructed me on the proper way to do a physics analysis, and especially to Greg Landsberg. Thanks are also due to Terry Heuring for answering a steady stream of questions with patience and consideration while I learned. The professors in the Florida State University group: Horst Wahl for his example, Harrison Prosper for his enthusiasm and knowledge of statistical analysis, Susan Blessing for demanding excellence of herself and those around her, and Steve Linn for his initial teaching. Many thanks are also due to Professor Mark Adams, who taught me how to do an analysis using muons, and whose questions helped me to understand my own work better. Because of him this work is far better than it would otherwise have been.

Finally, many thanks are due to the other students resident at D0. Especially to the other FSU students: Jeff McDonald, Russ Gilmartin, Brian Connolly, and Guoliang Wang. We have learned from each other and together, shared frustration, long nights of work, and triumphs. For their encouragement, compassion, understanding and fruitful discussions I am grateful.

CONTENTS

APPENDICES

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

We describe a search for the pair production of second generation leptoquarks that decay to muons plus other particles in 94 pb^{-1} of data taken with the DØ detector at the Fermilab Tevatron (center-of-mass energy $\sqrt{s} = 1.8$ TeV) from 1993–96. The search places limits on the cross sections and mass of second generation leptoquarks for various branching ratios and couplings. For both scalar leptoquarks decaying into a muon and a quark the mass limit is 200 GeV/c^2 while for one scalar leptoquark decaying into a muon and a quark with the other scalar leptoquark decaying into a neutrino and a quark the mass limit is 160 GeV/c^2 at the 95% confidence level.

CHAPTER 1

INTRODUCTION

> “If the universe is the answer, what was the question?”– Leon Lederman from his book *The God Particle*.

Particle physics is that branch of the physical sciences which concerns itself with discovering the most basic building blocks of nature. As a philosopher looks at the world and wonders what it all means, a particle physicist will wonder what it's all made of. This quest for knowledge has a long tradition in western thought. It began with Democritus, a Greek philosopher in the 5^{th} century BC. He believed that everything in the world was made up of small, indivisible chunks of matter called atomos, the Greek word for indivisible. This philosophy has influenced centuries of philosophers and scientists and the work continues today.

The idea that the world was constructed out of small indivisible units of matter received a large boost in 1872 when Dmitri Ivanovich Mendeleev [1] published the periodic table of the elements. Using his table he was able to predict the properties of yet undiscovered elements by empty places on the table. The elements became known as atoms, in homage to the Greek ideal of the indivisible unit.

Unfortunately, the atoms of 19^{th} century chemistry were not the ultimate building blocks of nature. By the turn of the century (1897) J. J. Thompson [2] had discovered the electron and with it the idea of atomic structure was born. Throughout the next several decades chemists and physicists worked to understand the structure of the atom. The structure was finally explained in light of a new theory, quantum mechanics, which began with the idea in 1900 by Max Planck [3] that energy comes in discrete bundles (quantization of energy). In 1911 Rutheford [4] found that the atom is composed of two pieces – the nucleus, and the electrons in orbitals about it. The shapes of the electron orbitals and the way in which they interacted with one another could be explained by the new quantum theory. By 1919 it was known that the nucleus was composed of particles known as protons, but the discovery in 1932 of the neutron by Chadwick [5] demonstrated the the the nucleus was composed of two particles – the proton and the neutron bound together to form the nucleus.

In 1927, Paul Dirac [6] put the quantum theory together with another great theory of 20^{th} century physics, special relativity [7], to form relativistic quantum theory. In doing so he also predicted the existence of antimatter, specifically the anti-electron or positron. The positron has the opposite electrical charge from the electron (positive rather than negative) but is otherwise the same. The positron was discovered by Carl Anderson [8] in 1932.

The advent of modern particle physics as a distinct field of study began in the 1940s with the work of Feynman, Schwinger, and Tomonaga [9], who developed quantum electrodynamics (QED). They used the quantum theory to quantize the electromagnetic field theory. That is, they explained electromagnetic phenomena at a basic level in terms of the exchange of photons (the particles of light).

As the 1940s and 1950s progressed, physicists began to discover more and more particles. The particle accelerator had been born and was in wide use. The particle accelerator uses powerful magnets and oscillating electric fields to accelerate particles and collide them with one another. The debris from the collision is examined and conclusions may be drawn about the structure of the original particles. In order to “look” into a particle (for example, a proton) one uses another particle (say, an antiproton). To penetrate deeply into the particles' structure one requires a high momentum for the collision. Through the relativistic relation $E^2 = p^2c^2 + m^2c^4$, high momentum translates into high energy. In fact, high energy physics is widely used as a synonym for particle physics and the two terms will be used inter-changeably throughout this work.

By 1960 the number of supposedly fundamental particles had proliferated without any apparent organizing principle. Then Murray Gell-Mann [10] proposed the $SU(3)$ group and the eight-fold way. The $SU(3)$ group was to particle physics what the periodic table of the elements had been to chemistry: a superb organizing principle built empirically without any underlying rational. Gell-Mann even predicted the existence and properties of an undiscovered particle based on an empty space in one of his octets, much as Mendeleev had predicted the existence of unobserved atoms 90 years before.

The organizing principle underlying the eight-fold way was discovered by Gell-Mann and George Zweig [11] a few years later when they independently proposed the quark model. The quark model proposed that all of the particles in the eight-fold way were combinations of three fundamental particles called quarks. They were known as the up, down and strange quarks. Some particles (baryons) were combinations of three quarks. Other particles (known as mesons) were combinations of a quark and an antiquark. Later, it was found that the quarks carried a new type of charge, analogous to the electric charge, known as color. Color charge came in three types as opposed to the two charges in electromagnetic theory. The color charges are red, green and blue (R,G,B). Although the quark model could explain the structure of the eight-fold way few physicists took it literally. Nobody had ever seen a quark or any evidence of structure in the baryons or mesons.

That changed in 1969 when experimental results began to suggest that the proton had substructure [12]. It was still unclear just what these particles were however. In 1974 a fourth quark was discovered [13] and the quark model took it's place as a fundamental theory in particle physics. Also in the 1970s, a new theoretical structure was being developed [14] called quantum chromodynamics (QCD). It was based on ideas taken from Feynmans' quantization of electromagnetism. The Standard Model of particle physics began to take shape.

Throughout the 1970s, 1980s and up to the present, physicists have worked to complete the Standard Model. The particles it predicts have been searched for and (except for one – the Higgs particle) have been found [15]. The theories have been put together into a self-consistent whole. However, there is also a considerable body of opinion among high energy physicists that the Standard Model is not the ultimate theory of particles and their interactions. There are a number of reasons for this, a few of which we will touch upon here.

There is a strong feeling among many in the high energy physics community that the Standard Model is incomplete. Some theorists (especially those for whom mathematical elegance is the primary recommendation of a theory) think that the Standard Model is an *ad hoc* collection of unrelated theories cobbled together without a unifying scheme. Another discouraging factor is that gravitation, the most important of the physical forces in the macroscopic world, is not included in the Standard Model. No one has yet succeeded in developing an experimentally verifiable theory of quantum gravity. There is also the so-called naturalness problem (discussed in section 2-4). For these reasons, among others, physicists have been motivated to search for physics beyond the Standard Model.

This work describes such a search. The search is for a hypothetical particle known as a leptoquark, because it has properties in common with both quarks and leptons.

The search was conducted at the DØ collider detector located at the Fermi National Accelerator Laboratory in Batavia, IL. This dissertation is organized as follows. This chapter has been a brief introduction to particle physics and its history. Chapter 2 will describe the Standard Model of particle physics in more detail, while chapter 3 will develop the necessary theoretical considerations to understand a search for leptoquarks. Chapter 4 will describe the experimental apparatus. Chapters 5 and 6 will give the specifics of the search. Chapter 7 will describe the combination of the various searches described in the preceeding two chapters and present overall results in a unified fashion. Chapter 8 will present some concluding thoughts. Finally, two appendices are included which present in detail two of the techniques used for the reader who is unfamiliar with them. Appendix A describes the calculation of confidence intervals in Bayesian statistics and Appendix B discusses neural networks.

CHAPTER 2

THE STANDARD MODEL

“It is a good idea not to put too much faith in experimental results until they are confirmed by theory.” – Sir Arthur Eddington

2.1 Introduction

From the particle physics point of view the elementary particles consist of the “leptons” and “quarks” having spin $= \frac{1}{2}\hbar$, the “bosons” having integral spin, and the “hadrons”. $\hbar$ is the reduced Planck constant $\hbar \equiv h/2\pi = 1.05 \times 10^{-34} Js$. The leptons enter the mathematical models as doublets; the electron and its associated neutrino (e^-, ν_e), the muon and muon neutrino (μ^-, ν_μ), and the tau and tau neutrino (τ^-, ν_τ). There are four vector bosons (spin $= \hbar$): the massless photon (γ), and the three massive vector bosons of the weak force, the W^+, W^-, and the Z^0. The hadrons are further divided into “mesons” (particles having integer spin $S = 0, \pm 1\hbar, \pm 2\hbar, \dots$) and the “baryons” (particles having half-odd integer spin $S = \pm\frac{1}{2}\hbar, \pm\frac{3}{2}\hbar, \dots$). This chapter presents a review of the theoretical basis of particle physics and relies heavily on references 16–18.

Leptons and vector bosons are believed to be fundamental. That is, they have no substructure down to a length scale of approximately $10^{-18}m$. In the mathematical models they are therefore treated as elementary fields in the lagrangian which describes their dynamical interaction. Hadrons, however, have been known for some time to have a definite size (of the order of $1fm = 10^{-15}m$) and to be composed of more elementary particles known as quarks. There are six known types (or flavors) of quarks, (u, d, s, c, b, t) for up, down, strange, charm, bottom, and top respectively.

The guiding principles for the construction of mathematical models in particle physics are Lorentz invariance, local gauge invariance and the principle of least action. Local gauge invariance is the requirement that the Dirac equation gives the same result when the quantum mechanical wave function (ψ) undergoes a phase change where the phase factor (δ) is a function of the coordinate system $(\delta = \delta(x))$. Thus, the Dirac equation must remain invariant when

$$\psi(\vec{x}, t) \rightarrow \psi\prime(\vec{x}, t) = e^{i\delta(x)}\psi(\vec{x}, t) \tag{2.1}$$

The Standard Model [19] of particle physics is a highly successful example of a minimal theoretical formulation based on these principles. The mathematical structure of the Standard Model proceeds from the mathematical theory of groups (group theory). Group theory deals with collections of objects (*i.e.* groups) which share similar properties under various transformations. The group with which the Standard Model is concerned is

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \tag{2.2}$$

In *eqn.* 2.2 S denotes "Special" (determinant = +1), U denotes "Unitary", the numbers in the parentheses refer to the dimension of the group and $\otimes$ is the direct product of the groups. So $SU(2)$ is a special unitary group of 2×2 matrices whose members have the property that $det(\mathbf{M}) = \mathbf{+1}$ and $\mathbf{M}\mathbf{M}^{\dagger} = \mathbf{M}^{\dagger}\mathbf{M} = \mathbf{I}, \forall\ \mathbf{M} \subseteq \mathbf{SU(2)}$, where $\mathbf{I}$ is the identity matrix. The subscripts on the groups refer to the physical property that connects the elements of the group. "C" refers to the color charge, "L" to the weak isospin and "Y" the weak hypercharge.

The $SU(3)_C$ gauge group is the symmetry group concerned with interactions proceeding via the strong nuclear force. This group acts on the quarks which are the elementary constituents of mesons and baryons, and is mediated (or transmitted) by the gluons (g) which are the gauge bosons of the group. The quarks and gluons are fields which carry the color charge which comes in three types, R, G, B as opposed to the two charges of the electromagnetic field ($\pm$). The "coupling" (strength of the force) between quarks and gluons is denoted by α_s, which may be greater than one. At high energy, α_s is sufficiently small that perturbation theory may be applied. α_s is not a constant term, but changes with the energy scale. As an example, when evaluated at the Z^0 mass (energy scale), $\alpha_s \approx 0.118$. The symmetry in $SU(3)_C$ is exact with the consequence that the gluons are massless. The theory concerned with strong interactions based on the $SU(3)_C$ group is known as Quantum Chromodynamics (QCD).

The $SU(2)_L \otimes U(1)_Y$ gauge group is the symmetry group of the unified electromagnetic and weak nuclear interactions. $SU(2)_L$ is the weak isospin group while $U(1)_Y$ is the hypercharge group. At low energies ($< 250\ GeV$) the $SU(2)_L \otimes U(1)_Y$ symmetry is broken into the $U(1)_{em}$ group describing electromagnetic interactions whose gauge boson is the massless photon and an associated coupling strength of

$$\alpha_{em} = e^2/c\hbar \approx 1/137 \tag{2.3}$$

at zero momentum, where e is the electric charge of the electron. The other residual group, $SU(2)_L$, describes the weak nuclear interactions. The mathematical method of symmetry breaking (known as the Higgs mechanism [20]) requires that the gauge bosons of this group acquire mass. The gauge bosons are the $W^\pm$ and the Z^0.

Table 2.1. The fermions of the Standard Model.

$q = \frac{2}{3}e$ quarks	$q = -\frac{1}{3}e$ quarks	$q = -e$ leptons	$q = 0$ leptons
u	d	e	ν_e
c	s	μ	ν_μ
t	b	τ	ν_τ

The Standard Model fermions (quarks and leptons) are summarized in table 2.1. The particles are grouped by electric charge (q) and mass (increasing down the table). Each member particle listed has a corresponding antiparticle. Quarks carry fractional electric charge (in units of the charge of the electron $-e$). The charges of the quarks are $\frac{2}{3}e$ (u,c,t) and $-\frac{1}{3}e$ (d,s,b). The leptons have integer charges $-e$ (e,μ, τ) or 0 (neutrinos).

The fermions interact with one another via the three forces contained in the Standard Model. The forces are: the strong nuclear force, the weak nuclear force, and the electromagnetic (EM) force. The interactions are described in terms of the exchange of fundamental vector bosons (spin 1, obeying Bose-Einstein statistics). The bosons – or quanta of the various fields – are the photon (γ) which mediates the EM force, the gluon (g) which mediates the strong nuclear force, and the $W^{\pm}$ and Z^0 bosons which mediate the weak nuclear force. The fourth of the physical forces, gravitation, is not contained in the framework of the Standard Model. Although it is important on a macroscopic scale it is very weak and does not play an appreciable part in interactions among fundamental particles.

The electromagnetic force is an interaction between particles having electric charge. It has an infinite range and is responsible for the attraction between electrons and the atomic nuclei.

Table 2.2. The fundamental interactions of physics

Interaction	Relative Strength	quanta	Mass (GeV/c^2)	Range
Strong	0.118	$8g$	0	$10^{-15}m$
EM	1/137	γ	0	∞
Weak	10^{-5}	$W^{\pm}/Z^0$	80.22/91.17	$10^{-18}m$
Gravity	10^{-38}	$G(?)$	0	∞

The strong nuclear force is a strong, short-range force that affects particles that carry color charge. Unlike the electric force, the strong charge comes in three types known as red, green, and blue (R,G,B). The strong force is responsible for binding together the nucleus of an atom and the quarks within the nucleons.

The weak nuclear force is a weak force with very short range. It exists between any of the leptons and quarks. It is responsible for things like the radioactive β decay of a nuclei. The strength of the weak nuclear force is approximately 10^{-5}. All of the forces are summarized in table 2.2. The following sections will discuss each of the Standard Model interactions individually.

2.2 Quantum Chromodynamics

Quantum Chromodynamics [21] (or QCD) was introduced in the early 1970s to describe the hadrons in terms of a set of more elementary particles, known as quarks. To fully describe what we observe in nature, six quarks are required. Like the leptons the quarks are pointlike (no observed structure) fermions having spin $S = \pm\frac{1}{2}\hbar$. Baryons are composed of three quarks $B = (q_a, q_b, q_c)$ and mesons a quark-antiquark pair $M = (q_a, \overline{q_b})$. Making these assignments one obtains the half-integer spin of the baryons and the integer spin of the mesons after accounting for the orbital angular momentum of the quarks within the hadrons. The quarks have electric charges in units of $\frac{1}{3}e$, so that the hadrons have only integer electric charge. The only problem at this point is that to describe the mass spectrum of the hadrons using a *baryonic* wave function requires a symmetry under the interchange of the quarks: $q_a \leftrightarrow q_b$. This symmetry violates Fermi-Dirac statistics, which applies to particles having half-odd integer spins. The problem is cured with the introduction of a new quantum number, the color, and the requirement that each quark comes in three colors $(q_a)^T = (q_a^1, q_a^2, q_a^3)$, where 1,2,3 are the color indices. The baryonic wave function is then totally antisymmetric under a color interchange

$$B = \sum_{i,j,k} (q_a^i, q_b^j, q_c^k) \tag{2.4}$$

while the *mesonic* wave function is constructed as

$$M = \sum_{i} (q_a^i, q_b^i) \tag{2.5}$$

All observed particles have the property that they are unchanged by a rotation in the $SU(3)_C$ group. Composite particles have no net color charge, a property known as confinement. This is analogous to the idea that atoms are neutral although they are composed of electrically charged particles. This colorless state has the immediate consequence that the only physically allowed combinations of quarks are those in which the color charges add to zero. Therefore, quarks may come in groups of three, one of each color for no net color, or in groups of two, a color and its complement. (Since larger groupings of quarks are simply multiples of 2 or 3 they are interpreted as combinations of non-fundamental particles rather than as extended groupings of quarks).

The generators of the $SU(3)_C$ group are 8 $(3 \times 3 - 1)$ matrices which are both traceless and hermitian. Traceless means that the sum of the diagonal elements is zero, while hermitian means that the complex transpose of the matrix is equal to the original matrix. The generators are denoted as λ_i with $i = 1, 2, \ldots, 8$ and are known as the Gell-Mann matrices. The diagonal matrices are taken as:

$$\lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \lambda_8 = \sqrt{\frac{1}{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \tag{2.6}$$

having simultaneous eigenvectors

$$R = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, G = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, B = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{2.7}$$

QCD is mediated by massless vector bosons (spin $= 1\hbar$) called gluons. The gluons come in eight color combinations:

$$R\overline{G}, R\overline{B}, G\overline{R}, G\overline{B}, B\overline{R}, B\overline{G}, \sqrt{\frac{1}{2}}(R\overline{R} - G\overline{G}), \ \sqrt{\frac{1}{6}}(R\overline{R} + G\overline{G} - 2B\overline{B}) \tag{2.8}$$

the $SU(3)_C$ singlet $\sqrt{\frac{1}{3}}(R\overline{R} + G\overline{G} + B\overline{B})$ carries no net color charge and hence is not a mediator in color interactions.

The strong coupling parameter α_s is not a constant, rather, it changes with the energy scale,

$$\alpha_s(|q^2|) = \frac{12\pi}{(11n - 2f)\ln(|q^2|/\Lambda^2)} \qquad (|q^2| \gg \Lambda^2) \tag{2.9}$$

where q is the momentum transfer, n = number of colors (3, in the Standard Model), f = number of flavors (6, in the Standard Model), and Λ is an energy scale,

$100 MeV < \Lambda < 500 MeV$. At low energies, α_s is nearly 1. However, at higher energies $\alpha_s < 1$ (as noted in table 2.2, $\alpha_s \approx 0.12$ when evaluated at the Z^0 mass, $q = m_{Z^0} = 91\ GeV/c^2$). Since this is so, QCD calculations may use perturbation theory for short distance (*i.e.* high energy) interactions. Thus, the theory is known as "asymptotically free". This means that at very short distances (*e.g.,* inside the radius of the proton) the quarks behave as if they were free particles. As the distance between them increases the strength of the force also increases. This has important consequences for this work. In $p\overline{p}$ collisions (such as those at the Tevatron) two quarks may be knocked out of the (anti)proton. As they move away from each other the potential energy between them increases with the distance. Thus it is soon energetically favorable for a new $q\overline{q}$ pair to be created from the vacuum rather than for the energy between the original pair to continue increasing. This process is known as hadronization. The same situation applies for gluons, which also carry a color charge. In the environment at the Tevatron, quarks and gluons are detected as highly collimated "jets" of hadrons. These jets will play an important role in this work.

2.3 Electroweak Interactions

The theory underlying the $U(1)_{em}$ symmetry group is Quantum Electrodynamics [18] (QED), first proposed by Dirac in 1927 and developed by Feynman, Schwinger, and Tomonaga in the 1940s. In QED a free electron is described by a four component spinor which corresponds to a spin state $J_z = \pm\frac{1}{2}\hbar$. The negative energy states are interpreted as antiparticles. The antielectron (positron) was discovered by Carl Anderson in 1932, validating the basic ideas of QED. The generator of the $U(1)_Y$ symmetry group is the weak hypercharge Y.

The $SU(2)_L$ symmetry group corresponds to the weak nuclear interaction. Its generators are the three components of the weak isospin (I). The generators are written as Pauli matrices $I_i = \sigma_i$ where

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.10}$$

with $i \equiv \sqrt{-1}$. While the weak and EM interactions appear unrelated at low energies (*e.g.*, $q^2 \ll M_W^2$, where q is the momentum transfer of the interaction), they become unified at higher energy scales (*i.e.* $q^2 \approx M_W^2$). This is represented by the Weinberg-Salam $SU(2)_L \times U(1)_Y$ group, known as the electroweak interaction. The fundamental vector bosons of the group are massless isovector triplets W_μ^i (i = 1, 2, 3) for the $SU(2)_L$ group and a massless isosinglet B_μ for the $U(1)_Y$ group. The Higgs mechanism [20] is used to spoil, or break, the symmetry of the group at low energies and to give mass to the vector bosons. Two of the W_μ^i acquire a mass and become the $W^\pm$; one linear combination of the B_μ and W_μ^3 becomes the Z^0, while another becomes the photon. This mechanism requires the introduction of a new particle into the Standard Model – the Higgs particle (H). The H is an isospin doublet of scalar mesons that generate the particle masses as a result of a self-interaction. The Higgs particle is the last remaining unobserved particle in the Standard Model.

The interaction of the fermions with the fields $\mathbf{W}_\mu$ and B_μ is described by a lagrangian density $\mathcal{L}$ and the fermionic currents

$$\mathcal{L} = g\mathbf{J}_\mu \mathbf{W}_\mu + g' J_\mu^Y B_\mu \tag{2.11}$$

where $\mathbf{J}_\mu$ and J_μ^Y are the isospin and hypercharge currents of the fermions respectively. g and g' are the couplings to the $\mathbf{W}_\mu$ and B_μ. They are related by

$$e = g\sin(\theta_W) = g'\cos(\theta_W) \tag{2.12}$$

where θ_W is known as the weak mixing angle ($\sin(\theta_W) \approx 0.23$).

The weak charged current interaction is parity violating while the electromagnetic interaction is parity conserving. Parity violating weak interactions connect left-handed states (*e.g.,* ν_μ and μ) whereas the parity conserving interaction applies to both left and right handed states. The lepton states consist of a left-handed doublet and a right-handed singlet:

$$\psi_L = \frac{1}{2}(1+\gamma_5)\begin{pmatrix}\nu_\mu \\ \mu\end{pmatrix}; (T = \frac{1}{2}, Y = -1) \tag{2.13}$$

$$\psi_R = \frac{1}{2}(1-\gamma_5)(\mu^-); (T = 0, Y = -2) \tag{2.14}$$

while the relevant states for the quarks are:

$$\psi_L = \frac{1}{2}(1+\gamma_5)\begin{pmatrix}u \\ d\end{pmatrix} \tag{2.15}$$

$$\psi_R = u_R \text{ or } d_R \tag{2.16}$$

In *eqns.* 2.11–14 above T and Y are the generators of the $SU(2)_L$ and $U(1)_Y$ gauge transformation groups. For low energies ($q^2 \ll M_W^2$) the masses of the bosons may be written as

$$M_{W^{\pm}} = \left(\frac{\sqrt{2}g^2}{8G_F}\right) = \left(\frac{\sqrt{2}e^2}{8G_F \sin^2(\theta_W)}\right) \tag{2.17}$$

$$M_{Z^0} = \left(\frac{M_{W^{\pm}}}{\rho \cos^2(\theta_W)}\right) \tag{2.18}$$

$$M_{\gamma} = 0 \tag{2.19}$$

where G_F is the Fermi constant ($G_F/(\hbar c)^3 = 1.17 \times 10^{-5} GeV^{-2}$), and ρ is a constant scale factor measured to be 1 in all experiments to date.

2.4 Beyond the Standard Model

Although the Standard Model has been tested repeatedly and found to be in good agreement with the experimental evidence, there are reasons to believe that it is not the ultimate theory of fundamental particles and their interactions. The Standard Model is an *ad hoc* choice of the gauge groups, it contains no explanation for the apparent symmetry between the quark and lepton sectors (*i.e.* the similarity between the family and generational structure seen in table 2.1). There is no reason within the Standard Model framework for the number of generations.

There is also the naturalness problem. This refers to the fine tuning necessary for the parameters in the Standard Model so that it agrees with experiment. As an example, consider the mass of the Higgs boson. This particle is required in the Standard Model to induce spontaneous symmetry breaking and to give mass to the other particles. The Higgs mass (m_H) is quadratically divergent; $m_H^2 = m_0^2 + \delta m_H^2 \approx m_0^2 - g^2\Lambda^2$ to lowest order in perturbation theory. m_0 is the "bare" Higgs mass, g is a dimensionless coupling constant and Λ is the energy scale. m_H is thought to be comparable to the electroweak scale ($O(250\ GeV/c^2)$). If $g^2 \approx 1$ and Λ is $O(M_{GUT})$ ($10^{16}\ GeV/c^2$) or $O(M_{Planck})$ ($10^{19}\ GeV/c^2$) then m_0 must be adjusted so that $m_0^2 - g^2\Lambda^2 \approx m_H^2$. This requires a fine-tuning over 30 orders of magnitude. This seems an unnaturally precise adjustment to make in a fundamental theory.

There is also the problem that gravity cannot be included in the Standard Model. As stated before, the gravitational interaction is sufficiently weak that it plays no role in fundamental particle interactions. Nonetheless, a theory which does not include one of the four forces of physics is probably not a fundamental theory.

Finally, there is the problem of unification. While there is no empirical evidence for it, most physicists believe that the four forces are different aspects of a single, unified force that would be apparent at sufficiently high energies scales (*e.g.,* M_{GUT} or M_{Planck}, the scale available at the moment of the big bang). Within the Standard Model, this is not possible. Thus, many physicists believe that a new theory is required which goes beyond the Standard Model. While we shall not be concerned with the particulars of these theories in this work, a brief description of some of them is included for completeness. To date, no conclusive experimental evidence for physics beyond the Standard Model has been observed.

2.4.1 Grand Unified Theories

Grand Unified Theories [22] (GUTs) propose a single interaction to describe the electromagnetic, weak and strong interactions which have a single coupling at some unification scale ($O(M_{GUT})$). They postulate a new spontaneous symmetry breaking analogous to that already present in electroweak theory to account for the different couplings seen in nature (at a much lower energy scale). The simplest of the GUT theories is based on the $SU(5)$ symmetry group which incorporates the $SU(2), SU(3)$, and $U(1)$ symmetry groups within it. However, the $SU(5)$ group has been ruled out because it predicts that protons will decay with a half-life which has been excluded by experiment. Other examples of GUT theories are string theory, which postulates the existence of strings of matter as the fundamental constituents of nature, and superstring theory, which incorporates a supersymmetry (see below) into string theory. These theories have the advantages of mathematical elegance and that they incorporate gravity in a natural fashion. They have the disadvantage that they require at least 11 dimensions to work.

2.4.2 Supersymmetry

The most popular theories beyond the Standard Model today are supersymmetric theories [23] (SUSY). This class of theories postulates a relation between the bosons and the fermions in the Standard Model. Each Standard Model particle has a superpartner (called a sparticle) with spin differing by $\frac{1}{2}\hbar$, but with the other quantum numbers remaining the same. Thus the Standard Model bosons have superpartner fermions and the Standard Model fermions acquire superpartner bosons. The superpartners cancel the quadratic divergences in the masses of the Standard Model particles thus yielding finite results to calculations without fine-tuning.

2.4.3 Leptoquarks

Leptoquarks [24], hypothetical particles that carry both lepton and baryon quantum numbers, arise in a natural fashion in several theories beyond the Standard Model. Particles which carry both lepton and baryon number are not allowed within the framework of the Standard Model, thus any search for leptoquarks (LQ) is necessarily a search beyond the Standard Model. Leptoquarks also carry fractional electric charge and $SU(3)_C$ color. Leptoquarks may be either scalar (spin = 0) or vector (spin = $1\hbar$) bosons. Vector LQ may have Yang-Mills couplings or anomalous couplings to the gauge bosons of the Standard Model (γ, g, $W^{\pm}$, and Z^0). Vector leptoquarks with Yang-Mills couplings are fundamental gauge bosons, while those with anomalous couplings may be bound states of a quark and a lepton. Leptoquarks offer a possible explanation for the symmetry between the quark and lepton sectors of the Standard Model.

The search described in this work is as model independent as a new phenomena search can be. This is because we look for LQ which are pair produced in $p\overline{p}$ collisions. The production of the LQ pair comes from gluon splitting ($g \rightarrow LQ\overline{LQ}$) and is therefore a strong interaction described by QCD rather than any new theory. The following chapter will set forth the relevant LQ phenomenology necessary to fully understand the search.

CHAPTER 3

LEPTOQUARK PHENOMENOLOGY

> “It is a capital mistake to theorize before one has data. Insensibly one begins to twist fact to suit theory instead of theory to suit fact.” – Sir Arthur Conan Doyle.

3.1 Overview

Within the Standard Model there is a symmetry observed between the quarks and the leptons. This symmetry is shown in table 2-1 at the beginning of the previous chapter. The structure of the particle groupings by mass and electric charge is the same. The rows in the table are known as families, the columns as generations. Thus the statement that the quarks and leptons in the Standard Model have the same family and generational structure. There is nothing in the Standard Model to account for this symmetry. However, the existence of such a symmetry hints at a possible underlying relation between the two types of particles at a fundamental level. This is a very attractive possibility offering the chance to express the fundamental particles of nature as differing aspects of a more fundamental type. Theories beyond the standard model which attempt to incorporate this symmetry almost universally predict the existence of some type of leptoquark. We are indebted to the work found in reference 17 for his excellent summary of leptoquark phenomenology.

Leptoquarks are particles that provide the common ground between leptons and quarks. They carry both lepton and baryon number, they couple both to leptons and quarks, carry fractional electric charge and integer spin (they are bosons). Leptoquarks are color triplets under $SU(3)_C$. In those theories where lepton and baryon number are conserved individually, leptoquarks could have masses of the order of the electroweak scale ($m_W \approx 80\ GeV/c^2$) without causing proton decay more rapidly than allowed by experimental constraints. Leptoquarks can take on many quantum numbers depending on the particular model used. They may have spin 0, 1, or 2; electric charge $-\frac{4}{3}, -\frac{1}{3}, \frac{2}{3}$, or $\frac{5}{3}$; isospin 0, $\frac{1}{2}$ or 1; baryon number $\pm\frac{1}{3}$; and lepton number $\pm$ 1.

3.2 Leptoquarks in Theoretical Models

Leptoquarks could exist as bound states of leptons and quarks in an extension to the Standard Model known as the strong-coupling confining standard $SU(2)_L \times U(1)_Y$ electroweak model [25]. In this model the conventional $SU(2)_L \times U(1)_Y$ lagrangian is used. Two additional assumptions are made. The first is that the $SU(2)_L$ coupling constant becomes large at a mass Λ ($\Lambda \approx O(250\ \mathrm{GeV}/c^2)$) which sets the energy scale for weak interactions. The second assumption is that no appreciable scalar (spin = 0) vacuum expectation value exists. This preserves the particle spectrum, as well as the charged and weak current interactions of the SM. At energies near or above the weak mass scale this theory diverges from the SM, with the strong coupling model allowing quark-lepton bound states.

In Grand Unified Theories [22], a leptoquark induces the transformation of a quark into a lepton resulting in a strong similarity between the leptons and the quarks as a naturally arising consequence of the unification.

Leptoquarks arise in the $SU(4)$ Pati-Salam [26] unification model where quarks are assumed to carry four colors. Three of these are the conventional colors of QCD while the fourth represents lepton number. The unification comes about by extending the $SU(3)$ gauge symmetry to an $SU(4)$ gauge symmetry. This results in leptoquarks as exotic gauge mesons carrying baryon and lepton quantum numbers.

The Technicolor [27] model attempts to deal with some of the problems of the SM such as the arbitrary choice of the scalar sector, the unknown decay modes of the scalars, the ambiguity of the elementary scalar solution, and the lack of experimental information about the number of $SU(2)_L$ weak-isospin doublets. Leptoquarks are color triplet technipions with $B = \frac{1}{3}$ and $L = -1$, which may decay into both quarks and leptons.

Leptoquarks also arise in $SU(5)$ unified theories. If the $SU(5)$ symmetry is a remnant of an E_6 invariance where particles are grouped in the 27-dimensional representations suggested by superstring theories [28], leptoquarks would occur in the 5 and 5* representations. The standard $SU(5)$ leptoquarks would then have to be very massive ($M_{LQ} \approx O(10^{10} - 10^{15}\ GeV/c^2)$) to avoid a rapid proton decay. This comes about because the $SU(5)$ invariance requires that leptoquarks also couple to quark pairs. Leptoquarks with relatively low masses ($O(m_W)$) can couple only to quarks and leptons. The coupling to quark pairs is forbidden. This violates the $SU(5)$ symmetry.

The introduction of leptoquarks generally leads to contributions to rare decays such as proton decay. Leptoquarks that have couplings that violate baryon or lepton number conservation must be very massive to avoid such complications. However they may still be as light as a few hundreds of GeV/c^2 while still avoiding such conflicts. Additionally, if leptoquarks coupled to all leptons they would give rise to flavor-changing neutral currents. Such interactions are severely restricted [29] by experimental data. Within this work we shall consider only leptoquarks which couple within a single generation of quarks and leptons.

3.3 Leptoquark Couplings

The most general interaction for a scalar leptoquark decaying to leptons and quarks is given by the lagrangian

$$\mathcal{L} = \overline{l}^i(\lambda_1^{ij} + \lambda_2^{ij}\gamma_5)q^j S + H.C. \tag{3.1}$$

where i, j are generation indices with $|\lambda_1^{ij}|^2 + |\lambda_2^{ij}|^2 = (\lambda^{ij})^2$ is the strength of the leptoquark coupling, l and q are the lepton and quark doublets, and $H.C.$ indicates the hermitian conjugate of the preceding term. The coupling λ^{ij} can be expressed in terms of the known electromagnetic coupling α_{em} by

$$(\lambda^{ij})^2/4\pi = k\alpha_{em} \tag{3.2}$$

where k is some unknown scaling factor. The existence of leptoquarks would lead to an s channel contribution in the process $e^+e^- \to q\overline{q}$ in addition to the exchange of γ and Z bosons. By requiring that the $q\overline{q}$ production cross section and the forward-backward asymmetry not deviate by more than 10% from the SM expectations, Hewett and Rizzo [30] have placed a limit on the value of $m_{LQ}(k)$. Leptoquarks with masses less than 150 GeV/c^2 are ruled out for $k > 0.5$ while for $k > 5$ leptoquarks with masses less than 450 GeV/c^2 are excluded.

Flavor-Changing Neutral Currents (FCNC) are tightly constrained by experimental data [29]. Limits on the rare meson decays such as $\pi^{\pm} \to e\nu$, $D^0 \to \mu^+\mu^-$ or $K^{\pm} \to e\nu$ are consistent with relatively small leptoquark masses if the leptoquarks are assumed to have chiral couplings. That is, a leptoquark could have either left-handed ($\lambda_1 = -\lambda_2$) or right-handed ($\lambda_1 = \lambda_2$) couplings. In addition, leptoquarks would have to satisfy two further conditions: 1) leptoquarks have three distinct generations as do the leptons and quarks, 2) leptoquarks couple only within the generation (so that

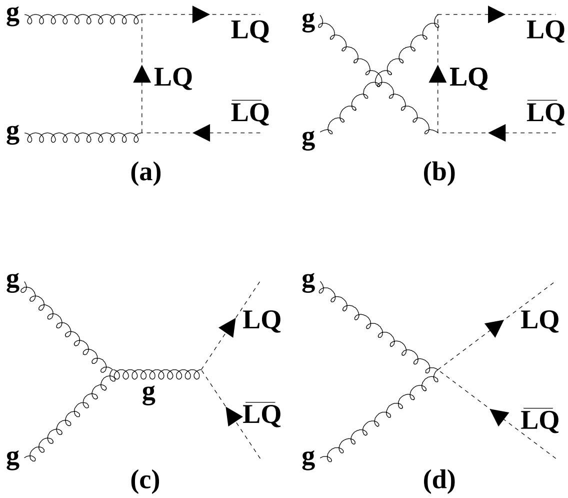

Figure 3.1. Feynman diagrams for the gluon-gluon sub-processes.

a first generation leptoquark would couple only to u and d quarks and e and ν_e leptons, and similarly for the other generations). If these conditions are satisfied, leptoquarks with masses in the currently accessible range of energies could exist within the experimental constraints.

In the search described herein we shall consider only leptoquarks which have dimensionless couplings to fermions and which conserve baryon and lepton numbers. Additionally, only leptoquarks which couple within a single generation and which are $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ invariant are considered. Leptoquarks with spin = 0 (scalar leptoquarks, S_{LQ}) and spin = 1 (vector leptoquarks, V_{LQ}) are considered. Finally, only leptoquarks with electric charge $-\frac{1}{3}$ or $\frac{2}{3}$ are considered.

3.4 Leptoquark Pair Production

At hadron colliders such as the Tevatron leptoquarks may be produced in pairs ($LQ\overline{LQ}$) in strong interactions via gluon-gluon fusion, $q\overline{q}$ annihilation or gluon-quark interactions. The leptoquark pair production is then a strong interaction (proceeding via the strong nuclear force). The dominant mode is gluon-gluon fusion [31–33].

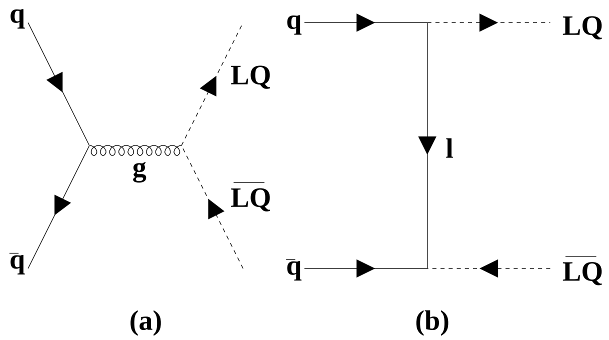

Figure 3.2. Feynman diagrams for the quark-quark sub-processes.

The Feynman diagrams for leptoquark pair production via gluon-gluon (gg) fusion are shown in fig. 3.1. The cross section (to leading order for S_{LQ}) is given by

$$\sigma_{gg}(gg \to LQ\overline{LQ}) = \frac{\pi\alpha_s^2}{6\hat{s}} \left[\left(\frac{5}{8} + \frac{31m_{LQ}^2}{4\hat{s}} \right) \varepsilon + \left(4 + \frac{m_{LQ}^2}{\hat{s}} \right) \frac{m_{LQ}^2}{\hat{s}} \ln\left(\frac{1-\varepsilon}{1+\varepsilon} \right) \right] \tag{3.3}$$

where α_s is the strong coupling constant, m_{LQ} is the mass of the leptoquark, $\sqrt{\hat{s}}$ is the center-of-mass energy of the parton collision, and $\varepsilon = \sqrt{1 - \frac{4m_{LQ}^2}{\hat{s}}}$.

The Feynman diagrams for the $LQ\overline{LQ}$ production via $q\overline{q}$ scattering are shown in fig. 3.2. The cross section is

$$\sigma_{q\overline{q}}(q\overline{q} \to LQ\overline{LQ}) = \frac{2\pi\alpha_s^2}{27\hat{s}}. \tag{3.4}$$

The contribution to the cross section from fig. 3.2 (b) is

$$\sigma'_{q\overline{q}}(q\overline{q} \to LQ\overline{LQ}) = \frac{2\pi\lambda_{LQ}^2}{9\hat{s}} \left[\left(1 - \frac{2m_{LQ}^2}{\hat{s}} \right) \ln \frac{s(1+\varepsilon) - 2m_{LQ}^2}{\hat{s}(1-\varepsilon) - 2m_{LQ}^2} - 2\varepsilon \right] \tag{3.5}$$

where λ_{LQ} is the unknown leptoquark coupling to the lepton-quark pair. The cross sections in *eqns.* 3.4 and 3.5 are both leading order for S_{LQ}. Since λ_{LQ} is much smaller than the strong coupling α_s, the contribution from *eqn.* 3.5 to the total cross section is insignificant and is ignored in the calculation of the total leading order (LO) scalar leptoquark cross section. For vector leptoquarks the diagrams are the same, but the cross sections are substantially larger due to the spin interactions.

The total LO cross-section for S_{LQ} is obtained by summing the individual sub-process cross-sections with the structure functions for the hadrons [34]. In the case of the Tevatron, the hadrons being collided are protons and anti-protons so we have

$$\begin{aligned}\sigma_{LO}(p\overline{p} \to LQ\overline{LQ}) = \sum_{a,b} \int_0^1 d(x_a x_b) \int_\tau^1 \frac{dx_a}{x_a} \\ \left[f_{a/A}(x_a) f_{b/B}(x_b) + (A \leftrightarrow B \text{ if } a \neq b) \right] \hat{\sigma}(\hat{s})\end{aligned} \tag{3.6}$$

where $\tau = x_a x_b$, $\hat{s} = x_a x_b s = \tau s$, A(B) represent the four-momentum of the proton (antiproton), a(b) the parton four-momentum, x_a and x_b are defined by a = x_aA, b = x_bB, and are the fraction of the total momentum of the proton(antiproton) carried by the parton a(b). $\sqrt{s}$ is the center-of-mass energy of the collision. The total cross section is then

$$\sigma(p\overline{p} \to LQ\overline{LQ}) = \sum_{i=1}^{3} \int_0^1 d\tau \int_\tau^1 \frac{dx}{x_a} [f_i(x,\hat{s}) f_i(\tau/x,\hat{s})] \sigma(\hat{s}) \tag{3.7}$$

where

$$f_1(x,\hat{s}) f_1(\tau/x,\hat{s}) = g(x,\hat{s}) g(\tau/x,\hat{s}) \tag{3.8}$$

$$\begin{aligned}f_2(x,\hat{s}) f_2(\tau/x,\hat{s}) = u(x,\hat{s}) u(\tau/x,\hat{s} + d(x,\hat{s}) d(\tau/x,\hat{s}) + \\ 2st(x,\hat{s}) st(\tau/x,\hat{s}) + 2sea(x,\hat{s}) sea(\tau/x,\hat{s})\end{aligned} \tag{3.9}$$

$$\begin{aligned} f_3(x,\hat{s})f_3(\tau/x,\hat{s}) = {} & u(x,\hat{s})u(\tau/x,\hat{s} + d(x,\hat{s})d(\tau/x,\hat{s}) + \\ & u(x,\hat{s})d(\tau/x,\hat{s}) + d(x,\hat{s})u(\tau/x,\hat{s}) + \\ & 4sea(x,\hat{s})sea(\tau/x,\hat{s}) \end{aligned} \tag{3.10}$$

with $g(x,\hat{s}), u(x,\hat{s}), d(x,\hat{s}), st(x,\hat{s}), \text{and}, sea(x,\hat{s})$ are *gluon, up, down, strange, charm*, and *sea quark* distributions for the protons. $u(x,\hat{s})$ and $d(x,\hat{s})$ are the sum of the *valence* and *sea* distributions. Equation 3.7 was integrated numerically using the CTEQ4M parton distribution functions. The cross section for leptoquark pair production at the Tevatron is shown in fig 3.3 for various couplings. On the plot, YM indicates the cross section for Yang-Mills vector couplings, MC the anomalous minimal vector couplings, and "min" an anomalous coupling ($\kappa_G = 1.3, \lambda_G = -0.2$) chosen to minimize the vector cross section. We shall not consider the "min" coupling further. Also shown on the plot for comparison is the LO cross section for scalar Yang-Mills couplings. The cross sections shown in the figure are the ones that are used in this work for vector leptoquarks. For scalar leptoquark pair production we shall use next-to-leading order (NLO) cross sections.

3.5 Next to Leading Order Cross Sections

For next to leading order calculations ($O(\alpha_s^3)$) the cross section dependence on the factorization scale becomes larger, increasing the uncertainty in the calculation. This is a peculiarity of the model, and not generally true. If one were to perform higher order calculations, the dependance on the factorization scale (denoted as μ) should become small. The factorization scale is a free parameter introduced into the theory, which may be thought of as the transition between large and small distance interactions.

At next to leading order, additional processes contribute to the total cross section. These are the quark-gluon subprocesses, shown in fig. 3.4. The Feynman

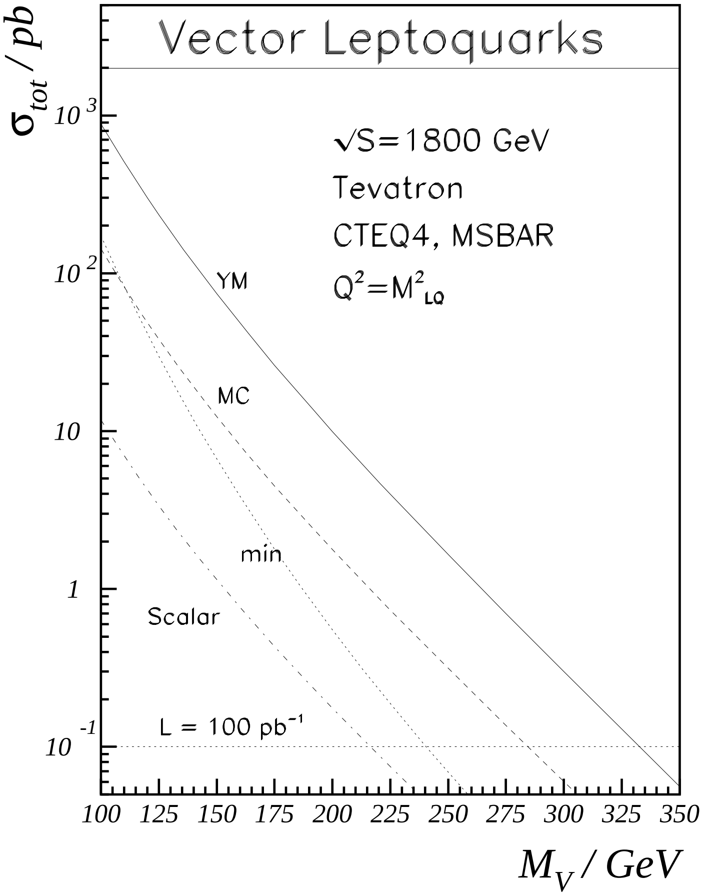

Figure 3.3. Leading order cross sections for LQ pair production at the Tevatron.

diagrams shown in fig. 3.4 represent additional contributions to the cross section over the leading order calculation. The NLO cross sections are larger than the LO cross sections. In fact, the NLO cross sections represent approximately a 15% increase over leading order. The NLO cross section was calculated using CTEQ4M parton distribution functions for $n_f = 5$ active quark flavors with an energy scale $\Lambda_{QCD} = 202 MeV$.

Vector leptoquark cross sections have not yet been calculated at next-to-leading order. Rather, we use the LO cross section evaluated at a scale $Q^2 = m^2_{VLQ}$. The previous LO cross sections were done at a scale $Q^2 = 4m^2_{VLQ}$. Changing the scale

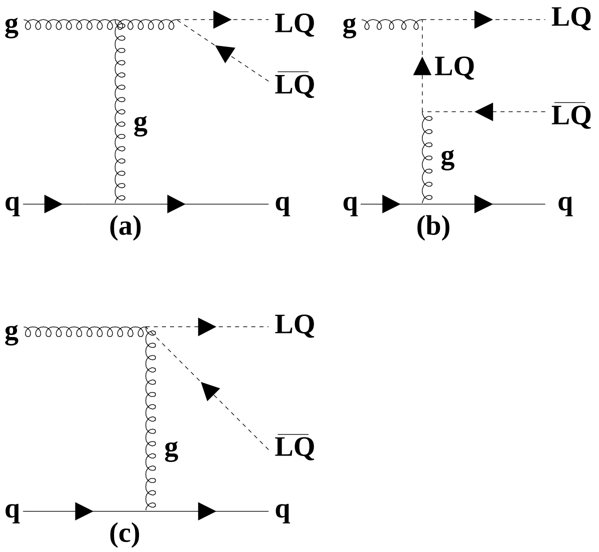

Figure 3.4. Feynman diagrams for the quark-gluon subprocesses.

in this fashion increases the LO cross section. For the scalar leptoquarks the LO cross section then becomes a good approximation of the NLO cross section (σ_{NLO}) minus the uncertainty in σ_{NLO} (*i.e.* the minimum of the NLO cross section). We have assumed that this will hold true for the vector couplings as well. Since the lower edge of the NLO cross section is what we shall use to set the mass limits in this work, this is a good approximation to use.

3.6 Search Signatures

The search described in this work is for second generation leptoquarks. We have assumed that leptoquarks couple only to a single generation due to experimental constraints on flavor-changing neutral currents [29]. We also assume that the leptoquark-lepton-quark coupling is sufficiently large ($\lambda_{LQ} > 10^{-12}$) so that the leptoquark would decay quickly enough to be seen in the DØ detector. The leptoquark would then decay to second generation leptons (μ, ν_μ) and quarks (c, s). The production and decay chain would therefore be

$$p\overline{p} \to LQ\overline{LQ} \tag{3.11}$$

with a subsequent leptoquark decay

$$\begin{aligned} LQ &\to \mu^- c \text{ or } \nu_\mu s \; (q_{LQ} = -\frac{1}{3}) \\ LQ &\to \mu^+ s \text{ or } \nu_\mu c \; (q_{LQ} = +\frac{2}{3}) \end{aligned} \tag{3.12}$$

Should both leptoquarks decay to a muon and a quark the signature would be $\mu\mu + jets$ (recall the description of hadronization from section 2-2). Neutrinos are not detected with the DØ detector, they are inferred from conservation of momentum (more detail will be presented in subsequent chapters). Neutrino signatures in the detector are therefore "missing" transverse energy (written as $\not{E}_T$, transverse indicates that the momentum of the particle, and thus the energy deposited in the DØ detector, are measured perpendicular to the direction of the $p\overline{p}$ beams). So a decay with one leptoquark going to a muon and a quark while the other decays to a neutrino and a quark will have the event signature $\mu + \not{E}_T + jets$ in the DØ detector. The final event signature, $\not{E}_T + jets$, is not dealt with directly in this work. We shall use previous searches in this channel when we discuss combining limits (chapter 7), but no direct search in this decay channel was conducted for this work.

3.7 Current Limits

Much theoretical and experimental work has been done since the first predictions of leptoquarks. Limits on the mass of leptoquarks have been published by both the DØ [35] and CDF [36] collaborations. Limits are established for different values of the possible branching ratios of leptoquarks into muons or neutrinos. A branching ratio is the fraction of the decay into a particular final state. So, if the leptoquark were to decay into a muon 70% of the time and into a neutrino 30% of the time, the branching ratios into muons(neutrinos) would be 0.7(0.3). The branching ratio of the leptoquarks is, of course, unknown. It is parameterized by β, the branching ratio into charged leptons (muons, for this work). So the branching ratio into neutrinos is then $(1-\beta)$. Results are generally published for the cases $\beta = 1, \frac{1}{2}, 0$ corresponding to decay only to muons ($\beta = 1$), equal decay into muons and neutrinos ($\beta = \frac{1}{2}$) and complete decay into neutrinos ($\beta = 0$). This is done both for easier comparison between experiments and because, in theoretical work, these are the branching ratios allowed by the chiral coupling of the leptoquarks to the neutrinos.

The limits in existence at the start of this work were all obtained from the study of the $\mu\mu + jets$ decay mode for scalar leptoquark couplings. The branching ratio is then β^2. The DØ limits are 184 GeV/c^2($\beta = 1$) and 140 GeV/c^2($\beta = \frac{1}{2}$). Current CDF limits are 202 GeV/c^2($\beta = 1$) and 160 GeV/c^2($\beta = \frac{1}{2}$). Here, GeV/c^2 is the unit of mass used. GeV is giga (10^9) electron volts divided by the square speed of light c (from the relativistic relation $E = mc^2$). 1 GeV/c^2 = 1.78 $\times 10^{-27}$ kg.

In this work we shall present a new search of the $\mu\mu + jets$ decay channel which extends the previous DØ search, as well as a search of the $\mu + \not\!\!E_T + jets$ decay channel. We shall present results for both scalar and vector couplings, as well as limits obtained by combining the searches for the various leptoquark decay signatures. Finally, we shall present exclusion contours in the β vs. m_{LQ} plane which exclude certain combinations of the parameters for leptoquarks.

Since any experiment is dependent upon the apparatus used, in the next chapter we shall present a description of the Tevatron accelerator and the DØ detector which were used to collect the data which have been analyzed for this work.

CHAPTER 4

THE DØ EXPERIMENT

> "Although my father could spot the problem with any experimental apparatus, it was usually best not to let him touch it" J.J. Thompson, Nobel Laureate, speaking of his father, also a Nobel Laureate.

4.1 Overview

The DØ experiment, located at Fermi National Accelerator Laboratory (Fermilab), in Batavia, IL, was proposed to study proton-antiproton ($p\overline{p}$) collisions with a center-of-mass energy $\sqrt{s} = 1.8$ TeV. The detector is a large (taller than three stories, weighing more than 5,000 tons) complex apparatus designed to investigate high transverse momentum (p_T) phenomena in hadronic collisions. These phenomena include the observation and measurement of the top quark, precision measurement of the W and Z bosons, production and study of b-quark hadrons, tests of perturbative QCD, and searches for new phenomena beyond the Standard Model. This chapter will describe the experimental setup of the Tevatron collider (relying heavily on reference 37) and the DØ detector (relying on reference 38), with particular attention to those systems which will weigh most heavily in the chapters that follow. We are particularly indebted to the work in references 17 and 39–41 for their excellent summaries of the experimental apparatus.

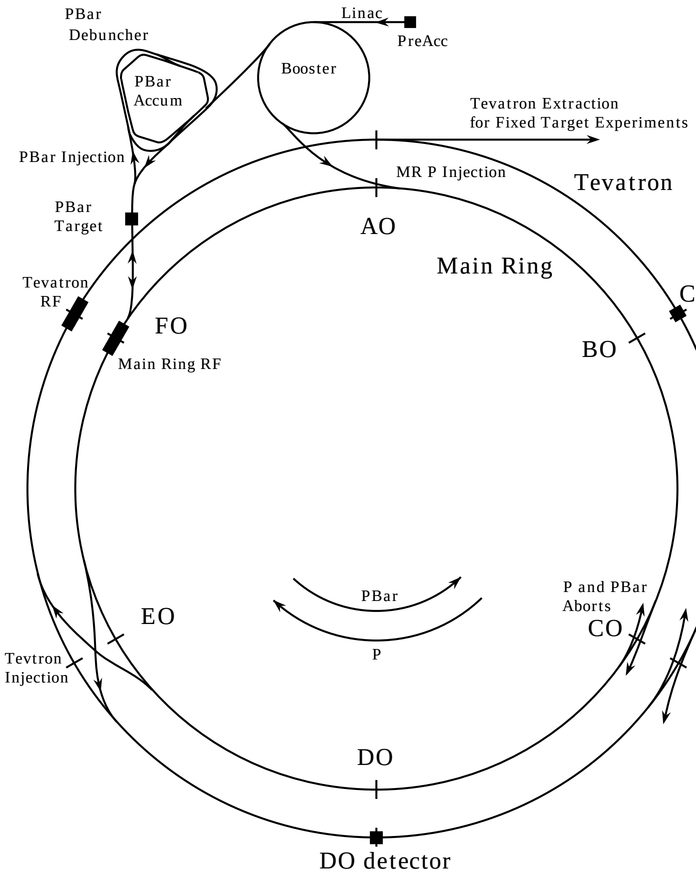

Figure 4.1. Schematic diagram of the Fermilab accelerator system.

4.2 The Fermilab Tevatron Collider

The Fermilab Tevatron [37], shown in fig. 4.1, accelerates protons and antiprotons to energies of 900 GeV. With these energies, the center-of-mass collision energy of a proton-antiproton pair is 1800 GeV, or 1.8 TeV. This makes the Tevatron the worlds most powerful particle accelerator. The acceleration is done in several stages using five different accelerators, of which the Tevatron is the final stage.

The first stage begins by accelerating negative hydrogen ions to an energy of 750 KeV using an electrostatic field produced by a Cockcroft-Walton generator. The accelerated ions are fed into a 400 MeV linear accelerator. The beam is passed through a thin carbon target which strips the electrons off of the hydrogen ions. The resulting beam is then composed of protons (positive hydrogen ions). This beam is accelerated to an energy of 8 GeV in the booster ring, and then to 120 GeV in the Main Ring. The Main Ring accelerator has a circumference of 3.7 miles, and functions as an injector for the Tevatron and a proton source for antiproton production.

Antiprotons are created by focusing the 120 GeV proton beam onto a copper/nickel target. Roughly 1 antiproton with an energy of 8 GeV is produced for every 100,000 protons incident upon the target. The antiprotons are selected by momentum, with those antiprotons having a large angular dispersion (or, equivalently, a large transverse momentum) deselected. The remaining antiprotons are focused into a beam and sent to the accumulator. The accumulator stores the antiprotons for several hours until approximately 10^{11} antiprotons have been collected. Within the antiproton accumulator the antiprotons are stochastically cooled [42]. Stochastic cooling is the process by which the antiproton beam is made orderly. The antiprotons leave the target with a wide range of energies, positions, and momenta. The randomness of the beam is equivalent to a high temperature. Because of this randomness the antiproton beam will not fit into the beam-pipe and the odds of a collision with a proton are very low. To cool the beam the antiprotons are sent into a storage ring with a circular orbit where a series of pickups and "kickers" alter the path of the particles. The pickups detect an "error signal" (the position or energy of the antiprotons) and the signal is amplified, inverted and sent to the kicker on the other side of the ring. The kicker then applies the signal received. This process tends to damp out the random motion of the antiprotons, resulting in a tight, well ordered beam which is suitable for injection into the next accelerator.

Protons and antiprotons are injected into the Main Ring in opposite directions (counter-rotating beams) and accelerated to 150 GeV before final injection into the

Tevatron. The Tevatron uses a series of super-conducting magnets to produce a 4 Tesla magnetic field to keep the particles in a circular orbit within the beam pipe. Electrostatic separators are used to prevent collisions between the two beams outside of the desired interaction regions inside the two detectors (DØ and CDF). At the desired interaction points, quadrapole magnets focus the beams for collision. The beams in the accelerator are grouped into six bunches of protons/antiprotons, and the beams cross the interactions regions with a period of 3.5 μs. A collider run can last up to 20 hours, during which time the antiproton accumulator continues storing antiprotons for the next injection.

The proton-antiproton flux is compared to the rate of $p\overline{p}$ collisions to compute the instantaneous luminosity ($\mathcal{L}$) of the collisions. It is a function of the density of the individual bunches and the periodicity of the crossing. During an individual run the luminosity decays exponentially from its initial value as the numbers of protons and antiprotons within each beam decline. During the course of the running period from which the data in this work is taken (1993 – 1996), the luminosity ranged from $0.2 - 20 \times 10^{30} cm^{-2} s^{-1}$. Luminositys (and cross sections) are typically measured with units related to "barns", 1 $barn = 10^{-24} cm^2$, where the factor of $10^{-24} cm^2$ comes from the square of the approximate radius of the proton $10^{-12} cm$. Luminosity is then in units of inverse barns. A luminosity of $1 \mu b^{-1} s^{-1}$ translates to a flux of 10^{30} particles per square centimeter per second.

4.3 Detector Coordinate Systems

Particle tracks in the DØ detector are described in up to four coordinate systems. The primary system used in this work will be a modified cylindrical coordinate system using the coordinates (z, η, ϕ). This system will be derived from other, more conventional systems in this section.

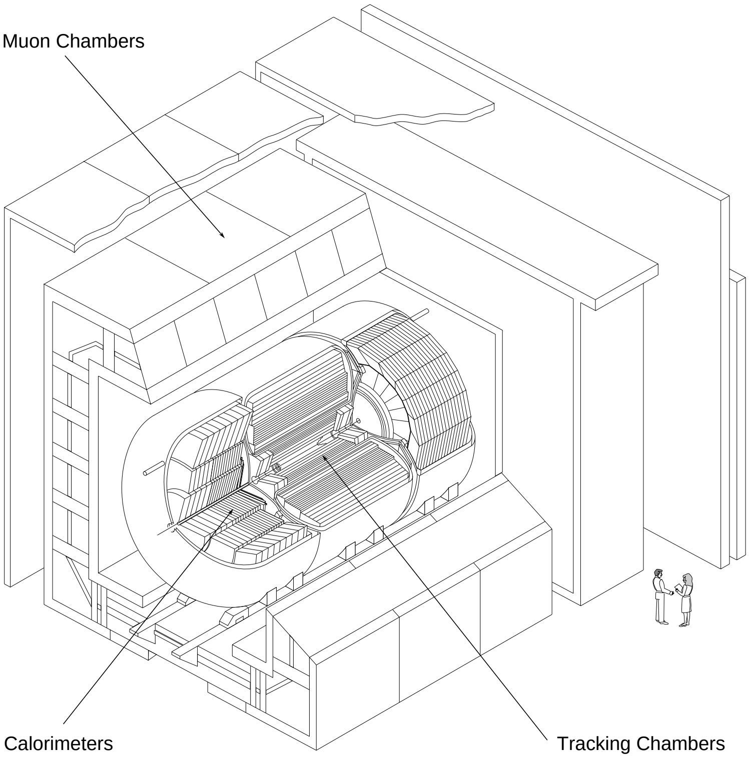

Figure 4.2. Cutaway view of the DØ detector.

We start from the simple Cartesian rectangular coordinate system. The beamline is taken as the z -axis, with the origin at the center of the DØ detector. The $x - y$ plane is then a slice of the detector itself. The detector, shown in fig. 4.2, is a cylinder whose axis is coincident with the beamline. In the lower right hand corner of fig. 4.2 human figures are shown for scale. The positive $z - axis$ is taken along the direction of the proton beam (to the left in fig. 4.2), the positive $y - axis$ is toward the top of the figure, and the positive $x - axis$ points radially outward. A cylindrical coordinate system may be derived from the rectangular coordinates by constructing a radial vector $r = +\sqrt{x^2 + y^2}$, and $\phi = \tan^{-1}(\frac{x}{y})$. We then have a coordinate system (r, ϕ, z). What is actually of interest in the detector, however, is not a particular point within the detector volume, but a particle track through the detector. We thus change to a modified cylindrical coordinate system (z, ϕ, θ). ϕ is the azimuthal angle of the track and θ the polar angle, or the angle between a track and the $z - axis$.

The final modification is to use the variable η instead of θ, where η is known as the psuedorapidity. One defines the quantity rapidity as

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z} \tag{4.1}$$

where E is the total energy and p_z the longitudinal component of the momentum of the particles. The advantage of rapidity is that it is invariant under a Lorentz transformation. This is advantageous since a longitudinal boost of the system will not affect the shape of the distribution in y. For energies at which the momentum is much greater than the particle mass (such as the Tevatron), $E \approx p$ and one may write

$$y \approx \frac{1}{2} \ln \frac{1 + E_z/E}{1 - E_z/E} \tag{4.2}$$

whereupon, using $\cos\theta = E_z/E$, the rapidity becomes

$$y \approx \frac{1}{2} \ln \frac{\cos^2(\theta/2)}{\sin^2(\theta/2)} \equiv \eta_d \tag{4.3}$$

The quantity η_d is known as the pseudorapidity. For massless particles $\eta_d = y$. The coordinates centered in the detector (z_d, η_d, ϕ_d) are known as the detector coordinates. Event vertices are not typically in the center of the detector, and the coordinates with their origin at the vertex of the event are known as the physics coordinates (z, η, ϕ). This is the coordinate system we shall use in this work.

Particle energy and momentum are most usually described with quantities taken transverse to the beamline. This is because there are no detectors along the beamline which could detect the passage of a particle traveling along the z axis at large η. The transverse energy(momentum) are defined by $E_T = E \sin\theta (p_T = p \sin\theta)$.

4.4 The DØ Detector

The DØ detector [38] is a large general purpose detector constructed for the identification and measurement of electrons, muons, and high p_T jets. The detector weighs 5500 tons, stands 13 meters in height and 20 meters in length. The detector may be divided into three broad systems: the central tracking system, the calorimeter, and the muon spectrometer. The sections which follow will describe each of the systems with a more detailed discussion for the calorimeter and the muon spectrometer, which are of more interest to this work.

4.5 Central Tracking

At the center of the DØ detector is the Central Tracking System. Consisting of four detector subsystems – the Vertex Drift Chamber (VTX), the Transition Radiation Detector (TRD), the Central Drift Chamber (CDC), and two Forward Drift Chambers (FDC) – the tracking system is positioned radially about the beamline with an inner radius of 3.7 cm and an outer radius of 78 cm. The central tracking system is shown in fig. 4.3.

The VTX, CDC, and FDC all use high electric fields inside an ionizable gas medium to measure the trajectories of charged particles. High energy charged particles from the beam collision ionize the gas as they pass through each chamber. Drift times of the ionization electrons to an anode wire are used to measure the spatial position of the original ionizing particle. By combining signals from each of the detectors the particle trajectories are reconstructed. Since DØ does not have a central magnetic field the reconstructed tracks are straight and contain no information about the particle momentum or charge.

The VTX is the innermost tracking chamber, consisting of three concentric layers of drift chambers designed to accurately measure the longitudinal position of the

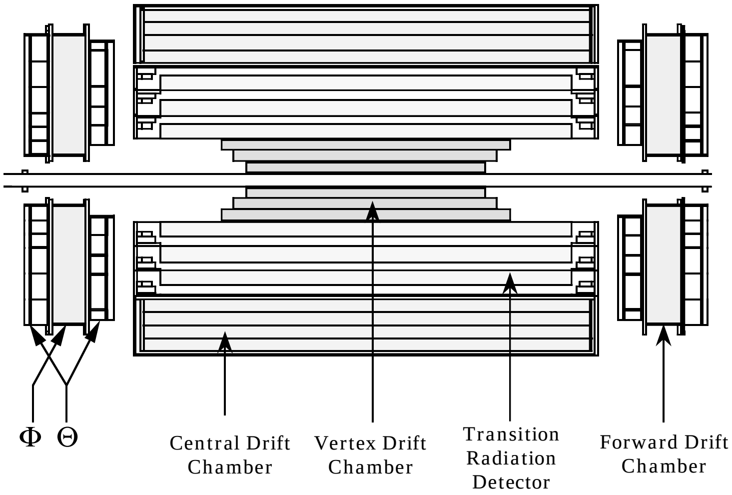

Figure 4.3. Components of the central tracking system

event vertex (the place where the $p\overline{p}$ collision happened). The VTX spans a range $|\eta| < 2$ in pseudorapidity. The resolution of the $r - \phi$ position is about 60 μm. The z position resolution is about 1.5 cm.

The TRD is used to discriminate between low mass objects (such as electrons) and higher mass objects (pions or other hadrons). The detector consists of several layers of different media. Particles passing through the detector produce radiation at the boundaries of these media which is inversely proportional to the particle mass. So low mass particles will produce more radiation in the TRD than will high mass particles of the same momentum.

The CDC surrounds the VTX and TRD and measures charged particle tracks within the range $|\eta| < 1.2$. It consists of four concentric layers of drift chamber modules, with each layer containing 32 azimuthal modules. The resolution of the $r - \phi$ position is about 180 μm, while the z position is resolved to about 3 mm.

The FDC, shown in fig. 4.4, is positioned on each end of the central tracking chambers and covers a range $1.0 < |\eta| < 3.2$. Each FDC consists of three layers of

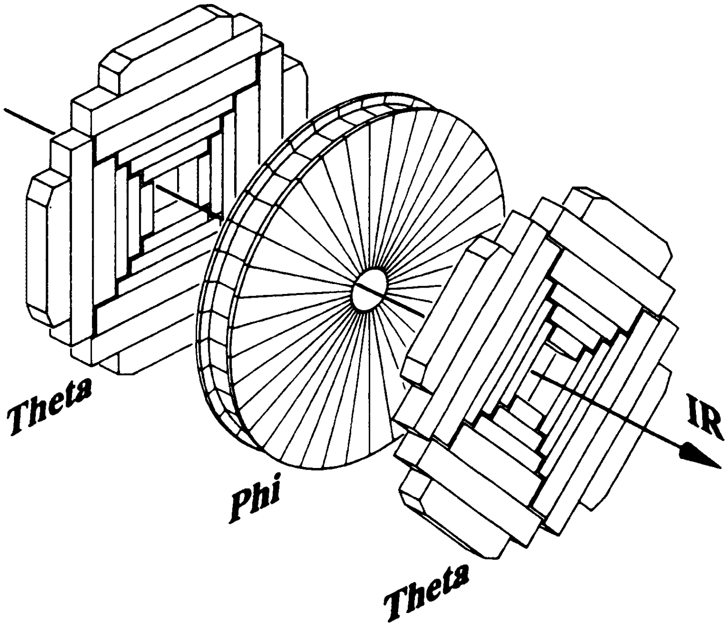

Figure 4.4. Exploded view of the forward drift chamber modules.

drift chamber modules and provides an $r-\phi$ resolution of 200 μm and a z resolution of 4 mm.

4.6 The Calorimeter

The DØ calorimeter is designed to measure the energies of particles produced in hadronic collisions. It provides kinematic information about electrons, photons and jets. Neutrinos are not detected with the DØ detector. Rather, their presence is inferred from conservation of momentum in the plane transverse to the beam. The energy deposited in the calorimeter is summed cell by cell during reconstruction. This sum is corrected for the detector geometry, and the quantity $\not{E}_T$ is calculated from the sum as the momentum vector which balances the energies deposited in the calorimeter. This $\not{E}_T$ could be due to a muon, which is a minimum ionizing particle and deposits small amounts of energy in the calorimeter, or from a neutrino, which interacts weakly (interactions proceed only via the weak nuclear force) and deposits no energy in the calorimeter. $\not\!\!E_T$ could also be produced due to mismeasurement of the energy in the calorimeter or from the passage of particles which did not originate from the $p\overline{p}$ collision (such as muons from cosmic rays). These parts will play a central role in the work described in the following chapters.

DØ has a sampling calorimeter. That is, the calorimeter uses a series of measurements to sample the energy produced by showers of particles in the calorimeter. The energy of the original particle is then reconstructed from the energies of the shower produced in the uranium plates. Figure 4.4 shows the DØ calorimeter and its major subsystems. In the figure the three individual cryostats, one central and two endcap, are shown and labeled.

The calorimeter is constructed of alternating layers of depleted uranium and liquid argon. The uranium constitutes a dense absorbing material and the liquid argon an ionizing medium. Particles traversing the calorimeter interact and produce secondary particles and also lose energy in the uranium plates. When passing through the uranium the particle will interact to produce low energy secondary particles (a process known as showering). These particles then ionize the liquid argon which measures the number of particles passing through it. The total number of particles is directly proportional to the incident energy. Thus the fractional energy and location (η, ϕ) may be measured.

High energy electrons and photons interact primarily with the uranium plates in complementary processes: pair production and bremsstrahlung. Pair production is $\gamma \rightarrow e^+e^-$. Bremsstrahlung is the process $e \rightarrow e\gamma$. With each interaction the number of particles (particle multiplicity) increases and the average energy per particle decreases. These are known as *electromagnetic* (EM) showers. The particle multiplicity increases exponentially so that the depth of the shower depends

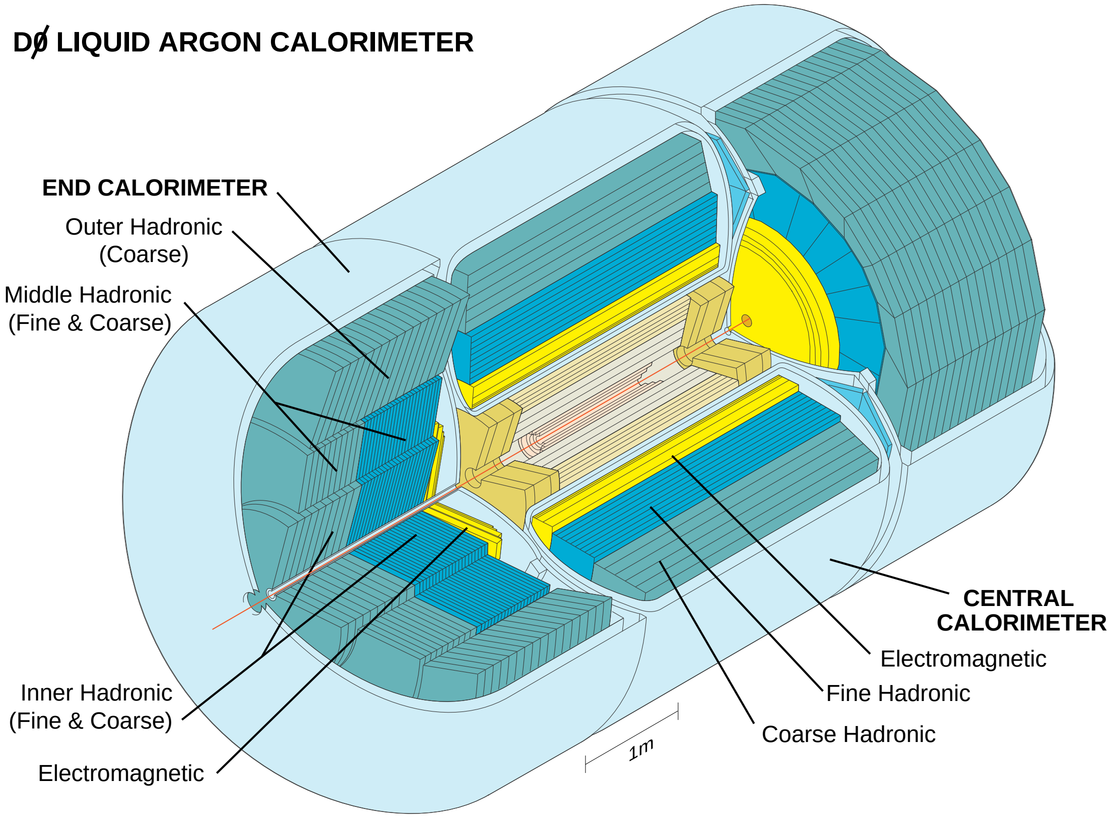

Figure 4.5. The DØ calorimeter.

logarithmically on the energy of the original particle. The probability that an electromagnetic particle will traverse a distance d without initiating a shower is given as

$$E(d) = E_0 e^{-\frac{d}{x_0}} \tag{4.4}$$

where d is the distance traveled before showering, E_0 is the initial energy of the particle, and x_0 is the "radiation length", which is characteristic of the medium. For uranium $x_0 \approx 3.2$ mm.

In contrast to electromagnetic particles, high-energy hadrons interact with the uranium nuclei with inelastic collisions proceeding via the strong nuclear force.

These collisions produce secondary particles, about $\frac{1}{3}$ of which are π^0's. The rest of the secondary particles also interact strongly, while the π^0's produce electrons and photons which interact electromagnetically. This type of shower is known as a "hadronic" shower, and develops more slowly (over longer distances) than electromagnetic showers. The probability of a hadron traversing a distance d without initiating a shower is also given by equation 4.6, but in this case $x_0 = \lambda_I \approx 10.5$ cm. λ_I is the "nuclear absorption length".

With EM showers developing over shorter distances than hadronic showers, the first several layers of the calorimeter are closely spaced and designed to measure accurately the energies of electromagnetic particles. The outer layers of the calorimeter use larger spacing for the measurement of hadronic showers. The hadronic calorimeter has two sections: a fine hadronic (FH) layer providing good hadronic shower energy resolution and a coarse hadronic (CH) layer used to contain hadronic showers so that they do not proceed into the muon spectrometer (a process known as "punchthrough").

The calorimeter cells are structured in layers. The layers are aligned in towers projecting back toward the interaction region (the lower left-hand corner of the diagram in fig. 4.6). The cell layer sizes are determined by the geometry and the shower depths. The EM layers are approximately 1-2 cm wide. Hadronic layers are approximately 10 cm wide. Each calorimeter layer is typically 0.1×0.1 in $\eta \times \phi$. This defines the resolution of the calorimeter. Fig. 4.6 shows a partial view of the calorimeter with the projective cell structure for successive layers. The numbers around the edge of the figure show the pseudorapidity of the lines. The liquid argon in the calorimeter is kept at a temperature of 78 K, which requires a stainless steel cryostat housing. The calorimeter is divided into three sections to facilitate access

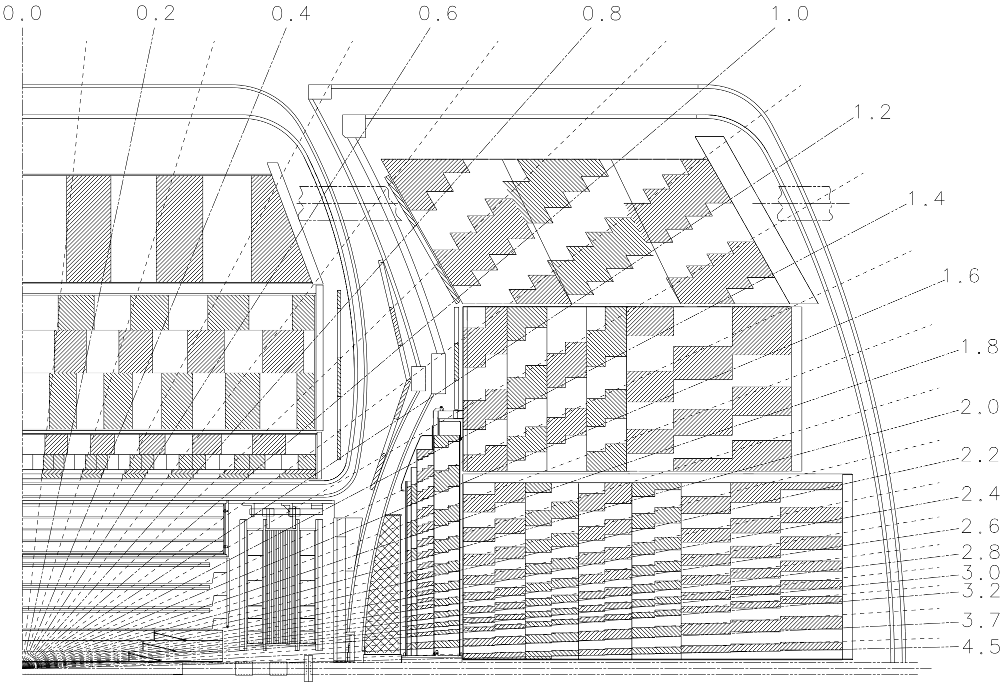

Figure 4.6. Side view of one quarter of the calorimeter.

to the tracking system. There is a central calorimeter (CC) cryostat located between two endcap calorimeter (EC) cryostats.

4.6.1 The Central Calorimeter

The Central Calorimeter (CC) is a cylinder providing coverage out to $|\eta| \leq 1.0$. The readout layers are stacked in r. The first four layers of the CC (counting outward from the beampipe) provide measurement of the EM showers at depths of 2, 4, 11, and 21 radiation lengths (x_0). Taken together these layers comprise 0.76 nuclear absorption lengths. The first, second and fourth layers have a cell segmentation of 0.1×0.1 in $\eta \times \phi$. The maximum of the shower development occurs in the third layer, so the segmentation in that layer is increased to 0.05×0.05 for a more precise measurement of the location and shape of the shower.

Beyond the EM layers are the three Fine Hadronic (FH) layers at depths of 1.3, 1.0 and 0.9 nuclear absorption lengths. All three layers have a cell segmentation of 0.1×0.1. After the FH is the Coarse Hadronic (CH) layer, providing a single readout layer having a depth of 3.2 nuclear absorption lengths.

4.6.2 The Endcap Calorimeters

The Endcap Calorimeters (EC) lie on either side of the CC, extending coverage out to $1.0 \leq |\eta| \leq 4.0$. The EM readout layers have a thickness 0.3, 2.6 7.9 and 9.3 radiation lengths comprising about 0.75 nuclear absorption lengths. For $|\eta| \leq 2.6$, the cell segmentation is the same as in the CC. For $2.6 < |\eta| \leq 3.2$ the segmentation in the third layer is decreased to 0.1×0.1. For $|\eta| > 3.2$, segmentation in all layers is decreased to 0.2×0.2 and continues to decrease until it is 0.4×0.4 for $|\eta| = 4.0$.

In the EC are three hadronic modules. Closest to the beam pipe is the inner hadronic module consisting of four fine hadronic (IFH) readout layers and one coarse hadronic (ICH) readout layer. The middle hadronic module surrounds the inner module in ϕ and has four fine hadronic (MFH) layers and a single coarse hadronic (MCH) layer. The outermost module is the outer hadronic (OH) module consisting of three coarse hadronic layers. In the range $0.7 < |\eta| < 1.1$ the EM and FH calorimeters are in the CC while the CH calorimeter is in the EC.

4.6.3 The Inner Cryostat Detectors

Due to the structure and support system of the calorimeter cryostats there exists a gap in the coverage between the Central and Endcap Calorimeters. The gaps span approximately $0.8 \leq |\eta| \leq 1.4$, with the result that there is only partial instrumentation of the EM and FH sections. This partial coverage creates a lack of uniformity in the energy response and acceptance within this region. The region has a substantial amount of absorption material with no energy sampling in the first radiation length. To supplement coverage in the region the Inner Cryostat Detector (ICD) is used. The ICD consists of scintillation counter arrays positioned on the inner walls of the EC cryostat to provide energy sampling in this region. Additional coverage is provided by the Massless Gaps (MG). These are detectors consisting of a readout cell having a signal board embedded in liquid argon positioned inside both the CC and EC cryostat walls to provide full coverage in ϕ.

4.6.4 Calorimeter Calibration

Prior to reaching the final stages of construction and assembly, the DØ calorimeter was tested and calibrated at the DØ test beam [43]. Single sets of modules from the various subsystems were installed in the test beam before assembly in the detector. Using beams of electrons and pions having known energies ranging from 2-150 GeV studies indicated that the response of the calorimeter is linear for electrons with energies greater than 10 GeV and pions greater than 20 GeV. For low energy particles the response is nonlinear which results in lower detection efficiencies and jet energy response. At low energies noise degrades signal resolution, while at higher energies resolutions are dominated by inhomogeneities and calibration uncertainties. Calorimeter resolutions have been measured to be $15\%/\sqrt{E}$ for electrons, $50\%/\sqrt{E}$ for pions, and $80\%/\sqrt{E}$ for jets, where E is in GeV.

The larger resolution uncertainty in the jet energy comes from several sources. First, jets are typically made up of a large number of low-energy particles for which the detector resolution is lower. Second, the calorimeter response is non-linear for the low-energy particles and therefore for the jets as well. So the jet energies must be corrected during event reconstruction by applying a “jet energy scale” correction,

Table 4.1. Typical calorimeter noise widths by section and cryostat.

Section	CC Pedestal Width (MeV)	EC Pedestal Width (MeV)
EM	10	10
FH	60	30
CH	75	55
ICD	8	15
MG	30	30

increasing the uncertainty of the jet energy. Finally, one must define what is to be included in the jet. DØ uses a cone algorithm $\delta\mathcal{R} \equiv \sqrt{\delta\eta^2 + \delta\phi^2}$. Jet cones are typically 0.3, 0.5, and 0.7. In this work we shall use a cone of radius 0.5 for jets.

Calibration of the DØ detector continued during the running period between collider runs. Pulsed signals were sent through each preamp in the calorimeter so that each readout channel could be calibrated. This calibration data was used to determine gain corrections for the preamps which were stored in a database used to correct each calorimeter readout channel during off-line event reconstruction.

4.6.5 Calorimeter Noise

Noise in the calorimeter may come from both the uranium plates and the electronics used to read the calorimeter information. These two noise sources may be distinguished by their distributions. Uranium noise results from β decay of the nuclei in the plates producing a Landau distribution with a long, high-energy tail. By contrast, noise in the electronics generates a symmetric Gaussian distribution.

During the running period the noise distribution (pedestal) in each calorimeter channel was measured regularly while no beam was present in the Tevatron. Fitting each distribution to a Gaussian established the mean and the width of the noise pedestal. Any channel within $\pm 2\sigma$ of its pedestal value was suppressed. This is known as *zero suppression* and greatly reduces the number of channels read out for each event.

Noise distributions differ for the various sections of the calorimeter due to the differing geometry of the sections. Electronic and uranium plate noise are proportional to the capacitance (area) of the readout cell and the number of gaps in the cell. Thus, smaller cells have less noise. Typical noise widths are listed in table 4.1.

4.7 The Muon Spectrometer

The Muon Spectrometer of the DØ detector consists of several layers of proportional drift tube chambers (PDTs) on either side of five toroidal iron magnets with a field strength of approximately 2 Tesla. These magnets are used to bend the tracks of muons passing through the spectrometer. The angular bend is used to determine the momentum of the muon and its charge. The toroids and associated PDT layers are shown in fig. 4.5.

Since muons are minimum ionizing particles (MIP) they are seldom absorbed in the calorimeter. Muons do deposit small amounts of energy in the calorimeter (a fact which shall play a critical role in this work). However, the muon mass together with a long lifetime ($\approx 2.2\mu s$), means that they usually escape the calorimeter without leaving a significant EM shower. Additionally, muons do not interact strongly and so induce no hadronic shower.

The incident trajectory of the muons entering the spectrometer is determined from a combination of the event vertex, information from the central tracking system, and a track vector seen in the first muon chamber. Comparing the initial and final

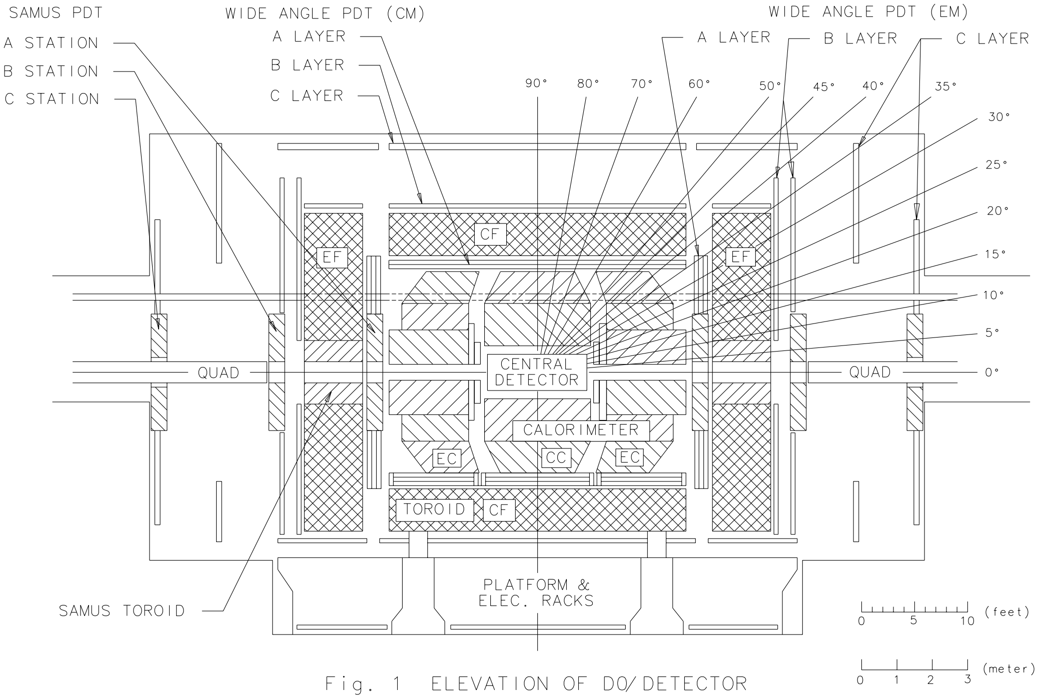

Figure 4.7. Elevated side view of the DØ detector.

directions of the muon provides information about the bend angle in the toroidal magnet. The muon momentum is directly related to this bend.

The Central Iron (CF) toroid covers the region $|\eta| \leq 1$, while the two Endcap Iron (EF) toroids cover the region $1 < |\eta| \leq 2.5$. These toroids together are known as the Wide Angle Muon System (WAMUS). There is also a Small Angle Muon System (SAMUS) covering the region $2.5 < |\eta| \leq 3.6$. The SAMUS is comprised of two toroids which fit into the central holes of the EF toroids.

WAMUS chambers are deployed in three layers: the “A” layer between the calorimeter and the toroid, and the “B” and “C” layers after the magnets. Each plane in the A layer chambers contains four PDTs, while the B and C layer chambers contain three PDTs each. The cell structure of the WAMUS system is uniform and consists of 164 individual chambers. The chambers are filled with a mixture of $Ar(90\%)/CF_4(5\%)/CO_2(5\%)$ gas.

The SAMUS consists of three stations: the A station prior to the SAMUS toroids and the B and C stations between the toroids and the low beta quadrapole for the DØ $p\overline{p}$ interaction region. Each station consists of three pairs of PDTs for a total of 5308 tubes in the SAMUS system.

The position resolution of the muon spectrometer is approximately ±3 mm corresponding to a momentum resolution of

$$\sigma(1/p) = 0.18(p-2)/p^2 \oplus 0.003 \tag{4.5}$$

where p is measured in GeV/c and $\oplus$ indicates addition in quadrature. Given the strength of the magnetic field in the muon spectrometer and the spatial resolution of the PDT chambers, muon momenta are well measured for p_T less than approximately 100 GeV/c. Above this value, however, the uncertainty in the transverse momenta of the muon is such that it is not possible to state the muon momenta with any certainty whatsoever. All that may be said is that the transverse momentum is greater than 100 GeV/c.

4.8 Triggering and Data Acquisition

At the interaction region $p\overline{p}$ beam crossings occur every 3.5 μ s. With the typical luminosity's from the run, at least one collision will occur every crossing (on average). This rate of collisions far exceeds the ability of current electronic and computer technology for recording the events. Therefore, DØ uses a multilevel triggering and filtering system designed to select apparently interesting events for recording, reconstruction and eventual analysis. There are three levels of triggers/filters increasing in sophistication and complexity through which each candidate event must pass before being recorded.

4.8.1 The Level 0 Triggers

The initial level of the triggering systems is known as the level 0 trigger. It is used to indicate that an inelastic collision has occurred and to measure the instantaneous luminosity in the interaction region. The level 0 triggers consist of an array of scintillation counters connected to photomultiplier tubes mounted on the inner surface of the endcap cryostats. The arrays provide partial coverage in ϕ between $1.9 < |\eta| < 2.3$ and full coverage in ϕ in the region $2.3 < |\eta| < 3.9$. Hits in both arrays coincident with a $p\overline{p}$ beam crossing indicate the occurrence of an inelastic collision. Additionally, the timing difference between the arrays is used to make an initial determination of the z position of the interaction vertex and to determine if multiple interactions occurred during a single beam crossing (the average number of interactions per crossing during the run was 2.6). Events which are flagged as inelastic collisions are passed along to the level 1 triggers.

4.8.2 The Level 1 Triggers

The level 1 triggering system is comprised of fast electronic hardware circuits controlled by software algorithms. The trigger uses coarse, rapidly digitized information from the calorimeter and the muon spectrometer to determine in less than 3.5 μs if a particular event meets specific criteria on energy deposits and topological requirements to merit further processing. Calorimeter cell information is grouped into trigger towers having a size of 0.2×0.2 in $\eta - \phi$ space. A level 1 trigger may, for example, require one or more triggers towers above some energy threshold in addition to hits in the muon chambers for the event to be passed along to the level 2 triggering system.

4.8.3 The Level 2 Triggers

The level 2 triggering system consists of software filters operating on a farm of 48 Vaxstation nodes operating in parallel to reconstruct events passing the level 1 system. Level 2 utilizes the fully digitized detector information from all detector components to reconstruct and identify specific objects such as jets, electron, photon, and muon candidates. The filters may select events passing requirements on the E_T, p_T, η, and ϕ as well as the number of objects of a specific type. In this work we used the set of level 2 filters which required one muon and at least one jet, and those requiring two muons and at least one jet. Further information on the level two requirements for this work will be presented in the next chapter.

4.9 Event Reconstruction

The events passing the level 2 filters are not in their final form since level 2 does not utilize detailed information from the full tracking system or the shower shape. Objects such as photons, electrons and jets are not well measured at this point due to restrictions on processing time and disk space. The tasks are performed off-line by a reconstruction package known as DØRECO. DØRECO applies calibration information to the raw data from each detector subsystem to correct the energy deposits and apply algorithms to the information to find physical particles.

4.9.1 Event Vertex

Since the E_T and η of jets depend on the location of the event vertex, an accurate determination of the vertex is necessary to properly reconstruct the jets. The vertex is determined using tracking information from the CDC (or from the FDC if the vertex is not found by the CDC). For each event every charged track is reconstructed and extrapolated to the axis of the beam to obtain a distribution of track intersections. A Gaussian is fit to the distribution of the intersections and the mean is used as the

z position of the vertex for the event. If the tracks are clustered into more than a single smooth distribution, each cluster of tracks is fit with a Gaussian and up to three event vertices may be found. The cluster with the largest number of tracks is called the primary vertex. If it is necessary to use information from the FDC to find the event vertex only one vertex will be found due to the high density of the tracks in the forward region. If no vertex is found using the reconstructed tracks the vertex position may be determined using the timing information from the level 0 scintillating arrays.

4.9.2 Jet Reconstruction

Following a hard scattering interaction the final state partons (the constituents of the (anti)proton) will carry a large momentum and a color charge. The partons will then undergo a hadronization process which serves to disperse the color charge. This process results in highly collimated showers of particles which are grouped into clusters of energy and defined as jets. DØ identifies jets using the hadronic calorimeter.

The energy response of the calorimeter varies with pseudorapidity and is not linear for particles below about 10 GeV. Since jets are collimated showers of low energy particles, the energy measured in the calorimeter is not the actual energy of the jet. The calorimeter response (known as the jet energy scale) is corrected to reflect the actual energy of the jet from the energy deposited in the calorimeter. The jet energy is also corrected for calorimeter noise (both electronic and uranium noise), signal pileup (energy remaining from the previous event), extraneous energy in the jet from an underlying event in multiple interactions, and leakage out of the jet cone due to shower width. The actual jet energy is given as

$$E^{jet}_{particle} = \frac{E^{jet}_{measured} - O}{(1 - S)R_{had}} \tag{4.6}$$

where O is an offset factor accounting for noise, pileup, and underlying events. S is the showering correction for the calorimeter, and R_{had} is the hadronic response of the calorimeter. The determination and values for these correction factors are well documented in refs [44].

4.9.3 Missing Transverse Energy Determination

The calculation of missing transverse energy ($\not\!\!E_T$) at DØ is based on energy deposits at the cell level in the calorimeter.

$$\begin{aligned}
(\not\!\!E_T)_x &= -\sum_i E_x^i \\
(\not\!\!E_T)_y &= -\sum_i E_y^i \\
(\not\!\!E_T) &= \sqrt{(\not\!\!E_T)_x^2 + (\not\!\!E_T)_y^2} \\
\phi_{\not E_T} &= \tan^{-1}(\not\!\!E_T)_x/(\not\!\!E_T)_y
\end{aligned} \tag{4.7}$$

where E_x^i and E_y^i are the x and y components of the energy deposited in the i^{th} cell of the calorimeter. Summation is over all cells in the calorimeter.

4.9.4 Muon Reconstruction

Data is taken from a series of triggers which require the presence of at least one muon candidate and at least one jet candidate per event. During the running period the names and requirements of these triggers changed frequently. To make the data set consistent we apply a series of loose selection criteria (also known as "cuts") to the data at the trigger/filter level. The data which pass these criteria will form our working data set. The selections are based on the muon and jet identification criteria to be applied in full at a later time.

We require the presence of at least one muon candidate passing the trigger with the following selection criteria. The candidate must have $|\eta_\mu| < 1.7$, and

$p_T^\mu > 5$ GeV$/c$. It is required to have a reconstructed track consistent with a muon originating from the interaction region within 25 cm of the reconstructed interaction vertex. The time of flight from the vertex to the muon spectrometer, calculated with information from the proportional drift tube chambers, must be less than 600 ns.

Cosmic ray muons are rejected using a veto on events which are back to back in η and ϕ. Cosmic ray muons which pass through the detector typically appear to be dimuon events. This is due to the fact that such muons often have very large energies. Thus, they enter the muon spectrometer and register as a muon candidate, pass through the detector and emerge as a muon candidate on the other side. They appear, therefore, as two muons which originated from the interaction region and emerged with back to back tracks.

Muon candidates are required to deposit energy in the calorimeter. The muon track through the calorimeter is reconstructed and the energy summed along all cells through which the track passes plus the "nearest neighbor" cells along the track. This energy is required to be $\frac{1}{2}$ GeV for muons with $|\eta_\mu| < 1.0$ and 1 GeV for muon candidates with $1.0 < |\eta_\mu| < 1.7$. Finally, we require muon candidates to have a reconstructed track which registered in at least 60% of the layers in the hadronic calorimeter.

Each event is also required to have the presence of at least two jet candidates with a transverse energy $E_T > 8$ GeV and $|\eta_j| < 4.0$. When these selection criteria are applied, we obtain a data set consisting of 4580 events representing an integrated luminosity of $94 \pm 5\ pb^{-1}$.

CHAPTER 5

THE $\mu\mu + JETS$ DECAY CHANNEL

“Get your facts first, and then you can distort them as much as you please.

- Mark Twain, quoted by Rudyard Kipling in *From Sea to Shining Sea*

5.1 Introduction

With the data set described in the previous chapter we begin a search for second generation leptoquarks by searching in the dimuon + jets decay channel [45–46]. Since we have assumed that leptoquarks couple only within a single generation of leptons and quarks, we do not consider decay modes such as $e\mu + jets$. The decay to a muon pair and jets may come about for any branching ratio except $\beta = 0$ (100% branching into neutrinos and jets). We shall consider in this chapter two possible branching ratios for leptoquarks into charged leptons. Our intermediate results at the end of this chapter will be stated in terms of a branching ratio of 100% into charged leptons and jets, corresponding to $\beta = 1$ (recall from chapter 3 the definition of β as the branching ratio to *charged* leptons) and for equal branching into charged leptons + jets and neutral leptons and jets ($\beta = \frac{1}{2}$). In the first case the branching ratio for dimuon events is simply 1, while for the second case the branching ratio for dimuon + jets events is $\frac{1}{4}$, branching ratio $= \beta^2 = \frac{1}{2}^2 = \frac{1}{4}$. The muon + neutrino + jets decay channel will be considered in the following chapter. Other possible values of β will be addressed in chapter 7.

5.2 Initial Event Selection

We apply the following initial selection criteria to the data set described in the previous chapter. We require at least two jets in each event with the following properties. The jet must have deposited between 5% - 95% of its energy in the electromagnetic calorimeter and less than 40% in the coarse hadronic calorimeter. We form the ratio of the two calorimeter cells which have the most deposited energy and require the ratio of the most energetic cell to the second most energetic cell to be less than 10. This is done to eliminate jets with inflated energies due to pileup. We also require that each jet have a transverse energy $E_T > 20$ GeV and $|\eta_j| < 3.0$. Throughout the remainder of this work we shall refer to jets and muons by number (*e.g.*, $E_T^{j_1}, p_T^{\mu_2}$, *etc*). The order of jets (muons) is by transverse energy (momentum). Thus the first jet (sometimes called the leading jet) in each event is the jet having the largest E_T, the second jet has the second highest E_T and so forth.

5.2.1 Muon Pair Identification

We shall require two muons in each event to pass the initial selection criteria. The identification of muons is described in full in this section. We shall accept muon pairs in three ways: first is two central ($|\eta| < 1.0$) muons, second is one central muon and one muon in the endcap ($1.0 < |\eta| < 1.7$) muon spectrometers and third is one central muon and one muon identified by a pattern of isolated energy deposited in the longitudinal segments of the hadronic calorimeter. These various identified muons will be referred to as μ_{Sp} (muons identified in the muon spectrometer) and μ_{Cal} (muons identified in the calorimeter).

The muon spectrometer is divided into five different sections as noted in section 4-6. We shall use only muons found in the Wide Angle Muon Spectrometer comprised of the central region and the two endcap plugs. Each of these elements of the spectrometer is divided into quadrants. Quadrants one through four refer to the central spectrometer, quadrants five through eight to the north endcap spectrometer and quadrants nine through twelve to the south endcap spectrometer. Initial selection will therefore require each event to have at least one muon with quadrant four or less.

Central muons are required to have a reconstructed track consistent with a minimum ionizing particle originating from the interaction region. They are required to have been detected in all three layers of the PDTs. This is done to eliminate events in which a jet penetrated through the hadronic calorimeter and into the muon spectrometer, thus registering in the detector as a muon candidate. Central muons are also required to have been detected in all layers of the hadronic calorimeter, or to have been detected in 70% of the layers in the hadronic calorimeter and to have deposited energy in the final layer of the calorimeter. They are also required to have deposited at least $\frac{1}{2}$ GeV in the calorimeter cells along the track plus one nearest neighbor cell. Muons are required to have a fully reconstructed track (described in full later on) in the calorimeter consistent with a minimum ionizing particle within 0.2 in $\Delta R \equiv \sqrt{\Delta\eta^2 + \Delta\phi^2}$ with the $\eta - \phi$ direction of the muon measured in the spectrometer. To eliminate muons originating from heavy quark decays, muons are required to be well isolated from any jets in the event passing the selection criteria listed in the previous section, $\Delta R_{\mu j} > 0.5$. All muons are required to have $|\eta| < 1.7$.

Finally, muons are required to have a transverse momentum 20 GeV/c < p_T^μ < 900 GeV/c. As noted in the previous chapter, the transverse momentum of the muons is measured by the curvature of the track through the PDTs in the magnetic field within the muon spectrometer. If a muon has a transverse momentum greater than about 100 GeV/c the track will not bend enough to be properly measured by the PDTs. All that may then be said about the muon p_T is that it is greater than

100 GeV/c. Muons with transverse momenta that are mismeasured in this way will form a very difficult background to suppress. We begin this process by requiring the muons to have a transverse momenta less than the beam energy of the Tevatron, 900 GeV.

An event that has at least one central muon may also have a second muon identified from the endcap muon spectrometer $5 \leq Quadrant \leq 12$ or, roughly, $1.0 < |\eta_\mu| \leq 1.7$. The selection criteria for this type of muon are somewhat more stringent than for central muons. This requirement is made for the following reason. During the running period the wires in the PDTs became contaminated by deposits from the gas that filled the chambers. This contributed to a degradation of the efficiency of the muon system. The degradation was particularly bad in the endcap spectrometers since this region was already inefficient. During the latter part of the running period it was discovered that the wires could be cleaned by applying a high voltage to the chambers. The resulting current would cause the deposits to be ejected from the wires (this process is known as "zapping" the PDTs). This process improved the efficiency of the muon system. Muons from the endcap spectrometers are accepted only from data taken after this process was begun (these are called "post-zap" muons).

Endcap muons are required to have a reconstructed track consistent with a minimum ionizing particle originating from the interaction region with more stringent criteria than central muons. They are required to have been detected in all three layers of PDTs in the muon spectrometer. Endcap muons are required to have tracks reconstructed in all layers of the hadronic calorimeter, or to have tracks which are reconstructed in at least 70% of the layers in the hadronic calorimeter and to have deposited some energy in the final layer of the electromagnetic calorimeter. Endcap muons are required to have deposited at least 1 GeV of energy along the reconstructed

Table 5.1. Muon identification criteria in the spectrometer.

Central Muons	Endcap Muons
Quadrant ≤ 4	$5 \leq$ Quadrant ≤ 12
	Post-Zap Muons only *
Three layer PDT *	Three layer PDT *
Hadronic Fraction *	Hadronic Fraction *
$E_{Cal}^{1nn} > \frac{1}{2}$ GeV *	$E_{Cal}^{1nn} > 1$ GeV *
$\Delta R_{\mu-Cal} < 0.2$ *	$\Delta R_{\mu-Cal} < 0.2$ *
$20\ \text{GeV}/c < p_T^{\mu} < 900\ \text{GeV}/c$	$20\ \text{GeV}/c < p_T^{\mu} < 900\ \text{GeV}/c$
$\lvert\eta_{\mu}\rvert < 1.7$	$\lvert\eta_{\mu}\rvert < 1.7$
$\Delta R_{\mu j} > 0.5$	$\Delta R_{\mu j} > 0.5$

track plus the nearest neighbor cells in the calorimeter, and the reconstructed track through the calorimeter must be within 0.2 in ΔR with the $\eta - \phi$ measured by the muon spectrometer.

Kinematic selection of the endcap muons is the same as for the central muons, $20\ \text{GeV}/c < p_T^{\mu} < 900\ \text{GeV}/c$, $|\eta_{\mu}| < 1.7$, and $\Delta R_{\mu j} > 0.5$. The muon selection criteria for central and endcap muons is summarized in table 5.1. In the table, a "*" indicates selections which were not applied during Monte Carlo studies. The quantity E_{Cal}^{1nn} indicates that the energy is summed over the calorimeter cell and the nearest neighbor cells.

5.2.2 Muon Identification in the Calorimeter

The DØ detector has excellent calorimetry. The calorimeter provides full coverage in η and ϕ and is finely segmented. In addition, there is no central magnetic field.

This means that particles passing through the calorimeter do not proceed along curved trajectories as would be the case if there was a central magnetic field with which they could interact. In addition, the calorimeter has sensitive energy detection and low noise, making it possible to resolve small signals in the calorimeter. For all of these reasons it is possible to reconstruct tracks from minimum ionizing particles in the calorimeter. DØ has developed a software package, known as Muon Tracking in the DØ Calorimeter [47] (MTC) package. Since a substantial part of the work presented here relies on the identification of muons with the calorimeter we shall present some of the necessary concepts for the process in this section.

The MTC package uses calorimeter information to identify and reconstruct energy deposits in the calorimeter which appear to lie along a track. The transverse segmentation of the calorimeter is relatively coarse, however, jets are generally well contained in the hadronic section of the calorimeter. Longitudinal segmentation of the calorimeter is sufficient to sample the layer by layer energy deposits of a track. This makes it possible to identify isolated muons. Calorimeter coverage insures that there is a uniform tracking efficiency as a function of η and ϕ with no coverage gaps. Muons travel along a straight line path due to the absence of a central magnetic field and they generally emerge from a known vertex measured in the central tracking chambers, providing a constraint to tracking muons with the calorimeter. Finally, muons are minimum ionizing particles, and they therefore deposit small amounts of energy in the calorimeter over many interaction lengths. Thus muons leave a distinctive signature in the calorimeter. The MTC package is designed to use these features of the DØ calorimeter to track muons.

The MTC package has two primary sections, each of which we use and shall treat in turn. First is the muon identification utility. This utility takes a given vertex, η and ϕ for a muon candidate and looks at energy deposits in calorimeter cells in a road surrounding the muon candidate. The results from this search are the muon calorimeter track, a χ^2 for the track fit, track energy deposits, and the calorimeter layer at which the muon seems to emerge from a jet. The information obtained from this utility is used in the muon identification criteria described in the previous section to obtain calorimeter track confirmation, hadronic fractions of muon detection in the calorimeter, and energy deposits in the calorimeter.

In addition, the MTC package has a muon finding utility. This utility performs a scan of the calorimeter on a cell by cell basis from a given vertex position to find candidate tracks emerging from the vertex. The information from this utility is used in this work to identify muons which failed to pass the identification presented in the previous section. The following requirements are imposed on the tracks obtained to identify muons in the calorimeter.

First, to insure that we are not double counting tracks associated with muons already identified, no tracks are accepted within $\Delta R_{\mu-Cal} < 0.2$ of any muons which pass the identification from the previous section. Additionally, we have found that the rate at which fake tracks are found is slightly dependant on the number of jets in the event (this will be more fully explored later on). To avoid accepting tracks which are due to jets, we do not accept tracks which are within 0.2 in $\Delta R_{\mu-Cal}$ of any muon candidates which failed identification because they were detected in only the first layer of the PDTs in the muon spectrometer, which could be due to punchthrough (a jet escaping the hadronic calorimeter and entering the first layer of the muon spectrometer).

Tracks passing the above criteria are processed as follows. As mentioned in section 4-6, neutrinos are not detected in the DØ calorimeter. They are inferred from conservation of momentum and the transverse energy is calculated by a sum over the calorimeter cells. Since muons are minimum ionizing particles, they do not deposit substantial energy in the calorimeter and thus they also contribute to the missing transverse energy. When a muon is identified by the criteria in section 5.2.1, the $\not\!\!E_T$ is corrected for the momentum of the muon. If there is still substantial $\not\!\!E_T$ in the event, it could be due to a muon which was not identified, or to a neutrino. For the search in the dimuon channel, we attempt to identify muons from tracks in the calorimeter which lie within 0.25 radians ($\approx 14^\circ$) in ϕ of the direction of the $\not\!\!E_T$. We require the track to have $|\eta_{Cal}| < 1.7$, and to have been detected in at least 90% of the layers in the hadronic calorimeter.

We require the track to be well isolated from any jets in the event ($\Delta R_{Cal-j} > 0.5$), and to have deposited at least $\frac{1}{2}$ GeV for tracks with $|\eta_{Cal}| < 1.0$ or 1 GeV for tracks with $1.0 < |\eta_{Cal}| < 1.7$. These selections are seen to be analogous to the muon identification criteria of the previous section. The final selection criteria for tracks in the calorimeter is $\chi^2_{Cal} < 10$, where χ^2_{Cal} is defined by

$$\chi^2_{Cal} \equiv \frac{1}{N_{Layers}} \sum_{i=1}^{N_{Layers}} \frac{(E_i - \lambda_i)^2}{\sigma_i^2} \tag{5.1}$$

where N_{Layers} is the number of calorimeter layers with a hit cell in the path of the muon candidate, E_i is the energy measured in layer i, λ_i is the most probable value of the energy for a calorimeter cell at the layer and η, and σ_i is the width of the energy distribution for a calorimeter cell at that layer and η. The calorimeter track identification is summarized in table 5.2 below. Calorimeter tracks which are found via the MTC software package and pass the selection requirements listed in the table are presumed to be muons.

We require events with a single, central, muon to have at least one track which passes the criteria listed in table 5.2. If there is more than one track in the event passing the criteria, the track with the lowest value of χ^2 is selected as the primary track. Muon kinematic quantities are calculated from the $\eta - \phi$ direction of the

Table 5.2. Calorimeter track selection criteria.

Calorimeter Track Selection Requirements
$\Delta R_{\mu j} > 0.2$
$\Delta\phi_{Cal-\not{E}_T} < 0.25$
$\lvert\eta_{Cal}\rvert < 1.7$
Hadronic Fraction > 90%
$\chi^2_{Cal} < 10$
$\Delta R_{Cal-jets} > 0.5$
$E_{Cal} > \frac{1}{2}(1)$ GeV ($\lvert\eta_{Cal}\rvert < (>) 1.0$)

track and the $\not{E}_T$. After this calculation, the muon candidate is required to have 20 GeV/$c < p_T^{\mu} <$ 900 GeV/c.

5.2.3 Concluding the Initial Event Selection

We now have a selection process which includes a good muon pair and two jets in each event. Since the muon identification is somewhat complicated, we present the following summary of the initial selection to this point. Each event in the data set is required to have at least two jets passing the criteria in section 5.2. We further require that each event has at least two muons. If there are more than two muons in the event which pass the selection (which happens only rarely), only the two leading muons are considered further. At least one of the muons in the event must be a central (CF) muon. The other muon in the event may be another central muon, a muon from the endcap spectrometer, or a muon identified in the calorimeter.

The final selection criteria in the initial selection process is a further requirement to eliminate cosmic ray muons. We do not accept events in which the dimuon pair

Table 5.3. Data reduction in the dimuon channel.

Selection Criteria	Events Passing
Trigger Selection	4580
$N_{jets} \geq 2$	2595
$N_{\mu} \geq 2$	70
Cosmic Veto	61

has $|\Delta\phi_{\mu\mu}| > 160^{\circ}$ and $|\Delta\eta_{\mu\mu}| < 0.5$. The effect of the various selections is shown in table 5.3.

The muon identification criteria which are applied to the data only are the same as the selection applied in the top quark search via the dimuon channel [48] with the addition of the calorimeter track confirmation. The efficiencies of the identification process are 71/89% for two central muons pre/post - zap. For one central and one endcap muon the identification efficiency is 86%. The calorimeter track confirmation has an efficiency of 94%. Muon triggers have an efficiency of 71% for central muons and 50% for endcap muons. The efficiencies were obtained from the preliminary leptoquark dimuon + jets search [49] (muon identification and triggers) and the $Z\gamma \rightarrow \mu\mu\gamma$ study [50] (calorimeter track match efficiency). All efficiencies were confirmed with $Z \rightarrow \mu\mu$ data.

5.3 Monte Carlo Generation

Signal Monte Carlo samples were generated using ISAJET v7.22 [51] for scalar second generation leptoquark pair production. For vector couplings the PYTHIA [52] event generator, modified for vector leptoquark pair production, is used. The detection efficiencies for scalar (S_{LQ}) and vector (V_{LQ}) leptoquark couplings are consistent within the differences due to choice of generator. We therefore do not distinguish between efficiencies obtained from scalar or vector Monte Carlo, but present a single set of experimental results. Furthermore, although the vector Monte Carlo was generated for "minimal vector" couplings ($\kappa_G = 1, \lambda_G = 0$), our experience [53] has been that in the region of interest ($m_{LQ} > 200$ GeV/c^2), detection efficiencies for various vector couplings are the same. We thus use the efficiencies obtained here to set limits on the pair production of vector leptoquarks with "Yang-Mills" couplings ($\kappa_G = 0, \lambda_G = 0$) as well as for minimal vector couplings.

Leptoquark pair production cross sections used to compare experimental results to theory are from recent next-to-leading order (NLO) calculations [54] done at a renormalization scale $\mu = m_{S_{LQ}}$ with uncertainties obtained from the variation of the renormalization/factorization scale from $2m_{S_{LQ}}$ to $\frac{1}{2}m_{S_{LQ}}$. The V_{LQ} cross sections [55] used are from leading-order (LO) calculations at an energy scale $Q^2 = m^2_{V_{LQ}}$. For a leptoquark mass of 200 GeV/c^2 the S_{LQ} cross section is $0.184^{+0.018}_{-0.026} pb$. For minimal vector couplings the cross section is 2.0 *pb*, while for Yang-Mills couplings it is 10 *pb*.

The dominant vector boson backgrounds, from $W + jets$, $Z + jets$, and WW production, are simulated using the VECBOS [56] Monte Carlo program with the HERWIG [57] program used for parton fragmentation. The fourth background considered, from top quark pair production and subsequent decay into one or more muons plus two or more jets, is simulated with the HERWIG [57] Monte Carlo program at a top quark mass of 170 GeV/c^2. All Monte Carlo files are processed through a detector simulation program based on the GEANT [58] package.

5.4 Background Studies

The Standard Model backgrounds considered in this work are $Z + jets$, $W + jets$, WW, and $t\bar{t}$ production. Each of these processes have decay modes to one or more muons as well as decays to taus with subsequent cascade decays to muons. In order to account for the Standard Model backgrounds we calculate the expected number of events for each of the processes considered after the initial selection. The $Z + 2\ jets$ cross section of 9.4 *pb* is used along with a factor to account for the branching ratio for $Z \rightarrow \mu\mu$ processes and the $Z \rightarrow \tau\tau$ with subsequent decay of the $\tau \rightarrow \mu\nu$. The calculation of the top quark pair production background uses the DØ measured cross section of 5.5 ± 1.8 *pb*.

For the selection criteria requiring $\mu_{Sp} - \mu_{Sp}$, the WW cross section of 10.4 *pb* is used. The $W + jets$ background in this selection is small and does not contribute in a significant fashion. We do not consider this background further for these selection processes.

For the $\mu_{Sp} - \mu_{Cal}$ a somewhat different selection process is required. First, in order to simulate a "lost" muon (that is, one that fails to pass the identification criteria), the second muon in dimuon events (*e.g.,* signal Monte Carlo, $Z + jets$ Monte Carlo) is dropped from the selection process. The MTC software package is then used to attempt to "recover" it. For backgrounds with single muon events (*i.e.* $W + jets$) this is unnecessary. These backgrounds contribute via an MTC fake. An MTC fake is defined as an event in which the calorimeter tracks happen to line up with missing transverse energy due to a neutrino and then pass all of the track selection criteria. This track will then be mistakenly identified as a muon.

These backgrounds are accounted for using a "fake rate" method. We do this for two reasons. First, we have found that the Monte Carlo does not model MTC fake tracks particularly well and second, by using a fake rate calculation we may make better use of the available Monte Carlo statistics.

To determine the fake rate we use $W + jets \rightarrow e\nu + jets$ data. We apply the jet selection and the muon identification and kinematic criteria outlined above, without the requirement that the muon be central. The event is required to have no muons passing the identification requirements. This is done to insure that any tracks in the event do not come from good muons, but are fake tracks and that the $\not\!\!E_T$ in the event is due to a real neutrino.

We have found that there is a small dependence of the fake rate on the jet multiplicity, with the fake rate increasing slightly with the number of jets in the event. We thus adopt the following method. The fake rate is calculated for $N_{jets} = 2, 3, \geq 4$ since the fake rate plateaus for $N_{jets} > 4$. Also, the number of events in the $W + jets \rightarrow e\nu + jets$ data sample with $N_{jets} > 4$ decreases significantly, so statistical errors become large for higher jet multiplicity's. The fake rate is determined from the number of events which have a calorimeter track passing the track selection divided by the total number of events with $N_{jets} = 2, 3, \geq 4$. We then average the fake rates obtained over the events in our data set with $N_{jets} = 2, 3, \geq 4$. As a cross check, we also calculate the fake rate on a per track basis. This is simply the number of tracks in the $W + jets \rightarrow e\nu + jets$ data sample which pass the track selection divided by the total number of tracks in the sample. We then average this over the track multiplicity in our data set. The results from both methods are consistent. We obtain a fake rate of (10.8±0.8)%, where the uncertainty is statistical.

To estimate the backgrounds which contribute via a fake calorimeter track we do not apply the track selection criteria to the Monte Carlo. We require that the event have at least one track with $|\eta| < 1.7$. The event is then accepted and the properties of the second muon are calculated from the $\not\!\!E_T$ and the $\eta - \phi$ direction of the track closest to the $\not\!\!E_T$. The estimate of the number of events in the data is then multiplied by 0.108 to account for the probability of obtaining a fake calorimeter track.

For the $W + jets$ background, the VECBOS [56] cross section of 1000 pb is a leading order result. Next-to-leading order processes may have a significant effect

Table 5.4. Expectation after initial selection.

Type	$\mu_{Sp}\mu_{Sp}$ *Selection*	$\mu_{Sp}\mu_{Cal}$ *Selection*	*Total*
$t\bar{t}$	0.30±0.1	1.7±0.6	2.0±0.6
$Z + jets$	2.6±0.5	3.0±0.7	5.6±0.9
$W + jets$	——	53±16	53±16
WW	2.0±0.8	0.4±0.2	2.4±0.9
Total Background	4.9±1.0	58±16	63±16
Signal	1.6±0.2	2.2±0.3	3.7±0.5
Data	6	55	61

on our estimation of the $W + jets$ background. To account for this we normalize the $W + jets$ background to the single muon data in a region where our signal expectation is very small (less than 0.05 expected events for a leptoquark mass of 200 GeV/c^2). We do this by normalizing in the lowest bins of a neural network $D_{nn} < 0.2$ discussed in detail in the following section (see fig. 5.4). We take the number of events in the data which passed the $\mu_{Sp} - \mu_{Cal}$ selection and subtract the expectations of the other background processes ($Z+jets$, WW, and $t\bar{t}$). The $W+jets$ expectation using the LO cross section is 34±10 events, while the expectation from the data minus the other backgrounds is 42 events. The $W + jets$ expectation is then normalized to account for this discrepancy. As a cross check we note that this is equivalent to inflating the $W + jets$ leading order cross section by 24% and is consistent with our expectations of the differences between LO and NLO cross sections. Our normalization is also consistent with the normalization obtained in the search for first generation leptoquarks [59] in the $e\nu + jets$ decay channel of 20%.

Table 5.5. Sources of systematic uncertainty.

Source	$t\bar{t}$	$Z+jets$	$W+jets$	WW	S_{LQ}
Energy Scale	6.6%	17%	17%	17%	6.6%
High p_T^μ *Eff*	6.4%	6.4%	6.4%	6.4%	6.4%
Multiple Interaction	4.7%	4.7%	4.7%	4.7%	3.1%
Cross Section	30%	10%	15%	15%	—-
Trigsim	5%	5%	5%	5%	5%
MC (PDF/Q^2/Generator)	10%	10%	10%	10%	10%
Overall Systematic	34%	24%	28%	28%	15%

With the backgrounds thus properly accounted for we have an expectation of 63±16 events in 94 pb^{-1} of data, consistent with 61 events in the data. The signal expectation after the initial selection is 3.7±0.5 events. The signal expectations given in this chapter are all calculated at the minimum of the NLO cross section ($\mu = 2\ m_{S_{LQ}}$) for scalar leptoquarks at a mass $m_{S_{LQ}} = 200$ GeV/c^2. The expectations for signal and background are shown in table 5.4. The table shows the expected number of signal and background events after the initial selection is completed. Expectations are broken down by selection process, and backgrounds are listed individually as well as collectively. The corresponding number of events in the data are shown for comparison.

Table 5.5 lists the sources of systematic error for each of the backgrounds individually as well as for the signal expectation. *Trigsim* is a software package which models the hardware triggering system. In the table the error assigned for *Trigsim* is an uncertainty in how well the software models the hardware triggering system. The high p_T^μ efficiency reflects uncertainties in the muon p_T resolution and

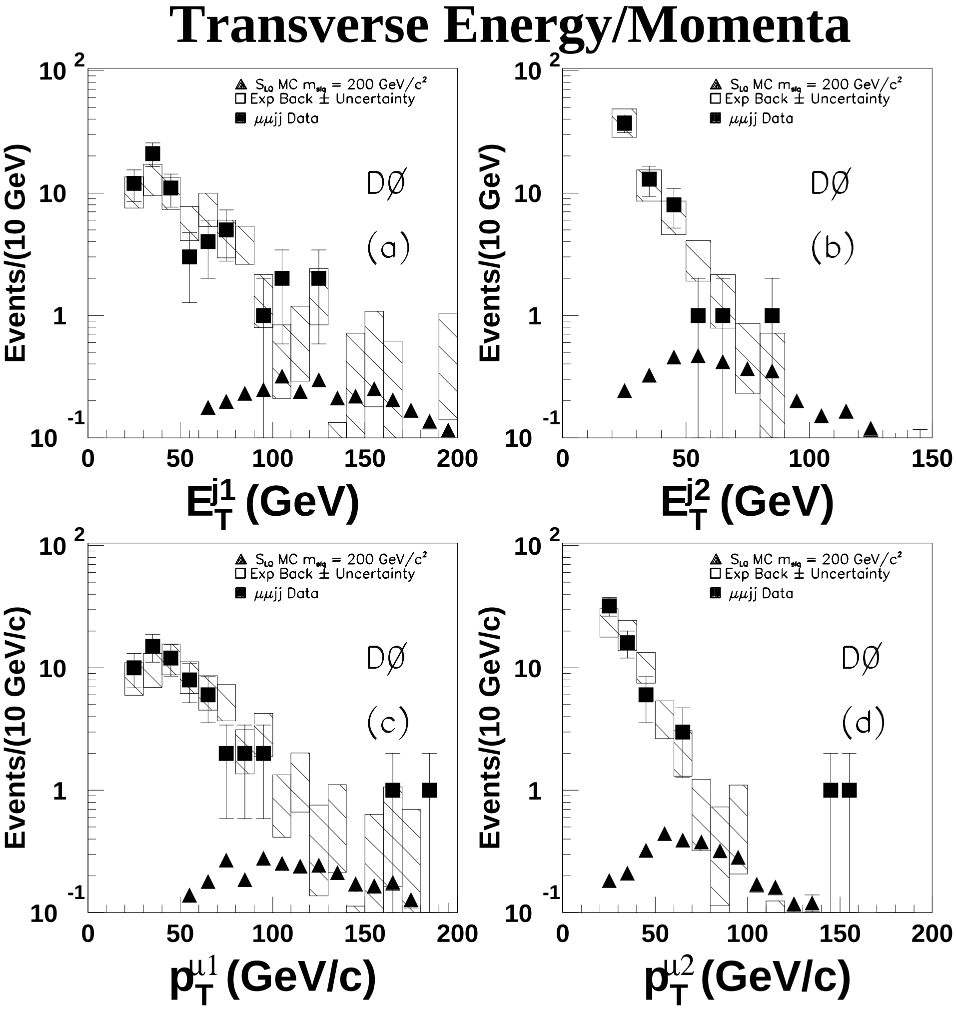

Figure 5.1. Transverse energy/momenta of the leading jets/muons.

trigger efficiencies. No uncertainty is assigned for the signal cross section since this variation will be used explicitly to obtain the results of the analysis. Muon p_T and jet E_T were varied by $\pm\sigma$ to test the effect of the variation on the results. Variations were small and found to be within the uncertainties assigned in table 5.5.

With the backgrounds properly accounted for we proceed to examine various distributions for evidence of a signal in our data set. The plots in fig. 5.1 show some of these distributions. Note that in the plots the hatched boxes show the background

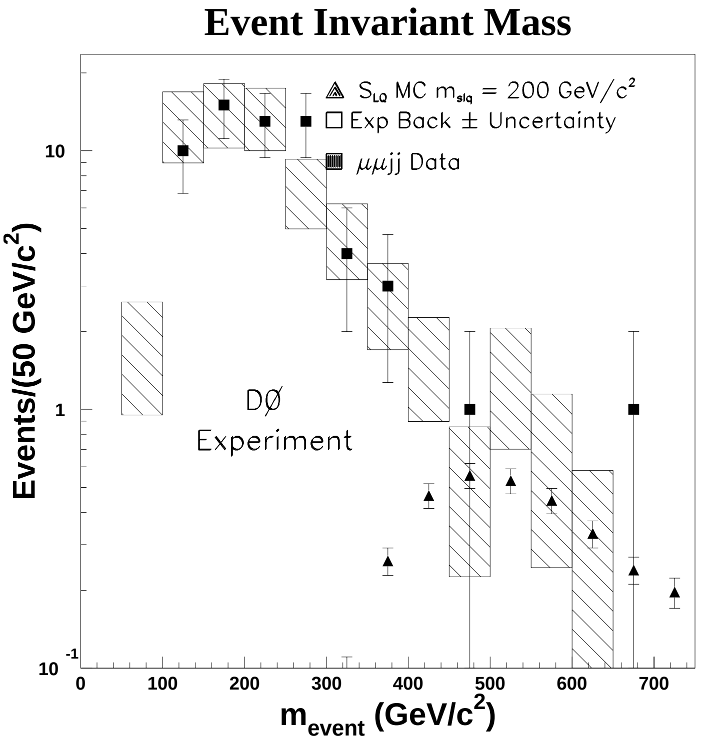

Figure 5.2. Event invariant mass.

plus the uncertainty on the background, the solid squares are the data with statistical uncertainties, and the triangles show the expected signal with uncertainties. For signal and background Monte Carlo both systematic (from table 5.5) and statistical uncertainties are accounted for in the overall uncertainty. No errors are shown on zero bins in the histograms. Although there are certainly statistical errors associated with these bins, the plots would become difficult to read given the scale if these errors were included.

5.5 Final Event Selection

Fig. 5.1(a) shows the E_T of the leading jet. Fig. 5.1(b) shows the transverse energy of the second most energetic jet. Figs. 5.1(c) and (d) show the transverse momentum of the leading and second leading muons respectively. On plots 5.1(c) and (d) there is one event in the data which is not seen on the scale presented. The p_T of the leading muon in this event is 899 GeV/c while the trailing muon has a p_T of 753 GeV/c. In fig. 5.2 is shown the event invariant mass. This quantity is the mass calculated from all muons and jets which pass the selection criteria. As in fig. 5.1 there is one data event which is off the scale of the plot (in fact, it is the same event). The event mass is 1844 GeV/c^2. From the plots in figs. 5.1 and 5.2 we see that the data are consistent with the background expectations.

With an understanding of the background we turn our attention to finding an appropriate set of cuts with which to separate the backgrounds from a possible signal. The first selection we apply is to the "sphericity" of each event. The sphericity, calculated in the center-of-mass frame and denoted as S_{CM}, is defined as follows. One forms the momentum tensor from the direct product of the three momenta of all of the muons and jets passing the initial selection in each event. The eigenvalues of the tensor are then calculated. The sphericity is then $S_{CM} = \frac{1.5}{\sum p^2}(\lambda_1 + \lambda_2)$ where $\lambda_1 < \lambda_2 < \lambda_3$ are the eigenvalues of the momentum tensor and p^2 is the square of the momenta of the various muons and jets.

As mentioned earlier, if the muon momentum is greater than about 100 GeV/c the value of the momentum obtained by the muon system is poorly determined due to insufficient curvature of the track through the magnetic field. However, given the current mass limits on second generation leptoquarks (184 GeV/c^2 from DØ and 202 GeV/c^2 from CDF), one expects a substantial fraction of the signal events to have muons with a p_T greater than 100 GeV/c. If one of the muons is severely mismeasured the event loses the expected spherical symmetry and would then tend to have a very low value of S_{CM}. The event sphericity is shown in fig. 5.3. A selection of $S_{CM} > 0.05$ is chosen.

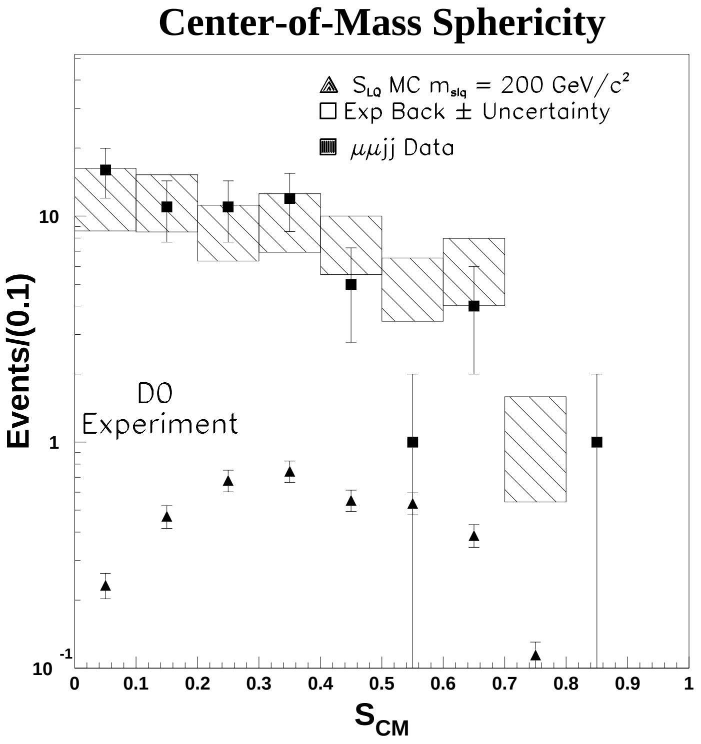

Figure 5.3. Event sphericity calculated in the center-of-mass frame.

We see from fig. 5.3 that, for very low values of the S_{CM}, there is little signal but a substantial number of expected background events. Thus, requiring a value of $S_{CM} > 0.05$ not only rejects difficult backgrounds with little loss in signal, but also improves the ratio of signal to background. With the selection requirement of $S_{CM} > 0.05$ we have 52 events remaining in the data consistent with an expected background of 53±13 events. The signal expectation remains 3.7±0.4 events. The data event noted in the discussion of figs. 5.1 and 5.2, with two obviously mismeasured muons has a center-of-mass sphericity of 0.02 and thus fails to pass this requirement.

We next turn our attention to finding an appropriate set of variables with which to eliminate the bulk of the backgrounds. The set of variables chosen for this purpose is $E_T^{j_1}$, $E_T^{j_2}$, $p_T^{\mu_1}$, $p_T^{\mu_2}$, $(E_T^{j_1} + E_T^{j_2})$, m_{event}, and $\frac{E_T^{j_1}+E_T^{j_2}}{\sum_{jets} E_T^{j_i}}$. These variables are then used as the inputs to a neural network (NN) [60]. The network structure is 7-15-1. The NN structure refers to the way in which the variables are connected to one another. In this case, there is a single layer of inputs with 7 nodes corresponding to the seven input variables. There is a "hidden" layer, so called because it is internal and is not seen by the user. The hidden layer has 15 nodes. Each node in the input layer is connected to each node in the hidden layer. The output of the network, known as the network discriminant (D_{nn}), is the output layer and contains a single node or value. This simply means that this network returns a single value as a result. The value may be loosely interpreted as the probability of a particular event being a signal event. The network output ranges from zero to one. For a full and detailed discussion of the construction, training and use of a neural network, the reader is encouraged to see Appendix B.

The network was trained on signal and background Monte Carlo samples with a set of loose cuts applied (not as strict a set of requirements as used in the initial selection). The background sample was a mixture of $Z+jets$, $W+jets$, and $t\bar{t}$ events. The WW Monte Carlo was not included since this background is small and training on $W + jets$ events accounts well for the WW backgrounds. The signal sample used in the training was for a scalar leptoquark mass of 200 GeV/c^2. Approximately 3,000 events of signal and background Monte Carlo were used in the training. This is slightly over twenty times the number of independent variables in the network (7*15 + 7 + 15 + 1 = 128) and is therefore a sufficient number of training events for the complexity of the network. The network output is shown in fig. 5.4. The plot shows the neural network discriminant, D_{nn} for signal/background/data. The data points are well modeled by the expected background and there is excellent separation between signal and background. We see from the distribution that the

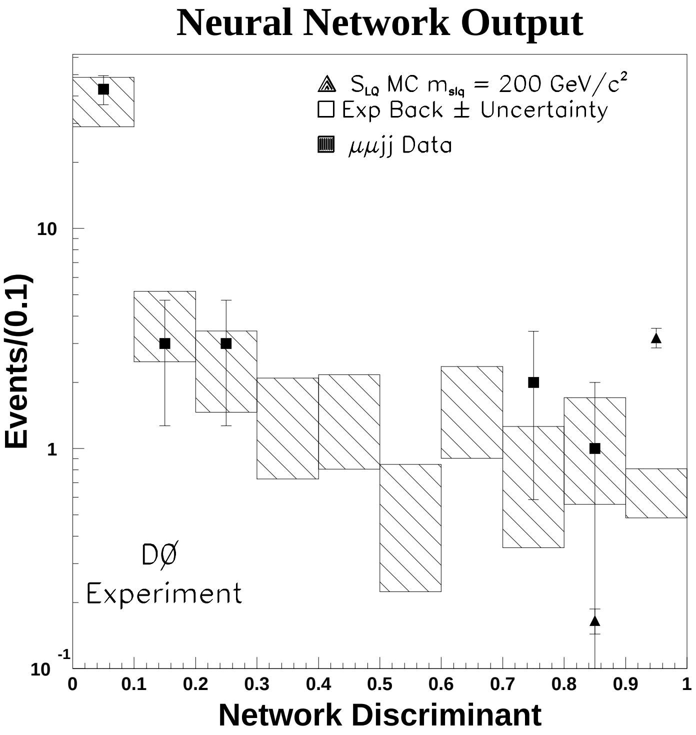

Figure 5.4. Neural Network output.

signal expectation is less than 0.1 events/bin until $D_{nn} > 0.8$, while the background is mostly well below this value.

We note here that it is this distribution which is used to normalize the $W + jets$ background (discussed in section 5.4). The normalization is performed in the two lowest bins shown on the plot in fig. 5.4, $D_{nn} < 0.2$. We do this because, in this region there is a very small signal expectation (less than 0.05 events expected) so that the normalization is not accidentally accounting for signal events in the data. Also, since we will ultimately place a selection requirement on D_{nn} which is greater than 0.2, the events used to normalize the $W + jets$ background are not in the final sample so that we are not biasing the results by performing the normalization.

In order to convince ourselves that we understand what the network is doing we have developed a set of simple selections on the input variables of the network in an attempt to see what values will produce a high network output. These selections are: $E_T^{j_1} > 65$ GeV, $E_T^{j_2} > 30$ GeV, $p_T^{\mu_1} > 55$ GeV/c, $p_T^{\mu_2} > 25$ GeV/c, $E_T^{j_1} + E_T^{j_2} > 115$ GeV/c, $m_{\text{event}} > 400$ GeV/c^2, and $\frac{E_T^{j_1}+E_T^{j_2}}{\sum_{jets} E_T^{j_i}} > 0.6$. These requirements select those events which tend to have high values of the network discriminant. When we apply these selections along with the S_{CM} Selection, there are no events left in the data consistent with a background expectation of 0.7±0.4 events. The signal expectation is 2.7±0.3 events.

We return for a moment to the selection requirement of the sphericity shown in fig. 5.3. We note here that there is a correlation between S_{CM} and D_{nn}. This correlation is what makes the sphericity selection so useful in conjunction with the neural network. Although the network will select events according to the requirements listed above, it is possible for the network output to be artificially high due to an extremely large value of a single variable (in this case, the muon p_T). However, when this is the case, the energy distribution in the event becomes unbalanced with the result that the sphericity tends to be very low. In this way the probability that mismeasurement of a single object can propagate all the way through the selection process and cause a background event to pass is lowered significantly.

Returning now to the network output shown in fig. 5.4, we see that we do not have evidence of a signal. We demonstrate this in the following manner. First, we apply a loose selection requirement to the events, $D_{nn} > 0.6$. For this requirement we have 3 events in the data consistent with 4.2±1.2 expected background events. All three events in the data are from the $\mu_{Sp} - \mu_{Cal}$ selection, for the backgrounds there are 3.7±1.2 events expected from the $\mu_{Sp} - \mu_{Cal}$ selection and 0.6±0.3 from the $\mu_{Sp} - \mu_{Sp}$ selection. The signal expectation is 3.6±0.4 events (2.1±0.3 from the $\mu_{Sp} - \mu_{Cal}$ selection and 1.5±0.2 from the $\mu_{Sp} - \mu_{Cal}$ selection). Table 5.6 shows a bin by bin breakdown of the D_{nn} distribution. The number of expected signal and

Table 5.6. Number of events in each D_{nn} interval by event type.

D_{nn}	$t\bar{t}$	$Z+jets$	$W+jets$	WW	Total Bkg	Sig	Data
0.0-0.1	0.3±0.1	4.2±0.7	32±10	2.0±0.8	39±10	0.02±0.01	43
0.1-0.2	0.2±0.1	0.4±0.1	3.2±1.3	0.0±0.2	3.8±1.4	0.02±0.01	3
0.2-0.3	0.2±0.1	0.2±0.04	2.1±1.0	0.0±0.2	2.4±1.0	0.03±0.01	3
0.3-0.4	0.2±0.1	0.1±0.02	1.2±0.7	0.0±0.2	1.4±0.7	0.02±0.01	0
0.4-0.5	0.2±0.1	0.1±0.03	1.2±0.7	0.0±0.2	1.5±0.7	0.03±0.01	0
0.5.0.6	0.1±0.1	0.1±0.04	0.3±0.3	0.0±0.2	0.5±0.3	0.04±0.01	0
0.6-0.7	0.1±0.1	0.1±0.02	1.2±0.7	0.3±0.3	1.6±0.7	0.05±0.01	0
0.7-0.8	0.1±0.1	0.1±0.03	0.6±0.4	0.0±0.2	0.8±0.5	0.06±0.01	2
0.8-0.9	0.2±0.1	0.1±0.02	0.9±0.6	0.0±0.2	1.1±0.6	0.2±0.02	1
0.9-1.0	0.5±0.2	0.2±0.03	0.0±0.3	0.0±0.2	0.7±0.4	3.2±0.3	0

background events are shown in the table for ten bins in the D_{nn} (the numbers shown correspond to the distribution in fig. 5.4). We find no evidence of an excess over Standard Model expectations. The highest value of the D_{nn} found in the data 0.82.

5.5.1 Selection Requirement Optimization

Since we have found no evidence of an excess over Standard Model expectations in the data set, we shall conclude the search in the dimuon channel by setting a limit on the pair production cross sections for leptoquarks. To maximize our limit we shall choose the final selection on the D_{nn} by maximizing a measure of sensitivity. We define a significance [61] as our measure of sensitivity;

$$S(D_{nn}) \equiv \sum_{k=0}^{n} P(k,b) M_A^{95\%}(k,b,s(m_{LQ})). \tag{5.2}$$

In *eqn.*. 5.2, k is some possible number of events in the data, b is the expected background, and s is the expected number of signal events for various values of the leptoquark mass. $P(k,b) = e^{-b}b^k/k!$ is a Poisson coefficient which gives the probability of finding k events for a given background expectation b. For example, for an expected background of 1.2 events, the probability of actually obtaining 1 event is $P(1,1.2) = e^{-1.2}1.2^1/1! = 1.2e^{-1.2} = 36\%$, while the probability of obtaining 2 events would be 22%. The Poisson distribution is used as a prior assumption on the distribution of Standard Model events in the data. $M_A^{95\%}$ is an approximate 95% confidence level mass limit calculated from the expected signal and background and the various possible number of data k. The approximation made is that the errors on the signal efficiencies and the background are not included. There is approximately a 1% difference between the approximate calculation and the exact mass limit calculated by including all the appropriate uncertainties. $S(D_{nn})$ is then simply the weighted average of the mass limits one would expect given $k = 0, 1, \ldots, n$ ($P(n,k) < 0.05$) events in the data. We calculate the significance for a range of D_{nn} and set the D_{nn} selection to coincide with the maximum of $S(D_{nn})$.

From the plot shown in fig. 5.5 we choose the selection requirement on D_{nn} to maximize the sensitivity $S(D_{nn})$. We see from the plot that the sensitivity has a broad plateau from $0.85 \leq D_{nn} \leq 0.95$. We set the final selection requirement in the center of this region at $D_{nn} > 0.9$ (which corresponds to keeping only the last bin in fig. 5.4). For this selection we have no events remaining in the data consistent with a background expectation of 0.7±0.3 events. The signal expectation is 3.2±0.3 events. For a background expectation of 0.7 events, the probability of seeing no events in the data (calculated from $P(k,b)$ defined above) is approximately 50%, or about the same as the probability of seeing any events in the data.

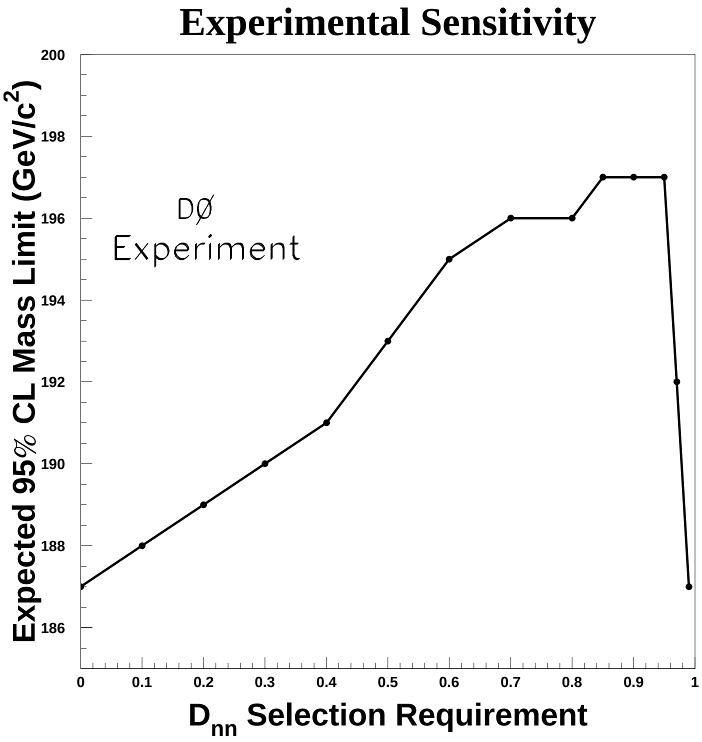

Figure 5.5. Experimental sensitivity in the $\mu\mu + jets$ decay channel.

Having now set our final selection it remains only to use this selection to obtain results for the leptoquark search in the dimuon channel. This is done in the next section.

5.6 Limits on Leptoquark Pair Production

With the selection process described in this chapter we find the signal detection efficiency for a range of leptoquark masses. From these efficiencies, the expected background of 0.7±0.3 and the fact that no data events remain after the selection is complete, we calculate the 95% confidence level cross section upper limit [62] for various possible leptoquark masses. It is important to note again that we make no distinction between the various possible couplings for leptoquarks, but present a single set of limits valid for any coupling. The signal detection efficiencies in the

Table 5.7. Results of the search for leptoquarks in the $\mu\mu + jets$ channel.

S_{LQ} Mass (GeV/c^2)	(eff±$stat$±sys) (%)	$\sigma^{95\%}_{Exp}$ (pb)	$\sigma_{S_{LQ}}$ (pb)	σ_{MV} (pb)	σ_{YM} (pb)
140	10.2±0.3±1.1	0.33	0.75	20	100
160	14.3±0.3±1.6	0.24	0.34	8.0	50
180	18.7±0.3±2.0	0.18	0.16	4.0	20
200	21.5±0.4±2.3	0.16	0.08	2.0	10
220	22.4±0.4±2.4	0.15	0.04	0.90	5.0
240	23.3±0.4±2.5	0.15	0.02	0.45	2.5
260	23.1±0.4±2.5	0.15	0.01	0.25	1.2
280	25.8±0.5±2.8	0.13		0.12	0.60
300	25.2±0.5±2.7	0.13		0.06	0.35
350	25.4±0.5±2.7	0.13			0.06
400	25.4±0.5±2.7	0.13			

dimuon channel, as well as the 95% confidence level upper limits are shown in table 5.7. Also shown for comparison are the minima of the NLO scalar cross sections and the LO cross sections for minimal vector and Yang-Mills couplings.

The limits shown in table 5.7 were calculated using a Bayesian [63] approach with a flat prior distribution for the signal cross section. The statistical and systematic uncertainties on the signal efficiency, the integrated luminosity, and the background estimation are included in the calculation with Gaussian prior distributions. For more detail on the calculation of cross sections limits see Appendix A.

From the cross section limits in table 5.7 we obtain a mass limit for second generation leptoquarks. We do this by comparing the experimental cross section

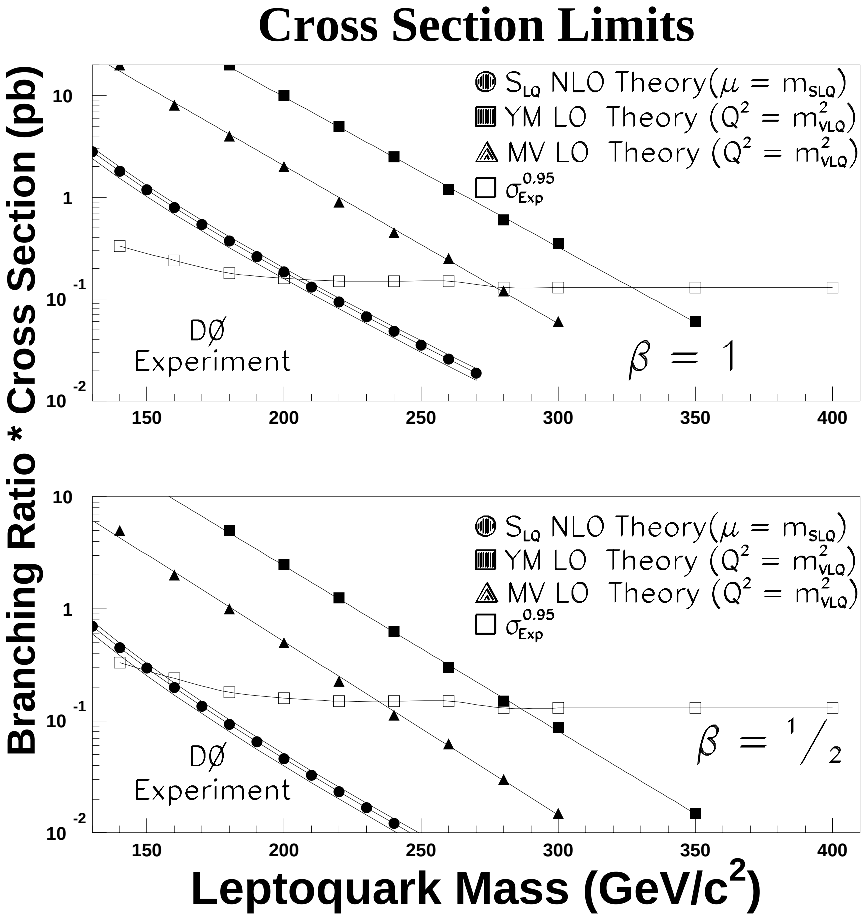

Figure 5.6. Mass limits at the 95% confidence level from the $\mu\mu + jets$ LQ search.

limits with the theoretical cross sections shown in the table. The comparison is shown in fig. 5.6. The intersection of the experimental curve with the lower edge of the theoretical cross section is taken as the mass limit.

The plots in figure 5.6 are shown for both scalar and vector couplings for two values of the branching ratio into charged leptons, $\beta = \frac{1}{2}, 1$. The mass limits for second generation leptoquarks obtained from these plots are 200, 275, and 325 GeV/c^2 for scalar, minimal vector and Yang-Mills couplings with $\beta = 1$. For

$\beta = \frac{1}{2}$ the results from the dimuon channel are 145, 230, and 285 GeV/c^2 for scalar, minimal vector, and Yang- Mills couplings respectively.

5.7 Summary of the Dimuon Search

Before continuing on to the second generation leptoquark search in the muon + neutrino + jets channel, we summarize here the results from the search in the dimuon channel. We have used the MTC software package to search for muons in the DØ calorimeter. Muons found in this fashion are added to muons found in the central region of the muon spectrometer as dimuon events. These are then combined with dimuon events where both muons were found in the spectrometer. We require these events to also have at least two good jets and to pass a selection designed to reduce the rate of cosmic ray muons in the data.

Having selected good dimuon + dijet events we apply a loose selection on the sphericity of the event, $S_{CM} > 0.05$. We next optimize our selection on a neural network with a sensitivity calculation and set our selection at $D_{nn} > 0.9$. With these requirements we find no events in the data consistent with 0.7±0.3 expected background events. The signal expectation, calculated at the minimum ($\mu = 2m_{S_{LQ}}$) of the NLO cross section for a scalar leptoquark mass of 200 GeV/c^2, is 3.2±0.3 events. From the signal detection efficiencies we obtained experimental limits on the pair production of second generation leptoquarks at DØ according to the prescription set forth in Appendix A.

These experimental cross section limits are turned into 95% confidence level mass limits by plotting the cross section limits and finding the value of the leptoquark mass for which the experimental limits intersect the theoretical cross sections. These mass limits are quoted as the results for the dimuon channel found by the search in this chapter.

We conclude this chapter on the dimuon search by considering the gain in efficiency obtained in this channel by using MTC muons. For a leptoquark mass of 200 GeV/c^2 the total efficiency is 21.5%. Of this, approximately 60% comes from including MTC muons. This is similar to the case for 180 GeV/c^2 scalar leptoquarks. By using MTC muons alone, and not including events with two muons found in the spectrometer, a mass limit of 182 GeV/c^2 would be obtained. This is almost the same as the previous DØ limit of 184 GeV/c^2.

Using the MTC package therefore allows us to extend the mass limit by 8% (from 184 GeV/c^2 to 200 GeV/c^2 for $\beta = 1$) by conducting a more complete search of the available data.

We turn our attention next to a search for leptoquark pair production in the $\mu\nu + jets$ decay channel. This work is described in the following chapter.

CHAPTER 6

THE $\mu + \nu + JETS$ DECAY CHANNEL

> "No scientific activity teeters on the brink between bravery and foolishness more than a search for undiscovered objects justified only by their necessity in theory". -Stephen J. Gould

6.1 Introduction

In this chapter we shall continue with the second generation leptoquark search by looking for leptoquark pair production in the $\mu\nu + jets$ decay channel [64–65]. This presents a complementary search to that described in chapter 5. Although the search described is complementary, that is, after the initial selection in this channel the data sets have no common events, many of the same techniques will be used. The reason we do this is that, in the previous search in the $\mu\mu + jets$ channel, we searched the single muon data and used the MTC [47] package to find additional muons. Therefore we already have a good technique with which to look for muons in the single muon data. Also, the backgrounds are the same.

Since there is so much in common between the two searches we shall adopt the following strategy. The data selection made at both the trigger level and in the initial selection process will not be described in full here. Rather, we shall note only the differences which exist between the two channels. The reader is to assume that anything which is not explicitly stated is the same as it was in the previous chapter.

6.2 Initial Event Selection

The event selection at the trigger level is the same as in the dimuon + jets selection with the following modification. We do not accept events which have two triggered muons. These we assume are dimuon events. This reduces our initial data sample from 4580 event to 4330 events, although the luminosity remains the same ($95 \pm 5\ pb^{-1}$), since the single muon and dimuon data sets were taken by different triggers.

In the initial selection we make the following modifications. In the jet selection we now require two or more jets with $E_T > 15$ GeV and $|\eta_j| < 2.0$. In addition, we require the leading jet have $|\eta_{j_1}| < 1.5$. The muon identification remains the same, but the kinematic selection is somewhat modified. We require that the muon have $p_T > 25$ GeV/c and $|\eta_\mu| < 0.95$. Additionally, we require only one central muon and no muons passing the requirements in the endcap spectrometers. The requirement that $p_T^\mu < 900$ GeV/c was removed since the same effect will be achieved with the selection for the neutrino. Also, we do not accept muons in the endcap spectrometers since in this region we expect a significant fall in the signal to background ratio. This was less serious in the previous channel since we were able to reduce it with the requirement of at least one central muon. Were we to allow single endcap muons, after the entire selection process which comprises this search was complete, we would gain approximately 9% signal efficiency at the cost of a 100% increase in the backgrounds.

The next step in the selection process is to look for $\not{E}_T$ which is consistent with a neutrino. As noted in chapter 4, neutrinos are not detected in the DØ detector. The presence of a neutrino is inferred from conservation of momentum in the plane transverse to the beamline. The calculation is performed in the calorimeter cells, and corrections are applied for the calorimeter geometry. We begin the $\not{E}_T$ selection by correcting the missing E_T from the calorimeter for any muons passing our selection criteria. This is done because muons are minimum ionizing particles and do not deposit substantial energy in the calorimeter. They therefore contribute to the missing transverse energy.

We require that $\not\!\!E_T > 30$ GeV. This selection is more stringent than the muon or jet energy requirements. It is done for two reasons. First, from the results of the previous chapter the mass limit for leptoquarks with a branching ratio $\beta = \frac{1}{2}$ is 145 GeV/c^2. We are now searching in a channel which has it's maximum contribution at $\beta = \frac{1}{2}$, and so we expect to be looking for leptoquarks in the mass region above 150 GeV/c^2. Therefore we expect that the neutrino from a leptoquark decay will be substantially more energetic than a neutrino from the decay of one of the vector boson (W, Z) backgrounds. Second, requiring $\not\!\!E_T > 30$ GeV reduces the probability that there is missing transverse energy due to jet energy mismeasurement or from uncertainties in the jet energy scale.

The selection requirement for the $\not\!\!E_T$ is that we do not accept events which have the missing transverse energy within $\pi \pm 0.1$ radians of the muon in ϕ. That is, we reject events where $\Delta\phi_{\mu\nu}\epsilon\ [\pi - 0.1, \pi + 0.1]$. This selection is made because of the possibility of mismeasurement in the muon p_T. If the muon p_T is significantly mismeasured, then the $\not\!\!E_T$ will point directly away from the muon track after the correction of the calorimeter based $\not\!\!E_T$ for the muon passing the muon selection requirements. If this happens, we are not able to say that there was another particle which contributed substantially to the $\not\!\!E_T$. We are thus unable to conclude that there is a neutrino in the event. This selection requirement is the one which allows us to eliminate the requirement $p_T^\mu < 900$ GeV/c in the muon selection. It accomplishes the same thing, but it is slightly less efficient. However, it is necessary here because we are looking specifically for a neutrino.

We next take the $\not\!\!E_T$ and use the MTC [47] package to search the calorimeter for tracks which are consistent with a muon. This is done in the same manner as in the previous chapter. As before, if we find a track which passes the track selection criteria, it is presumed to be a dimuon event. Since we are interested in a $\mu\nu + jets$ signature however, in this case we reject these events.

The final steps in the initial selection are to impose two requirements on the event topology. We require that the opening angle in ϕ between the muon and the missing transverse energy be greater than 0.3 radians. Additionally, we require that the separation in the $\eta - \phi$ plane between the two leading jets be greater than 1.4. The selections are then, $\Delta\phi_{\mu\nu} > 0.3$, and $\Delta R_{j_1 j_2} > 1.4$.

These topological requirements are made for the following reasons. In a leptoquark pair decay the two leptons (μ, ν) come from different parent particles. One leptoquark decays to μq, the other to νq. For this reason the leptons in the event are at most loosely correlated (there may be small correlations due to the spin for vector couplings). However, in background events, notably the $W + jets$ backgrounds, this is not so. The leptons in a W decay both come from the W $(W \rightarrow \mu\nu)$. One thus expects that the leptons will emerge back-to-back in the W rest frame. If the leptons are close together then we may infer that the W which produced them had a significant momentum (Lorentz boost). This boost is passed along to the daughter particles and they thus acquire a significant $p_T/\not\!\!E_T$. The resulting transverse energy distributions will appear in the detector to look very much like a signal event. Similar reasoning holds for the jets in the event. Applying these topological requirements has the effect of reducing those backgrounds which are most difficult to separate from signal events.

With the initial selection described we have 109 events remaining in the data. The data reduction is shown in table 6.1.

Table 6.1. Data reduction in the muon neutrino channel.

Selection Criteria	Events Passing
Trigger Selection	4330
$N_{jets} \geq 2$	3016
$N_{\mu} = 1$	473
$\not{E}_T$ Selection	142
$\Delta\phi_{\mu\nu} > 0.3$	137
$\Delta R_{j_1 j_2} > 1.4$	109

6.3 Background Studies

The Standard Model backgrounds which we consider here are the same as in the previous chapter. We consider $Z + jets$, $W + jets$, WW, and $t\bar{t}$ production. Each of these processes have decay modes to one or more muons as well as decay modes into taus with subsequent cascade decays to muons. Dimuon backgrounds cannot be neglected since the probability of one of the muons failing to be identified is not negligible.

We account for the Standard Model backgrounds in a fashion similar to that in the dimuon search. The $Z + jets$ cross section of 9.4 *pb* is used, along with the WW cross section of 10.4 *pb*. The calculation of the top quark pair production background uses the DØ measured cross section of 5.5±1.8 *pb*.

From our previous work in the dimuon channel, we know that the MTC package does not model fake tracks particularly well. We therefore do not apply the MTC selection to $\mu\nu$ Monte Carlo (*e.g.*, $W + jets$, signal or WW Monte Carlo). We use instead the fake rate of 10.8% from the dimuon search and reduce the efficiencies of these Monte Carlo events by this amount.

Table 6.2. Signal and background expectation after the initial selection.

Type	Events
$t\bar{t}$	2.4±0.8
$Z + jets$	2.7±0.7
$W + jets$	101±30
WW	1.5±0.6
Total Bkg	108±30
Exp Signal	4.9±0.7
Data	109

The $W + jets$ Monte Carlo is normalized to the data in the same fashion as in the dimuon search. We again look at a neural network [60] (to be introduced shortly, see figure 6.2) with $D_{nn} < 0.2$. The signal expectation in this region at a leptoquark mass of 160 GeV/c^2 is approximately 0.1 events. After the normalization the background expectation is 108±30 events.

We again note that the normalization of the $W + jets$ background is equivalent to inflating the cross section of 1000 pb by 19% and is therefore consistent with the normalizations found previously of 20% (leptoquarks in the eν channel [59]) and 24% (leptoquarks in the dimuon channel).

The expected signal and backgrounds are shown in table 6.2. For comparison, the number of events remaining in the data after the initial selection is shown. Note that all signal expectations in this chapter are presented for a leptoquark mass of 160 GeV/c^2. The calculation is performed at the minimum of the NLO cross section ($\mu = 2m_{S_{LQ}}$) for a branching ratio $\beta = \frac{1}{2}$. Systematic errors are the same as in the dimuon search and are shown in table 5-5.

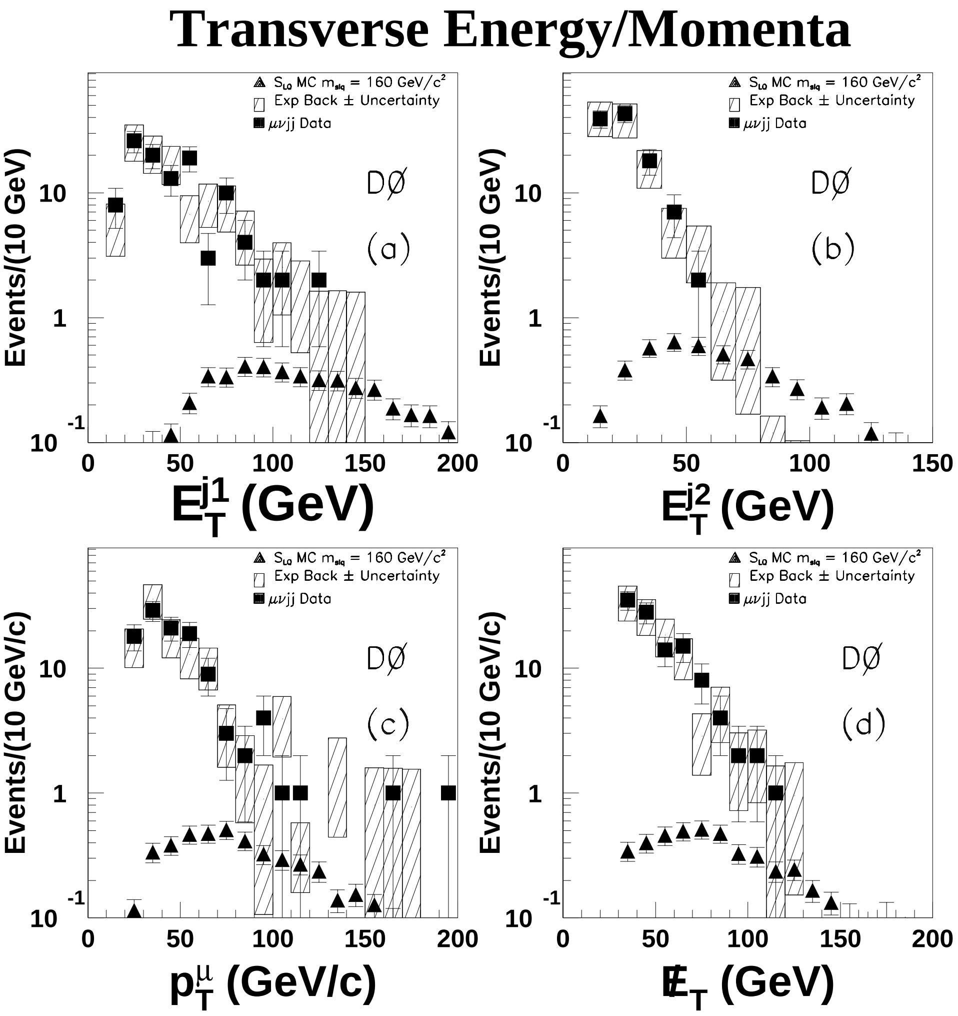

Figure 6.1. Transverse energy/momenta of the four leading objects.

The events passing the initial selection are shown in fig. 6.1. In the figure plot (a) shows the transverse energy of the leading jet, (b) shows the transverse energy of the second jet. Plot (c) is the transverse momentum of the muon while (d) shows the missing transverse energy (the transverse energy of the neutrino). As in the previous search, the muon p_T and jet E_Ts were varied by $\pm\sigma$ to test the effect of the variation on the acceptances. Variations were small and found to be accounted for by the systematic uncertainties list in table 5-5.

6.4 Final Event Selection

With an understanding of the backgrounds we turn our attention to finding and optimizing an appropriate set of selection requirements with which to separate signal and backgrounds. We take as our variable set the transverse energies/momenta of the four leading objects; $E_T^{j_1}$, $E_T^{j_2}$, p_T^{μ}, and $\not\!\!E_T$.

These four variables are used as the inputs to a neural network [60] with a structure 4-9-1. The network was trained on a signal Monte Carlo sample at a mass of 160 GeV/c^2. The sample contained approximately 1,600 events. The background sample used in the training was a mixture of $W+jets$, $Z+jets$, and $t\bar{t}$ events. Again, the WW Monte Carlo was not used in the training since training on the $W+jets$ sample is sufficient to account for the WW backgrounds. The background sample contained approximately 2,600 events. The number of independent variables in the network is 50, thus the size of the signal and background training samples is sufficient for the complexity of the network. The training was done on events passing the initial selection, but with the muon and neutrino requirements relaxed to $p_T^{\mu} > 10$ GeV/c and $\not\!\!E_T > 10$ GeV respectively. Also, the event topology requirements were not imposed for the training samples. This was done to insure that the sample size was sufficient for the complexity of the network, and so that the events to which the network is applied for the final selection is not the same set as the events it was trained on. The signal sample was an independently generated Monte Carlo set. The network discriminant is shown in figure 6.2.

From the distributions in fig. 6.2 we see that the data are well modeled by the backgrounds. Additionally, there is separation between signal and background. The signal expectation is less than 0.2 events per bin until approximately $D_{nn} > 0.6$, while the backgrounds are mostly well below this value. The distribution shown in the figure is that used to normalize the $W+jets$ Monte Carlo to the data. The

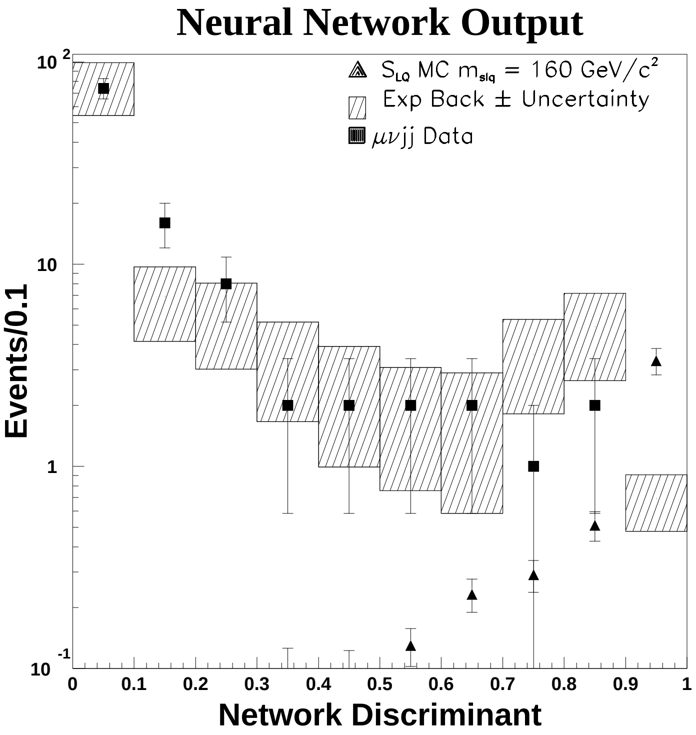

Figure 6.2. Neural Network output for events passing the initial selection.

events used are in the two lowest bins in the plots, $D_{nn} < 0.2$. The normalization is done before the event topology selections are applied (the last two selections in table 6-1).

Since the number of background events is still large. We again turn to a closer examination of the network distribution. This breakdown of the events remaining after the initial selection is shown in table 6.3.

From the distribution shown in fig 6.2 and the table 6.3, we see that the data is well modeled by the background and that we have no evidence for a signal in the data. We therefore proceed to set a limit on the pair production of second generation leptoquarks in the muon + neutrino + jets decay channel for a branching ratio $\beta = \frac{1}{2}$.

Table 6.3. Number of events in each D_{nn} interval by event type.

D_{nn}	$t\bar{t}$	$Z+jets$	$W+jets$	WW	Total Bkg	Sig	Data
0.0-0.1	0.2±0.1	1.9±0.5	74±22	0.6±0.3	77±22	0.04±0.01	74
0.1-0.2	0.2±0.1	0.3±0.1	6.1±2.8	0.4±0.3	6.9±2.8	0.1±0.02	16
0.2-0.3	0.2±0.1	0.1±0.03	5.3±2.5	0.0±0.1	5.5±2.5	0.1±0.02	8
0.3-0.4	0.2±0.1	0.1±0.03	3.0±1.7	0.1±0.1	3.4±1.8	0.1±0.02	2
0.4-0.5	0.1±0.1	0.1±0.03	2.3±1.5	0.0±0.1	2.5±1.5	0.1±0.02	2
0.5-0.6	0.1±0.1	0.1±0.03	1.5±1.2	0.1±0.1	1.9±1.2	0.1±0.03	2
0.6.0.7	0.2±0.1	0.1±0.02	1.5±1.2	0.0±0.1	1.7±1.2	0.2±0.04	2
0.7-0.8	0.2±0.1	0.1±0.02	3.0±1.8	0.3±0.2	3.6±1.8	0.3±0.1	1
0.8-0.9	0.3±0.1	0.1±0.03	4.5±2.3	0.0±0.1	4.9±2.3	0.5±0.1	2
0.9-1.0	0.6±0.2	0.1±0.03	0.0±0.9	0.0±0.1	0.7±0.9	3.3±0.5	0

6.4.1 Selection Requirement Optimization

We optimize our final selection requirement in the same fashion as was done in the $\mu\mu+jets$ decay channel. That is, we maximize the measure of sensitivity defined in *eqn*. 5-1 [61]. Fig. 6.3 shows the sensitivity, defined in *eqn* 5-1, as a function of the D_{nn}. The D_{nn} selection is chosen to coincide with the peak at $D_{nn} = 0.9$.

From the plot in fig. 6.3 we see that the sensitivity has a well defined maximum for a network discriminant value of 0.9. We set the network selection requirement to coincide with this maximum $D_{nn} > 0.9$ (corresponding to keeping only the final bin in fig. 6.2). For this selection there are no events remaining in the data consistent with an expected background of 0.7±0.9 events. The signal expectation is 3.3±0.5 events. We again note that, assuming a Poisson distribution for the backgrounds,

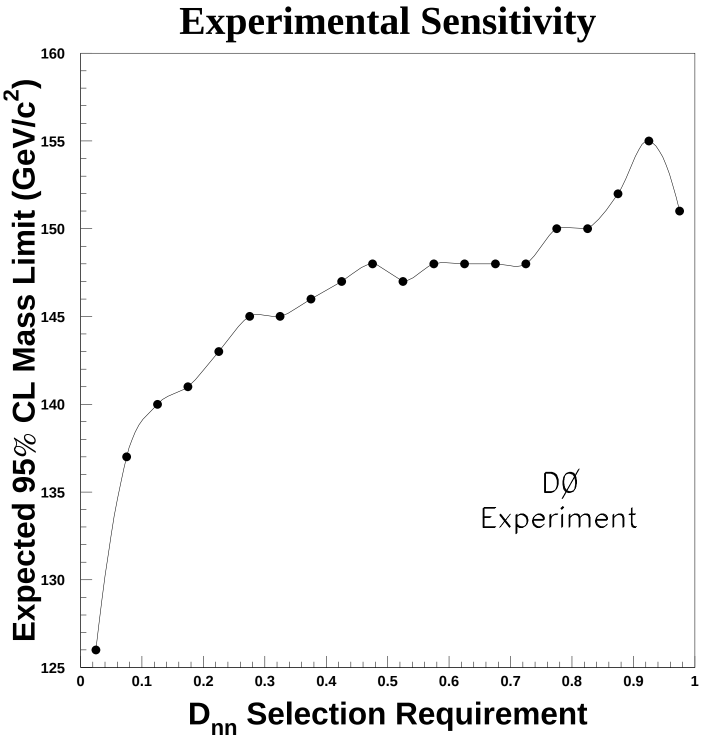

Figure 6.3. Experimental sensitivity in the $\mu\nu + jets$ decay channel.

the probability of obtaining no events in the data for a background of 0.7 expected events is approximately 50%.

6.5 Limits in the $\mu\nu + jets$ Decay Channel

The signal detection efficiency and experimental cross section limits [62] at the 95% confidence level are tabulated in table 6.4 for a wide range of leptoquark masses. As before, the minima of the theoretical cross sections are shown for scalar, minimal vector, and Yang-Mills couplings. The theory cross sections are shown for a branching ratio of $\beta = \frac{1}{2}$. Recall that the branching ratio β is defined as the branching ratio into charged leptons. Thus the branching ratio to neutrinos is given by $1 - \beta$. The branching ratio for a leptoquark pair to decay into $\mu\nu + jets$ is then $2\beta(1-\beta)$, where

Table 6.4. Results of the search for leptoquarks in the $\mu\nu + jets$ channel.

S_{LQ} Mass (GeV/c^2)	$(eff \pm stat \pm sys)$ (%)	$\sigma^{95\%}_{Exp}$ (pb)	$\frac{1}{2}\sigma_{S_{LQ}}$ (pb)	$\frac{1}{2}\sigma_{MV}$ (pb)	$\frac{1}{2}\sigma_{YM}$ (pb)
100	3.8±0.2±0.6	0.94	2.8	53	425
120	5.0±0.2±0.8	0.72	2.2	23	145
140	7.3±0.3±1.1	0.49	0.75	10	50
160	10.4±0.3±1.5	0.34	0.34	4	25
180	12.2±0.3±1.8	0.30	0.16	2	10
200	13.5±0.3±2.0	0.26	0.08	1	5
220	14.2±0.3±2.1	0.25	0.04	0.45	2.5
240	15.3±0.3±2.3	0.23	0.02	0.23	1.3
260	15.6±0.3±2.3	0.22	0.01	0.13	0.6
280	16.4±0.3±2.4	0.21		0.06	0.3
300	15.8±0.4±2.3	0.22		0.03	0.18
350	16.5±0.4±2.4	0.21			0.03
400	17.3±0.4±2.6	0.20			

the 2 arises because there are two particles which may decay in either mode. For $\beta = \frac{1}{2}$ the branching ratio to $\mu\nu + jets$ is also $\frac{1}{2}$. For details on the construction of confidence limits see Appendix A.

The cross section limits from table 6.4 are converted to a mass limit in the same manner as in the $\mu\mu + jets$ channel. Fig. 6.4 shows the branching ratio times cross section as a function of the leptoquark mass. The intersection of the experimental curve with the minima of the theoretical cross sections are the mass limits at the 95% confidence level. From the plot in fig. 6.4 we find the mass limits for leptoquark pair

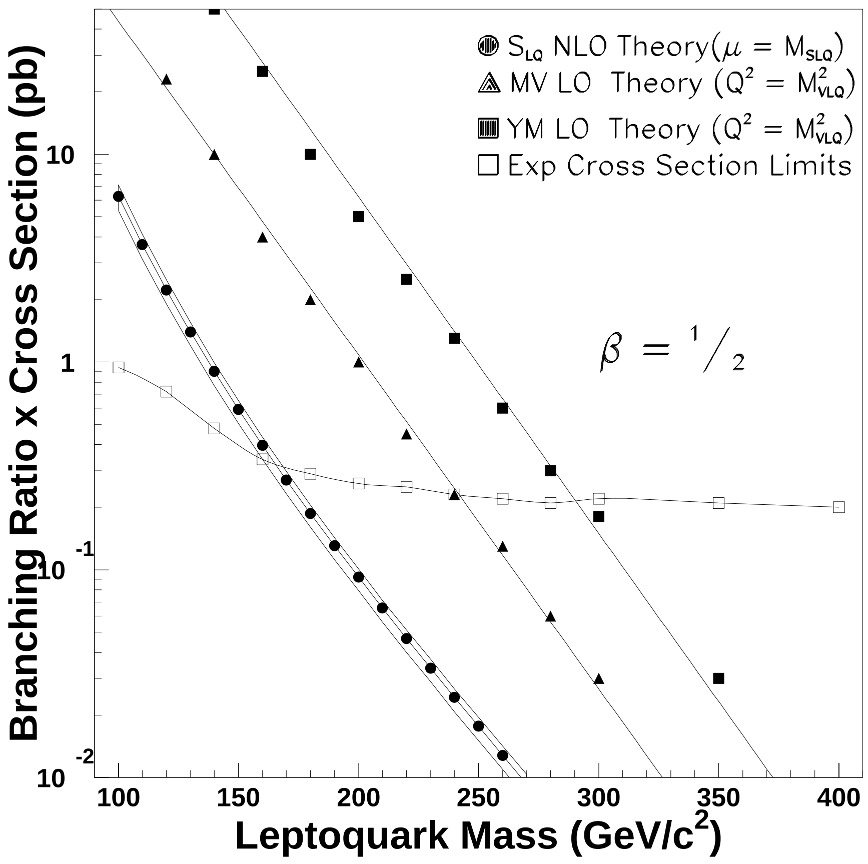

Figure 6.4. Mass limits at the 95% confidence level from the $\mu\nu + jets$ LQ search.

production in the $\mu\nu + jets$ decay channel for equal branching ratios to μq and νq to be: 160, 240, and 290 GeV/c^2 for scalar, minimal vector, and Yang-Mills couplings respectively.

We now have in hand everything which is needed to complete the search. In the next chapter we present the combination of the two searches described in this chapter and the previous chapter, along with the results from a previous leptoquark search in the $\nu\nu + jets$ ($\beta = 0$) decay channel. We shall then proceed to obtain the final results of this analysis.

CHAPTER 7

RESULTS

> "One should never bet against anything in science at odds of more than about 10^{12} to 1 against."–Ernest Rutheford

7.1 Introduction

We proceed to take the results obtained in chapters 5 and 6 and combine them to obtain the overall mass limits of the second generation leptoquarks. To see why this combination is both possible and useful, we consider the following case. Take, as an example, an equal branching ratio for μq and νq decays for leptoquarks. We shall then have $\beta = \frac{1}{2}$. Although the $\mu\mu + jets$ and $\mu\nu + jets$ data sets are complementary (that is, the intersection of the two sets is the empty set), both channels make a contribution to the total cross section of a leptoquark decay. To see that this is so we note again the branching ratios for the three possible decay modes for leptoquark pairs. The branching into $\mu\mu + jets$ is β^2. For $\mu\nu + jets$ the branching ratio is $2\beta(1-\beta)$, and for the $\nu\nu + jets$ decay mode the branching ratio is $(1-\beta)^2$. Although we have not conducted a search for leptoquarks in the $\nu\nu + jets$ channel, such a search has been done at DØ [66]. We shall use these results for the combination of channels. If one adds the three branching ratios the sum is 1, as it must be, regardless of the value of β. However, the relative contribution from each channel to the total varies as a function of β. For $\beta = \frac{1}{2}$ there is a 50% contribution from the $\mu\nu + jets$ channel, and 25% contributions from each of the other channels.

7.2 $\nu\nu + jets$ Decay Channel

For the combination we use the DØ results of the $\nu\nu{+}jets$ search. This search set limits on the pair production of scalar leptoquark pairs to be 79 GeV/c^2. Note that, since DØ does not directly detect neutrinos, no conclusion may be drawn about what type of neutrino it is. This limit is therefore independent of the leptoquark generation. The detector signature for this search was $\not\!\!E_T + jets$. It extended to leptoquarks with masses of 160 GeV/c^2. Monte Carlo signal events were later generated [67] to a leptoquark mass of 200 GeV/c^2 and the search was applied to these signal events. This demonstrated that the signal detection efficiency had reached a plateau for masses above 160 GeV/c^2. We therefore use the efficiencies from the $\not\!\!E_T + jets$ search, and for all masses above 160 GeV/ c^2 the efficiency at 160 GeV/c^2 is used.

7.3 Common Systematic Errors

The three channel combination assumes a 13% common systematic error [70]. Although each of the three data sets is complementary, all data taken by DØ comes from the same place, the Tevatron. Therefore the uncertainty on the total luminosity is common to each channel. Additionally there are common systematic errors due to the detector and the software used to reconstruct the events and so forth. Assuming a 13% common systematic uncertainty is a typical value when combining channels at DØ. Again, for details on the construction of confidence limits, see Appendix A.

7.4 Combining Channels at $\beta = \frac{1}{2}$

The 95% confidence level [62] cross section limits obtained by combining the three decay channels in the fashion described above are listed in table 7.1.

Table 7.1. Combined 95% CL cross section limits for leptoquarks with $\beta = \frac{1}{2}$.

S_{LQ} Mass (GeV/c^2)	$\sigma^{95\%}_{Exp}$ (pb)	$\sigma_{S_{LQ}}$ (pb)	σ_{MV} (pb)	σ_{YM} (pb)
200	0.29	0.16	2.0	10
220	0.26	0.08	0.90	5.0
240	0.25	0.04	0.45	2.5
260	0.24	0.02	0.25	1.2
280	0.22		0.12	0.60
300	0.23		0.06	0.35
350	0.23			0.06
400	0.22			

Note that the cross sections shown in table 7.1 are not multiplied by a branching ratio since we are accounting for all of the decay modes for leptoquark pairs. Thus the branching ratio is simply 1.

The cross section limits in table 7.1 are shown in fig. 7.1. The plot shows the experimental cross section limits from the three channel combination, as well as the three theory cross section curves for the couplings considered in this work. From the figure we obtain the 95% confidence level mass limits for leptoquark pair production at $\beta = \frac{1}{2}$ from the combined searches. The combined limits are 180, 260, 310 GeV/c^2 for scalar, minimal vector, and Yang-Mills couplings respectively.

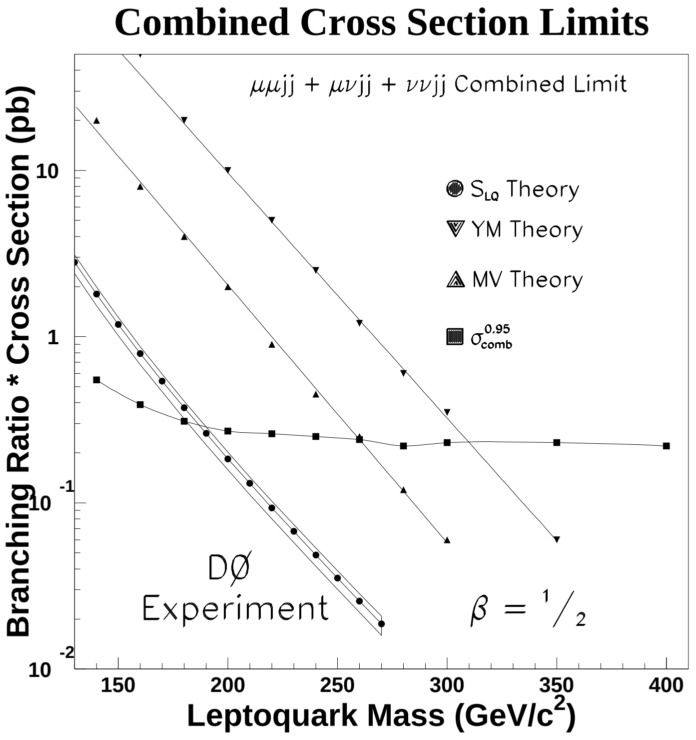

Figure 7.1. Combined leptoquark mass limits for $\beta = \frac{1}{2}$.

7.5 Results of the search

To obtain the results of this search, we follow the same procedure outlined in the previous sections. We present, in table 7.2, a summary of the mass limits obtained by this search for $\beta = 0, \frac{1}{2}, 1$.

We then compute mass limits for various values of β between 0 and 1. This has the effect of altering the relative contribution of each channel to the total cross section. For each value of β we obtain a mass limit for each of the couplings under consideration. These limits are then plotted in the β vs. m_{LQ} plane. The result is an exclusion contour which shows the restrictions placed on simultaneous values of the branching ratio to charged leptons and the mass of a second generation leptoquark. These exclusion contours are shown in figs. 7.2, 7.3, and 7.4 for scalar, Yang-Mills,

Table 7.2. 95% confidence level mass limits for $\beta = 0, \frac{1}{2}, 1$.

β	Scalar(GeV/c^2)	MV(GeV/c^2)	YM(GeV/c^2)
1	200	275	325
$\frac{1}{2}$	180	260	310
0	79	160	205

and minimal vector couplings respectively. They represent the results of the analysis. The area to the left of the outer most curve in each of the figures is excluded by the this search.

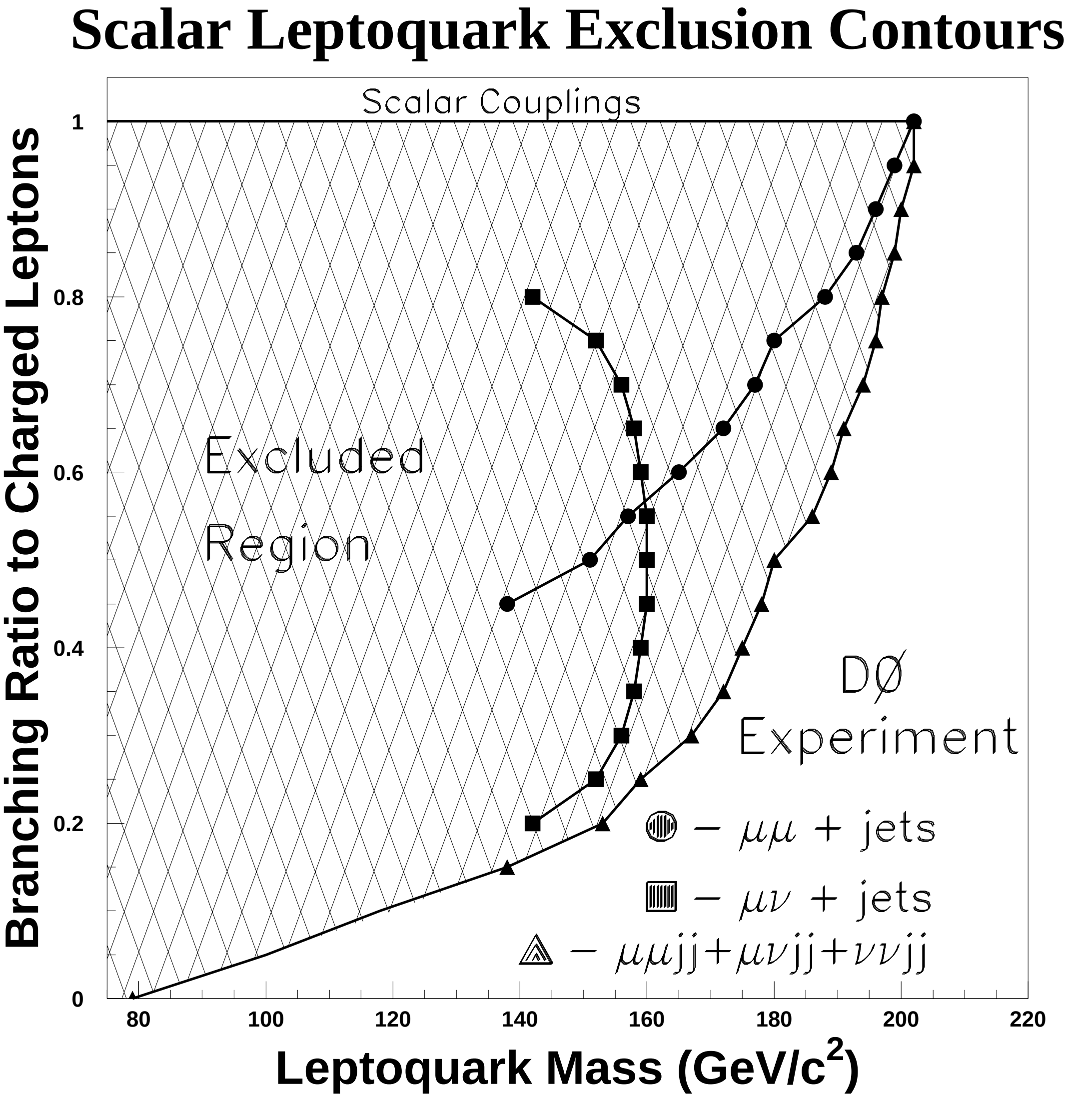

Figure 7.2. Exclusion contours for leptoquarks with scalar couplings.

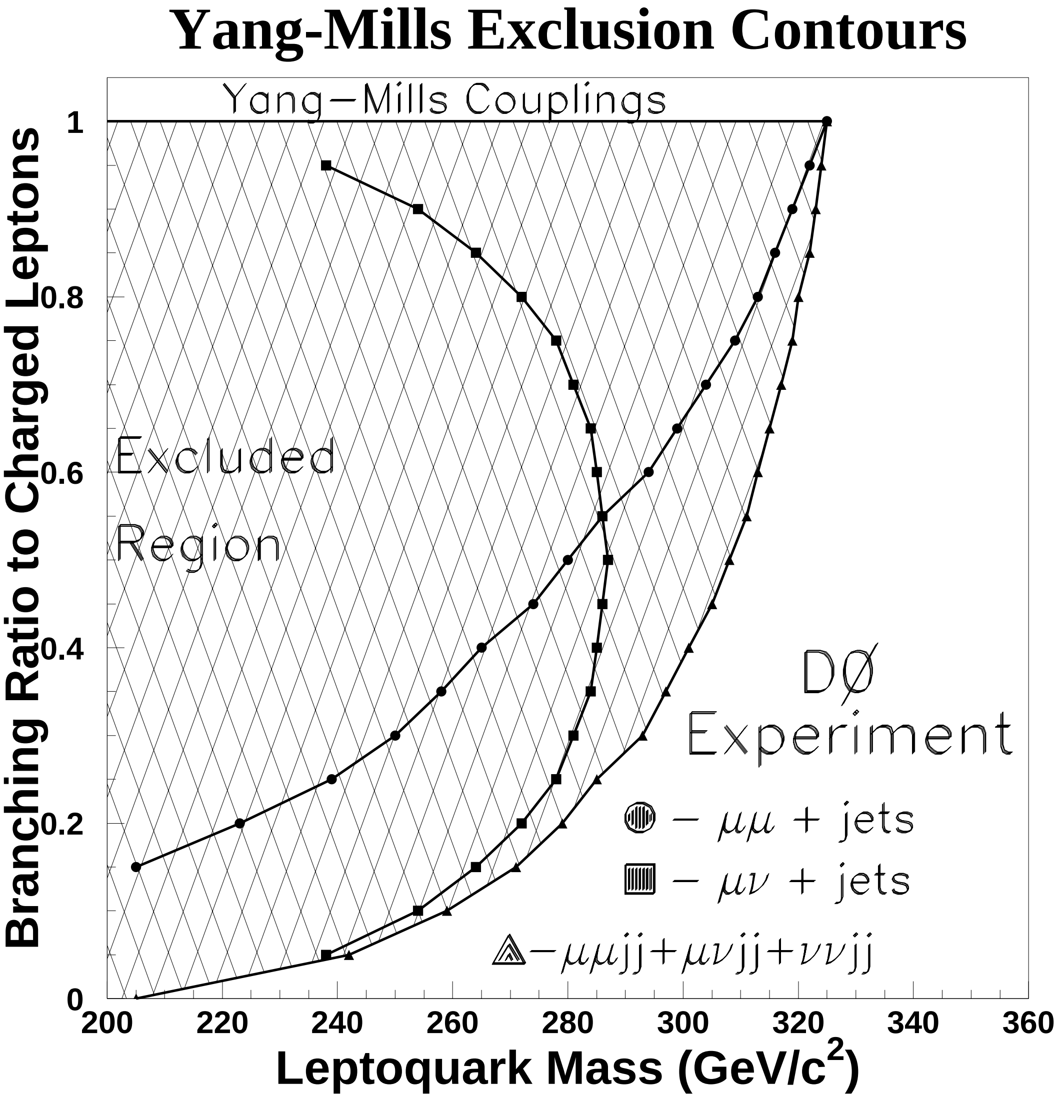

Figure 7.3. Exclusion contours for leptoquarks with Yang-Mills vector couplings.

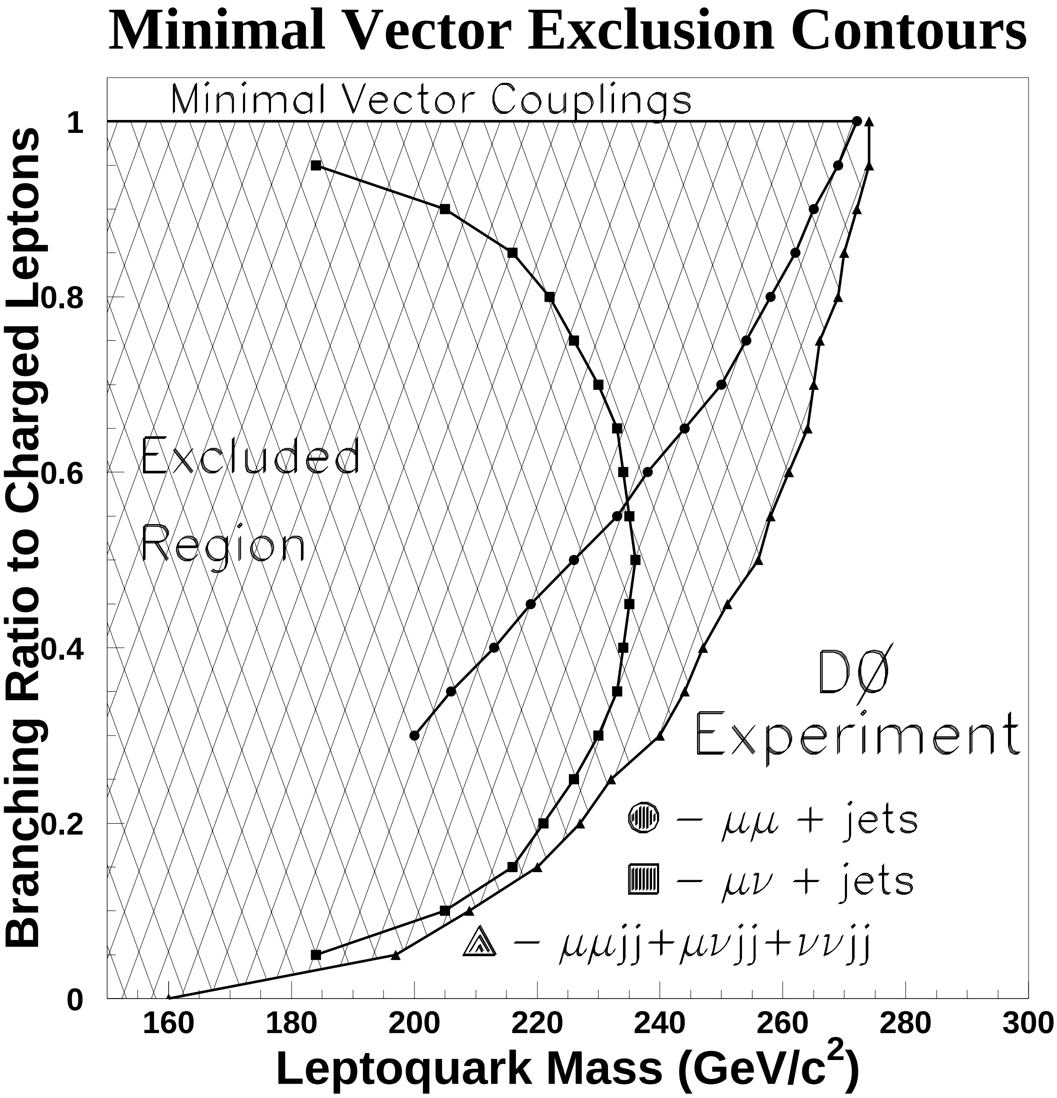

Figure 7.4. Exclusion contours for leptoquarks with minimal vector couplings.

CHAPTER 8

CONCLUSIONS

> “The most exciting phrase to hear in science, the one that heralds the most discoveries, is not ’Eureka!’ (I found it!) but ’That’s funny...?’ – Isaac Asimov

We conclude this work by presenting a brief review of the work described in the preceding chapters. We began a search for second generation leptoquark pairs with the DØ detector at the Fermilab Tevatron. The Tevatron accelerator has a center-of-mass energy $\sqrt{s} = 1.8$ TeV. We search for the pair production of second generation leptoquarks in 94 pb^{-1} of data collected by the DØ collaboration from 1993-96.

We considered leptoquarks with scalar, minimal vector ($\kappa_G = 1, \lambda_G = 0$), and Yang-Mills ($\kappa_G = 0, \lambda_G = 0$) couplings. We searched for leptoquark pairs decaying into muons and jets, as well as muons, neutrinos and jets.

In the dimuon + jets decay channel we use the excellent calorimetry available at DØ to search for muons within the calorimeter, in addition to muons identified from information gathered from the muon spectrometer. Using a complementary selection process that considered both event topology and energy, we find no evidence for a leptoquark signal in the data.

We next proceeded to search for leptoquark pair production in the muon, neutrino and jets decay channel. We again use the MTC software package. In this case, however it is used as a veto on dimuon events. Again using a complementary selection process accounting for event topology and energy, we find no evidence of a leptoquark signal within the data.

Having found no evidence for leptoquarks within the DØ data, we then proceed to set limits on the pair production cross sections and mass of the leptoquark for both scalar and vector couplings. The result is a set of exclusion contours in the β vs. m_{LQ} plane (figs. 7.2,3,4). We quote as our results the mass limits for three specific values of the branching ratio for leptoquarks into charged leptons, $\beta = 0, \frac{1}{2}, 1$. (The $\beta = 0$ result is not derived in the work, rather we use the results of a previous search in the $\nu\nu + jets$ decay channel). For $\beta = 0$ we find mass limits of 79, 160, 205 GeV/c^2 for scalar, minimal vector, and Yang-Mills couplings respectively. For $\beta = \frac{1}{2}$ we obtain limits of 180, 260, 310 GeV/c^2 for scalar, minimal vector, and Yang-Mills couplings. For a branching ratio $\beta = 1$, the limits are set at 200, 275, 325 GeV/c^2 for scalar, minimal vector, and Yang-Mills couplings, respectively.

These results extend the previous DØ limits by 8% ($\beta = 1$ for scalar Leptoquarks) and 22% ($\beta = \frac{1}{2}$). The dramatic increase in the mass limit for $\beta = \frac{1}{2}$ comes about because we have considered the $\mu\nu + jets$ decay channel, which the previous work did not include. The current limits on second generation leptoquarks from CDF are 202 GeV/c^2 for $\beta = 1$ and 160 GeV/c^2 for $\beta = \frac{1}{2}$. Both of these limits are obtained from a study of the $\mu\mu + jets$ decay channel, and did not consider the $\mu\nu + jets$ decay mode. Additionally, neither of these previous searches considered second generation Leptoquarks with vector couplings. Additionally, in order to compare the results of DØ and CDF for $\beta = 1$, it is necessary to note that the two collaborations set mass limits in slightly different ways. DØ sets a mass limit where the experimental cross section limits are equal to the minimum of the theoretical cross section ($\mu = 2m_{S_{LQ}}$). This is a matter of making a conservative claim. CDF sets mass limits at $\mu = \frac{1}{2}m$. For scalar leptoquarks the difference in the NLO cross section between $\mu = \frac{1}{2}m_{S_{LQ}}$ and $\mu = 2m_{S_{LQ}}$ is approximately 8 GeV. Had we set the limits in this work in the way that CDF does, the result for scalar leptoquarks at $\beta = 1$ would be 208 GeV/c^2 instead of 200 GeV/c^2. Thus, while the DØ result is slightly less than the limit from CDF, it represents a more stringent criteria.

We have thus presented in this dissertation results on searches for second generation leptoquarks which are more complete than the previously existing work. In those cases where limits previously existed from the collider experiments at Fermilab, the results presented herein are better than or comparable to the existing limits.

APPENDIX A

CALCULATING LIMITS

> "There are an essentially infinite numbers of stars in the universe. There are, however, only a finite number of habitable planets. Not all of the planets are actually inhabited, so the number of inhabited planets is an even smaller finite number. Since any finite number divided by infinity is zero, the population density of the universe is exactly zero, and anyone you meet is figment of your imagination."–Douglas Adams, from *The Hitchhikers Guide to the Galaxy.*

Herein we outline the procedure used by the DØ collaboration to construct 95% confidence level cross section limits. This procedure has been set forth in reference 62 on which we have relied for the outline in this appendix. For a complete treatment of the theory of probability the reader is recommended to consult reference 68. Further information about the construction of confidence limits may be obtained from references 69 and 70.

For some discrete variable, A, the probability of the proposition A, given that another proposition, B, is true is $P(A|B)$. For continuous variables, x, the probability becomes a probability density function denoted in the same fashion, $P(x|B)$. The probability of X having it's value between X and $X + dx$, given B, is $P(x|B)dx$. To denote all of the relevant information known prior to the calculation of interest, we use the variable I.

The DØ collaboration uses a Bayesian technique to set cross section limits. That is, the calculation is based upon Bayes' Theorem.

$$P(A|BC) = \frac{P(B|AC)P(A|C)}{P(B|C)} \tag{A.1}$$

Bayes' theorem relates the prior information to the knowledge of the parameters obtained after the data have been studied. The relationship proceeds through the likelihood function and some normalization constant determined from

$$\sum_{allA} P(A|BC) = 1. \tag{A.2}$$

We shall relate the general form of Bayes' theorem to the specific case under discussion later in the text. The calculation follows several steps, which we outline below.

The first step is to define the model. For a new particle search the accepted model is the total number of expected events from signal and background, μ. μ is related to the signal cross section σ, the signal detection efficiency ϵ, the integrated luminosity $\mathcal{L}$, and the expected background b.

$$\mu = b + \mathcal{L}\epsilon\sigma \tag{A.3}$$

For various cases the expected background may be a more complicated expression consisting of many different processes, each with it's own acceptance, cross section, branching ratios, *etc.*

The second step is to determine the likelihood function for the data, given the model. For a counting experiment such as this work the accepted likelihood function is a Poisson distribution with some expectation value (mean) μ. The probability of observing k events in the data given an expectation value μ is

$$P(k|\mu, I) = \frac{e^{-\mu}\mu^k}{k!} \tag{A.4}$$

where I is again all of the relevant prior information. We see here a similarity between *eqn.* A-2 and the definition of $P(k,b)$ from *eqn.* 5-2. Indeed, these were the considerations which lead to the development of *eqn.* 5-2 as a useful form of determining optimal selection criteria. Given the model described above, the likelihood function is

$$P(k|\sigma,\mathcal{L},\epsilon,b,I) = \frac{e^{-(b+\mathcal{L}\epsilon\sigma)}(b+\mathcal{L}\epsilon\sigma)^k}{k!} \tag{A.5}$$

giving the probability of observing k events in the data given $\sigma, \mathcal{L}, \epsilon$, and b.

We next assign the prior probabilities for each of the parameters. The available information available prior to the analysis of the data (*e.g.,* the integrated luminosity within some error) is used to determine the prior probabilities for each parameter in the likelihood function. For this case there are no correlations between the parameters and the probability density may be factorized into a product of independent prior functions. $P(\sigma,\theta|I) = P(\sigma|\theta I)P(\theta|I) = P(\sigma|I)P(\theta|I)$, assuming that σ is not dependant on θ so that $P(\sigma|\theta I) = P(\sigma|I)$, where $\theta = \theta(b,\mathcal{L},\epsilon)$ represents all of the parameters *except* for σ, the signal cross section.

Our knowledge of the efficiency, integrated luminosity and the background estimation are included as Gaussian prior distributions.

$$P(x|I) = \begin{cases} \frac{1}{\sigma_x\sqrt{2\pi}}e^{-\frac{(x-\overline{x})^2}{2\sigma_x^2}} & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases} \tag{A.6}$$

where x may be ϵ, $\mathcal{L}$, or b.

The signal cross section is more problematic since the theory of "uninformative priors" is the least developed portion of Bayesian probability. We use a finite flat prior of the form

$$P(\sigma|I) = \begin{cases} 1/\sigma_{max} & \text{if } 0 \leq \sigma \leq \sigma_{max}; \\ 0 & \text{if } \sigma \leq 0 \end{cases} \tag{A.7}$$

where σ_{max} is chosen sufficiently large that the probability of $\sigma > \sigma_{max}$ is negligible.

We next apply Bayes' theorem to find the posterior probability. That is, Bayes' theorem is used to invert the likelihood function from a probability of finding k events in the data given I, to a likelihood for a signal cross section given $k|I$. The connection of the abstract statement of Bayes' theorem in *eqn.* A-1 to the case in hand is as follows.

1.) A $\equiv$ signal cross section between σ and $\sigma+d\sigma$, integrated luminosity is between $\mathcal{L}$ and $\mathcal{L}+d\mathcal{L}$, signal detection efficiency between ϵ and $\epsilon+d\epsilon$ and background is between b and $b+db$;

2.) B $\equiv$ k events are observed in the data;

3.) C $\equiv$ all relevant prior information, including the description of the knowledge of the parameters $\sigma, \mathcal{L}, \epsilon,$ and b along with the assumptions used in building the model.

Bayes theorem then becomes

$$P(\sigma, \theta|k, I) = N\frac{e^{-(b+\mathcal{L}\epsilon\sigma)}(b+\mathcal{L}\epsilon\sigma)^k}{k!}P(\sigma|I)P(\mathcal{L}, \epsilon, b|I) \tag{A.8}$$

where the constant, N, is determined from the normalization condition

$$\int_0^\infty d\sigma \int_0^\infty d\mathcal{L} \int_0^1 d\epsilon \int_0^\infty db P(\sigma, \mathcal{L}, \epsilon, b|k, I) = 1 \tag{A.9}$$

Next we integrate out the nuisance variables. Since we are interested in the cross section and not the integrated luminosity, the efficiency, or the background these are known as nuisance variables. We integrate over the appropriate variables to remove them. The resulting posterior distribution is then

$$P(\sigma|k, I) = \int_0^\infty d\mathcal{L} \int_0^1 d\epsilon \int_0^\infty db P(\sigma, \mathcal{L}, \epsilon, b|k, I) \tag{A.10}$$

$P(\sigma|k, I)$ is now the probability distribution for a signal cross section given the observed number of events in the data, k, and the relevant prior information I. The 95% confidence level cross section upper limit is obtained by requiring that

$$0.95 = \int_0^{\sigma_{UL}} d\sigma P(\sigma|k, I) \tag{A.11}$$

where $P(\sigma|k, I)$ is the normalized posterior probability distribution and σ_{UL} is the cross section upper limit which we seek.

In practice the posterior distribution is calculated using a program called **LIMIT**. The program performs a Mote Carlo integration to solve the integrals and accounts for correlations between the errors.

Finally, we consider the combination of channels. For n independent channels (as we have in this work for $n = 3$), the probability for observing k events in the data is simply

$$P(k|\mu, I) = \prod_{i=1}^{n} P(k_n|\mu_n, I) \tag{A.12}$$

The rest of the calculation follows directly from the work already presented. For this case we take common systematic errors to be 13% since, although the channels are independent of each other, all of the data taken ultimately comes from the same source and is reconstructed by the same software, *etc*.

APPENDIX B

NEURAL NETWORKS

> "A neural network is just a computer program. That's all it is. It'll do what you tell it. So, if you tell it the wrong thing, it'll do the wrong thing. The old rule of GIGO applies" – Dan Karmgard, response to a question during a presentation.

B.1 Neural Network Overview

Since the analysis presented in this dissertation is heavily dependent on neural networks and they are all to often treated as "black boxes", we shall endeavor in this appendix to set forth the information necessary so that the reader may understand how the network operates. It will be useful to keep the quote at the beginning of the chapter in mind. Neural Networks are much like any other computer program and GIGO (Garbage In, Garbage Out) applies. Neural Networks are not magic bullets and their use does not mean that no thought is required.

They do have several strengths. First, they provide an efficient means of mapping several variables onto a single variable. This makes the process of picking a selection criteria much easier. Second, the neural network is nonlinear. This means that the selection criteria is not necessarily a box in the space of the variables used. To see this, imagine that we wish to select events on the basis of two variables, (x, y). If we make a plot of the events in our sample in these two variables we see some distribution in the scatter plot. Requiring $x > x_0$ and $y > y_0$ will impose a box on the scatter plot. Using a neural network one could have a curve. Imagine for a moment that the curve is a circle of radius r. Then one could require that only events which lie within the circle in the $x-y$ plane such that $x^2+y^2 \leq r^2$ are accepted. This may be a more efficient means of keeping signal and rejecting background than a simple box. The network will probably not result in a simple circle, but the idea is the same. Additionally, the network allows a simple extension into higher dimensional spaces. Using three or four or more variables in a network is no more difficult than using two, it is just much more difficult to see.

We shall begin this appendix by introducing the necessary terminology and jargon, and describe the type of network used in the analysis presented in chapters 5 and 6. We shall then proceed to work out the first few cycles of the training for a simple neural network by hand, so that the reader may become familiar with the way in which the network functions. Although the network we use as an example will necessarily be simple, generalization to more complex network structures is straightforward.

For reference, the network we have used in the analysis in chapters 5 and 6 is a modification of the Lund network using JETNET [71]. Additionally, we have found reference 72 to be an excellent work for the understanding of neural networks.

B.2 Neural Network Basics

Neural networks are an attempt to mimic, with a computer program, the way in which learning takes place in humans. A stimulus is presented to the senses (the *Input Layer*), the information is processed internally (the *Hidden Layer*), and a conclusion is drawn (the *Output Layer*). The goal of the network is pattern recognition, something at which people excel but computers have problems with. The networks used in this work are all trained with a process known as *Supervised*

Learning. Events are presented to the network at the input layer during training. The network is told what the event is to represent (signal or background) and what the desired result is (output of 1 for signal and output of 0 for background). The network processes the inputs through the hidden layer and then compares the result at the output layer with the result that is desired. This comparison is then used to modify the processing earlier in the network to bring the actual output closer to the desired output. The process continues until the updating results in changes that are less than some specified amount. That is, the network output converges to some particular value, within some given tolerance.

Once convergence has taken place, the network structure and the values it uses for internal processing are written out as a FORTRAN program which may be used on other events. These new events are presented at the input layer and a value for the output is calculated. The output for the networks in this work is between 0 and 1. The value of the network output may be thought of as the probability that a particular event is a signal event. That is, for output values which are close to 1, the event in question fits the pattern of a signal event very well, while values close to 0 imply that the event does not fit the hypothesis of a signal event well.

Although a neural network may have any structure desired, the networks we use are all of the same basic type. They are known as *Feed-Forward Back-Propagating* networks. Back-propagation refers to the manner in which the network updates the internal values during training. Feed-forward refers to the direction in which values are propagated through the network. We shall treat each of these in turn.

A network is composed of *Layers*, the layer structure is shown in fig. B-1. The diagram is for a 2-2-1 neural network. The circles on the left represent the inputs (labelled x and y). These *Neurons* constitute the input layer. The two circles in the middle represent the hidden layer, and the rightmost circle represents the

Neural Network Schematic

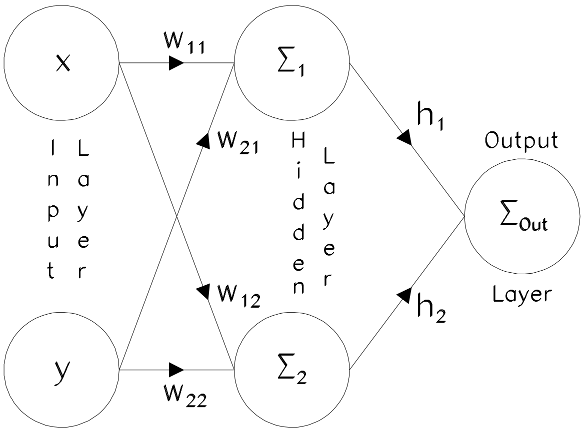

Figure B.1. Network diagram showing layering, connectivity, and the direction of information flow of a feed-forward neural network.

output layer. Each of the circles represents a value input to or calculated within the network. These values are known as neurons or nodes (nodes comes from the computer terminology to refer to the existence of a particular computer on a network). The arrows on the diagram show the connections between the neurons. Note that each neuron in a particular layer is connected to all of the neurons in the subsequent layer but that none of the neurons is connected to any other neuron within the same layer. This is what is meant by feed-forward. The connections are all in the direction of the flow of information through the network.

Back-propagation is the term used to describe how the network learns. On completing a calculation the network has a value at the output layer. This value is compared to the desired value (*e.g.,* 1 for signal events). A calculation is then made (to be done explicitly later) which takes into account the difference between the desired and actual output, the current weights in the network, the values output by the various neurons and the learning rate parameter. The learning rate is a number which is set for the network that governs how much of a change to introduce on each pass during the training. Once the error is determined and propagated backwards through the network a new cycle in the training begins. This continues until the values at the output layer change by less than a set amount between subsequent cycles, or until a specific number of training cycles have been completed.

In fig. B-1 each arrow has associated with it some number (w_{11}, *etc*). These numbers are the *weights*. The weights are the numbers which actually determine how the network calculates the output. When a given input is presented, the value at the first neuron in the hidden layer is calculated from $xw_{11} + yw_{21}$, while the value at the second neuron is calculated from $xw_{12}+yw_{21}$ [1]. These weights are the numbers which are changed during the training.

From the values which are passed into the network, the output of each of the neurons is calculated from a *Sigmoid Function*. There are many ways to calculate the output of the neurons of course, but this is what is used in the networks which we work with in chapters 5 and 6. The sigmoid function we use is defined as $g(x) = 1/(1 + e^{-2x})$, where x is the value of all connections into a neuron in the subsequent layer. Thus, the value coming into the first neuron in the hidden node is $1/(1 + e^{-\frac{2}{T}(xw_{11}+yw_{21})})$ and similarly for the second neuron. As a final matter, the input is scaled by a constant called the temperature T shown in the formula above. The temperature is a parameter which is set for the network and does not change.

---

[1] The network may also contain a threshold and a bias. The actual value at the first neuron is $g(xw_{11} + yw_{21} + b_1)\Theta_1(xw_{11} + yw_{21} + b_1 - \theta_0)$ with $\Theta(x - \theta_0)$ a step function and b_1 a constant. In order for a neuron to *fire* (pass its value to the next neuron) it must have a minimum value, and regardless of the inputs the minimum value will be b_1. Thresholds and biases are also determined during the training of the network. Our networks contain biases but no threshold functions. Also, in the LUND network, the biases are known as thresholds.

Putting the entire calculation together, the value at the output for a feed-forward network using a sigmoid function is

$$O_i(\vec{x}) = g\left[\frac{1}{T}\sum_j h_j g\left(\frac{1}{T}\sum_k w_{jk}x_k + b_j\right) + b_i\right] \tag{B.1}$$

where the w_{jk} and h_j are the weights and the b_i, b_j are possible bias values on the various connections. Again, we shall not use biases in the simple example network for this appendix. The subscript i in *eqn.* B-1 refers to the value of the i^{th} output node. In this work we use only networks with a single value at the output, so the subscript and summation may be safely ignored.

In the following sections we shall calculate explicitly the adjustment to the weights which takes place during training and the output of a completed network.

B.3 Network Training

There are a number of things to consider when designing and training a network. The first is the size of the network, which determines the complexity. The number of free parameters in a network is given by $I \times H + I + H + O$, where I = number of input neurons, H = number of neurons in the hidden layer, and O = number of neurons in the output layer. In order to have a sufficiently trained network one requires at least an order of magnitude more events on which to train than there are free parameters in the network. So, for the 2-2-1 network shown in fig. B-1 the number of free parameters is 9. Hence one requires a minimum of 90 events each of signal and background with which to train the network. A related question is how many neurons to use in the hidden layer. This is a subject for which there is no clear answer. We have used the rule of thumb that the number of hidden neurons should be in the range $2I \pm 1$ for the networks in this analysis. (The network we use as an example in this appendix uses only 2 hidden neurons for simplicity). The reason for the rule of thumb quoted above is that, for $H < 2I - 1$ there may not be a sufficient number of degrees of freedom with which to take advantage of the correlations that exist between the input variables. For $H > 2I + 1$ the network may attempt to use correlations that don't actually exist. In general, there is typically about a 1% difference in performance between networks with $2I - 1$, $2I$, and $2I + 1$ hidden neurons.

The next thing to note is that networks function best if the inputs presented are between the values of 0 and 1. Thus, the inputs to the network are normalized. In practice what this means is that one looks at the distributions of all events which are to be used in the training and divides the inputs by the highest value of the variables. So, in the network from chapter 5, the input of the leading jet E_T is actually $E_T/490$, this being the highest value of the leading jet E_T in the signal and background samples used in the training. In this way all of the inputs to the network are constrained between 0 and 1.

Finally, one must take care that the network is not biased by the training sample. There are two components to this. First, one requires a sufficient number of events on which to train. This way the network is not artificially influenced by individual events. Second, an independent sample is used for the training. In this fashion the network is not being used on the same events on which it was trained. In our case, the signal sample was a different Monte Carlo set. For the backgrounds, no such independent set was available. We therefore chose to train the network on events which had looser requirements than the initial cuts detailed in the analysis chapters. In this way the network training samples were not the same events as those on which it is used to determine expected signal and background estimates.

Finally, one must determine the temperature and learning rate parameters to use. As with the number of hidden neurons, these parameters are the subject of much passionate discussion. We state without justification that the values we used are

T = 1.0 and 0.01 for the learning rate. Additionally, we have used a training which requires 5,000 cycles and does not end early even if the network is converging.

B.4 Example of Network Training

We now turn to working out an example based on the network in fig. B-1. We shall take our desired outputs to be 1 and 0 for the training vectors (0.85,0.74) and (0.09,0.12) respectively. The weights are randomly assigned initially. The training is used to adjust the weights. The network contains 6 weights

$$[W] = \begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix} \tag{B.2}$$

for the weights between the input and hidden layers and

$$\vec{h} = (h_1, h_2) \tag{B.3}$$

for the weights between the hidden and output layers. The other three degrees of freedom are the biases which we have set to 0. The random values assigned at the beginning of the training are

$$\begin{aligned} [W] &= \begin{pmatrix} -0.1 & 0.18 \\ 0.42 & 0.8 \end{pmatrix} \\ \vec{h} &= (-0.1, 0.18) \end{aligned} \tag{B.4}$$

where we note that, for simplicity the initial weights assigned in $\vec{h}$ are the same as the upper row in $[W]$. The training of the network is broken down into *cycles* and *patterns*. A pattern is one of the input vectors being used for training, along with the desired network output for that vector. A cycle is one complete pass through all training patterns. Our training here will use 1,000 cycles over the two input patterns with one update (adjustment of the weights) per pattern. The first two cycles will be fully detailed and we shall then present the fully trained network.

In order to understand how the network is trained we shall require the following equations,

$$\begin{aligned}\Delta_{NN} &= \text{desired network output - actual network output}\\ b_i &= \Delta_{NN} \times h_i \times O_i^1 \times (1 - O_i^1)\\ h_i &= h_i + \Delta_{NN} \times \beta \times O_i^1\\ w_{ij} &= w_{ij} + I_i \times \beta \times b_j\end{aligned} \tag{B.5}$$

where Δ_{NN} is the error on the network output after a specific pass, b_i is the error to be back-propagated through the network, β is the learning rate (which we have set to 0.2 in order to speed up the learning process), O_i^1 the input into the i^{th} neuron in the hidden layer, and I_i is the i^{th} element of the input vector. Additionally, we shall denote the network output as a function of the input vector by $O_{NN}(x, y)$.

On the first pass of the training, the input into the hidden layer is calculated from the input vector $(0.85, 0.74)$. The input into the first neuron in the hidden layer is $1/(1 + e^{-2(I_1 w_{11} + I_2 w_{21})}) = 1/(1 + e^{-0.4516}) = 0.6110$. The input into the second neuron is $1/(1 + e^{-2(I_1 w_{12} + I_2 w_{22})}) = 0.8161$. From the hidden layer the output of the network is $1/(1 + e^{-2(0.6110 h_1 + 0.8161 h_2)}) = 0.5428 = O_{NN}$. From this we may calculate the quantities we require to adjust the weights. $\Delta_{NN} = 1.0 - O_{NN} = 0.4572$. The adjustments to the weights between the hidden and output layers are calculated from *eqns.* B-5 as

$$\begin{aligned}h_1 &= h_1 + \Delta_{NN} \times \beta \times O_1^1\\ &= -0.1 + 0.4572 \times 0.2 \times 0.6110\\ &= -0.04413\\ h_2 &= h_2 + \Delta_{NN} \times \beta \times O_2^1\\ &= 0.18 + 0.4572 \times 0.2 \times 0.8161\\ &= 0.2546.\end{aligned} \tag{B.6}$$

In order to calculate the adjustments to the weights between the input and hidden layers we must first calculate the errors which are to be back-propagated. From the definitions in *eqns.* B-5 we have

$$\begin{aligned} b_1 &= \Delta_{NN} \times h_1 \times O_1^1 \times (1 - O_1^1) \\ &= 0.4572 \times (-0.1) \times 0.6110 \times (1 - 0.6110) \\ &= -0.0108 \\ b_2 &= \Delta_{NN} \times h_2 \times O_2^1 \times (1 - O_2^1) \\ &= 0.4572 \times 0.18 \times 0.8161 \times (1 - 0.8161) \\ &= 0.0124. \end{aligned} \tag{B.7}$$

We may now obtain the adjustments to the weights in $[W]$.

$$\begin{aligned} w_{11} &= w_{11} + I_1 \times \beta \times b_1 \\ &= -0.1 + 0.85 \times 0.2 \times (-0.0108) \\ &= -0.1018 \end{aligned} \tag{B.8}$$

$$\begin{aligned} w_{12} &= w_{12} + I_1 \times \beta \times b_2 \\ &= 0.18 + 0.85 \times 0.2 \times 0.0124 \\ &= 0.1821 \end{aligned} \tag{B.9}$$

$$\begin{aligned} w_{21} &= w_{21} + I_2 \times \beta \times b_1 \\ &= 0.42 + 0.74 \times 0.2 \times (-0.0108) \\ &= 0.4184 \end{aligned} \tag{B.10}$$

$$\begin{aligned} w_{22} &= w_{22} + I_2 \times \beta \times b_2 \\ &= 0.8 + 0.74 \times 0.2 \times 0.0124 \\ &= 0.8018. \end{aligned} \tag{B.11}$$

After training on a single pattern $(0.85, 0.74)$ with desired output 1 the weights are adjusted to be

$$[W] = \begin{pmatrix} -0.1018 & 0.1821 \\ 0.4184 & 0.8018 \end{pmatrix} \qquad \text{(B.12)}$$

$$\vec{h} = (-0.0441, 0.2546)$$

whereupon the network proceeds to the next pattern. We shall continue to detail the training of the network, but henceforth we shall not detail the arithmetic involved in each calculation. The training is an iterative process and the formulae are the same for each update, only the weights and the quantities which are calculated from them change.

The next pattern presented to the network for training is the vector $(0.09, 0.12)$ with a desired output of 0. The inputs into the hidden layer are calculated from the vector and the updated weights from the previous pattern. The input into the first hidden neuron is 0.5205 and the input into the second hidden neuron is 0.5561. The network output is then 0.5590. From this we calculate the error in the network output to be $\Delta_{NN} = 0 - 0.5590 = -0.5590$. The errors to be back-propagated are $b_1 = 0.0062$ and $b_2 = -0.0351$. From these we find the updates to the weights to be made as a result of this pattern to be

$$[\Delta W] = \begin{pmatrix} 0.0011 & -0.0006 \\ 0.0001 & -0.0008 \end{pmatrix} \qquad \text{(B.13)}$$

$$\Delta\vec{h} = (-0.0582, -0.0622)$$

from which we obtain the new weights

$$[W] = \begin{pmatrix} -0.1017 & 0.1815 \\ 0.4185 & 0.8010 \end{pmatrix} \tag{B.14}$$
$$\vec{h} = (-0.1023, 0.1925)$$

which completes the first training cycle of the network. We now move to the second cycle of network training by revisiting the first input pattern with the revised weights. The network output from the first pattern in the second cycle is 0.5472 resulting in an error $\Delta_{NN} = 0.4528$. The back-propagated errors are $\vec{b} = (-0.0110, 0.0130)$ which result in a new set of weights

$$[W] = \begin{pmatrix} -0.1036 & 0.1837 \\ 0.4169 & 0.8029 \end{pmatrix} \tag{B.15}$$
$$\vec{h} = (-0.0471, 0.2664)$$

which are then used in the training of the network on the next pattern. The network output in the second cycle of pattern 2 is $O_{NN} = 0.5615$ which gives a network error of $\Delta_{NN} = -0.5615$. The errors to be back-propagated are $\vec{b} = (0.0066, -0.0369)$. From these quantities we calculate the updated weights to be

$$[W] = \begin{pmatrix} -0.1035 & 0.1830 \\ 0.4171 & 0.8020 \end{pmatrix} \tag{B.16}$$
$$\vec{h} = (-0.1055, 0.2039)$$

which completes the second cycle. We shall not present any more detail on the training of the network. We simply note that, after the full training of 1,000 cycles the weights are

$$[W] = \begin{pmatrix} -1.672 & 0.8216 \\ -0.1514 & 1.050 \end{pmatrix} \tag{B.17}$$
$$\vec{h} = (-7.589, 2.351)$$

and the output errors on the network output are 0.0220 for the first pattern and -0.0294 for the second pattern. Thus the network returns an output of 0.9780 for the training vector $(0.85, 0.74)$ and an output of 0.0292 for the training vector $(0.09, 0.12)$. We can see that, although the changes in any one training cycle are small, after many cycles the effect is quite dramatic.

Before proceeding to an examination of the network a final comment on training is in order. Having worked through the first two cycles of network training by hand we are now in a position to quantify the statement made at the beginning of this appendix that the use of a neural network does not mean that no thought is required. If one examines the updates that are made to the weights during the two training cycles we see that the weights tend to oscillate further and further apart as each pattern is used. This is because the training vectors were sufficiently different to be good discriminators on their own. Had this not been the case, the weights would have oscillated back and forth without achieving separation. That is, the changes to the weights induced by one pattern would have been largely undone by the next. In the end, the network would not have achieved significant separation between the patterns.

We may thus draw this conclusion, when used with a set of variables which are themselves good at discriminating signal from background, a neural network can improve the efficiency of retaining signal and rejecting background (this is usually quantified by attempting to maximize some form of $S/(S+B)$, where S stands for the expected signal and B the expected background). However, if the set of variables input to the network are not sufficient to discriminate signal from background on their own, the network will not be of any help. Thus the use of a neural network in an analysis requires sufficient consideration to determine a proper set of variables, as well as the appropriate constants which govern the training and operation of the network.

B.5 Neural Network Usage

We now proceed with the neural network which we trained in the previous section. We have run 10 vectors through the network and the network output for each vector is shown below.

$$\begin{aligned}
O_{NN}(0.91, 0.83) &= 0.9817 \\
O_{NN}(0.23, 0.32) &= 0.2671 \\
O_{NN}(0.01, 0.15) &= 0.0104 \\
O_{NN}(0.78, 0.42) &= 0.9642 \\
O_{NN}(0.50, 0.50) &= 0.8772 \\
O_{NN}(0.00, 0.00) &= 0.0053 \\
O_{NN}(1.00, 1.00) &= 0.9854 \\
O_{NN}(0.45, 0.76) &= 0.8821 \\
O_{NN}(0.75, 0.75) &= 0.9710
\end{aligned} \tag{B.18}$$

We now have a properly trained neural network which may be used to discriminate between two types of vectors. The “signal” vectors which are represented by $(0.85, 0.74)$ are those which will occupy the upper right-hand corner of the normalized $x-y$ plane and the “background” vectors, represented by $(0.09, 0.12)$ which lie in the lower left-hand corner of the $x-y$ plane. The discriminating power of the network is very good, but that is because the training patterns were selected purposefully so that they could be easily separated. Also note that we used only two training patterns for simplicity when the complexity of the network actually calls for training over approximately 200 patterns. Due to the simplicity of separating the two vectors used for the training however, we were able to achieve good results with a very small training set.

We now draw an analogy which will hopefully make clear the sort of thing that the network is doing. We are using the network to search a phase space defined by the $x-y$ plane between 0 and 1. Those vectors in *eqn.* B-15 which have a high value of the network output are also those which are closer to the "signal" vector. For example, consider the vector (0.5,0.5). The distance in the $x-y$ plane between (0.5,0.5) and (0.85,0.74) is 0.42 (from the Pythagorean theorem $d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$) while the distance to the "background" vector is 0.56. Thus the vector (0.5,0.5) is closer to the pattern represented by the signal vector and as such receives a high output from the network. Those vectors which are closer to the background vector all receive a low value of the network output. This is a very simple example of the pattern matching which was mentioned at the beginning of this appendix.

Suppose now that the network had been trained on a large sample of signal and background vectors. The adjustment of the weights in the network would have corresponded to selecting a representative pattern for signal and another to background. Each vector which was then run through the network would receive an output which classified how well it matched the patterns against which it was being compared. This would correspond to calculating the distance between the test vectors and the representative patterns within the network. The calculation of the distance is analogous to the operation of the network, but the network does internal calculations in a much more sophisticated manner.

B.6 Conclusion

We note that the network which we have explored in this appendix achieves good separation between signal and background due to the simplicity of the problem. In a real application, the networks used are much more complex. In addition, the problem of separating signal and background is much more complex. Most of the time the signal and background are not obviously separated in the phase space of the variables used in the network. Rather, there is some overlap.

To visualize this problem, imagine that the set of signal vectors in the $x - y$ plane was contained in a semicircle of radius 0.7 centered on the point $(1, 1)$ and the background occupied a circle of radius 0.6 centered about the origin. In each case we should also imagine that the density of points for each type is highest toward the center of the respective circles and falls slowly as one looks out toward the edges. This is a more realistic situation.

The signal and background may be separated, but there is significant overlap between them. The training of the network will pick representative vectors against which test vectors may be compared, and the network version of distance will be calculated. But, because of the overlap between signal and background, the network output will form a continuous curve with some background events receiving high scores from the network while some signal events receive low values. There will not be a way to place a selection cut on the network output without losing some signal and retaining some background as there is in the example presented here. Neural networks are not magic bullets, but used properly and with caution, they allow a significant increase in our ability to find a small signal in a large noise.

We conclude this appendix with the following example. People are exceptionally good at pattern recognition. When we encounter a chair which has a design different from any we have seen before, we are not confused. We recognize it immediately as a chair. We do this through pattern recognition. Our eyes report to us that this new object conforms to the pattern we know as chairness. It has some number of legs and a part on which to sit and (usually) a piece against which we may rest our back.

Neural networks are the attempt to teach a computer to do the same thing. The training routine of supervised learning (which corresponds to our childhood, supervised by our parents) presents to the network some number of patterns which represent different things (signal or background). The network learns by adjusting the weights between the various neurons. On completion of this learning process the network is used to attempt to match other vectors to the patterns it learned during the training period.

While neural networks may seem to be exotic “black boxes”, they are in fact just one of many tools to be used in data analysis. The process of learning is complex though not difficult to understand, but their use is fairly simple and straightforward. We have presented, in this appendix, the knowledge necessary to understand how a neural network functions and the reader is now better able to judge the analysis presented in this dissertation.

BIBLIOGRAPHY

[1] I. Mendeleev, *Annalen der Chemie und Pharmacie VIII*, Supplementary Volume for 1872, p. 511

[2] J. J. Thompson, *Philos. Mag.* **44**, 293 (1897).

[3] M. Planck, *The Theory of Heat Radiation*, 2^{nd} ed., P. Blakiston Son & Co., (1914).

[4] E. Rutheford, *The Scattering of the α and β Rays and the Structure of the Atom*, Proceedings of the Manchester Literary and Philosophical Society, IV, 55, pp. 18-20, (1911).

[5] J. Chadwick, *The Existence of a Neutron*, *Roy. Soc., Proc*, **136** pp. 692-708, June 1, 1932.

[6] P. A. M. Dirac, *The Principles of Quantum Mechanics*, 4^{th} ed., Oxford Clarendon Press, (1958).

[7] A. Einstein, *On the Electrodynamics of Moving Bodies*, *Annalen der Physik*, **17** (1905).

[8] C. Anderson, *The Positive Electron*, *Phys. Rev.*, **43**, 491 (1933).

[9] J. Schwinger, *Selected Papers on Quantum Electrodynamics*, Dover Publications, New York, (1958); R.P. Feynman, *QED: The Strange Theory of Light and Matter*, Princeton University Press, Princeton, NJ (1985).

[10] M. Gell–Mann, Phys. Rev. **125** 1067 (1962); M. Gell–Mann and Y. Ne'eman, *The Eightfold Way*, Benjamin Publishers, New York (1964).

[11] *Developments in the Quark Theory of Hadrons*, D. B. Lichtenberg and S. P. Rosen, eds., Nonanatum, Hadronic Press (1980).

[12] E.D. Bloom *et. al.*, *High Energy Inelastic e-p Scattering at 6° and 10°. Phys. Rev. Lett.*, **23**, 930 (1969); M. Breidenbach *et. al.*, *Observed Behavior of Highly Inelastic Electron Proton Scattering. Phys. Rev. Lett.*, **23**, 935 (1969).

[13] J. J. Aubert *et. al.*, *Experimental observation of a heavy particle J. Phys. Rev. Lett.*, **33**, 1404 (1974); J. E. Augustin *et. al.*, *Discovery of a narrow resonance in e^+e^- annihilation. Phys. Rev. Lett.*, **33**, 1406 (1974); C. Bacci *et. al.*, *Preliminary result of Frascati (ADONE) on the nature of a new 3.1 GeV Particle Produced in e^+e^- Annihilation. Phys. Rev. Lett.*, **33**, 1408 (1974); G. S. Abrams *et. al.*, *Discovery of a Second Narrow Resonance in e^+e^- Annihilation. Phys. Rev. Lett.*, **33**, 1453 (1974).

[14] S. Weinberg, *Phys. Rev. Lett.* **27**, 1264 (1967).

[15] R. Cahn and G. Goldhaber, *Experimental Foundations of Particle Physics*, Cambridge University Press, London (1989).

[16] P. Aurenche, *The Standard Model of Particle Physics*, hep-ph/9712342, Dec., 1997.

[17] G. Wang, Ph.D. Dissertation, The Florida State University, 1997 (Unpublished).

[18] D. Griffiths, *Introduction to Elementary Particles*, John Wiley & sons (1987), pp. 213-256.

[19] F.E. Close, *An Introduction to quarks and partons*, Academic Press (1979), London; F. Halzen and A.D. Martin, *Quarks and leptons*, John Wiley & sons; C. Quigg, *Gauge Theories of the Strong, weak, and electromagnetic interactions*, Benjamin-Cummings Publishing Company (1983), London.

[20] P.W. Higgs, *Spontaneous Symmetry Breakdown Without Massless Bosons*, *Phys. Rev.* **145**, p. 1156 (1966).

[21] T. Muta, *Foundations of Quantum Chromodynamics*, World Scientific (1987), Singapore; R.K. Ellis, W.J. Stirling, and B.R.Webber, *QCD and Collider Physics*, Cambridge University Press (1996).

[22] G. Ross, *Grand Unified Theories*, Benjamin-Cummings Publishing Co. (1995), London; A.D. Collins and E.J. Squires, *Particle Physics and Cosmology*, John Wiley & sons (1989).

[23] H.P. Nilles, *Phys. Rep.*, **110**, 1 (1984); X. Tata, *An Introduction to Supersymmetry and Supersymmetry Phenomena*, UH-511-713-90 (Unpublished).

[24] J.C. Pati and A. Salam, Phys. Rev. D**10**, 275 (1994); E. Eichten *et. al.*, Phys. Rev. D**34**, 1547 (1986); W. Buchmüller and D. Wyler, Phys. Let, B**177**, 377 (1986); E. Eichten *et. al.*, Phys. Rev. Let. **50**, 811 (1983); H. Georgi and S. Glashow, Phys. Rev. Let. **32**, 438 (1974)

[25] L.F. Abbott and E. Farhi, *Phys. Lett.*, B**101**, 69 (1981), *Nucl. Phys.* B**189**, 547 (1981).

[26] J.C. Pati and A. Salam, *Phys. Rev.* D**10**, 275 (1974).

[27] E. Eichten *et. al.*, *Phys. Rev.* D**34**, 1547 (1986).

[28] W. Buchmüller and D. Wyler, *Phys. Lett.* B**177**, 377 (1986); E. Witten, *Nucl. Phys.* B**258**, 75 (1985).

[29] M. Leurer, Phys. Rev. D**49**, 333 (1994); J.L. Hewett and S. Pakvasa, *Phys. Rev.* D**37**, 3165 (1988).

[30] J.L. Hewett and T.G. Rizzo, *Phys. Rev.* D**36**, 3367 (1987).

[31] S. Dawson, E. Eichten and C. Quigg, *Phys. Rev.* D**31**, 1581 (1985).

[32] P.R. Harrison and C.H. Llewellyn-Smith, *Nucl. Phys.* B**213**, 223 (1983).

[33] J. Blümlein, E. Boos and A. Kryukov, *Z. Phys.* C**76**, 137 (1997).

[34] V. Barger and R.J.N. Phillips, *Collider Physics*, Addison-Wesley (1987).

[35] DØ Collaboration, S. Abachi *et. al.*, Phys. Rev. Lett. 75 3618 (1995); DØ Collaboration, B. Abbott *et. al.*, *Search for Leptoquarks at DØ*, XVIII International Symposium on Lepton Photon Interactions, Hamburg FDR (1997).

[36] CDF Collaboration, F. Abe *et. al.*, Fermilab-pub-98/121-E (submitted to *Phys. Rev. Lett.*) (1998).

[37] J. Thompson, *Introduction to Colliding Beams at Fermilab*, Fermilab-TM-1909 (1994).

[38] DØ Collaboration, S. Abachi *et. al.*, Nucl. Instrum. Methods Phys. Res. A **338**, 185 (1994).

[39] J. Perkins, Ph.D. Dissertation, The University of Texas at Arlington, 1998 (Unpublished).

[40] J. Krane, Ph.D. Dissertation, The University of Nebraska, 1998 (Unpublished).

[41] D. Norman, Ph.D. Dissertation, The University of Maryland, 1995 (Unpublished).

[42] D. Möhl, G. Petrucci, L. Thorndahl and S. Van der Meer, Phys. Rep. C **73** (1980). For a brief overview of stochastic cooling see http://adrfpc01.fnal.gov/DaveMc/misc/uw600/sld062.htm

[43] N. Hadley, *Cone Algorithm for Jet Finding*, DØ Internal Note 904 (1989) (Unpublished).

[44] D. Elvira *et. al.*, *Jet Energy Scale at DØ*, DØ Internal Note 3287 (1997) (Unpublished); Jet Energy Scale Group, *Jet Energy Scale (CAFIX 5.1)*, DØ Note 3139 (1996) (Unpublished).

[45] D. Karmgard, *Search for Second Generation Scalar Leptoquarks in the $\mu\mu jj$ Channel*, DØ Internal Note 3518 (1998) (Unpublished); B. Connolly and D. Karmgard, *Second Generation Vector Leptoquarks at DØ*, DØ Internal Note 3546 (1998) (Unpublished).

[46] DØ Collaboration, B. Abbott *et. al.*, *Search for Second Generation Leptoquark Pairs Decaying to $\mu\mu + jets$ in $p\overline{p}$ Collisions at 1.8 TeV*, submitted to Phys. Rev. Lett..

[47] E. Gallas, *The MTC Package (Muon Tracking in the DØ Calorimeter)*, DØ Internal Note 2066 (1995) (Unpublished).

[48] J.W. Bantly and R.E. Hall, *Top Counting in the Dimuon Channel: DØ Analysis of Run I*, DØ Internal Note 3088 (1997) (Unpublished).

[49] J.W. Bantly *Second Generation Leptoquarks in Run Ib/c at DØ*, DØ Internal Note (1995) (Unpublished).

[50] DØ Collaboration, B. Abbott *et. al.*, *Phys. Rev.* D**57** 3817 (1998)

[51] F. Paige and S. Protopopescu, BNL Report No. 38304, 1986 (unpublished). We used ISAJET v7.22 with CTEQ2L parton distribution functions.

[52] T. Sjöstrand, Comp. Phys.Comm. **82**, 74 (1994), release v5.7.

[53] A. Boehnlein, FERMILAB-CONF-98-337-E.

[54] M. Krämer, T. Plehn, M. Spira, and P.M. Zerwas, *Phys. Rev. Lett.*, **79**, 341 (1997).

[55] J. Blümlein, E. Boos, and A. Kryukov *Z. Phys. C*, **76** 137 (1997).

[56] F.A. Berends *et. al.*, *Nucl. Phys.*, **B357**, 32 (1991).

[57] G. Marchesini *et. al.*, *Comp. Phys. Comm.*, **67**, 465 (1992), release v5.7.

[58] R. Brun and F. Carminati, CERN Program Library Writeup W5013 (1993) (Unpublished). We used GEANT v3.15.

[59] P. Bhat *et. al.*, *DØ Optimized Search for First Generation Leptoquarks in the $e\nu_e jj$ Channel with Run I Data*, DØ Internal Note 3308 (1997) (Unpublished).

[60] DØ Collaboration, P.C. Bhat, in *Proceedings of the* 10^{th} *Topical Workshop on Proton-Antiproton Collider Physics*, edited by R. Raja and J. Yoh, p. 308 (AIP Press, 1995), and references therein. We used JETNET v3.0.

[61] B. Knuteson, *Algorithms for Computing Significance*, DØ Internal Note 3345 (1997) (Unpublished).

[62] I. Bertram *et. al.*, *A Recipe for the Construction of Confidence Limits*, DØ Internal Note 2775 (1995) (Unpublished).

[63] Particle Data Group, R.M. Barnett *et. al.*, *Phys. Rev.* D**54**, 166 (1996).

[64] D. Karmgard, *Search for Second Generation Scalar Leptoquarks in the* $\mu\nu jj$ *Channel*, DØ Internal Note 3545 (1998) (Unpublished); B. Connolly and D. Karmgard, *Second Generation Vector Leptoquarks at DØ*, DØ Internal Note 3546 (1998) (Unpublished).

[65] B. Abbott *et. al.*, *Search for Second Generation Leptoquark Pairs Decaying to* $\mu\nu + jets$ *in* $p\overline{p}$ *Collisions at 1.8 TeV*, submitted to Phys. Rev. Lett..

[66] DØ Collaboration, B. Abbott *et. al.*, *Phys. Rev Lett.* **81** 38 (1998).

[67] E. Boos *et. al.*, *DØ Optimized Search for First Generation Vector Leptoquarks*, DØ Internal Note 3416 (1998) (Unpublished).

[68] W. Weaver, *Lady Luck - The Theory of Probability*, Anchor Books, Garden City, NY (1963).

[69] Particle Data Group, C. Caso *et. al.*, *Particle Physics Booklet*, extracted from the *European Physical Journal* C**3**, 1 (1998), pp.213-231.

[70] G. Landsberg and C. Grosso-Pilcher *Combined Limits on First Generation Leptoquarks from the CDF and DØ Experiments*, DØ Internal Note 3436/CDF Internal Note 4575, (1997) (Unpublished).

[71] C. Peterson, T. Rögnvaldsson and L. Lönnblad *JETNET 3.0 - A Versatile Artificial Neural Network Package*, CERN-TH.7135/94 (1994)

[72] V. Rao and H. Rao *C++ Neural Networks and Fuzzy Logic*, 2^{nd} Edition, MIS:Press, New York, NY, (1995).

BIOGRAPHICAL SKETCH

Daniel J. Karmgard

Date of Birth

June, 18th, 1963

Education

1995-1999 The Florida State University, Ph.D., Physics

1993-1995 California State University, Long Beach, M.S., Physics

1991-1993 University of California, Los Angeles, B.S., Physics, B.S., Math

1987-1991 El Camino College, A.A. General Education, A.S., Physics

Research and Teaching

1996-1999 Research Assistant, Fermi National Accelerator Laboratory,

DØ Experiment, Florida State University Group

1995-1996 Teaching Assistant, Physics, Florida State University

Professional Positions

1985-1995 Applied Technology Assoc., Inc. – Programer/Analyst

1981-1985 United States Air Force

Invited and Plenary Talks

1. *"A parameterization of electro-weak mass matrices for quarks."*, D. Karmgard and S. Rajpoot, Proc Beyond the Standard Model IV, Lake Tahoe, Dec. 1994, California State University at Long Beach Preprint, Jan. 1995

2. *"The Search for First Generation Leptoquarks with the D0 Detector"*, American Physical Society Spring Meeting, Columbus, OH April, 1998

3. *"The Search for First Generation Leptoquarks with the D0 Detector"*, New Perspectives Conference Batavia, IL July, 1998

4. *"Search for Second Generation Leptoquarks with the DØ Detector"*, American Physical Society, Division of Particles and Fields, Los Angeles, CA, January, 1999

5. *"Leptoquark Searches at DØ"*, University of Notre Dame, High Energy Physics Seminar, Notre Dame, IN, February, 1999.

Publications in Refereed Journals

1. *"A parameterization of electro-weak mass matrices for quarks."*, D. Karmgard, S. Rajpoot, and C. Pilot Z. Phys. C **72** 469-478 (1996)

2. *" Search for Scalar Leptoquark Pairs Decaying to Electrons and Jets in $\overline{p}$-p Collisions."*, B. Abbott *et. al.*, Phys. Rev. Letters **79**, 4321 (1997) , Fermilab-Pub-97/252-E, hep-ex/9707033.

3. *" Search for the Trilepton Signature from Associated Gaugino Pair Production."*, B. Abbott *et. al.*, Phys. Rev. Letters **80**, 1591 (1998), Fermilab-Pub-97/153-E, hep-ex/9705015.

4. *"Determination of the Mass of the W boson Using the DØ detector at the Tevatron"*, B. Abbott *et. al.*, Phys. Rev. D **58**, 12002 (1998), Fermilab-Pub-97/328-E, hep-ex/9710007.

5. *"$Z\gamma$ Production in $\overline{p}$-p Collisions at $\sqrt{s}$=1.8TeV and Limits on Anomalous $ZZ\gamma$ and $Z\gamma\gamma$ Couplings."*, B. Abbott *et. al.*, Phys. Rev. D **57**, 3817 (1998), Fermilab-Pub-97/363-E, hep-ex/9710031.

6. *"Search for First Generation Scalar Leptoquark Pairs in $\overline{p}$-p Collisions at $\sqrt{s}$ = 1.8 TeV."*, B. Abbott *et. al.*, Phys. Rev. Letters **80**, 2053 (1998), Fermilab-Pub-97/344-E, hep-ex/9710032.

7. *"A Measurement of the W Boson Mass."*, B. Abbott *et. al.*, Phys. Rev. Letters **80**, 3000 (1998), Fermilab-Pub-97/423-E, hep-ex/9712028.

8. *“A Measurement of the W Boson Mass.”*, B. Abbott *et. al.*, Accepted by Phys. Rev. D, Fermilab-Pub-97/422-E, hep-ex/9712029.

9. *“Direct Measurement of the Top Quark Mass.”*, B. Abbott *et. al.*, Phys. Rev. D **58**, 052001 (1998), Fermilab-Pub-98/031-E, hep-ex/9801025.

10. *“Search for the Decay* $b \to X_s\mu^+\mu^-$*.”*, B. Abbott *et. al.*, Phys. Lett. **B423**, 419, 1998; Fermilab-Pub-98/033-E, hep-ex/9801027.

11. *“Measurement of the Shape of the Transverse Momentum Distribution of W Bosons Produced in* $\overline{p}$*-p Collisions at* $\sqrt{s}$*= 1.8 TeV.”*, B. Abbott *et. al.*, Phys. Rev. Letters **80**, 5498 (1998), Fermilab-Pub-98/075-E, hep-ex/9803003.

12. *“Limits on* $WW\gamma$ *and* WWZ *Couplings from W Boson Pair Production.”*, B. Abbott *et. al.*, Phys. Rev. **D58** Rapid Communications 051101, Fermilab-Pub-98/076-E, hep-ex/9803004.

13. *“Search for Charge 1/3 Third Generation Leptoquarks in* $\overline{p}$*-p Collisions at* $\sqrt{s}$*= 1.8 TeV.”*, B. Abbott *et. al.*, Phys. Rev. Letters **81**, 38 (1998), Fermilab-Pub-98/081-E, hep-ex/9803009.

14. *“Limits on Anomalous* $WW\gamma$ *and* WWZ *Couplings.”*, B. Abbott *et. al.*, Phys. Rev. **D58** Rapid Communications 031102, July 1998, Fermilab-Pub-98/094-E, hep-ex/9803017.

15. *“Search for Heavy Pointlike Dirac Monopoles.”*, B. Abbott *et. al.*, Phys. Rev. Letters **81**, 524 (1998), Fermilab-Pub-98/095-E, hep-ex/9803023.

16. *“The DØ Detector at TeV33.”*, B. Abbott *et. al.*, D0note 3410, Fermilab-Pub-98/124-E, hep-ex/9804011.

17. *“Determination of the Absolute Jet Energy Scale in the DØ Calorimeters.”*, B. Abbott *et. al.*, Submitted to Nucl. Instr. and Meth. A, Fermilab-Pub-97/330-E,hep-ex/9805009.

18. *“The Dijet Mass Spectrum and a Search for Quark Compositeness in* $\overline{p}$*-p Collisions at* $\sqrt{s} = 1.8$ *TeV.”*, B. Abbott *et. al.*, Submitted to Phys. Rev. Letters, Fermilab-Pub-98/220-E, hep-ex/9807014.

19. *“The Inclusive Jet Cross Section in* $\overline{p}$*-p Collisions at* $\sqrt{s} = 1.8$ *TeV.”*, B. Abbott *et. al.*, Submitted to Phys. Rev. Let., Fermilab-Pub-98/207-E,hep-ex/9807018.

20. *“Small Angle J/Psi Production in* $\overline{p}$*-p Collisions at* $\sqrt{s}$*=1.8TeV.”*, B. Abbott *et. al.*, Submitted to Phys. Rev. Letters, Fermilab-Pub-98/237-E, hep-ex/9807029.

21. *"Measurement of the Top Quark Mass in the Dilepton Channel."*, B. Abbott *et. al.*, Submitted to Phys. Rev. D , Fermilab-Pub-98/261-E, hep-ex/9808029.

22. *"Search for Squarks and Gluinos in Single-Photon Events with Jets and Large Missing Transverse Energy in $\overline{p}$-p Collision at $\sqrt{s}$=1.8 TeV."*, B. Abbott *et. al.*, Submitted to Phys. Rev. Letters, Fermilab-Pub-98/243-E, hep-ex/9808010.

23. "*Measurement of the Top Quark Pair Production Cross Section in $\overline{p}$-p Collisions using Multijet Final States.*", B. Abbott *et. al.*, Submitted to Phys. Rev. D , Fermilab-Pub-98/130-E, hep-ex/9808034.

24. *"Probing Hard Color-Singlet Exchange in $\overline{p}$-p Collisions at $\sqrt{s}$=630 GeV and 1800 GeV."*, B. Abbott *et. al.*, Submitted to Phys. Rev. Letters, Fermilab-Pub-98/285-E, hep-ex/9809016.

D0 Experiment Internal Papers (authored or co-authored)

1. *"DØ Optimized Search for First Generation Leptoquarks with Run I Data"*, P. Bhat *et. al.*, DØ Note 3208, May, 1997

2. *"DØ Optimized Search for First Generation Leptoquarks in the $e\nu_e jj$ Channel"*, P. Bhat, *et. al.*, DØ Note 3308, Oct., 1997

3. *" Silicon Track Trigger Performance Studies"*, T. Heuring, *et. al.*, DØ Note 3415, Oct., 1997

4. *"DØ Optimized Search for First Generation Vector Leptoquarks"*, E. Boos, *et. al.*, DØ Note 3416, Mar., 1998

5. *"Search for Second Generation Scalar Leptoquarks in the $\mu\mu jj$ Channel"*, Dan Karmgard, DØ Note 3518, Sep., 1998

6. *"Search for Second Generation Scalar Leptoquarks in the $\mu\nu_\mu$ jj Channel"*, Dan Karmgard, DØ Note 3545, Oct., 1998

7. *"Second Generation Vector Leptoquarks at DØ"*, Brian Connolly and Dan Karmgard, DØ Note 3546, Oct., 1998

ERRATUM

After publication of this dissertation it was found that one of the selection criteria listed was not applied in the way in which it was reported. The cut in line 6 of table 5.2, is listed as $\Delta R_{Cal-jets} > 0.5$. In fact, the cut that was applied was not $\Delta R = \sqrt{\Delta\phi^2_{Cal-jets} + \Delta\eta^2_{Cal-jets}} > 0.5$ but $\sqrt{(\phi_{Cal} - \phi_{jets})^2 + (\eta_{Cal} - \phi_{jets})^2} > 0.5$.

While this is not a physically meaningful selection, it was applied consistently to both data and Monte Carlo. Therefore the numbers reported for backgrounds, data, and signal efficiency are valid.

The analysis was changed, however, for the internal notes and the journal publication to remove the physically meaningless selection. With the MTC track selection as it was, the MTC fake rate is 10.8%. If this error is corrected the MTC fake rate is reduced to approximately 4%. If this cut is removed entirely then the MTC fake rate becomes 11.3%. This is what was done for publication. With the new fake rate , the normalization for the $W + jets$ cross section is calculated to be 18%. These changes introduce small changes into the analysis. For example, in the $\mu\mu + jets$ search there are 62 events after the initial selection instead of 61, while in the $\mu\nu + jets$ search there are 107 rather than 109. The backgrounds become 63±16 ($\mu\mu + jets$) and 106±30 ($\mu\nu + jets$) after the initial selections.

The signal efficiencies change only slightly. Final efficiencies are within 0.1% of the values listed in chapters 5 and 6. The limits on the cross sections do not change appreciably, and the mass limits remain the same.

Thus, while the details of the search have changed slightly, the changes are minor and the results remain consistent.