A MODEL INDEPENDENT SEARCH FOR NEW PHYSICS IN FINAL STATES CONTAINING LEPTONS AT THE DØ EXPERIMENT

By

Joel M. Piper

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Physics and Astronomy

2010

ABSTRACT

A MODEL INDEPENDENT SEARCH FOR NEW PHYSICS IN FINAL STATES CONTAINING LEPTONS AT THE DØ EXPERIMENT

By

Joel M. Piper

The standard model is known to be the low energy limit of a more general theory. Several consequences of the standard model point to a strong probability of new physics becoming experimentally visible in high energy collisions of a few TeV, resulting in high momentum objects. The specific signatures of these collisions are topics of much debate. Rather than choosing a specific signature, this analysis broadly searches the data, preferring breadth over sensitivity.

In searching for new physics, several different approaches are used. These include the comparison of data with standard model background expectation in overall number of events, comparisons of distributions of many kinematic variables, and finally comparisons on the tails of distributions that sum the momenta of the objects in an event.

With 1.07 fb^{-1} at the DØ experiment, we find no evidence of physics beyond the standard model. Several discrepancies from the standard model were found, but none of these provide a compelling case for new physics.

For Kirsten

Acknowledgements

First, I would like to thank my thesis adviser, Jim Linnemann. As I was struggling to find a research position upon my return from France, Jim was willing to take a chance on me. He provided me specific guidance when I needed it and at the same time gave me the independence and responsibility I needed to mature as a scientist. Throughout all of the ups and downs of this analysis, I never had any question that Jim was behind me, supporting me every step of the way.

I would like to thank Jim Kraus, who took as long as I needed to answer all of my daily questions about physics and research methods, while at the same time providing a steady stream of obscure historical facts and cultural minutiae.

I would like to thank Chip Brock, who introduced me to the world of high energy physics. It was intimidating at first, but I got the hang of it.

I would like to thank the other current and former members of my guidance committee: Sekhar Chivukula, C.P. Yuan, Gary Westfall, and Carlo Piermarocchi. I never failed to leave a committee meeting without a clear idea of how much more I had to learn.

I would like to thank all current and former members of the L2 team. Linnemann, Bob, Kraus, Miroslav, Abaz, Roger, and Reinhard all helped change my response to a 2 A.M. page from sheer panic into mild annoyance. Thanks also to Mandy, Shannon, Ernest, Mikolaj, Ike, James, Weigang, Emmanuel, and Enrique.

I would like to thank all of the current and former members of the MIS group: Linnemann, Kraus, Nayeem, Peter, Prolay, Anatoly, Oleksiy, and especially Serban, who taught me how to work through each step of a high energy physics analysis. I would also like to thank the members of the editorial board, the NP conveners, physics coordinators and DØ spokespeople, who always took an active interest in the analysis, particularly Todd, Arnaud, Horst, Aurelio, and Jean-Francois. Also, thanks to Bruce for building the foundations for this search with his Run I work and for his

early technical advice in this analysis.

I would like to thank everyone at MSU, especially Brenda and Debbie for all of their hard work in making the group and the department function. Everything would fall apart without you. Also, thank you Reinhard and Reiner, for your help in my transition to DØ.

I would like to thank Betta, Jorge, and Monica for making the hour-long commute from Chicago bearable, sometimes even fun. Also, to my roommates, Gorge and Mango, thanks for dragging me out of the house every now and then.

I would like to thank my in-laws for inviting me into their family and for their generosity while Kirsten and I figure out our next steps.

I would like to thank the rest of my family, especially Mom, Dad, and Mary. When immediately upon receiving my undergraduate degree in environmental engineering, I had the crazy idea of getting a Ph.D. in particle physics, they didn't blink. Knowing that I always have their love and support has given me the courage to live life on my own terms, making decisions based on hope rather than fear.

Finally, I would like to thank Kirsten. To get this degree, we spent years in Lansing and years in Chicago with months of separation. My research often required long hours and intense stress, but I was able to enjoy even the most difficult times because I knew that every night when I came home, the stress of the day would melt away, and I'd be happy again. I really couldn't have done this without you. Thank you for going through this with me. Thank you for your love and support. Thank you for making my life so much happier. I love you.

Table of Contents

List of Tables

List of Figures

Images in this dissertation are presented in color.

Chapter 1

Introduction

As the Fermilab Tevatron Collider heads into its final years, the two major high energy experiments there, DØ and CDF, will finish collecting data. The energy frontier in high energy physics is now beginning its transition to the Large Hadron Collider outside of Geneva, Switzerland and the major high energy physics experiments there, ATLAS and CMS. As these experiments begin, it is fair to ask one central question from the Tevatron. Are the results of the DØ and CDF experiments consistent with the standard model? The standard model is the single theoretical framework that has successfully predicted all new fundamental particles discovered after its inception, with the exception of the Higgs boson, a particle still within the reach of the Tevatron.

While many analyses check the precision of the standard model using measurements of known particles and others check for well-motivated extensions, the focus of this analysis is to determine if the standard model is well-described as broadly as possible given the constraints of our detector and limitations in modeling the standard model. Most searches for new physics at the Tevatron have focused on a specific model, often molding the search to be sensitive to a particular signal. However, the proposed extensions to the standard model include so many possible signals, there are too many areas of phase space to conduct dedicated searches. We find that the major constraints in these searches is the sensitivity of the detector to a particular

area of phase space and the ability of our Monte Carlo and detector simulation to properly model it.

While this analysis strives for model independence, several assumptions about physics beyond the standard model are necessary to provide some sensitivity. We assume that the new physics will manifest itself by containing objects with reasonably high transverse momentum. And, we check final states containing leptons because this is where our detector has been most heavily tested and where we believe we can have enough sensitivity to detect deviations.

Our focus is on three methods to test agreement between data and the standard model background. We divide our data and background into final states based on object content, then check event counts in data against our expectation. We plot many different event distributions to see if there are any large disagreements. Then, we focus specifically on one distribution, the sum of object momenta and the missing transverse momentum in each event, and search for large data excesses in the tails of these distributions.

Many other quantities could also be checked for disagreement, but we believe these three tests should provide us with a good sense of whether there is new physics for which our experiment and current background simulation could be sensitive.

The focus of this dissertation is to describe such a search, a process that leads from opening a container of hydrogen gas to probing the edges of scientific understanding. The description of this process is divided into four basic parts. The first part contains three chapters which describe background material that is not specific to this particular analysis. This includes an explanation of the standard model (the group of theories that we are testing) with a brief overview of the current landscape of potential extensions. Then, the chain of events that turns hydrogen gas into 1.96 TeV proton-antiproton collisions is briefly discussed, followed by a description of the procedure DØ uses to identify the remnants of collisions and turns into a comprehen-

sive understanding of the underlying physics processes. The last part of this section discusses the selection and storage of those collisions (events) which are considered the most useful for scientific understanding. The second part involves a description of the selected events specific to this analysis, the simulation of events representing the standard model expectation needed as background for this analysis, and the necessary additional corrections needed to modify the simulation to account for known defects and oversimplifications. The third part will discuss the specific analysis strategy and the details of the procedure in comparing the selected data with the expectations of the standard model. Finally, the last two chapters present the results of the comparison along with their interpretation.

Chapter 2

Symmetry, the Standard Model and Beyond

Physics is a science of symmetries. In classical, relativistic, and quantum theory, symmetries provide profound insights into the laws of nature. Many of the symmetries used in classical physics are intuitive, but deeper symmetries in quantum field theory helped to produce the current theory of all interactions observed at the quantum scale, known as the standard model. While the gauge symmetries of quantum field theory are fundamentally different than the space-time symmetries of classical physics, they share a basic commonality. The space-time symmetries involve the invariance of physical laws to translations, rotations, etc. The internal phase symmetries of the standard model are invariant to "rotations" within the space of the interactions. The $SU(3)_C$ invariance of the strong force, for instance, simply means that the strong force is invariant with respect to the color of the quarks or gluons that it is acting on. If one "rotates" the quarks in this nonphysical space, the strong force will act in exactly the same way. While these symmetries are no longer in a physical space and simply represent a redundancy in the theory, they still obey the dictates of Noether's theorem, discussed in Section 2.1.1, and possess a conserved current. The understanding of the redundancies that describe the spaces of the three interactions,

led to the creation of the standard model, whose interactions are determined by the gauge group, $\mathrm{SU}(3)_C \otimes \mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$. It has been extensively tested and has often shown remarkable agreement with nature.

Despite its success, the standard model has a limited reach. This leads to the idea that new physics lies beyond the currently accepted theory. Searches for physics beyond the standard model, such as the one described in this thesis, are now common at the major collider experiments. These searches look for extensions to the standard model that encapsulate it within a more general theory which can maintain its validity to higher energy scales. Currently, there are theoretical and experimental hints that we are currently at the energetic limit of the standard model and we will be able to measure properties of new physics at the LHC and possibly even at the Tevatron. The form of this new physics is not a settled topic. Theorists have proposed many different strategies to extend the standard model. These theories are not only consistent with all of the current data, but they are also typically analogues of current processes and build on their observed properties. The most well-accepted extensions tend not to predict one specific signature in a particle detector. They are more general with many different signatures and provide little reason for one signature to be expected over another. (The choice of which possibility to search for is often made based on the acceptance and sensitivity of the detector.)

The approach in this dissertation is to look across many final states to check for any disagreements with the standard model, rather than focus on any of these extensions specifically. This is done in three stages. First, events with high momenta are selected and some small corrections are applied. Then, the data are searched broadly for large discrepancies. Finally, the tails of the transverse momentum distributions are compared against expectation in specific final states.

The symmetry discussion in this chapter follows discussions in [30] and [31]. The rest borrows heavily from the books [1], [32], and a series of lectures given at the

Hadron Collider Summer School of 2008 [33].

2.1 Symmetry

Using the Lagrangian form of particle motion, basic conservation laws can be derived by finding symmetries in nature. This is true in Newtonian and relativistic classical physics in properties such as conservation of energy and momentum. In quantum field theory, besides the spatial quantities, additional phase symmetries suggest new conserved quantities that lead to the basic interactions of particle physics.

The formulation of the standard model can be found by modifying basic free particle Lagrangians to be invariant under certain phase transformations in a space that is not physically observable. The presence of massive particles, however, prevents these symmetries from being complete explanations of the model. To account for particle masses, a form of symmetry breaking is introduced to preserve the symmetric structure while matching the observed experimental results.

2.1.1 Symmetry and Conservation Laws

The content of an introductory physics course can be summarized from a handful of observations about nature. These are typically given in the form of Newton's three laws, but particle motion can also be described by the principle of least action, where a particle chooses a path such that the action integral S is minimized in $S = \int_{t_1}^{t_2} \mathcal{L} dt$. The function $\mathcal{L}$ is called the Lagrangian density and leads to the equations of motion given in Equation 2.1.

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) - \frac{\partial \mathcal{L}}{\partial q_i} = 0 \ (i = 1, 2, ...s) \tag{2.1}$$

The list of q_i's are the position coordinates and the $\dot{q}_i$'s are the velocities associated with the q_i's. There is one equation for each degree of freedom in the system. From

these basic equations, assumptions about symmetries of space and time can lead to conserved quantities. For example, assuming the homogeneity of time leads to the conservation of energy. The homogeneity of space leads to the conservation of momentum, and the isotropy of space leads to the conservation of angular momentum.

When particles are considered relativistically, the same arguments can be used to account for the possible symmetries in the four dimensional space-time of special relativity (Minkowski space): translations, time displacements, rotations and Lorentz transformations.

In quantum theory, we abandon specific kinematic predictions for a probabilistic wave function using statistics to determine particle properties. The nonrelativistic Schrödinger wave equation is shown in Equation 2.2.

$$i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r},t) = -\frac{\hbar^2}{2m}\nabla^2\Psi(\mathbf{r},t) + V(\mathbf{r})\Psi(\mathbf{r},t) \tag{2.2}$$

where $\hbar$ is the reduced Planck constant, Ψ is the wavefunction, a probability amplitude for a particle to have a position $\mathbf{r}$ at a time t, m is the mass of the particle, and V is the time-independent potential energy of the particle at $\mathbf{r}$.

The simple deterministic equations of motion have been replaced by the probabilistic interpretation of quantum theory. The Schrödinger equation describes how a particle moves when in the presence of a force described by the potential V.

The Schrödinger equation does not, however, satisfy the requirements of special relativity. To incorporate special relativity, the relativistic Dirac equation is used, Equation 2.3. This version of the Dirac equation is not the simplest notationally, but it shows the direct predictions of antimatter partners (ϕ_+ , ϕ_-), and the relation to spin in the Pauli matrices, σ. The $\phi_{+,-}$ are two component spinors representing the wave functions of a particle and its antiparticle. The Dirac equation does not include the potential in the Schrödinger equation. Interactions in quantum field theory are determined by the interaction of separate field equations rather than the simplification

of a separate potential.

$$\begin{pmatrix} mc^2 & c\sigma \cdot p \\ c\sigma \cdot p & -mc^2 \end{pmatrix} \begin{pmatrix} \phi_+ \\ \phi_- \end{pmatrix} = i\hbar \frac{\partial}{\partial t} \begin{pmatrix} \phi_+ \\ \phi_- \end{pmatrix} \tag{2.3}$$

The Euler-Langrange equations are still effective field theory descriptions of probabilistic particle motion, and in this context take the form of Equation 2.4.

$$\frac{\partial \mathcal{L}}{\partial \phi} - \frac{\partial}{\partial x^\mu} \left[\frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \phi \right)} \right] = 0 \tag{2.4}$$

Noether's theorem relates a symmetry in a physical law to a conserved quantity, called a charge [34]. The theorem is typically cast in the language of classical electromagnetism, but it can be used to classify any of the symmetries that will be discussed in this chapter. In any system, the quantity that is conserved must flow continuously across the system. In the case of momentum, this means that the momentum of one particle must be transferred to that of another, or to some other part of the system. A particle that loses momentum in one part of the system, while a completely unrelated particle gains momentum in another part of the system would satisfy an overall conservation of the momentum "charge", but it could not be described as a continuous momentum flow and would fail the requirements of Noether's theorem. The flow of this quantity is termed generally, as a current, J, and the quantity itself is generally called a charge Q. Noether's theorem is shown in Equation 2.5.

$$\frac{d}{dt} \int J^0_\nu d^3x = \frac{Q_\nu}{dt} = 0 \tag{2.5}$$

This equation can describe the energy, momentum, angular momentum, as well as all of the quantities that will be introduced later in the chapter.

2.1.2 Symmetry and Particle Interactions

Any measurable property of the free particle depends on $\left|\overline{\psi}\psi\right|$[1] rather than on the wave function (ψ) itself. Therefore, there is a freedom in picking an absolute phase of the wave. While the phase choice has no effect on predictions about the properties of the wave itself, it determines the interference effects when the wave interacts with another field. In order for the Dirac equation to properly describe the interaction, the phase of the second wave must be chosen to be consistent with the first. This is true even if the waves are not causally connected. In order to preserve causality, it was suggested that the phase choice for an individual wave should be independent of the choice of other waves; it should be locally invariant. This principle was formalized with non-Abelian gauge theories by Yang and Mills [35]. In order to maintain invariance, the Dirac Lagrangian (using a simplified notation from Equation 2.3) would need to be modified as in, Equation 2.6.

$$\begin{aligned} \mathcal{L}_{tot} &= \overline{\Psi}\left(i\hbar c\gamma^{\mu}\partial_{\mu} - mc^2\right)\Psi \rightarrow \\ \mathcal{L}_{tot} &= \overline{\Psi}\left(i\hbar c\gamma^{\mu}\partial_{\mu} + ieA_{\mu} - mc^2\right)\Psi - \frac{1}{4\mu_0}F^{\mu\nu}F_{\mu\nu}. \end{aligned} \tag{2.6}$$

When this equation is compared to the previous Dirac equation, the partial derivative ∂_{μ} is replaced by what is known as the covariant derivative, $D_{\mu}\Psi = \left(\partial_{\mu} + ieA_{\mu}\right)\Psi$. The covariant derivative was introduced explicitly to maintain the local gauge invariance. However, it is found that the additional term describes the interaction between the particle and an electromagnetic field. The $F_{\mu\nu}F^{\mu\nu}$ term represents the kinetic energy of the electromagnetic field itself.

[1] $\overline{\psi}$ is the adjoint spinor $\psi^{\dagger}\gamma^0$ and $\gamma^0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$.

2.1.3 Symmetry and the Standard Model

The interactions necessary to preserve the phase symmetry described above can be put into the language of group theory. The symmetry arising from the electromagnetic interaction satisfies the conditions of a U(1) group. If this process is repeated for $SU(2)_L$, the weak interactions can be obtained, and for $SU(3)_{\text{color}}$, the strong interactions. While the strong interactions obey an exact symmetry, it is found that the electromagnetic and weak interactions together form an $SU(2)_L \otimes U(1)_Y$ symmetry. It is a linear combination of the gauge bosons predicted in this theory that are responsible for interactions involving weak isospin and electromagnetic charge. This symmetry is broken, and at low energies becomes the familiar U(1) symmetry of electromagnetism.

2.1.4 Symmetry, Broken

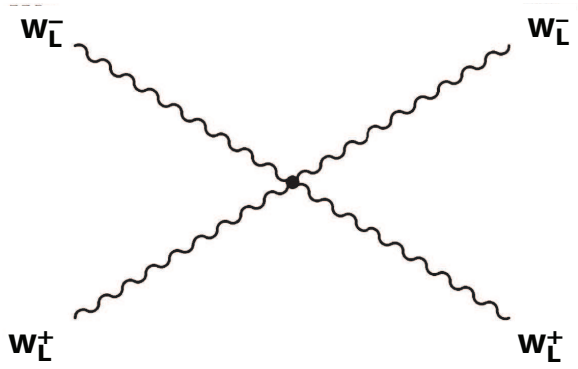
The $SU(2)_L \otimes U(1)_Y$ symmetry of the electroweak interactions is only satisfied if the gauge bosons are nearly massless (loop corrections would provide a mass of the Z boson of 35 MeV). Additionally, fermion mass terms would also break this gauge invariance and must also be missing from a gauge-invariant theory. Since the interactions satisfy the electroweak symmetry, the symmetry must be broken to provide particle masses [36]. The mechanism of mass generation through Goldstone bosons that provide mass to the W and Z bosons, the Higgs mechanism, has been experimentally verified. The specific dynamics of the Higgs mechanism, however, are still unknown [4].

Theories predicting the dynamics of the Higgs mechanism fall into two general categories. First, is the addition of a weakly-interacting self-coupled elementary scalar. The other option is to add additional strong-interaction dynamics among new fermions. The simplest form of the Higgs mechanism in the electroweak sector is the addition a single scalar doublet, which is the form that is currently incorporated

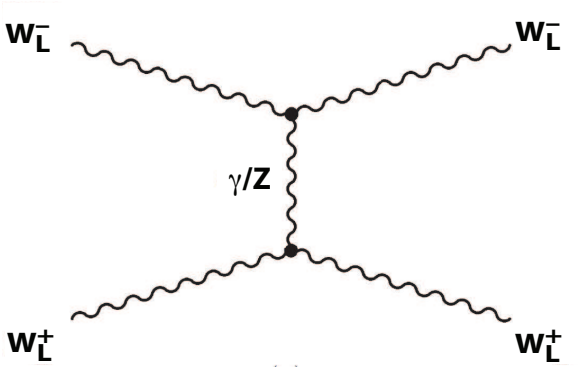
into the standard model. One of the experimental consequences of this form is a single observable scalar particle, the Higgs boson. Even if one abandons the concept of $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ symmetry and simply adds the mass terms directly, the theory would diverge for some interactions involving the weak gauge bosons, such as the longitudinally-polarized W boson scattering process as seen in Figure 2.1. This divergence violates the fundamental principle of unitarity in quantum theory, predicting total probabilities greater than one. This violation is shown to be universal and to all orders, with a critical energy of ~ 1.2 TeV.

The standard model Higgs mechanism uses the fact that the electroweak interaction symmetry can be preserved if the vacuum is defined at a nonzero value incorporating what is called a vacuum expectation value. This lowest nonzero energy of the vacuum is an energy density that permeates all space. The $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ symmetry is valid for the overall system until a specific vacuum ground state is chosen, then the symmetry is spontaneously broken. The ground state of the system could have a minimum among a continuum of possible values. Nature must choose one of these possibilities and once the choice is made, the symmetry of the system is gone. Additional diagrams due to the Higgs boson, the observable scalar particle from the vacuum, cancels divergent terms in gauge boson scattering. The diagrams in Figure 2.2 are complementary to those in Figure 2.1.

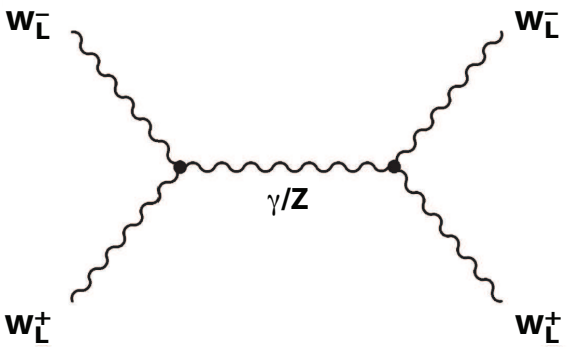
This idea is often visualized through looking at simpler analogous symmetries. For a pseudoscalar wave equation with charge symmetry (symmetric in $\phi' \rightarrow -\phi$), the minimum energy ground state has two symmetric possibilities. Nature would only be able to choose one of them, breaking the symmetry of the overall equation. This can be extended to a complex scalar field where the vacuum ground state energy can now choose among points on a circular minimum as seen in Figure 2.3, known as the Mexican hat potential. The standard model symmetry is slightly more complicated, but the idea is a generalization of the previous examples. The field must choose a



(a)

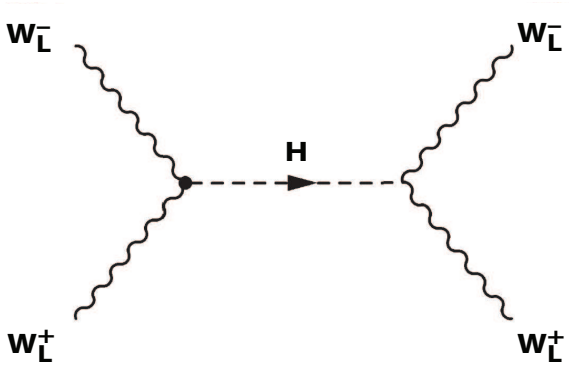


(b)

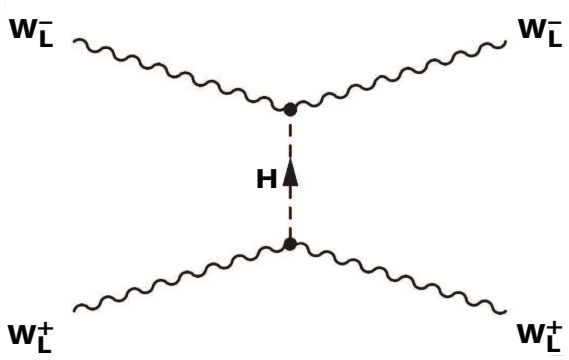


(c)

Figure 2.1: Longitudinal W boson scattering. This is one of the processes that would have a divergent cross section without additional diagrams. In this thesis, the time axes in Feynman diagrams run left to right.



(a)



(b)

Figure 2.2: Additional diagrams involving the Higgs boson that cancel divergences of longitudinal W boson scattering in the standard model.

direction in SU(2) space breaking the symmetry. This creates a Higgs doublet in SU(2) and a singlet in U(1). Hints of the direction are shown in the experimental observation of electromagnetic charge conservation.

Even if the Higgs mechanism does not satisfy the simple Higgs doublet assumption of the standard model, if new physics are at considerably higher energies, the lightest Higgs boson introduced in theories that use weak-coupling will mimic the properties of the single Higgs boson of the standard model. The lack of evidence for physics beyond the standard model from electroweak precision data (see Section 2.2.2) hint that it may be unlikely to find low-mass new physics, and a Higgs boson associated with high-mass new physics would show many of the same properties as the standard model Higgs.

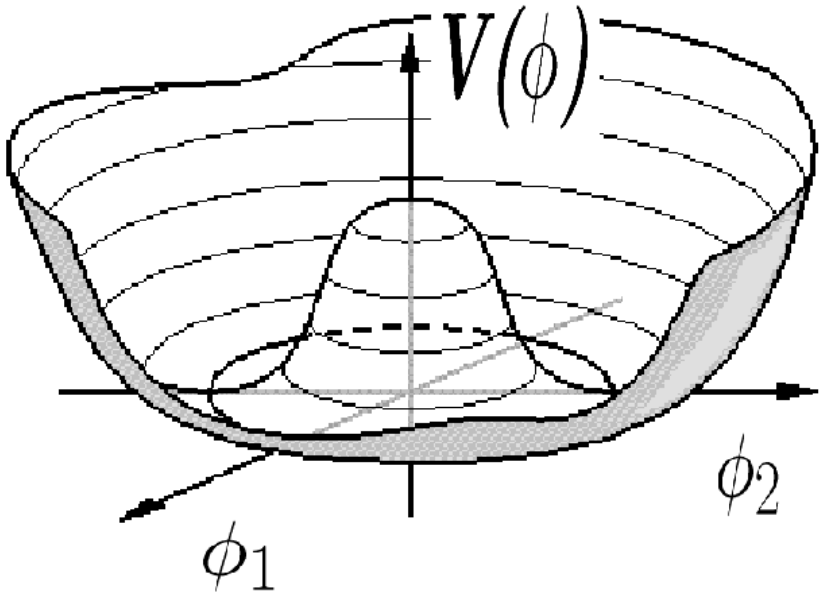


Figure 2.3: The Mexican hat potential shown as an analogue of the spontaneous symmetry breaking of the standard model. The overall potential is completely symmetric, but nature must choose a specific minimum within the potential for the vacuum expectation value [4].

2.2 The Standard Model

While the ideas above are useful in understanding how nature preserves basic symmetries, the standard model is typically described by the experimentally observed particles and interactions. The couplings among the particles of the standard model are determined from the gauge symmetries. Still, twenty-six parameters are determined from experiment. The similarity among these hint at possible greater underlying symmetries of a more general theory. The standard model would then be just an effective low energy theory of this more general model.

2.2.1 Particles and Interactions

We observe particle interactions as forces that change a particle's measurable properties. Three types of interactions are measurable at the quantum scale: the electromagnetic, weak, and strong forces. These forces are shown in Table 2.1, with an approximation of their strength and range. Gravitation is included for completeness. The particles observed in experiment and their properties are listed in Table 2.2.

Table 2.1: The four fundamental forces with their approximate interaction distances and strengths [1].

Force	Relative Strength[a]	Range (m)	Carriers
electromagnetic	10^{36}	infinite	photon
weak	10^{25}	10^{-18}	$W^{\pm}$, Z
strong	10^{38}	10^{-15}	gluons
gravitation	1	infinite	graviton[b]

[a] The relative strengths are approximate and vary depending on the particles involved.
[b] The graviton has not been observed.

The intrinsic property of spin is used to differentiate two classes of particles. Particles with integer spin are known as bosons, while those with half-integer spin are called fermions. The known fundamental matter particles are spin-1/2 fermions while the particles that mediate the interactions between the matter particles are spin-1 bosons, called gauge bosons since they arise from the phase invariance of the interaction Lagrangians.

The $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ electroweak theory predicts two charged and two neutral gauge bosons. From experiment, it is seen that only one of them is involved with the known electromagnetic force seen at low energies. This is the massless photon. The other types of particle exchange allowed in the electroweak theory are suppressed because the gauge bosons are massive. Once the interaction energies approach the

Table 2.2: The fundamental particles with their force-related quantum numbers and masses [2].

Particle	EM Charge	Spin	Colored?	Number of Particles	Mass
e	-1	1/2	No	4	0.511 MeV
${\nu_e}^a$	0	1/2	No	2?	< 2 eV^b
μ	-1	1/2	No	4	106 MeV
${\nu_\mu}^a$	0	1/2	No	2?	< 0.19 MeV
τ	-1	1/2	No	4	1.78 GeV
${\nu_\tau}^a$	0	1/2	No	2?	< 18.2 MeV
u	2/3	1/2	Yes	12	$1.5 - 3.3$ MeV
d	$-1/3$	1/2	Yes	12	$3.5 - 6.0$ MeV
c	2/3	1/2	Yes	12	1.27 GeV
s	$-1/3$	1/2	Yes	12	105 MeV
t	2/3	1/2	Yes	12	173.1 GeV
b	$-1/3$	1/2	Yes	12	4.20 GeV
$W^\pm$	± 1	1	No	2	80.4 GeV
Z	0	1	No	1	91.2 GeV
γ	0	1	No	1	0
gluon	0	1	Yes	8	0
graviton^c	0	2	No	1	0

[a] The neutrino mass eigenstates are heavily mixed from the flavor eigenstates. This means the masses quoted here will be mixtures of the various neutrino flavors, much more so than the quarks.

[b] This assumes CPT invariance. The limit comes from the antineutrino. There are much weaker limits on the neutrino.

[c] The graviton has not been observed.

masses of the gauge particles, these additional exchanges are observed. The charged W bosons and neutral Z boson are the weak components observed from the electroweak theory. The weak force is only visible to left-handed particles. This particle property is related to the parity of the particle. These left-handed particles contain a "charge" known as weak isospin which is mediated in much the same way as the electromagnetic charge.

The interactions arising from the SU(3) symmetry are known as strong interactions. They are mediated by a massless gluon. The exchange of color "charge" has a couple complications that differentiate it from the electromagnetic charge. The color charge comes in three types, and the charge carriers are colored objects as well. It is found that colored objects cannot be directly observed in experiment. Each colored object is drawn to create neutral "white" colored objects by proper combinations of the individual colors. The strong force increases with distance, so as colored particles move away from each other, eventually the energy will produce other colored objects to create overall neutral measurable objects. As these colored objects are split apart and form other objects, they produce a stream of colorless objects known as jets. These are bound states of two quarks (one color with its anticolor) or three quarks (one of each of color or one of each anticolor). The two quark states are known as mesons while the three quark states are known as baryons. The proton and neutron are examples of baryons (proton- $[uud]$, neutron- $[udd]$). The residual strong forces from the proton and neutron are what hold together atomic nuclei, similar to how residual electromagnetic forces in atoms hold together molecules.

Due to the additional interactions discussed in the above paragraphs, the table of particles is actually incomplete. The expansion of the up quark which is subject to all of the fundamental forces is shown in Table 2.3.

Table 2.3: Fundamental particles, revised. This is an example of the full particle content of the standard model. Each particle listed has a corresponding antiparticle with opposite electromagnetic charge. Furthermore, each particle has left-handed and right-handed members to determine if they interact through the weak force. Additionally, quarks come in three types of colors. The gluons carry color combinations and total eight separate particles.

Particle	EM Charge	Spin	Color	Weak Isospin
u_L^{red}	$+2/3$	1/2	red	Yes
$\overline{u}_L^{\text{red}}$	$-2/3$	1/2	antired	Yes
u_R^{red}	$+2/3$	1/2	red	No
$\overline{u}_R^{\text{red}}$	$-2/3$	1/2	antired	No
u_L^{blue}	$+2/3$	1/2	blue	Yes
$\overline{u}_L^{\text{blue}}$	$-2/3$	1/2	antiblue	Yes
u_R^{blue}	$+2/3$	1/2	blue	No
$\overline{u}_R^{\text{blue}}$	$-2/3$	1/2	antiblue	No
u_L^{green}	$+2/3$	1/2	green	Yes
$\overline{u}_L^{\text{green}}$	$-2/3$	1/2	antigreen	Yes
u_R^{green}	$+2/3$	1/2	green	No
$\overline{u}_R^{\text{green}}$	$-2/3$	1/2	antigreen	No

2.2.2 Experimental Confirmation

The theoretical picture of the standard model has been accepted after rigorous testing by many types of experiments over many different channels. The most profound verification of the picture was the successful prediction of the W and Z bosons, and the relation of their masses. Additionally, the standard model predicted the existence of the gluon, charm and top quarks before their eventual discovery.

The most current measurements of the electroweak sector of the standard model have been compiled and evaluated by the LEP Electroweak Working Group [5]. The

plot in Figure 2.4 shows the overall consistency of each of the complementary measurements compared to their best fit. The overall agreement is amazingly consistent. The table includes masses of the W and Z bosons and the top quark, the heavy boson widths (Γ), the hadronic cross section, the weak mixing angle, the hadronic contribution to the running QED coupling constant at the Z-pole, and various asymmetry measurments, A, (such as in charge or polarization) and decay width ratios, R, $\left(\text{such as in } \frac{\Gamma(b\overline{b})}{\Gamma(\text{hadrons})}\right)$.

2.2.3 Difficulties with the Standard Model

The standard model is not a fundamental theory. Despite its success and its compelling derivation from symmetries, much of the information in the standard model comes from experimental measurements. There are twenty-six parameters whose values must be added to the standard model by hand, and it would be preferred to have a theory where these were theoretically determined. The most obvious shortcoming of the theory is its failure to incorporate gravity. While the current energy scales probed are not sensitive to this interaction, it obviously must be incorporated into a full theory. Surprisingly, the standard model has held up incredibly well in experiments. Its imminent failure keeps being delayed as it has shown itself to be able to make precision predictions beyond the point where it might be expected to fail. Most of this discussion was adapted from a lecture by Guido Altarelli at the 2008 Hadron Collider Summer School [37].

2.2.3.1 Theoretical Difficulties

The most fundamental shortcoming of the standard model is the lack of an explanation of gravitation. Although the gravitational interaction is so weak that its effects are not measurable at current experiments, it can be calculated when the gravitational force would contribute noticeably to measurements. It is found that the center

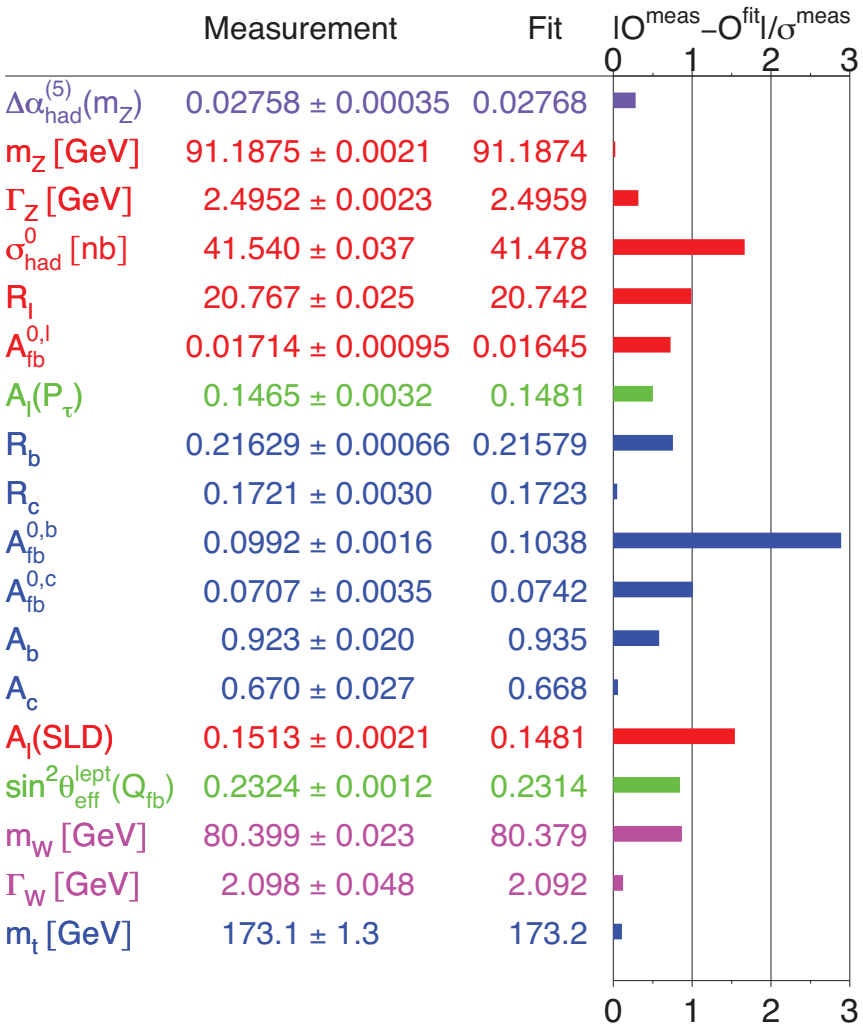


Figure 2.4: Many experiments measure values that are interconnected by the standard model. This figure shows how much each measurement pulls on the overall fit. Most of the measurements show excellent agreement. The value of $\Delta\alpha^{(5)}_{\rm had}(m_Z)$ is taken from low energy experiments; the next five are LEP I line shape and lepton asymmetries; $A_\ell(P_\tau)$ is from LEP I tau polarization; the next six are from LEP I and SLD heavy-flavor measurements; $\sin^2\theta^{\rm lept}_{\rm eff}(Q_{\rm fb})$ is from LEP I $q\overline{q}$ asymmetry; the two W boson measurements are from both the Tevatron and LEP II, and the top mass is only from the Tevatron [5].

of mass energy of an accelerator needed to to probe the gravitational sector would need to be on the order of 10^{19} GeV, which is known as the Planck mass (M_{Planck}). This scale is impossible to probe with any foreseeable technology, and even the highest energy cosmic rays are 10^{14} GeV. Information about this sector can only be gleaned indirectly from observing the large-scale structure of the universe.

Another difficulty is known as the hierarchy problem. This refers to the difference between the scales where the weak and gravitational interactions become important. When calculating the loop corrections to the Higgs mass, there are quadratic divergences that must cancel at the level of new physics (the Planck mass in the standard model). While this is not explicitly forbidden, it seems unnatural.

Additionally, flavor physics is not well-described. It is found that there are three representations of the fundamental particles with different masses. Each of the mass eigenstates of these flavors are combinations of conserved interaction eigenstates. The amount of mixing is now fairly well measured but not well-described. The standard model contains no compelling explanation for the observed three generations of particles that seem identical in their interactions but vastly differing in mass.

2.2.3.2 Experimental Difficulties

As has been mentioned, the experimental confirmation from particle colliders has been exacting. The precision measurements in the electroweak sector have put constraints on many new physics models. There are a handful of experiments that show some disagreement, such as the muon anomalous magnetic moment, $(g-2)_\mu$, the forward-backward assymetry for bottom quark production, $A_{fb}^{0,b}$, and the frequency of a B_s^0 meson spontaneously oscillating to its antiparticle $\overline{B}_s^0$, B_s mixing. However, these disagreements do not point to a consistent compelling argument for new physics, and it remains to be seen if these indirect inconsistencies are due to difficulties in prediction, experiment or actually are the effects of new physics. Most of the

unsettling experimental results are not based on a particular experiment but through the interpretation of the data in general, some of which are described below. These factors simply point to behavior that deviates from what would be expected from our current understanding of nature or observed phenomena that the standard model cannot describe.

First, the fundamental forces vary their interaction strength as a function of energy. This can be thought of as a property of the vacuum. As energies get higher and higher, they can probe closer and closer to the particle, penetrating fluctuations from the vacuum which can screen the charge of the object. The electromagnetic charge becomes weaker as e^-e^+ pairs are created out of the vacuum and the dipole moment of the pair screens the overall charge seen from the electron. The opposite happens in the strong force where the charge screening creates not only quark-antiquark pairs but also gluons. The gluons are aligned such that the overall color charge seen increases with distance. In the standard model, the weak, strong, and electromagnetic interactions begin to approach the same strength with an increase in energy. It is unusual, however, that while the forces approach similar coupling strengths, they do not seem to converge.

The recent discovery that neutrinos are massive also points to a theory beyond the standard model. The neutrinos are found to be so much lighter than the other particles of the standard model that the mass hierarchy is difficult to explain. One explanation for this phenomenon is that neutrinos are Majorana particles (particles that are their own antiparticles) and get their masses through interactions that do not conserve lepton number. These actions, however, seem to be suppressed by the GUT scale (the grand unification scale where the electromagnetic, weak, and strong forces are merged into a unified field theory).

Astronomical observations have shown that most of the matter in the universe is a material that has not yet been seen in experiments (or anywhere else) on earth. This

matter is detected through its gravitational interactions with visible matter. The standard model provides no particles that can account for the amount of dark matter seen in the universe.

Additionally, the vacuum expectation value would create an energy density of the universe which is ~ 49 orders of magnitude above what is actually measured. The constant of the vacuum energy is actually arbitrary, so this factor is not completely inconsistent. However, again, as with the cancellation of terms in the Higgs mass, the value seems unnatural.

2.3 Beyond

Knowing that the standard model can not be a final theory, it is reasonable to predict what a more fundamental theory may be that still satisfies all of the observed experimental data. These theories look to explain the dynamics of the Higgs mechanism that lead to electroweak symmetry breaking [38].

Three methods of doing this will be briefly described. The first is to reduce the unnaturalness of the loop corrections to the Higgs mass, as in supersymmetry. The second is to eliminate the hierarchy problem by introducing extra dimensions in which gravity propagates, making the scales only appear different in our 4-D world. The final class of theories that will be explored are the technicolor models. These introduce a new force which follows the pattern seen in the development of QCD, where the interaction was originally thought to be mediated by pion exchange.

2.3.1 Supersymmetry

The hierarchy problem is a fundamental difficulty highlighting the enormous difference between the scale of electroweak symmetry breaking and the GUT or Planck scale. If the standard model is correct to gravitational energies, then all terms of this scale

can enter into loop corrections to the Higgs mass. A natural explanation to the hierarchy problem is that there are fermionic partners for the standard model bosons and bosonic partners for the standard model fermions. These partners allow the cancellation of divergent loop corrections and would allow a natural Higgs mass up to the GUT scale [39].

The incorporation of these effects could also give a light Higgs boson within current electroweak constraints, the unification of gauge coupling strengths at the GUT scale, and a possible cold dark matter candidate. Other attempts to solve the naturalness problem in the Higgs sector can yield difficulty in maintaining consistency in the Yukawa couplings, but with SUSY, all of the observed masses are consistent.

It has so far not been possible to consistently explain experimental data with only particles coming from the minimal supersymmetric standard model (MSSM). The addition of a hidden sector that does not interact with the standard model particles or the expansion of the theory to extra dimensions provide an additional variable in considering how supersymmetry might be naturally realized. The phenomenology of the underlying theory depends upon how the visible MSSM sector communicates with the hidden sector (or across the bulk between the branes) [2].

2.3.2 Extra Dimensions

Extra dimensions in various guises have the ability to bring down the fundamental scale of particle physics from M_{Planck} to M_{EW} (the scale of electroweak symmetry breaking) by allowing gravity to propagate in additional space-time dimensions. This would allow the strength of gravity to be on the order of the other forces in the universe and only appear weak in the 4-D brane in which we do experiments [2].

In flat extra dimensions, only the graviton propagates outside of the observable 4-D brane. The size of the additional extra dimensions can be determined by moving the Planck mass to a scale that is adequate for electroweak symmetry breaking. The

graviton propagating in a compactified dimension will have a tower of possible energy states. For a large number of extra dimensions, the number of these states (Kaluza-Klein Modes) that would be observable is small and the experimental signature would be difficult to see. If the number of dimensions is small $\lesssim 8$, then these states would have small mass splitting and would act like a massive non-interacting particle. This process can be searched for directly in monojet and monophoton states. Additionally the large number of states can be checked indirectly by looking at differential cross sections in dilepton production [40].

Warped extra dimensions act similarly. Gravitons originate in a separate Planck-brane in 5-D space, and the strength of gravity is suppressed by a warp factor which reduces its overall strength when it reaches the 4-D standard model brane. With the extra dimensions now not needing to be small, the Kaluza-Klein modes can have greater spatial differences. The lowest mode would only couple with the strength of gravity and be unobservable. The first excited state could be produced at the TeV scale and could be relevant to collider searches, coupling to diphotons and dileptons [41].

2.3.3 Technicolor

There seems to be a rather simple analogue to the difficulties seen in the electroweak sector at 1 TeV, which are the difficulties with the pion description of QCD at energies ~ 1 GeV. This would signal a new force that is not easily seen in particle interactions at detectors [2, 38].

The W bosons would be the analogues of the charged pions, and the Z boson would be the analogue of the neutral pion. The trouble with this framework is that there is no simple way to incorporate observed fermion masses. Extended technicolor couples the fermion masses to technifermions at a scale much higher than TeV. This is broken to simple Technicolor at energies at the TeV scale. This would explain the

masses of flavors but does not give a reason why no flavor-changing neutral currents are observed. A variation of this called "Walking Technicolor" allows some terms to be enhanced because of techniparticle interactions. This means that the couplings in the theory must run, but not in an analogous way to QCD. The fundamental particles of technicolor cannot become asymptotically free at high energies. This explanation satisfies everything naturally except the top quark masses.

One final variant yields "Top-Assisted Technicolor" which has the top quark interacting with a new strong interaction. This would make the standard model top quark part technifermion. This additional interaction would also predict massive gluons, known as top gluons that have been the subject of collider searches.

2.3.4 Experimental Signatures

Each of these theories can yield different types of experimental signatures. In an R-parity conserving supersymmetry[2], for example, large amounts of missing transverse energy would be expected. However, most of the signatures are very model-dependent within their overarching framework. One important feature of all of these models is that in order to bring naturalness to electroweak symmetry breaking, the scale of these new phenomena must be around the TeV scale. This leads to new particles that would tend to decay to standard model particles with very high momenta. This is the singular common feature, which makes it difficult to make a specific prediction. Often many assumptions have to be made to reduce the parameter space of the theory being presented.

[2]In supersymmetry, baryon and lepton numbers are no longer conserved. Since the conservation of these quantities has been tested very precisely, R-parity is introduced to supress these violating processes. It is defined as $R = (-1)^{2j+3B+L}$, where j is the spin, B is the baryon number, and L is the lepton number. In supersymmetric extensions of the standard model, standard model particles have R-parity of 1, and supersymmetric particles have R-parity of -1.

2.4 Analysis Strategy

With the overabundance of experimental signatures, this thesis describes an attempt to search for physics beyond the standard model, attempting to minimize the assumptions about the nature of that physics. Generally, the extensions of the standard model include particles decaying to high energy particles related to the electroweak symmetry breaking mechanism. The standard model background modeling simulations at DØ that are currently most developed are for final states containing leptons, which will be the subject of this dissertation. The strategy involves three basic steps: (1) the selection of high-p_T events and addition of correction factors, (2) the comparison of overall event counts and histogram shapes, and (3) a check of the high-p_T tails of distributions. These steps are accomplished in (1) the DØ MIS (model independent search) analysis packages, and the experiment-independent (2) Vista and (3) Sleuth algorithms.

2.4.1 DØ MIS Analysis Packages

The DØ MIS (model independent search) analysis packages are responsible for object selection and implementation of the necessary correction factors. Events that contain isolated high-p_T leptons are selected. Final states are then defined based on the objects occurring in these final states. The simulation used for the prediction does not properly account for events arising from multijet processes. These events are modeled by reversing certain object selection cuts used to define the electron, muon, and tau. The selected events are separated into seven nonoverlapping final states that are dominated by a particular standard model process. These seven were chosen based on the possible dominant lepton or lepton pairs in an event (e, μ, ee, $\mu\mu$, eτ, $\mu\tau$, μe). The states are inclusive in jets and use a series of cuts to establish a single dominant process. For these states, a fit is performed to find individual scale factors

that are unaccounted for in the simulation. These fits are based on three fundamental observables (p_T, η, ϕ). After the scale factors are determined, additional check plots are used to make sure that the fits using simple variables match the more complicated parameters of the events.

2.4.2 VISTA

The events and scale factors from the MIS analysis packages are then passed to the experiment-independent VISTA program. The VISTA program attempts to see if the selected data can be accommodated by the standard model background developed for this analysis. This algorithm focuses on significant discrepancies in exclusive final states and agreement in 1-D histogram shapes. In VISTA, the final states are defined by the full object content in the event. A final state with one jet would be placed in a different final state than one with two jets. Overall consistency among many histograms and final states assures us that we can pass this information to the more narrowly focused SLEUTH algorithm. VISTA could find new physics if there were a general and broad discrepancy with the standard model that could not be explained by detector or simulation problems. More information on VISTA can be found in Section 8.1.

2.4.3 SLEUTH

SLEUTH combines various final states to improve sensitivity and calculates the event p_T sum. This is the sum of the transverse momenta of all of the objects in the event, $\sum_{obj} \left| p_{T,obj} \right|$ and the missing transverse energy, $\left| \not\!\!E_T \right|$. The missing transverse energy ($\not\!\!E_T$) is the negative of the vector sum of the observed energy of the objects in the event. In a hadron collider, the transverse component of the energy should be conserved, so invisible particles carrying energy can be partially reconstructed by calculating the transverse energy imbalance. Each event in these final states will have

exactly one value of this quantity. These values are put into increasing order and compared to the Monte Carlo prediction. For each event in data, the number of data events with $\sum p_T$ equal to or larger than the $\sum p_T$ for that event are counted and compared to the number of weighted Monte Carlo events found in the same region. If there is a data excess, then the probability for the Monte Carlo to fluctuate up to or beyond a value as large as that seen in the data is calculated from the Poisson distribution. This is done for each event. The region of the $\sum p_T$ distribution found to have the largest data/background discrepancy is chosen. Then, to quantify the probability of seeing a discrepancy as large as what is seen in data from statistical fluctuations in the background, the experiment is repeated, by creating Poisson fluctuations in each of the bins of the background distribution. The difference between the event $\sum p_T$ in this pseudoexperiment is compared to the actual background, and the region of maximum discrepancy is found again. This procedure is repeated to determine how many pseudoexperiments would need to be run to see a fluctuation as large as what is seen in data. This process is repeated for all of the final states. If the probability of any point fluctuating up to what is seen in the data is less than 0.001, then the state is marked for further study. Additional information on SLEUTH can be found in Section 8.2.

Chapter 3

The DØ Experiment at the Fermilab Tevatron Collider

The Fermilab Tevatron Collider was the highest energy collider in the world until the Large Hadron Collider produced collisions at energies above those of the Tevatron for the first time on December 8th, 2009. During the time that data was collected for this analysis, the Tevatron was the highest energy collider. The Tevatron collides protons with antiprotons at two locations on a one kilometer radius ring. Two high energy physics experiments, CDF and DØ, sit at the two interaction points as can be seen in Figure 3.1. The first section will describe the acceleration process which leads to the collisions seen at DØ. The second section discusses the DØ detector. The detector consists of material and electronics used to measure characteristics of particles emanating from collisions. Each particle has its own unique signature and the detector may measure the path, charge, energy, momentum, and/or vertex of the particle to try to determine the kinematics of the event corresponding to the collision. All of the information from the collisions must be filtered to reduce the rate of incoming events and the overall data size to a manageable level. This data must be collected and stored, and other qualities of the detector environment must be measured to make the data collected meaningful.

There have been two major data taking periods at the Tevatron. The first ran from 1992 to 1996 and is referred to as Run I. The second began in March 2001 and is ongoing. This period is known as Run II. Furthermore, additional upgrades were performed in 2006. This splits Run II into the period before the upgrades, Run IIa, and the period after, Run IIb [13].

3.1 The Fermilab Tevatron Collider

There are three major stages of the accelerator that lead to the $p\overline{p}$ collisions at the Tevatron. The first is the creation and acceleration of H^- ions. Second, the electrons are stripped off and the remaining proton is accelerated for either eventual injection into the Tevatron for collisions or toward the $\overline{p}$ target for antiproton production. The third major process is the creation, debunching, and storage of the antiprotons [20].

3.1.1 Creation and Acceleration of H^-

The early stages of acceleration are completed by H^- ions. Using an ion of opposite charge from the final product allows easier accumulation of protons in the Booster. The ions are created in the preaccelerator source and accelerated in the Cockcroft-Walton preaccelerator. They then undergo further acceleration in the Linac through the low energy drift-tube Linac (DTL) and the side-coupled Linac (SCL). The Linac is the last stage of the H^- acceleration where it then enters the Booster and is stripped of its electrons [8].

3.1.1.1 Preaccelerator

The preaccelerator begins with a 30 ft^3 bottle of H_2 that contains enough hydrogen for around six months of Tevatron operation [42]. This source is released into a magnetron in an electrically-charged dome [7]. A magnetron uses a magnetic field to

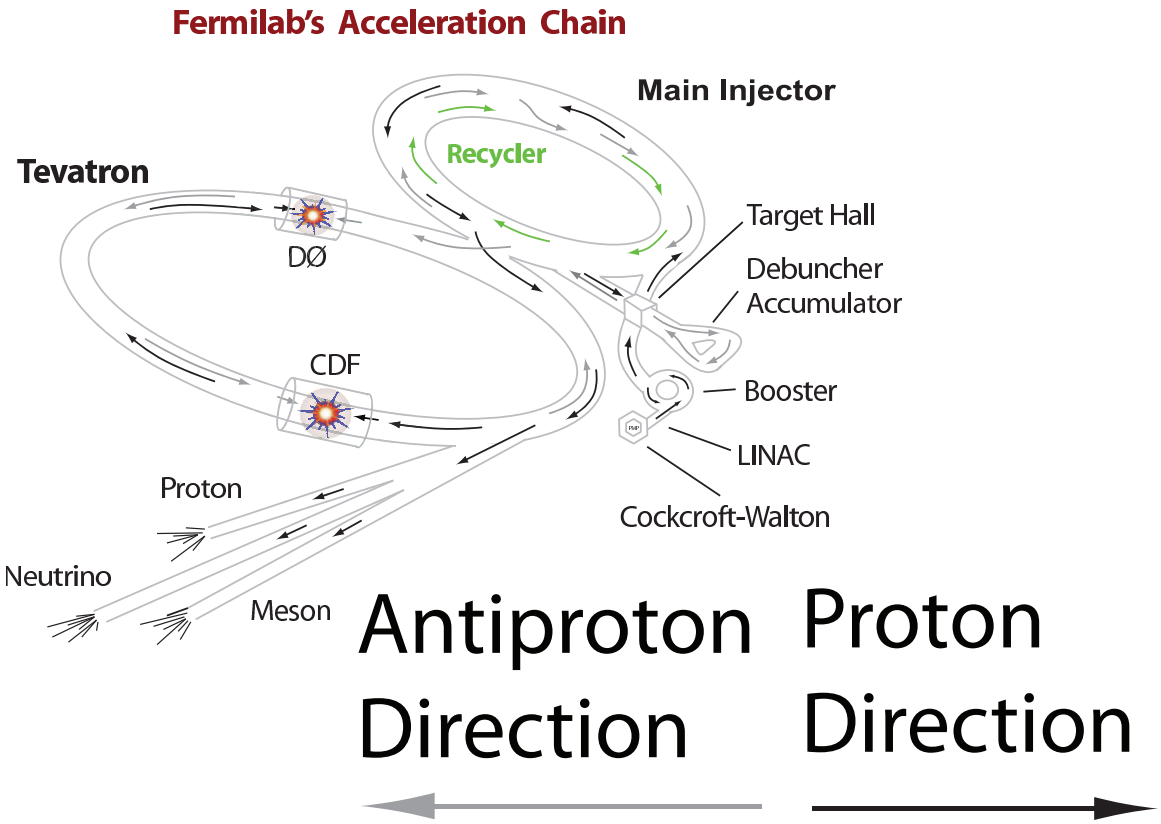


Figure 3.1: The Fermilab accelerator chain [6].

cause the light electrons to spiral around a cathode which is enclosed within an anode. The heavier charged particles will be pulled into the cathode or anode, while neutral particles will be hit with a barrage of electrons. A diagram of the magnetron is shown in Figure 3.2. Some protons will pick up two electrons from the dense plasma, pulling the newly formed ions toward the anode. The main mechanism is from sputtering off hydrogen atoms from the surface of the cathode. The addition of cesium vapor raises the probability that the hydrogen atom will pull off the necessary electrons to form the H^- ions. Some of these ions will be pulled through an aperture in the anode. Once through the aperture, there is a magnetic right-angle bend which selects H^- ions while the electrons and other particles of different charge/mass ratio that also happen to pass through the aperture are sent into a dump and lost. The gas goes from the dome containing the magnetron to a grounded wall where it reaches a final

energy of 750 keV and enters a transfer line to the Linac [8]. The static field of 750 kV is created by a Cockcroft-Walton accelerator using a smaller 75 kV source which is then multiplied several times by a system of capacitors. The total field strength is limited by the size of the area where the acceleration is to take place and the electrical breakdown point. This is the only point in the accelerator chain where static fields are used for acceleration. At higher energies, static fields are too difficult to maintain to be of practical use.

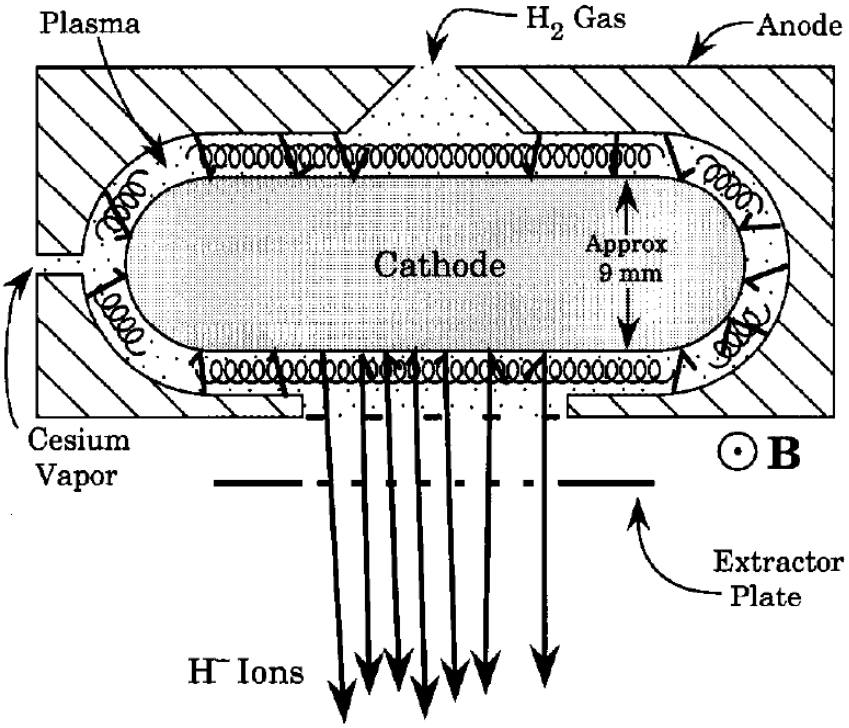


Figure 3.2: The Magnetron: Creating H^- [7].

3.1.1.2 Linac

The Linac takes the 750 keV H^- ions and accelerates them to 400 MeV over 79 m. This is done in two sections, a low energy drift-tube Linac, and a higher energy side-coupled Linac. The drift-tube Linac contains five radio frequency stations, while the side-coupled Linac uses Klystrons for acceleration. The drift-tube Linac (DTL)

uses a single varying B-field to produce a fluctuating E-field. The ions are exposed to the field when it pushes the ions forward and are shielded from it when it pushes in the opposite direction. As the ions gain energy, the length of shielded pipe must increase to compensate for the fact that the ions cover a larger distance over the same period of time. A diagram of the DTL is shown in Figure 3.3. The increasing particle energy from this process is shown in Figure 3.4.

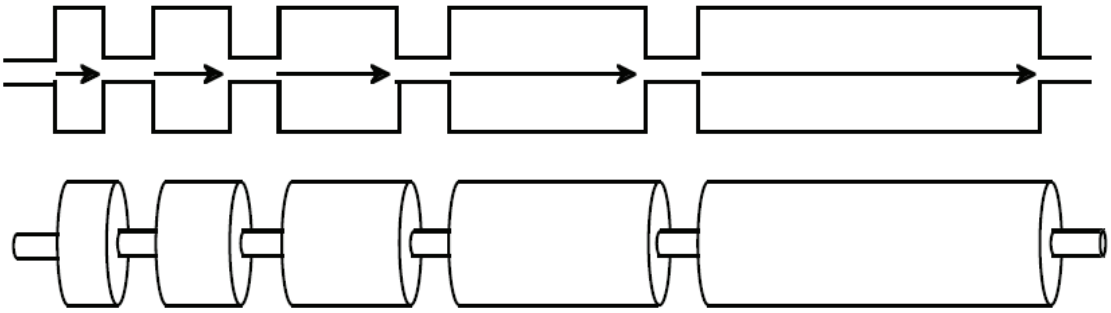


Figure 3.3: The Drift Tube Linac [8]

The side-coupled Linac (SCL) uses Klystrons to produce the electric fields used for acceleration. The Klystron produces a flow of electrons which are bunched in cavities, and then accelerated [43]. These electron bunches excite microwaves in an output cavity that flow into a waveguide. These waves are used to produce the electric field that is seen by the ions in the SCL. The electrons used to generate these waves are then absorbed. While each chamber in the DTL uses the same fluctuating magnetic field to produce the electric field seen by the ions, the SCL cavities are separated using different generated field strengths. Upon leaving the linear accelerator, the ions are next sent to the Booster.

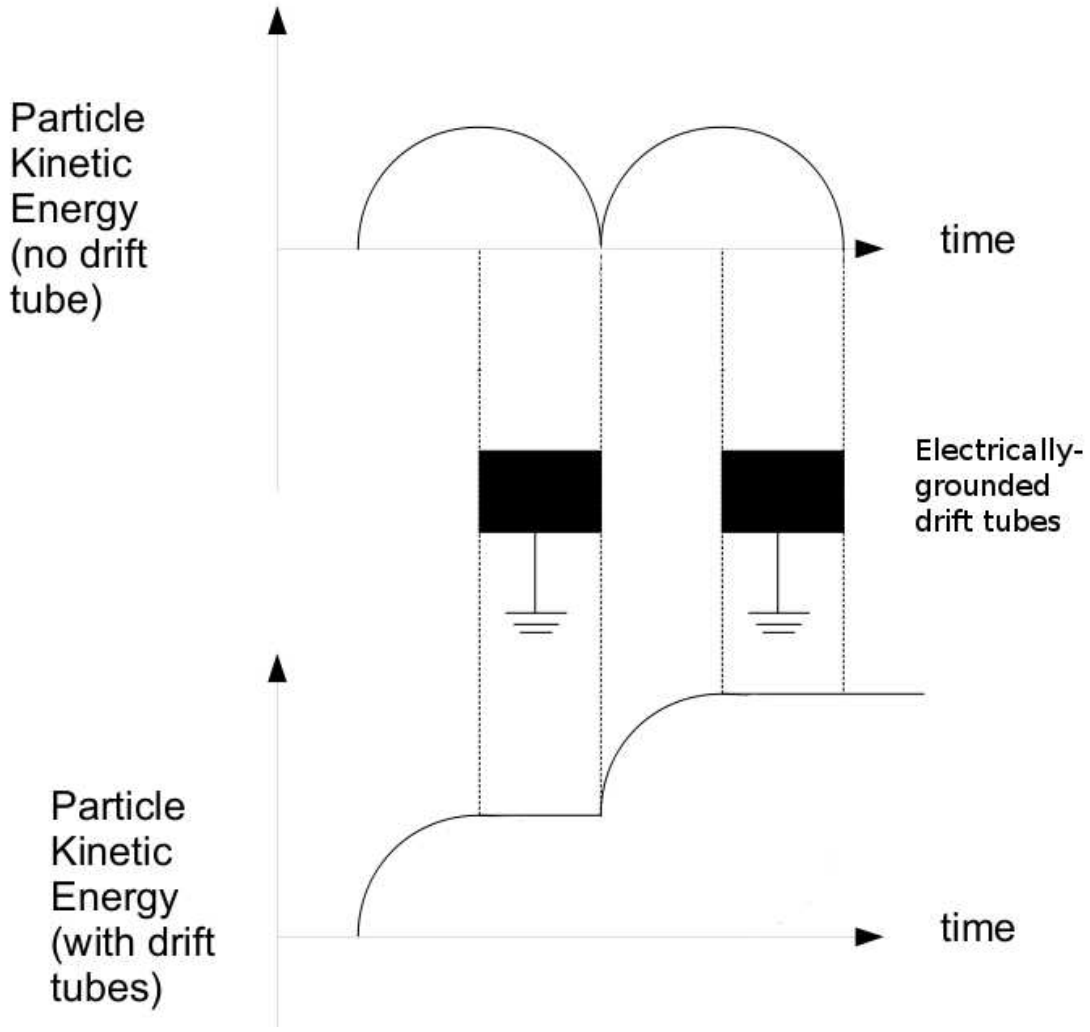


Figure 3.4: The energy of particles as they travel through the DTL [8].

3.1.2 Creation and Acceleration of the Proton Beam

Once the ion reaches the Booster, it has reached an energy of 400 MeV. At the Booster it is stripped of its electrons, and the remaining proton is accelerated to 8 GeV. After this, the proton is sent to the Main Injector where it can be stored for injection into the Tevatron, sent out a beam line to fixed target experiments, or diverted to a target for the production of antiprotons.

3.1.2.1 Booster

The Booster is a synchrotron that takes 400 MeV H^- ions, strips off both electrons producing protons, and accelerates the protons to 8 GeV. The Booster is 75 m in radius and accelerates the protons with 17 RF cavities before sending them to the Main Injector [44]. When filled, the Booster contains $3 \cdot 10^{12}$ protons. Particles must be aligned in a way such that they experience an accelerating E-field at the same time. At any time, each of the 17 RF locations could be used to accelerate particles. The possible particle acceleration paths are known as buckets. If the bucket contains particles, it is known as a bunch.

In the circular synchrotron machines, each particle arrives at an individual RF cavity many times, each time with increasing energy. In order to ensure that an accelerating field is found inside the cavity, the radio frequency of the field needs to be modified.

Non-ideal particles will not be accelerated as expected, and each particle that is slightly ahead of the ideal particle in phase will get less of an increase in E-field, and each particle behind the ideal will get a larger increase. This causes the non-ideal particles to oscillate around the ideal particle trajectory in what are called synchrotron oscillations. Similarly, the restorative forces of quadrupoles used to focus the beam will redirect wayward particles toward the ideal path, but it is necessary to continually correct them to keep them in the beam. This type of oscillation due to the focusing elements of the detector are called betatron oscillations [8].

In the Booster, particle energies reach a point where the stable synchrotron oscillations discussed above are no longer valid. As particle momenta are increased, the velocities of the particles approach the speed of light, and there is little difference in speed across the particle bunch. This means that higher momenta particles will still receive the increase in energy leading to a larger radius to traverse in the Booster. With the same velocity, this requires a longer time than the synchronous

particle to complete a cycle. This means that the higher energy particles that were arriving early, begin arriving late as speed approaches the speed of light and relativistic considerations dominate. At this point, the former restorative forces become destabilizing, and the fields are modified to anticipate higher energy particles arriving late and lower energy particles arriving early. This transition occurs ~ 4.2 GeV, and is passed through quickly to minimize instabilities.

The H^- ions enter the booster where a magnetic field draws them toward an already spinning proton beam. As the two beams are brought together, the H^- beam hits a carbon foil where electrons are stripped producing additional protons. Next, the entire beam is subjected to the same magnetic field producing a dogleg for the protons and putting them back into the normal path of the Booster. The field will cause the remaining H^- ions to be cast into a beam dump while neutral hydrogen atoms continue along the original path and are subject to the same fate [7]. Using an H^- ion beam, allows charge-exchange with the neutral carbon foil. Since the charge-exchange is nonconservative, the conservation of phase space necessary to satisfy Liouville's theorem is not a necessary condition, and the new protons created from H^- ions can be fully merged with the existing proton beam. This was the primary motivation for using the H^- beam rather than immediately creating and accelerating protons. The process of electron stripping and merging of the beams can be seen in Figure 3.5. Once particles pass through the Booster synchrotron loop 24 times, they reach an energy of 8 GeV and are sent to the Main Injector.

3.1.2.2 Main Injector

The main injector performs several functions [9]. Some of the 8 GeV incoming protons are accelerated to 150 GeV and injected into the Tevatron for collisions. Others are ramped up to 120 GeV and sent toward the antiproton target. Antiprotons also enter the main injector and are similarly ramped to 150 GeV for insertion into the Tevatron.

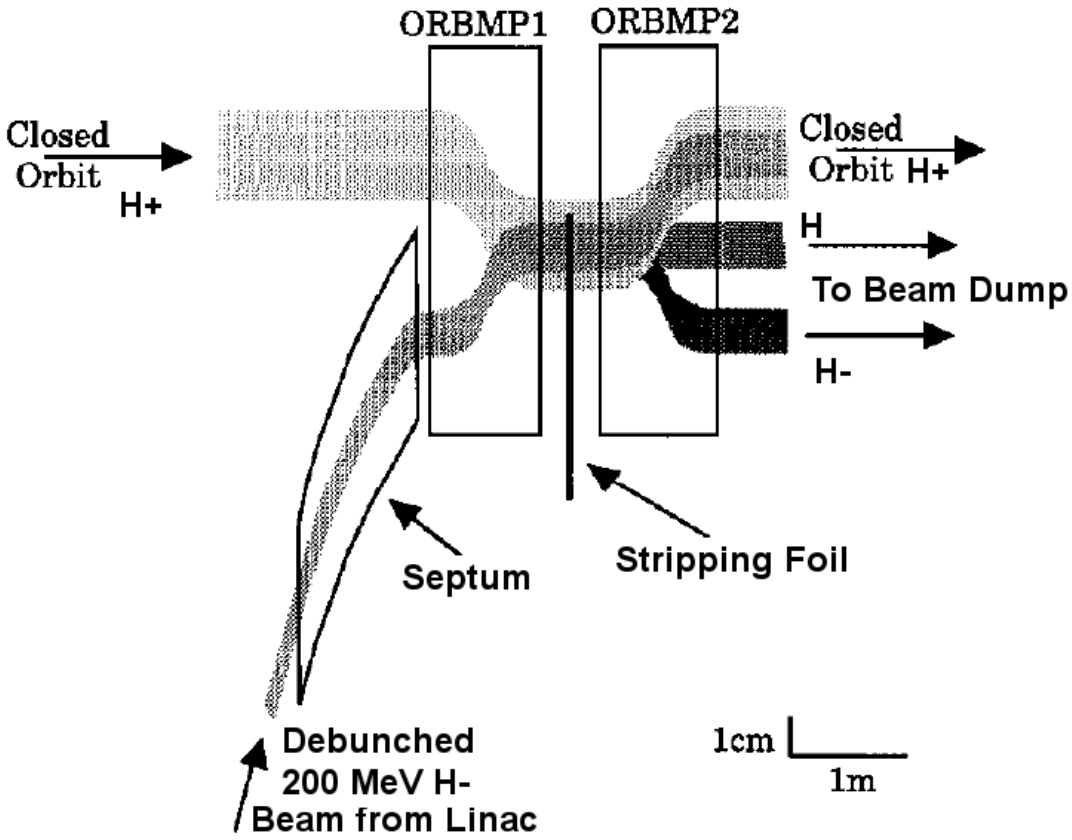


Figure 3.5: The H^- ion is stripped of its electrons and merged with an existing beam of protons using a carbon foil and dogleg magnet [7].

Both the protons and antiprotons are coalesced at flattop before moving to the Tevatron. Flattop is the accelerator condition where the current in the accelerator is maintained as constant and the accelerating voltage is dropped to nearly zero. This releases some of the restorative bunching forces in the cavity, so bunches are able to drift. Special RF cavities then make several bunches (7 for protons, 4 for antiprotons) coalesce into a single bunch. The RF voltage is then turned back up with a newly coalesced bunch structure. For proper insertion, the Main Injector and the Tevatron are set to have the same RF frequencies and phase. To line up a particular MI bunch with a Tevatron bucket, the RF is changed slightly in the MI until the desired MI bunch is aligned with the target Tevatron bunch. The RF frequency is then restored

and the transfer from the Main Injector to the Tevatron is made. This process is known as transfer cogging and is described in the Accelerator Concepts Rookie Book, [8], as follows: "Imagine two large gears meshed together. The Main Injector gear has 588 teeth (RF buckets), and the Tevatron gear has 1113. Once these gears are synched up with each other, they are locked into position relative to each other as well, and particle transfers can occur between them. We want to send protons in a given MI bucket into any Tevatron bucket. The solution is to change the RF frequency in the MI slightly, making the two machines out of phase with each other for a time. While the two gears are out of phase with each other, they will rotate at different speeds, causing different sets of teeth to come near to one another. If the MI frequency were changed back to its original value at the appropriate time, any MI bucket could line up with any Tevatron bucket." The loading of the Tevatron from the Main Injector can be seen in Figure 3.6.

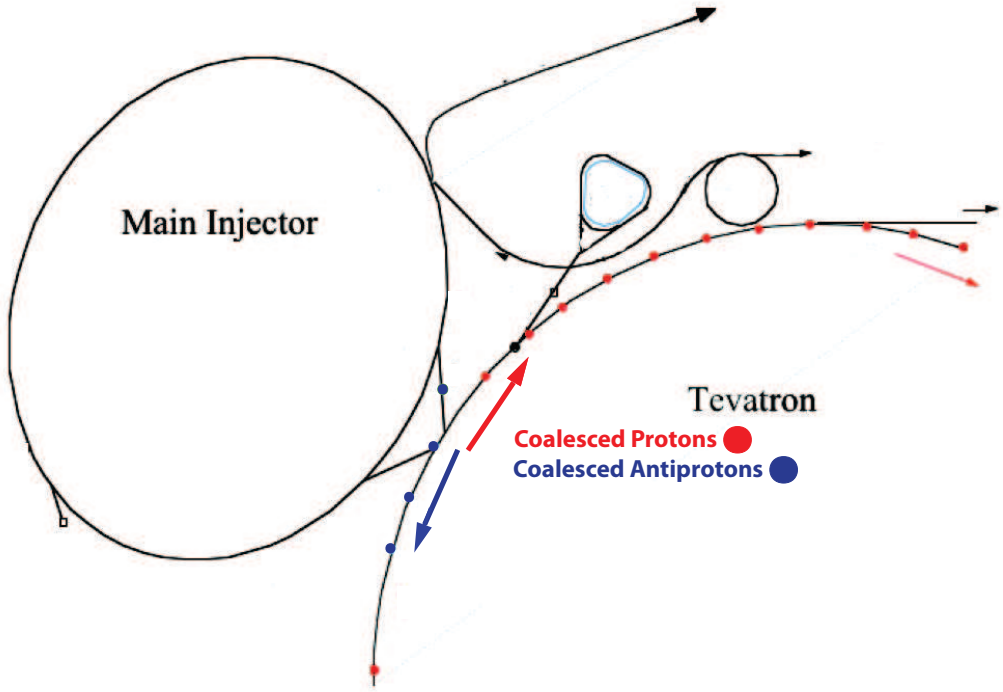


Figure 3.6: Loading the Tevatron from the coalesced protons and antiprotons in the Main Injector [9].

Some protons are sent down a beam line that leads to fixed targets and analysis by other experiments at Fermilab. The rest are used in the production of antiprotons. These 120 GeV protons are sent from the Main Injector, through a beam line near the Tevatron and toward the Inconel (a nickel-ion alloy) antiproton target.

3.1.3 Antiproton Production and Storage

The Fermilab Tevatron Collider is a $p\overline{p}$ collider. The $\overline{p}$'s are created at the Tevatron using accelerated protons. The 120 GeV protons from the Main Injector are directed into a target, the antiprotons from this collision are peeled off, debunched, stored, and finally injected into the Tevatron for use in collisions. There are four parts of the antiproton system: the target, Debuncher, Accumulator, and Recycler. The Recycler, which was originally planned for storage of unused antiprotons from the Tevatron, has instead become the final step in antiproton storage and cooling before transfer into the Tevatron.

3.1.3.1 Target

Energies of 120 GeV in the Main Injector were chosen specifically to best produce antiprotons at 8 GeV. It takes approximately 50,000 protons to produce ~ 1 antiproton [10].

The target is made of a single cylinder of Inconel, a nickel-ion alloy, chosen because of its ability to withstand high stresses due to rapid beam heating.

Since the momentum spread is not important for protons about to hit the target, the protons undergo bunch rotation reducing the time spread of the particles at the expense of increased momentum spread. A lithium lens focuses the incoming antiprotons in the x and y planes with a very strong magnetic field. The lithium lens was used rather than a traditional quadrupole because of its ability to focus in both transverse planes and produce a very strong magnetic field. It has the disadvantage of

losing $\sim 18\%$ of the antiprotons to absorption because the beam must pass through beryllium end plates and the lithium conductor. Lithium was specifically chosen because of its low density, to minimize the absorption and scattering effects. A pulsed dipole then selects 8 GeV antiprotons. This is shown is Figure 3.7.

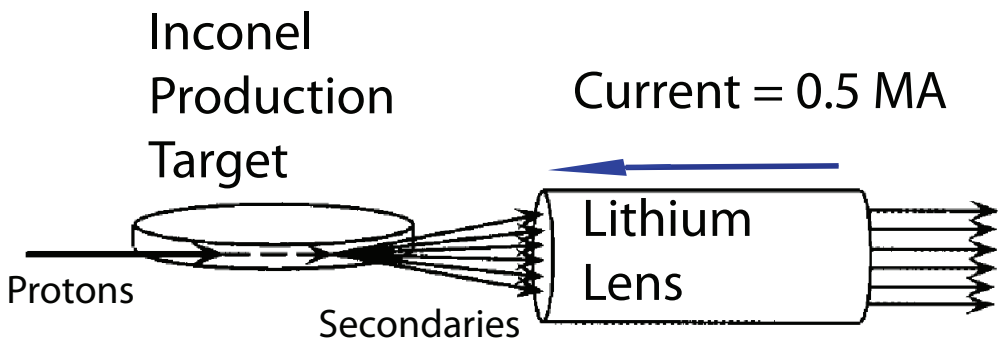


Figure 3.7: A figure showing the $\overline{p}$ target and the lithium lens used to select antiprotons at 8 GeV [7].

From here, the antiprotons follow a beam line to the Debuncher.

3.1.3.2 Debuncher

When antiprotons enter the Debuncher, they have a wide momentum spread. A dipole was used to select antiprotons of ~ 8 GeV, but the momentum spread of entering $\overline{p}$'s is still large [7].

The Debuncher uses bunch rotation to reduce the $\overline{p}$ momentum spread. Bunch rotation is the same process that was used on the protons before hitting the target, but in the opposite direction with the opposite goal. By reducing the momentum spread and broadening the time structure (phase), smaller magnetic apertures are effective and stochastic cooling works much better.

After passing through the Debuncher, the $\Delta p/p$ is reduced from 4% to 0.2% or around 18 MeV. This principle is outline in Figure 3.8.

After the reduction in the momentum spread, the particles remain an additional

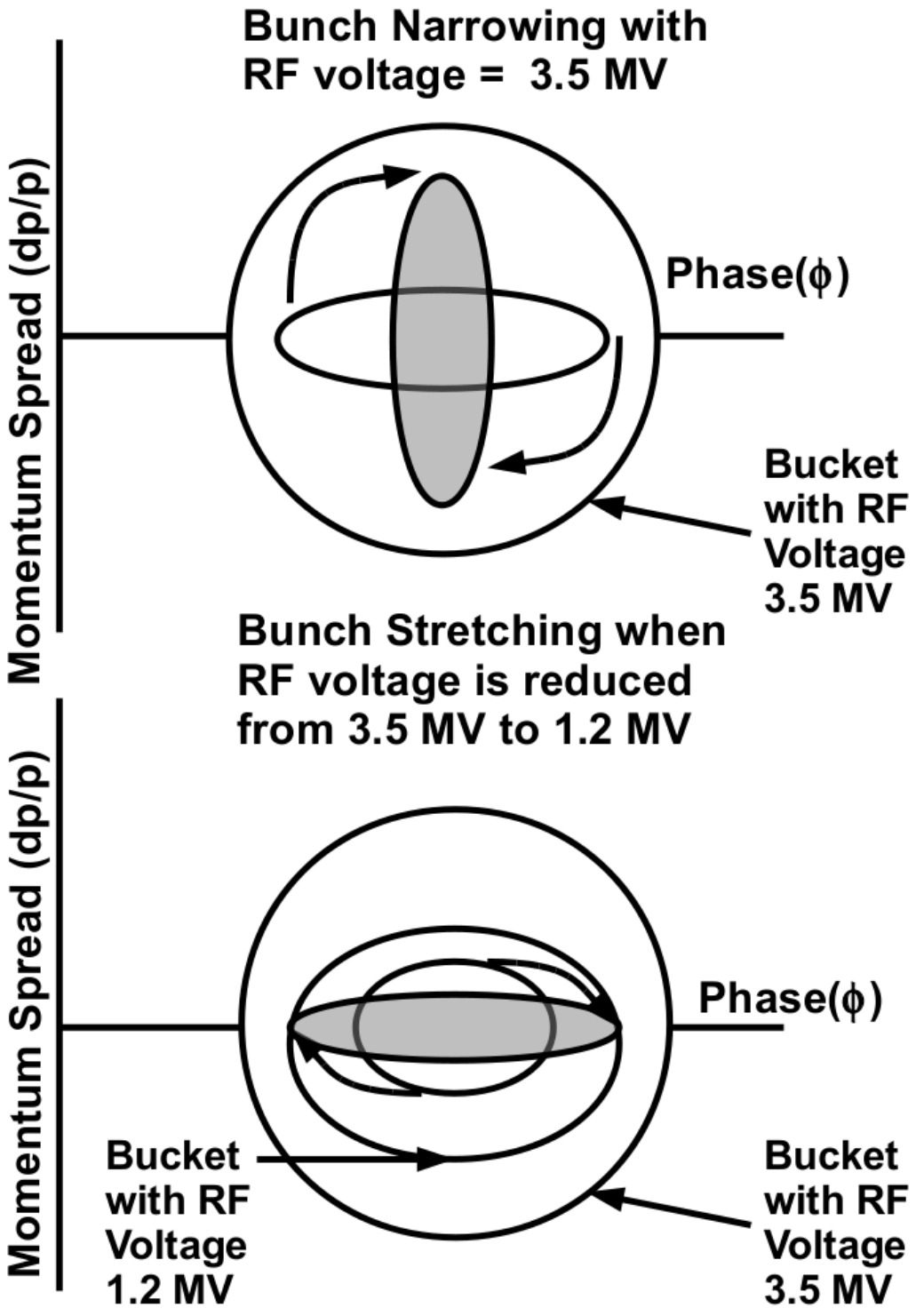


Figure 3.8: This shows the process of bunch rotation. The phase of individual particles is sacrificed to get a more consistent momentum in the beam [10].

two seconds in the Debuncher where they undergo stochastic cooling. Stochastic cooling is the process where the transverse position of a particle is found, related to a betatron oscillation, and sent a corrective signal to dampen the oscillation. The magnitude of betatron oscillations drops by a factor of around two in the Debuncher.

3.1.3.3 Accumulator

The purpose of the Accumulator is to accumulate and store antiprotons. First, 8 GeV $\overline{p}$'s are injected into the Accumulator [10]. The injected beam remains 80 mm outside of central $\overline{p}$ orbit. Then, the beam is decelerated by 60 MeV to move it to the stacktail (the edge of the central orbit). The RF is then turned off there, so the beam is adiabatically debunched, and the momentum of the particles drops by a total of 150 MeV from the injection point to the central Accumulator energy. After 20 minutes the antiprotons reach the core of the beam where they undergo momentum and betatron cooling before transfer to the Recycler. The path of the $\overline{p}$'s within the orbit of the Accumulator can be seen in Figure 3.9.

3.1.3.4 Recycler

The Recycler runs 47 inches above the Main Injector in the same tunnel. It cools the antiproton beam and stores it before injection in the Tevatron, functioning much like a larger and more complex version of the Accumulator. There are four stochastic cooling systems within the Recycler, two horizontal, one vertical, and one longitudinal.

The Recycler also uses electron cooling to cool the antiproton beam [11]. Electron cooling works by sending a beam of electrons parallel to the antiprotons. The $\overline{p}$'s undergo Coulomb scattering with the electrons and lose energy until they reach a thermal equilibrium. The cooling process is shown in Figure 3.10.

Once the antiprotons have been cooled in the Recycler, they are accelerated to 150 GeV in the Main Injector and sent into the Tevatron.

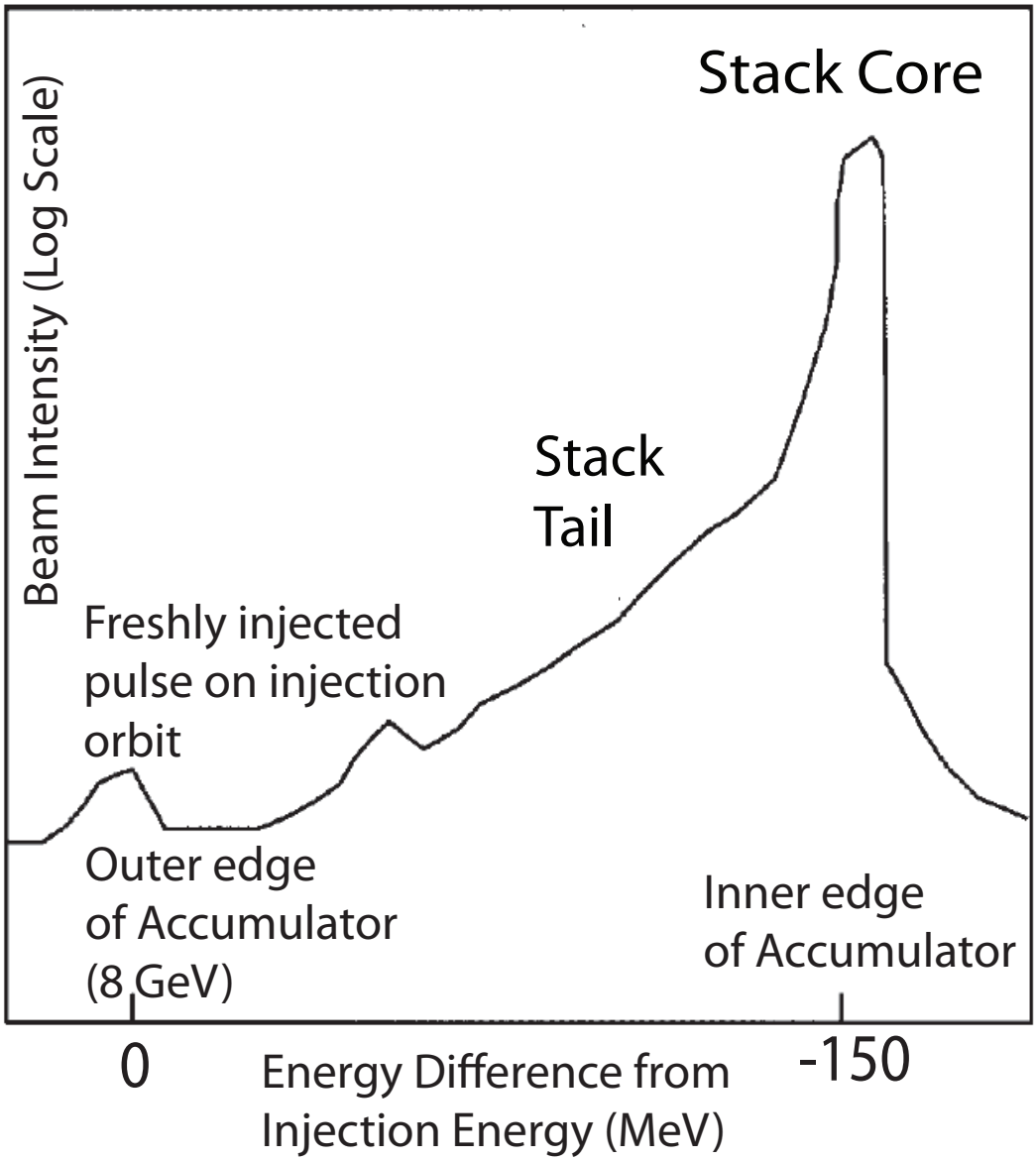


Figure 3.9: The figure shows the path of the antiprotons within the Accumulator as a function of energy. As the particle begins to lose energy, it slowly moves into the center of the orbit [7].

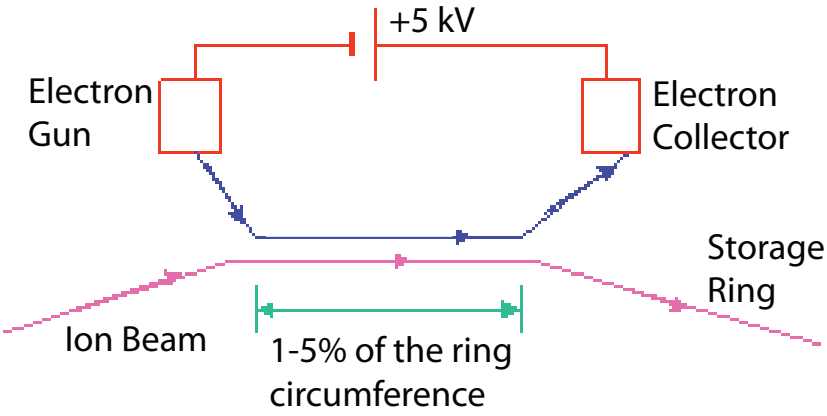


Figure 3.10: The process of electron cooling. A stream of electrons is pushed over the antiproton beam, absorbing energy until the antiprotons are at thermal equilibrium with the cool electron beam [11].

3.1.4 The Tevatron

The Tevatron is a one kilometer radius, 1.96 TeV proton-antiproton collider. It uses only superconducting magnets kept at 4.6 K with liquid helium. The beam pipe is kept at 10^{-9} torr. The beam in the Tevatron is accelerated in eight separate cavities, four are used for antiprotons and four for protons.

Protons and antiprotons are sent to the Tevatron with energies of 150 GeV, and are ramped using the same magnets to 980 GeV. Once the energies in the beam have reached their goal, the Tevatron begins what is called the low β squeeze.

The position of a particle will deviate from that of the ideal. The area of transverse phase space that is occupied by the particle beam is known as the emittance. The amplitude of the beam spread is proportional to a term known as the β function. The value of this function is typically on the order of meters (this is proportional to the beam spread which is on the order of 100's of microns). Focusing quadrupoles at the interaction regions reduce the value of the β function at these areas (known as β^*) to 35 cm. This is equivalent to a beam spread of 10's of microns.

Once the beam size is reduced, collisions commence, but the unstable portion of the beam still needs to be removed. The part of the beam that falls outside of the stable region is known as the beam halo. Collimators are pushed near the beam to remove the unstable beam halo.

A period of collisions, typically lasting between one half to one full day is known as a store. Thirty-six separate bunches are collided, divided into three superbunches of 12 bunches with a 2.6 μs spacing between the superbunches. Collisions at the Tevatron happen every 396 ns within a superbunch. The bunch spacing at the Tevatron is shown in Figure 3.11.

The luminosity is a measure of the number of collisions occurring in a unit time. The average luminosity during Run IIa was on the order of $81 \cdot 10^{30}$ cm^{-2} s^{-1} while this has recently increased to $200 \cdot 10^{30}$ cm^{-2} s^{-1} or more. The average number of collisions in each crossing have gone from an average of around 2.3 early in Run IIa to 5.8 at higher luminosities.

3.2 The DØ Experiment

The DØ detector was proposed in 1983 for $p\overline{p}$ collisions at an energy 1.8 TeV. The first run of the Tevatron took place from 1992 to 1996, leading to the discovery of the top quark among many other significant achievements. The second run began in 2001 with an increase in energy to 1.96 TeV and decreased bunch spacing producing more collisions and provided greater sensitivity to rare physics processes [13].

All the physics detectors at DØ rely on an understanding of how high energy particles from the $p\overline{p}$ collisions interact with matter. The Bethe equation shown in Equation 3.1 describes charged particles interacting with matter through ionization for mid to very high energy particles. The variables are defined in [2] with units in MeVg^{-1}cm^2. A common example is the interactions of a muon traveling through copper, shown in Figure 3.12. Electrons also interact through ionization, but high

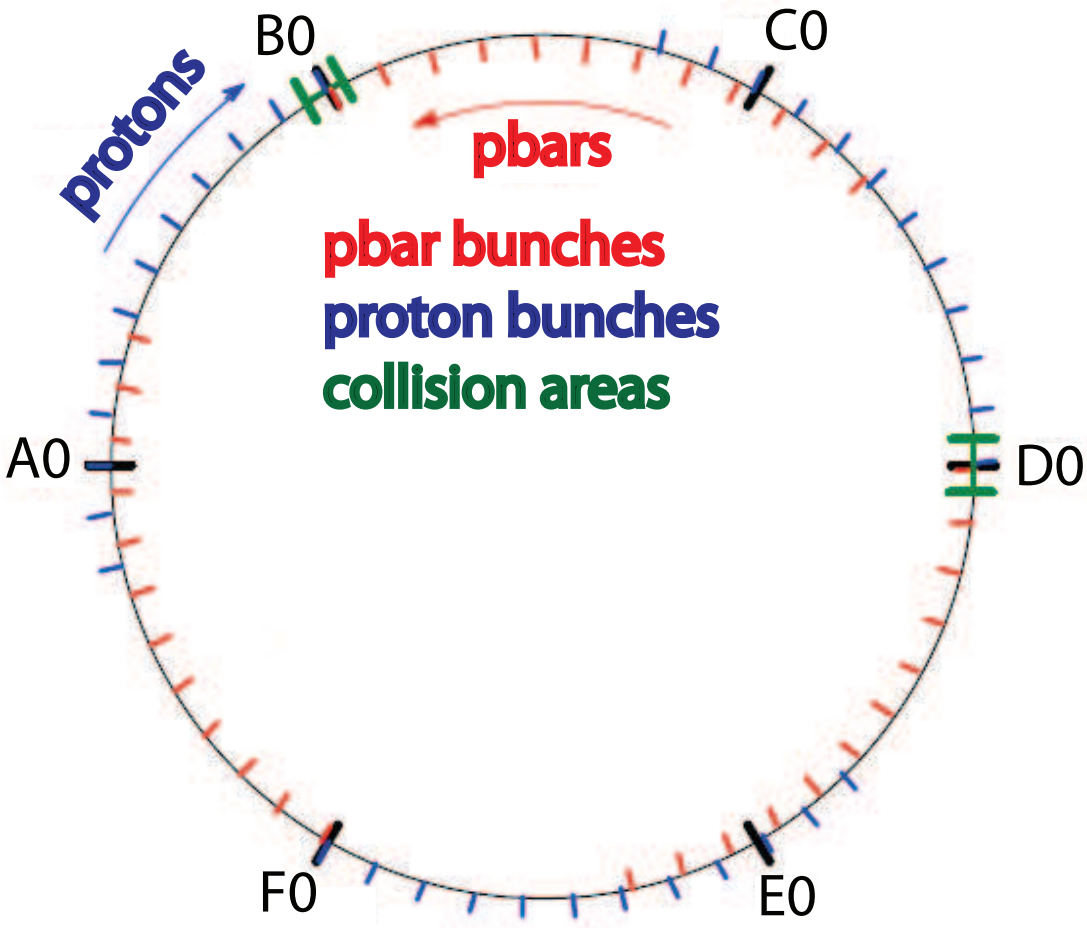


Figure 3.11: The Tevatron bunch structure. Collisions happen every 396 ns within a superbunch. There are three "Trains" of twelve bunches with an abort gap between "Trains" of 2.617 μs [12].

energy electrons at DØ lose most their energy through bremsstrahlung emission of a photon. The relative fraction of energy an electron loses in lead is shown in Figure 3.13. Photons at high energies typically interact through pair production.

$$-\frac{dE}{dx} = Kz^2\frac{Z}{A}\frac{1}{\beta^2}\left[\frac{1}{2}ln\frac{2m_ec^2\beta^2\gamma^2T_{max}}{I^2} - \beta^2 - \frac{\delta\left(\beta\gamma\right)}{2}\right] \tag{3.1}$$

As mentioned above, the interaction of a particle and material is dependent upon the interactions that influence that particle. A charged particle is sensitive to electro-

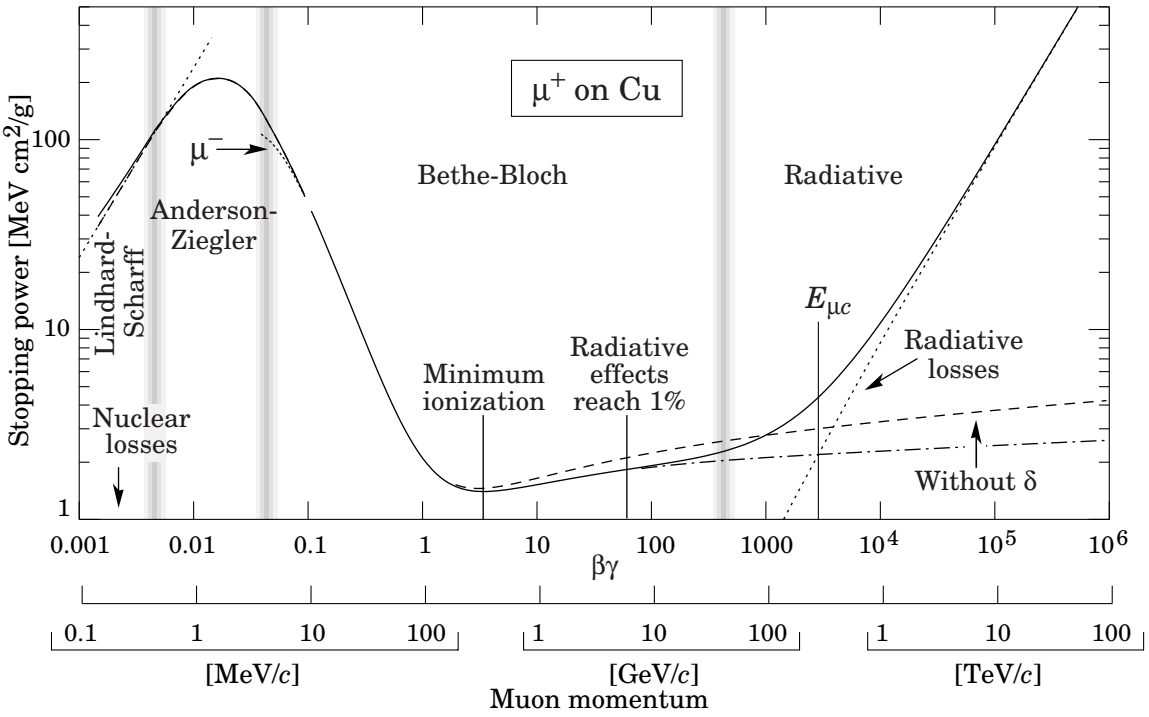


Figure 3.12: The Bethe equation showing the stopping power for a muon traveling through copper. The solid line represents the total energy loss [2].

magnetism, and through ionization leaves tracks in the DØ central tracking system. Electrons also interact through bremsstrahlung with nuclei in the material of the detector. Once the material reaches the density of that in the EM calorimeter, the electrons can lose most of their energy. Photons similarly lose energy in dense materials through pair production. This ties the decays and energy measurements of these two types of particles together. A high energy electron can emit a high energy photon through bremsstrahlung, which will then pair produce an electron and a positron, which can then emit additional photons. This chain of events can continue until average photon energy drops below the pair creation threshold, after which time, Compton scattering is the dominant process. This will ionize molecules by kicking electrons from their bound states. At this point, the shower stops growing. The average energy lost by an electron or photon will be measured in this analysis in radiation lengths (χ_0). This is the amount of material for an electron energy to be reduced $1/e$

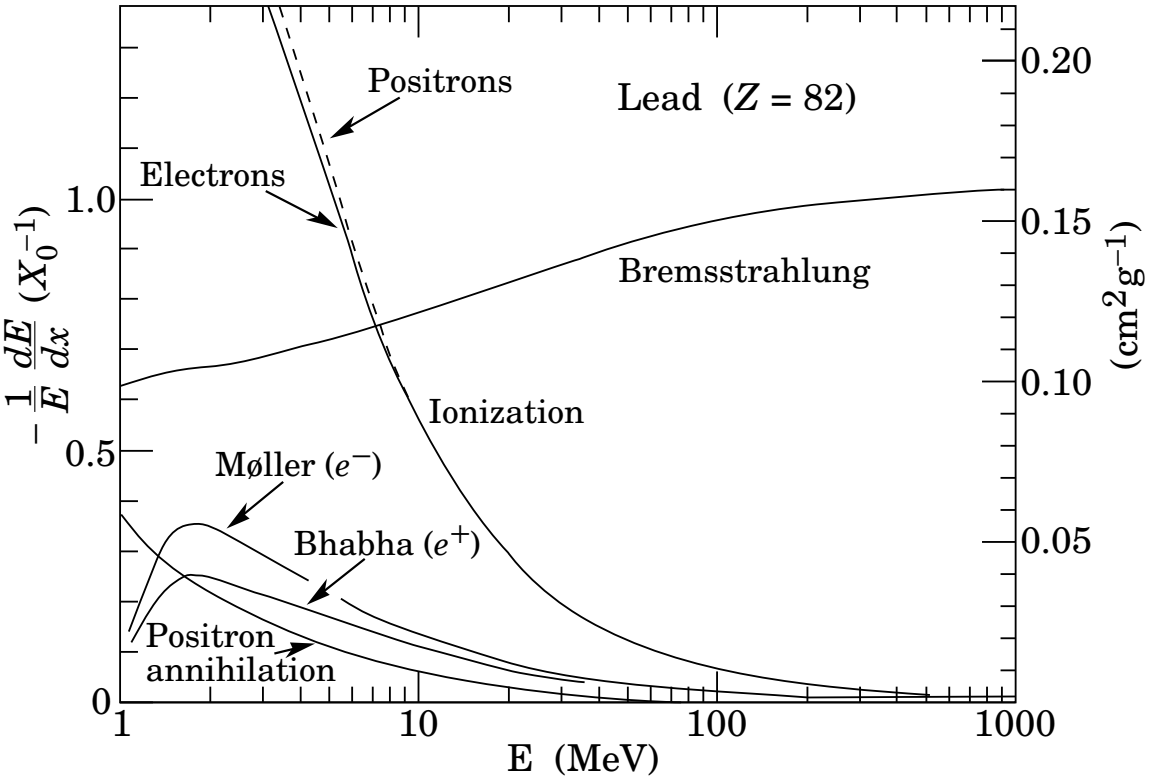


Figure 3.13: High-energy electrons and positrons lose most of their energy from bremsstrahlung. The relative energy loss for an electron or positron in lead per radiation length is plotted against the electron or positron energy [2].

and also 7/9 of the mean free path for pair production of a high energy photon.

The high mass of the muon makes all but the most energetic at the Tevatron below the threshold of significant energy loss through bremsstrahlung. The muon will still lose some energy through ionization, but this is typically of the order of a few GeV. Since muons also do not interact hadronically, and their decay time is considerably longer than it takes to exit the detector, the main way of identifying muons is the fact that they get through the calorimeter to produce a path in the muon system. This path can the be tied to a track in the central tracking system and their signature within the calorimeter of minimum energy loss through ionization. This type of minimum ionizing particle is known as a MIP.

Charged hadrons are also susceptible to ionization within the tracking system, but

the lower cross section of nuclear interactions allows them to pass through the EM calorimeter without losing all of their energy. The hadronic calorimeter was designed to provide enough material for hadronic particles to interact inelastically with atomic nuclei to the point that most of their energy is lost. Hadronic particles shower in a way similar to electrons and photons. As they interact with nuclei they decay into less energetic particles which can decay again and again in a hadronic shower. The π^0 decays into two photons which can decay electromagnetically and provide an electromagnetic component to the hadron showers. The energy grows until the lightest hadronic particles, the pions, can no longer be produced. Particles that interact hadronically have an analogue to the radiation length, called the nuclear interaction length (λ_A). This accounts for energy losses by all types of nuclear interactions.

There are three major detector subsystems: the central tracking system, the calorimeter, and the muon system at the outside of the detector. The main subdetector components and relative detector size can be seen in Figure 3.14. The primary sources used for the explanation of the detector physics were [45, 46], and the primary resources for the DØ implementation of these devices were [13, 47, 48].

DØ uses a right-handed cylindrical coordinate system with positive z oriented along the proton direction and positive y pointing straight up. Given the right-handed coordinate system, the x-axis points out from the center of the Tevatron ring. Several other variables are used when measuring position with the DØ detector. The azimuthal angle, ϕ, is measured from the x-axis in the xy-plane. The polar angle, θ, is measured from the z-axis in the yz-plane. The perpendicular distance from the z-axis, r is defined as $r = \sqrt{x^2 + y^2}$. The polar angle is typically not used in favor of η, the pseudorapidity. This is defined as

$$\eta = -ln\left[tan\frac{\theta}{2}\right]. \tag{3.2}$$

The pseudorapidity approximates the true rapidity,

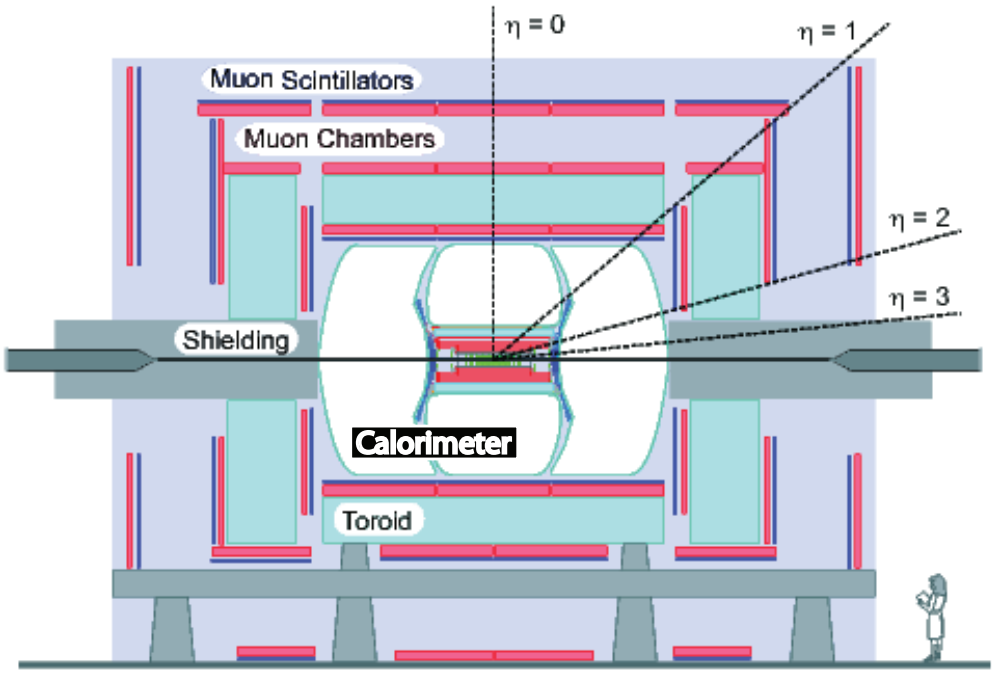


Figure 3.14: The DØ Detector- The z-axis is in the direction of the proton beam, the y-axis is straight up, and the x-axis points out, away from the center of the Tevatron. The central tracking system is within the calorimeter [13].

$$y = \frac{1}{2} ln \left[\frac{E + p_z c}{E - p_z c} \right], \tag{3.3}$$

for finite angles as $mc^2/E \rightarrow 0$. Rapidity is a Lorentz invariant quantity under longitudinal boosts. The pseudorapidity is a more useful quantity than the polar angle both for its invariance properties as well as the fact that particle flux is rather evenly distributed in pseudorapidity so that it is a convenient way to divide the detector in the polar direction.

3.2.1 Central Tracking, Solenoidal Magnet, and Preshower

The central tracking system operates on the principle of fundamental particles minimally interacting with detector components. The tracking system attempts to mea-

sure particle position without interacting strongly enough to change the particle direction significantly or absorbing a non-negligible fraction of the particle's energy. When these detectors are layered, the position measurements of each layer can be combined to reconstruct tracks. Only particles with charge interact enough to provide position measurements. The tracking system lies within a solenoidal magnet (causing the charged particles to bend) allowing for charge and momentum measurements. Outside of the solenoidal magnet, preshower scintillators allow for quick energy sampling to help identify electrons and to assist in tracking before the particles hit the calorimeter, which will be discussed in Section 3.2.2.1. A view of these central detector components can be seen in Figure 3.15.

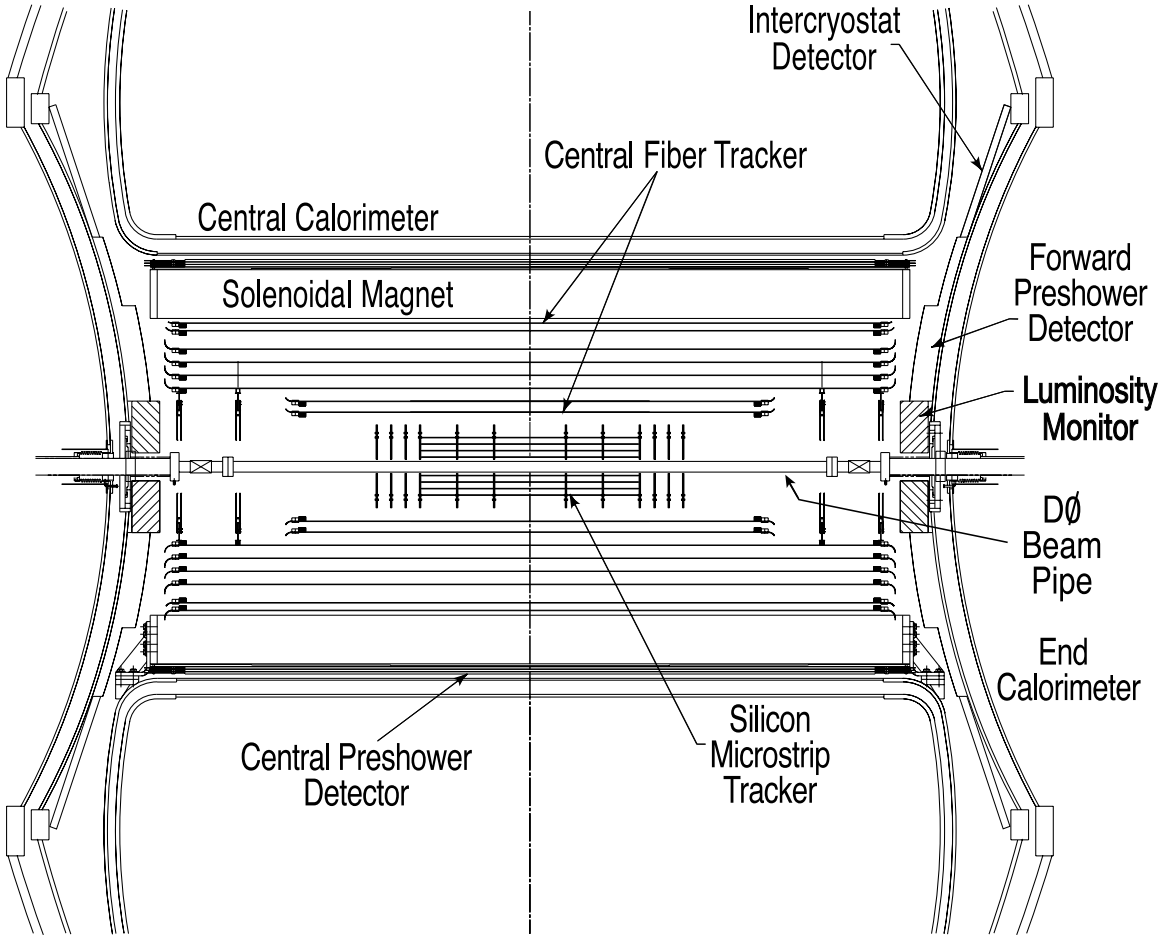


Figure 3.15: The inner tracking system showing the SMT, CFT, the solenoidal magnet and preshower detectors [13].

3.2.1.1 Silicon Microstrip Tracker

The SMT is a silicon microstrip tracker that provides tracking and vertexing information over the η range needed for objects detected in the calorimeter or muon system. Additional information about the silicon detector comes primarily from [14]. In order to produce hits at normal incidence over a range of η values across the extended interaction region, $\sigma \approx 25$ cm, the SMT system uses a series of 12.4 cm long barrels interspersed with disks. Each barrel contains four concentric layers, the closest to the beamline at a radius 2.6 cm and the furthest with a radius of 10 cm. This has been complemented in Run IIb with an additional layer of silicon, layer 0, which resides 1.7 cm from the beamline [49]. Layer zero required a new beryllium beampipe of 1.5 cm radius onto which the detector was attached directly.

The SMT tracks position through the use of 300 μm thick silicon wafers. Two of these 6 cm wafers are placed together in what is called a ladder. The silicon is slightly n-doped, but as it encounters radiation, donor states are removed and acceptor states are created leading to type inversion, and the bulk becomes p-doped. Type inversion allows the tracker to function for longer under heavy radiation than a design for only a single type of doping. The depletion voltage needed to bring charge to the surface of the each of the strips is decreased as the radiation adds impurities to the bulk. This will reduce the depletion voltage until the type inversion, after which it will steadily grow until the microstrip becomes unusable. Only the first layer of the silicon appears to eventually cross the utility threshold in the expected lifetime of the Tevatron.

There are three types of sensor design, single-sided, double-sided, and doubled-sided double-metal. The single-sided modules are used in the Layers 1 and 3 of the outer two barrels and only provide axial information. The double-sided modules are used in Layers 2 and 4 for all barrels and the double-sided double-metal for layers 1 and 3 in the inner four barrels. The double-sided sensors have small stereo angles of 2° while the double-sided double-metal are at larger angles of 90°, for gathering

3-D information for the primary vertex and secondary vertex finding, respectively. Each of these has 50 μm pitch strips on the p-side, while the pitch of the n-side of the double layers varies depending on the stereo angle. Layer 0 has a pitch of 75 μm and a total of 256 channels which are readout outside of the active detector region to minimize the mass that particles must travel through.

The SMT was designed to maximize the number of detector layers each particle went through and to have some particles pass these layers perpendicularly to get the best hit resolution. The SMT location close to the beampipe allows measurements of secondary vertices. These are used in the identification of b-quarks which briefly form B mesons. The B mesons live long enough to have a distinct secondary vertex. Since the barrels will not measure forward particles well, disks have been included to sample these particles with higher η. There are twelve of these F-disks located at $|z| = 12.5, 25.3, 38.2, 43.1, 48.1, 53.1$ cm. The disks use twelve double-sided wedge detectors. Additionally, in Run IIa there were four H-disks located at $|z| = 100.4, 121.0$ cm for very high-η particles. The SMT consists of nearly 800,000 strips providing a spatial resolution of ~ 10 μm. The full barrel-disk structure used in Run IIa can be seen in Figure 3.16.

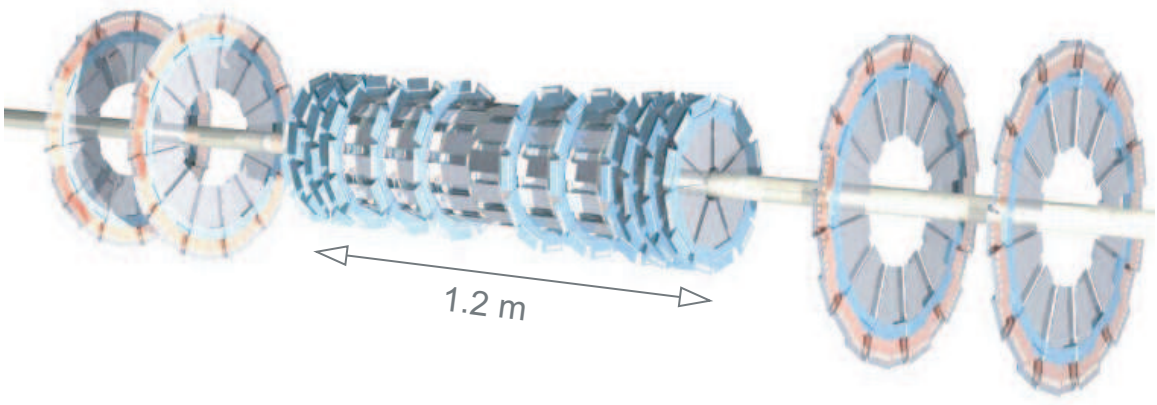


Figure 3.16: The layout of the silicon microstrip detector [13].

Estimates of the momentum and impact parameter resolution are shown in Figures

3.17, 3.18. The addition of layer zero improved the impact parameter resolution by $\sim 55\%$ and expected to improve heavy flavor tagging $\sim 15\%$ [50]. The z resolution is shown to be 35 μm for 90° stereo, and 450 μm for 2° stereo [51]. This is adequate for the goals of primary vertex finding with the small angle stereo and secondary vertex finding with large angle stereo.

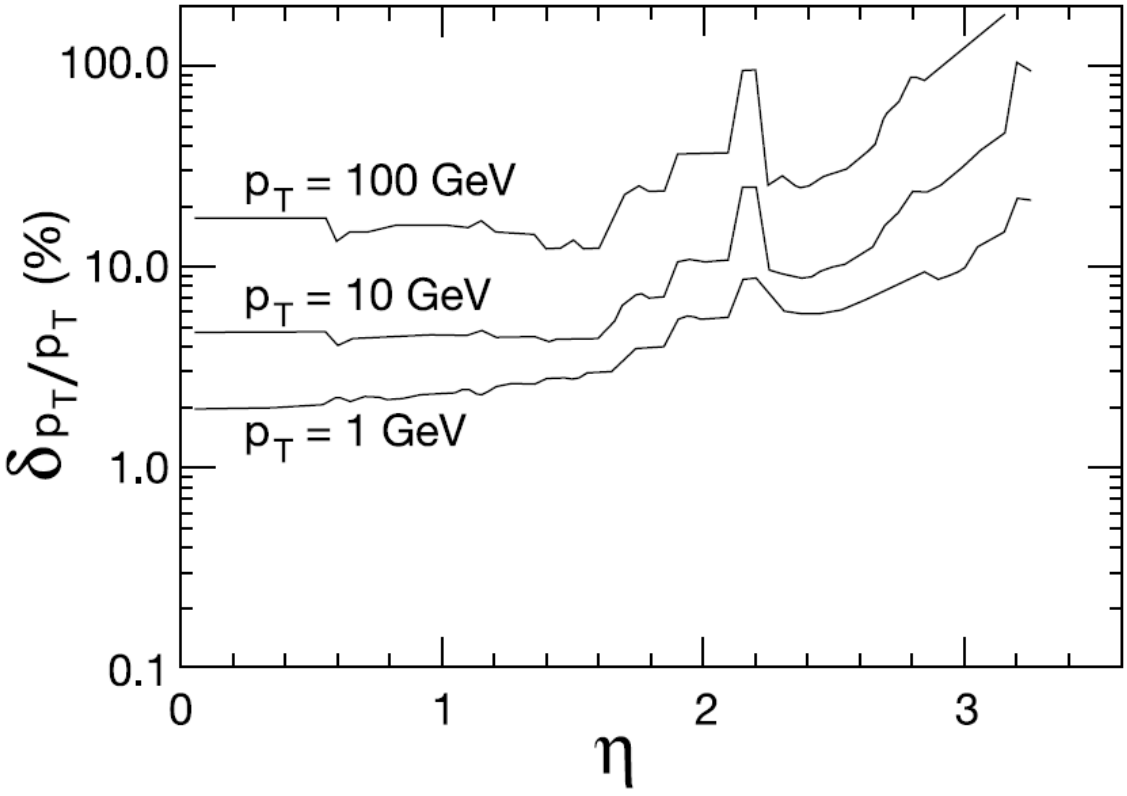


Figure 3.17: The p_T resolution expectation with respect to η for different particle momenta [13].

The currents from the SMT are readout with low-mass Kapton cables using 128-channel SVXIIe chips. The input for one train of beam collisions ($\sim$ 12) is integrated, and then reset during the gaps between the superbunches. On a Level 1 accept signal, the pedestal values are subtracted, and the signals are sent to Wilkinson ADCs to digitize the signal. The first use of the SMT data is in the Level 2 trigger.

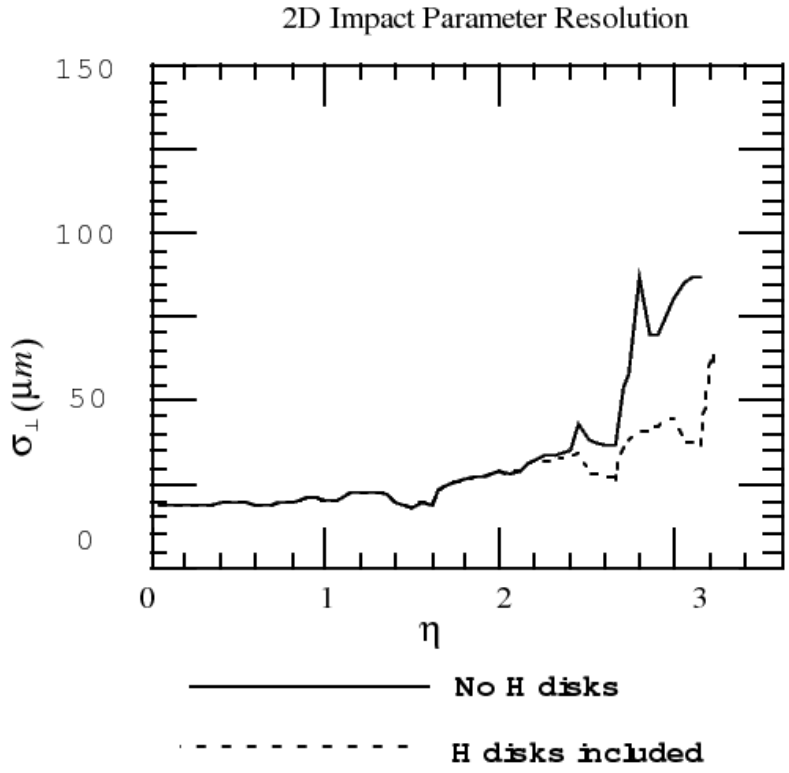


Figure 3.18: The impact parameter resolution expectation from the SMT technical design report [14].

3.2.1.2 Central Fiber Tracker

The central fiber tracker uses a scintillating plastic to determine the position of charged particles through eight concentric cylinders of two doublet layers. The radius of the innermost cylinder is 20 cm, and the outermost is at 52 cm. The inner cylinders are 1.66 m long to allow room for the H-disks in the SMT, and the outer cylinders are 2.52 m providing $|\eta|$ coverage to 1.7. One doublet layer in each cylinder is aligned axially while the second alternates between $\pm 3^\circ$ to give a stereo measurement in the z-plane. The layout of the CFT within the tracking system is shown in Figure 3.19.

The 835 μm diameter fibers consist primarily of polystyrene which when excited transfers energy to paraterphenyl by dipole-dipole interactions. The paraterphenyl emits light at 340 nm, which would be quickly absorbed in the polystyrene. Therefore

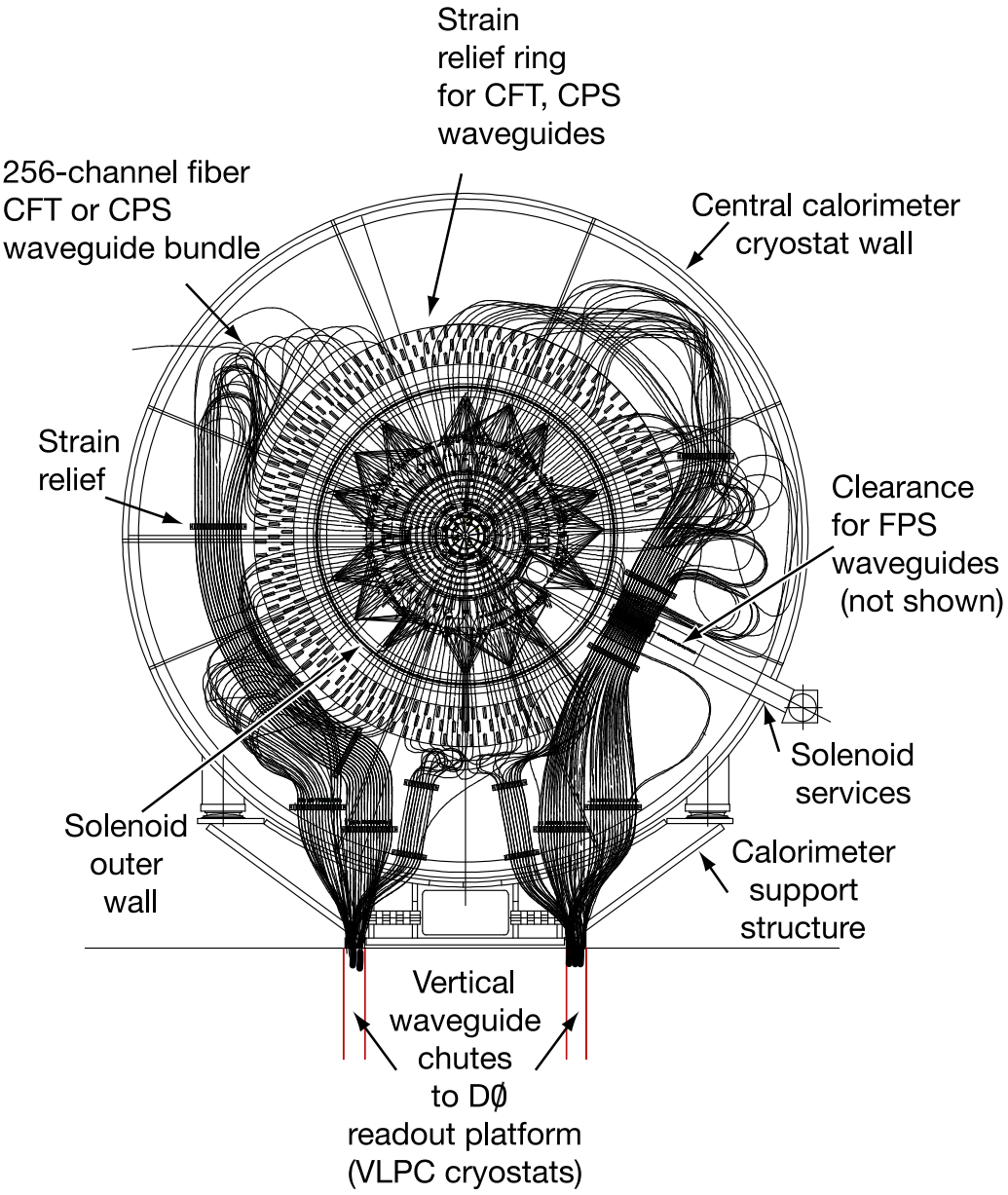


Figure 3.19: The central fiber tracker with supports within the solenoid [15].

an additional agent, 3-hydroxyflavone, is added which absorbs the 340 nm signal and re-emits at 530 nm. This wavelength transmits easily through the polystyrene. Each fiber contains two layers of cladding to maximize internal reflection, and is stopped at one end with sputtered aluminum providing 90% reflectivity. The attenuation length in the scintillating fibers is 5 m. The fibers are connected to clear waveguides fibers which transfer it out from the central tracking region through gaps in the calorimeter to the housing of the VLPC (visible light photon counters) in a cryostat below the central calorimeter. This covers a distance of 7.8 to 11.9 m. The attenuation length of the waveguides is 8 m. All of the fibers total 0.0028 radiation lengths, with the carbon supports 0.0032 χ_0 and the various glues 0.0030 χ_0 for each of the eight CFT layers. This allows high energy electrons to pass the CFT layers without losing a large fraction of their energy.

The VLPC is an avalanche photodetector. It consists of impurity-band silicon with the entering photons creating electron-hole pairs. The holes drift through a depletion zone and into an impurity band colliding with neutral donors and releasing an electron. The electron begins an avalanche by impact ionization with the neutral donor impurities. The gain saturates at $\sim 10^4$. The process producing the electron avalanche is shown in Figure 3.20. The VLPC boards are also used for the central and forward preshower detectors discussed in Section 3.2.1.4.

3.2.1.3 Solenoidal Magnet

The solenoidal magnet was installed to improve the momentum resolution of charged tracks passing through the tracking system. The magnet is 2.73 m in length and 1.42 m in diameter, determined by the available space within the calorimeter.

The magnet was designed to operate in both polarities, provide a uniform field, maximize the tracking area, minimize the materials used, and have adequate safety mechanisms in place in the case of a quench. A field of 2 T was found to be the

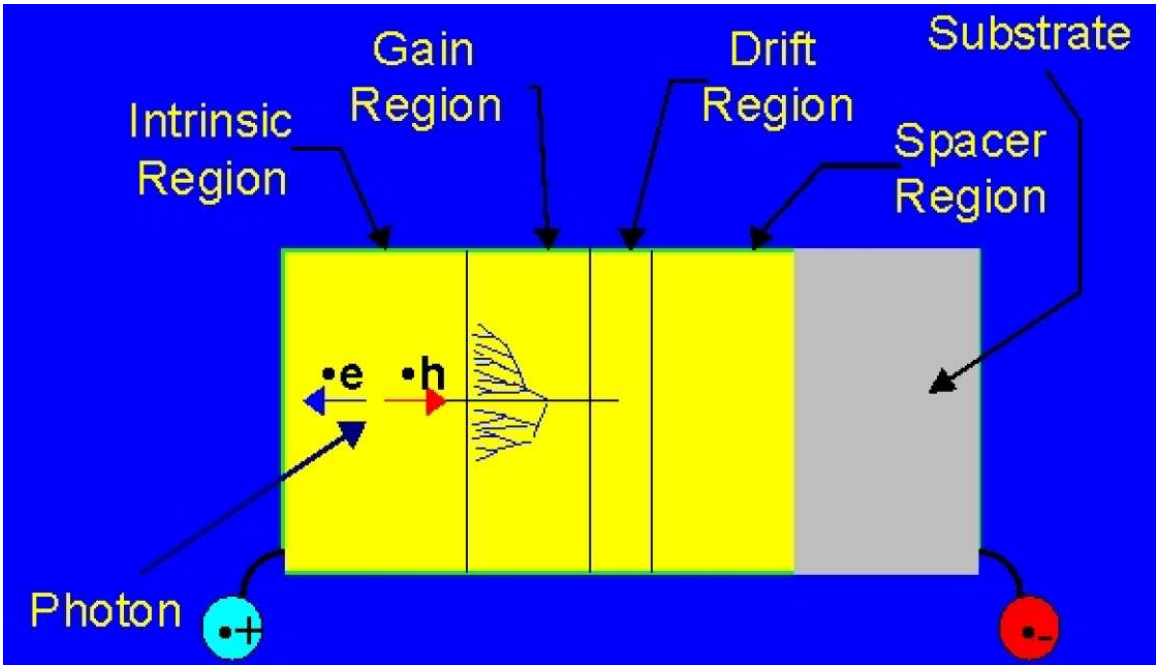


Figure 3.20: The process used to create the electron avalanche from the incoming scintillated light in the VLPC. A photon enters the intrinsic region of undoped silicon creating an electron-hole pair. The hole moves to the drift region where it removes an electron from an atom. The electron accelerates through the gain region freeing more electrons from atoms. The current from these freed electrons is then collected to record the presence of the initial photon [15].

optimal field to best satisfy the above conditions.

The magnet is constructed using strands of Cu:NbTi in a ratio of 1.34:1 and stabilized with aluminum. Each strand is 0.848 mm in diameter with 18 strands in each conductor. The material in the magnet totals 0.87 χ_0, and it is kept superconducting with liquid helium. The full magnetic field is shown in Figure 3.21.

3.2.1.4 Preshower Detectors

Outside of the solenoid, in the 5 cm gap before the calorimeter lie the central and forward preshower detectors. These function in some ways similar to the tracking detectors and others to the calorimeters. Measurements from the preshower can help with electron identification and background rejection by correcting EM shower

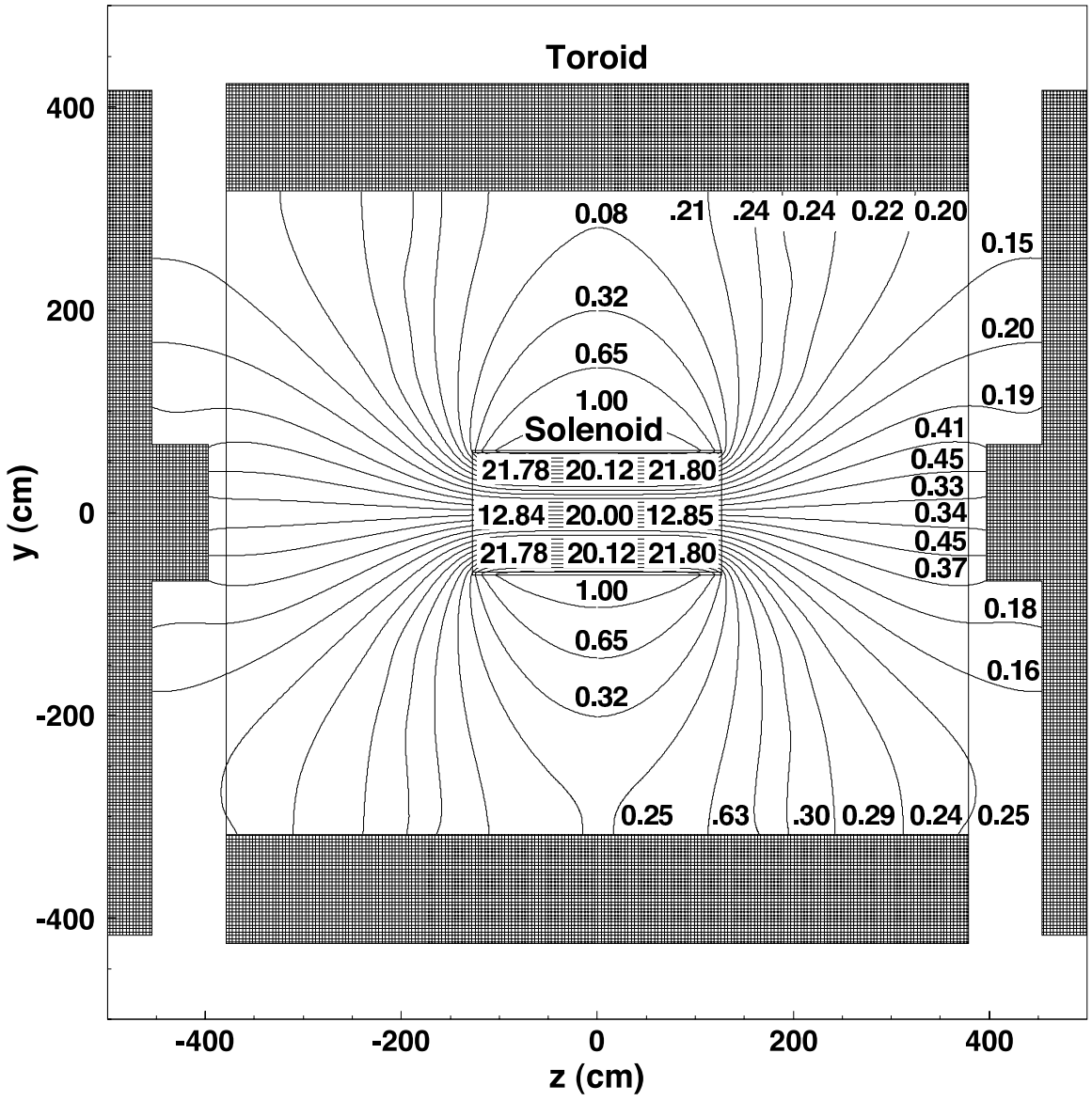


Figure 3.21: The magnetic field seen by particle traveling through the DØ experiment in kG [13].

measurements for energy losses in the solenoid and other upstream material. The particle signals in the preshower detectors are measured quickly enough to allow their inclusion in the Level 1 trigger.

The preshower scintillators are made of triangular strips of polystyrene, as with the CFT. This is mixed with small amounts of p-terphenyl and diphenyl stilbene to allow for transfer of the photon through the scintillator to wavelength-shifting fibers located at the middle of the triangle. These fibers are attached to clear waveguides and are sent to VLPCs, just as in the case of the CFT. The fibers are 835 μm in diameter. The triangular scintillators are shown for each of the detector types in Figures 3.22, 3.23, and 3.24.

The central preshower consists of three cylindrical layers. Before the detectors is roughly one radiation length of lead, with the thickness varied to provide $\sim 2\ \chi_0$ before reaching the preshower detectors for incoming particles in all directions. The central preshower provides coverage of $\eta < 1.3$. Each of the CPS layers consists of 1280 separate scintillation strips.

The forward preshower is located between the luminosity monitor and the inter-cryostat detector. It consists of two layers of two planes of scintillator strips. The first two layers are known as the MIP layers, referring to the minimum ionizing particle. In these two layers, light particles are still not expected to shower too much, only depositing the minimum ionizing energy. The MIP layers are made of 206 scintillator strips. The outer layers are called shower layers. Between each two layers, $2\ \chi_0$ of lead-stainless steel absorber material are placed to induce showering. They are created with 288 scintillator strips. Electrons easily shower in the absorber while heavier charged particles tend to leave MIP signals both in the MIP and shower layers. Photons usually leave no signal in the MIP layer while depositing energy in the shower layer. Each pair of FPS layers are at a 22.5° stereo angle from each other. The FPS covers $1.5 < |\eta| < 2.5$.

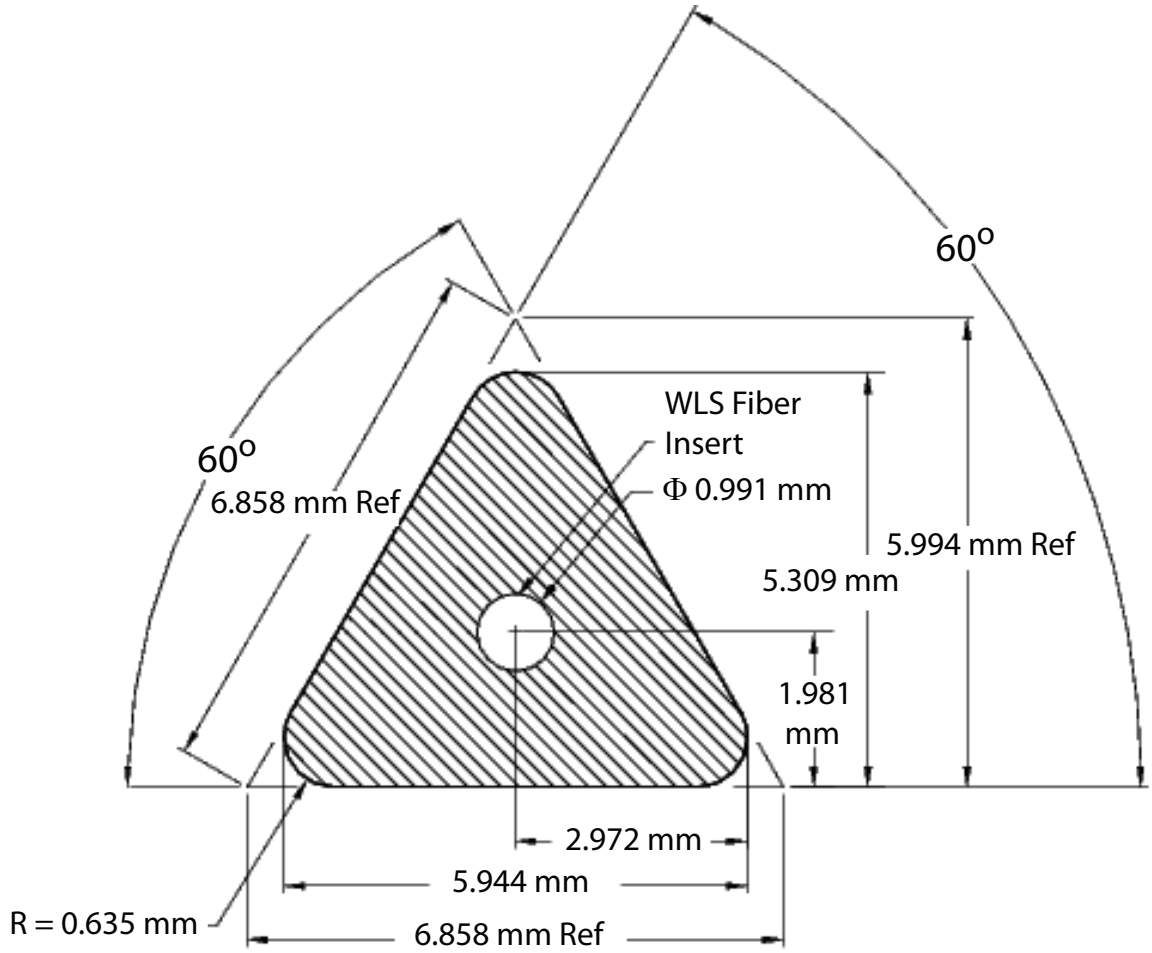


Figure 3.22: The general scintillator geometry for the central and forward preshower system [13].

3.2.2 Calorimeters and Intercryostat Detectors

The calorimeter is the primary tool for particle energy measurement at DØ . A series of absorber plates interspersed with liquid argon and signal boards induce and sample the showering of electrons, jets, photons, and taus. The calorimeter information can also provide shower shape identification of these particles as well as muons. Also, when combined with muon p_T measurements outside the calorimeter, information about non-interacting particles can be inferred from a transverse energy imbalance.

The DØ experiment uses three sampling calorimeters with some additional detec-

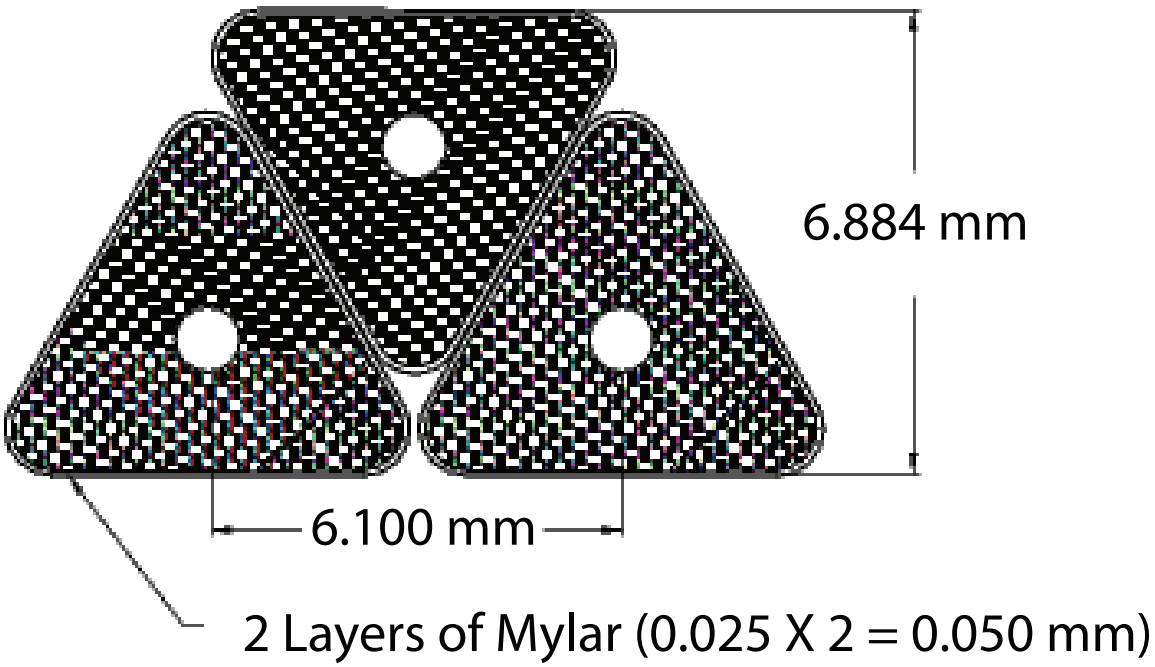


Figure 3.23: The arrangement of scintillation tiles in the central preshower [13].

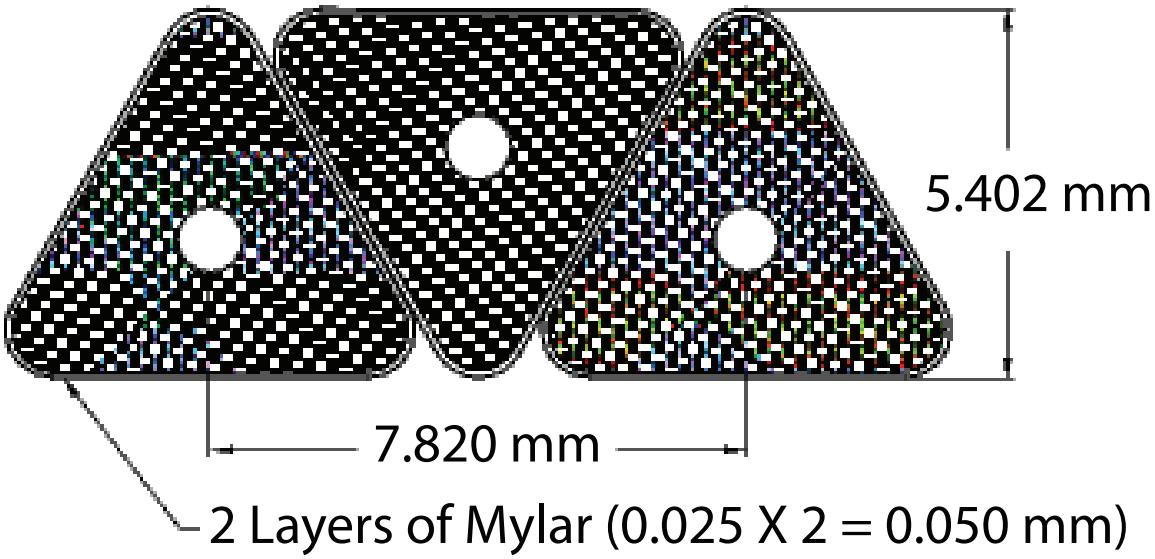


Figure 3.24: The arrangement of scintillation tiles in the forward preshower [13].

tors placed between separate calorimeter cryostats. The cryostats maintain temperatures of 90 K necessary for the liquid argon to be the most effective as an active material. The central calorimeter provides coverage for $|\eta| < 1$ while the end calorimeters extend that to $|\eta| < 4$. All of the calorimeters are segmented into electromagnetic, fine hadronic and coarse hadronic layers. The EM and fine hadronic use a uranium absorber, while the coarse hadronic use copper or stainless steel, and all of the layers use liquid argon for energy sampling. The layout of the calorimeters is shown in Figure 3.25.

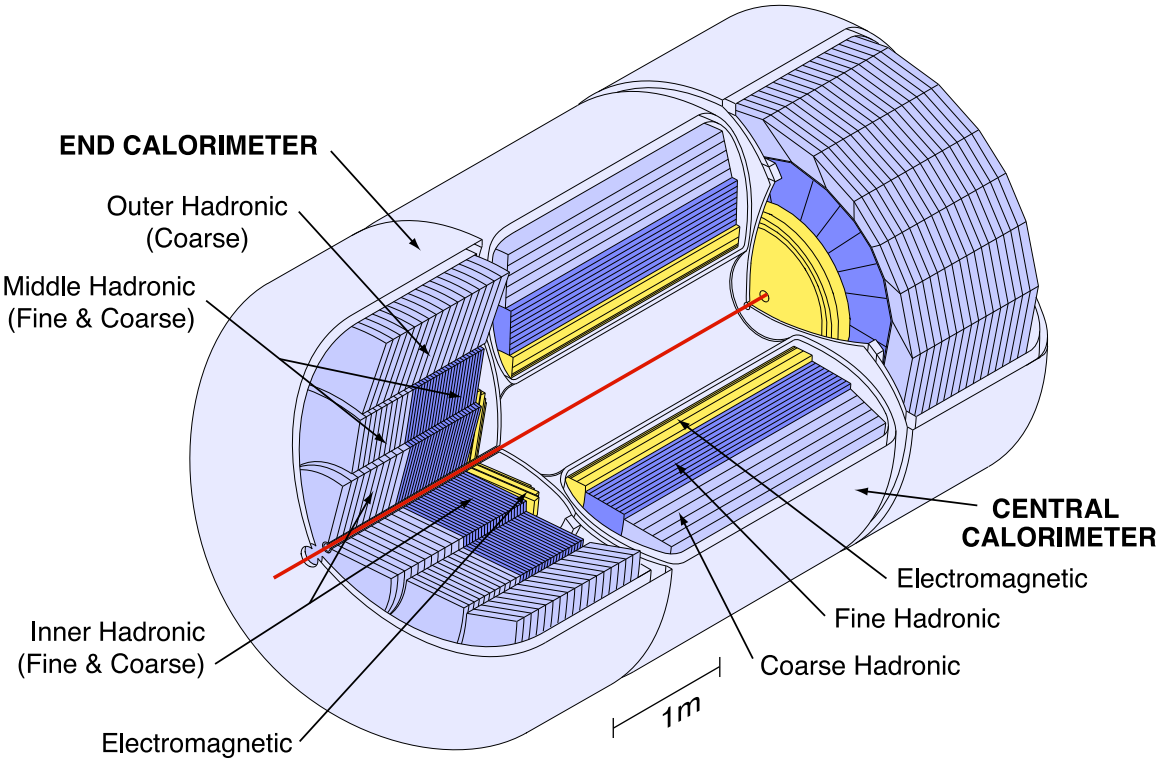


Figure 3.25: The three DØ calorimeters showing the division into layers [13].

3.2.2.1 Calorimeters

The EM calorimeter uses thin plates of 3 or 4 mm of nearly pure depleted uranium. The fine hadronic calorimeter is 6 mm thick with uranium and 2% niobium alloy. Uranium is used because it is a dense material and energy loss is compensated by

nuclear fission. The coarse hadronic is 46.5 mm to ensure an energy measurement of the most energetic particles using copper in the central calorimeter and stainless steel in the end calorimeters. In all cases, liquid argon is chosen as the active material because it is radiation hard, dense and its response is uniform and linear.

In each of the layers, the absorbers are kept grounded while the signal boards have a voltage of 2.0 kV applied to them. Electrons drift across 2.3 mm of liquid argon in about 450 ns. An example of a calorimeter cell is shown in Figure 3.26. The first two EM layers are around 2.0 χ_0. This close spacing is used to help differentiate photons from neutral pions. It is this early shower shape that shows the largest contrast between signatures. Before reaching the calorimeter, a particle would be subjected to about 4.0 χ_0 at $\eta = 0$ and 4.4 χ_0 at $|\eta| = 2$. The total EM calorimeter thickness is $\sim$ 20 χ_0. The central calorimeter is a total of 6.9 λ_A at $\eta = 0$, and the end calorimeters are 9.5 λ_A at smallest angles. The amount of material in each layer is shown in Table 3.1. The end calorimeter outer hadronic is not included. It is entirely coarse hadronic stainless steel and $\sim$ 6.0 λ_A thick.

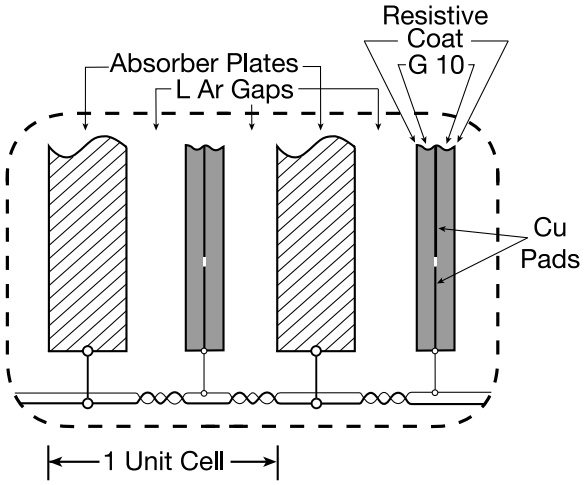


Figure 3.26: A calorimeter cell showing absorber plates, liquid argon and signal boards [13].

The transverse size of the readout cells were chosen to match the transverse shower size- 1-2 cm in EM and 10 cm for the hadronic cells. All of the readout towers were

Table 3.1: Amount of material in each of the calorimeter layers measured in radiation lengths, χ_0, and nuclear interaction lengths, λ_A. The outer hadronic is $\sim 6.0\ \lambda_A$ thick

Cen Cal	χ_0	λ_A	ECal IH	χ_0	λ_A	ECal MH	χ_0	λ_A
<EM 1	4.0		<EM 1	4.4		<EM 1	4.4	
EM 1	1.4		EM 1	1.6		EM 1	1.6	
EM 2	2.0		EM 2	2.6		EM 2	2.6	
EM 3	6.8		EM 3	7.9		EM 3	7.9	
EM 4	9.8	tot 0.76	EM 4	9.3	tot 0.95	EM 4	9.3	tot 0.95
FH 1		1.3	FH 1		1.1	FH 1		0.9
FH 2		1.0	FH 2		1.1	FH 2		0.9
FH 3		0.76	FH 3		1.1	FH 3		0.9
FH 4		X	FH 4		1.1	FH 4		0.9
CH 1		3.2	CH 1		4.1	CH 1		4.4

divided into sections of $\Delta\eta = 0.1$ and $\Delta\phi = 0.1$ except the third EM layer with $\Delta\eta \times \Delta\phi = 0.05 \times 0.05$. At Level 1 and Level 2 triggers, triggers are formed based on a $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ tower sizes.

3.2.2.2 Intercryostat Detectors and Massless Gaps

The calorimeters have gaps in coverage due to their separate cryostats. This causes incomplete coverage in the range $0.8 < |\eta| < 1.4$. This is somewhat mitigated by additional detectors in these regions to sample the energy there.

Single cell massless gaps are signal boards installed within the central and end cryostats and are readout in front of the uranium.

The intercryostat detectors are attached to the exterior surface of the end cryostats. They provide coverage of $1.1 < |\eta| < 1.4$. Cabling from the central tracking system

does not allow a more complete coverage area.

The ICD is made of $0.5''$ thick scintillating tiles of Bicron BC-400 housed in light-tight aluminum. Each tile covers $\Delta\eta \times \Delta\phi$ of 0.3×0.4 and is divided into 12 subtiles to match the calorimeter divisions of 0.1×0.1. Each of the subtiles is read out by two wavelength-shifting fibers on the outside edges. These fibers are connected to clear optical fibers which lead to a photomultiplier tube.

The photomultipliers operate by directing the incoming photons to a photocathode which excites an electron via the photoelectric effect. An applied voltage directs the electron toward a dynode, where it transfers its energy to electrons there. Some of these secondary electrons are emitted and travel toward another dynode, and the process is repeated. This causes an electron cascade, similar to that described for the VLPCs in the CFT section. An anode at the bottom collects the current and amplifies it for analysis.

3.2.3 The Muon System

The calorimeter system is able to measure energy depositions for most particles observed in collisions of interest at DØ . Certain particles such as the neutrino interact so weakly with material that observation of these particles is currently impossible within the constraints of the experiment (or any non-specialized experiment for that matter). Muons are charged particles, so they ionize the material they pass through. The muons of most interest have large momenta. This, combined with its relatively high mass compared to the electron cause the muons to only leave a small amount of their energy in the detector after passing through the calorimeter. Once the muons approach 200 GeV, they begin emitting photons through bremsstrahlung as with electrons, but this energy is rarely reached in collisions at the Tevatron. Therefore, the only muon signals that are seen are ionization tracks in the tracking system. For this reason, additional tracking is added outside the calorimeter. This provides a verifi-

cation of the tracks seen in the central tracking system, identifies the particle as a muon (mostly due to its survival through the thick absorber), and provides separate p_T and charge measurements. These measurements benefit from the much cleaner environment and could allow for better p_T measurements of muons with $p_T > 100$ GeV. The magnetic field in the muon system is similar in strength to that of the tracking system (1.8 T vs. 2.0 T). The greater distance between layers in the muon system perpendicular to the magnetic field (1-2 meters vs. 52 cm in tracking system) will make the bend of a high-p_T muon (and thus its momentum) easier to measure. For lower momenta muons, the broader granularity and greater multiple scattering in the muon system make the central tracking measurements more accurate. Currently, the momentum measurements are taken exclusively from the central tracking system, but a move to use the p_T from the muon system for high-p_T muons is currently under consideration.

The system is divided into central and forward systems similar to the calorimeter. The central muon system has coverage of $|\eta| < 1.0$, and the forward muon system extends to $|\eta| = 2.0$. Both of these systems measure tracks on either side of a toriod.

3.2.3.1 Toroidal Magnets

The muon toroids allow a separate p_T measurement of the muons outside of the central tracking system. The separate measurement allows a quick muon p_T measurement to allow a p_T cutoff in the L1 muon trigger, reject muons from pion and kaon decays, and allow for cleaner matching of the muon to a track in central tracking.

The central toriod is a square annulus 109 cm thick. The inner surface of the toroid is 318 cm from the beamline. It is made up of 20 coils of 10 turns.

The end toroids are located $454 \leq |z| \leq 610$ cm. Each of the end toroids have eight coils of eight turns. The magnet current is 1500 A providing a field of 1.8 T.

3.2.3.2 Central Muon

The central muon system consists of proportional drift tubes for accurate position measurement, cosmic cap and cosmic bottom for time correlations with the beam crossing, and $A\phi$ scintillation counters for fast triggering and additional position measurements.

3.2.3.2.1 Muon Proportional Drift Tubes The central muon drift tubes consist of three layers of drift tubes, one inside the toroid (the A layer) and two outside (B and C). The B and C layers are separated by 1 m. Approximately 55% of the fiducial area is covered by three layers and 90% has at least two. The individual chambers are 2.6×5.6 m^2 created from extruded aluminum. The A layer has four decks except at the bottom which has three. The B and C layers have three decks throughout their coverage. Each chamber consists of 72 or 96 cells, each of which is 10.1 cm wide. An anode wire is fed through the center of the cell. Vernier cathode pads are attached above and below the wire to provide information about the hit position.

Each cell is ganged with a partner. The arrival time at one wire is compared to the arrival time of the partner. Using the time difference between the two hits, a the location of the particle can be inferred. Additionally, the charge distributions are checked for a more precise measurement. The resolution in the PDTs is 1 mm. The charge division method is only used in the A layer. The B and C layers only use this method in 10% of the cells for monitoring purposes. It was found that these additional measurements would have minor improvement on the resolution at high cost.

The PDTs use a gas mixture of 84% argon, 8% methane, and 8% CF_4. The anode wires are kept at 4.7 kV, the cathode pads at 2.3 kV, and the aluminum case is grounded. The drift velocity of an electron in the gas was found to be 10 cm/μs

which gives a maximum drift time of 500 ns. There are a total of 164 proportional drift chambers for 11,386 anode wire cells. The layout of all of the wire chambers used in the central and forward regions can be seen in Figure 3.27.

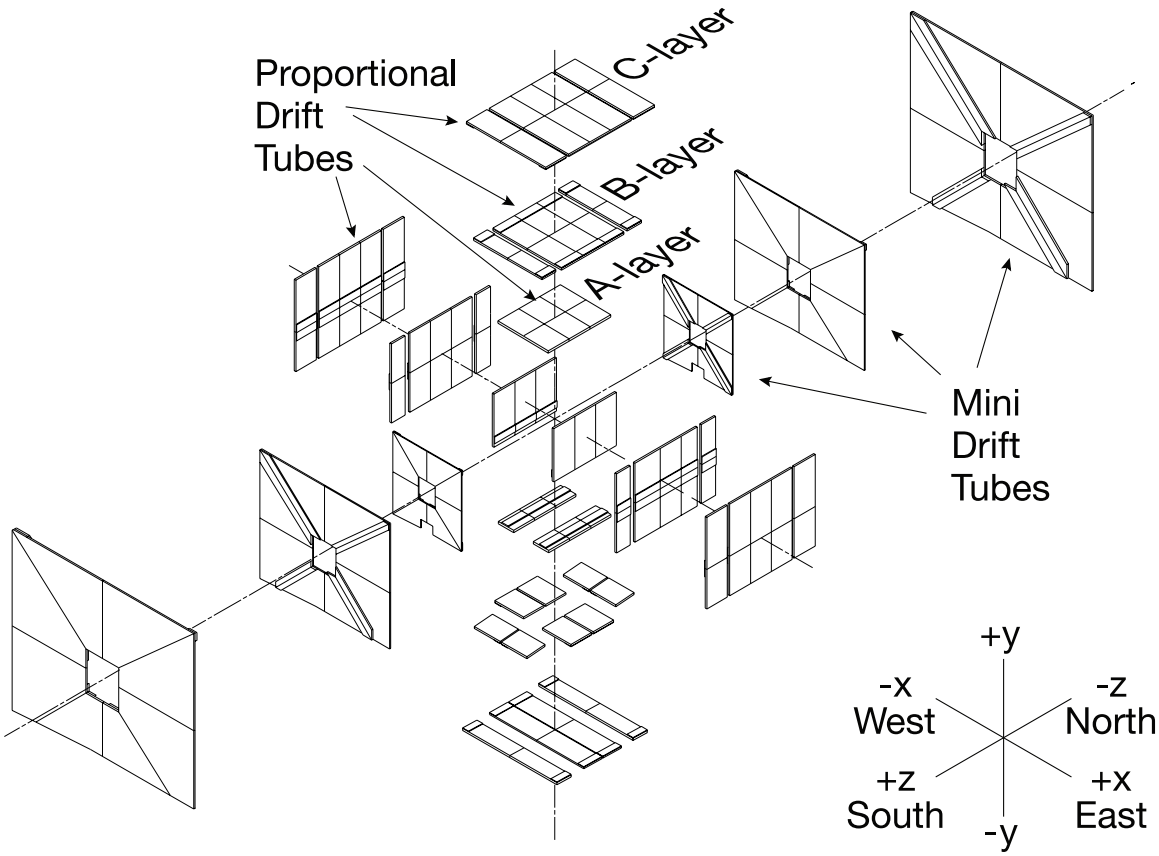


Figure 3.27: The layout of the wire chambers used in the DØ muon system [13].

3.2.3.2.2 Cosmic Cap and Cosmic Bottom The cosmic cap and bottom are installed on all sides of the detector to provide timing information for scintillation hits. These detectors associate a signal in the PDT with a bunch crossing to discriminate against cosmic muons. The scintillators use $0.5''$ Bicron 404A and are readout with a PMT. The layout of all of the scintillators in the central and forward muon system are shown in Figure 3.28.

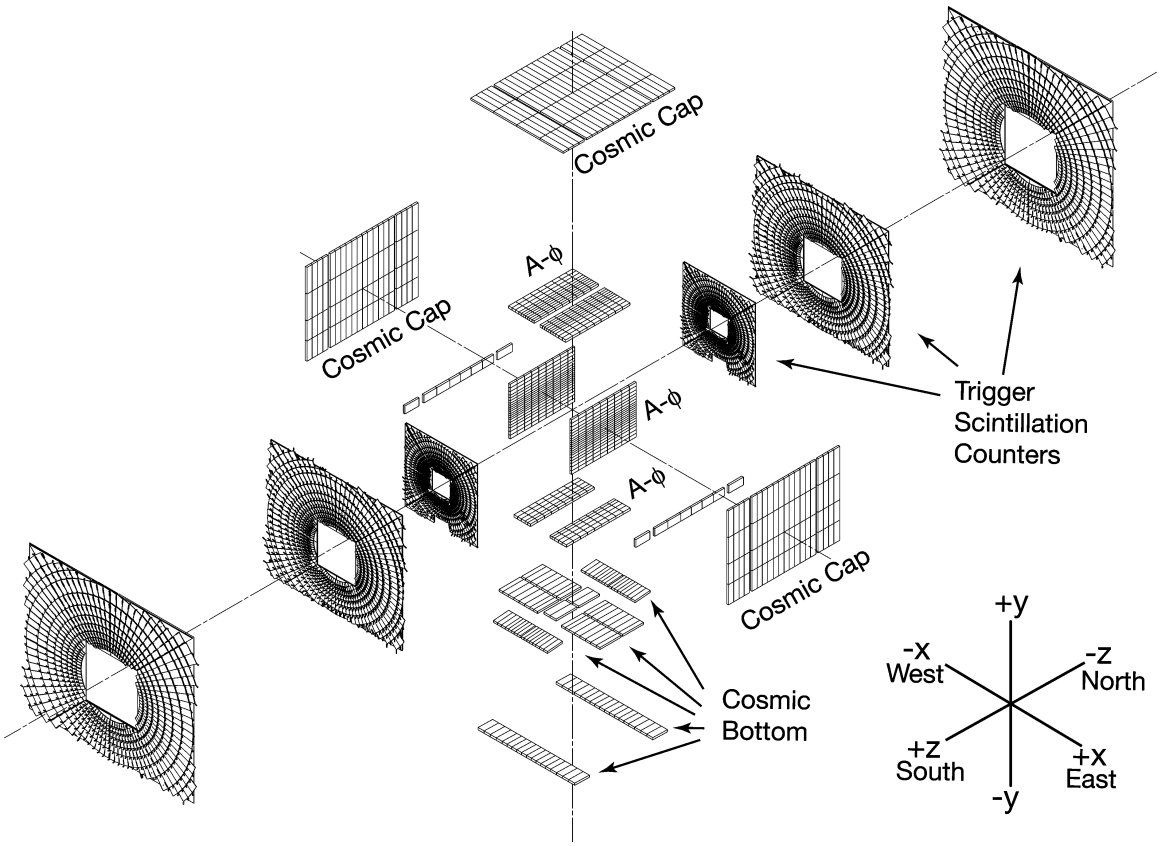


Figure 3.28: The layout of the scintillators used in the DØ muon system [13].

3.2.3.2.3 $A\phi$ **Scintillation Counters** The second layer of scintillation counters is used inside the A layer for triggering and rejection of backscatter from the forward system. This information is matched with CFT tracks for Level 1 triggering of single high-p_T muons and lower p_T dimuons. These counters are segmented by 4.5° in ϕ to match the CFT segmentation. There are nine counters along the z direction. These scintillators also use Bicron 404A and are connected to a PMT. The average muon signal produces 50-60 photoelectrons.

3.2.3.3 Forward Muon

The forward muon system provides coverage up to $|\eta| < 2.0$. The forward system consists of three layers of small proportional drift tubes called MDTs, 3 layers of trigger scintillation counters, and shielding of the beam pipe to reduce energy depositions

from p and $\overline{p}$ fragments and beam halo.

3.2.3.3.1 Mini Drift Tubes The mini drift tubes follow the same principle as the PDTs but with a shorter drift time and slightly better resolution. The drift tubes in the forward region are smaller to account for the fact that muon fluxes are fairly constant in rapidity, so the forward regions need smaller cells to maintain segmentation for the highest η values.

Three layers of MDTs are divided into octants, each of which contains three or four planes of tubes. There are a total of 48,640 anode wire cells with a maximum tube length of 5.8 m. Each MDT is divided into eight cells of 9.4×9.4 mm^2. The wires are made of tungsten and gold with a diameter of 50 μm. As with the PDTs, the MDTs are made from aluminum extrusion combs and covered with stainless steel foil inside of a PVC sleeve. The MDTs use a different gas mixture of 90% CF_4 and 10% methane. The longest drift time in an MDT is 60 ns nearly a factor of ten shorter than in the PDTs.

A voltage of -3.2 kV is applied to the cathode and the anode wire is grounded. Each wire is connected to an amplifier and discriminator. The amplifier discriminator boards link to 32 channels and can detect signals of 2.0 μA. All of this collected information is sent to DAQ. The stand alone resolution of a 40 GeV muon is 20%, and gives a better muon p_T resolution than the central tracking after 100 GeV or in the region $1.6 < \eta < 2.0$ where there are fewer CFT layers. An example of the MDT cells is shown in Figure 3.29.

3.2.3.3.2 Trigger Scintillation Counters The forward trigger scintillation counters are located on each MDT layer both inside and outside the end toroids. Each layer is divided into octants of 96 counters each. The ϕ segmentation is 4.5° matching the CFT. These scintillators also use 0.5" thick Bicron 404A cut into trapezoids. Wavelength-shifting bars are attached to the side of the plate and attach to a 1"

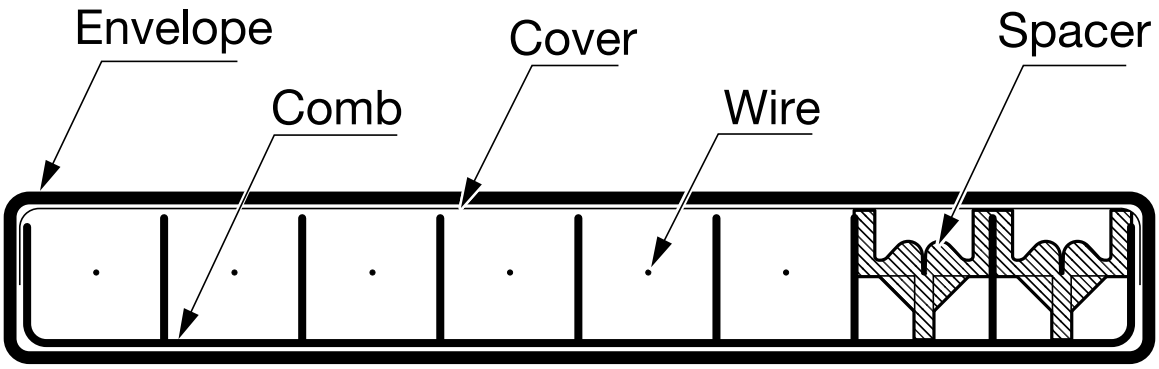


Figure 3.29: A diagram of the individual MDT cells [13].

phototube with 15% quantum efficiency at 500 nm and a gain of $\sim 10^6$. After amplification, the signal is sent to a 10-bit ADC and to a discriminator. These signals are passed to the Level 1 trigger and a scintillator front end TDC. After digitization, the amplitude and time information is sent to the Level 2 trigger and the data acquisition system.

3.2.3.3.3 Beam Pipe Shielding Three sources deposit significant amounts of energy and can limit the lifetime of the muon system without proper shielding. These include the following:

1. p, $\overline{p}$ fragments from interacting with end calorimeter and beam pipe measured in the A layer,

2. p, $\overline{p}$ fragments interacting with the low β quadrupole sending hits to the B and C layers, and

3. beam halos from the tunnel.

The beam pipe is covered with 16" of iron, 6" of polyethylene and 2" of lead to reduce this background. The iron is a strong absorber of hadronic and electromagnetic particles with $\lambda_A = 16.8$ and $\chi_0 = 1.76$. The polyethylene absorbs neutrons because of its high hydrogen content, and the lead absorbs the high energy γ ray photons.

This provides a factor of 50-100 reduction in energy deposition in the muon detector elements, reducing aging effects and limiting interference in particle detection.

3.3 The Trigger System, Data Acquisition, and Luminosity Measurement

The interactions of the particles with the detector have been discussed, but for physics analysis, this information must be passed to permanent storage, reduced in size, and properly interpreted. With an average of six inelastic collisions every half of a microsecond at current common luminosities, we would have to read out and store information about $1.7 \cdot 10^6$ events every second. Full detector readout at this level would be impossible, and the amount of information that would need to be stored would be cost prohibitive and unmanageable. Additionally, even highly unlikely situations in which the detector can mimic a process that has interesting physical properties become relevant. This (as well as the quantum nature of the processes) necessitates the use of statistics to differentiate the observation (or non-observation) of an interesting physics process with a detector effect. In order to anticipate the rate of interesting processes, it is necessary to understand the number of expected interesting collisions. The luminosity system performs this task.

3.3.1 Trigger System

The proton and antiproton beams at the Tevatron cross at DØ at a rate of 1.7 MHz. In order to read out, reconstruct, and store adequate information for analysis, it was found that the rate needed to be reduced to $\sim$ 100 Hz. Most of the physics processes of interest for DØ analyses happen at rates much smaller than the storage rate, so if the 100 events/second are tuned to only store events which might be of interest, very little useful information will be lost. The reduction in rate is the purpose of the

triggering system. The DØ trigger system uses three layers. The first uses hardware to reduce the rate to 2 kHz, a second level system uses firmware and simple software to drop that in half, and a third layer of more complex software provides the final reduction to 100 Hz. The basic layout of the trigger system is shown in Figure 3.30.

3.3.1.1 The Level 1 Trigger

The first level trigger is divided into four sections corresponding to different parts of the detector which are loosely brought together by the trigger framework system. The trigger framework reads information from the subdetector triggers and the accelerator and makes the decision to accept or reject the event.

Incoming events are stored in buffers giving the L1 system 3.5 μs to make a decision. This is roughly a factor of ten larger than the beam crossing rate within a superbunch.

3.3.1.1.1 Trigger Framework The trigger framework is responsible for making accept or reject decisions for Level 1. The framework itself does not provide any further processing. It simply performs a logical OR of all of the trigger terms it receives from the subsystems while accounting for beam conditions that are necessary for each trigger to pass. The Level 1 system has 128 possible triggers, each with its own beam condition requirements, which make a total of 256 terms that the framework checks. On top of this, the framework also monitors trigger rates and deadtime, coordinates trigger vetoes, and handles trigger prescaling if the trigger of interest would fire at a rate too high for the triggering or readout system to handle.

3.3.1.1.2 Level 1 Calorimeter Trigger The Level 1 Calorimeter Trigger monitors 12 EM and 1280 Hadronic towers to look for energy patterns of interest and make sure that events that have large or unusual energy signals are saved for further analysis. The calorimeter is divided into $\Delta\eta \times \Delta\phi$ of 0.2×0.2 for triggering. This is

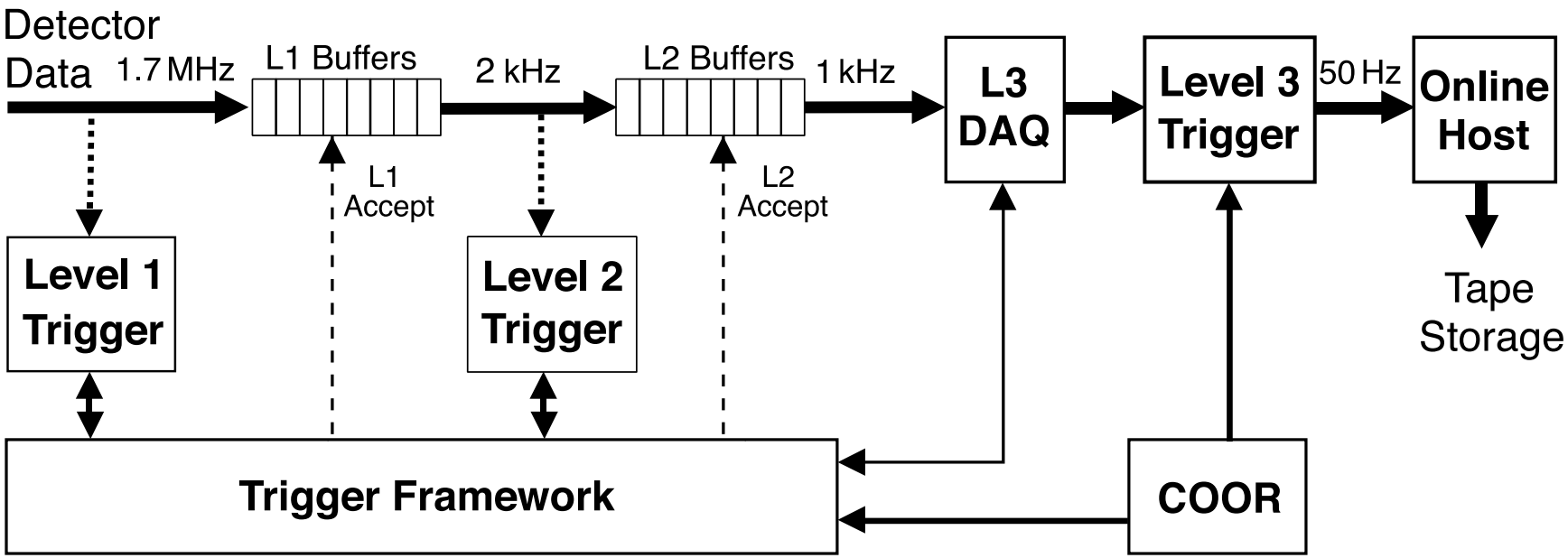


Figure 3.30: The DØ trigger system and basic communication layout [13].

coarser than the subdivisions used in reconstruction.

The triggers include the total sum of transverse energies and the missing transverse energy, both at four thresholds. Also, individual towers are monitored, if a certain number show transverse energy above a limit provided in the trigger list, the event will pass. Additionally, 4×8 towers in $\Delta\eta \times \Delta\phi$ are also checked. This roughly corresponds to the energy deposited by a hadronic jet.

3.3.1.1.3 Level 1 Central Tracking Trigger The Level 1 Central Tracking Trigger uses fast discriminator data to look for matched hit patterns in three scintillator-based systems, the CFT, CPS, and FPS. In all of these systems, the discriminator bits from the analog front-end boards (AFEs) are read in and sent to the digital front-end boards. After this point the data is handled differently depending on where the information is coming from. One path handles the CFT and CPS axial information, another deals with the CPS stereo, and a third processes the FPS.

3.3.1.1.3.1 CFT/CPS axial The axial system compares the discriminator information with thousands of predefined tracking equations and looks for a match. Each digital front-end board (DFE) unpacks the CFT data and stores the six highest p_T tracks. This is done in four separate FPGAs with a fifth that sorts the tracks, matches them with CPS clusters, counts the tracks and total p_T and calculates the sector occupancy.

These tracks are then sent over a coaxial cable to the Level 1 muon system which is discussed below. The tracks are also sent to a board which combines 10 sectors into an octant and finds which sector had the most fiber hits, and whether there were any isolated tracks. This information is passed to another board which generates the trigger terms to send to the trigger framework.

If the AFE receives the L1 accept signal, the fiber data is digitized. The digitized CFT signals then travel to Level 2 and Level 3 to be used as seeds for track lists.

3.3.1.1.3.2 CPS stereo This is the information from the two CPS stereo layers providing three-dimensional information for triggering. Here the digital front-end boards (DFEs) store the discriminator bits but do not begin processing until after receiving a Level 1 accept. The processing consists of a search for hit clusters, and then sorting those clusters to be sent to the Level 2 preshower (L2PS) crate and the Level 3 readout.

3.3.1.1.3.3 FPS The forward preshower is processed in three steps, first the clusters are found, then they are combined, and finally the trigger terms are generated. The DFE finds the clusters and saves the list for use in Level 2 (not implemented). The cluster counts are then summed and passed to another board where the trigger terms are produced.

On a Level 1 accept, the AFE digitizes the fiber data, and the DFEs extract the cluster lists. The FPS then sorts the clusters to send to L2PS and the Level 3 readout.

3.3.1.1.3.4 STT On an L1 accept, the L1CTT seed tracks are reformatted for the L2STT system. First is a check for track overlaps, then each individual sextant is checked for tracks. These tracks are then in the proper format to be sent to L2STT by optical fiber for use as seed tracks.

3.3.1.1.4 Level 1 Muon Trigger The Level 1 Muon trigger follows the same principle as the central tracking trigger. The trigger looks for patterns that match the input L1 CTT tracks, wire hits from the MDT and PDT, and the scintillation counters. Scintillator trigger cards, MTC05, match tracks to muon scintillator hits, and separate wire trigger cards, MTC10, match the scintillator information to track stubs in the wire chambers.

Decisions are made for each octant of the system, and this information is correlated at the muon trigger crate manager. The manager forms 256 L1 Muon triggers and

sends 32 of them to the trigger framework.

3.3.1.1.5 Level 1 Forward Proton Trigger The FPD trigger also follows the CTT and Muon L1 triggers in operational principle. Discriminator signals are sent to three DFEs. These discriminator signals are matched against predefined hit patterns. If a match is found, the event is saved. Events that have very large hit multiplicities are ignored because this is most likely due to beam halo.

3.3.1.2 The Level 2 Trigger

The Level 2 trigger bridges the gap between the hardware for each of the detector subsystems at Level 1, and the software algorithms using the full detector readout at Level 3. The input rate to Level 2 is $\sim$ 2.0 kHz, determined by the digitization rate of the central fiber tracker. Level 2 cuts this rate approximately in half for the input limit of 1 kHz needed for full calorimeter digitization needed at Level 3. Level 2 is the first place that information from all different subsystems is combined globally, and the first place that silicon tracking information is used.

The Level 2 system consists of five different preprocessors, each of which create basic objects such as tracks, EM objects, jets, etc., and send these to the Level 2 Global processor which combines information of these objects to make the trigger decision. The L2 trigger system uses two buses, the VME backplane associated with the crate, and a 128-bit custom MBUS. The MBUS can handle up to 320 Mbits/s. The VME bus is used for the readout of the L2 crates to be sent to Level 3 and the data acquisition system, as well as communication with the run coordination system and monitoring. The MBUS is used to pass inputs to the L2 processors.

Each Level 2 crate contains several types of common components: one single board computer for controlling the readout of the L2 output over the VME backplane, one dual-port memory for communication between the run coordination system and the Level 2 processors, one MBT (Magic Bus Transceiver) card for collecting input

to be sent to the processors, queuing that information, collecting signals from the serial command link (run and event numbers and information from run coordination system), and transmitting information from the preprocessors to the global crate. Additionally, most crates have fiber input converters and VME transition cards to convert from optical fibers to the Hotlinks used in the L2 system. The *βeta* card is the card used for L2 processing. It is a dual 1 or 2 GHz processing card, with one processor used exclusively for the L2 executable and a second for utilities and monitoring. The STT and Muon crates have additional specialized cards that are specific to those preprocessors. They will be discussed in the section on the given preprocessor.

3.3.1.2.1 Preprocessors Each of the major subsystems sends partial readout information to the Level 2 system. The Level 2 system analyzes this information using preprocessor crates to form simple objects. These objects are sent to the Global processor where they are refined and combined for more complicated decisions.

3.3.1.2.1.1 L2CAL The Level 2 calorimeter preprocessor creates jets, EM objects and missing transverse energy out of 2560 trigger towers. The L2 system receives separate energy information for the electromagnetic and hadronic calorimeters. Jets are formed out of 5×5 towers clustered around seed towers, which are defined as towers with $E_T > 2$ GeV. The EM objects use EM towers within $E_T > 1$ GeV combined with the neighboring tower of the greatest energy. The missing transverse energy calculates the vector sum of E_T from towers. The $\not\!\!E_T$ calculation can set different limits for the minimum tower E_T used in the calculation with η ranges defined in configuration files.

3.3.1.2.1.2 L2MUC and L2MUF Two preprocessors, L2MUC and L2MUF, are responsible for creating preprocessor objects from the central and forward muon

systems. Unlike other preprocessors, the inputs for the muon are coaxial cables that are sent to a CIC (Cable Input Converter) for conversion into the standard Hotlinks format. The muon systems also have an extra stage of processing. The muon sectors are first sent to 800 200 Hz DSPs where an initial stage of processing is done. Each of the DSPs searches for track segments in a small region of the detector. The DSPs are spread over 11 central and five forward VME boards. After the track segments are created, they are sent to the βeta processors where the segments are used to make muon candidates with p_T and quality information.

3.3.1.2.1.3 L2PS The Level 2 preshower processor takes Central Preshower axial clusters and combines them into quadrants. The CPS cluster centroid looks for clusters that match in three layers. These output clusters are checked for a track, and then tagged as either electrons or photons. Currently, the forward preshower information is not sent to Level 2, and the central preshower information is not sent to L2 Global (so not used in the trigger decision).

3.3.1.2.1.4 L2CTT The L2CTT processor was designed to take input from L1CTT and L2STT. Currently, only the L2STT information is used. Three different variables are used to define the STT tracks: initial azimuthal angle ϕ_0, the azimuthal angle at the third EM layer of the calorimeter, ϕ_{em3}, and the isolation. This information is used to provide two track lists sent to Global, one sorted in p_T, and the other sorted by impact parameter.

3.3.1.2.1.5 L2STT Each event that passes Level 1 sends its L1CTT information to the STT. The track from L1CTT is established as the seed. A road is established around the track into the SMT, and hits that are within the defined road are associated with the track. The L2STT only looks at axial strips for the trigger decision. The inner and outer CFT layers with at least three of the four SMT layers

are used to define the track parameters.

The L2STT processing takes place in three stages. First the Fiber Road Card, gets the inputs from L1CTT and TFW and sends it the other modules while managing the data buffers. Then, the Silicon Trigger Card receives SMT data, checks the SMT clusters with the roads defined from the L1CTT. Then, the Track Fit Card makes the final hit selection and applies the fit. The L2STT layout is shown in Figure 3.31.

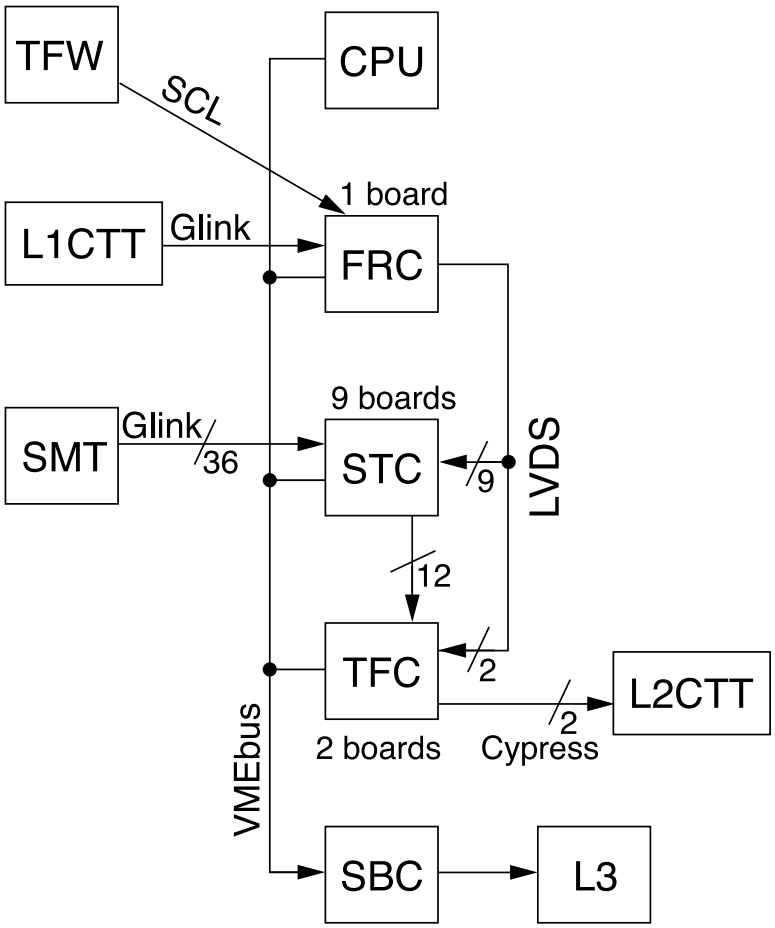


Figure 3.31: Data flow in L2STT [13].

Heavy-flavor events are selected by measuring the impact parameter, and to avoid p_T dependence, the impact parameter significance. This more complex parameter incorporates multiple scattering effects.

3.3.1.2.2 Global Processor The Level 2 Global processor is the first part of the trigger system to look across all subsystems and the object relations among them. The decisions are based on the incoming list of 128 triggers decided at Level 1. Global uses this list to determine which algorithms to run. All of the preprocessor information is available to use, and the Global processor can further refine or combine the objects sent from the preprocessors to make a list of Global objects. The Global objects are then used in the trigger algorithms to determine if the event should pass Level 2. More information on the global processor can be found in Appendix A.

3.3.1.3 The Level 3 Trigger

The Level 3 decision is based on a set of candidate objects or relations that use algorithms called filter tools. The tools unpack the data, look for hits, create clusters, and reconstruct the objects. The parameters for running these processes are stored in COOR-defined lists called refsets. All of the objects or relations that pass a particular filter tool are cached in case they are needed in the future. As with the L2 trigger, the Level 3 trigger has a set of algorithms associated with a particular L2 bit. If the L2 bit is set, only then are these filters run.

3.3.1.3.1 Level 3 Jets and Electrons The jet tool at Level 3 uses a simple cone algorithm, and the ability to suppress hot cells. It has more precise readout than at Level 2, and uses the primary vertex position. Electrons use a jet cone of radius 0.25 in η and ϕ along with E_T, EM fraction and shower shape cuts. They can also require a preshower match.

3.3.1.3.2 Level 3 Muons Level 3 muons use wire and scintillator hits as with the other levels, however, Level 3 can also access information from the inner tracker and the calorimeter. Additionally cosmic ray vetoes are applied using out-of-time information and tracks that penetrate outside of a particular candidate. The muon

tracks are extrapolated to the central tracker, and the track is determined by the fit that minimizes χ^2. These candidates are further mapped to a MIP signal in the calorimeter.

3.3.1.3.3 Level 3 $\not\!\!E_T$ The $\not\!\!E_T$ works by creating intermediate pseudorapidity sums. The $\not\!\!E_T$ is calculated along with the ϕ value of the $\not\!\!E_T$, the scalar E_T, and the $\not\!\!E_T$ significance.

3.3.1.3.4 Level 3 Tracking The tracking works by first fitting a circle through the axial layers, then using a link-and-tree method to join clustered hits from different layers. It starts from the outer layer and works in. The track is then fit to a helix and the smallest χ^2 is found. The CFT vertex and beam spot info are also used for a full 3-D vertex in every event.

3.3.1.3.5 Level 3 Relation Filters Additional filters can be added on top of the individual object filters. Examples include the invariant mass, acoplanarity, and the H_T of the event.

3.3.2 Data Acquisition

The data acquisition system transports data from the VME readout for each crate and transfers it to the Level 3 farms. These are then sent to the online host which uses the information for logging and monitoring. The COOR system controls triggering and data acquisition. The farm data is sent to the collector which then directs each event to a data logger associated with the event's output stream. A copy of the event is also sent to the distributor where it is used for online monitoring in the trigger and physics examines.

The datalogger writes the data to files based on their stream and also creates metadata for storage in the database. DLSAM monitors local data buffers and re-

quests file storage in the ENSTORE tape storage system, $\sim$ 3 km from DØ . The data path from the L3 farm nodes to storage on tape and the examines is shown in Figure3.32.

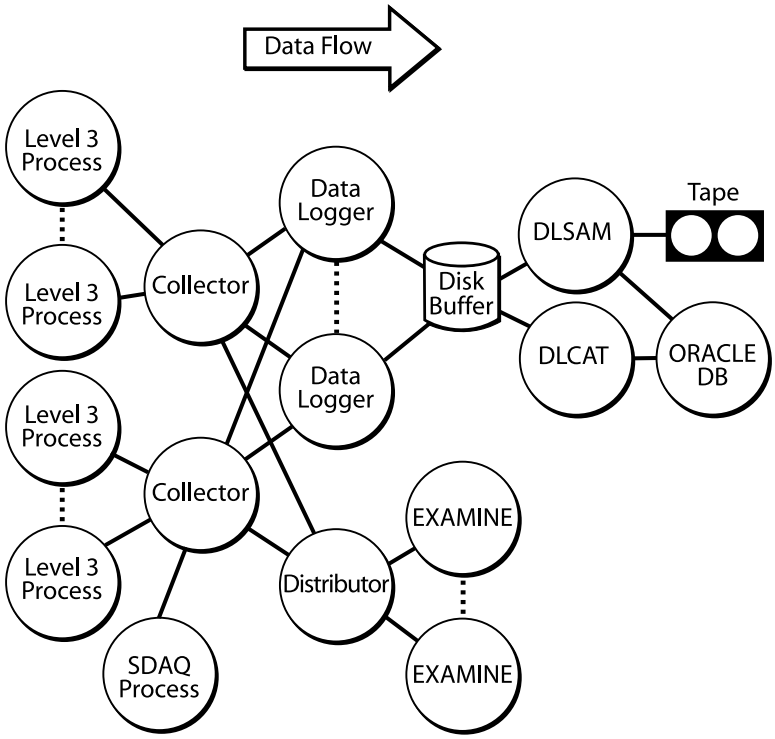


Figure 3.32: Data path from the L3 farm nodes to tape storage and the online examines [13].

3.3.3 Luminosity

The luminosity system is responsible for determining the number of hard collisions in the DØ interaction region measured as a rate of particles per interaction cross section per unit time. This value determines the likelihood of observing a particular process. As the experiment sees more luminosity, increasingly rare processes may be observed.

The luminosity is derived by determining the number of inelastic collisions seen with two scintillation counters at very high values of η. It consists of two arrays of 24 scintillation counters attached to a photomultiplier. These detectors are found at

$|z| = 140$ cm in front of the end calorimeters between the beam pipe and the forward preshower detector covering $2.7 < \eta < 4.4$. The detector setup in the rz-plane can be seen in Figure 3.33.

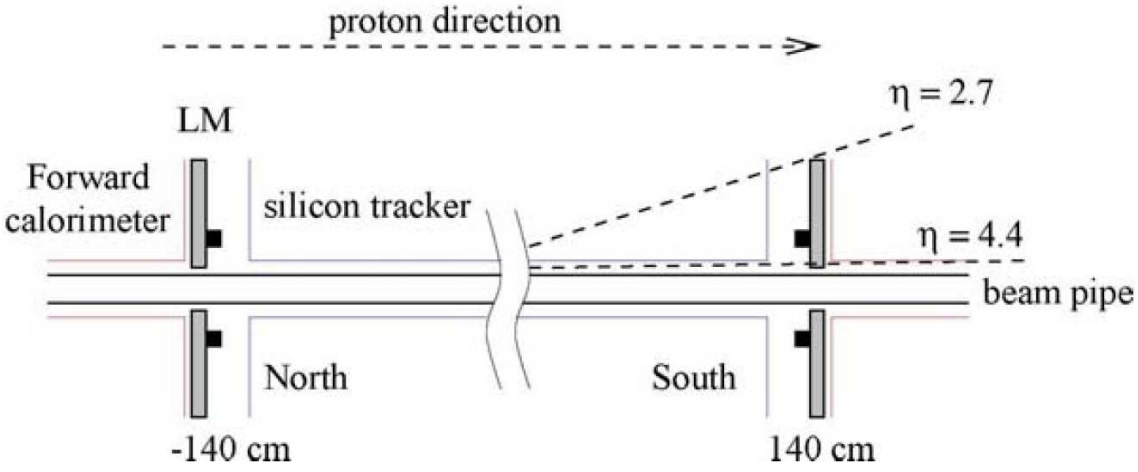


Figure 3.33: The placement of the luminosity detector as seen in the rz plane [16].

The luminosity system must differentiate between hits originating from the collisions in the detector and scattered particles coming from the beam pipe (beam halo). It does this by calculating the time difference between hits at either end of the detector. Using this information it determines an interaction vertex using,

$$z = \frac{1}{2}c(t_{z-} - t_{z+}). \tag{3.4}$$

If z is measured to be less than 100 cm than a collision is assumed because halos will typically show a z vertex of $\sim$ 140 cm [16].

Additionally, the possibility of multiple interactions in a single crossing must be considered. To determine the number of multiple interactions at a given instantaneous luminosity, the number of zero interaction crossings are counted, and a Poisson probability distribution is assumed to determine the average number of interactions.

The luminosity is measured with the following formula,

$$P(0) = e^{\frac{-\sigma_{eff} * L}{f}} * \left(2e^{\frac{-\sigma_{ss} L}{2f}} - e^{\frac{-\sigma_{ss} L}{f}} \right) \tag{3.5}$$

where P(0) is the measured quantity that is determined separately for each of the 36 bunch crossings. These values are determined over the course of a minute so that the measurement uncertainty drops to $< 1\%$ while the change in instantaneous luminosity is negligible.

With the total inelastic cross section at 1.96 TeV of 60.7 ± 2.4 mb, the effective cross section, σ_{eff}, is found to be 48.0 mb and the single side cross section, σ_{ss}, is 9.4 mb.

Luminosity is a one of the largest sources of uncertainty in precision measurements made at DØ . It has a 6.1% uncertainty mostly stemming from the σ_{eff} measurement (5.4%), half of that from the inelastic cross section and the other half from acceptance and efficiency.

Chapter 4

Event Reconstruction and Object Identification

Experimental particle physics tests assumptions about the basic laws of particle interactions. In order to perform these tests, physicists must translate the mathematical predictions of the theory into energy clusters and ionization tracks in the detector. Similarly, the signatures observed in the detector must be translated to determine the underlying physics process between the partons within the colliding proton and antiproton.

Fortunately, in fundamental physical interactions, there are relatively few contributing particles. Of the fundamental particles, electrons, photons, muons, light jets, heavy jets and taus and can be reasonably well distinguished from each other. The light jets include gluons and up, down, strange and a substantial fraction of charm and bottom quarks while the heavy jets are focused on bottom quarks, with a substantial contribution from charm. In order to reconstruct these objects, some basic preliminary information is first calculated. Tracks, the interaction vertex, and secondary vertices are useful in distinguishing among these objects, and they are complex enough that they are often used as independent objects rather than just parameters associated with distinct particle types.

4.1 Tracks and Vertices

The tracking system registers small energy deposits as ionizing charged particles pass through the detector material. When these energy deposits reach a predefined threshold in the tracking system (either in the silicon strips of the SMT or the scinitillating fibers of the CFT), they are registered as hits and saved in the raw data. A combination of these hits are strung together to create the basis for particle tracks through the tracking system (and extrapolated beyond). These tracks can then be traced back to their point of origin providing the initial collision point, the primary vertex, or delayed decay vertices (secondary vertices).

4.1.1 Tracks

Particle tracks are found using the Lorentz force equation to determine particle motion in a magnetic field, Equation 4.1.

$$\frac{d\mathbf{p}}{dt} = q\mathbf{v} \times \mathbf{B}. \tag{4.1}$$

When the magnetic field is uniform, the equation describes a helix, with radius, $r = \frac{Bc}{p_T}$, where c is the speed of light and p_T is the transverse momentum of the particle. The reconstruction system propagates tracks across detector surfaces based upon their geometry and material composition. A detailed look at the propagation of tracks and their error matrices can be found in [52], and the addition of multiple scattering and energy loss effects from material is described in [53].

Track reconstruction at DØ uses two algorithms, the AA and HTF algorithms, and one algorithm to improve the prediction of the path of the track, the Kalman filter.

4.1.1.1 Alternate Algorithm (AA)

Track-finding with the *AA* involves looking for three axial hits in the SMT [54]. The hits in the SMT are checked from the inside, closest to the beamspot and propagated out. The algorithm begins by looking at all hits in a given layer. For each initial axial hit, the SMT layers outside the one containing the initial hit are checked for a second axial hit within $\Delta\phi < 0.08$ of the first. If at least one second hit is found, then a third axial hit in a layer outside of the second must define a circle with $r > 30$ cm, which corresponds to $p_T > 180$ MeV. The track is kept if the impact parameter is less than 2.5 cm, and the overall fit has $\chi^2 < 16$ for the three points. The initial track stub finding can be seen in Figure 4.1.

Once these initial tracking hypotheses are determined, the hits are extrapolated out to the remaining SMT and into the CFT. If an associated hit adds < 16 to the χ^2 value, then the hit is added. If more than one hit satisfies this condition, the track hypothesis is split, and all are kept. These axial hits can correspond to many different stereo projections. As more hits are added, only certain stereo projections are feasible. Tracks with several possible stereo projections may be part of a track hypothesis with the stereo projection only determined after the determination of the primary vertex.

After all of the tracking layers have been checked for hits, the track hypotheses are reduced further by forcing the tracks to satisfy the following conditions:

1. At least 4 hits contain stereo and axial information
2. No more than three layers are missed between any two hits
3. No more than 6 misses in the extrapolation region
4. No more than two misses between layers in the SMT
5. At least five times as many hits as misses

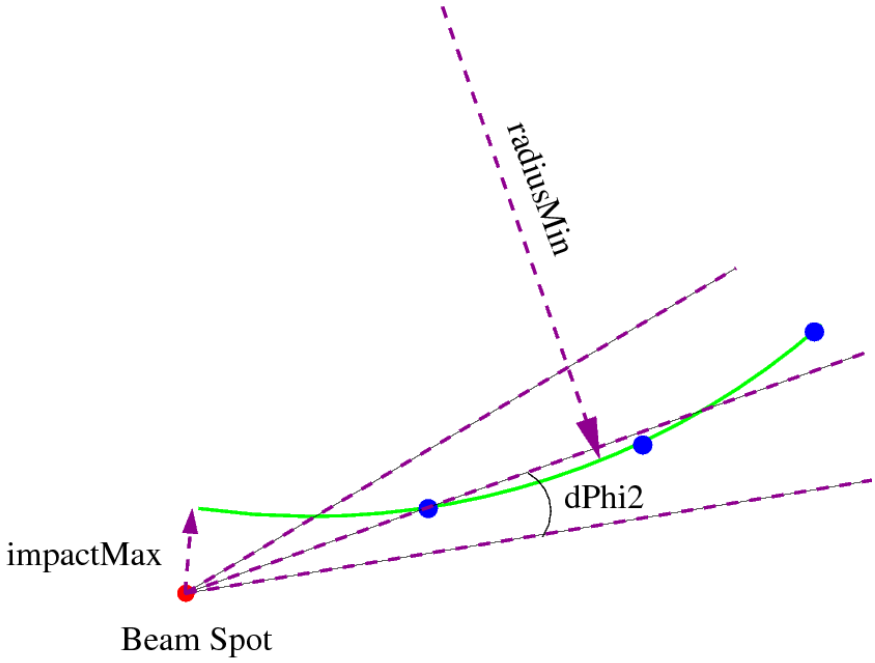


Figure 4.1: The Alternative Algorithm looks for at least three hits in the SMT and extrapolates outward to the CFT [17].

6. If at least one miss between layers, no more than four total misses combining between layers and outward extrapolation, and no more than three total misses combining between layers and inward extrapolation

After this, the final determination of AA tracks is done by eliminating tracks that have too many shared hits.

Using this final AA SMT-based track list, primary vertices are determined from these tracks. Another round of fitting then begins with CFT-based tracks that follow the same seeding procedure except they must connect to the primary vertex with $d_{\text{axial}} < 1.5$ cm and $d_{\text{stereo}} < 1.5$ cm. The tracks are connected through the CFT, and then inward into the SMT.

4.1.1.2 Histogramming Track Finder

In a homogeneous magnetic field with no material, a track can be specified by three parameters in a plane perpendicular to the field [17]. The HTF algorithm specifies these as $\rho,\ \phi,\ dca$ while assuming that the dca (distance of closest approach) is small. This reduces the track parameters to two variables which can then be plotted in a histogram. If each pair of points were plotted, then the tracks could be found from peaks in the histogram with an expected $\frac{n(n-1)}{2}$ entries for a track with n hits.

The HTF method, however, reduces the number of calculations by instead looking at each individual hit, and plotting all possible $\rho,\ \phi$ values that could produce that particular hit. This will produce a line of values in the 2-D histogram. Each hit in the track will produce its own line of values, and the final track parameters can be determined by finding which of these histogram bins has the most hits. The actual track values occur where the lines in $\rho,\ \phi$ space for all of the individual hits intersect. The steps showing the transformation from tracks to peaks in histograms is shown in Figure 4.2.

4.1.1.3 Kalman Filter and Fit

Once the tracks from the AA and HTF algorithms have been determined, they are combined, and duplicates are removed. The points corresponding to each track are then reworked to find the best fit using a Kalman filter and fit [55]. The Kalman filter begins with an individual point in the track and using information from the material composition of the tracking elements (accounting for energy loss and multiple scattering), provides best fit estimates of the tracking parameters, and the associated error matrix. As the track prediction extends inside to out, point-by-point, more and more data points are added to the track parameter determination. The last point should then have the most accurate tracking parameters. After this filtering is finished, the filter is then run in the opposite direction, from the outside in. The

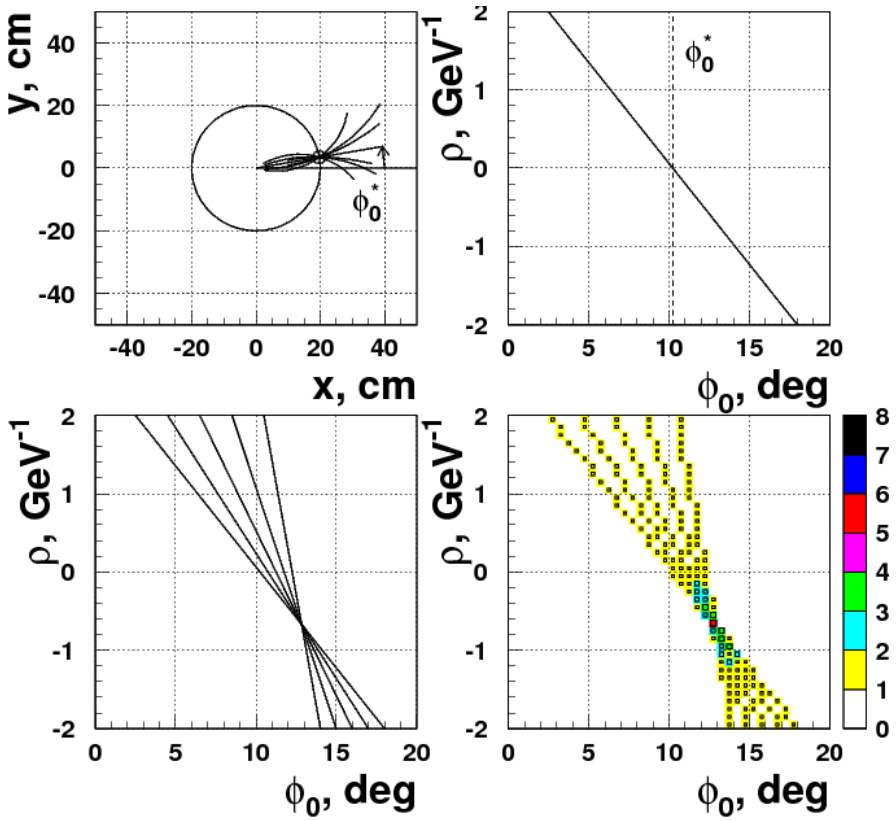


Figure 4.2: The Histogramming Method looks for at peaks in 2-D histograms plotting ρ and ϕ. Histograms with the most hits define tracks [17].

final parameters associated with each point are determined by a fit between those determined from both of these Kalman filtering procedures. This is known as the Kalman fit.

4.1.2 Primary Vertices

Primary vertices are selected by following the tracks back toward their origin and finding where multiple tracks intersect. This is done in several stages [18]. First, only tracks that have $p_T > 0.5$ GeV are chosen. These tracks are then classified as either within or outside of the SMT fiducial region ($\sim |z| < 36$ cm). For tracks within the fiducial region, two SMT hits are required, while those outside have no such requirement. All selected tracks within 2 cm of each other are then clustered.

Then, all of the selected tracks are combined to find a best fit. If the overall fit has a χ^2 per degree of freedom greater than ten, then the track contributing the largest χ^2 is removed until the value drops below the threshold. The remaining tracks are then subject to a cut in the impact parameter significance, $\frac{dca}{\sigma(dca)}$, of five. Once this cut is applied, the remaining tracks enter the adaptive vertex fitting algorithm, the heart of primary vertex selection.

The adaptive fitting algorithm was developed to replace a Kalman filter, which when applied directly, pulled the vertex parameters toward secondary vertices, and when used with a strict cut in χ^2 or impact parameter resolution, lost too many tracks originating from the primary vertex. The adaptive algorithm begins with a normal Kalman filter pass using all of the remaining tracks associated with a particular vertex. Once the best fit is found, a weight is assigned to each of the tracks based on Equation 4.2,

$$w_i = \frac{1}{1 + e^{\frac{\chi_i^2 - \chi_{cutoff}^2}{2T}}}, \tag{4.2}$$

where χ_{cutoff}^2 is tunable and set to 10, and T is set to 1. This equation with T set to zero would give the Kalman filtering procedure with an additional χ^2 cutoff. The given weights for various values of χ^2 and impact parameter resolution are shown in Figure 4.3.

Once all of the weights are determined, another iteration is run to determine the new best vertex position. The weights are then recalculated and the procedure is repeated until the weights converge. This way all of the tracks can contribute to a particular vertex, and if primary and secondary vertices are found simultaneously, each track can contribute fractionally to each vertex.

Finally, after all of the vertices have been determined, one must be selected as the primary vertex [56]. This last stage is done in the following four steps:

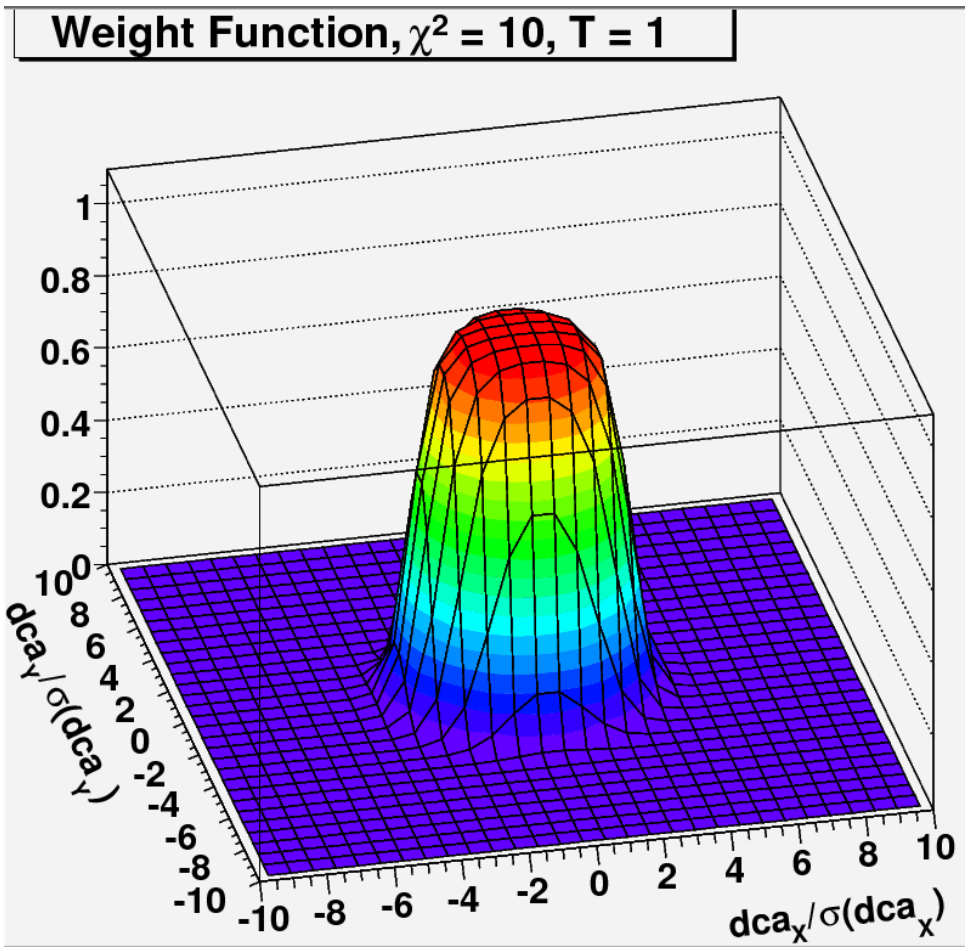


Figure 4.3: Values for the weighting function of a given track to its contribution to a particular primary vertex given a fixed χ^2 and various impact parameter resolutions of the track-vertex system [18].

1. Tracks are clustered within 2 cm.

2. The vertex with highest multiplicity within the cluster is selected.

3. All vertices are given a minimum-bias probability based on the $log_{10}(p_T)$ of the associated tracks. The probability of an individual track originating from a minimum-bias is shown in Equation 4.3, and the probability that the vertex is associated with a minimum-bias interaction is shown in Equation 4.4.

4. The vertex with smallest minimum-bias probability is selected as the primary vertex

$$P(p_T) = \frac{\int_{log10(pt)}^{\infty} F(p_T) dp_T}{\int_{log10(0.5)}^{\infty} F(p_T) dp_T}, \tag{4.3}$$

where F is the minimum bias track $log_{10}(p_T)$ spectrum distribution obtained from a Monte Carlo simulation.

$$PMB = \prod \sum_{k=0}^{N-1} \frac{(-ln \prod)}{k!}, \tag{4.4}$$

where $\prod$ is the product of the individual probabilities of the tracks associated with the vertex.

The only assumption made in determining the probability is that tracks from a hard scatter have higher p_T tracks than those from minimum-bias events. The probability of a track with a particular p_T to originate from a minimum-bias event is determined by integrating over a minimum-bias distribution. The probabilities of all of the tracks in the event are multiplied together, and then the minimum-bias probability for the vertex is determined after removing the track-multiplicity dependence from the probability. The study of the efficiency of this method has been determined for the dataset in this analysis in [57].

4.1.3 Secondary Vertices

Third generation particles are important in the study of many proposed new physics signals as well as rare physics signals in the standard model [58]. Jets that originate from the decay and hadronization of b-quarks are identifiable from a short-lived B-meson that exists long enough to isolate a decay vertex different from the primary vertex. The reconstruction of secondary vertices uses a Kalman-filtering technique that is accomplished in five steps. The specific parameters used in identifying displaced vertices may vary depending on the efficiency/mis-ID ratio determined for a particular analysis. For the MIS search, we look to minimize misidentification at the

expense of efficiency and choose a fairly tight list of parameters to identify b-quark jets.

The five steps in secondary vertex identification are as follows:

1. Find track clusters of 5 GeV within a cone of $\mathcal{R} < 0.5$.

2. Select tracks not associated with the primary vertex.

3. Find vertices by including all tracks within a cluster that add less than χ_0^2 for the χ^2 of the vertex fit, where for tight b-tags, $\chi_0^2 = 3$.

4. Additional vertex selection cuts are made on impact parameter significance, dca, decay length, etc. The parameters used to determine b-jets will be explained in more detail in Section 4.4.2.3.

5. If more than one vertex share a particular track, only the best vertex (based on smaller opening angle and χ^2/dof) will be kept. This is done until all tracks are associated with a single vertex.

4.2 Electromagnetic Objects

Electrons and photons are objects that react in similar ways in the calorimeter. Electrons emit photons through bremmstrahlung, which in turn pair produce electron-positron pairs, each of these again producing bremmstrahlung radiation, with this process repeating until the energy drops below the photon pair production threshold. The photon produces a nearly identical signature in the calorimeter, where it will initially pair produce an electron and positron which produce bremmstrahlung radiation, producing photons that pair produce also giving an electromagnetic shower.

Fortunately, electrons and photons can be distinguished in the tracking system because the electrons ionize the tracking material while the photons can pass through undetected. Additionally, the preshower detector can produce tracking and early

shower signals that can differentiate photons from neutral pions which decay to two photons.

The first step in electron and photon identification is the same, electromagnetic cluster reconstruction, and is outlined in [59]. The experiment uses two methods to identify clusters, the *Simple Cone Algorithm* and the *Single Cell NN*. The algorithm which is most commonly used is the *Simple Cone Algorithm*, which is also the one implemented in this analysis.

The *Simple Cone Algorithm* is based on towers in the calorimeter. These towers are defined by the three electromagnetic layers as well as the first layer of fine hadronic material. For EM objects reconstructed in the central calorimeter, the first step is to find the layer with the highest E_T. Then, all adjacent towers with $E_T > 50$ MeV within a cone of $\mathcal{R} < 0.4$ are added to the initial tower.

In the end calorimeters, the EM clusters are sets of adjacent cells with transverse direction < 10 cm from the initial cell with the highest energy in the third EM layer. This layer is chosen because it has a segmentation of 0.05×0.05 rather than the coarser 0.1×0.1 in the other layers.

After the candidate EM clusters have been determined, they are tested against several criteria to determine whether are not the will be accepted as EM objects. The four conditions are listed below.

1. Cluster $E_T > 1.5$ GeV

2. 40% of cluster energy must be concentrated in central tower

3. The electromagnetic fraction, defined as the energy in the electromagnetic layers divided by the energy in all layers except the coarse hadronic, must be at least 0.9 ($f_{EM} = \frac{E_{EM}}{E_{tot}} > 0.9$).

4. The cluster must also be isolated, as the electrons and photons have narrow shower shapes compared to hadronic jets. A cut of 0.2 is put on the variable

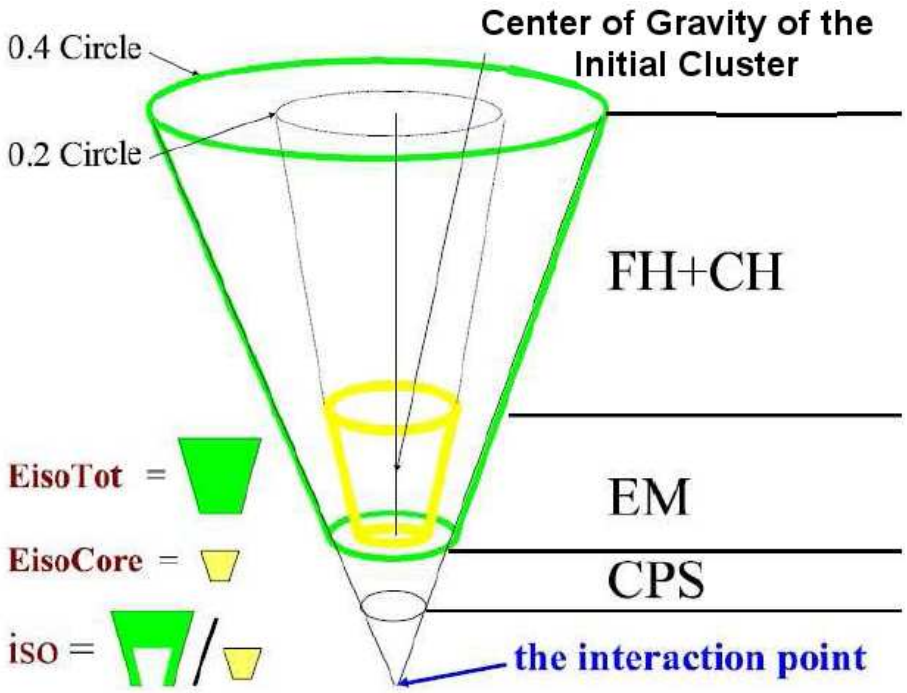


Figure 4.4: The isolation of EM objects is determined by looking at the fraction of total energy in a cone of $\mathcal{R} < 0.4$ minus the amount of energy in the EM calorimeter in a cone of $\mathcal{R} < 0.2$ normalized to the EM energy. The CPS is the central preshower detector [19].

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

The isolation algorithm variables are displayed in Figure 4.4.

The centroid of the deposit is determined from the energy deposition in the third EM layer. A weighted average of energy depositions in this layer is combined with primary vertex information to determine the kinematic quantities associated with this EM object. After this, the preshower layers are checked for hits in areas near the calorimeter cells. If they fall within $\Delta\eta \times \Delta\phi$ of 0.05×0.05 of the weighted average, the energy is recalculated, and the centroid adds the preshower information to determine the direction.

Electrons are then separated from photons by searching for a track. This process uses a hits-on-the-road method. It checks a road in 0.05×0.05 in $\Delta\eta \times \Delta\phi$ between

the cluster and primary vertex. If a track with $p_T > 1.5$ GeV is found, the object is considered an electron. Otherwise, it is considered a photon.

4.2.1 Electrons

The method used above is the most basic way to identify an electron. Depending on the analysis, several types of electron definitions may be chosen according to how strictly electrons are to be differentiated from photons and jets (typically neutral pions). The definitions of many of these separation variables is outlined in [60].

The f_{EM} and f_{iso} defined above can help distinguish between electrons and jets. Several other quantities are also useful in making the differentiation.

A variable called the H-Matrix looks at longitudinal and transverse shower shapes and studies the covariance matrix to determine if the given shower shape is consistent with an electron. The seven variables considered in the matrix are the EM energy fractions in each of the EM layers (showing the longitudinal shower development), the $r - \phi$ cluster width in EM3 (showing the transverse development), the log of the total shower energy and the log of the longitudinal position of the primary vertex. There is a separate matrix for each ring of calorimeter cells with a particular η. The shower shape is classified by its χ^2 value. The matrix cut requires the χ^2 value less than a given cut.

Another variable is the track match χ^2. This is based on the difference between expected ϕ and z values in the track and the cluster value in the third EM layer. The χ^2 value of the track match is defined as $\chi^2_{\text{spatial}} = \left(\frac{\delta\phi}{\sigma_\phi}\right)^2 + \left(\frac{\delta z}{\sigma_z}\right)^2$ where the σ values are resolutions of the quantities.

Finally, the electron likelihood combines several variables [60]. The seven variables used in the likelihood for $p17$ are the following:

1. $P(\chi^2_{\text{spatial}})$, the probability of the track-match χ^2.

2. $\frac{E_T}{p_T}$: does the energy in calorimeter match a certain track?

3. *dca* (distance of closest approach) of track: is track associated with primary vertex?

4. The H-matrix explained above.

5. The EM fraction explained above.

6. The number of tracks within $\mathcal{R} < 0.05$ of the candidate electron track.

7. The total p_T of the tracks within $\mathcal{R} < 0.4$ of the candidate track

The likelihood uses these quantities to determine an overall jet separation. The tighter the cut, the more electron-like the object, and the less jet-like. The likelihood is constructed from a sample of real electrons and jets or photons misidentified as electrons (fake electrons). For each of the seven variables probability distributions for real electrons ($P_S(x_i)$) and fake electrons ($P_B(x_i)$) are determined [61]. With the variables assumed to be independent, an overall real electron and fake electron probability can be determined as shown in Equation 4.5. The likelihood is the ratio of the probability that the electron comes from a real electron over the probability that it comes from either a real or fake electron, as shown in Equation 4.6.

$$P_S(\vec{x}) = \prod_{i=1}^{7} P_{i,S}(x_i), P_B(\vec{x}) = \prod_{i=1}^{7} P_{i,B}(x_i) \tag{4.5}$$

$$\mathcal{L}_e(\vec{x}) = \frac{P_S(\vec{x})}{P_S(\vec{x}) + P_B(\vec{x})} \tag{4.6}$$

The energy of the electrons taken directly is found to underestimate the expected Z-peak in Drell-Yan distributions. The energy scale and offset is then determined to best match the known Z boson peak [19].

In this analysis, we use electrons that are defined as *Top Tight*, which includes the following:

- likelihood > 0.85,
- $f_{iso} < 0.15$,
- $f_{EM} > 0.9$,
- H-matrix $\chi^2 < 50$,
- track $p_T > 5$,
- Calorimeter energy to track momentum ratio < 2.5.

The specific analysis cuts can be found in Section 7.1.

4.2.2 Photons

The selection of photons is similar to the electron selection. The track-match probability can be reversed to discriminate between electrons and photons. The tighter the cut put on the reverse track-match probability, the stronger the discriminating power of the variable. The isolation, electromagnetic fraction, and the H-matrix are also used to differentiate between the photon and jets, just as they are for the electron. Photon identification is discussed in [62].

The photon also uses several other variables. The IsoHC4 finds the scalar sum of tracks' momenta in a hollow cone between $0.05 < \mathcal{R} < 0.4$, with the direction measured in EM3. Only tracks with $p_T > 0.4$ GeV and $|z_{vtx} - z_{trk}^{dca}| < 1.0$ cm are included. Additionally, the preshower is used to determine the difference in shower shapes between a photon produced in direct production and the decay of a neutral pion. The cps_{rms} finds the spatial differences between the energy deposits in the preshower and the those in EM3. This discriminates against clusters with wide energy

deposits. A second CPS variable is the $cps_{rms_{sq}}$ which takes the square of the ϕ and measures the difference in the energy squared distributions. This discriminates against multiple peaks in the preshower as would be expected in the diphoton decay of the neutral pion.

The photon energy also needs to be slightly modified from that expected from electrons due to a slightly different longitudinal energy deposition in the calorimeter. This correction varies from 2% in photons of 30 GeV to 0.1% in 150 GeV photons.

4.3 Muons

Muon objects at DØ are reconstructed based on information in three subdetector systems: the muon system, the calorimeter, and the central tracking system. All other known standard model objects either shower in the calorimeter or leave no trace in the detector at all. Only muons survive the calorimeter to deposit energy in the outer muon tracking system. The main problem with reconstructing events with muons is then less involved with object identification and more with the provenance of the muon and the quality of the associated properties.

Each muon is defined by three types of object definitions described in [22]. The first regards the quality of the muon based on information from the local muon system. There are four definitions. In decreasing order of quality, these are:

1. Tight
2. Medium Nseg 3
3. Medium
4. Loose.

This analysis uses $Medium\ Nseg\ 3$. The Nseg = 3 refers to the three segments of the muon system that must contain hits for the local tracks. The A layer is located

within the muon toroidal magnet while the B and C layers lay outside of the toriod. For a muon to be referred to as Nseg3, there must be hits on either side of the toroid. For a $Medium\ Nseg$ 3, there must be at least two hits in the A layer drift tube wires, at least 1 hit in an A-layer scintillator, at least 2 hits in the B or C layer drift tubes, and at least one hit in the B or C layer scintillators (unless there are four hits in the BC-layer drift tubes).

The second object definition is based on the quality of the track in the central tracking detector. This analysis uses the tight track definition, but there are also medium and loose options. For tight tracks, the distance of closest approach to the vertex of the matched track must be < 0.02 cm. Additionally, the track must satisfy a $\chi^2/dof < 4$. Also, there must be at least one hit in the SMT detector.

The final parameter is the isolation of the muon. The physics of most interest in leptonic final states arise from muons that are the result of object decays from heavy bosons. These objects produce isolated muons, while muons coming from the decay of heavy-flavor jets are produced within the cone of the decaying jet. To determine if the muon is sufficiently isolated, several isolation definitions are given. This analysis uses an isolation definition known as $NPTight$.

For a muon to be $NPTight$, it must satisfy conditions for the track halo and the calorimeter halo. The track halo is defined as $|\sum^{\text{tracks}} p_T|$ in a $\Delta R(\text{track}, \text{muon track}) < 0.5$ cone. For the calorimeter halo, $|\sum^{\text{cells}} E_T|$, the calorimeter energy is calculated in a cone of $0.1 < \Delta R < 0.4$. In $NPTight$, the track halo must be less than 2.5 GeV, and the calorimeter halo must be less than 2.5 GeV as well. The cone used for muon isolation is shown in Figure 4.5.

There are additional cuts to reject cosmic muons by restricting the time for the muon to propagate out to the A, B, C layers to 10 ns.

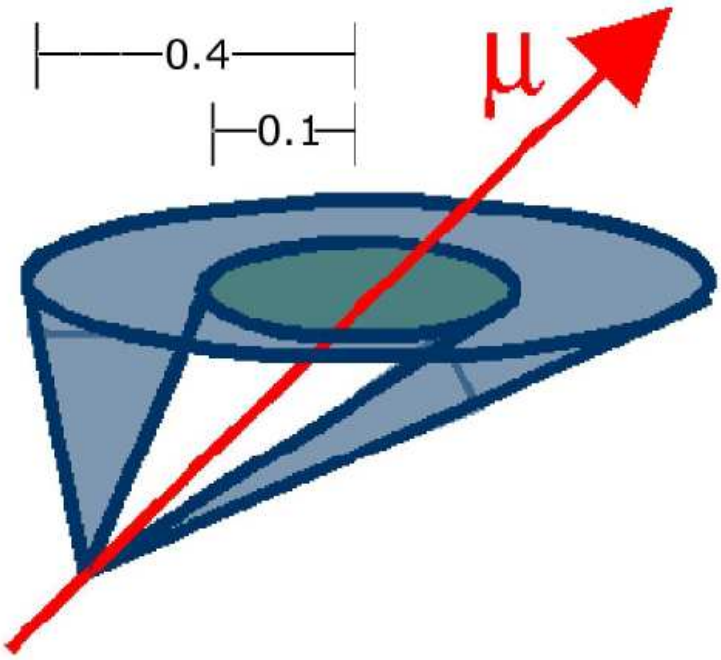


Figure 4.5: The muon isolation cone for calorimeter isolation is a hollow cone of $0.1 < \mathcal{R} < 0.4$ [20].

4.4 Hadronic Objects

Hadronic objects produce hadrons and have object properties that are primarily determined within the calorimeter. The two detector objects of this type that are separately defined in DØ analyses are hadronic taus and jets.

4.4.1 Taus

Taus can decay to electrons, muons, and hadrons. It is difficult to determine whether light leptons in an event final state originated from a tau, but the hadronic signature from a tau differs fairly significantly from that of a jet. Taus' calorimeter quantities are determined from two algorithms, the $Cal\ Cluster$ and $EM\ Sub-Cluster$. The $Cal\ Cluster$ uses a simple cone of $R = 0.3$, and an isolation cone of $R = 0.5$. The $EM\ Sub-Cluster$ is used to find π_0's. This uses a nearest neighbor algorithm in

the third EM layer. If any clusters are found, cells in other layers are combined with preshower information to determine the properties of the tau object. Tau reconstruction and identification is discussed in detail in [63].

Next, the calorimeter clusters are matched to tracks in the central tracking system. A tau typically produces three tracks or less. The best way to suppress a track from jets is to ensure that the tau has no more than three tracks associated with it. The track matching procedure is described below.

1. All tracks within $|R = 0.5|$ cone are sorted in p_T
2. If a track has $p_T > 1.5$ GeV, then an attempt is made to match it to a calorimeter cluster
3. Up to two more tracks may be added if they are within 2 cm in the z-direction from the base track
4. A second associated track may be added if the invariant mass of the two tracks is < 1.1 GeV.
5. A third track may be added if the invariant mass of the three tracks is < 1.7 GeV (the mass of the τ) and the total charge of the tracks is one or negative one.

Three types of hadronic tau decays define the three tau types at DØ :

1. $\tau^{\pm} \rightarrow \pi^{\pm}\nu_{\tau}$
2. $\tau^{\pm} \rightarrow \rho^{\pm}\nu_{\tau} \rightarrow \pi^{0}\pi^{\pm}\nu_{\tau}$
3. $\tau^{\pm} \rightarrow \pi^{\pm}\pi^{\mp}\pi^{\pm}$

Each of these taus with their unique decay properties are found in the detector with different algorithms. The three types of tau decays are shown in Figure 4.6.

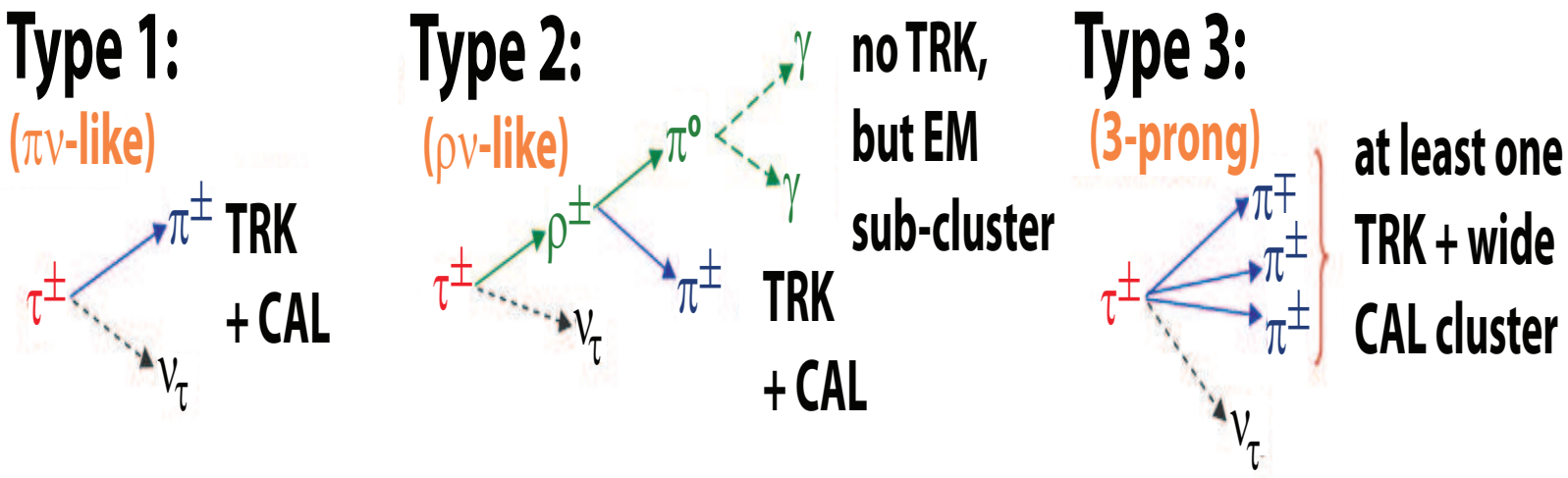


Figure 4.6: Hadronic tau objects at DØ are defined by three types of decays. This analysis identifies taus that undergo any of these decays.

A type 1 tau is based on an algorithm searching for the first listed decay type. There are two properties used to distinguish the tau from the jet in this algorithm. The first is the ΔR between the track and the calorimeter cluster. The second is the ratio $\frac{E_T}{p_T}$. For a type 1 tau, the energy deposited in the calorimeter should be equal to the momentum of the associated track since there is only one detectable object associated with the tau.

A type 2 tau uses the ΔR between the track and the hadronic part of the cluster, the em3iso $= \frac{E_T(EM3)^{cluster}}{E_T(EM3)^{total}}$. The final parameter is the mass of the track combined with the EM3 cluster. The EM3 cluster should correspond to the π^0 and the track to the charged pion. These two should never have a mass larger than the known 1.7 GeV mass of the tau.

Type 3 taus use the ΔR between the second track and the calorimeter cluster. Also the number of tracks between 10° and 30° is checked to ensure there are not other associated tracks expected in a jet. Also the energy deposited in the calorimeter is compared to the sum of track momenta. Each of the tracks should be associated with the calorimeter deposit with no additional energy from the τ in the calorimeter.

Each of the objects then uses a neural network to cut on several variables associated with each tau type. A neural network is a multivariate process that takes a vector of n inputs and maps them onto m outputs without knowing the functional form of the mapping. The neural network works by repeatedly presenting inputs associated with certain outputs. In high energy physics, training signal and background samples are presented to the network. With this sample the network is able to learn how known inputs are mapped to outputs, and the neural network can be trained to maximize its discrimination between signal and background. The trained neural network can then be used on a sample with unknown signal and background content. More information can be found in [64].

Type 1 Neural Net [63]:

1. $tauprf \rightarrow \frac{E_T^1+E_T^2}{\sum^i E_T^i}$ This variable compares the energy in the two highest towers to the overall energy deposited. Since tau objects have narrower signatures, the fraction of energy deposited in the highest towers for taus should be higher than for jets.

2. $tauiso \rightarrow \frac{E(0.5)-E(0.3)}{E(0.3)}$ This determines the tau energy isolation. It measures the energy deposited in a circle around the centroid. This again uses the fact that taus have narrow signatures. Most of the tau energy should be deposited within $R = 0.3$ in η, ϕ. Additional energy outside of this region points to a more jet-like object.

3. $tauEM12isofr$ The ratio of the transverse energy in the first two layers of the calorimeter to the total transverse energy in a cone of $\mathcal{R} < 0.5$ centered at the centroid of the calorimeter deposition.

4. $tauett1/taupt$ This is just the ratio of the calorimeter energy deposition to the momentum of the track.

5. $tauettr/(tauettr + tauett1 + tauett2 + tauett3)$ The total transverse momentum of all of the tracks in a cone of $\mathcal{R} < 0.5$ that are not associated with the tau divided by the total transverse momentum of all tracks.

Type 2 and 3 Neural Net:

1. $tauprf$ Same as above.

2. $tauiso$ Same as above.

3. $tauett1/tauEt_iso$ The transverse momentum of the leading track as a fraction of the calorimeter energy in $\mathcal{R} < 0.5$.

4. $taue1e2/taupt$ The square root of the product of the initial track p_T and the E_T of the electromagnetic cluster.

5. $taudalpha/pi \rightarrow$ The opening angle between highest p_T track and corresponding EM cluster divided by the $\sin\theta$ where $\sin\theta$ is the sine of the azimuthal angle of the calorimeter cluster centroid.

6. $tauettr/(tauettr + tauett1 + tauett2 + tauett3)$ Same as above.

The type two taus are very close to the signature of an electron with a single track and energy deposition in the EM calorimeter. To remove these, the H-matrix χ^2 (see Section 4.2.1) is required to be less than 30 within an $\mathcal{R}$ of 0.4. Additionally, the $f_{EM} < 0.8$, $\frac{E}{p} - 1 < 1$, and $|\Delta\phi(MET, \tau_{trk})| < 0.5$ to account for the energy deposition of the charged pion in the hadronic calorimeter and the missing energy of the tau neutrino that will be associated with the decay.

4.4.2 Jets

Discrimination among decay signatures of different quark flavors and between gluons and quarks is quite difficult. Heavy-flavor jets of b-quarks and to a lesser extent c-quarks have hadronization signatures that allow some minimal discrimination. With the exception of bottom quarks, all quarks and gluons will be considered identical. Predictions of decay properties from theory are determined from cross sections using perturbative QCD. The predictions involve partonic objects that have yet to hadronize and decay. Some Monte Carlo programs incorporate showering and fragmentation models to predict showers of hadrons known as particle jets. At DØ , measurable properties of jets correspond to energy depositions in the calorimeter. These define the three types of jets necessary to bridge fundamental theories to particle signatures: partonic objects, particle jets, and detector jets as seen in Figure 4.7.

4.4.2.1 Jet Reconstruction

As discussed in Section 3.2.2.1, the calorimeter is divided into cells of η and ϕ. The cells are put together into clusters that extend out roughly along rays from the center

of the detector. A group of these cells is a geometrical tower. In jet reconstruction algorithms, each cell is treated kinematically as a massless object with its own energy and momentum determined by the energy deposition and its projection from the detector center. The full jet reconstruction procedure is outlined in [65].

The first step in reconstructing jets is to create energy towers out of the geometrical towers. The energy in each calorimeter cell is measured and compared to the width of the energy signal due to noise. If the energy is 2.5 σ^{cell}, the cell's energy can be added to the overall tower energy. Generally, noise from electronics and radioactivity can cause isolated cells with high energy. If the cells are sufficiently isolated they may be removed with the $NADA$ algorithm [66]. Additionally, the $T42$ algorithm removes any cell with an energy $< 4\ \sigma^{cell}$ that does not have a neighbor with energy $> 4\ \sigma^{cell}$. The details of the $T42$ algorithm can be found in [67]. The cells that survive these noise cuts become part of the final reconstructed tower. For each tower, one then calculates full 4-momentum values, from which its energy, momentum, and directional properties can be calculated. These towers are then fed into the *Simple Cone Algorithm*.

The *Simple Cone Algorithm* takes individual energy towers and loops over them creating preclusters. Any tower with with a transverse momentum of > 0.5 GeV will be checked as a seed for precluster construction. If the lead (highest energy) cell comes from a cell in the coarse hadronic calorimeter or is part of the end cap massless gap, an additional condition must be met. Since these cells are typically noisy, the total tower p_T minus the lead cell p_T still must be > 0.5 GeV.

Once the precluster seed has been determined, then all of the remaining towers are looped over. If $\Delta\mathcal{R} < 0.3$ between the tower and the precluster, and the p_T of the tower is > 1 MeV, then the tower is combined with the precluster using the full 4-momentum. In the creation of preclusters, the pseudorapidity is used in $\mathcal{R}$. In later stages, the actual rapidity is used in $\mathcal{R}$ calculations. Any precluster with $p_T > 1$

GeV and with more than one tower is used in the jet creation algorithm.

Several jet reconstruction algorithms have been proposed, but this analysis and nearly all others have relied on one algorithm called *Run II Cone Algorithm* [65]. This algorithm follows three steps to produce the final jets that are used in analyses: clustering, finding midpoint protojets, and merging/splitting.

4.4.2.1.1 Clustering The clustering method takes each of the preclusters as seeds to form protojets and the list of all of the towers to recalculate energy values when the preclusters are turned into protojets. First, the preclusters are ordered in p_T. The distance $\Delta\mathcal{R}$ between the precluster and any already created protojet must be > 0.25. The protojet candidate then goes through an iterative process to find its final configuration as a protojet. A cone of $\mathcal{R} = 0.5$ is created and all cells within the cone are combined to form a new protojet candidate. The values associated within the new protojet are recalculated. A new cone is formed and the process continues. The iteration stops when any of the following conditions are met.

1. The transverse momentum of the protojet candidate is < 3 GeV, in which case the candidate is discarded.

2. The cone stabilizes with $\Delta\mathcal{R}$ between successive iterations of the protojet candidates < 0.001. This protojet passes to the next stage.

3. The number of iterations reaches 50. Again, this protojet will be passed to the next stage of reconstruction.

4.4.2.1.2 Finding Midpoint Protojets This step is new for Run II. It was found that calculations from perturbative QCD with infrared and collinear cut-offs were unstable without the inclusion of midpoint protojets. In this step, pairs of protojets with $0.5 < \Delta\mathcal{R} < 1.0$ are considered. The midpoints in p_T of these jets are found and the same clustering algorithm as above is run with two differences. First,

there is no condition put on the minimum $\Delta\mathcal{R}$ between the precluster and another protojet. Second, there is no removal of duplicated jets.

The midpoint protojets are then added to the list of the protojets created directly from preclusters as discussed above.

4.4.2.1.3 Merging and Splitting The energy depositions in the calorimeter must only be used once. To ensure this, the list of protojets is checked for overlaps. All protojets are looped over and checked for overlaps with other protojets. If a protojet shares a calorimeter cell with another jet, then the fraction of energy of the lower p_T jet that is shared with the neighboring protojet is calculated. If over half of the energy is shared, then the two protojets are merged into one. If under half, the protojets split the energy by their distance from the cell $\Delta\mathcal{R}$, where $\mathcal{R}$ is now using the true rapidity rather than the pseudorapidity. In either case, the new jet or jets are added to the protojet list, and the process is begun again. This is repeated until there are no protojets that share calorimeter cells. A final cut of $p_T > 6$ GeV is put on the protojets, and the resulting list consists of the jets that are used in analyses.

4.4.2.2 Jet Energy Scale

Unlike other detector objects, jets are considerably removed in their properties from the initial decay objects. To better understand the physics processes that led to a particular detector signature, the detector jets are corrected to determine the energy and direction of the original object that entered the detector. When this information is determined, the detector signature can be better compared to the original physics process of interest. The determination of the jet energy scale for p17 data can be found in [68]. The transition from partons to detector jets is shown in Figure 4.7.

The particle jet energy can be related to the measured jet energy by the following equation,

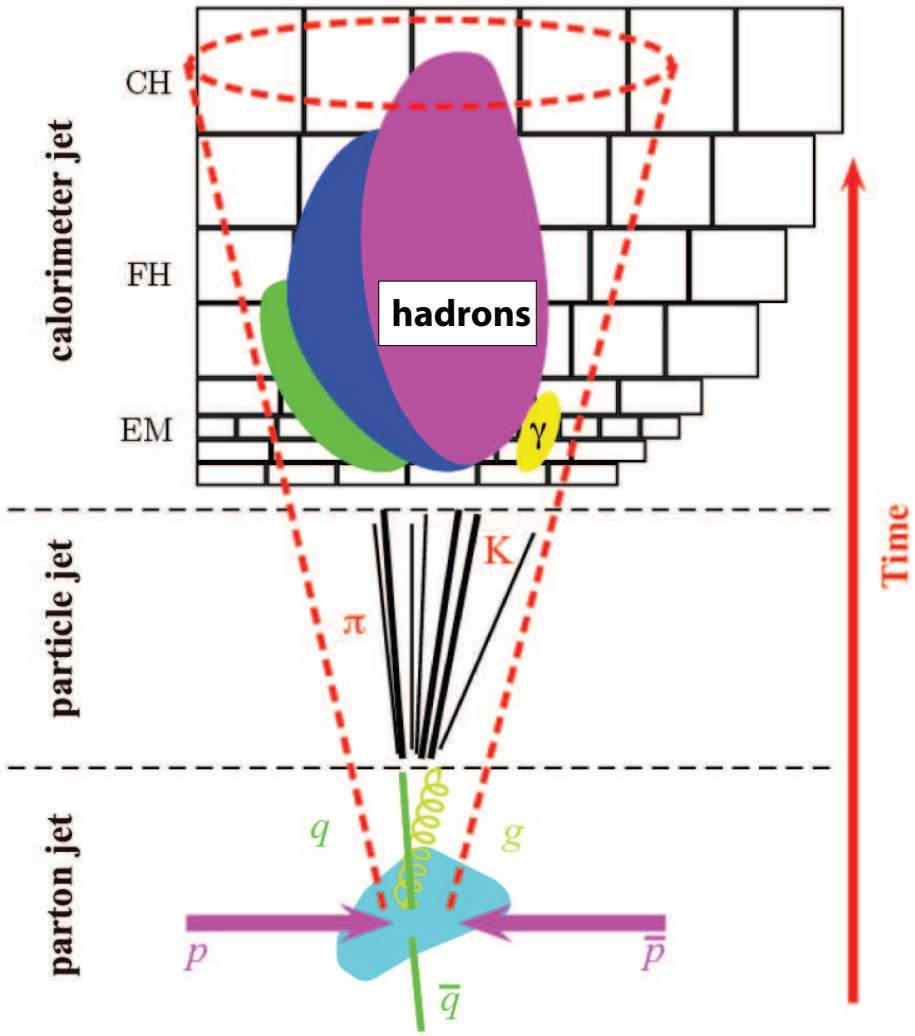


Figure 4.7: The partons of the initial physics processes decay and hadronize to particle jets which then leave tracks in the inner tracking system and energy in the calorimeters.

$$E_{\text{jet}}^{\text{particle}} = \frac{E_{\text{jet}}^{\text{measured}} - E_O}{R_{\text{jet}} S_{\text{jet}}}. \tag{4.7}$$

The rest of the section will describe each of these variables.

E_O is the offset energy. This energy can be due to electronic noise signals and the radioactive decay of the uranium absorber plates. Additionally, there can be more than one $p\overline{p}$ interaction in an event and also energy left over from previous beam crossings.

The amount of offset energy will depend on several variables:

- The size of the cone used to create the jet, because a larger cone size is more likely to include depositions from unrelated processes.
- The pseudorapidity of the jet because the granularity changes in η, and the response of different parts of the calorimeter are nonuniform.
- The number of vertices changes the probability that there will be additional collisions for a particular beam crossing.
- The instantaneous luminosity also will determine the likelihood of multiple collisions and the amount of energy that may left over from a previous crossing.

The variable R_{jet} is the fractional response in the calorimeter to a particle jet with a particular energy. This value tends to be less than one because of energy lost in the calorimeter, the uninstrumented regions between detector modules, the lower response the calorimeter has to hadrons compared to electrons and photons, and inhomogenieties among the modules. The calorimeter response will depend on the energy of the jet, the cone size used, and the pseudorapidity.

The last variable is S_{jet}. This determines the fraction of energy deposited within the cone defined for the jet. This value would be less than one from parts of the jet energy that falls outside of the defined cone. Additionally, energy from other jets that falls into the jet cone, may push the S value above one.

A difficulty with this method is that the true values cannot be determined from data. The offset energy will depend on the fraction of energy within cone, and the response will also depend on the other factors. Monte Carlo can be used to get an idea what additional correction factors need to be included to properly translate the energy of a jet back to its particle state.

The offset energy is determined by special triggers that collect events that do not have a hard scatter process. The calorimeter response is determined from $\gamma + jet$ events. The photon response is better measured and calibrated, so the jet energy measurements can be determined by the photon energy depositions. The in-cone fraction uses $\gamma + jet$ events in both data and Monte Carlo. When these events are back-to-back, the amount of energy that falls within the cone can be determined.

Since the values of R and E_O are determined directly from data, they contain biases. Additional multiplicative factors, k_O and k_R can be determined from Monte Carlo to correct the estimated factors in data to the simulated values determined from the Monte Carlo.

Further discussion of jet corrections (shifting, smearing, and removal) for Monte Carlo are discussed in 6.2.1.

4.4.2.3 Jets from Bottom Quarks

Many new physics processes preferentially decay to heavy quarks. These events can be distinguished from light jet events because bottom quarks form B-hadrons with lifetimes long enough to provide a identifiable vertex distinct from the main vertex in the event. The heavy quark jets can be identified from the secondary vertex and the association of calorimeter jets to "track jets", which are groups of tracks in the tracking detectors. Track jets are defined as tracks within $\mathcal{R} < 0.5$ of a seed track of $p_T > 1.0$ GeV with at least 2 SMT hits. The track jet must have at least two tracks with combined $p_T > 1.0$ GeV. Several different algorithms have been used for b-jet

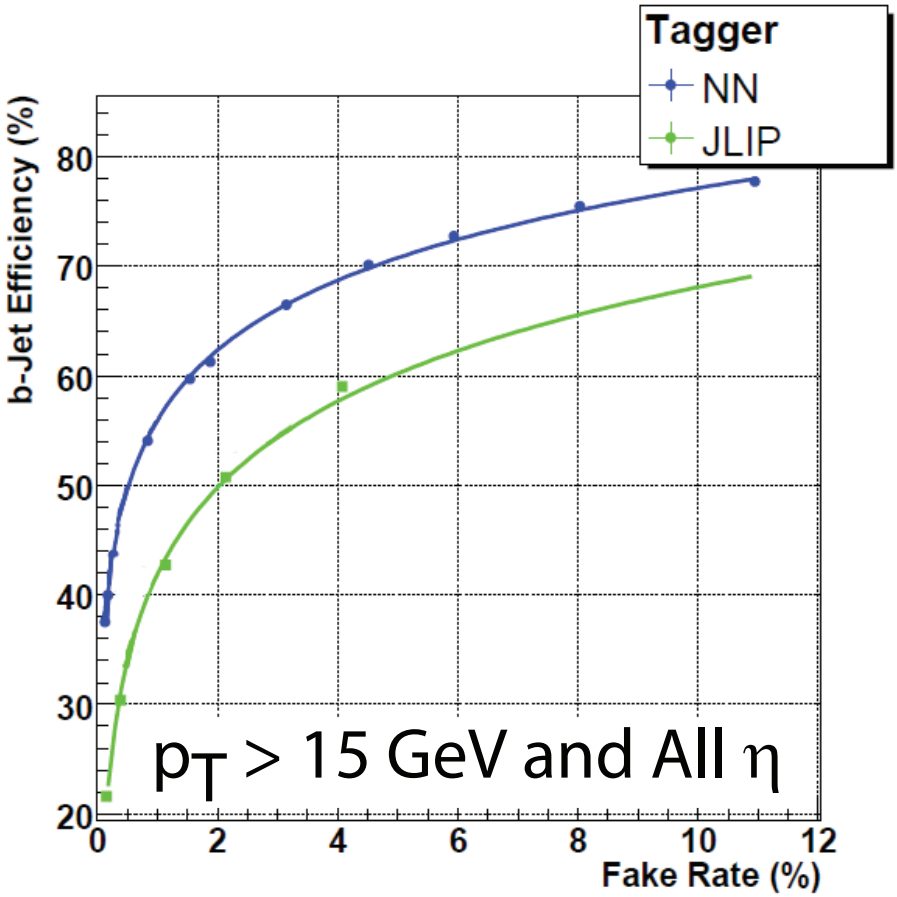


Figure 4.8: Efficiency versus fake rate for various operating using the neural net and jet lifetime probability tagger.

identification. Three of these, the Jet Lifetime Probability (JLIP), Counting Signed Impact Parameter (CSIP), and the Secondary Vertex Tagger (SVT) have been used in a neural network improving the discrimination between b-jets and light jets better than any of the individual algorithms. The description of this process is described in detail in [69]. The performance benefits over the JLIP tagger are shown in Figure 4.8.

The seven variables used in the neural network are:

1. SVT_{SL} DLS: Decay length significance of the secondary vertex.

2. $CSIP_{Comb}$: A weighted combination of impact parameter significance of tracks

associated with jet.

3. $JLIP\ Prob$: Probability that a jet originated from the primary vertex. The closer to zero, the more likely a b-jet. If there is not enough information to make this determination, the variable is set to one.

4. $SVT_{SL}\ \chi^2_{dof}$: Chi-squared per degree of freedom for the secondary vertex.

5. $SVT_L\ N_{trks}$: Number of tracks used to reconstruct the secondary vertex.

6. $SVT_{SL}\ Mass$: Mass of the secondary vertex. This is the combined rest mass of the tracks, assuming they are all pions.

7. $SVT_{SL}\ Num$: Number of secondary vertices found in the event.

The subscript L in SVT refers to the Loose operating point for that algorithm. The SL refers to super loose which is an operating point which was not used when SVT was used as an individual tagger. For the neural network, more information provides greater discriminating power, so the cut was lowered to allow more multijet background to pass the cuts.

The efficiency of the neural net to properly identify a b-quark is divided into two parts. The first is the efficiency for the b-quark to be reconstructed as a matched track jet. If the track jet is identified, the jet is defined as "taggable". The taggability differs between the data and the Monte Carlo, so that a scale factor must be applied to the Monte Carlo to appropriately simulate the data events. The b-tagging efficiency is based solely on the ability of the taggable jet to pass a certain b-tagging operating point of the neural network. This is discussed in more detail in Sections 6.2.2 and 6.2.3.

4.5 Missing Transverse Energy ($\not\!\!E_T$)

The initial longitudinal momentum of an event is not known a priori due to the fact that the colliding partons are part of the larger hadronic particles being accelerated, and the distribution of that energy among the partons is not well known. The transverse energy of the system, however, should be approximately balanced. If an event shows a large amount of transverse energy, it can be inferred that a non-interacting particle passed through the detector. In the standard model, only the three neutrinos are non-interacting. Many theories of new physics include massive non-interacting particles that would show up in the detector as large amounts of missing transverse energy, well beyond that expected from W boson or Z boson decays.

The missing transverse energy is calculated based on the common T42 algorithm, as are all energy calculations of calorimeter objects. This means that the energy for any calorimeter reconstructed object only uses the subset of calorimeter cells that have positive energy more than 4σ from the width of their noise, and neighboring cells in all three spatial dimensions that have energy levels greater than 2σ.

The missing energy incorporates all of these cells into its energy calculations except for those of the coarse hadronic calorimeter. The cells from the CH are only used if they are included in the reconstruction of the jet. Additionally, muons deposit only a small fraction of their energy in the calorimeter. The calorimeter-based $\not\!\!E_T$ must be adjusted to account for muon energy determined from the muon momentum. Furthermore, many objects undergo further corrections to relate them to the energies of the initial particles before they enter the detector. Therefore, for consistency, these corrections must also be applied to the $\not\!\!E_T$. These include corrections for the jet, electron, and tau energy scales.

Chapter 5

Data and Monte Carlo Samples

5.1 Data Sample

This thesis is based on the DØ Run IIa data set which ran from 2002 to 2006. Each data-taking session at DØ is separated into several hour runs with approximately the same running conditions. The run numbers associated with this analysis begin with run 151817 and finish with run 219000. The full analysis code used at DØ is also categorized by versions, so that the full analysis structure can be understood based on a single version number. The data from RunIIa used in this analysis was processed with version p17.09.03 and converted to the standard root-tree-based analysis format (CAF trees) using version p18.05.00. After the data is converted into the CAF tree analysis format, it is divided into skims based on the object content of the event. This analysis uses two skims: MUinclusive and EMinclusive which are defined by an event that passes certain basic object criteria. For a single muon, this is typically a loose object definition and a p_T cut of 8 GeV. For a single em object (electron or photon), a loose object definition is also used but with a p_T cut of 20 GeV. Each of the skims also have options of looser object and p_T requirements if the event contains other objects. We also apply an additional condition on the data from the EMinclusive skim so that it is not also included in the MUinclusive. This prevents the possibility

of double counting events. The original data reconstruction, root tree production, and skimming was performed by the DØ Common Analysis Format Group [70].

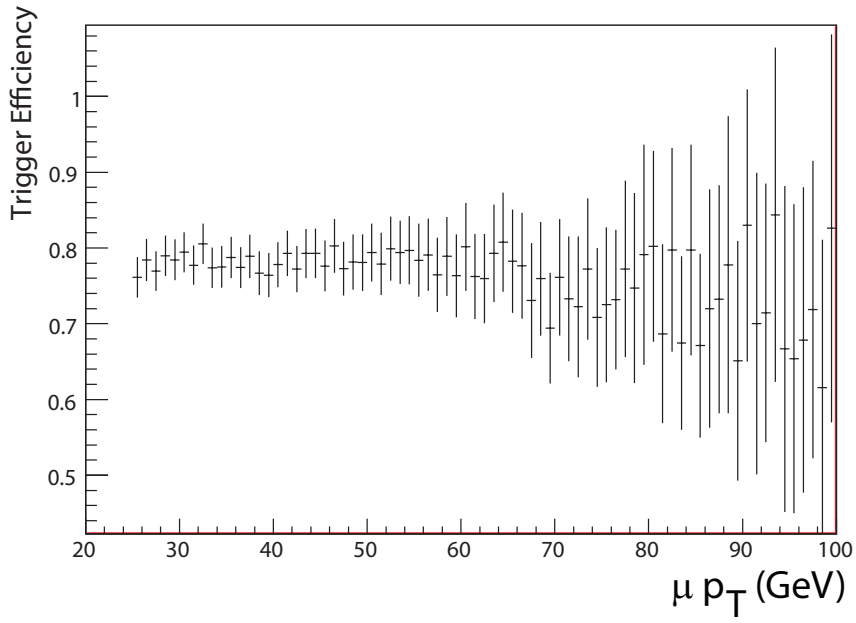
Once the data are collected, they are checked by subdectector groups to ensure data quality. Each group will mark particular runs or luminosity blocks as bad if the detector components were unable to provide reliable information during that time. All runs and luminosity blocks that are marked as bad are removed from the run. The total luminosity used in this analysis after data quality checks is 1.1 fb^{-1}.

In our analysis, we focus on events that contain objects with high values of transverse momentum. In order to quickly incorporate changes to our analysis framework, we perform a subskim on all data and Monte Carlo choosing only those with high-p_T objects, and saving them in a condensed format. The details of this format and skim can be found in Appendix B.

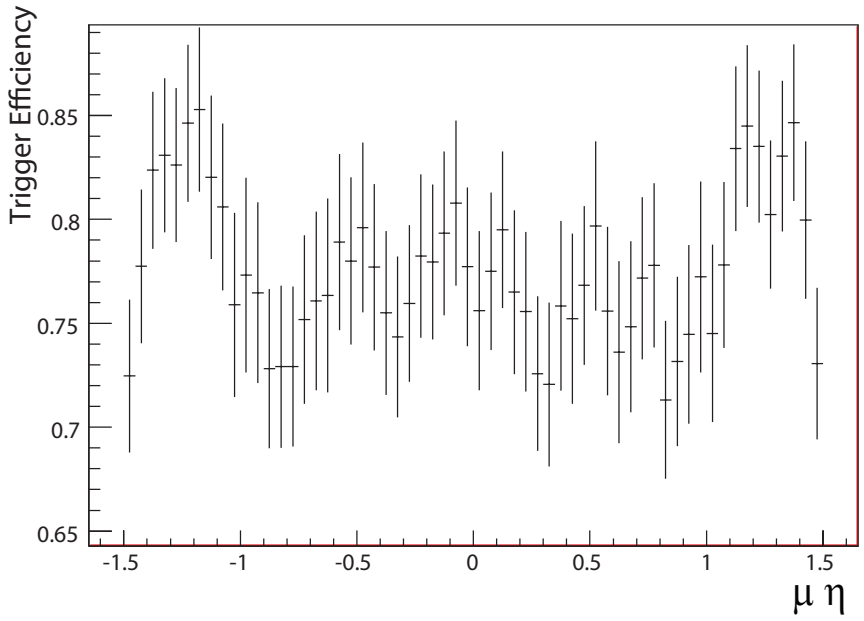
The events entering the high-p_T data set have no specific trigger requirements. We choose events at the MIS level beyond the trigger turn-on threshold, so that there is no specific momentum or energy dependence in the trigger efficiency. The p_T dependence of each of the plots are shown in Figures 5.1(a), 5.2(a). The η dependence is shown in Figures 5.1(b), 5.2(b). These plateaued efficiencies allow the incorporation of the trigger efficiencies in the normalization fits performed at the MIS level as discussed in Chapter 7. The exception is μ η dependence, which along with a jet multiplicity dependence led us to include trigger efficiencies directly in later versions of the analysis.

5.2 Monte Carlo Samples

The primary way we represent our standard model background is with Monte Carlo generators. With a model-independent search we must incorporate many different background processes to properly model the data. We primarily use two generators for this purpose, ALPGEN for producing processes where we need to accurately incor-

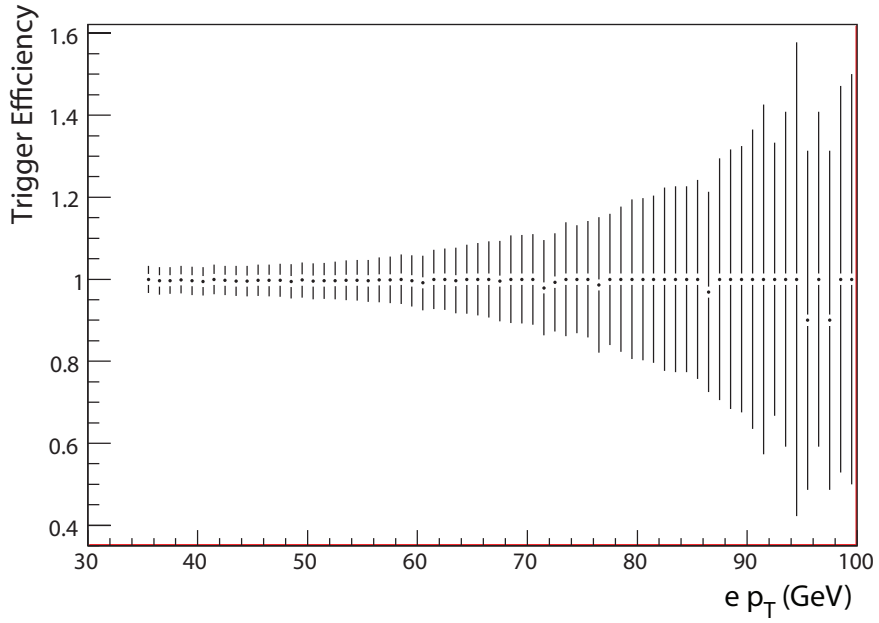


(a)

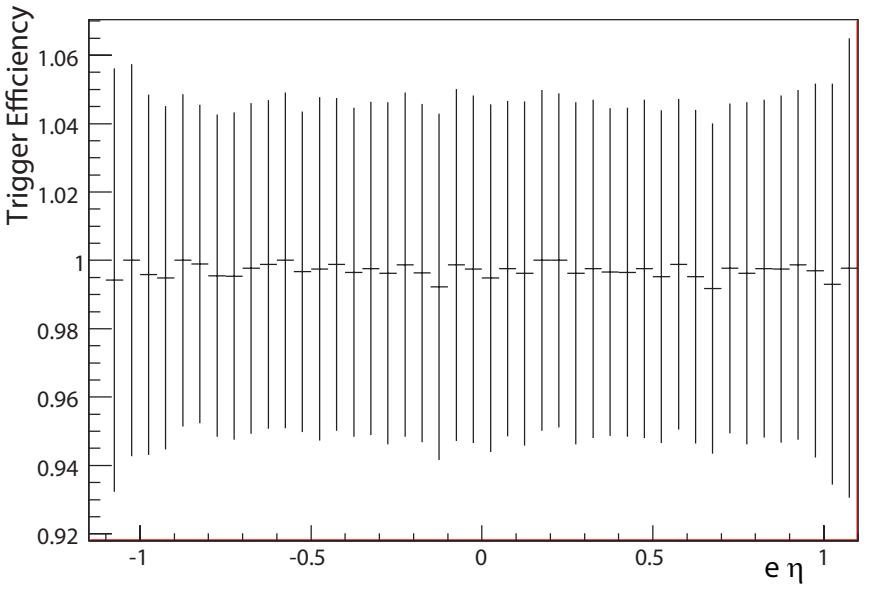


(b)

Figure 5.1: Trigger efficiencies for single muon triggers in single μ final state. The μ η distribution is shown to not be completely flat. This, along with a multiplicity dependence led us to incorporate trigger efficiencies directly in later analysis runs.



(a)



(b)

Figure 5.2: Trigger efficiencies for single electron triggers in single e final states.

porate jets produced in the hard scatter, and PYTHIA where these are less important and our focus is on accurate hadronization and showering. When using ALPGEN, we match the jets produced in the hard-scatter to PYTHIA for appropriate hadronization and showering.

5.2.1 Monte Carlo Production Process

For some objects, independent programs provide more accurate simulation of particle processes and decays. Specifically, TAUOLA is used for τ decays [71], and EVTGEN is used for the decay of b hadrons [72]. More details on these programs can be found in the references.

5.2.1.1 PYTHIA

PYTHIA includes hard and soft interactions, parton distributions, initial and final state parton showers, multiple interactions, fragmentation and decay [73]. Because of its completeness, it is the first choice to model the $p\overline{p}$ collisions at the Fermilab Tevatron Collider. While much of the particle collision is modeled in PYTHIA, it does not incorporate the parton content of the proton and antiproton. For these measurements, a parton distribution function determined by the CTEQ collaboration is used [74]. For all of the processes used in this analysis, the CTEQ6L1 library is used. These are the leading order parton distributions specifically designed for Monte Carlo event generators. It uses leading order hard cross sections with NLO α_s values which are appropriate for the event generation methods used in PYTHIA and ALPGEN.

5.2.1.2 ALPGEN

ALPGEN calculates exact matrix elements at leading orders for QCD and electroweak interactions [75]. The benefit of using ALPGEN comes from its exact leading order calculations for processes that include high jet multiplicities. ALPGEN produces

parton-level events with information on color and flavor, so that it may be matched to PYTHIA for showering and hadronization.

Matching of partons from the ALPGEN calculation to PYTHIA showering has the fundamental difficulty of trying to separate the hard interaction from the showering process. We use the MLM matching scheme which showers all events and looks for an appropriate $\Delta\mathcal{R}$ match between the partons and the jets, rejecting those events without a match and also those with an additional unmatched jet [76]. This eliminates double-counting from collinear partons with overlapping matrix elements and partons that are too soft to produce their own jet. PYTHIA controls the partonic showering and hadronization.

Additionally, final states with high jet multiplicities and heavy flavors (b-quarks and c-quarks) are particularly interesting in the search for new physics. Since the MIS analysis considers events with b-tagged jets in separate final states from those with light quark jets, it is particularly important to ensure that the standard model background has enough events to accurately reflect the expected background in these rare processes. To ensure this, processes that include additional heavy-flavor quark pairs are produced separately from the light flavor contributions. In order to avoid double counting the events, the light quark jets need to be scanned for b quark and c quark content, and the b quark events need to be searched for additional c quarks [77]. These checks are required because of overlaps in phase space. Two charmed quarks may be produced in tree level diagrams and during the PYTHIA parton showering. When these events are removed, then all of the processes can be combined into a single input with appropriate weights without fear of double-counting. The following is summary of the skimming performed on each ALPGEN input process that includes the production of additional radiative jets.

- $X + N$ light partons- Events with (parton level) $c\overline{c}$ or $b\overline{b}$ added (by PYTHIA) in the initial state radiation (ISR) or final state radiation (FSR) are removed.

- $X + c\overline{c} + N$ light partons- All events are kept.
- $X + b\overline{b} + N$ light partons- Events with $c\overline{c}$ in ISR or FSR are removed.

The events produced from this combination of generators are then passed through the DØ detector simulation and combined with minimum bias events taken from actual data. The detector simulation, d0gstar, is based on GEANT 3.2.1 [78]. The d0sim program tracks particles through these detector elements, simulates hits on the detector, and then digitizes those hits. The events are then reconstructed and converted to the standard DØ analysis root-based trees (CAF-trees) using the same chain of processors as data.

5.2.1.3 Additional MC Corrections

A handful of additional corrections are made after the production of the CAF trees. These corrections typically involve parameters that may be updated frequently and would require a lengthy resimulation if performed at an earlier stage. These corrections include muon smearing (to correct for better resolution in the simulation), calculation of the process weights, and jet smearing and shifting. More details about these processes can be found in Chapter 6.

5.2.2 Monte Carlo Samples

The following physics background standard model processes are currently considered, where j is a light jet (gluon, u quark, d quark, s quark) and lp is a light parton:

1. $W + Nj$
2. $W + c\overline{c} + Nj$
3. $W + b\overline{b} + Nj$
4. $Z/\gamma^* + Nj$

5. $Z/\gamma^* + c\overline{c} + Nj$

6. $Z/\gamma^* + b\overline{b} + Nj$

7. $t\overline{t} \rightarrow (2\ell + 2\nu + 2b) + Nj$

8. $t\overline{t} \rightarrow (\ell\nu + 2b + 2j) + Nj$

9. WW

10. WZ

11. ZZ

5.2.2.1 W boson + Jets

W boson + jets samples are grouped by the parton content of each event. In order to ensure that there are no overlaps with heavy-flavor samples, the samples are heavy flavor skimmed, as described in Section 5.2.1.2. The W boson + 5 light parton, W boson + $b\overline{b}$ + 3 light partons, and W boson + $c\overline{c}$ + 3 light partons samples are inclusive in jet content. The W boson Monte Carlo uses ALPGEN v2.11. All of these samples use PYTHIA v6_413 for showering and hadronization. A list of W Monte Carlo samples with number of events, leading log cross sections, and effective exposure can be seen in Table 5.1.

The factorization scale used for W boson + jets is set by the interaction energy scale of the process,

$$Q^2 = M_W^2 + \sum_{\text{jets}} p_T^2(j). \tag{5.1}$$

The factorization scale is the scale chosen to divide the hard scatter process calculated by ALPGEN and the showering and hadronization by PYTHIA [76].

Table 5.1: W + jets samples

Monte Carlo Sample	Number of Events	Production Release	σ_{LL}	Effective Exposure fb^{-1}
$W \to \ell\nu + 0lp$ exclusive	2914k	p17.09.07	4520 pb	0.645
$W \to \ell\nu + 1lp$ exclusive	8478k	p17.09.06	1277 pb	6.64
$W \to \ell\nu + 2lp$ exclusive	4964k	p17.09.06	304.8 pb	16.3
$W \to \ell\nu + 3lp$ exclusive	2443k	p17.09.06	72.4 pb	33.7
$W \to \ell\nu + 4lp$ exclusive	1718k	p17.09.06	16.49 pb	104
$W \to \ell\nu + 5lp$ inclusive	521k	p17.09.07	4.95 pb	105
$W \to \ell\nu + c\overline{c} + 0lp$ exclusive	1175k	p17.09.08	23.96 pb	49.0
$W \to \ell\nu + c\overline{c} + 1lp$ exclusive	598k	p17.09.08	13.35 pb	44.8
$W \to \ell\nu + c\overline{c} + 2lp$ exclusive	237k	p17.09.08	5.38 pb	44.1
$W \to \ell\nu + c\overline{c} + 3lp$ inclusive	248k	p17.09.08	2.50 pb	99.2
$W \to \ell\nu + b\overline{b} + 0lp$ exclusive	1041k	p17.09.08	9.34 pb	111
$W \to \ell\nu + b\overline{b} + 1lp$ exclusive	663k	p17.09.08	4.26 pb	156
$W \to \ell\nu + b\overline{b} + 2lp$ exclusive	285k	p17.09.08	1.55 pb	184
$W \to \ell\nu + b\overline{b} + 3lp$ inclusive	349k	p17.09.08	0.74 pb	471

5.2.2.2 Drell-Yan + Jets

The Drell-Yan process is the primary standard model production mechanism for dilepton events in the high-p_T processes considered in the MIS search. These are also grouped by light parton content with $Z/\gamma^* + 3$ light partons, $Z/\gamma^* + c\bar{c} + 2$ light partons and $Z/\gamma^* + b\bar{b} + 2$ light partons inclusive in light parton content. The Drell-Yan samples also use ALPGEN v2.11. A list of Drell-Yan Monte Carlo samples can be seen in Table 5.2. The choice of factorization scale for this production process is shown in the equation below,

$$Q^2 = M_Z^2 + p_T^2(Z). \tag{5.2}$$

Table 5.2: Drell-Yan Monte Carlo samples

Monte Carlo Sample	Generated Z/γ^* mass	Number of Events	σ_{LL}(pb)	Eff. Exp. fb^{-1}
$Z/\gamma^* \to ee + 0lp$ exclusive	15-60 GeV	562k	311	1.81
$Z/\gamma^* \to ee + 1lp$ exclusive	15-60 GeV	427k	35.1	12.2
$Z/\gamma^* \to ee + 2lp$ exclusive	15-60 GeV	164k	8.79	18.7
$Z/\gamma^* \to ee + 3lp$ inclusive	15-60 GeV	78k	2.49	31.3
$Z/\gamma^* \to ee + 0lp$ exclusive	60-130 GeV	1025k	131	93.2
$Z/\gamma^* \to ee + 1lp$ exclusive	60-130 GeV	177k	40.0	4.43
$Z/\gamma^* \to ee + 2lp$ exclusive	60-130 GeV	83k	9.40	8.83
$Z/\gamma^* \to ee + 3lp$ inclusive	60-130 GeV	77k	2.84	27.1
$Z/\gamma^* \to ee + 0lp$ exclusive	130-250 GeV	94k	0.887	106
$Z/\gamma^* \to ee + 1lp$ exclusive	130-250 GeV	84k	0.346	243

Continued on Next Page...

Table 5.2 (cont'd)

Monte Carlo Sample	Generated Z/γ^* mass	Number of Events	σ_{LL}(pb)	Eff. Exp. fb^{-1}
$Z/\gamma^* \to ee + 2lp$ exclusive	130-250 GeV	87k	0.0881	988
$Z/\gamma^* \to ee + 3lp$ inclusive	130-250 GeV	75k	0.0466	1610
$Z/\gamma^* \to ee + 0lp$ exclusive	250-1960 GeV	98k	0.0686	1430
$Z/\gamma^* \to ee + 1lp$ exclusive	250-1960 GeV	88k	0.0349	2520
$Z/\gamma^* \to ee + 2lp$ exclusive	250-1960 GeV	88k	0.0105	8380
$Z/\gamma^* \to ee + 3lp$ inclusive	250-1960 GeV	74k	0.00548	13500
$Z/\gamma^* \to ee + c\overline{c} + 0lp$ exclusive	60-130 GeV	47k	3.05	15.4
$Z/\gamma^* \to ee + c\overline{c} + 1lp$ exclusive	60-130 GeV	43k	1.07	40.2
$Z/\gamma^* \to ee + c\overline{c} + 2lp$ inclusive	60-130 GeV	22k	0.424	51.9
$Z/\gamma^* \to ee + b\overline{b} + 0lp$ exclusive	60-130 GeV	230k	0.965	238
$Z/\gamma^* \to ee + b\overline{b} + 1lp$ exclusive	60-130 GeV	48k	0.350	137
$Z/\gamma^* \to ee + b\overline{b} + 2lp$ inclusive	60-130 GeV	21k	0.132	159
$Z/\gamma^* \to \mu\mu + 0lp$ exclusive	15-60 GeV	552k	309	1.79
$Z/\gamma^* \to \mu\mu + 1lp$ exclusive	15-60 GeV	423k	34.3	12.3
$Z/\gamma^* \to \mu\mu + 2lp$ exclusive	15-60 GeV	163k	8.64	18.9
$Z/\gamma^* \to \mu\mu + 3lp$ inclusive	15-60 GeV	77k	2.52	30.6
$Z/\gamma^* \to \mu\mu + 0lp$ exclusive	60-130 GeV	985k	133	7.41
$Z/\gamma^* \to \mu\mu + 1lp$ exclusive	60-130 GeV	198k	39.6	5.00
$Z/\gamma^* \to \mu\mu + 2lp$ exclusive	60-130 GeV	93k	9.32	9.98
$Z/\gamma^* \to \mu\mu + 3lp$ inclusive	60-130 GeV	86k	2.77	31.0
$Z/\gamma^* \to \mu\mu + 0lp$ exclusive	130-250 GeV	101k	0.885	114

Continued on Next Page...

Table 5.2 (cont'd)

Monte Carlo Sample	Generated Z/γ^* mass	Number of Events	σ_{LL}(pb)	Eff. Exp. fb^{-1}
$Z/\gamma^* \to \mu\mu + 1lp$ exclusive	130-250 GeV	91k	0.345	264
$Z/\gamma^* \to \mu\mu + 2lp$ exclusive	130-250 GeV	86k	0.0885	972
$Z/\gamma^* \to \mu\mu + 3lp$ inclusive	130-250 GeV	73k	0.0455	1600
$Z/\gamma^* \to \mu\mu + 0lp$ exclusive	250-1960 GeV	93k	0.0678	1370
$Z/\gamma^* \to \mu\mu + 1lp$ exclusive	250-1960 GeV	88k	0.0351	2510
$Z/\gamma^* \to \mu\mu + 2lp$ exclusive	250-1960 GeV	82k	0.0105	7810
$Z/\gamma^* \to \mu\mu + 3lp$ inclusive	250-1960 GeV	77k	0.00559	13800
$Z/\gamma^* \to \mu\mu + c\overline{c} + 0lp$ exclusive	60-130 GeV	47k	3.05	15.4
$Z/\gamma^* \to \mu\mu + c\overline{c} + 1lp$ exclusive	60-130 GeV	43k	1.07	40.2
$Z/\gamma^* \to \mu\mu + c\overline{c} + 2lp$ inclusive	60-130 GeV	23k	0.412	55.8
$Z/\gamma^* \to \mu\mu + b\overline{b} + 0lp$ exclusive	60-130 GeV	267k	0.967	276
$Z/\gamma^* \to \mu\mu + b\overline{b} + 1lp$ exclusive	60-130 GeV	48k	0.351	137
$Z/\gamma^* \to \mu\mu + b\overline{b} + 2lp$ inclusive	60-130 GeV	22k	0.132	167
$Z/\gamma^* \to \tau\tau + 0lp$ exclusive	15-60 GeV	535k	310	1.73
$Z/\gamma^* \to \tau\tau + 1lp$ exclusive	15-60 GeV	431k	34.3	12.6
$Z/\gamma^* \to \tau\tau + 2lp$ exclusive	15-60 GeV	167k	8.73	19.1
$Z/\gamma^* \to \tau\tau + 3lp$ inclusive	15-60 GeV	76k	2.48	30.6
$Z/\gamma^* \to \tau\tau + 0lp$ exclusive	60-130 GeV	868k	133	6.52
$Z/\gamma^* \to \tau\tau + 1lp$ exclusive	60-130 GeV	193k	39.7	4.86
$Z/\gamma^* \to \tau\tau + 2lp$ exclusive	60-130 GeV	87k	9.70	8.97
$Z/\gamma^* \to \tau\tau + 3lp$ inclusive	60-130 GeV	78k	2.78	28.1

Continued on Next Page...

Table 5.2 (cont'd)

Monte Carlo Sample	Generated Z/γ^* mass	Number of Events	σ_{LL}(pb)	Eff. Exp. fb^{-1}
$Z/\gamma^* \to \tau\tau + 0lp$ exclusive	130-250 GeV	100k	0.888	113
$Z/\gamma^* \to \tau\tau + 1lp$ exclusive	130-250 GeV	90k	0.352	256
$Z/\gamma^* \to \tau\tau + 2lp$ exclusive	130-250 GeV	80k	0.0915	874
$Z/\gamma^* \to \tau\tau + 3lp$ inclusive	130-250 GeV	71k	0.0451	1570
$Z/\gamma^* \to \tau\tau + 0lp$ exclusive	250-1960 GeV	93k	0.0680	1370
$Z/\gamma^* \to \tau\tau + 1lp$ exclusive	250-1960 GeV	88k	0.0351	2510
$Z/\gamma^* \to \tau\tau + 2lp$ exclusive	250-1960 GeV	82k	0.0104	7880
$Z/\gamma^* \to \tau\tau + 3lp$ inclusive	250-1960 GeV	76k	0.00569	13400
$Z/\gamma^* \to \tau\tau + c\bar{c} + 0lp$ exclusive	60-130 GeV	39k	3.05	12.8
$Z/\gamma^* \to \tau\tau + c\bar{c} + 1lp$ exclusive	60-130 GeV	43k	1.08	39.8
$Z/\gamma^* \to \tau\tau + c\bar{c} + 2lp$ inclusive	60-130 GeV	21k	0.420	50.0
$Z/\gamma^* \to \tau\tau + b\bar{b} + 0lp$ exclusive	60-130 GeV	93k	0.967	96.1
$Z/\gamma^* \to \tau\tau + b\bar{b} + 1lp$ exclusive	60-130 GeV	182k	0.351	519
$Z/\gamma^* \to \tau\tau + b\bar{b} + 2lp$ inclusive	60-130 GeV	87k	0.132	659

5.2.2.3 $t\bar{t}$

The $t\bar{t}$ samples also use the ALPGEN/PYTHIA generation method and are grouped by light parton content and the decay results of the W bosons produced in the top quark decay. A top quark mass of 172 GeV was assumed. All of the $t\bar{t}$ samples were produced with ALPGEN v2.11 and PYTHIA v6_413. The samples are listed in Table 5.3. For $t\bar{t}$, the choice of factorization scales is given by,

$$Q^2 = M^2_{\text{top}} + \sum_{\text{jets}} p^2_T(j). \tag{5.3}$$

Table 5.3: $t\bar{t}$ samples

Monte Carlo Sample	Number of Events	σ_{LL}(pb) * K_{factor}	Eff. Exp. fb^{-1}
$t\bar{t} \to 2b + 4lp$ exclusive	97k	1.91	50.8
$t\bar{t} \to 2b + 5lp$ exclusive	90k	0.792	114
$t\bar{t} \to 2b + 6lp$ inclusive	24k	0.389	61.7
$t\bar{t} \to \ell\nu + 2b + 2lp$ exclusive	518k	1.83	283
$t\bar{t} \to \ell\nu + 2b + 3lp$ exclusive	98k	0.761	129
$t\bar{t} \to \ell\nu + 2b + 4lp$ inclusive	93k	0.374	249
$t\bar{t} \to 2\ell + 2\nu + 2b + 0lp$ exclusive	368k	0.438	840
$t\bar{t} \to 2\ell + 2\nu + 2b + 1lp$ exclusive	236k	0.183	1290
$t\bar{t} \to 2\ell + 2\nu + 2b + 2lp$ inclusive	242k	0.0899	2690

5.2.2.4 Diboson

The dibosons were produced using PYTHIA v6_413. The samples were produced inclusively in parton content and decay and are listed in Table 5.4.

5.3 Multijets Background from Data

The multijets background is determined based on the inclusive final state considered. For $e + jets$ and $e\mu$ the jet background is determined from an electron background sample. For $\mu + jets$, the background comes from non-isolated muons. For $\mu\tau$ and $e\tau$ the loose tau objects are used to provide the appropriate backgrounds. The multijet

Table 5.4: Diboson samples

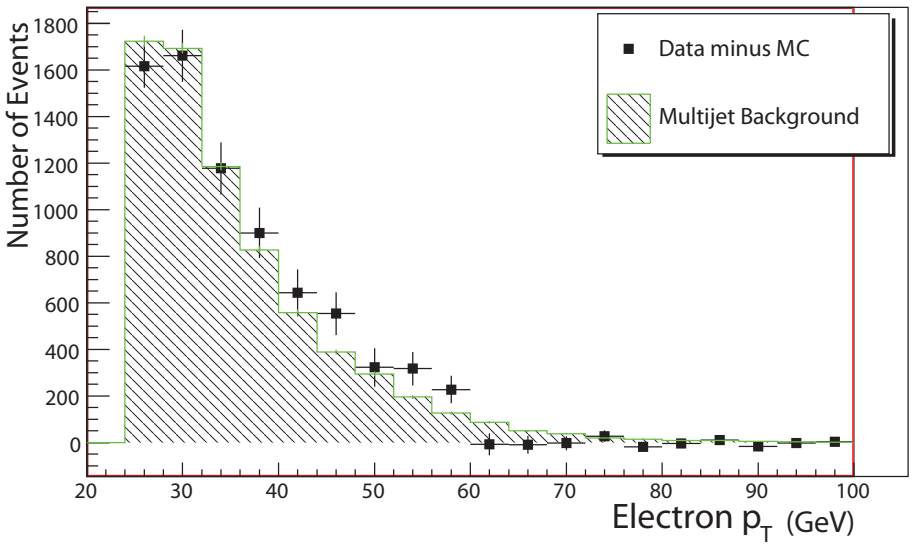
Monte Carlo Sample	Number of Events	σ_{LO}(pb)	Effective Exposure fb^{-1}
WW inclusive	2460k	11.6	212
WZ inclusive	602k	3.25	185
ZZ inclusive	593k	0.425	444

contribution in $\mu\mu$ and ee final states are insignificant, with the given event selection cuts.

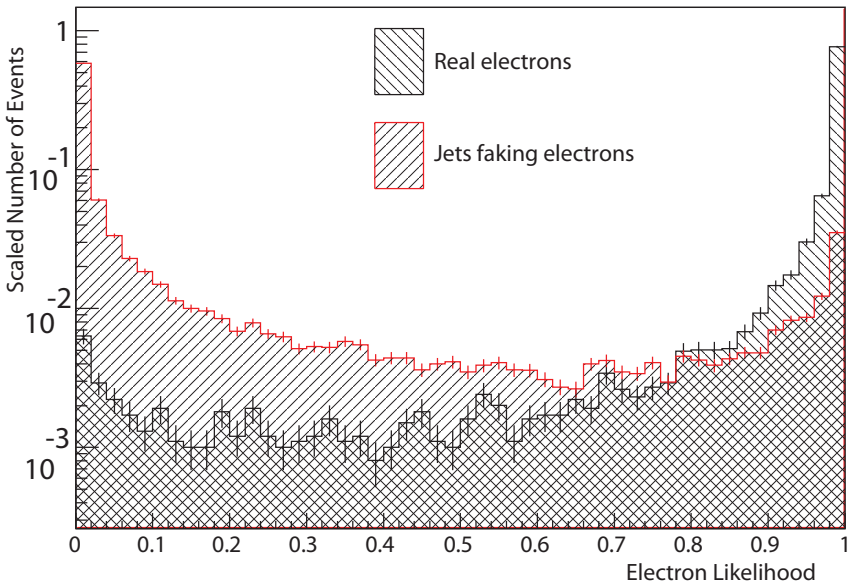
5.3.1 Multijets Background Using Loose Electrons

In order to estimate the multijets background arising from jets misidentified as electrons, a sample of loose electrons is chosen by using loose electron likelihood criteria. The sample uses the same cuts as those for Monte Carlo and data except for the likelihood. The additional cut is determined by reversing the selection cuts used for the MIS electron objects in $e + jets$ and $e\mu$ final states. Based on the plots shown in Figures 5.3(a), 5.3(b), it was determined that the best reflection of electron-like jets could be found using an electron selection with likelihood values between 0.2 and 0.8, in contrast to electrons in the dielectron inclusive final states which require likelihoods greater than 0.85 and electrons in the $e + jets$ final state which require likelihoods greater than 0.95.

The contamination of this background from real electrons decaying from the W boson was found to be only about 0.5% of the sample.



(a)



(b)

Figure 5.3: A comparison of the multijets background to the data minus MC for the electron p_T in (5.3(a)) and the likelihood for electrons from the Z peak vs. those from back-to-back electron-jet in (5.3(b)).

5.3.2 Multijets Background Using Non-isolated Muons

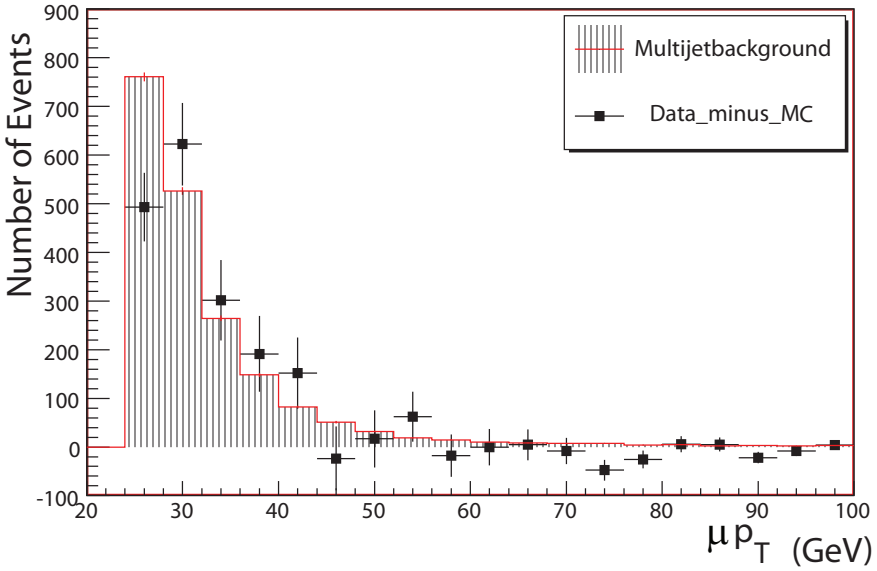
The main contribution of multijets backgrounds that involve muon final states come from real decays of heavy quarks decaying to muons. These muons are typically close to the jets they decay from, so loosened isolation criteria are used to identify this contribution.

The muon object criteria are the same as for other objects with the exception of the E_T track cone and E_T calorimeter halo variables. The criteria for these in the $\mu + jets$ and $\mu\mu$ final states is < 2.5 GeV. To best sample jets that contribute to these final states, these cuts are changed to $4.0 < X < 8.0$ GeV, where X represents both the track cone and the calorimeter halo. The distributions associated with the background cuts can be seen in Figures 5.4(a), 5.4(b).

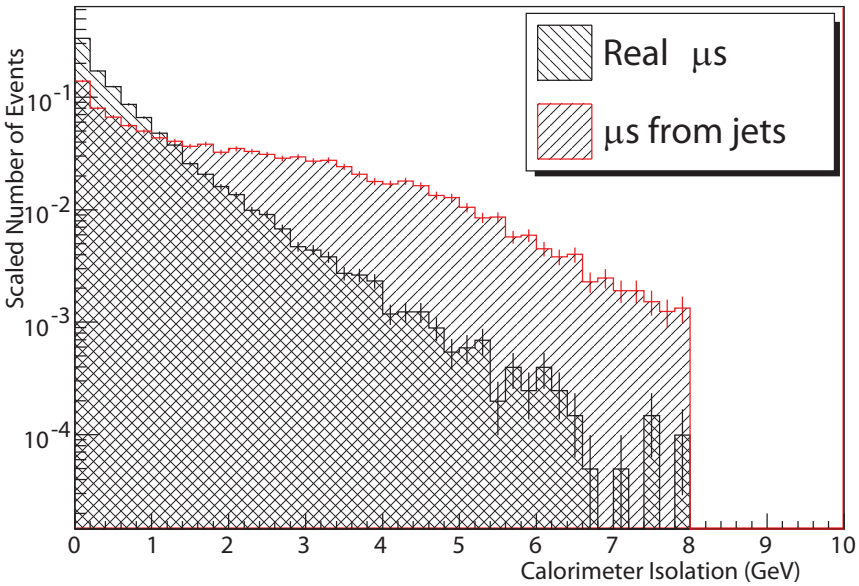
The pollution in this sample from real muons is substantial. This will affect the overall normalization factors as the W process Monte Carlo values will decrease because some of this background is in the multijets selection. This affects the normalization values but should only slightly affect the distributions' overall sensitivity.

5.3.3 Multijets Background Using Loose Taus

In $e\tau$ and $\mu\tau$ final states, multijets background is primarily from jets misidentified as τ leptons. Based on Figures 5.5(a), 5.5(b), the best values of the tau NN to identify jets are between 0.3 and 0.8. These best represent the shapes seen with higher NN values and minimize the pollution from real tau leptons from W boson and Drell-Yan decays.

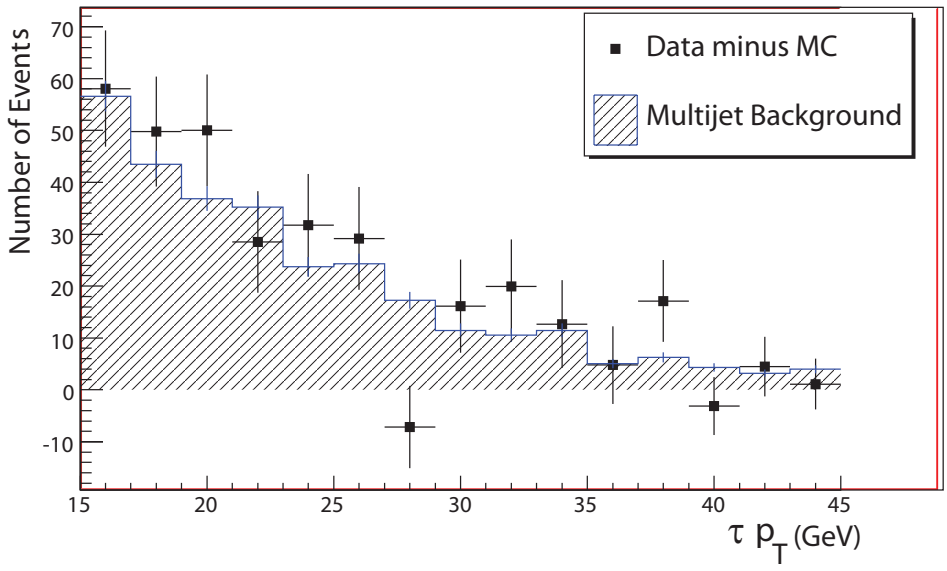


(a)

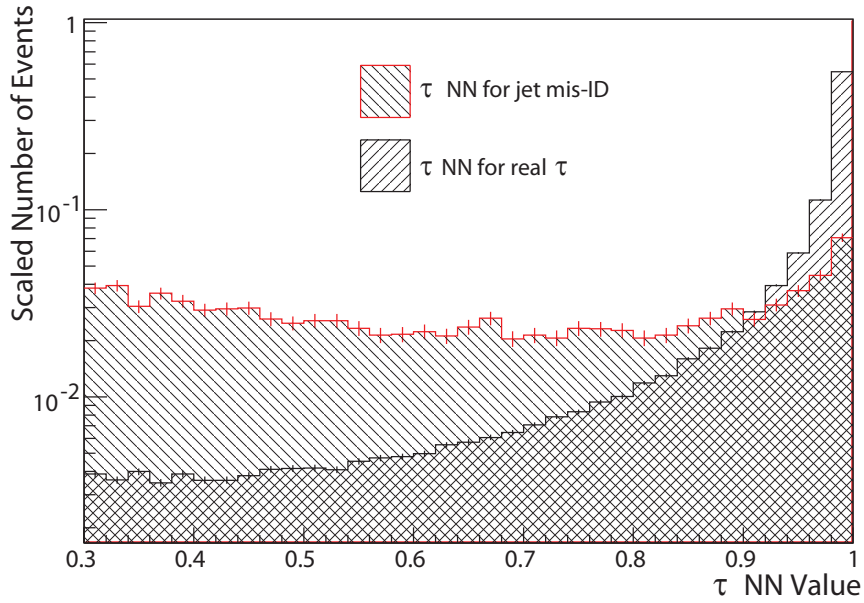


(b)

Figure 5.4: A comparison of the multijets background to the data minus MC for the muon p_T in (5.4(a)) and the calorimeter halo for muons from the Z peak vs. those from back-to-back muon-jet in (5.4(b)).



(a)



(b)

Figure 5.5: A comparison of the multijets background to the data minus MC for the tau p_T in (5.5(a)) and the neural network output value for Monte Carlo taus vs. Monte Carlo jets reconstructed as taus (5.5(b)).

Chapter 6

Corrections to Monte Carlo

The Monte Carlo event generators are required to model finite approximations of processes and GEANT detector simulations must accurately model an extremely complex detector that is not completely static. Given the complexity of the task, some additional corrections must be added to the simulation in order to properly model the data.

6.1 Lepton Smearing and Efficiency

The Monte Carlo simulation of leptons shows higher reconstruction and identification efficiency and better energy resolution than is seen in the data. The efficiency is corrected by applying a scale factor to the events to reduce the event weight. The efficiency factor is chosen by comparing actual and Monte Carlo efficiency of events in the Z-peak. The resolution factor comes from matching the invariant mass spectrum of the Z boson in each of the dilepton states.

6.1.1 Muon Smearing

The muon resolution in data is found to not exactly match that estimated in the Monte Carlo, so an additional "smearing" parameter is applied. "Smearing" convolves the

calculated spectrum with additional resolution effects to match observation. This is determined by looking at data and Monte Carlo events from $Z \rightarrow \mu\mu$ and $J/\Psi \rightarrow \mu\mu$ processes. The widths of these distributions can be seen in Table 6.1. It is found that the correction in the charge to transverse momentum ratio can best be adjusted using a double Gaussian. The equation used is shown below,

$$\frac{q}{p_T} \rightarrow \frac{q}{p_T} + AG_1(0,1) + \frac{B\sqrt{cosh\ \eta}}{p_T}G_2(0,1), \tag{6.1}$$

with A found to have the best fit at 0.007 GeV^{-1} and B at 0.009, where G_1 and G_2 are two independent Gaussian random numbers of mean 0 and $\sigma = 1.0$.

Table 6.1: The width of the peaks using a best fit to a double Gaussian. In both the Z boson and J/Ψ peaks, the Monte Carlo has a better energy resolution than the data. This table was taken from [3].

	Data	Monte Carlo
$\sigma(Z \rightarrow \mu^+\mu^-)$	6.85 GeV	4.84 GeV
$\sigma(J/\psi \rightarrow \mu^+\mu^-)$	0.0651 GeV	0.0647 GeV

The J/ψ distribution before smearing can be seen in Figure 6.1. The Z and J/ψ peaks after smearing can be seen in Figures 6.2(a), 6.2(b).

The resolution is shown to vary from 2.5% from a muon of $p_T = 5$ GeV to 9% at $p_T = 40$ GeV, after requiring $|\eta_{CFT}| = 0$ and at least one SMT hit.

6.1.2 Electron Smearing

Electrons also show better resolution in Monte Carlo than in the data [21]. The general form of the electron smearing is given in [21], but a simplification of this general form is found to be sufficient to match our data in the $Z \rightarrow ee$ distributions, as shown in Equation 6.2.

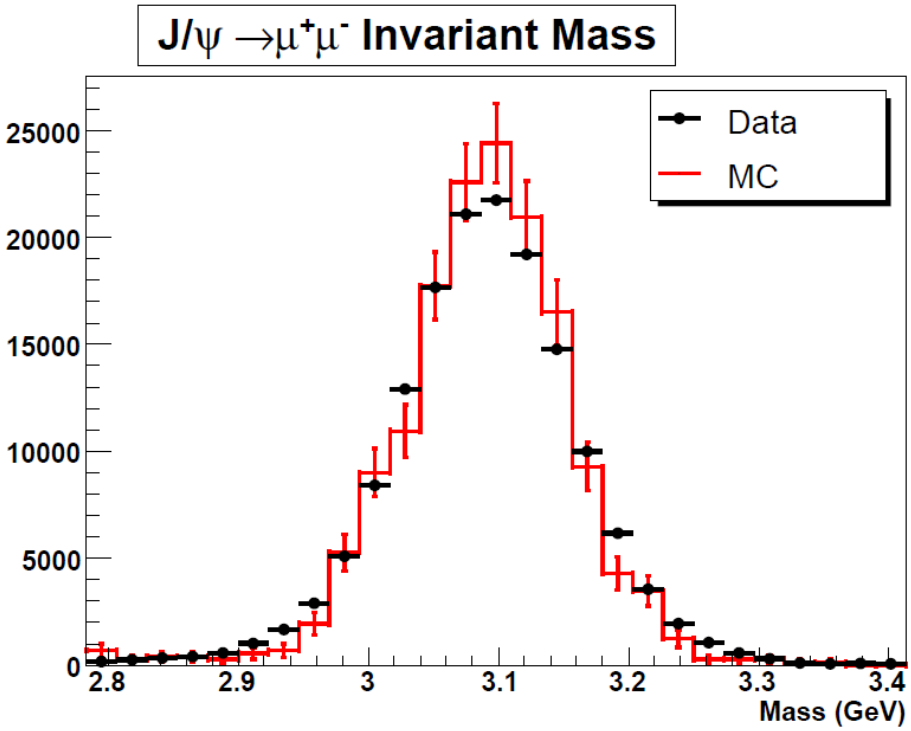


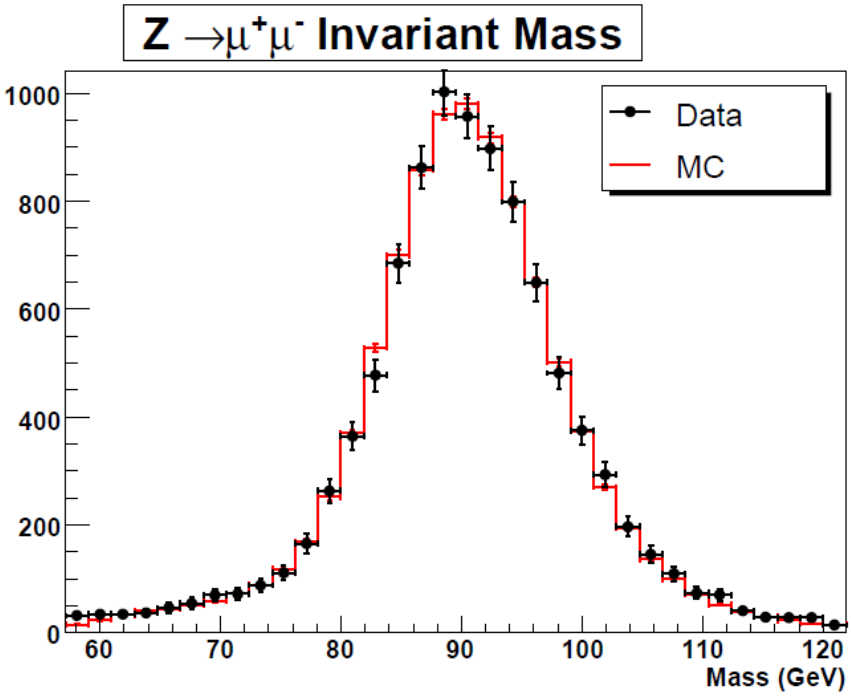
Figure 6.1: The J/ψ peak before applying muon smearing [3].

$$E' = EG(0, \alpha\ c) \tag{6.2}$$

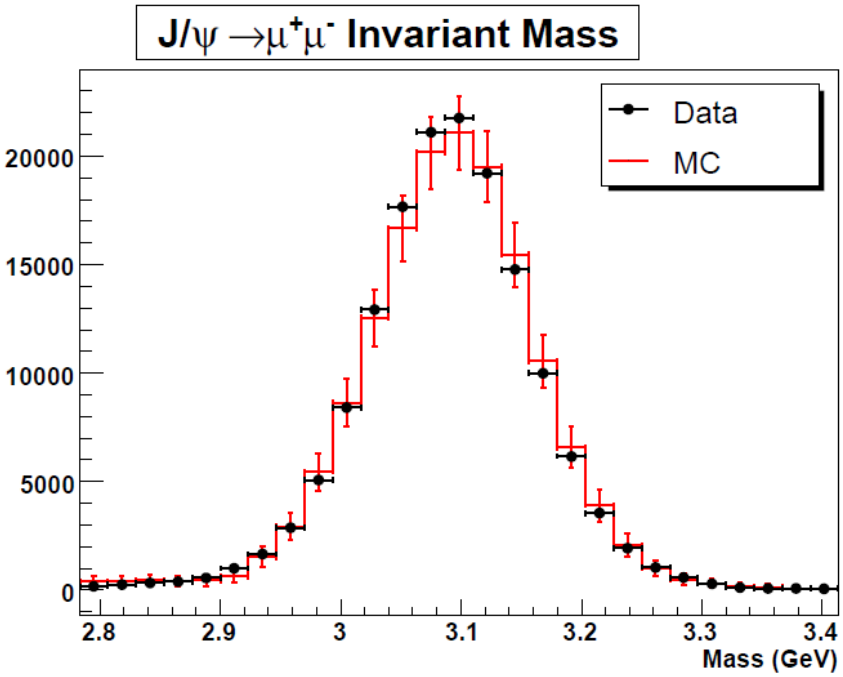
After the fit, the parameters were found to be $\alpha = 1.004$ and $c = 0.0305$ for the fiducial region, where G is again an independent Gaussian random number with mean of 0 and $\sigma = \alpha$c. A figure showing the Z peak after application of the smearing is shown in Figure 6.3.

6.1.3 Muon Efficiency

The muon reconstruction and identification were discussed in Chapter 4. Each of three muon identification criteria (local muon ID, track type, isolation) introduce scale factors with respect to the Monte Carlo [22]. These efficiencies are calculated for each muon in the event, so the overall event weight may incorporate factors for multiple muons.



(a)



(b)

Figure 6.2: μ smearing effects in the Z and J/ψ peaks. Figures taken from [3].

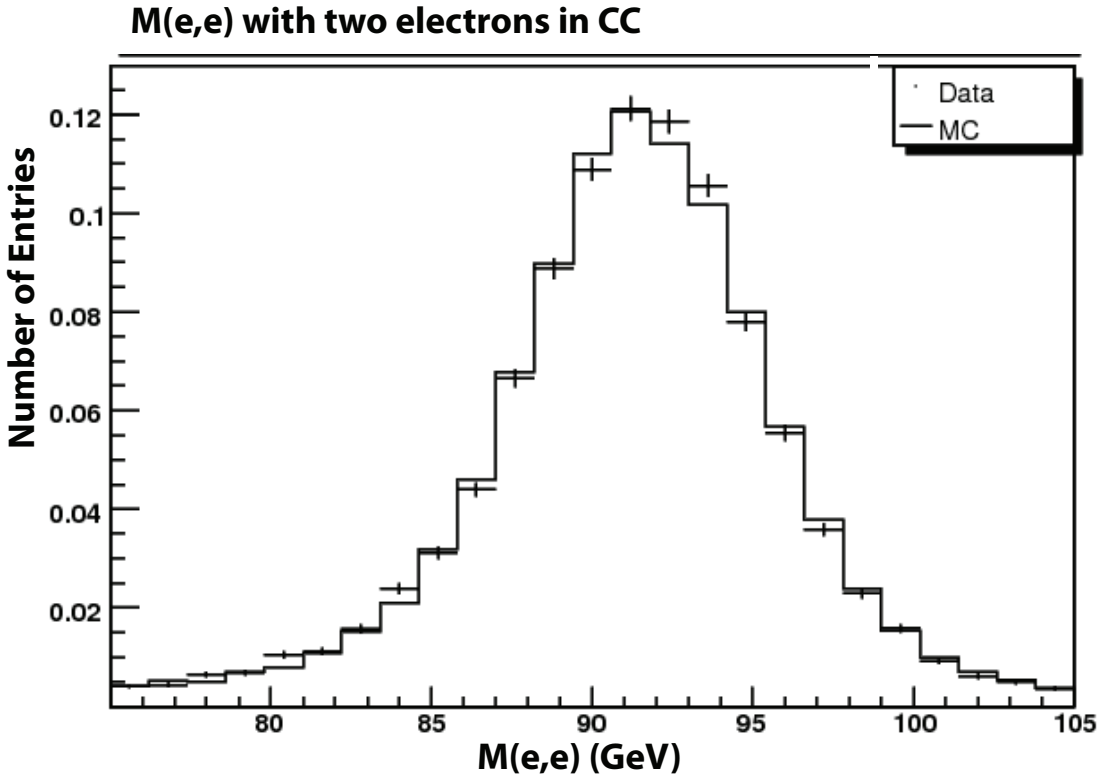


Figure 6.3: Electron smearing effects in the Z peak. Figure taken from [21].

The efficiencies for each of the muon identification criteria are determined by the tag-and-probe method [22]. This method uses tight reconstruction cuts on one muon and loose parameters on a second while choosing events that correspond to the Z boson peak. The muon with tight cuts is called the tag leg, and for data events, this object must have been able to trigger the event. The efficiencies for each of the criteria are determined by looking at the other muon that was only required to pass loose cuts. The fraction of objects passing each of these criteria for the probe leg determines the efficiency.

For the local muon identification efficiency, the detector η and ϕ are used to parameterize the efficiency of the $Medium\ NSeg3$ muons. This averages to a data efficiency of 81.5% and a necessary scale factor of 0.91 (since the Monte Carlo efficiency is 89.6%). For tracks, the parameters z and η_{CFT} define the efficiency for

tight tracks used in this analysis. The efficiency of the tight tracks is around 80.2% with a scale factor of 0.97. Finally, the $NPTight$ isolation efficiency is parameterized by the particle η. This is an average efficiency of 0.92 and also a scale factor of about 0.92. The efficiencies for each of these can be seen in Figures 6.1.3, 6.1.3, 6.1.3.

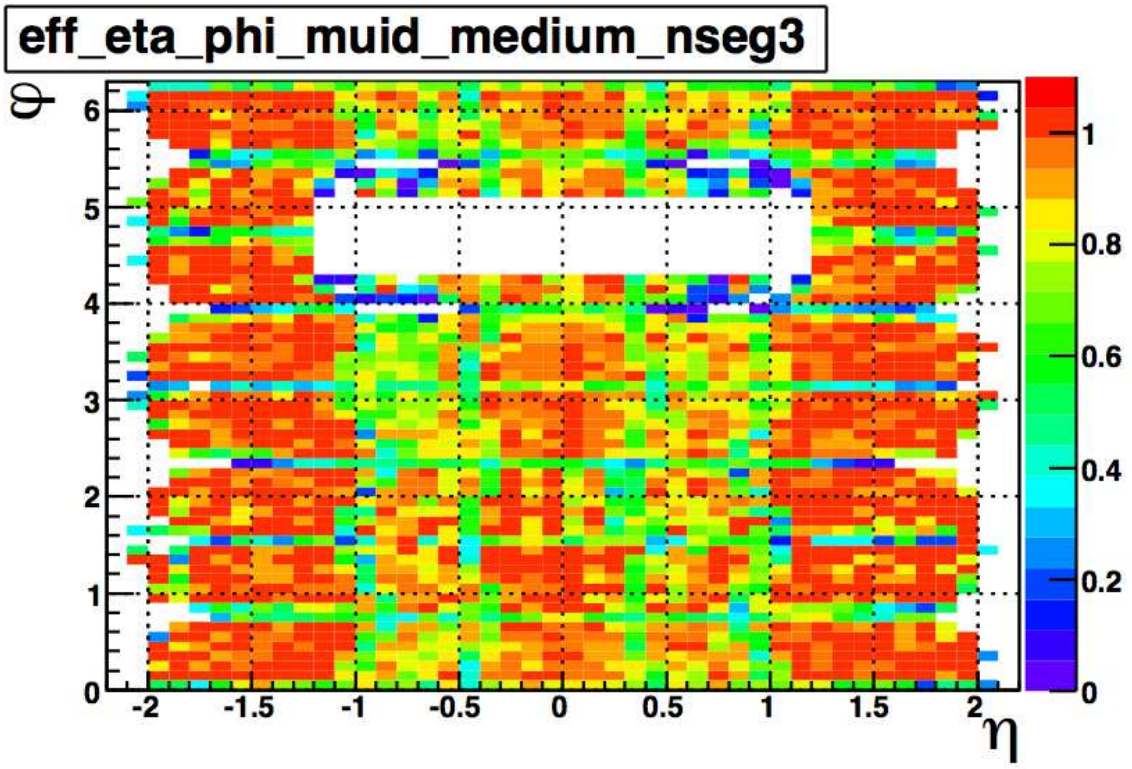


Figure 6.4: Combination μ efficiency for local muon system, cosmic veto, and track match. The hole is due to supports for the calorimeter preventing placement of muon chambers. Figure taken from [22].

6.1.4 Electron Efficiency

The electrons from the simulation also show a higher reconstruction efficiency than that seen in data. The efficiency calculations for electrons use the same tag-and-probe method as was described above for the muons.

The electron efficiencies are measured for two sets of criteria [23]. The first is the preselection where the efficiency of the probe electron is measured for object type,

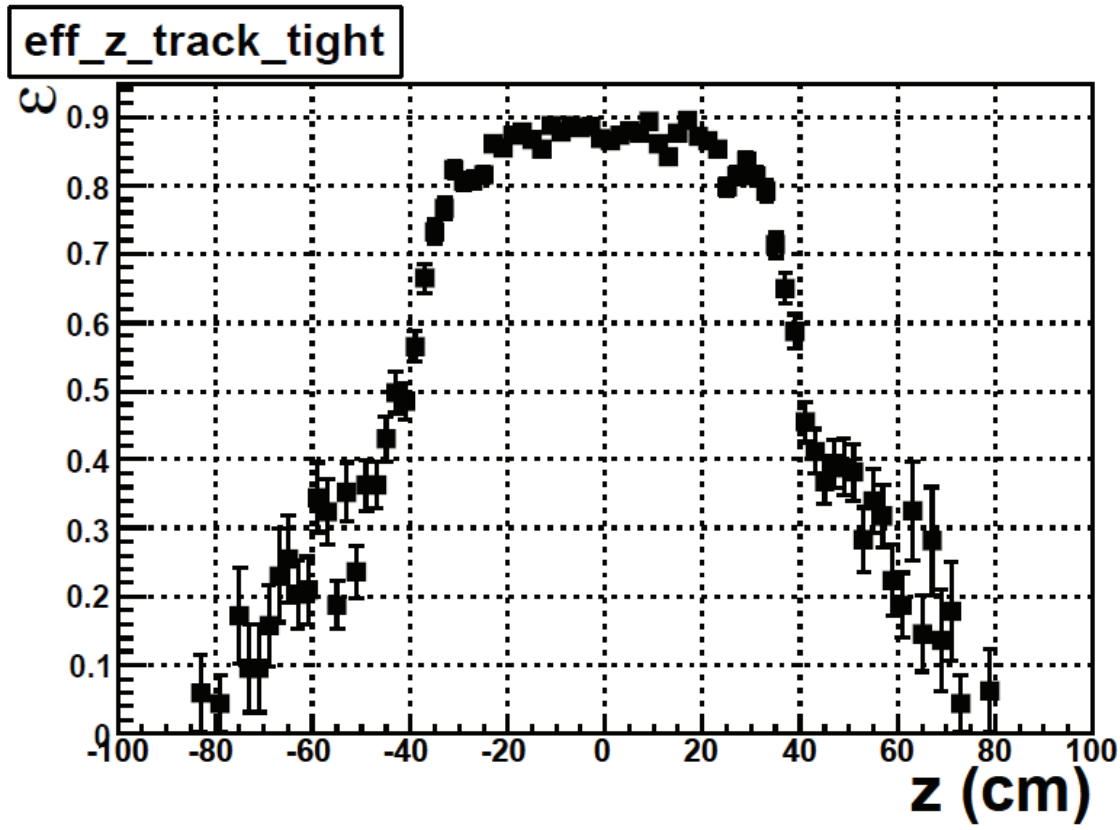


Figure 6.5: μ efficiency for *tight* track reconstruction. Figure taken from [22].

isolation and electromagnetic fraction quantities. The preselection is parameterized in η_{det}, with an average efficiency of around 97%. The *top tight* electron selection used in this analysis, has a further efficiency relative to the preselection. The *top tight* definition includes information on the H-matrix and likelihood. Its efficiency is parameterized in η_{det} and ϕ_{det} and is found to average 73%. The efficiencies are shown in Figures 6.7(a), 6.7(b).

6.2 Jet Weights

Jets at DØ have the same type of difficulties as the leptons: the jet efficiency and energy resolution are better in Monte Carlo than in data. Additionally, the probabil-

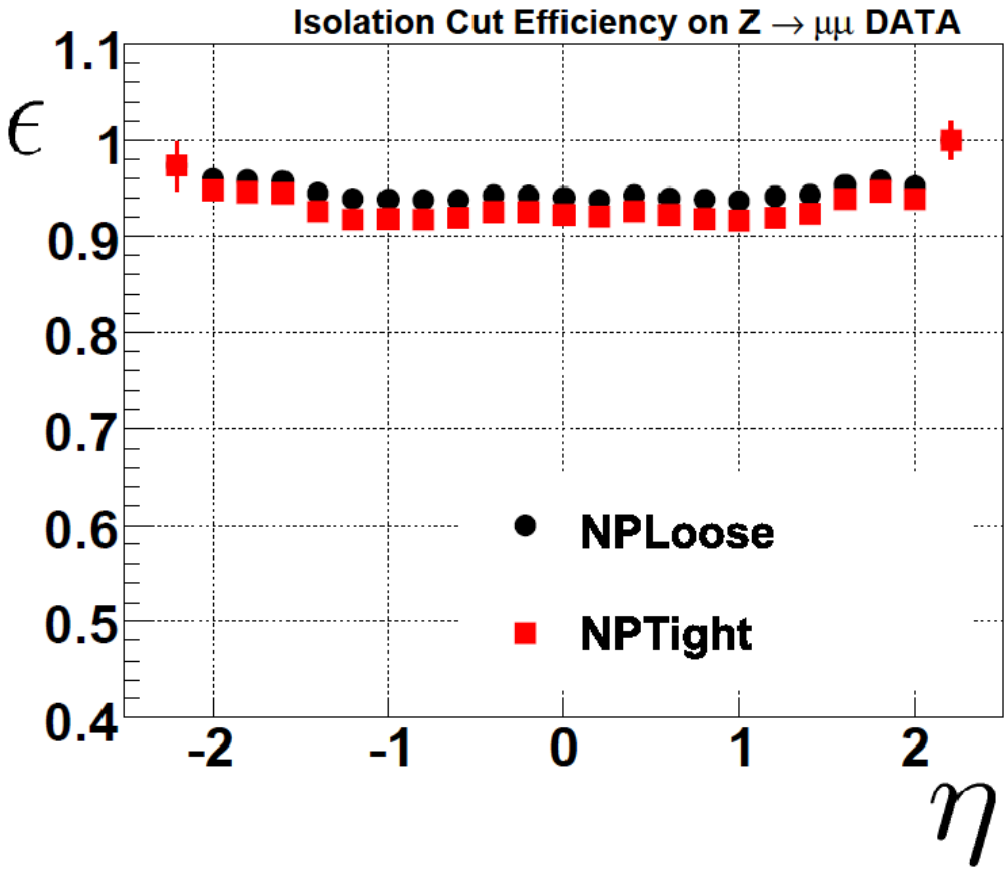
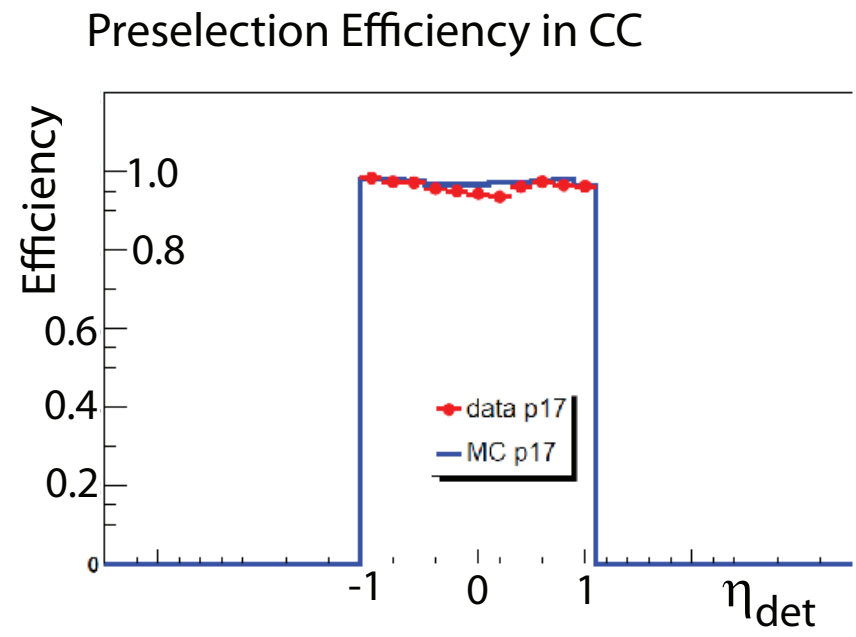


Figure 6.6: μ isolation efficiency for $NPTight$ isolation requirement. Figure taken from [22].

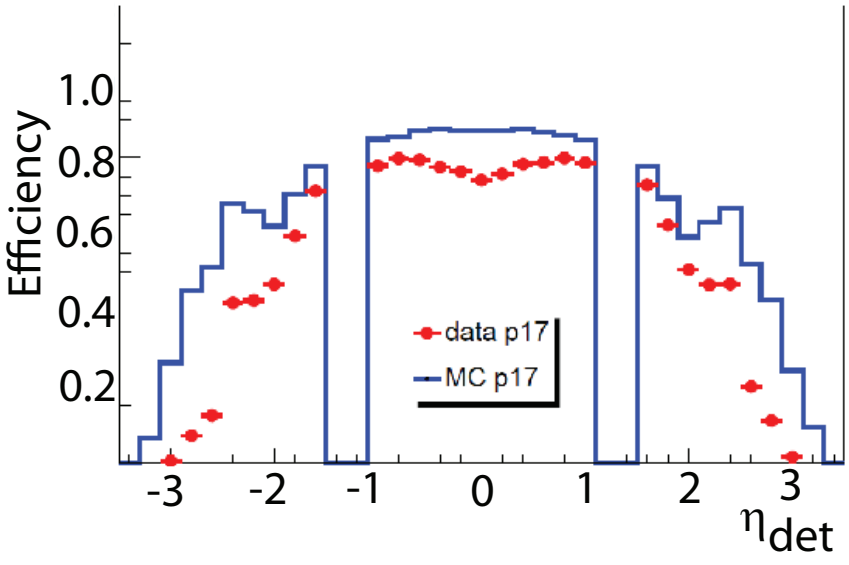
ities to find a track jet (taggability), and to identify a b-quark are also higher in the Monte Carlo.

6.2.1 JSSR

The JSSR correction stands for jet smearing, shifting and removal [24]. It has been found that additional corrections are needed when considering jet modeling beyond those discussed in the section on the jet energy scale. These are the relative jet energy scales, jet energy resolutions, and efficiencies.



(a)



(b)

Figure 6.7: Electron efficiencies for preselection and *top tight*. Figures taken from [23].

These values are determined by looking at $\gamma + jet$ and $Z/\gamma^* + jet$ data and Monte Carlo. In both cases most of the relevant information can be found in the momentum imbalance between the photon and the jet.

$$\Delta S = \frac{p_T^{jet} - p_T^{\gamma/Z}}{p_T^{\gamma/Z}} \tag{6.3}$$

The differences in the momentum distributions between data and Monte Carlo show the additional corrections to Monte Carlo that need to be added to give appropriate jet distributions.

At high values of γ p_T, these ΔS distributions are Gaussian. At lower values, there are some threshold effects. The two distributions can be deconvolved into a Gaussian and error function representing the threshold effects. The turn-on curves plateau by 15 GeV, below which jets are removed. Additional η-dependent efficiencies are applied where the efficiency plateaus. This is only relevant for jets falling in the InterCryostat Region where the efficiency is noticeably lower than 100%. The ΔS distribution showing the convolved Gaussian and error function is shown in Figure 6.8.

The differences in the means of the ΔS Gaussian fits are used to determine the shifting in the Monte Carlo, and the widths of the Gaussians are used as an additional shifting parameter. These functions were derived based on the p_T values of the γ or Z boson, so the reconstructed jet energies must be mapped to approximations of the recoil energy of the boson-jet system.

The final p_T values of the jets are modified according to the following equation after the removal of jets below 15 GeV.

$$p_T^{new} = p_T + p_T^{\gamma/Z}[shift(p_T^{\gamma/Z}) + G(0, smear(p_T^{\gamma/Z}))], \tag{6.4}$$

where $shift$ is the relative jet energy scale, $smear$ is the oversmearing factor, and

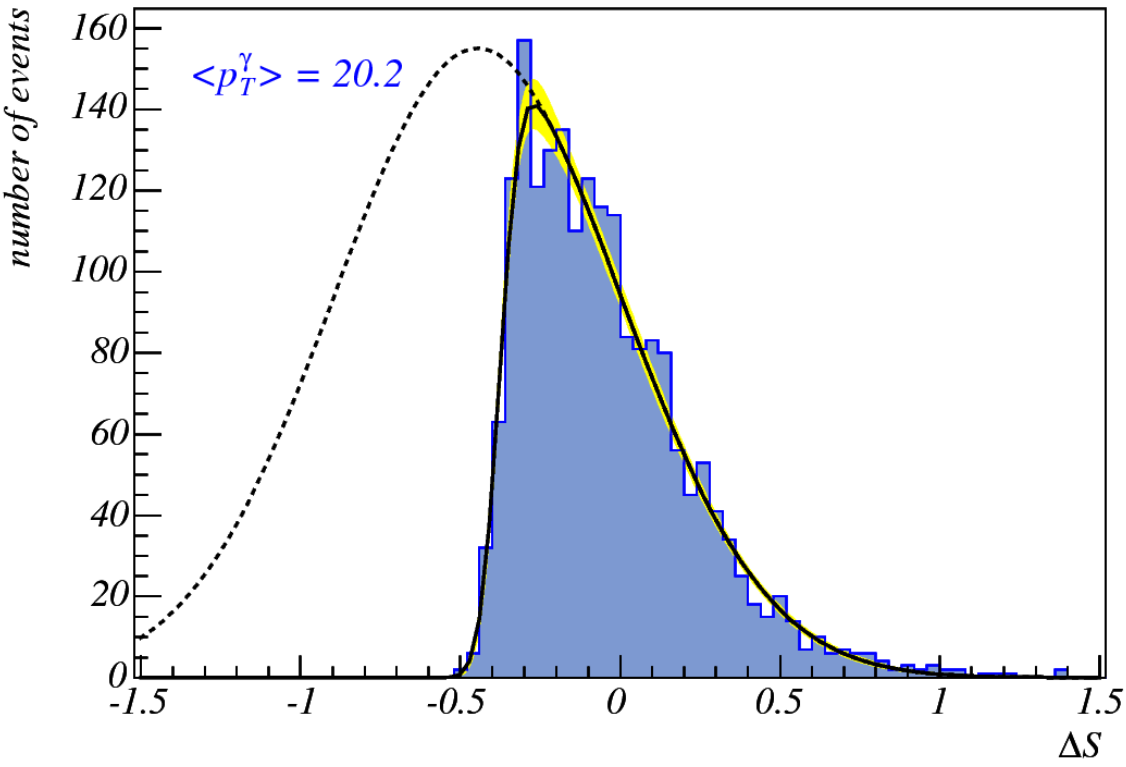


Figure 6.8: The difference in the energy measurements between the photon and a jet in back-to-back events. The distribution can be approximated by convolving a Gaussian distribution with an error function [24].

G is a Gaussian random number with mean of 0 and $\sigma = smear(p_T^{\gamma/Z})$. If jets fail our jet criteria after adjusting the p_T values, these jets are removed. We used the common RunJSSR CAF processor to implement these modifications.

6.2.2 *b*-tagging Rate Correction

The *b*-tagging in this analysis uses direct tagging of the Monte Carlo jets to determine the overall *b*-tagging rate. Direct tagging makes a one-to-one correspondence between a jet and *b*-quark. Each jet will have a probability to originate from a *b*-quark, and based on that probability a certain fraction will be labeled *b*-quark jets. This differs from the tagging rate function method, in which the fractional probability for each jet to originate from a *b*-quark is kept, and each jet is effectively a fractional *b*-quark jet.

The probability to tag a b-quark jet in Monte Carlo is higher than that seen in data [25]. The b-tagging algorithm group looked at samples tagged with the neural network tagger used in this analysis and compared it to rates seen using a soft lepton tagger, which is independent of the neural network. It set up a system of eight equations with eight unknowns, two of which were the tagging efficiency of each of the taggers. The "Tight" operating point used in this analysis has an average data efficiency of 47%, which requires a Monte Carlo scale factor of approximately 0.87. The parameterized efficiency and scale factor can be seen in Figure 6.9.

Sometimes jets are tagged as b-jets, which are not actually jets from b-quark fragmentation. The rate of mistagging depends strongly on the flavor of the actual parton with which the tagged jet is associated. The jet is matched to a quark or gluon within a $\mathcal{R} < 0.4$. The flavor-dependent scale factor is then determined using the following equation,

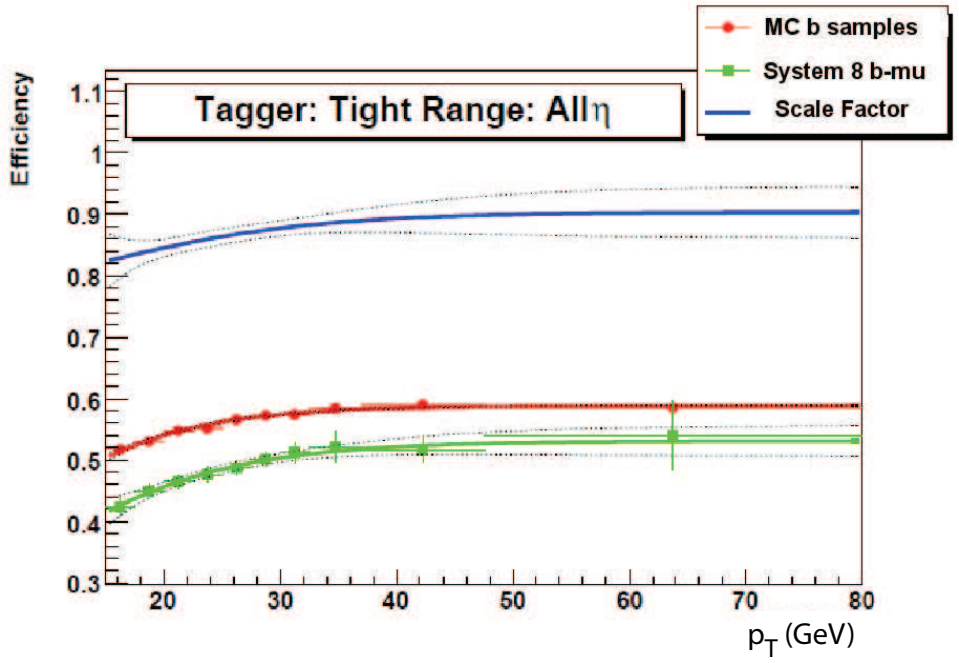
$$SF_{b-jet} = \frac{TRF^{data}(p_T, \eta, flavor)}{TRF^{MC}(p_T, \eta, flavor)} \tag{6.5}$$

where the TRF are the b-tagging rates in data and Monte Carlo. The average fake tag rate with the "Tight" operating point is 0.55%. The parameterized fake tag rate can be seen in Figure 6.10.

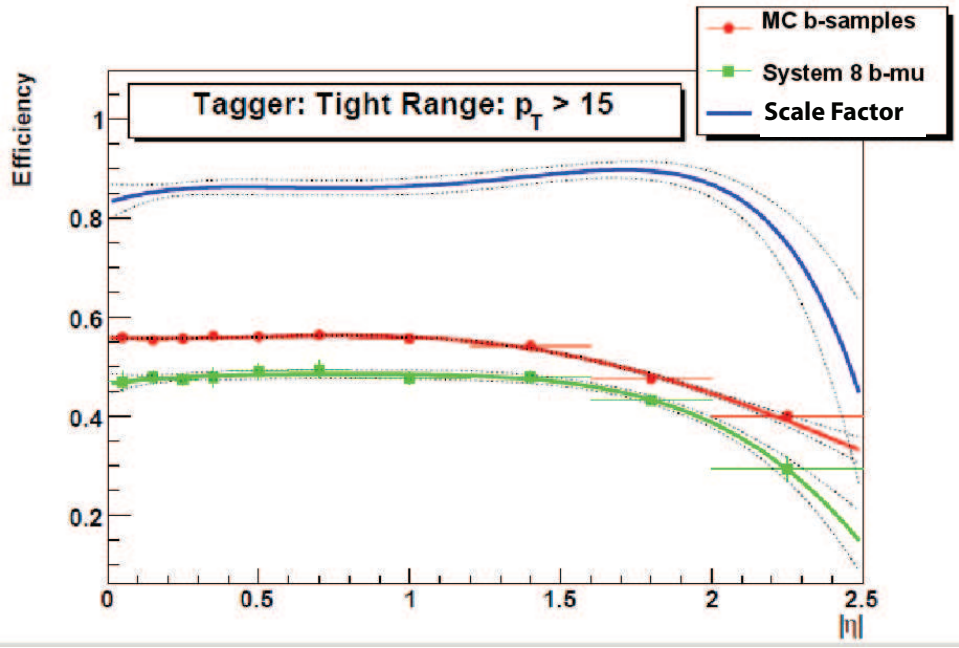
6.2.3 Track Jet Finding (Taggability) Scale Factors

The efficiency to tag a b-quark is divided into two parts. The probability for a particular tagger to tag a jet, and the probability for a calorimeter cone jet to qualify as a track jet ("taggability"). The Monte Carlo is found to have a significantly higher efficiency than data, so a scale factor must be added to properly model the data.

In this analysis, the taggability (track jet-matching efficiency) is derived based on a parameterization in η, p_T, and z vertex position developed by Yuji Enari for the Higgs and W to dilepton, neutrino, and b-jet analysis [79]. The reconstructed



(a)



(b)

Figure 6.9: The efficiency and scale factor necessary to apply to the Monte Carlo for the "Tight" operating point using the NN b-tagger [25].

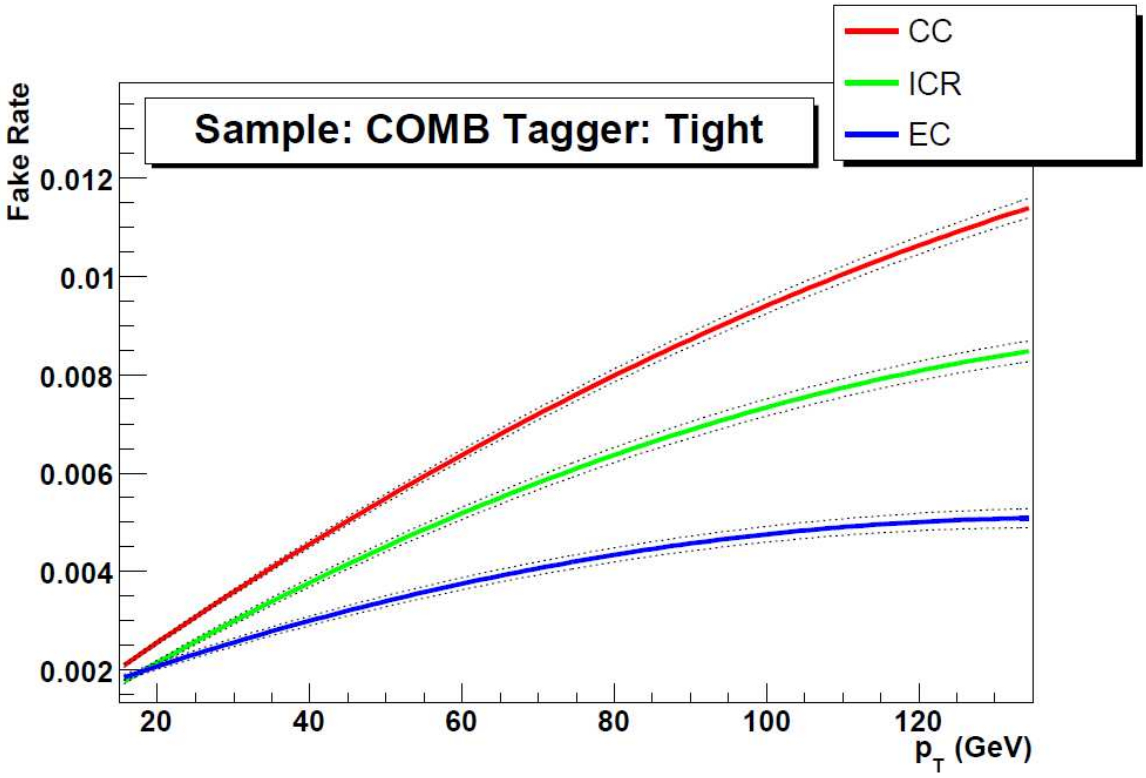


Figure 6.10: The fake tag rate for the "Tight" operating point using the NN b-tagger [25].

cone jet is required to be within $\Delta\mathcal{R} < 0.5$ of a track jet. The track jet requires a hit in the SMT system or on an F-disk, so the taggability is highly dependent on the (longitudinal) z-vertex position and pseudorapidity. The scale factor used in this analysis for W and Drell-Yan processes including heavy quarks, which contribute to single muon final states is shown in Figure 6.11.

6.3 Common Analysis Reweighting

Several Monte Carlo reweighting functions were implemented because the integration of the necessary changes in the Monte Carlo and detector modeling algorithms is nontrivial. These reweighting functions are implemented in a standard way across the collaboration and include fixes to the weak gauge boson spectrum, a reweighting

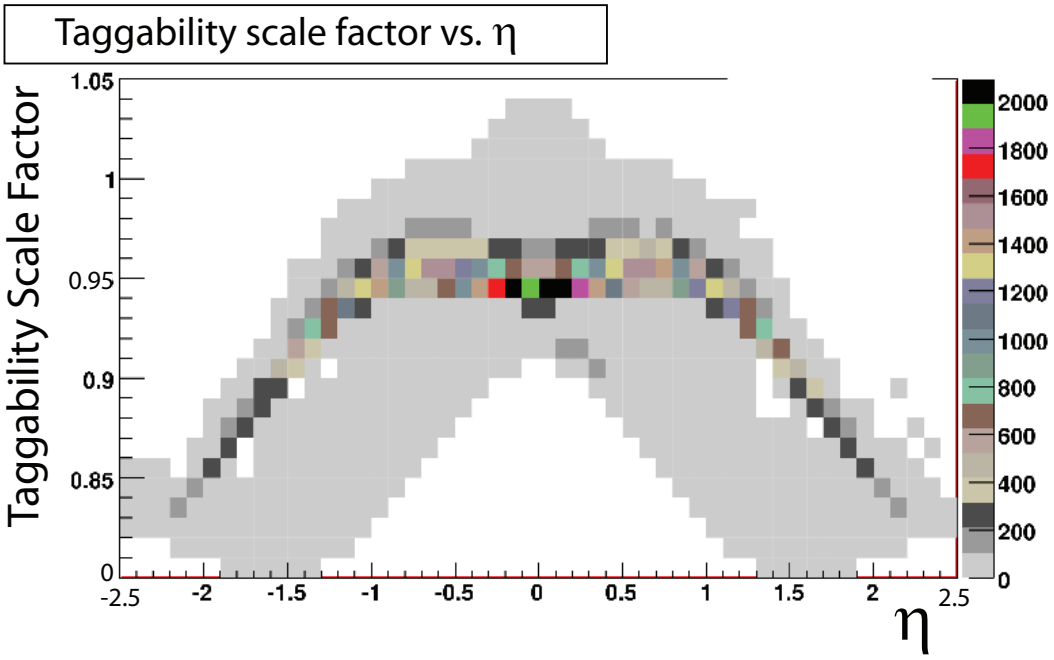


Figure 6.11: The taggability used in this analysis for W and Drell-Yan heavy-flavor processes that are binned as single muon plus jets final states.

for luminosity and one for the z-vertex position.

6.3.1 Weak Gauge Boson p_T Reweighting

The Monte Carlo method of using ALPGEN matched to PYTHIA is inconsistent with data in the Z and W boson p_T spectra at low values of boson p_T [80]. Because of this, a Z p_T reweighting is performed to carefully match the behavior seen in the measured Z p_T distribution from the $Z \rightarrow ee$ process [81].

The Z boson p_T reweighting is carried over to the W boson p_T by utilizing the theoretical ratio of the W/Z p_T spectra [82, 26]. The scale factor used in the W p_T reweighting is shown in Figure 6.12.

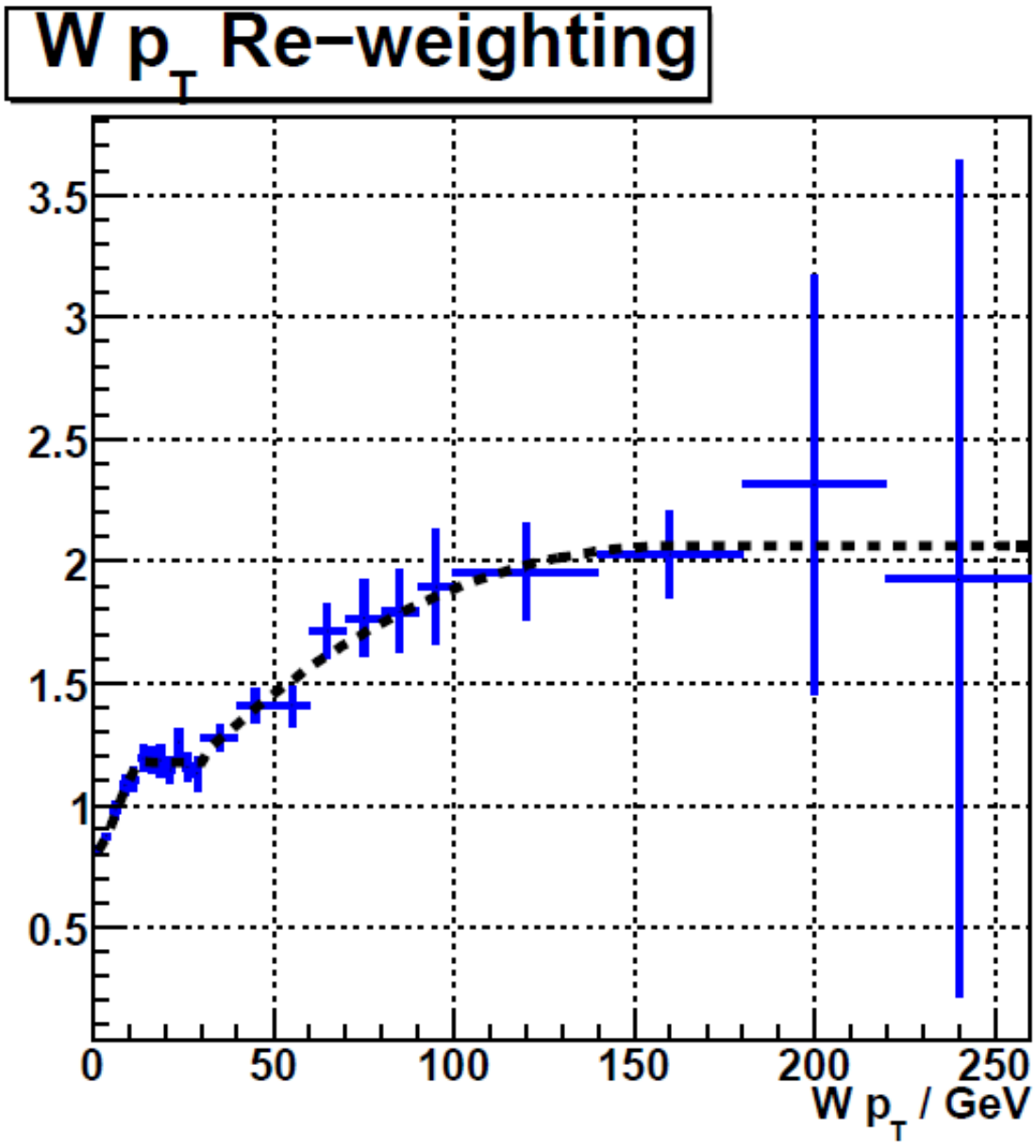


Figure 6.12: The generator-level reweighting function extracted from the cross-section ratio of W boson to Z boson production [26].

6.3.2 Luminosity and z Vertex Reweighting

The Monte Carlo at DØ uses real zero-bias events to model the beam background. Zero-bias events are collected to add the additional beam-interaction background to the hard process modeled in the simulation. In early Run II, there were an average of 2.3 collisions in each beam crossing, and by the end of Run II, the average increased to 5.8. The zero-bias events, however, were taken over a different luminosity distribution than that for the data. Therefore, both of these distributions are plotted, and the Monte Carlo is reweighted to match the actual luminosity distribution of the data [83].

Also, the z vertex position is assumed to be Gaussian in the Monte Carlo. The data shows the vertex position is slightly non-Gaussian, so the Monte Carlo is again reweighted to match the vertex distribution found in data [84].

6.4 Analysis-specific Weights

After applying all of the efficiencies and collaboration weights, the MIS group found it necessary to add two more weighting distributions, a same-sign correction and $\Delta\phi$ correction.

6.4.1 Same Sign Reweighting

Electrons in this analysis have their energy measured in the calorimeter. However, the energy deposition does not give the sign of the electron. The sign is determined by the direction that the associated track bends in a magnetic field. This is not a problem for low momentum, well-defined tracks, but for high-momentum or muddled tracks, it may be difficult to determine the direction of the bend. If the curve is assumed to be in the wrong direction, the electron will end up with the wrong sign. This difficulty is also present for hadronic taus (and to a lesser extent muons). The pion used in

the tau reconstruction also has its energy measurement in the calorimeter and sign determined from the tracking system. Muons also get their sign from the tracking system, but this is then confirmed in the outer muon system to ensure that there is agreement. This last condition makes sign misidentification significantly lower in the muon system.

The rate of sign misidentification will be directly related to the resolution in the tracking system. The problem is not that there is some sign misidentification. The difficulty lies in the fact that the amount of sign misidentification is not properly modeled in the detector simulation. In order to properly model the data, we therefore add another scale factor to the electron and tau Monte Carlo to approximate the appropriate rate of sign mis-ID.

The incorporation of this scale factor is difficult because a direct fit to the full data sample would bias us in our search for new physics. We therefore restrict our sample to dielectron events that have invariant masses in the Z boson peak [27]. For this study, we are only looking at the electron calorimeter energies, and we do not use the track p_T or look at the sign of the electrons. In the mass range used in this study, 60 GeV $< M_{inv} <$ 120 GeV, we see very little contribution from multijet processes. We can therefore assume that the events in this region come exclusively from Drell-Yan production. Figure 6.13 shows the invariant mass distributions of the same-sign and opposite-sign data and Monte Carlo.

The Monte Carlo scale factors for same-sign and opposite-sign events is determined by looking at the data/Monte Carlo ratios of the above distributions. The following equation was used to determine the scale factor.

$$\begin{aligned} \frac{N_{SS,Data}}{N_{SS,Data}+N_{OS,Data}} = SF_{SS}\frac{N_{SS,MC}}{N_{SS,MC}+N_{OS,MC}} \Rightarrow \\ SF_{SS} = \frac{N_{SS,Data}}{N_{SS,MC}} \times \frac{N_{SS,MC}+N_{OS,MC}}{N_{SS,Data}+N_{OS,Data}} \end{aligned} \tag{6.6}$$

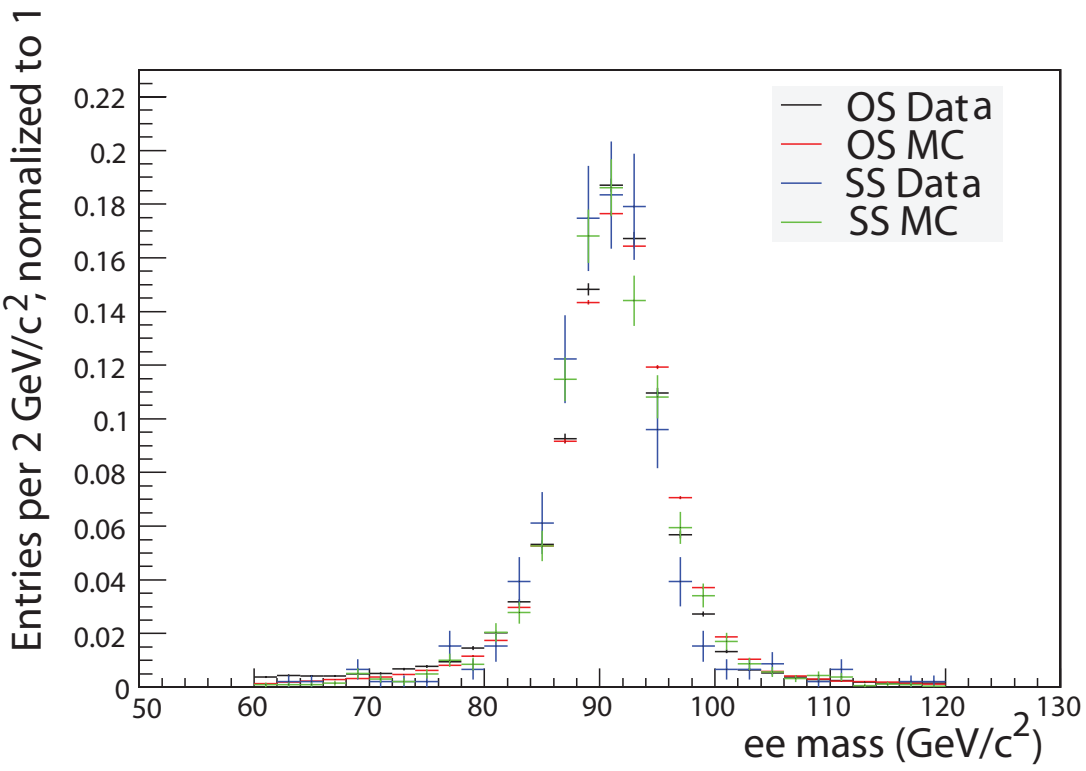
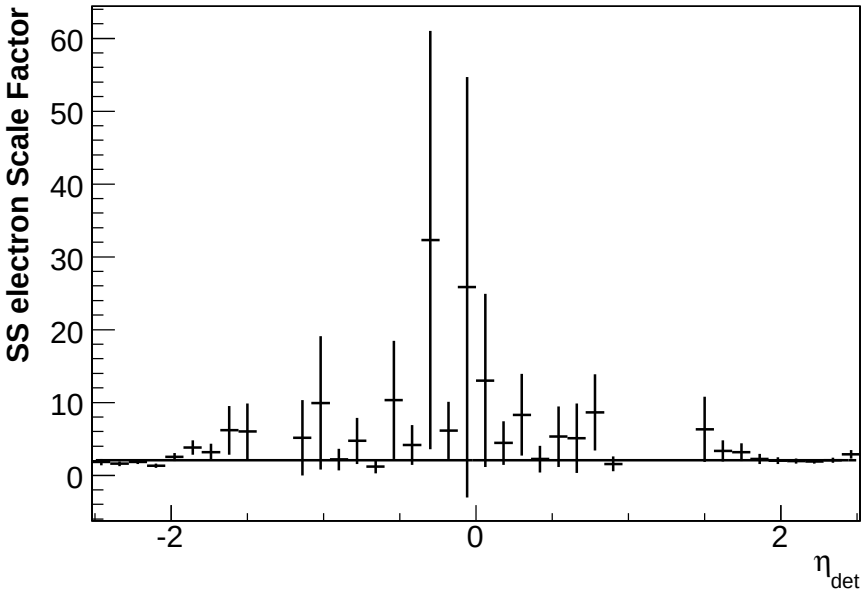


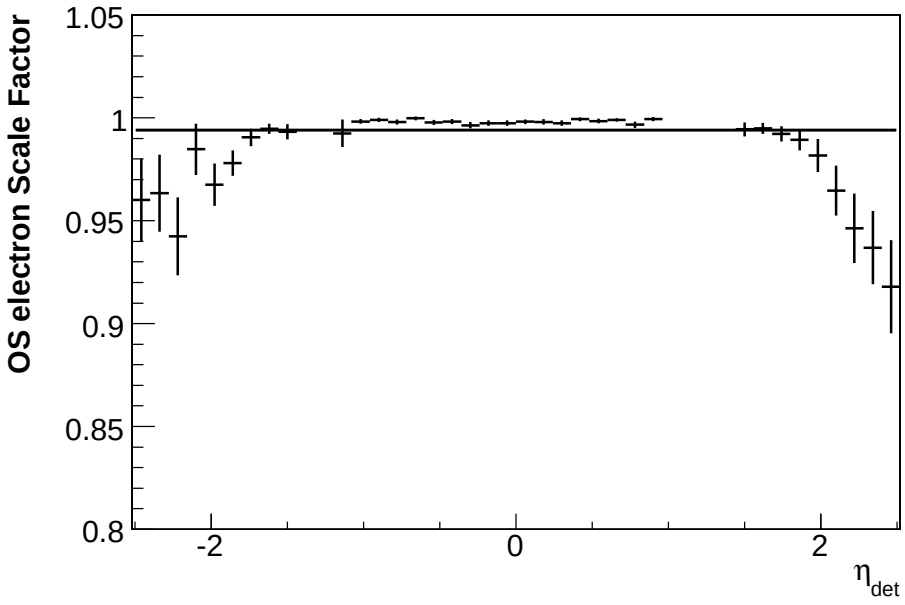
Figure 6.13: The dielectron invariant mass for oppositely charged data (black), oppositely charged MC (red), same sign data (blue), and same sign MC (green). The plot integrals are all normalized to 1.0. These plots are for electrons with $\eta < 2.5$ [27].

with a similar equation for SF_{OS}. The overall scale factor found for same-sign and opposite-sign dielectron events was found to be $SF_{SS} = 2.049$ and $SF_{OS} = 0.994$. The scale factor's detector η dependence is shown in Figures 6.14(a), 6.14(b). While some η dependence was found, the factors were driven by the large number of events in the high-η bins. A separate scale factor for the central bins was not particularly well-motivated, given the few same-sign events with central electrons.

The pions are expected to follow the same sort of distribution because the pions in this analysis are also based on single tracks. Therefore, the same scale factors are applied to the $e + \tau$ final states. For $e + \mu$ and $\tau + \mu$, the same rate of sign misidentification is assumed for scale factors of 1.52 for same-sign states and 0.997 for opposite-sign states. These assume the same probability of a charge flip of the



(a)



(b)

Figure 6.14: Same sign 6.14(a) and opposite sign 6.14(b) scale factors vs. η_{det}. The black line shows the scale factors that are incorporated into the analysis.

electron or tau, but reduce the scale factor in half to account for a single electron or tau.

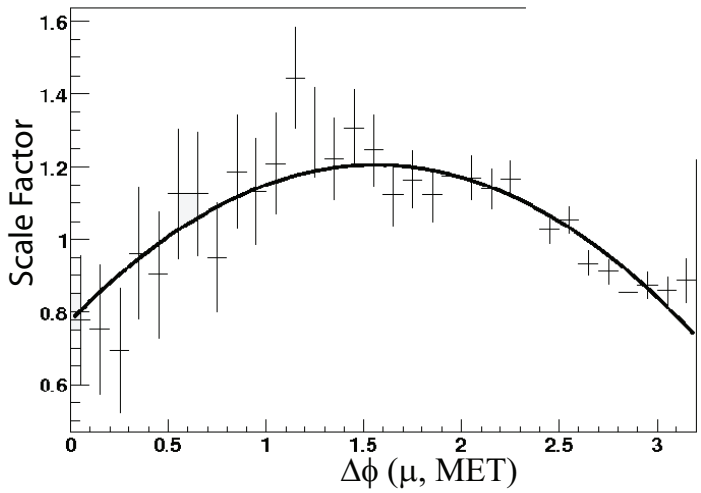
6.4.2 $\Delta\phi$ Correction

The Monte Carlo also shows some significant differences between the data and Monte Carlo $\Delta\phi$ distributions between objects in events. Large discrepancies were seen between leptons in dilepton final states and between the lepton and $\not\!\!E_T$ in single lepton + jets final states. We must make an assumption here: that these distributions are not expected to reveal new physics, but rather indicate a modeling deficiency. With that assumption, we apply a reweighting scheme to the $\Delta\phi$ distributions. This reweighting affected not only the $\Delta\phi$ distributions themselves, but also other quantities that depend on the spatial distribution of particles such as W boson p_T.

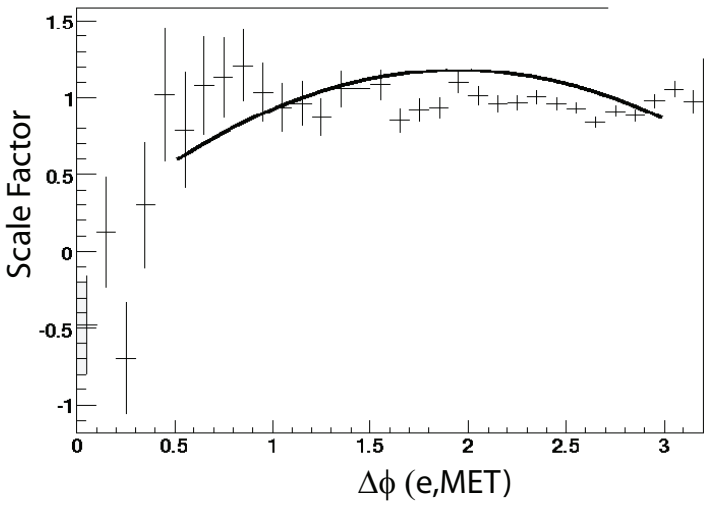
For the μ + jets state, we first applied an $\not\!\!E_T$ cut of 20 GeV. The $\not\!\!E_T$ direction for events with $\not\!\!E_T$ less than 20 GeV is not well known and should not be used in calculating the $\Delta\phi$ weights between μ and $\not\!\!E_T$. The results of $\Delta\phi(\mu, \not\!\!E_T)$ is shown in Figure 6.15(a). This correction is fit to a parabola to $\Delta\phi < 2.5$ where it is smoothed into a linear fit, and continuity is forced at the transition.

In the e + jets final state, we also have an $\not\!\!E_T$ cut of 20 GeV, so the fit can be applied directly. The result of the e + jets fit is shown in Figure 6.15(b). The fit here is not obviously different from flat, so no additional scale factor is applied.

The dilepton final states only need adjustments at $\Delta\phi$ values that imply nearly back-to-back leptons. This can be seen in Figures 6.16(a) and 6.16(b). In $\mu\mu$, a weight of 1.1 is used for $2.8 < \Delta\phi < 3.0$ and 0.93 for $\Delta\phi > 3.0$. For ee, the weight is 1.06 for $2.8 < \Delta\phi < 3.0$ and 0.93 for $\Delta\phi > 3.0$.

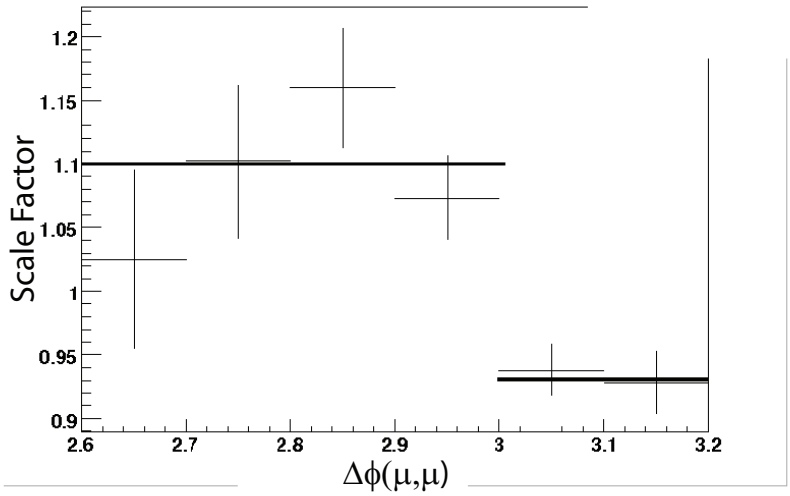


(a)

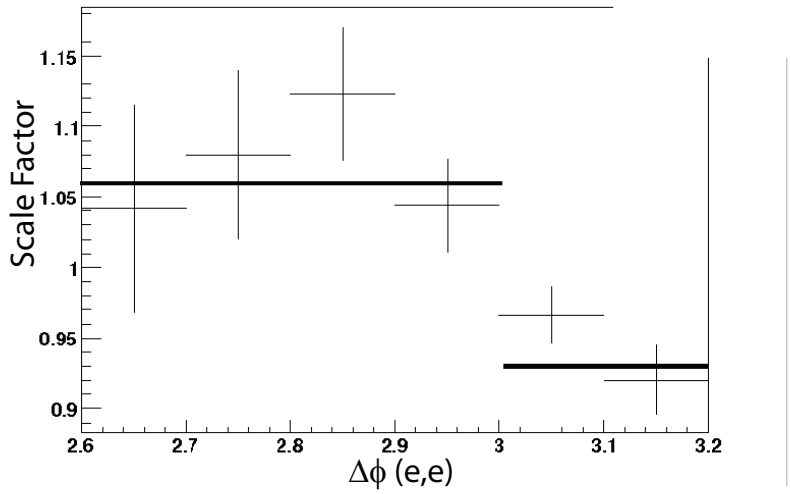


(b)

Figure 6.15: Ratio of $\Delta\phi$ distributions between data and Monte Carlo in inclusive single lepton final states.



(a)



(b)

Figure 6.16: Ratio of $\Delta\phi$ distributions between data and Monte Carlo in inclusive dilepton final states.

Chapter 7

MIS Analysis Packages

After all known Monte Carlo corrections are applied, the whole data set is divided into seven non-overlapping final states. The overall Monte Carlo normalization weights are adjusted for input processes that contribute to each of the inclusive final states. The fit process uses several histograms of basic object quantities (simple, single-object kinematic variables) and fits a single parameter for each of the input processes, so that the χ^2 probability is minimized for the combined fit. Once the fit values are found, the histograms are plotted again, taking into account the values obtained from the general fit. Several other quantities are also plotted before and after the application of the fit parameters to see how the general fit affects more complex distributions. Note: Most of this chapter is adapted from the internal DØ analysis note, and from the publicly available conference note [27, 85].

7.1 Inclusive Final States

The seven states are inclusive in jets and additional objects, as specified in Table 7.1, where each state is listed with the objects that define it and the associated object cuts. The additional objects (X in the table) require cuts as seen in Table 7.2.

Table 7.1: Table of final state object cuts: The seven inclusive final states that are being considered, along with their basic object cuts.

MIS Final State	Object	Min p_T (GeV)	Max $\|\eta\|$
e + jets + X[a]	e	35	1.1
	jet	20	2.5
	$\not\!\!E_T$	20	NA
μ + jets + X[b]	μ^h	25	1.7
	jet	20	2.5
	$\not\!\!E_T$	20	NA
ee + X[c]	e	15	1.1
$\mu\mu$ + X[d]	μ^h	15	2.0
$\mu\tau$ + X[e]	μ^h	15	2.0
	τ	15	2.5
$e\tau$ + X[f]	e	15	2.5
	τ	15	2.5
μe + X[g]	μ^h	15	2.0
	e	15	2.5

[a] X $\neq e$, μ, τ, γ
[b] X $\neq e$, μ, τ
[c] X $\neq \mu$, τ
[d] X $\neq e$, τ
[e] X $\neq e$
[f] X may be any object
[g] X $\neq \tau$
[h] Muons have an additional maximum p_T cut of 300 GeV.

Table 7.2: Table of object cuts required for inclusion as additional objects (X) in one of the seven final states listed in Table 7.1.

Object	Min p_T (GeV)	Max $\|\eta\|$
e	15	2.5
μ[a]	15	2.0
τ	15	2.5
jet	20	2.5
γ	15	1.1

[a] Muons have an additional maximum p_T cut of 300 GeV.

7.1.1 μ + jets

The μ + jets final state is dominated by the production and decay of W bosons. This state is defined by exactly one muon with $p_T > 25$ GeV and with $\eta < 1.7$. In order to reduce the amount of multijet background, at least one jet with $E_T > 20$ GeV is also required, as well as $\not\!\!E_T > 20$ GeV. The muon must satisfy the $Medium\ Nseg3$ conditions for the local muon system, the *tight* track requirements, and the $NPTight$ isolation requirements as discussed in Chapter 4. The final state is inclusive in jets and photons, but any other additional objects would push the event into a different final state.

7.1.2 e + jets

The electron + jets final state parallels the muon + jets final state. Jets are more easily misidentified as electrons, so the cuts on this final state are slightly tighter to minimize the contribution from the multijet final states. The electron p_T cut is at 35 GeV with $\eta < 1.1$ and $\not\!\!E_T > 20$ GeV. The likelihood cut, $\mathcal{L}_e > 0.95$, is tighter than the default *top tight* definition.

7.1.3 $\mu\mu$

The dimuon final state requires at least two muons with the same quality definitions as the muon from the μ + jets final state. The muon p_T cut is dropped to $p_T > 15$ GeV because of the smaller contribution from multijet background. If this state includes jets, jet p_T must be > 20 GeV. It is inclusive in jets and muons, but an additional e or τ would place the event in the μe or $\mu\tau$ final states.

7.1.4 ee

The dielectron final states require each electron to have e $p_T > 20$ GeV and likelihood > 0.85. The electrons are also confined to the central calorimeter and use the same jet cuts as the other final states. The end calorimeters were excluded for this analysis because of inconsistencies between electrons measured in the central calorimeter and electrons measured in the end caps when attempting to fit histograms in the dielectron final state normalization fit.

7.1.5 $\mu\tau$

The $\mu\tau$ states contain at least one muon and one tau. It is inclusive in all objects except electrons, which would move the state into the μe final state. The requirements are μ $p_T > 15$ GeV and τ $p_T > 15$ GeV. The τ $NN_\tau > 0.9$, and the τ of type 2 has an additional electron separation cut of $NN_e > 0.2$.

7.1.6 $e\tau$

The electron + tau objects are inclusive in all objects. The electron and tau p_T cuts are $p_T > 15$ GeV. The electron likelihood is at > 0.95, for this state to differentiate it from a large multijet background (since many apparent τ's are misidentified jets). The hadronic NN cuts are the same as $\mu\tau$, but the electron separation NN is set to

0.8 to remove the dielectron events.

7.1.7 μe

The μe final state requires p_T cuts of 15 GeV for the muon and the electron. It is inclusive except for τ's which would fall into the $e\tau$ final state. The electrons can be identified normally as *top tight* with likelihood > 0.85 or as misidentified τ's with the electron separation NN < 0.2.

7.2 The MIS Fit

The fits for normalization use several histograms of basic object quantities to determine a scale factor, altering the overall normalization of each input process, so that the χ^2 probability is minimized for the combined fit. In order to avoid fitting to the high-p_T tails that will eventually be searched for new physics, we check each object in the event to see if the object p_T is outside the bulk of the distribution. Basic histograms like $\not\!\!E_T$, p_T, η, $cos(\phi_{obj} - \phi_{\not\!\!E_T})$ for the leptons and jets are used to fit while we reserve more complex variables to check the fit quality. These more complex variables include the mass or transverse mass of two or more objects, jet multiplicities, $\Delta\phi$ between two objects, inclusive jet p_T, W and Z p_T, etc. If an event contains any object in the tails, then none of the objects in the event will be used in the fit.

A full list of the processes which are normalized based on these inclusive fits, along with the final states that are used to determine their values, are shown in Table 7.3. A slightly simplified example using the electron + jets + X final state (X is not an e, μ or τ) would work as follows. We know this state to be dominated by the W processes, but it also has a significant contribution from multijet and Drell-Yan. We use a constant normalization factor for the Drell-Yan process, determined by a separate fit to the ee + X final state (X $\neq \mu$ or τ). This parameter will be held fixed

in the e + jets fit, along with other rare processes which have contributions which would be too small to fit accurately. Then, the W and multijet contributions will find the best agreement to fit the given histograms and two scale factors will be used to give an overall weight to the $W \rightarrow e\nu$ and $multijet \rightarrow e + jets$ contributions. Once the fit values are found, the histograms are plotted again taking into account the values obtained from this fit. The total background contribution for a particular final state would be,

$$B = \sum_{i}^{\#bkg} SF_i B_i \tag{7.1}$$

where the scale factors (SF_i) for each background process (B_i) are determined only once in the final state in which their contribution is the most important, and then held fixed in any other final state to which they contribute.

Our simplified modeling implementation does not directly account for certain normalization factors due to such things as trigger efficiencies and some K-factors (corrections for Monte Carlo leading order or leading logarithm cross sections to the observed cross section for a process). In order to avoid gross errors in normalization, we perform a fit, described below, for each of these states to obtain the scale factors which reproduce the distributions of the selected data with the background from standard model Monte Carlo and multijet background determined from data. These seven states were selected so that each is dominated by a specific standard model process. Since the seven states are non-overlapping, they can be combined as an input to the VISTA algorithm without fear of double-counting.

The fit itself minimizes the negative logarithm of the likelihood function for each set of parameters and converts this value to a χ^2. The fit is performed with the Minuit fitter [86]. It minimizes the χ^2 of the fitting histograms by looking at the differences, bin-by-bin, between the data and the standard model background. The floating parameters are modified until a minimum is found. Only two or three parameters for each of the final states are modified. The plots are not varied explicitly by shape.

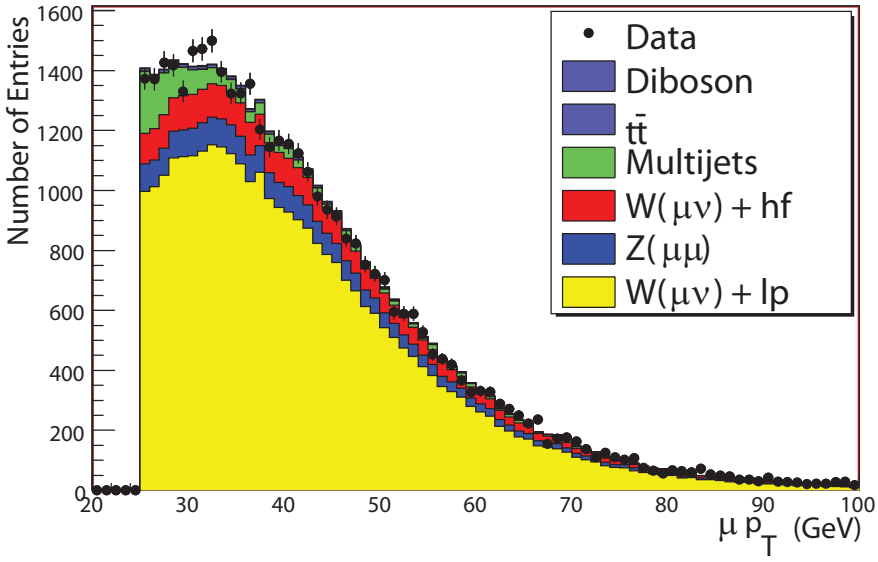
Table 7.3: Table of input processes for which the normalization is determined from inclusive final state fits along with the final states that are used in determining its value.

Input Process	Final States
$W \to e\nu$ + light partons	e + jets
e multijet background (e + jets)	e + jets
$W \to \mu\nu$ + light partons	μ + jets
μ multijet background (μ + jets)	μ + jets
$Z/\gamma \to ee$ + light partons	ee
$Z/\gamma \to \mu\mu$ + light partons	$\mu\mu$
Heavy flavor/light flavor content	e + jets, μ + jets, ee, $\mu\mu$
Z/γ: >0 light partons/0 light partons	ee, $\mu\mu$
$Z/\gamma \to \tau\tau$ + light partons ($e\tau$)	$e\tau$
τ multijet background ($e\tau$)	$e\tau$
$Z/\gamma \to \tau\tau$ + light partons ($\mu\tau$)	$\mu\tau$, τ types (1,2) and 3
τ multijet background ($\mu\tau$, τ types 1,2)	$\mu\tau$, τ types 1,2
τ multijet background ($\mu\tau$, τ type 3)	$\mu\tau$, τ type 3
$Z/\gamma \to \tau\tau$ + light partons (μe)	μe
e multijet background (μe)	μe

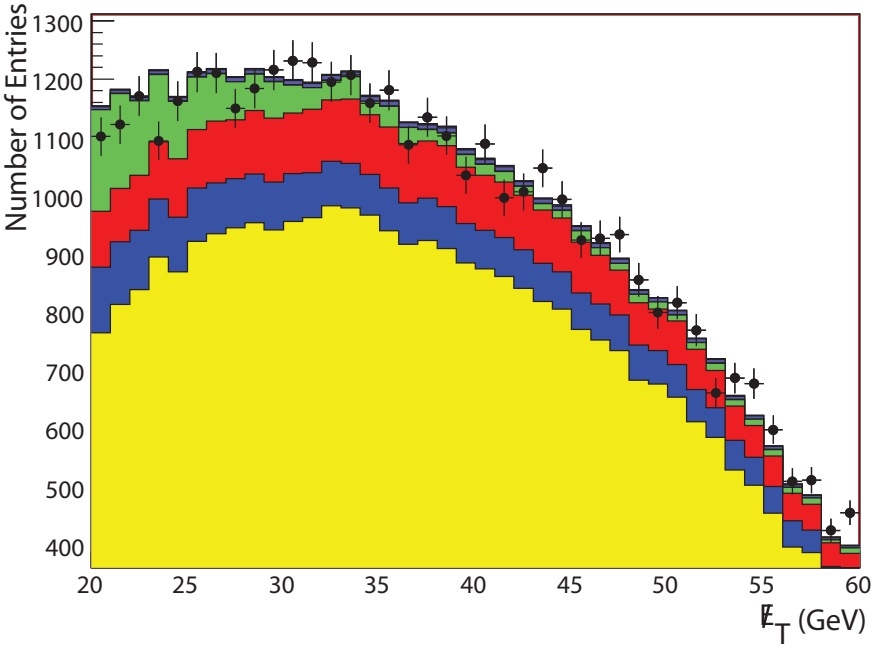
The final results are the best fit parameters corresponding to the normalization of the floating input processes. For the single lepton states and the τ final states, the multijet background is a significant contribution. It is assumed that the contribution from the other Monte Carlo samples to the multijet background is small. The input process scale factors for the Monte Carlo should already include the contributions of the process to the multijet background from data. The main effect of process contributions to the multijet background would be that the multijet state would resemble the process Monte Carlo, making it difficult for the fit to reliably find the multijet contribution.

The fits are then checked for qualitative agreement. The main purpose of the MIS normalization process is to make sure that the fundamental processes are well-modeled. A Kolmogorov-Smirnov (KS) probability is determined for each of the histograms to provide a quantitative check for comparison. Additionally, the overall scale factors are checked to compare to those of other analyses. If all normalization factors were properly included in the Monte Carlo, the scale factors would all fit to 1.0. No specific cut is required for the KS probability because the main quantitative analysis will be done at the later VISTA and SLEUTH stages. Two histograms that are included in the overall fit and two checking histograms that are not part of the fit are shown for each of the seven final states in Figures 7.1, 7.2, 7.3, 7.4, 7.5, 7.6, 7.7, 7.8, 7.9, 7.10, 7.11, 7.12, 7.13, 7.14.

After the first stage of fitting is finished, several constraints are imposed upon the fitted condition. The ratio of light-parton to no-parton content that is fit for the dimuon and dielectron states is constrained to the same value. We expect this ratio to be independent of flavor, so a weighted average of the two states is taken and then fixed for a final fit. The same is done for the heavy-flavor states. The ratio of heavy-flavor to light-parton content is fixed for the dilepton and single lepton final states. An example equation showing the averaging process is shown in Equation 7.2.

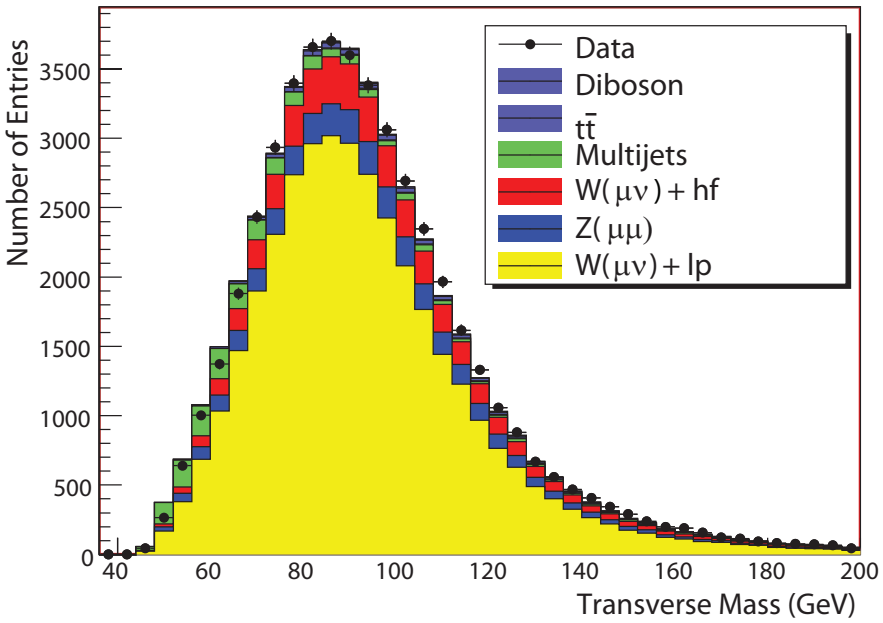


(a)

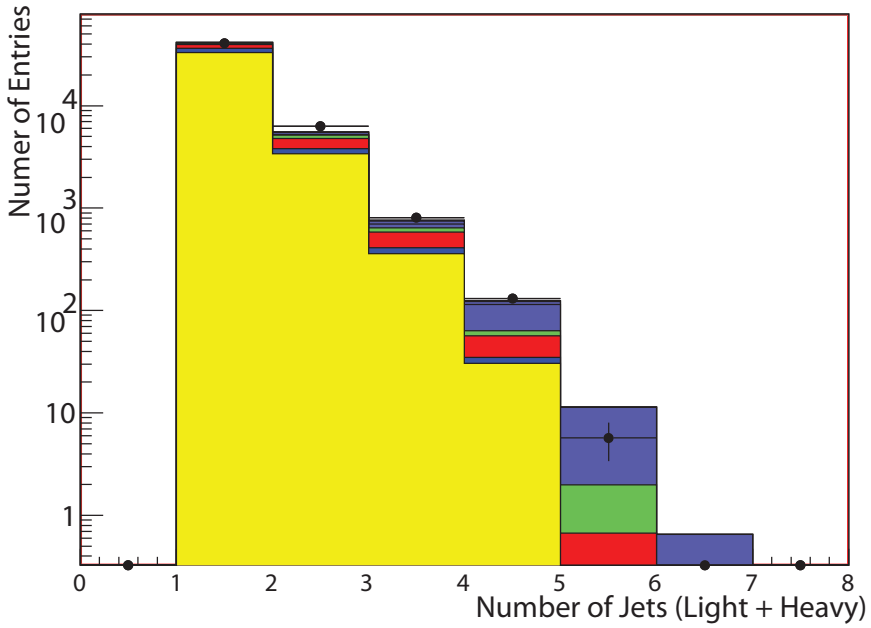


(b)

Figure 7.1: μ + jets final state fitting histograms: μ p_T, $\not\!\!E_T$.

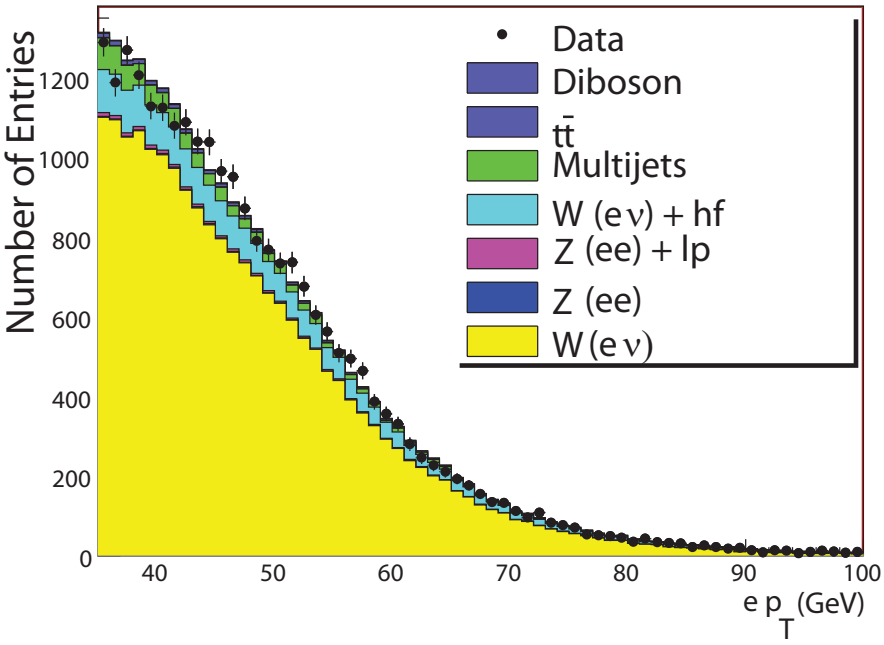


(a)

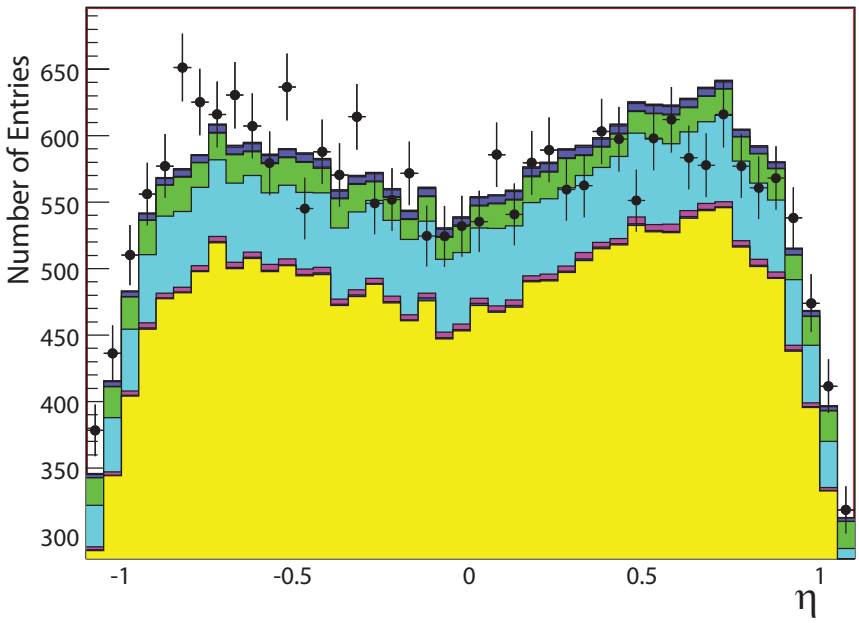


(b)

Figure 7.2: μ + jets final state checking histograms: transverse mass ($\mu, \not\!\!E_T$), number of jets.

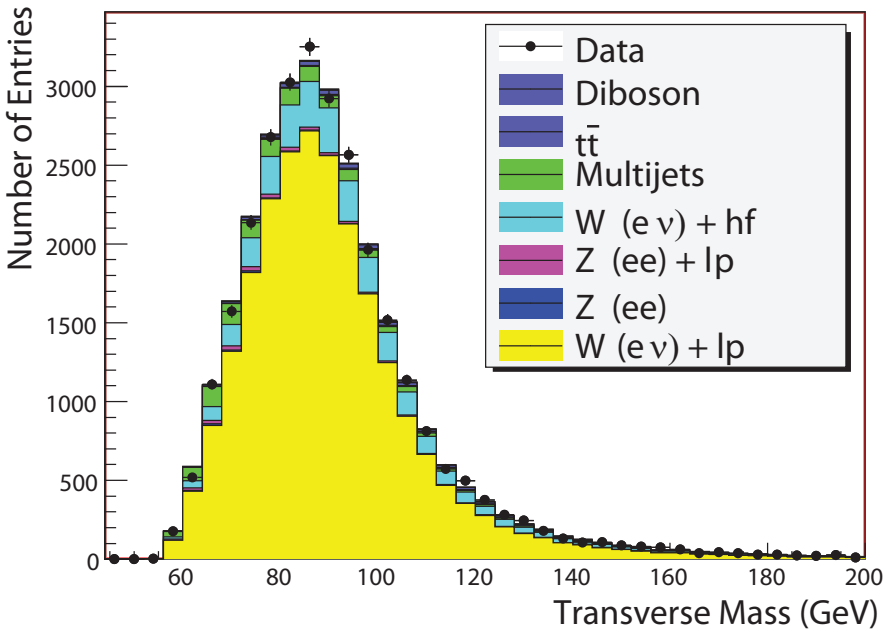


(a)

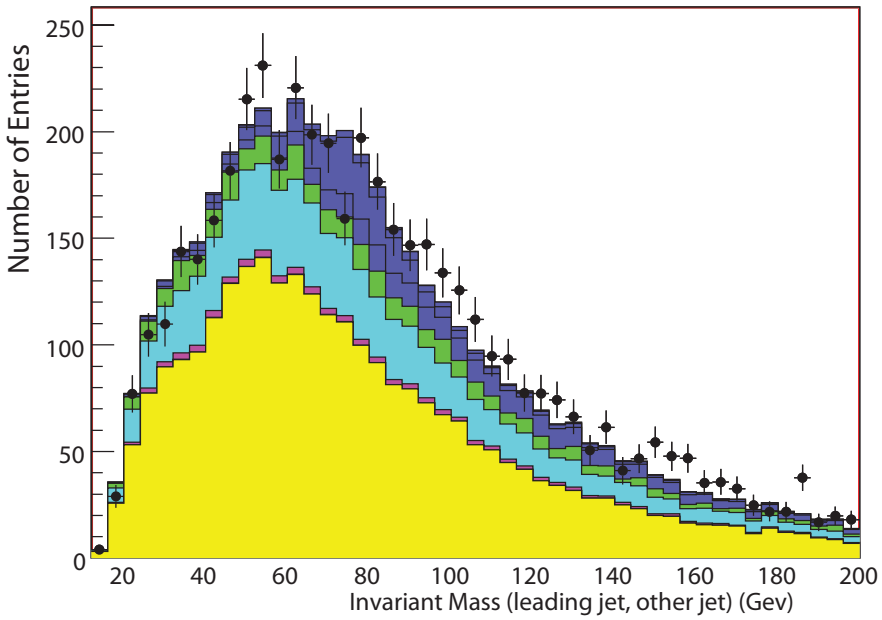


(b)

Figure 7.3: e + jets final state fitting histograms: e p_T, e η.

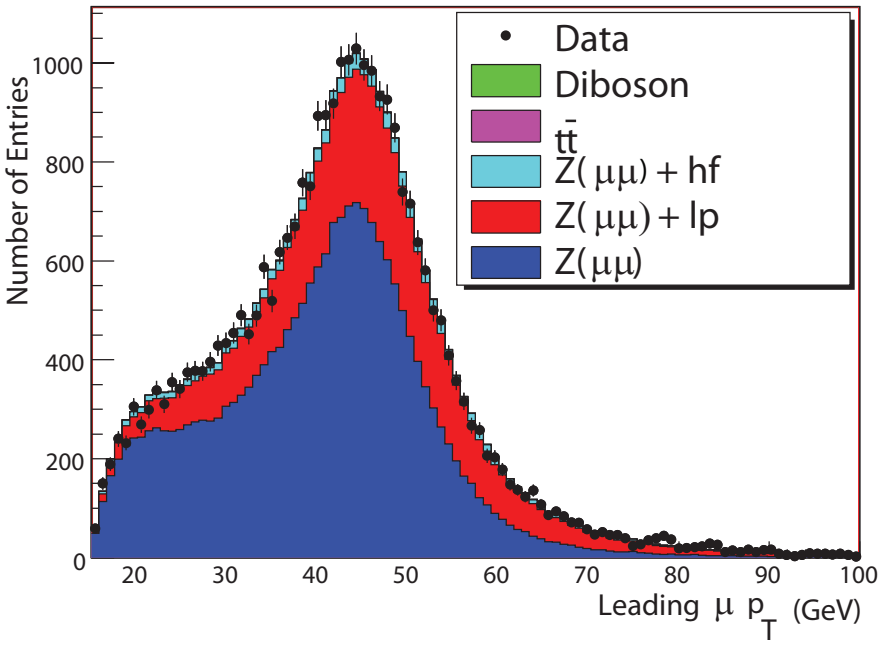


(a)

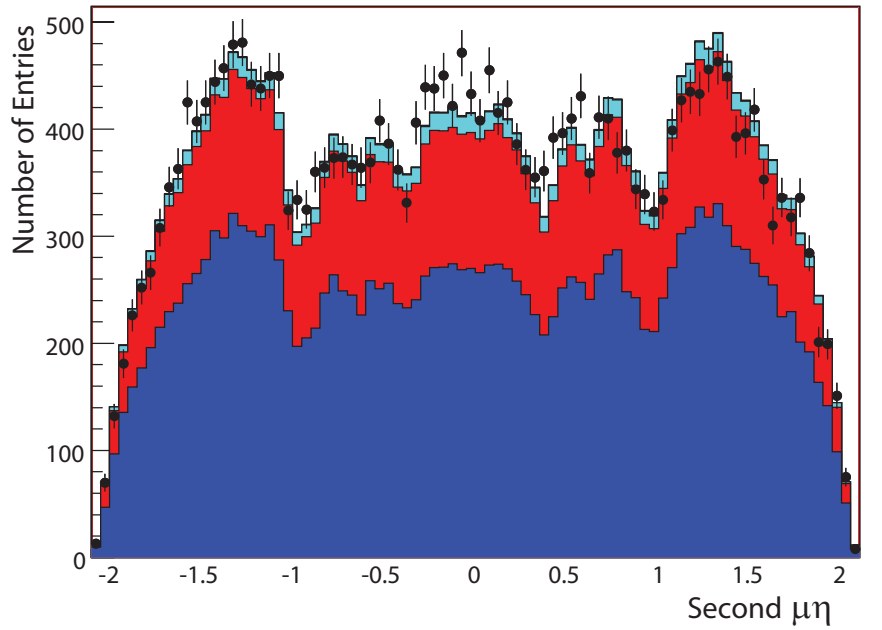


(b)

Figure 7.4: e + jets final state checking histograms: transverse mass (e, $\not\!\!E_T$), invariant mass of leading jet with other jets in the event.

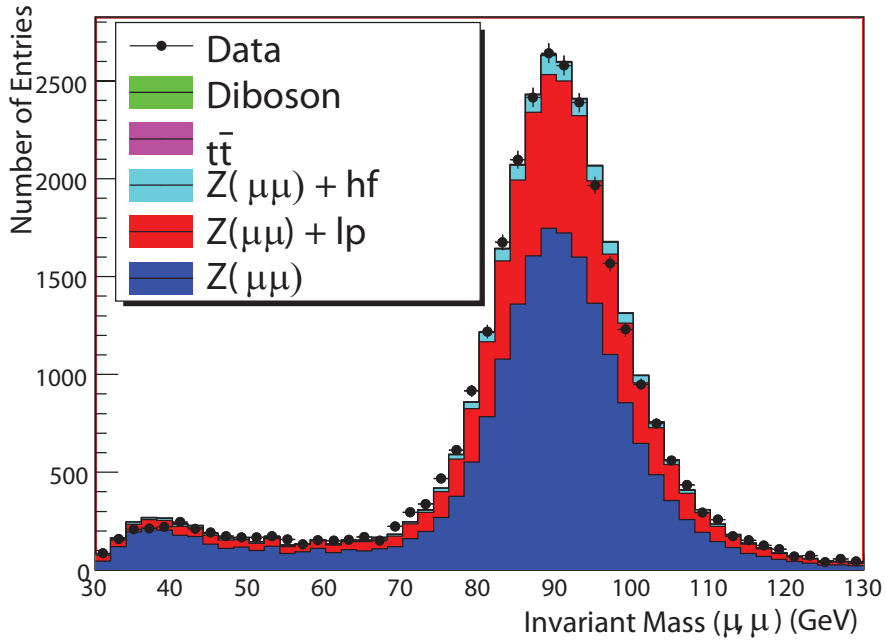


(a)

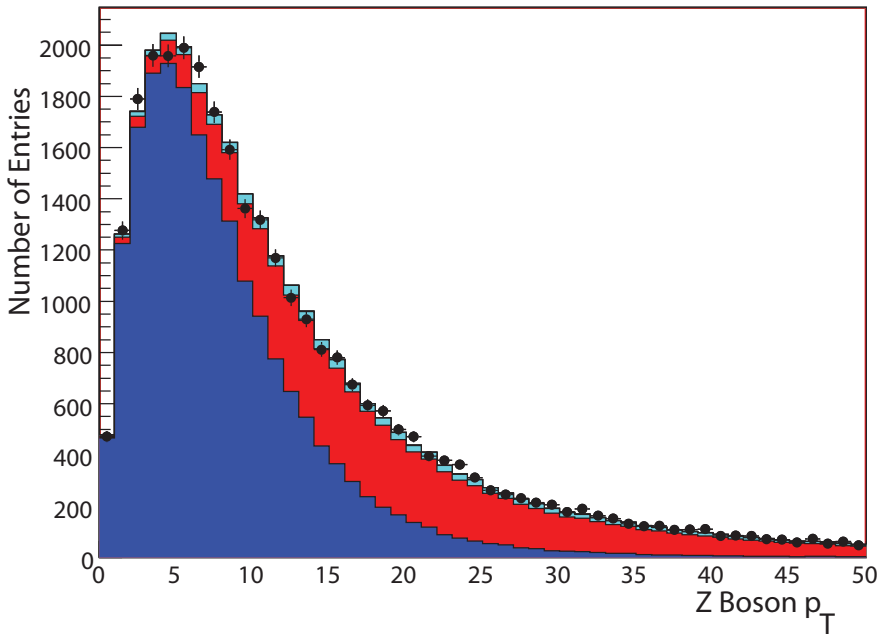


(b)

Figure 7.5: $\mu\mu$ final state fitting histograms: leading μ p_T, second μ η.

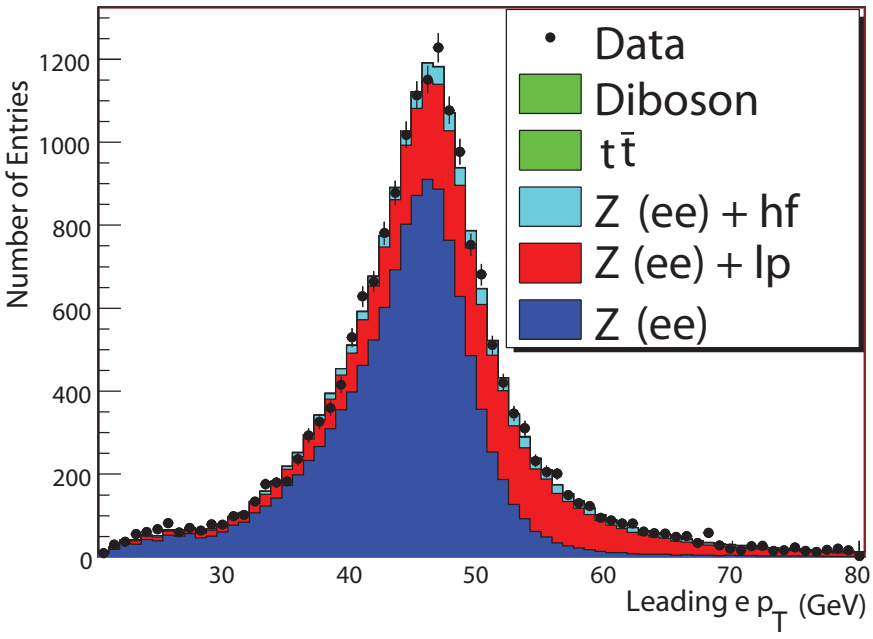


(a)

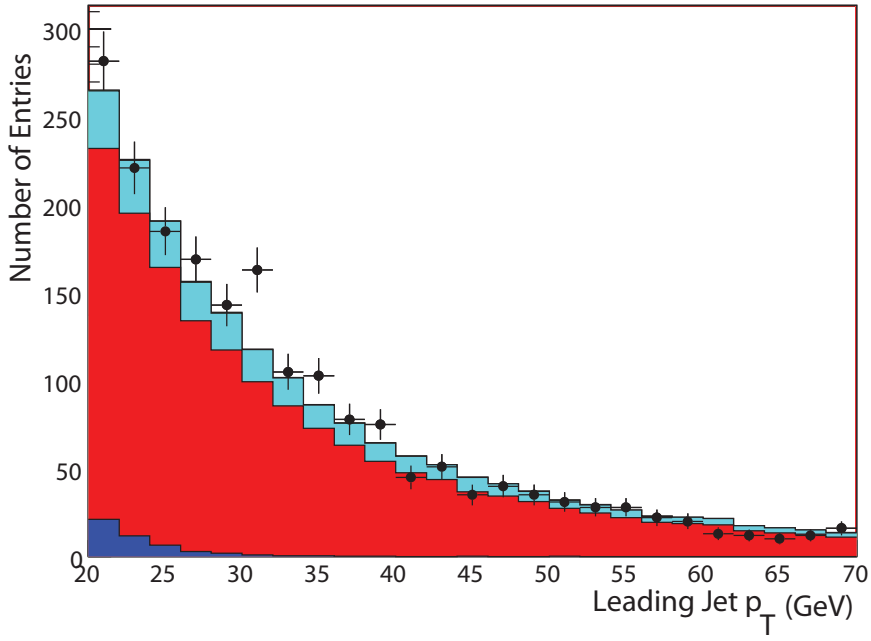


(b)

Figure 7.6: $\mu\mu$ final state checking histograms: invariant mass (μ, μ), Z boson p_T.

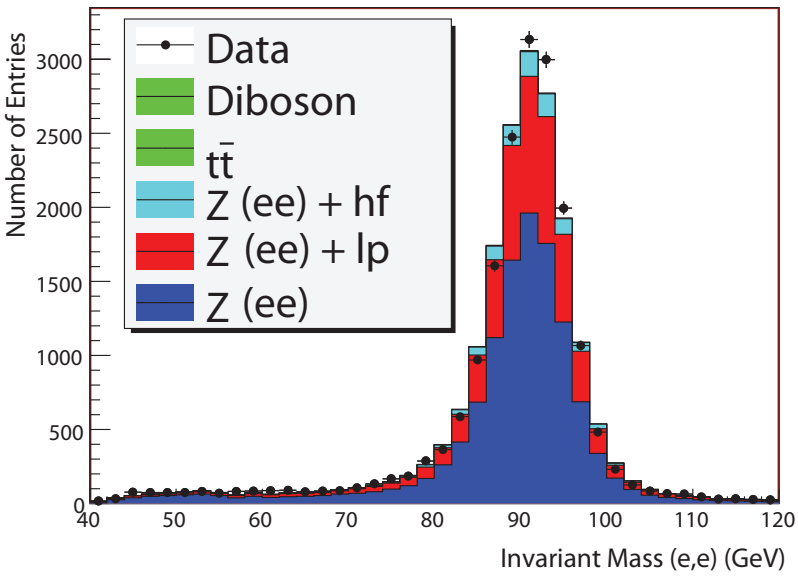


(a)

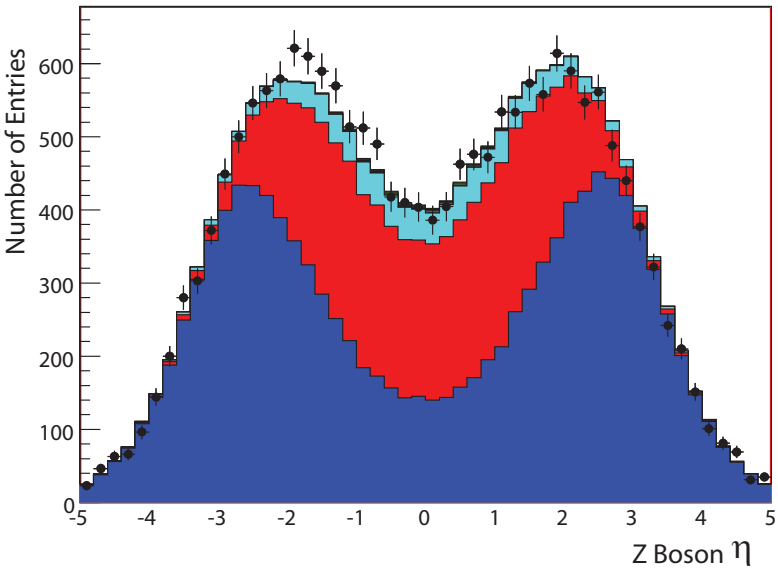


(b)

Figure 7.7: ee final state fitting histograms: leading e p_T, leading jet p_T.

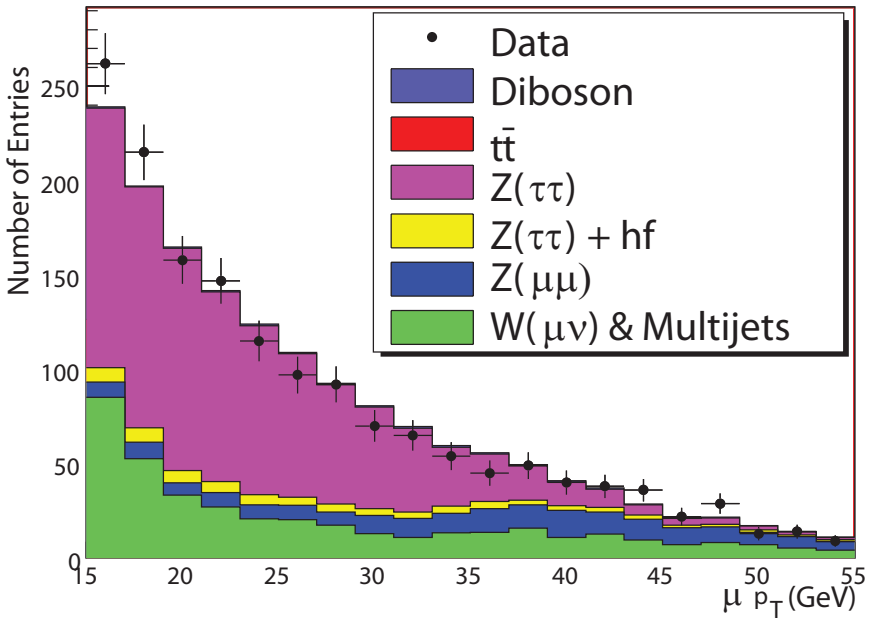


(a)

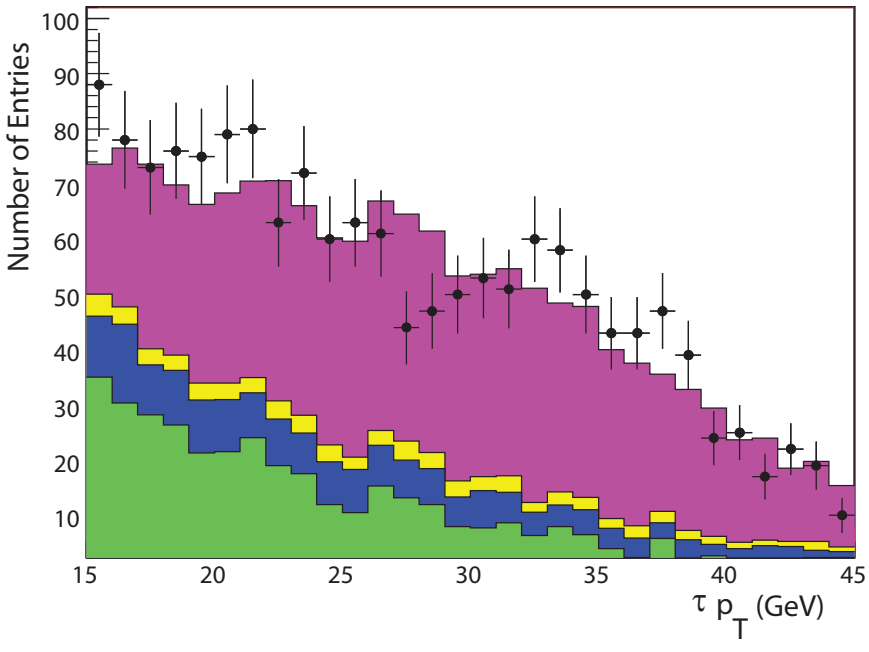


(b)

Figure 7.8: ee final state checking histograms: invariant mass (e,e), Z boson η. Several distributions, such as the Z η show some several bin discrepancies. When the trials-corrected probabilities are determined for these discrepancies, the significance is shown to be at the level of one sigma. We work to generally improve the standard model background modeling, but the focus is on statistically significant discrepancies.

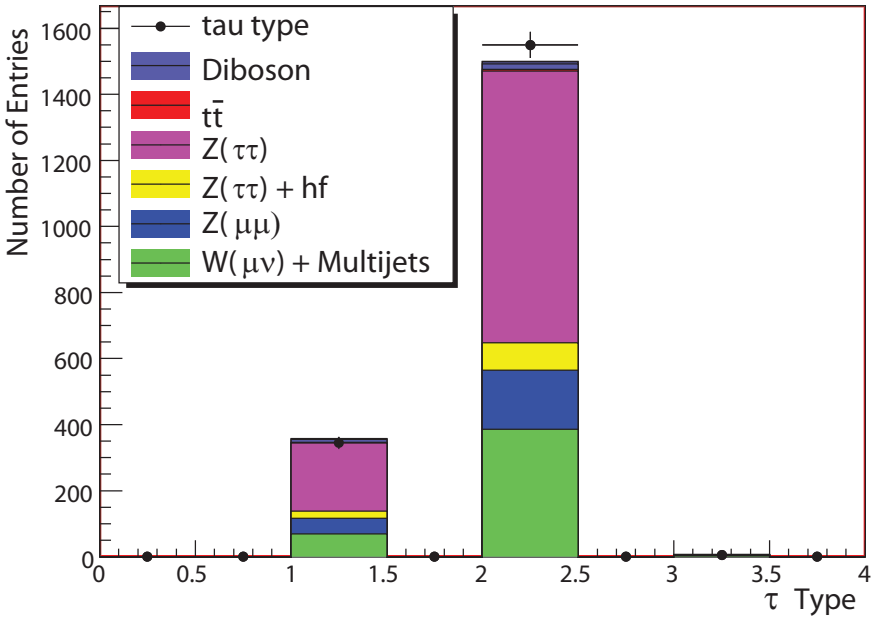


(a)

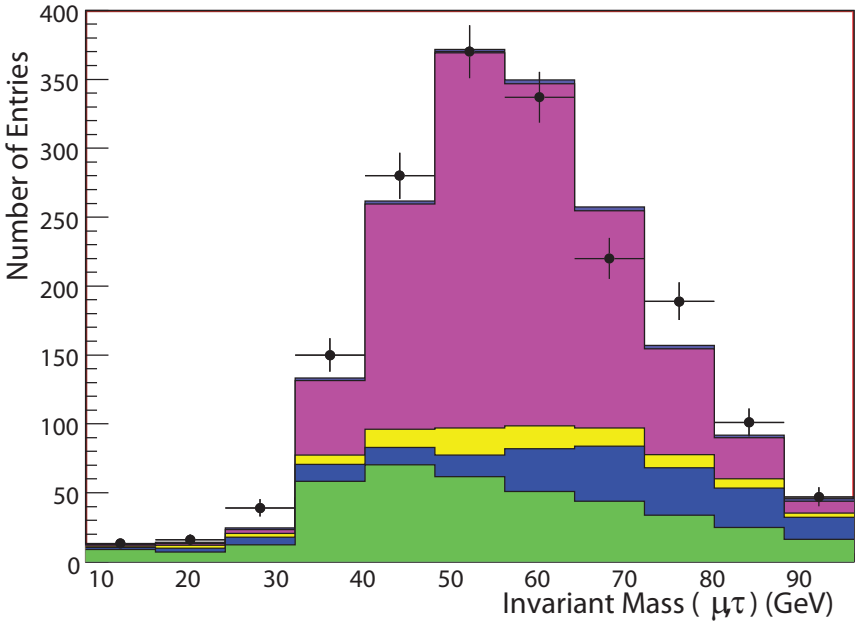


(b)

Figure 7.9: $\mu\tau$ final state fitting histograms: μ p_T, τ p_T.

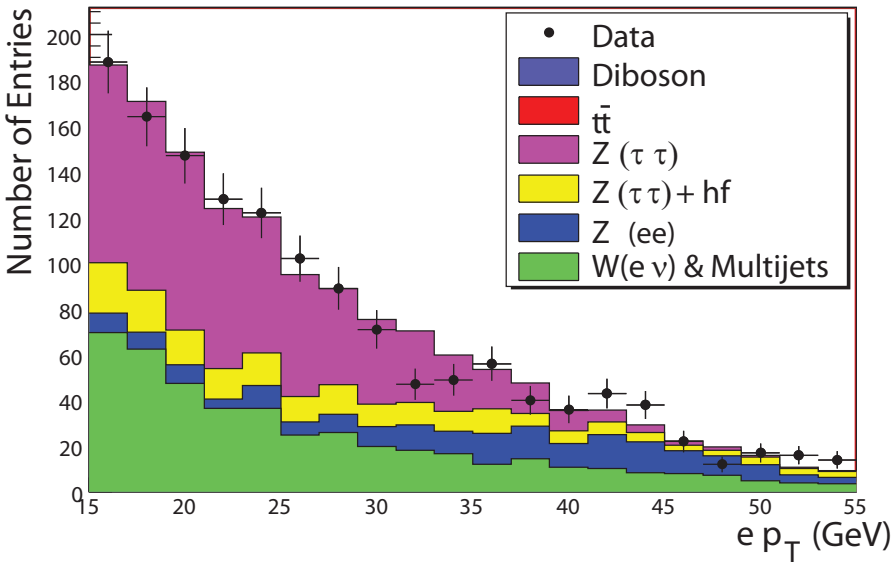


(a)

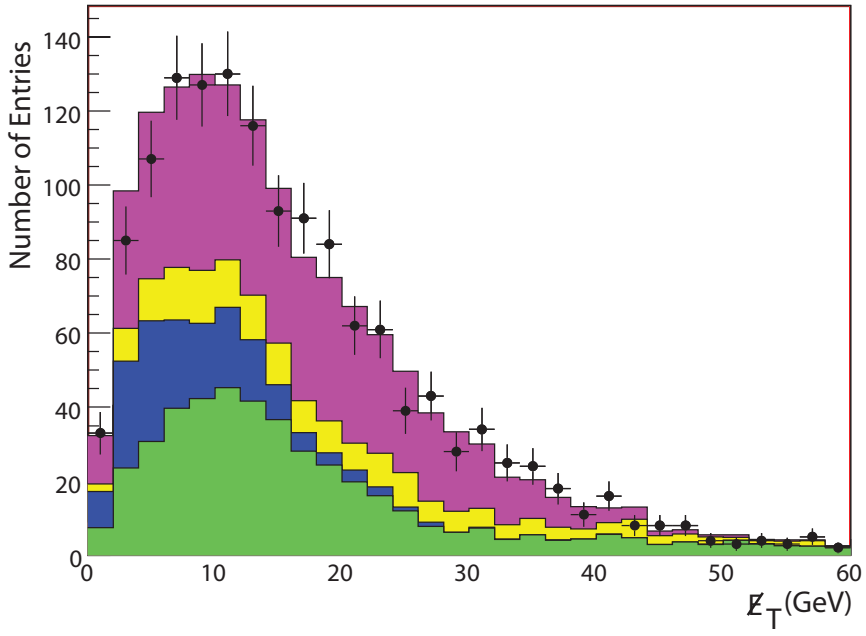


(b)

Figure 7.10: $\mu\tau$ final state checking histograms: τ type, invariant mass (μ, τ). Low values of invariant mass show single-bin discrepancies in the three final states that are dominated by the Drell-Yan $\tau\tau$ process. These are related to p_T threshold issues with the taus, and when accounting for trials are not statistically significant.

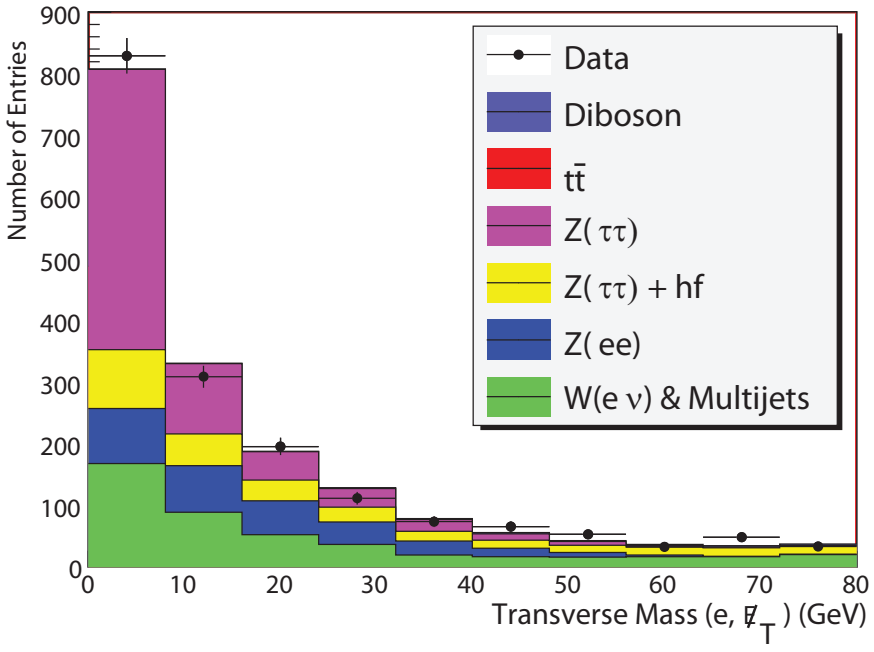


(a)

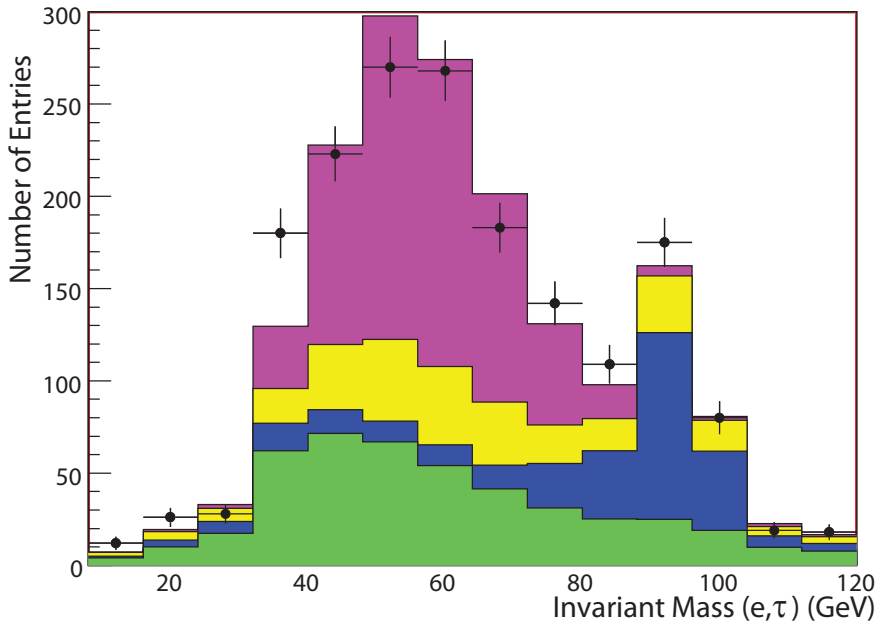


(b)

Figure 7.11: eτ final state fitting histograms: e p_T, $\not\!\!E_T$.

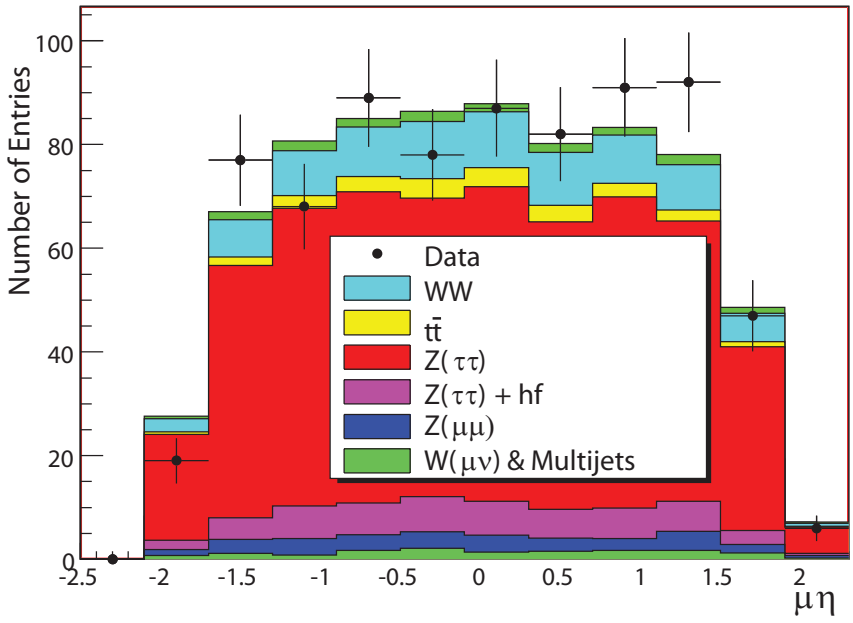


(a)

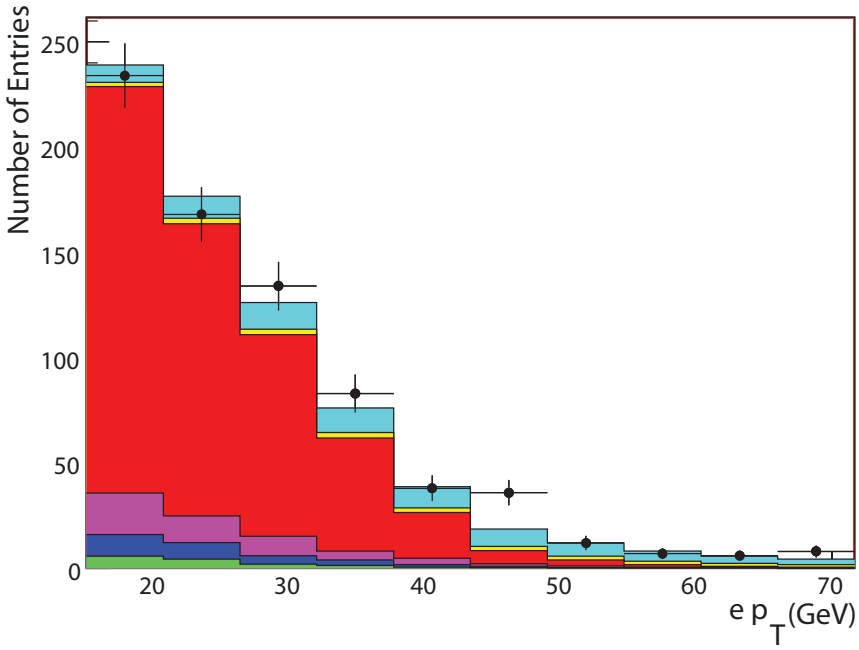


(b)

Figure 7.12: eτ final state checking histograms: transverse mass (e,$\not\!\!E_T$), invariant mass (e,τ).

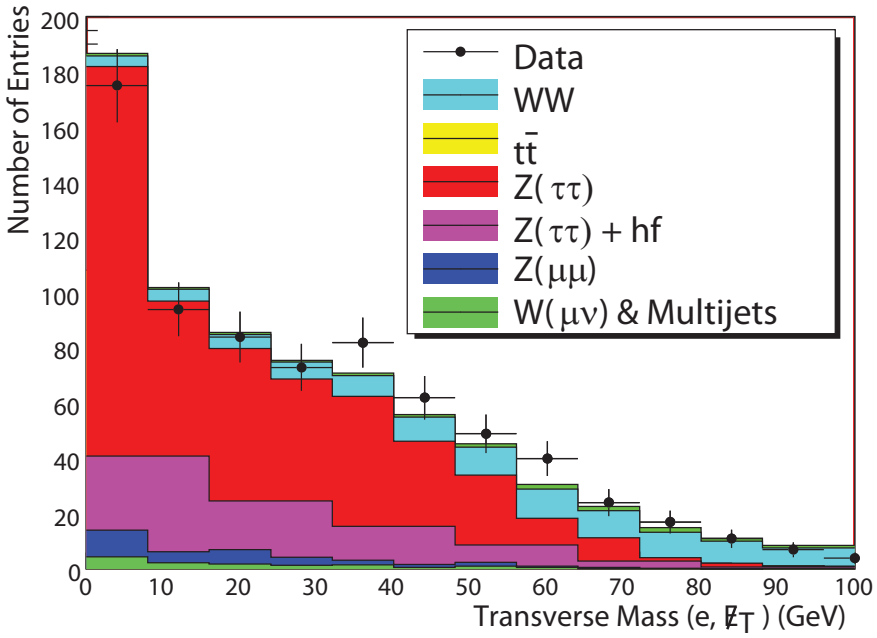


(a)

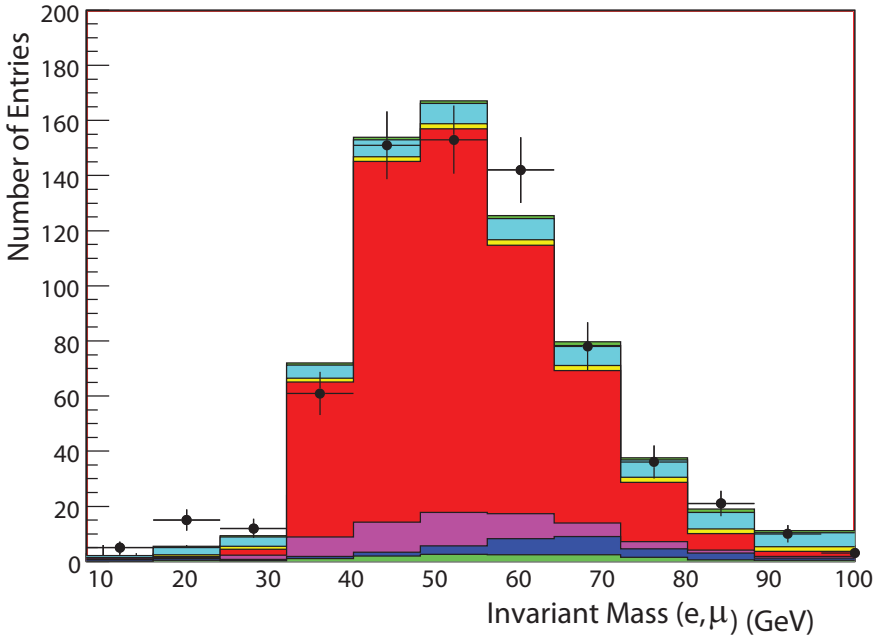


(b)

Figure 7.13: μe final state fitting histograms: μ η, e p_T.



(a)



(b)

Figure 7.14: μe final state checking histograms: transverse mass (e,$\not\!\!E_T$), invariant mass (e,μ).

A final fit with these conditions imposed can be seen for the dimuon state in Figure 7.15.

$$SF = SF_{ee}\frac{\#\text{events}(ee)}{\#\text{events}(ee+\mu\mu)} + SF_{\mu\mu}\frac{\#\text{events}(\mu\mu)}{\#\text{events}(ee+\mu\mu)} \tag{7.2}$$

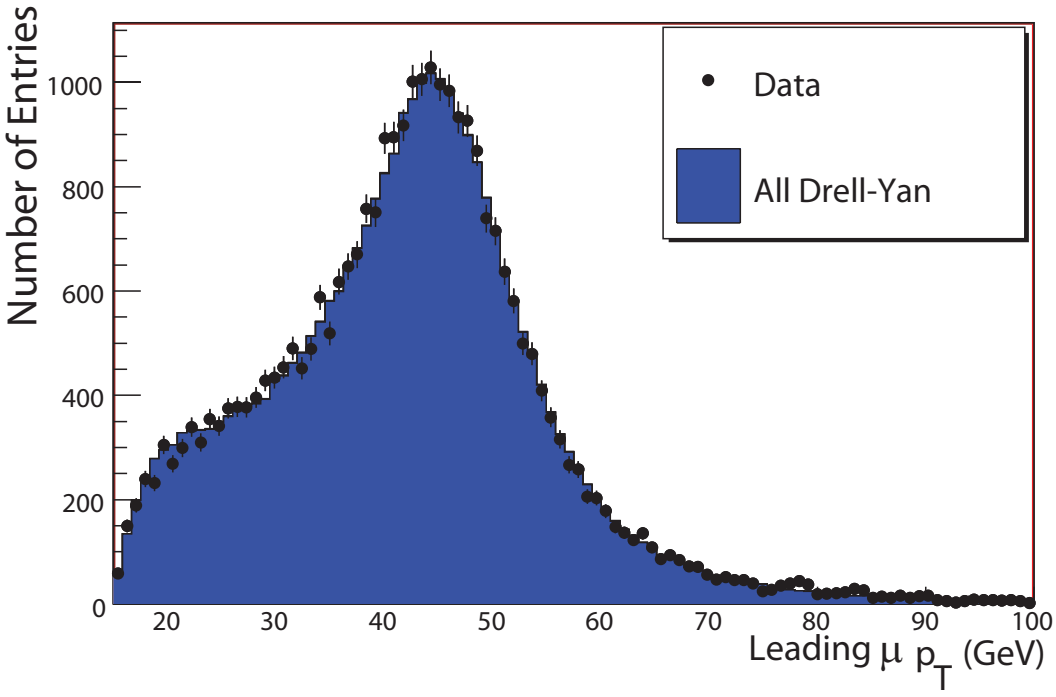


Figure 7.15: One final fit is performed after fixing the ratios of light-parton to no-parton and heavy-flavor to light-parton. This figure shows the dimuon final state after these ratios are averaged with the dielectron and then fixed.

A final check is made to see the effect of the inclusion of the rare final states, $t\bar{t}$ and diboson. The $t\bar{t}$ contribution can be checked in histograms such as the 4-jet HT as shown in Figure 7.16. The diboson contribution can be seen in many of the histograms in the μe final state.

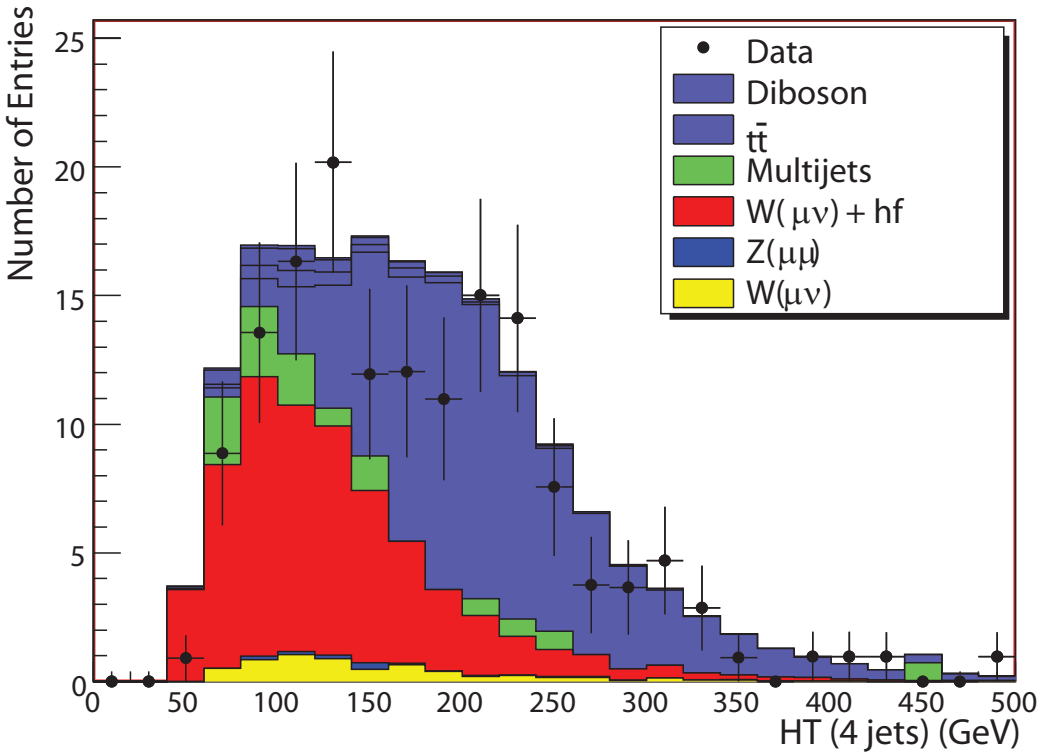


Figure 7.16: The scalar sum of the transverse energy of jets in μ + jets events with at least 4 jets. This final state shows the necessity of $t\bar{t}$ Monte Carlo to properly describe the data.

7.3 Text File Production

Once all of the normalization weights have been determined and the input processes checked for agreement, the input files for the experiment-independent VISTA and SLEUTH algorithms are created. These algorithms take text file inputs which only contain the most basic information about the objects. The overall event weight, run/event number and vertex position are kept along the object p_T, η and ϕ. Using this simple information, the algorithms quantify the overall agreement.

The text files are created in the same way that the histograms were created for the fit. The same computer code is used in their production, with the addition of one input weight that comes from the inclusive state normalization fits.

An example of one line of a $\mu\tau$ text file can be seen in Figure 7.17.

```
mutau12_ztt 1.113 0.000239137  ppbar  1960  409.057 13.9419
mu- 61.563 0.135986 139.865  tau+ 162.489 -1.20636 317.824  j 0
21.319 -1.64867 131.604  uncl  0 41.058 0 -37.7612    ;
```

Figure 7.17: The figure shows one line of a $\mu\tau$ text file used as input into the VISTA algorithm. Only the run and event numbers, the vertex position, weight, and the object p_T, η, ϕ information are kept for each event. In the figure, each object is shown in a different color.

Chapter 8

VISTA and SLEUTH

Once all of the event selection cuts and correction factors are determined in the MIS analysis packages, VISTA and SLEUTH consolidate this information and use it to quantify the agreement between the data and the standard model background. The SLEUTH algorithm was developed for DØ Run I [87, 88, 89] and later modified at the H1 experiment [90]. The VISTA algorithm was developed at CDF [91, 28]. All of the algorithms are discussed in detail in [92].

8.1 VISTA

VISTA is an experiment-independent program developed by the CDF experiment to compare event counts and 1-D histogram shapes between data and the standard model expectation, while explicitly taking into account the trials factor associated with the number of places checked. While the original algorithm provided the capability to make refinements in object identification, event selection, and correction factors, the DØ version of VISTA removes this functionality and only provides the quantitative comparison between the samples.

VISTA checks overall event counts and Kolmogorov-Smirnov probabilities, so it is most sensitive to differences in the central parts (not tails) of distributions. This

would provide sensitivity to new physics with large cross-sections or modeling issues affecting variables across final states. The SLEUTH algorithm is most sensitive to the possibility of new physics in the tails of distributions.

8.1.1 Exclusive Final States

The seven final states used to create the input text files described in the last chapter are fully defined in the input to VISTA. The objects defined in each event are then used to place each of the events into a particular final state. For example, if the event contained a muon and two jets, this would go into a separate final state from an event with a muon and three jets. The object content completely defines the exclusive final states used for event counts and histogram checking in VISTA. The full list of final states with the event counts for each final state and trials corrected discrepancies measured in units of equivalent Gaussian standard deviation are shown in Table 8.1.

Table 8.1: The 180 VISTA final states.

VISTA Final State	Number of Data Events	Expected Background	σ
4 j $\mu^{\pm}$ $\not\!\!E_T$ $\tau^{\mp}$	1	1.5±1.3	0
4 j $\mu^{\pm}$ $\not\!\!E_T$	99	83.7±3.2	0
4 j $\mu^{\pm}$ $\mu^{\mp}$ $\not\!\!E_T$	2	1.7±1.2	0
4 j $\mu^{\pm}$ $\mu^{\mp}$	2	1.5±1.2	0
3 j $\mu^{\pm}$ $\tau^{\mp}$	1	0.8±1.1	0
3 j $\mu^{\pm}$ $\not\!\!E_T$ $\tau^{\mp}$	3	10.6±2	0
3 j $\mu^{\pm}$ $\not\!\!E_T$	750	684±8.5	0
3 j $\mu^{\pm}$ γ $\not\!\!E_T$	6	3.3±1.8	0
3 j $\mu^{\pm}$ $\mu^{\mp}$ $\not\!\!E_T$	15	21.9±1.6	0
3 j $\mu^{\pm}$ $\mu^{\mp}$ γ	1	0.2±1	0

Continued on Next Page...

Table 8.1 (cont'd)

VISTA Final State	Number of Data Events	Expected Background	σ
3 j $\mu^{\pm}$ $\mu^{\mp}$	25	24.9±1.7	0
2 $\mu^{\pm}$ $\not\!\!E_T$	11	5.6±1.6	0
2 $\mu^{\pm}$ $\mu^{\mp}$ $\not\!\!E_T$	1	1.2±1.1	0
2 $\mu^{\pm}$	1	1.4±1.2	0
2 j $\mu^{\pm}$ $\tau^{\pm}$	2	5.3±1.9	0
2 j $\mu^{\pm}$ $\tau^{\mp}$	11	10.4±1.9	0
2 j $\mu^{\pm}$ $\not\!\!E_T$ $\tau^{\mp}$	55	46.3±3	0
2 j $\mu^{\pm}$ $\not\!\!E_T$ $\tau^{\pm}$	14	19.8±2.8	0
2 j $\mu^{\pm}$ $\not\!\!E_T$	6213	5362±29.2	+9.3
2 j $\mu^{\pm}$ γ $\not\!\!E_T$ $\tau^{\mp}$	1	0.8±1.1	0
2 j $\mu^{\pm}$ γ $\not\!\!E_T$	34	18±2.4	0
2 j $\mu^{\pm}$ $\mu^{\mp}$ $\not\!\!E_T$	118	136.6±2.5	0
2 j $\mu^{\pm}$ $\mu^{\mp}$ γ $\not\!\!E_T$	2	0.7±1.1	0
2 j $\mu^{\pm}$ $\mu^{\mp}$ γ	1	1.3±1.1	0
2 j $\mu^{\pm}$ $\mu^{\mp}$	216	225±3.2	0
2 $e^{\pm}$ 2 j γ $\not\!\!E_T$	1	0±1	0
2 $e^{\pm}$ 2 j	2	0.9±1.1	0
2 $e^{\pm}$ $\not\!\!E_T$	3	3.4±1.4	0
2 $e^{\pm}$ γ	2	1.6±1.4	0
2 $e^{\pm}$ j	5	7±1.5	0
2 $e^{\pm}$ $e^{\mp}$ $\not\!\!E_T$	1	1.2±1.1	0
2 $e^{\pm}$ $e^{\mp}$ γ $\not\!\!E_T$	1	0±1	0
2 $e^{\pm}$	28	24.6±2.2	0

Continued on Next Page...

Table 8.1 (cont'd)

VISTA Final State	Number of Data Events	Expected Background	σ
2 b 2 j $\mu^{\pm}$ $\not\!\!E_T$	11	17.5±2	0
2 b $\mu^{\pm}$ $\not\!\!E_T$	26	29.3±2.1	0
2 b $\mu^{\pm}$ γ $\not\!\!E_T$	1	0.2±1	0
2 b j $\mu^{\pm}$ $\not\!\!E_T$	15	15.3±1.3	0
2 b $e^{\pm}$ 2 j $\not\!\!E_T$	6	14.4±1.4	0
2 b $e^{\pm}$ $\not\!\!E_T$	18	20.5±1.7	0
2 b $e^{\pm}$ $\mu^{\mp}$ $\not\!\!E_T$	5	3±1	0
2 b $e^{\pm}$ j $\not\!\!E_T$	11	13.2±1.4	0
$\mu^{\pm}$ $\tau^{\pm}$	194	138.8±6.1	+2
$\mu^{\pm}$ $\tau^{\mp}$	828	843.5±10.1	0
$\mu^{\pm}$ $\not\!\!E_T$ $\tau^{\pm}$	273	285.7±7.9	0
$\mu^{\pm}$ $\not\!\!E_T$ $\tau^{\mp}$	1239	1208.8±12.4	0
$\mu^{\pm}$ γ $\tau^{\mp}$	39	29.7±2.5	0
$\mu^{\pm}$ γ $\not\!\!E_T$ $\tau^{\pm}$	6	5.1±2	0
$\mu^{\pm}$ γ $\not\!\!E_T$ $\tau^{\mp}$	35	23.7±2.5	0
$\mu^{\pm}$ $\mu^{\mp}$ 2 γ	1	0.2±1.1	0
$\mu^{\pm}$ $\mu^{\mp}$ $\tau^{\pm}$	18	29.2±2.7	0
$\mu^{\pm}$ $\mu^{\mp}$ $\not\!\!E_T$ $\tau^{\pm}$	19	30±2.8	0
$\mu^{\pm}$ $\mu^{\mp}$ $\not\!\!E_T$	3559	3195.3±21.4	+4.7
$\mu^{\pm}$ $\mu^{\mp}$ γ $\not\!\!E_T$	29	13.4±1.8	+0.6
$\mu^{\pm}$ $\mu^{\mp}$ γ	178	111.6±3.5	+4.2
$\mu^{\pm}$ $\mu^{\mp}$	22801	23048.6±136.8	0
j 2 $\mu^{\pm}$ $\not\!\!E_T$	1	0.8±1.2	0

Continued on Next Page...

Table 8.1 (cont'd)

VISTA Final State	Number of Data Events	Expected Background	σ
j 2 $\mu^{\pm}$	1	0.1±1	0
j $\mu^{\pm}$ $\tau^{\mp}$	83	90.7±3.9	0
j $\mu^{\pm}$ $\tau^{\pm}$	33	29.1±3.2	0
j $\mu^{\pm}$ $\not\!\!E_T$ $\tau^{\pm}$	77	97.2±4.9	0
j $\mu^{\pm}$ $\not\!\!E_T$ $\tau^{\mp}$	225	260.4±6.1	0
j $\mu^{\pm}$ $\not\!\!E_T$	41154	41816.7±222.5	0
j $\mu^{\pm}$ γ $\not\!\!E_T$ $\tau^{\mp}$	7	5.3±1.6	0
j $\mu^{\pm}$ γ $\not\!\!E_T$ $\tau^{\pm}$	2	1.6±1.6	0
j $\mu^{\pm}$ γ $\not\!\!E_T$	197	101.8±4.5	+6.6
j $\mu^{\pm}$ $\mu^{\mp}$ $\tau^{\pm}$	2	7.4±1.9	0
j $\mu^{\pm}$ $\mu^{\mp}$ $\not\!\!E_T$ $\tau^{\pm}$	2	8.1±1.9	0
j $\mu^{\pm}$ $\mu^{\mp}$ $\not\!\!E_T$	871	758.6±5.7	+2
j $\mu^{\pm}$ $\mu^{\mp}$ γ $\not\!\!E_T$	3	3.8±1.3	0
j $\mu^{\pm}$ $\mu^{\mp}$ γ	29	12±1.5	+1.6
j $\mu^{\pm}$ $\mu^{\mp}$	2070	2124±13.6	0
$e^{\pm}$ 4 j $\not\!\!E_T$	69	69.4±2.3	0
$e^{\pm}$ 3 j $\not\!\!E_T$ $\tau^{\mp}$	2	2.8±1.4	0
$e^{\pm}$ 3 j $\not\!\!E_T$	457	439.7±5	0
$e^{\pm}$ 3 j $\mu^{\mp}$ $\not\!\!E_T$	1	1.7±1.1	0
$e^{\pm}$ 2 j $\tau^{\mp}$	8	8.7±1.9	0
$e^{\pm}$ 2 j $\tau^{\pm}$	3	3.9±1.8	0
$e^{\pm}$ 2 j $\not\!\!E_T$ $\tau^{\pm}$	5	6.3±2	0
$e^{\pm}$ 2 j $\not\!\!E_T$ $\tau^{\mp}$	21	19.5±2.3	0

Continued on Next Page...

Table 8.1 (cont'd)

VISTA Final State	Number of Data Events	Expected Background	σ
$e^{\pm}$ 2 j $\not\!\!E_T$	3627	3479.8±18.5	0
$e^{\pm}$ 2 j γ $\tau^{\mp}$	1	0.5±1.1	0
$e^{\pm}$ 2 j γ $\not\!\!E_T$ $\tau^{\mp}$	1	1.2±1.3	0
$e^{\pm}$ 2 j $\mu^{\mp}$ $\not\!\!E_T$	7	11.6±1.4	0
$e^{\pm}$ 2 j $\mu^{\pm}$ $\not\!\!E_T$	2	0.4±1.1	0
$e^{\pm}$ 2 j $\mu^{\mp}$	1	1.8±1.2	0
$e^{\pm}$ $\tau^{\pm}$	167	121.6±5.6	+1.1
$e^{\pm}$ $\tau^{\mp}$	612	651.1±9.3	0
$e^{\pm}$ $\not\!\!E_T$ $\tau^{\mp}$	556	532.5±8	0
$e^{\pm}$ $\not\!\!E_T$ $\tau^{\pm}$	111	96.4±5	0
$e^{\pm}$ γ $\tau^{\pm}$	5	1.5±1.5	0
$e^{\pm}$ γ $\tau^{\mp}$	38	38.3±3	0
$e^{\pm}$ γ $\not\!\!E_T$ $\tau^{\mp}$	8	10.9±1.9	0
$e^{\pm}$ γ $\not\!\!E_T$ $\tau^{\pm}$	1	1.4±1.5	0
$e^{\pm}$ $\mu^{\mp}$ $\tau^{\mp}$	1	0.7±1.2	0
$e^{\pm}$ $\mu^{\pm}$ $\tau^{\pm}$	1	0.4±1.3	0
$e^{\pm}$ $\mu^{\mp}$ $\not\!\!E_T$ $\tau^{\mp}$	1	1±1.2	0
$e^{\pm}$ $\mu^{\mp}$ $\not\!\!E_T$	204	208.6±4.4	0
$e^{\pm}$ $\mu^{\pm}$ $\not\!\!E_T$	19	8.5±1.6	0
$e^{\pm}$ $\mu^{\mp}$ γ $\not\!\!E_T$	5	5.5±1.6	0
$e^{\pm}$ $\mu^{\pm}$ γ $\not\!\!E_T$	3	0.7±1.2	0
$e^{\pm}$ $\mu^{\mp}$ γ	11	11.6±2	0
$e^{\pm}$ $\mu^{\pm}$ γ	1	0.3±1.2	0

Continued on Next Page...

Table 8.1 (cont'd)

VISTA Final State	Number of Data Events	Expected Background	σ
$e^{\pm}\ \mu^{\pm}\ \mu^{\mp}\ \not\!\!E_T$	5	5.3±1.4	0
$e^{\pm}\ \mu^{\pm}\ \mu^{\mp}$	8	9.1±1.6	0
$e^{\pm}\ \mu^{\mp}$	343	350.3±6.3	0
$e^{\pm}\ \mu^{\pm}$	16	6.2±1.6	0
$e^{\pm}$ j $\tau^{\mp}$	65	71.3±3.6	0
$e^{\pm}$ j $\tau^{\pm}$	30	24.2±3	0
$e^{\pm}$ j $\not\!\!E_T\ \tau^{\pm}$	39	26.6±3.1	0
$e^{\pm}$ j $\not\!\!E_T\ \tau^{\mp}$	112	116.8±4.2	0
$e^{\pm}$ j $\not\!\!E_T$	24482	24817.5±128.3	0
$e^{\pm}$ j $\gamma\ \tau^{\mp}$	2	5±1.4	0
$e^{\pm}$ j $\gamma\ \not\!\!E_T\ \tau^{\mp}$	4	3.5±1.4	0
$e^{\pm}$ j $\mu^{\mp}\ \not\!\!E_T$	64	54.8±2.3	0
$e^{\pm}$ j $\mu^{\pm}\ \not\!\!E_T$	6	2.7±1.3	0
$e^{\pm}$ j $\mu^{\mp}\ \gamma\ \not\!\!E_T$	2	1.7±1.2	0
$e^{\pm}$ j $\mu^{\mp}\ \gamma$	1	0.3±1.1	0
$e^{\pm}$ j $\mu^{\pm}\ \mu^{\mp}$	3	0.9±1.1	0
$e^{\pm}$ j $\mu^{\mp}$	14	16.3±1.8	0
$e^{\pm}$ j $\mu^{\pm}$	1	0.7±1.2	0
$e^{\pm}\ e^{\mp}$ 4 j $\not\!\!E_T$	1	0.6±1.1	0
$e^{\pm}\ e^{\mp}$ 4 j	1	2.4±1.2	0
$e^{\pm}\ e^{\mp}$ 3 j $\not\!\!E_T$	4	4.3±1.3	0
$e^{\pm}\ e^{\mp}$ 3 j γ	1	0.5±1.1	0
$e^{\pm}\ e^{\mp}$ 3 j	25	27±1.6	0

Continued on Next Page...

Table 8.1 (cont'd)

VISTA Final State	Number of Data Events	Expected Background	σ
$e^{\pm}\ e^{\mp}$ 2 j $\not\!\!E_T$	23	21±1.7	0
$e^{\pm}\ e^{\mp}$ 2 j γ	4	5.3±1.3	0
$e^{\pm}\ e^{\mp}$ 2 j	242	247.4±3.5	0
$e^{\pm}\ e^{\mp}\ \not\!\!E_T$	180	169±5.1	0
$e^{\pm}\ e^{\mp}\ \gamma\ \not\!\!E_T$	6	6±1.8	0
$e^{\pm}\ e^{\mp}\ \gamma$	254	270.8±5.6	0
$e^{\pm}\ e^{\mp}\ \mu^{\pm}\ \not\!\!E_T$	1	2.2±1.2	0
$e^{\pm}\ e^{\mp}\ \mu^{\pm}$	1	0.6±1.1	0
$e^{\pm}\ e^{\mp}$ j $\not\!\!E_T$	84	69.2±2.4	0
$e^{\pm}\ e^{\mp}$ j $\gamma\ \not\!\!E_T$	3	2±1.3	0
$e^{\pm}\ e^{\mp}$ j γ	35	38.9±2.1	0
$e^{\pm}\ e^{\mp}$ j	1854	1880.8±10.3	0
$e^{\pm}\ e^{\mp}$	16152	16083.1±105.1	0
b 3 j $\mu^{\pm}\ \not\!\!E_T$	31	37.5±2.1	0
b 3 j $\mu^{\pm}\ \mu^{\mp}\ \not\!\!E_T$	0	0.5±1	0
b 2 j $\mu^{\pm}\ \not\!\!E_T$	76	83.2±3.7	0
b 2 j $\mu^{\pm}\ \mu^{\mp}\ \not\!\!E_T$	2	3.2±1.2	0
b 2 j $\mu^{\pm}\ \mu^{\mp}$	1	2.7±1.2	0
b $\mu^{\pm}\ \not\!\!E_T$	620	702.1±10.6	0
b $\mu^{\pm}\ \gamma\ \not\!\!E_T$	3	4.2±2.1	0
b $\mu^{\pm}\ \mu^{\mp}\ \not\!\!E_T$	12	18.3±1.5	0
b $\mu^{\pm}\ \mu^{\mp}\ \gamma$	1	0.3±1.1	0
b $\mu^{\pm}\ \mu^{\mp}$	35	36±1.8	0

Continued on Next Page...

Table 8.1 (cont'd)

VISTA Final State	Number of Data Events	Expected Background	σ
b j $\mu^{\pm}$ $\not\!\!E_T$	266	306.3±7.3	0
b j $\mu^{\pm}$ γ $\not\!\!E_T$	3	1.2±1.1	0
b j $\mu^{\pm}$ $\mu^{\mp}$ $\not\!\!E_T$	7	13.1±1.3	0
b j $\mu^{\pm}$ $\mu^{\mp}$	8	12.8±1.4	0
b $e^{\pm}$ 3 j $\not\!\!E_T$	31	32±1.6	0
b $e^{\pm}$ 2 j $\not\!\!E_T$ $\tau^{\mp}$	0	0.7±1	0
b $e^{\pm}$ 2 j $\not\!\!E_T$	67	56.8±2	0
b $e^{\pm}$ 2 j $\mu^{\mp}$ $\not\!\!E_T$	6	1.7±1	0
b $e^{\pm}$ $\tau^{\mp}$	0	1.7±1	0
b $e^{\pm}$ $\tau^{\pm}$	0	0.5±1	0
b $e^{\pm}$ $\not\!\!E_T$ $\tau^{\mp}$	2	3.1±1.4	0
b $e^{\pm}$ $\not\!\!E_T$ $\tau^{\pm}$	0	0.6±1	0
b $e^{\pm}$ $\not\!\!E_T$	414	423.2±6.6	0
b $e^{\pm}$ $\mu^{\mp}$ $\not\!\!E_T$	5	3.5±1.1	0
b $e^{\pm}$ $\mu^{\mp}$	1	0.3±1	0
b $e^{\pm}$ j $\tau^{\mp}$	0	0.6±1	0
b $e^{\pm}$ j $\not\!\!E_T$ $\tau^{\mp}$	6	2.9±1.5	0
b $e^{\pm}$ j $\not\!\!E_T$	187	187.5±3.7	0
b $e^{\pm}$ j $\mu^{\mp}$ $\not\!\!E_T$	1	6.7±1.1	0
b $e^{\pm}$ $e^{\mp}$ 2 j	0	2.7±1	0
b $e^{\pm}$ $e^{\mp}$ $\not\!\!E_T$	5	3.6±1.2	0
b $e^{\pm}$ $e^{\mp}$ γ	0	0.7±1	0
b $e^{\pm}$ $e^{\mp}$ j $\not\!\!E_T$	0	5±1	0

Continued on Next Page...

Table 8.1 (cont'd)

VISTA Final State	Number of Data Events	Expected Background	σ
b $e^{\pm}$ $e^{\mp}$ j	6	16.5±1.4	0
b $e^{\pm}$ $e^{\mp}$	50	36.8±1.7	0

8.1.2 Final State Populations

VISTA checks the agreement in final state populations by determining the Poisson probability for the background estimation to fluctuate to what is seen in data. The calculation of the probability of data fluctuating from a perfectly known background is shown in Equation 8.1, with p_d the probability of getting d or more background events, d the number of data events, and b the weighted number of background events. This probability then needs to take into account the fact that many final states were searched by reducing the significance of any individual fluctuation. The probability formula is directly derived from the binomial probability. If the probability of a fluctuation is x, and there are N_s distributions that could have a fluctuated to that probability, then the probability that any of the distributions would have fluctuated to x follows the binomial probability formula, shown in Eqs. 8.2 and 8.3.

$$p_d = \sum_{i=d}^{\infty} \frac{b^i}{i!} e^{-b} \tag{8.1}$$

$$P(k \; out \; of \; N_s) = \frac{N_s!}{k!(N_s-k)!}(p_{fs}^k)(1-p_{fs})^{N_s-k} \tag{8.2}$$

where k would be the number of distributions more significant than p_{fs}. We are interested in a situation where any of the distributions is more significant than p_{fs}, so the probability would equal $p = 1 - P(0 \; out \; of \; N_s)$ [1],

[1]For states that show a data excess, p_{fs} is given for the probability of data to fluctuate up to the background. In cases where there is a data deficit, the Poisson probability is calculated in the

$$p = 1 - \Big(1 - p_{fs}\Big)^{N_s} \tag{8.3}$$

The reduced probability can then be converted into units of standard deviation by solving for σ in Equation 8.4.

$$\int_{\sigma}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = p \tag{8.4}$$

with the value in σ positive for data events exceeding standard model background and negative for the opposite situation.

As an example of determining the agreement consider a final state with 167 data events and 121.6 $\pm$ 5.6 events predicted from the standard model background. The statistical error in the background is determined by the number of events in background sample. Since the number of events is determined from a variety of sources, the individual contributions from each of the bins must be combined to determine expected statistical errors.

The probability for a precisely known background of 121.6 to fluctuate up to 167 or more is 1.572 10^{-5}. Since the background is not known precisely, the Poisson distribution is convolved with a Gaussian with a width of the expected background uncertainty, 5.6 events to give a probability of 8.12 10^{-4}. This probability is then adjusted to take into account the number of trials (in this case 180) to arrive at final probability of 0.136, which corresponds to 1.1 standard deviations. Since this value is below the 3σ threshold, it is not reported as significantly discrepant.

The final state populations with the σ values after trials factor for each of the final states can be found in Table 8.1.

opposite direction. It is the probability of seeing d or less data events given the background b, so the sum in Equation 8.1 would run from $i = 0$ to d, rather than from d to ∞.

8.1.3 Histogram Shapes

VISTA also considers 1-D histograms to quantify data/background agreement. The probability of the data being a statistical fluctuation from the expected background distribution uses a probability determined from the Kolmogorov-Smirnov statistic. The Kolmogorov-Smirnov test is based on the empirical cumulative probability distributions of the data and background and finds the largest deviation in the distribution. After determination of the initial probability, it follows the same basic method as the event counts discussed above, reducing significance by the trials factor and converting into units of σ.

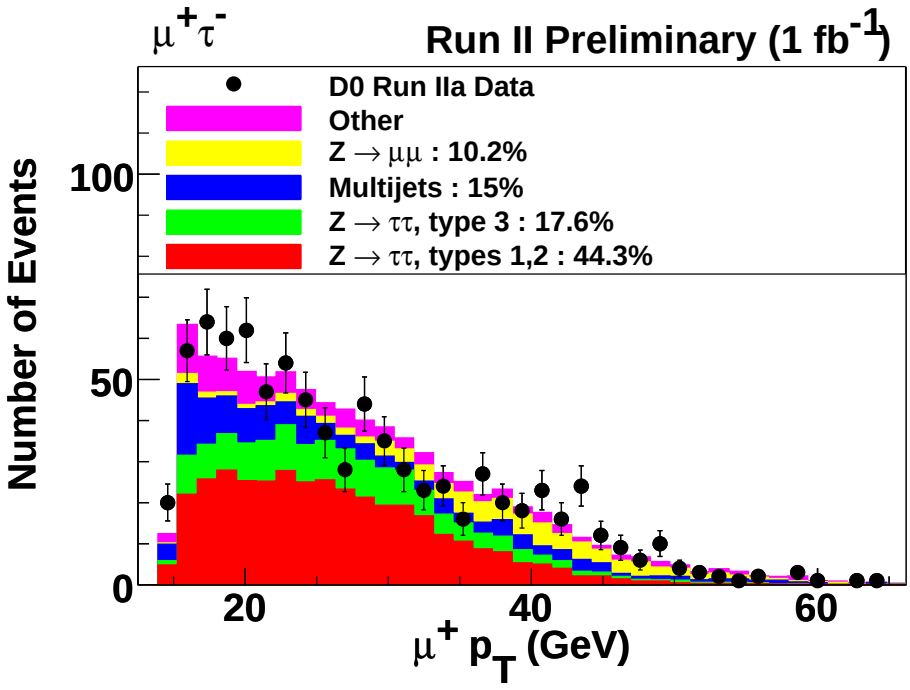
The histograms plotted for each final state include the p_T, η, and ϕ for each object in the event, $\not\!\!E_T$, spatial differences between each pair of objects in the event, $\Delta\phi$, $\Delta\mathcal{R}$, invariant mass among all object combinations, transverse mass among all object combinations with $\not\!\!E_T$, and a few other specialized variables.

An example of a full set of histogram shape plots for one final state with only two objects is shown in Figures 8.1, 8.2, 8.3, 8.4, 8.5, 8.6, 8.7, 8.8, 8.9, and 8.10.

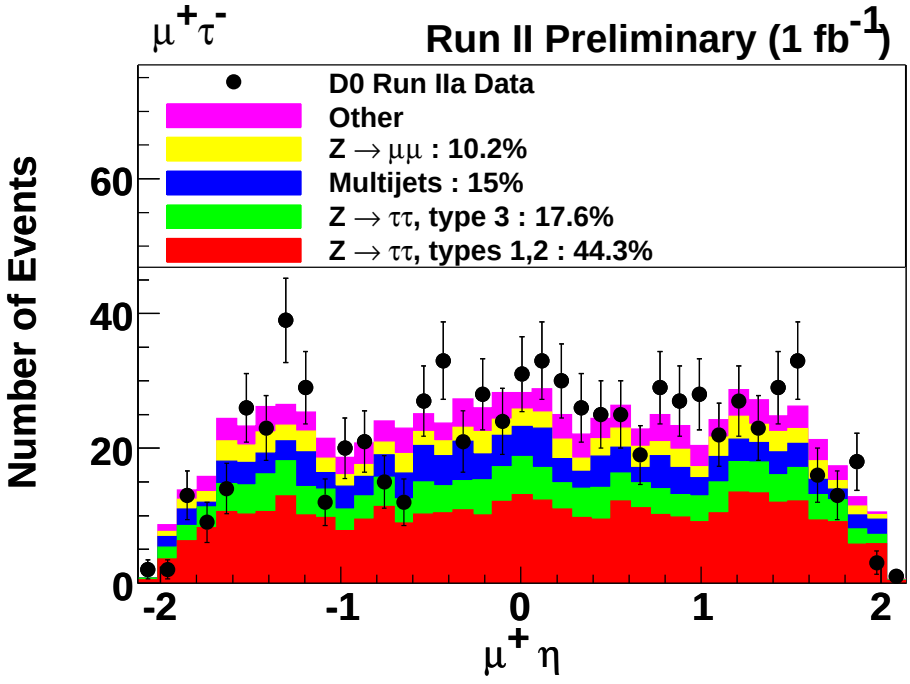
8.2 SLEUTH

SLEUTH is a check of the high-p_T tails of final states. This will find any new physics that are in accord with the basic SLEUTH assumptions.

- The new physics final states have objects with high-p_T relative to the standard model and instrumental backgrounds.
- The new physics occurs in a small subset of final states.
- The new physics occurs as excesses of data over standard model background.

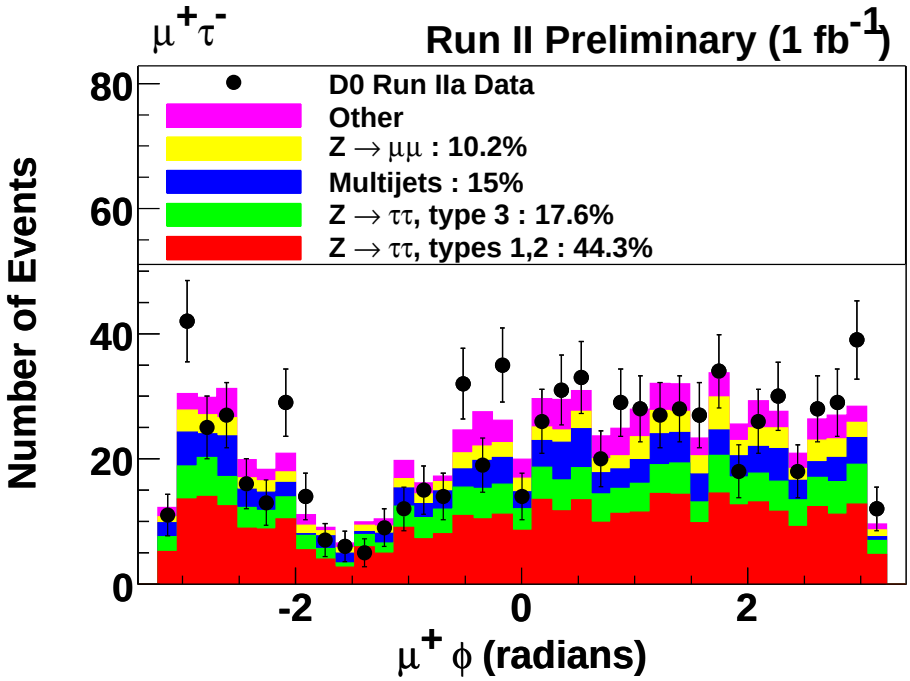


(a)

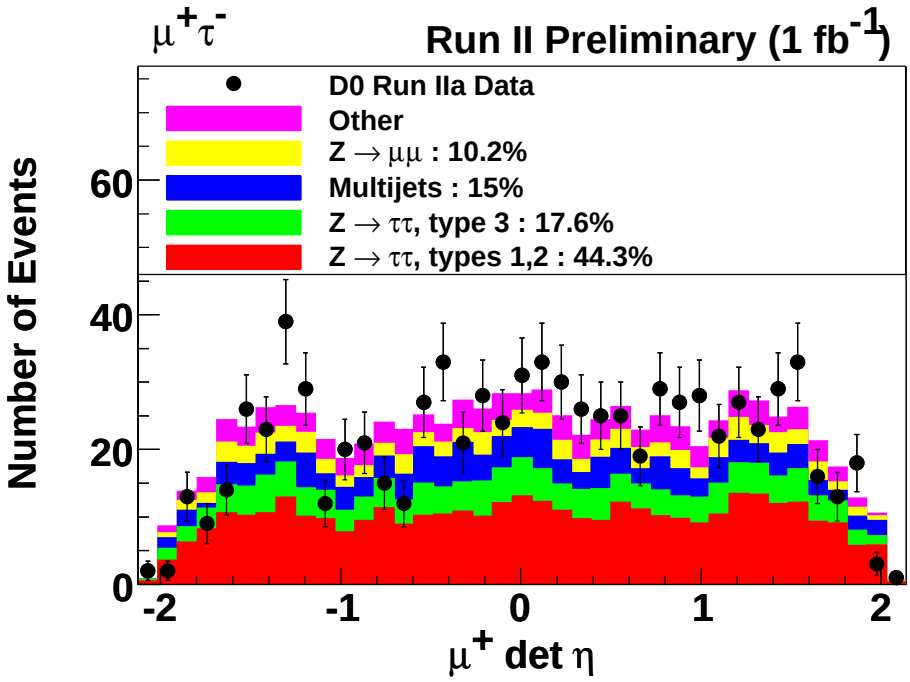


(b)

Figure 8.1: All of the histograms plotted for the VISTA state with one muon and one tau. The μ p_T and μ η distributions.

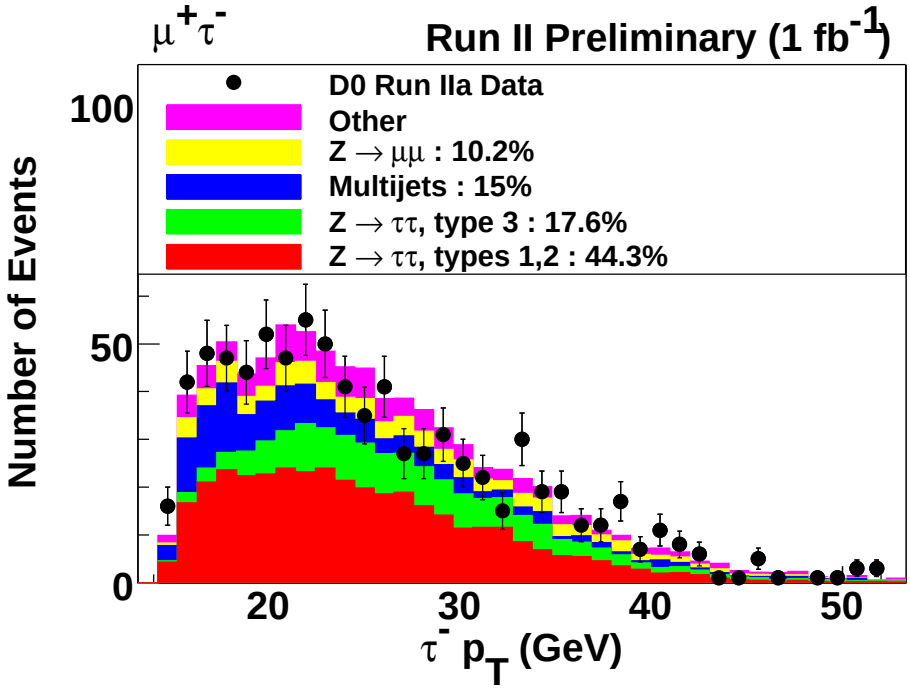


(a)

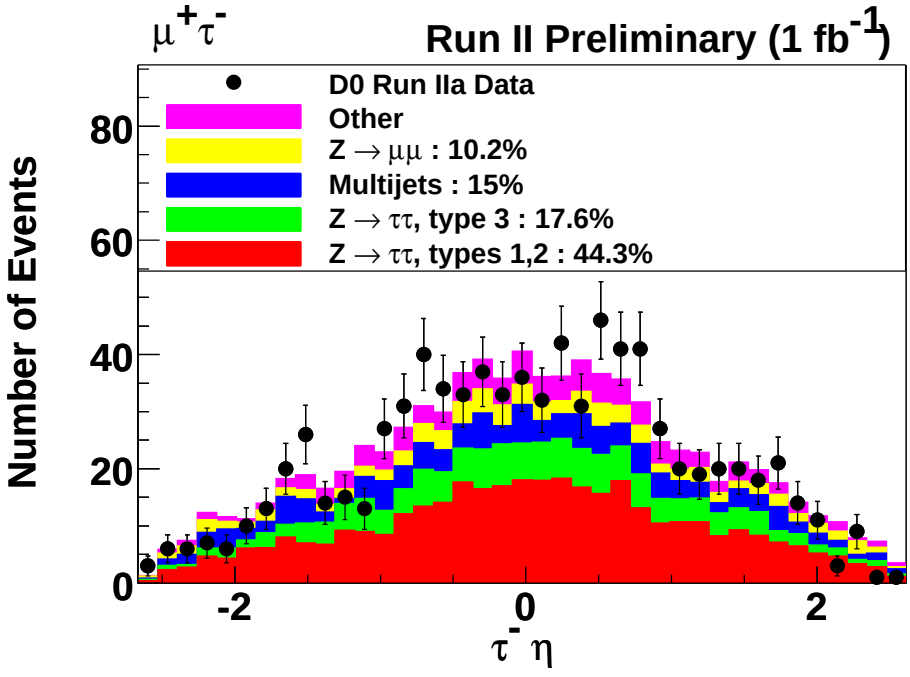


(b)

Figure 8.2: All of the histograms plotted for the VISTA state with one muon and one tau. The μ ϕ and μ detector η distributions.

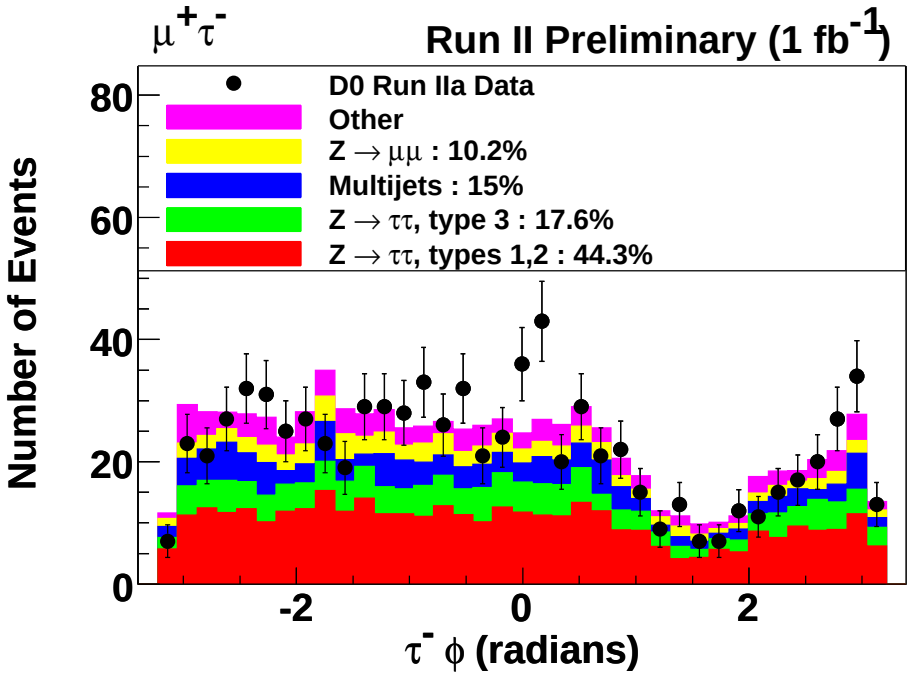


(a)

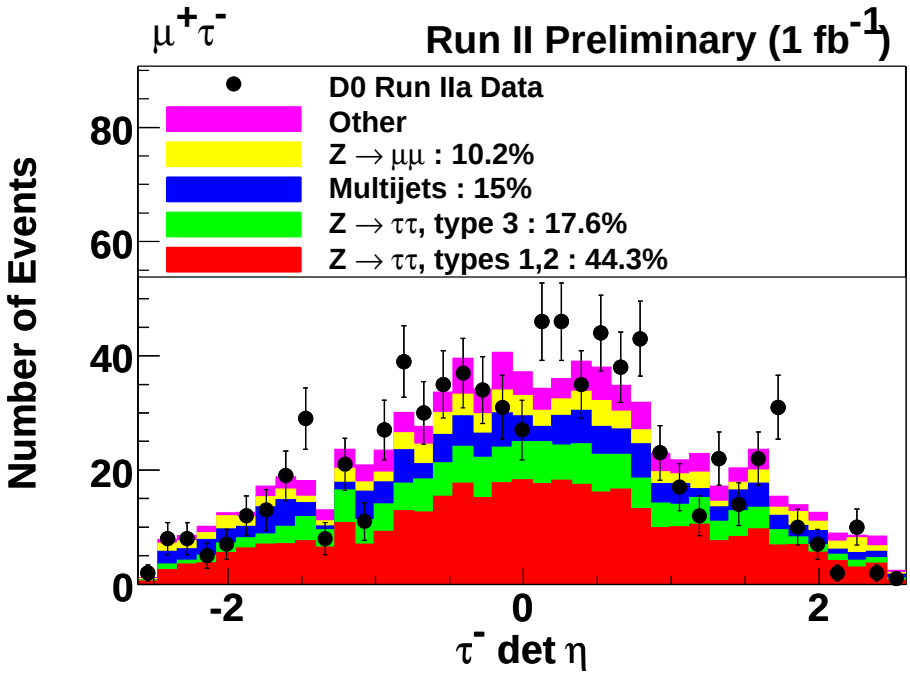


(b)

Figure 8.3: All of the histograms plotted for the VISTA state with one muon and one tau. The τ p_T and the τ η distributions.

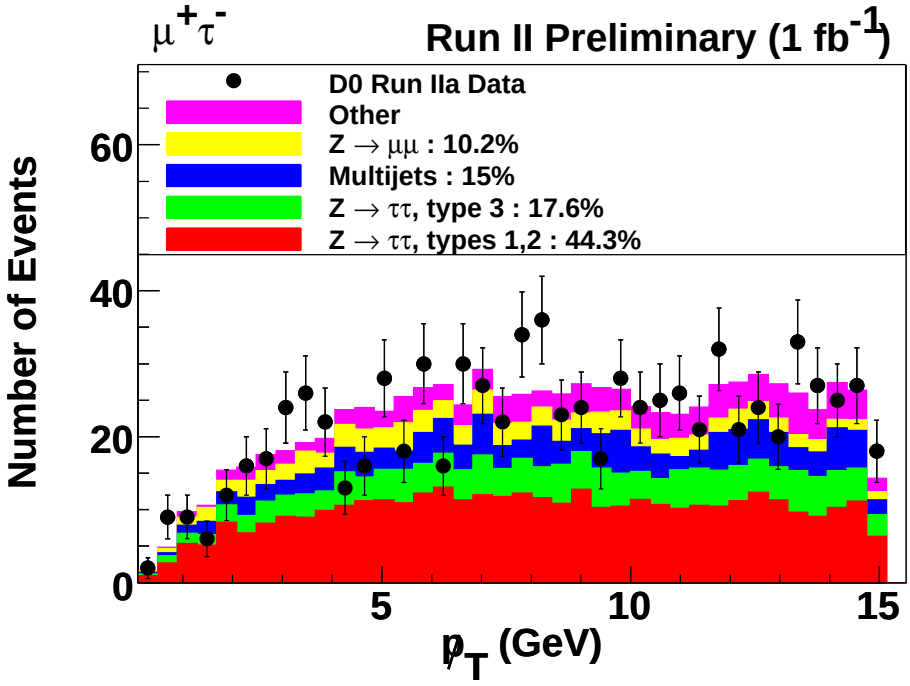


(a)

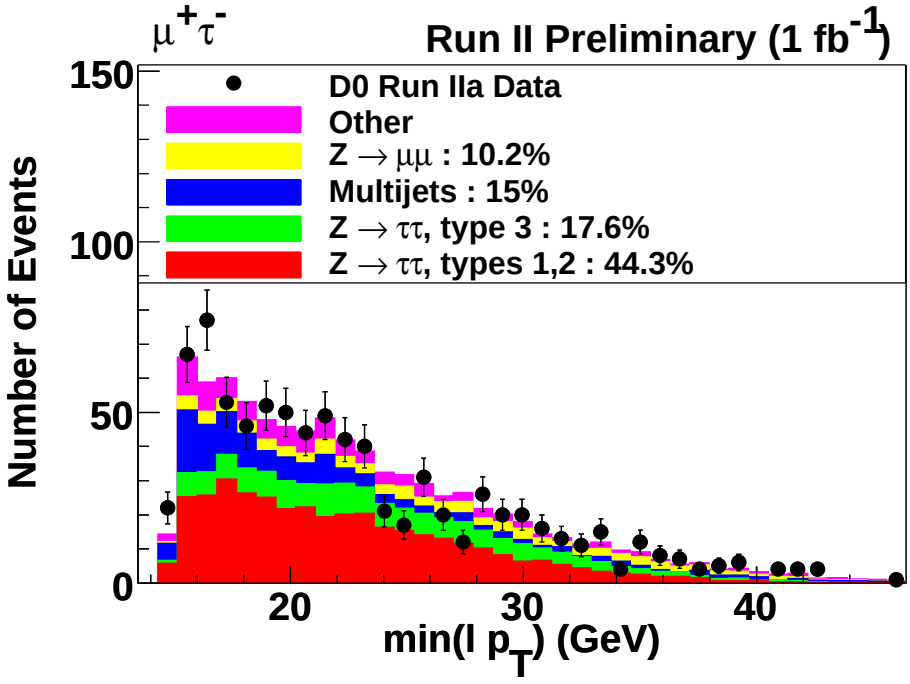


(b)

Figure 8.4: All of the histograms plotted for the VISTA state with one muon and one tau. The τ ϕ and the τ detector η distributions.

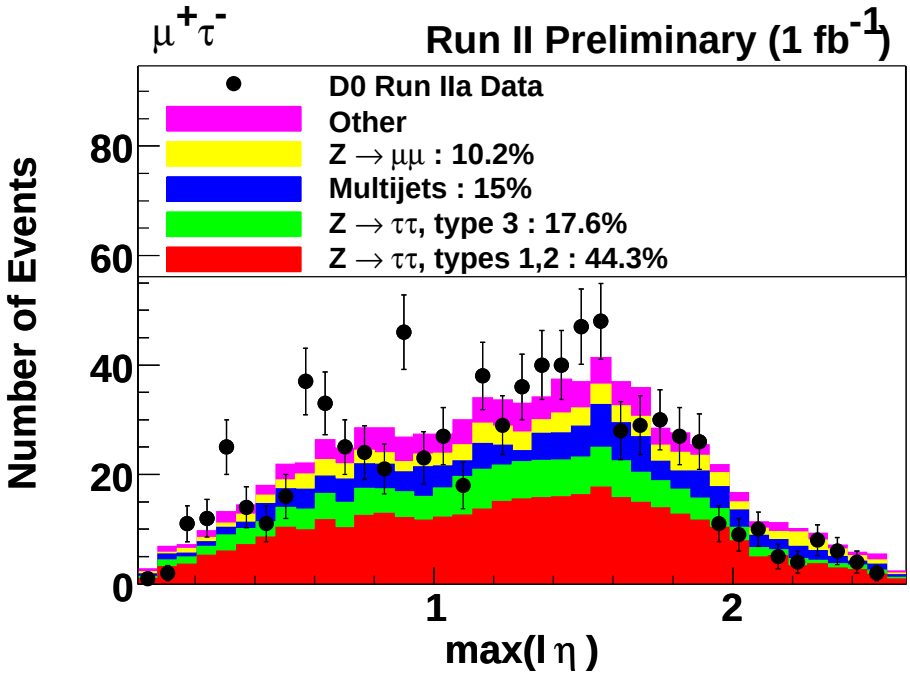


(a)

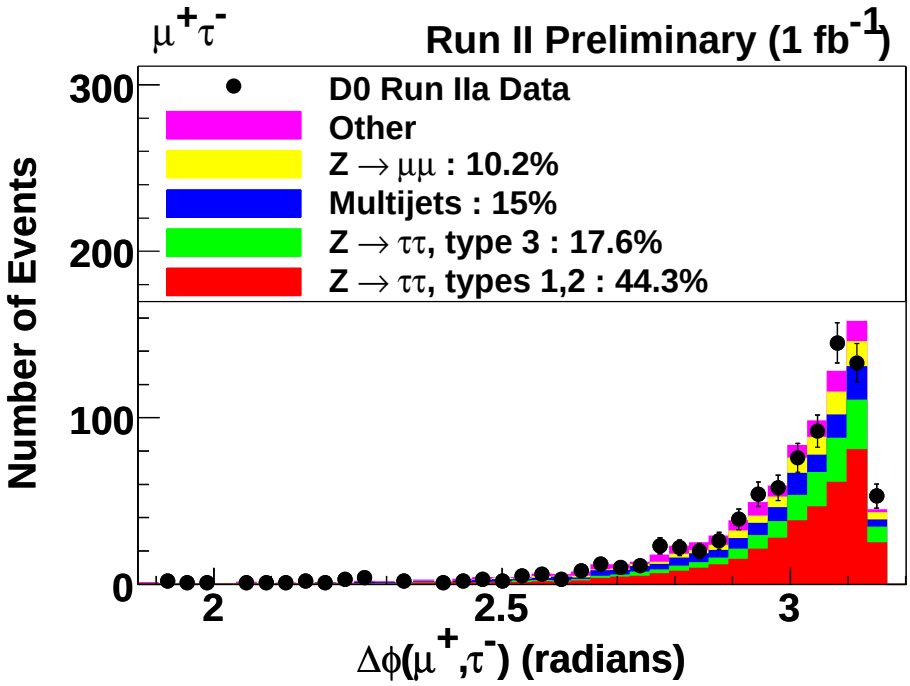


(b)

Figure 8.5: All of the histograms plotted for the VISTA state with one muon and one tau. The $\not\!\!E_T$ and minimum p_T of the μ and τ.

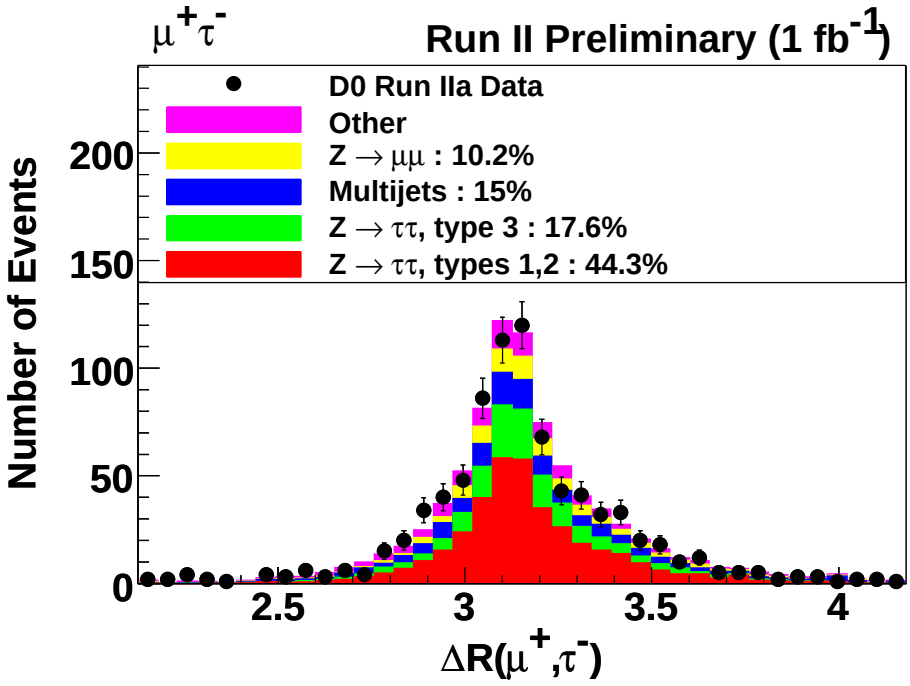


(a)

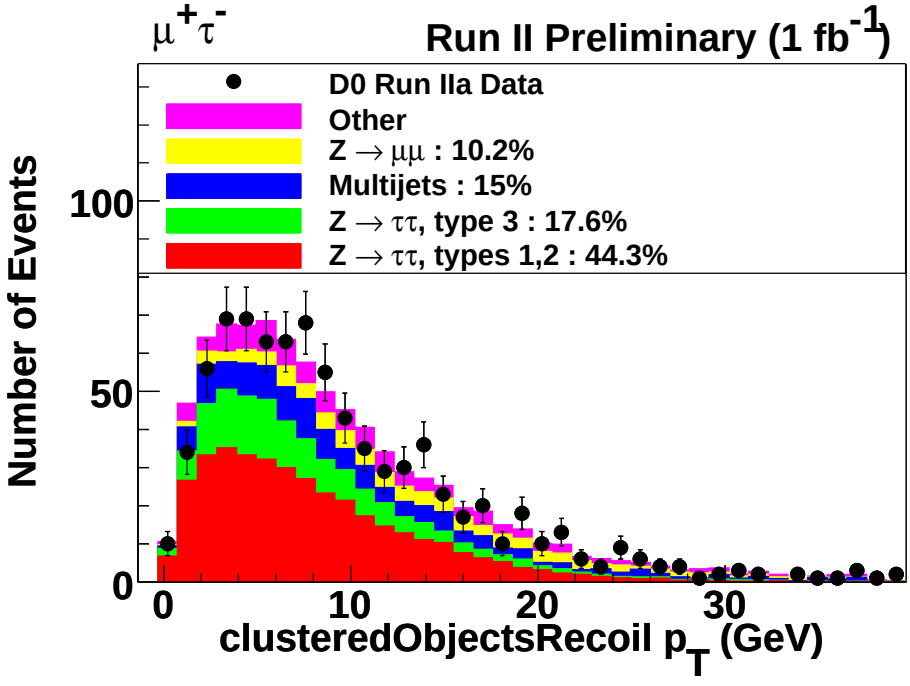


(b)

Figure 8.6: All of the histograms plotted for the VISTA state with one muon and one tau. The maximum η of the μ and τ.

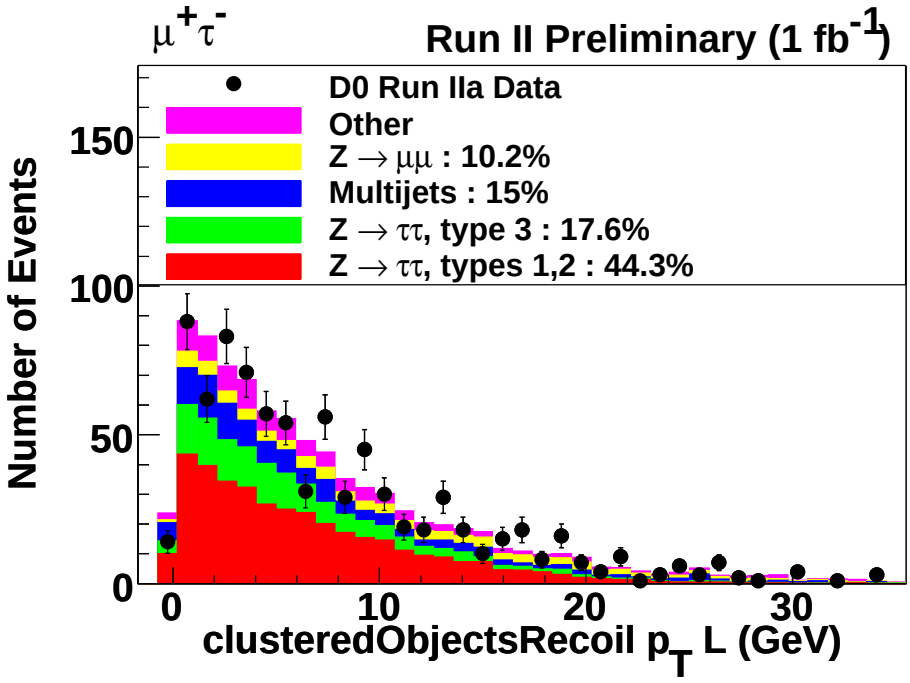


(a)

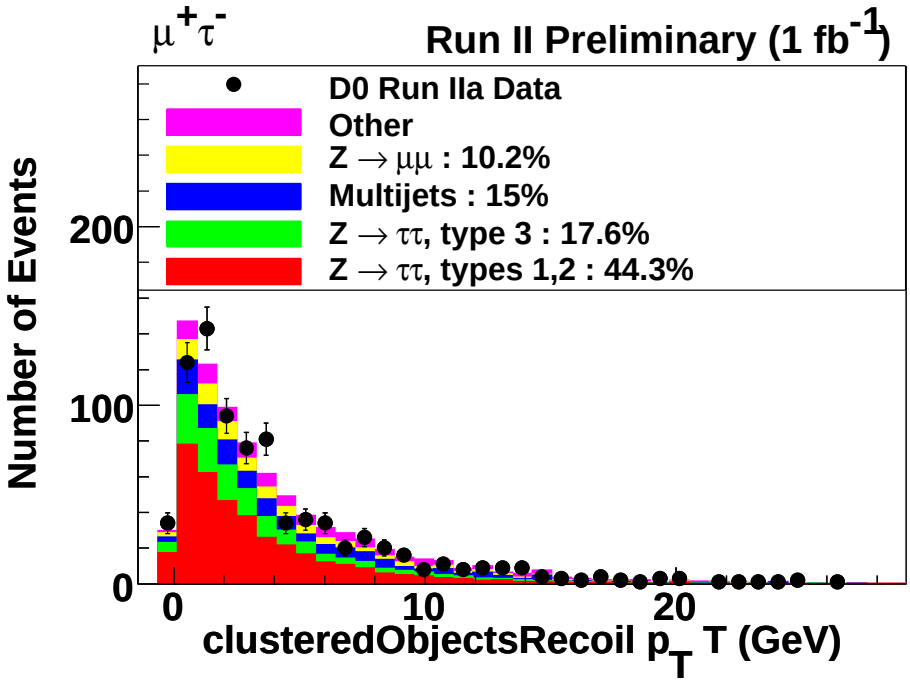


(b)

Figure 8.7: All of the histograms plotted for the VISTA state with one muon and one tau. The $\Delta\mathcal{R}$ between the μ and τ. The clustered object recoil is the vector sum of the $\not\!\!E_T$ and unclustered energy.

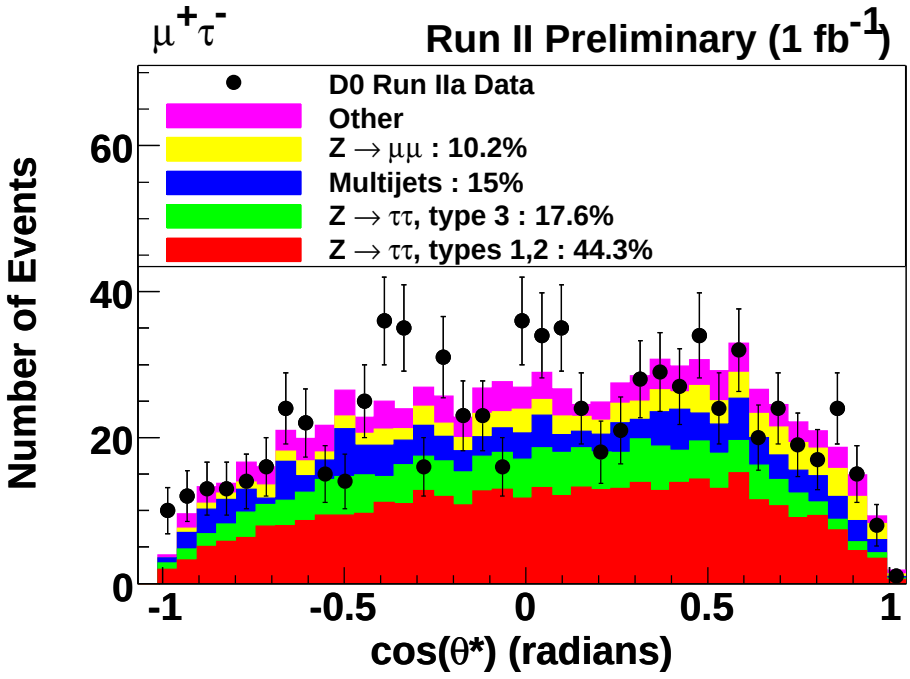


(a)

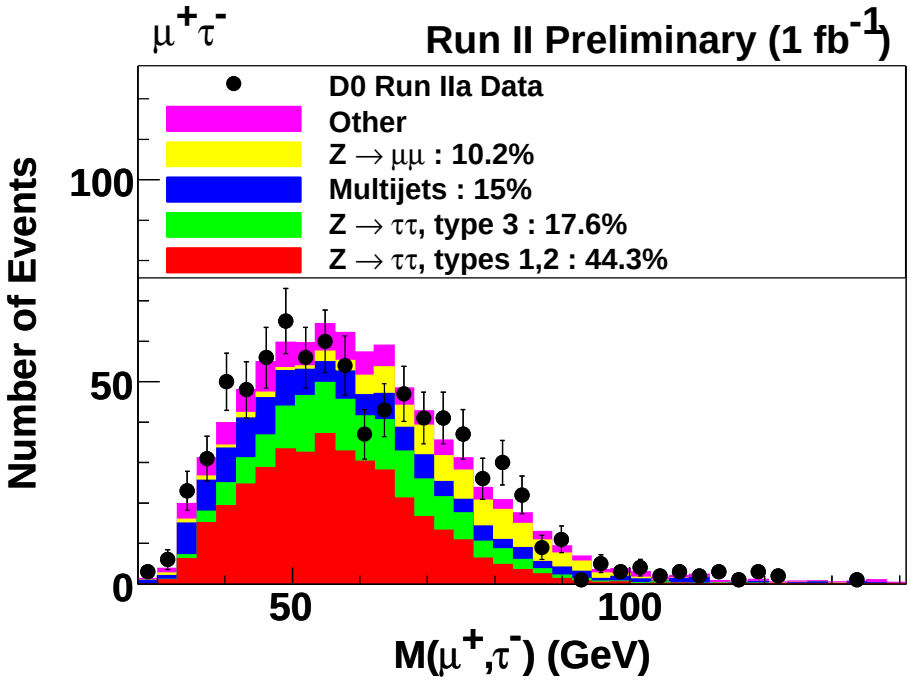


(b)

Figure 8.8: All of the histograms plotted for the VISTA state with one muon and one tau. A thrust axis is defined as the vector sum of the two objects in the event. The clustered object recoil is then determined for the transverse and longitudinal components with respect to the thrust axis.

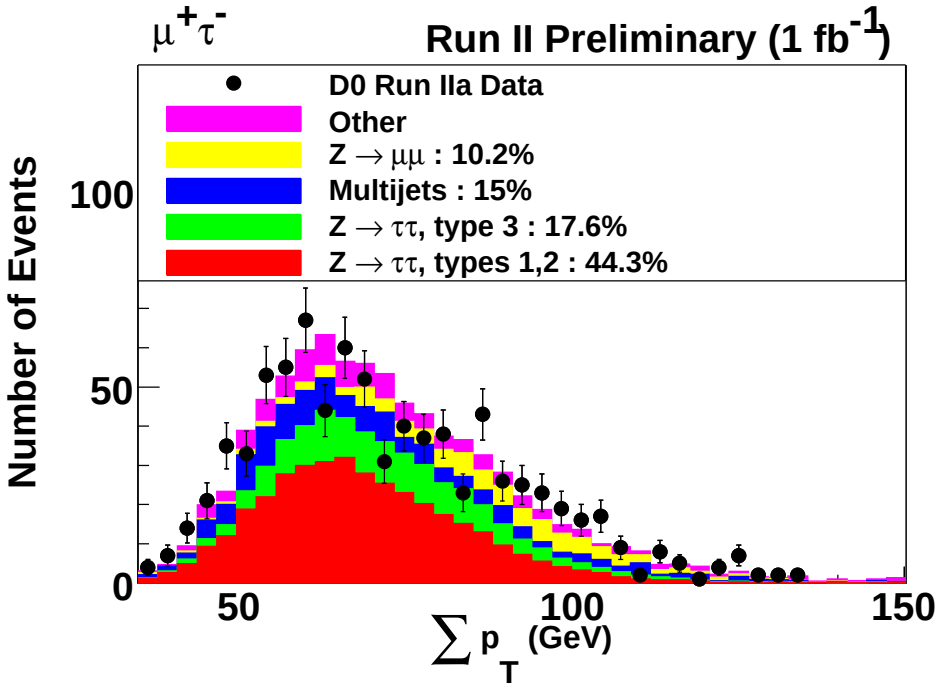


(a)



(b)

Figure 8.9: All of the histograms plotted for the VISTA state with one muon and one tau. The plot $cos(\theta^*)$ shows the cosine of the angle between the positively-charged lepton and the reconstructed Z boson in the frame of the Z boson. Also, the invariant mass of the μ and τ.



(a)

Figure 8.10: All of the histograms plotted for the VISTA state with one muon and one tau. The scalar sum of the transverse momenta of all of the objects in the event plus the missing transverse energy.

8.2.1 SLEUTH Final States

SLEUTH reduces the number of final states searched to lower the overall trials factor that needs to be applied. Four basic principles are applied in the final state reduction.

- Allow global charge conjugation, which means that an event with each of the object charges flipped, would go into the same final state. For example, an event that contains a positively-charged muon and negatively-charged electron would go into the same final state as a negatively-charged muon and positively-charged electron. Or, two positively charged electrons would go into the same final state as two negatively-charged electrons.

- Apply 1^{st} and 2^{nd} generational equivalence, which means that events where each electron is switched to a muon and each muon to an electron would belong

to the same final state. This would mean an event with two oppositely-charged electrons would go into the same final state as two oppositely-charged muons but would be in a different final state than one muon and one electron with opposite charges.

- Jets in the hard scatter are produced in pairs. This point assumes that if an event has an odd number of jets, the unpaired jet is a gluon from initial or final state radiation. The physics involving this extra gluon would not be directly tied to the new physics process, so it is not used to define a separate final state. Therefore, an event with a muon and two jets would go into the same final state as a muon with three jets.

- b-quarks are produced in pairs. This assumes that if there are an odd number of jets that have been b-tagged in an event, then there is probably another jet that also originated from a b-quark. This would put an event with one b-tagged jet and one light jet in the same final state as an event with two b-tagged jets.

8.2.2 SLEUTH Algorithm

The SLEUTH algorithm starts with the final states described above. For each final state, one variable is calculated: $\sum p_T$,

$$\sum p_T = \sum_{i}^{\#\ objs} |\vec{p}_i| + \left|\not\!\!E_T\right| \tag{8.5}$$

which is determined from the scalar addition of all of the object transverse momenta in the event plus the missing transverse energy. At each data $\sum p_T$ value, SLEUTH counts the number of data and expected background events with a $\sum p_T$ equal to or greater than the $\sum p_T$ of the data event. The probability associated with this comparison is determined in the same way as for the raw event counts in VISTA. That is, a minimum $\sum p_T$ cut is chosen which maximizes the excess of data over

background for each final state. The most significant difference in those event counts is chosen for each final state. Pseudoexperiments are performed to find the overall probability that the data seen could result from only statistical fluctuations in the in the standard model background for that final state. The number of events in the fake data sample is determined from a Poisson fluctuation of the total number of background events, and then the fake data points are distributed by drawing random numbers between zero and one, sorting these numbers, and then associating them with a percentage of background content that falls before each particular fake data point. With this new fake data sample, the SLEUTH algorithm is run again. The number of pseudoexperiments needed to see something as interesting as what is seen in the data is calculated for each final state. Finally, the lowest probability final state is chosen, and the significance of that state is reduced by the number of final states checked by the same algorithms used for other multiple comparisons.

An example plot for opposite-sign muons or electrons with two or three jets is shown in Figure 8.11.

8.2.3 $t\bar{t}$ Sensitivity Test

We now perform a test of the sensitivity of the SLEUTH search, by testing whether $t\bar{t}$ would have been discovered in this data sample. For the $t\bar{t}$ test, the full background sample except for the $t\bar{t}$ Monte Carlo was pushed through the analysis. The main concern would be whether other final states would be able to compensate for the missing Monte Carlo, and SLEUTH would not be sensitive to $t\bar{t}$ in the data.

From Figure 8.13, one can clearly see the difference between including and removing the Monte Carlo. With a threshold of 0.001, the SLEUTH test with the $t\bar{t}$ Monte Carlo included has a statistical fluctuation probability of 0.69, but the Monte Carlo without $t\bar{t}$ has a probability of only $< 1.6\ 10^{-7}$.

The test was repeated using only a randomly selected 10% of the data (100 pb^{-1}).

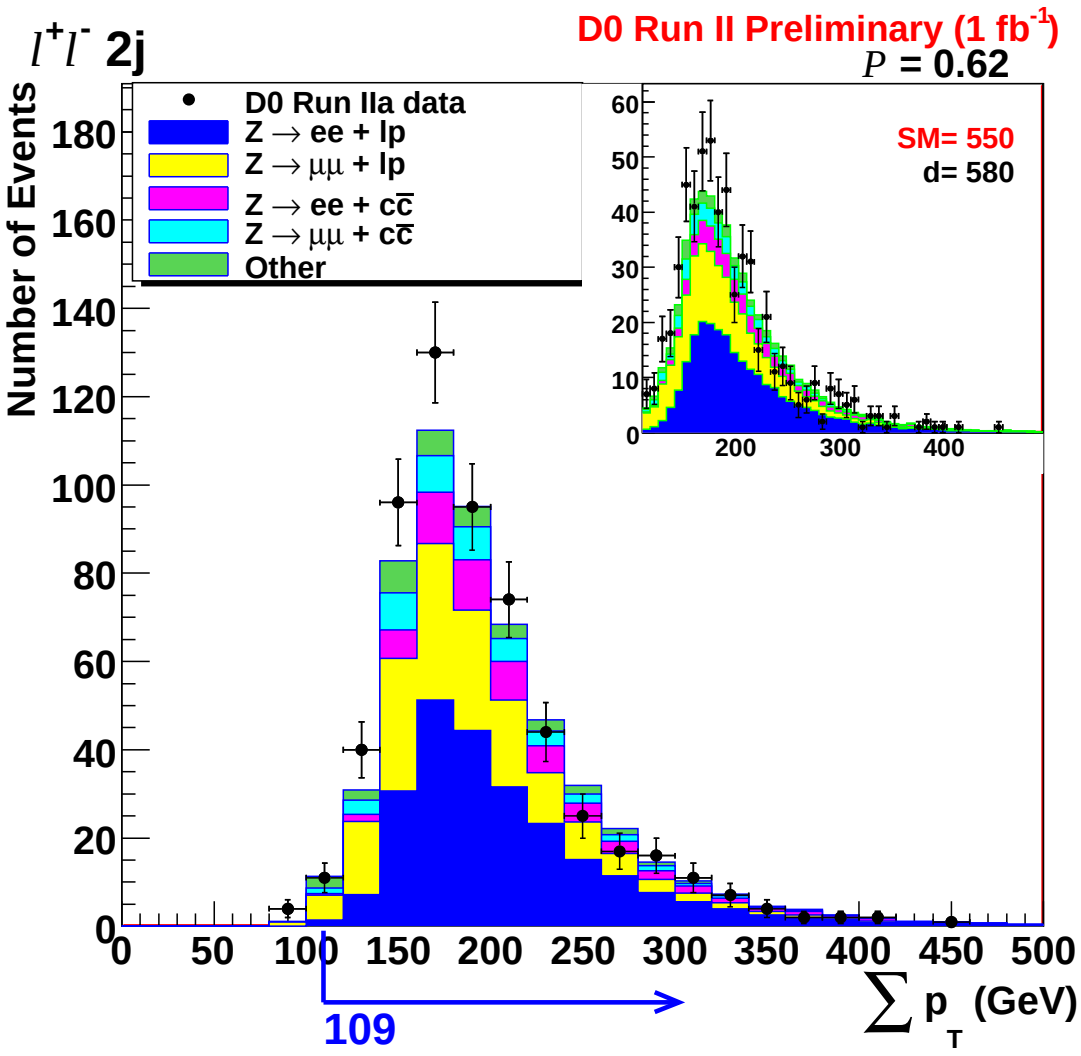


Figure 8.11: An example SLEUTH plot for the opposite-sign light dilepton (dimuon or dielectron) final state with two or three additional jets (not b-tagged). In this figure, the $\sum p_T$ cut that maximizes the discrepancy is at 109 GeV, which encompasses almost the entire distribution. This region is enlarged in the plot in the upper right, showing 580 data events compared 550 predicted from the standard model background providing a probability of a statistical fluctuation of 0.62.

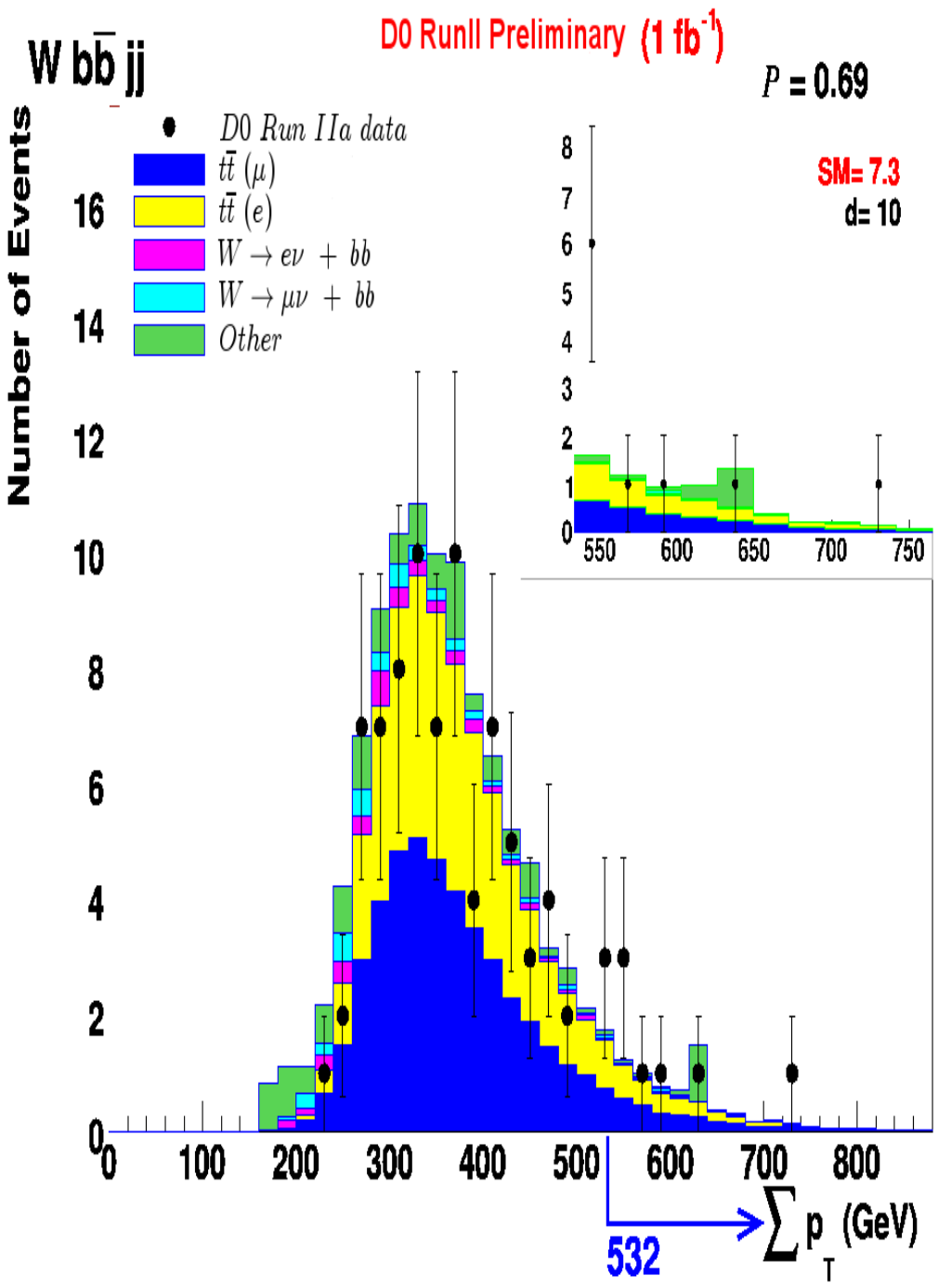


Figure 8.12: Sensitivity test for $t\bar{t}$. In this figure the $t\bar{t}$ Monte Carlo is included, and there are only minor differences between data and standard model background.

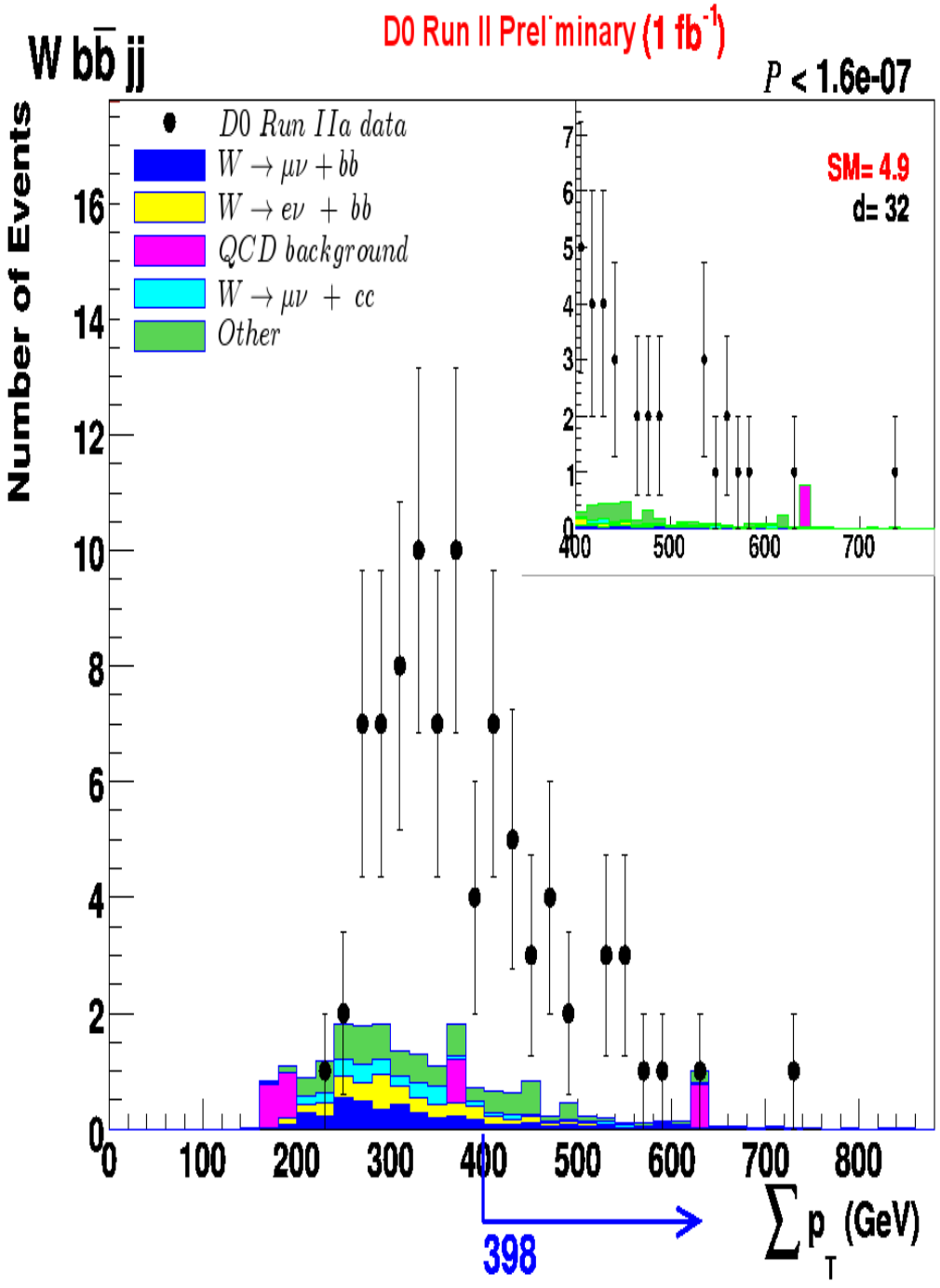


Figure 8.13: Sensitivity test for $t\bar{t}$. The figure shows the results of pushing through the entire analysis procedure without the $t\bar{t}$ Monte Carlo. In this case, SLEUTH easily passes the criterion of interest at 0.001 for this common $t\bar{t}$ final state.

This test was also shown to be successful, although starting from a low statistics fluctuation of probability = 0.12 and a fixed contribution from heavy-flavor (too few statistics for fitting). The 10% sensitivity test is shown in Figure 8.15.

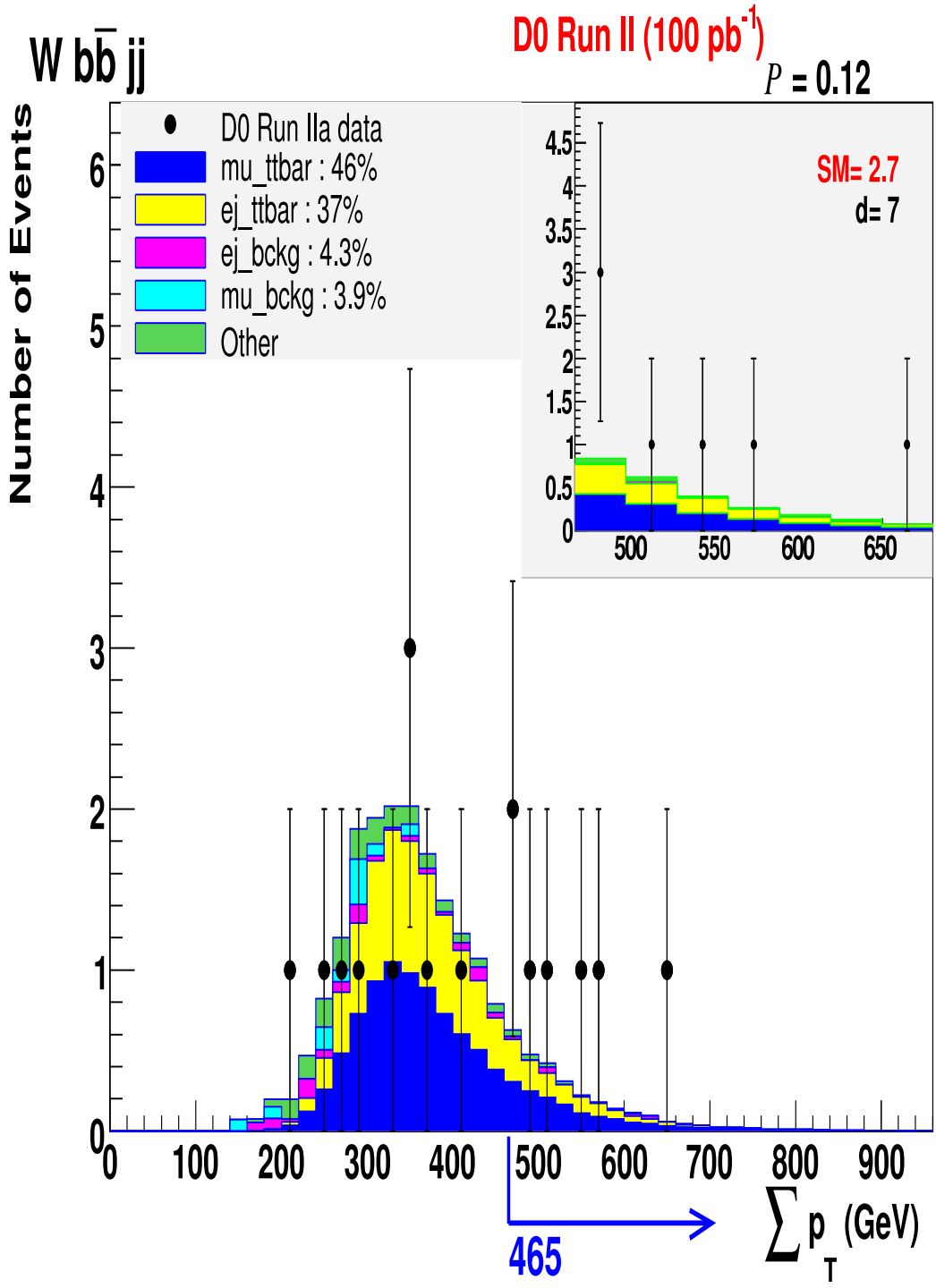


Figure 8.14: Sensitivity test for $t\bar{t}$ in 100 pb^{-1}. This figure includes the $t\bar{t}$ Monte Carlo, and the differences between data and standard model background are again minor.

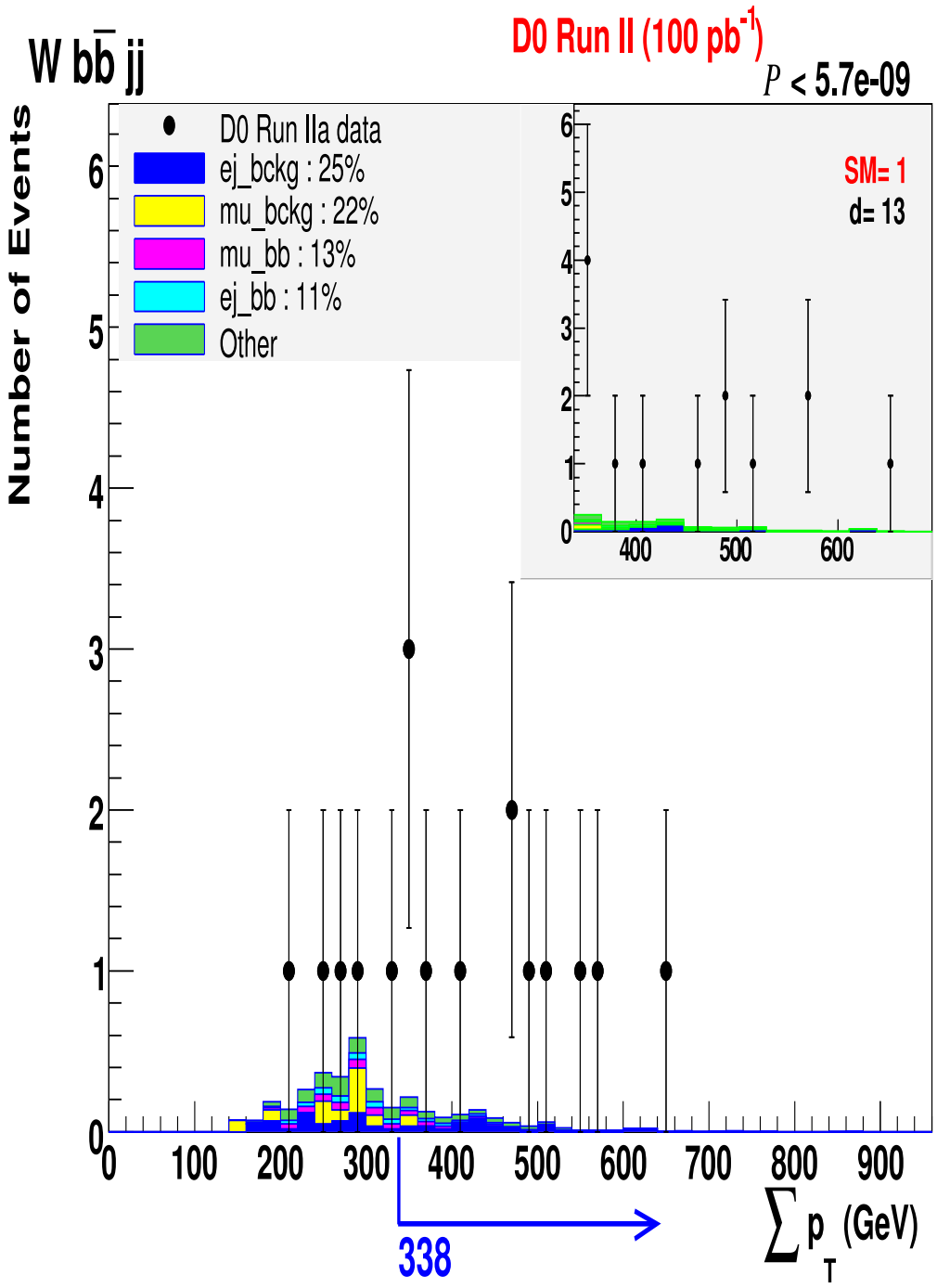


Figure 8.15: Sensitivity test for $t\bar{t}$ in 100 pb^{-1}. This figure shows the results of running the full analysis procedure using 10% of the Run IIa dataset when the $t\bar{t}$ Monte Carlo is removed. Even with this smaller sample, the SLEUTH algorithm still crosses the threshold.

Chapter 9

Results

The process of comparing the data to the expected standard model background in this analysis was done in three steps. First, the data were separated into seven final states, checked for qualitative agreement, and fit for normalization factors using the MIS analysis packages. All data cuts and correction factors were applied at this level. The second step involved dividing the seven final states into 180 states defined by the object content using VISTA. VISTA checked the overall event counts in each final state and plotted over 9000 1-D histograms. Each of the histograms was checked for agreement using Kolmogorov-Smirnov probabilities. VISTA then reported the agreement of all final states and histograms, putting special focus on those with statistically significant differences that disagree at a level of 3σ, as discussed in Section 8.1.2. The final step was to combine some of the VISTA final states and check the tails of $\sum p_T$ distributions using SLEUTH. SLEUTH checked for discrepancies in the tails that correspond to probabilities < 0.001 that the standard model background distributions would fluctuate to a distribution as discrepant as what is seen in data. The value of 0.001 is equivalent to approximately 3σ after incorporation of the trials factor.

9.1 Model Independent Search Normalization Fits

The purpose of the MIS analysis packages is to ensure that the primary standard model processes used for the background estimate are well-modeled. Seven final states were considered, each with one dominant standard model process, and the overall normalization for these processes was determined from a fit.

All of the processes considered in each of the seven final states with the number of histograms used in the fits, the number of histograms used to check agreement, the overall normalization scale factors with their uncertainties, the number of events, and the χ^2 of the fit are shown in Table 9.1.

9.2 VISTA

In VISTA, the separation of the input dataset into final states (completely defined by the object content of the event) yields a total of 180 unique final states. For these 180 final states, the probability of the data distribution resulting from a statistical fluctuation of the background sample is determined from $p = 1-(1-p_{fs})^{180} \approx 180\, p_{fs}$(for p_{fs} small), where p_{fs} is the probability that the number of events predicted in the standard model background would fluctuate up to or down to what is observed in data (before applying the correction for the 180 trials). This is then converted into units of standard deviation using $\int_{\sigma}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = p$. This procedure is described in more detail in Section 8.1.2. The final state probabilities converted into standard deviations before adding the trials factor correction are shown in Figure 9.1. This distribution shows most final states near the center, with some excess at the tails. Of the 180 distributions, four show significant discrepancy. These are the final states μ + 2 jets + $\not\!\!E_T$ with a converted probability of 9.3σ after trials correction, $\mu + \gamma + 1$ jet + $\not\!\!E_T$ with 6.6σ, $\mu^+\mu^- + \not\!\!E_T$ with a discrepancy of 4.4σ and $\mu^+\mu^- + \gamma$ at 4.1σ.

Table 9.1: The results of the MIS inclusive fits for all inclusive final states. Ignoring k-factors and trigger efficiencies, all Monte Carlo samples should fit to 1.0 for 1.0 fb^{-1} of data. The dominant standard model process is listed first for each final state.

MIS State	Input Sample	Weight	Uncer-tainty	# Fit Hists	# Check Hists	Fit χ^2/dof	# Events
e + jets	$W \to e\nu$	0.921	0.004	19	22	1297/1022	25k
	QCD e fakes	0.266	0.040				
	hf/lp ratio[a]	1.94	0.121				
μ + jets	$W \to \mu\nu$	0.712	0.004	19	21	1861/977	39k
	QCD μ fakes	0.684	0.055				
	hf/lp ratio[a]	1.94	0.043				
ee	$Z \to ee$ + 0lp	1.00	0.007	22	18	1243/883	18k
	lp/0lp ratio[b]	1.11	0.016				
	hf/lp ratio[a]	1.94	0.099				
$\mu\mu$	$Z \to \mu\mu$ + 0lp	0.782	0.007	22	22	1321/987	30k
	lp/0lp ratio[b]	1.11	0.011				
	hf/lp ratio[a]	1.94	0.070				
$e\tau$	$Z \to \tau\tau$	1.02	0.035	13	8	537/497	1.4k
	W/QCD τ fakes	0.185	0.030				
	hf/lp ratio[a]	1.94	fixed				
$\mu\tau$	$Z \to \tau\tau$	0.686	0.025	13	8	593/497	1.4k
	W/QCD τ fakes	0.206	0.049				
	hf/lp ratio[a]	1.94	fixed				
μe	$Z \to \tau\tau$	1.31	0.016	13	6	436/467	0.74k
	W/QCD τ fakes	0.02	0.285				
	hf/lp ratio[a]	1.94	fixed				

[a] Heavy-flavor quark (c, b) to light parton (g, s, u, d) radiative jet ratio. It is determined by allowing the heavy-flavor fraction in each of the final states to float. An average was taken for the Drell-Yan and the W final states. These numbers were found to be very similar, so a final averaging was done incorporating all of the final states. This averaged ratio was then fixed in all of the final states, and the other parameters were found from a second fit.

[b] The zero light parton to greater than zero light parton ratio. This is the ratio of Drell-Yan production with additional radiative jets to Drell-Yan production without these additional partons and is determined for the dielectron and dimuon states. Since this factor is expected to be flavor-independent, an average of the values was determined, and then fixed for each of the final dilepton fits.

The μ + 2 jets + $\not\!\!E_T$ final state discrepancy shows an excess of events with a muon at $\eta > 1.0$ as seen in Figure 9.2(a). The excess points to an oversimplification in our approach to trigger efficiencies. The proportion of events selected by single muon vs. muon plus jets triggers changes significantly as we increase jet multiplicity. These triggers introduce η-dependent efficiencies which are not properly incorporated into our simple fits. The dimuon with missing transverse energy final state shows an excess of data compared to the standard model Monte Carlo prediction. A study of the track curvature of data and MC muons, and of the associated resolution, has shown that an additional smearing should be applied in the Monte Carlo to appropriately simulate very high p_T muons. The prime signature of these muons is an excess of $\not\!\!E_T$ because of the lack of compensation for the mismeasured, unbalanced track. The $\Delta\phi$ distribution of the muon and $\not\!\!E_T$ can be seen in Figure 9.2(b), where the excess tends to be with events where the missing transverse energy is pointing opposite to a muon. The other two states are directly related to an oversimplified modeling of the photon misidentification rate. This can be seen in Figures 9.3(a) and 9.3(b). There are three reasons for the discrepancy in the photon states. First, the Monte Carlo generators are known to poorly reproduce these processes. Second, the rate of jets misidentified as photons are not modeled well in the detector simulation. Finally, the $Z \rightarrow \tau\tau$ contribution with an electron misidentified as a photon overestimates the tracking efficiency, so that the Monte Carlo will have fewer of these events than the data. All of these may contribute to these plots. Hard jets are more easily misidentified as photons, which may explain part of the reason the data spectrum is harder, but there could be many contributing factors.

The 180 final states contribute a total of 9335 individual 1-D histograms in various variables, and a shape comparison is performed for each. The trials factor adjusted probability is determined with $p = 1-(1-p_{shp})^{9335}$, where p_{shp} is the KS probability to observe an individual shape discrepancy (before applying the correction for 9335

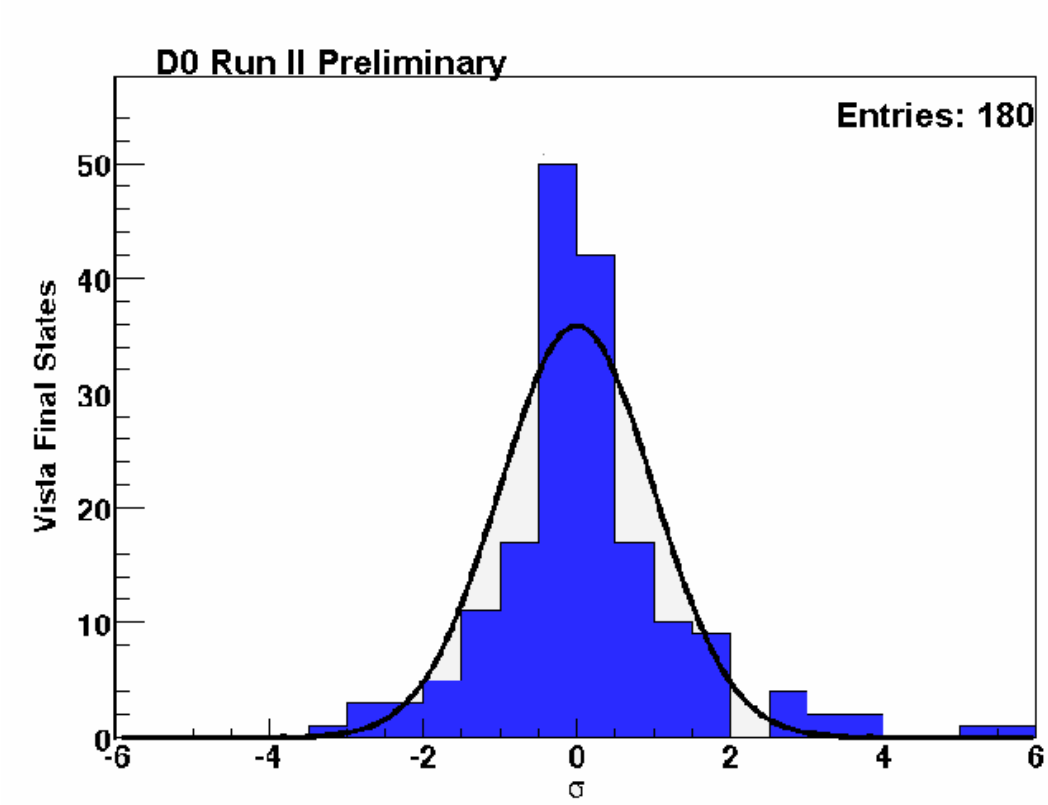
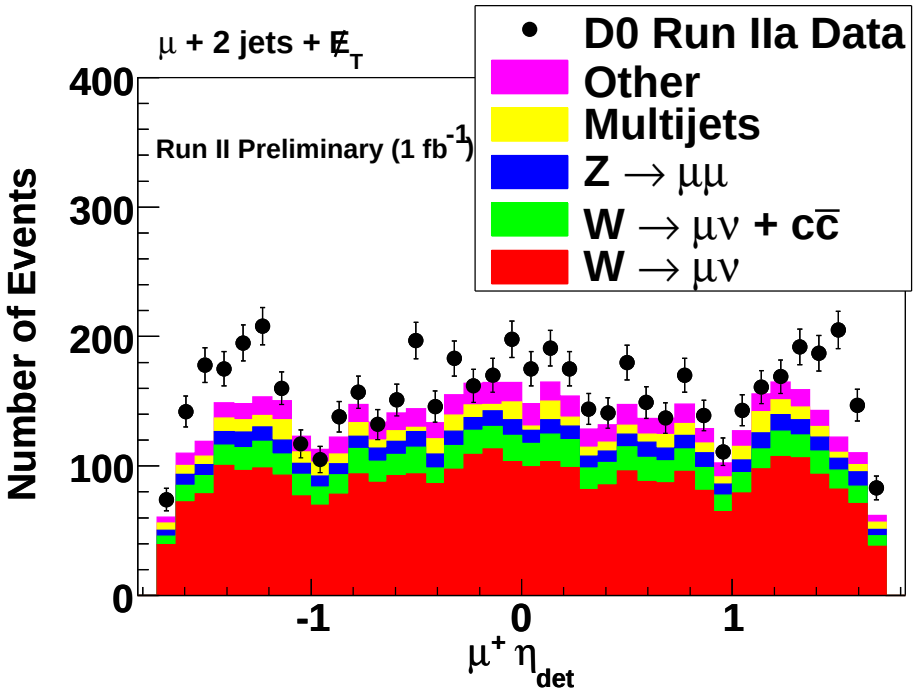


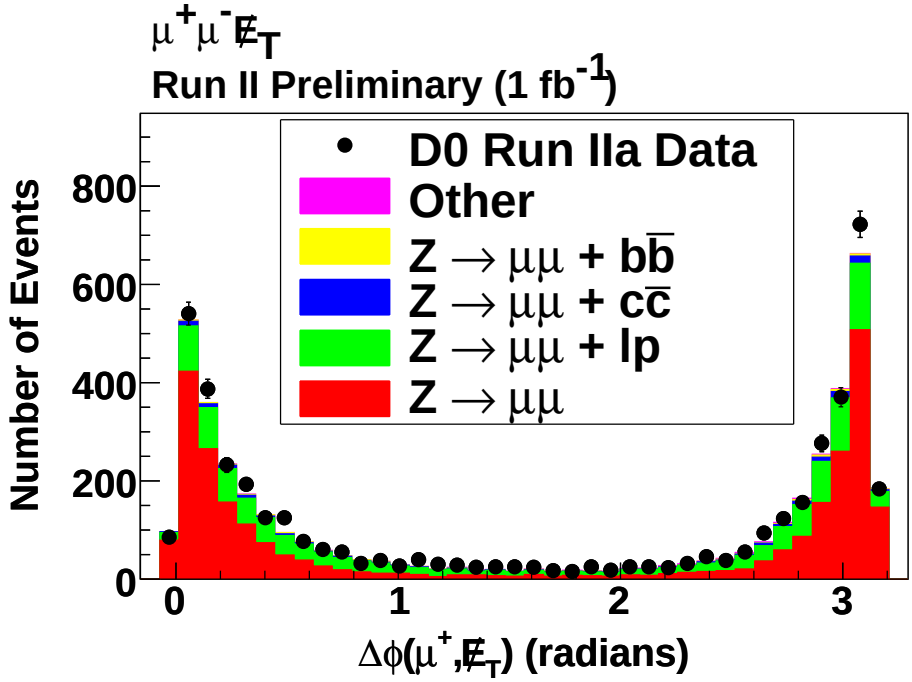
Figure 9.1: VISTA final state σ distribution for Run IIa sample before accounting for the trials factors. The curve represents a Gaussian distribution centered at zero to guide the eye. The event count distributions are expected to obey Poisson statistics, which is why the distribution is narrower than the curve.

trials). As with the probability for a final state discrepancy, the probability for a shape discrepancy is converted into units of standard deviation and the discrepancies are shown. For the histogram shapes, any deviation $>3\sigma$ is considered discrepant. The distribution of standard deviations before trials correction is shown in Figure 9.4. This distribution approximates a slightly shifted Gaussian of the expected width, but several distributions appear in the tails. The shift to the right (toward poor agreement) is expected because we introduce scale factors only for the most important discrepancies (minor systematic discrepancies are not individually treated, and these contribute preferentially towards bad agreement).

A total of 23 distributions are found to be discrepant at the 3σ level after trials

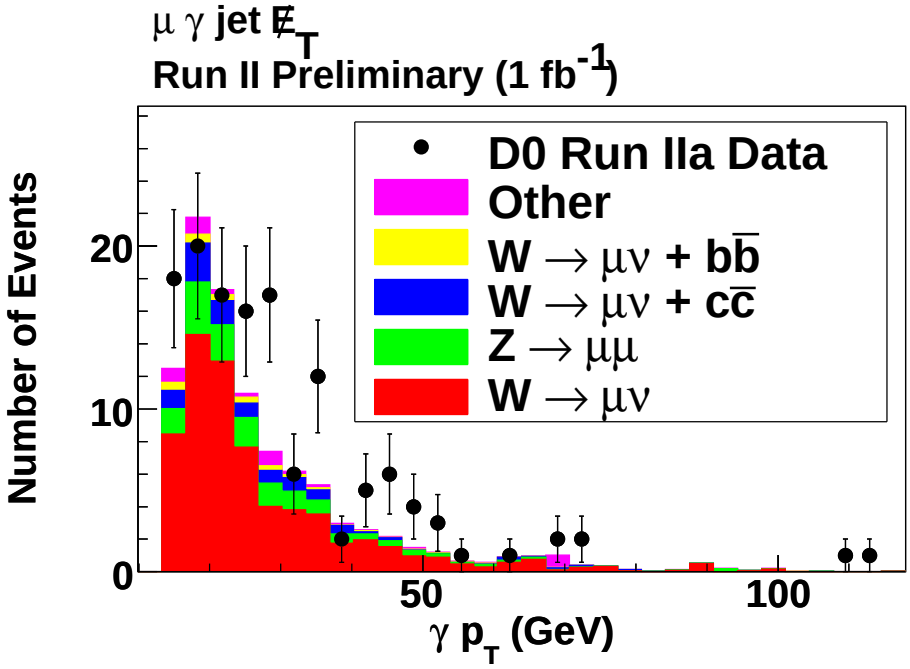


(a)

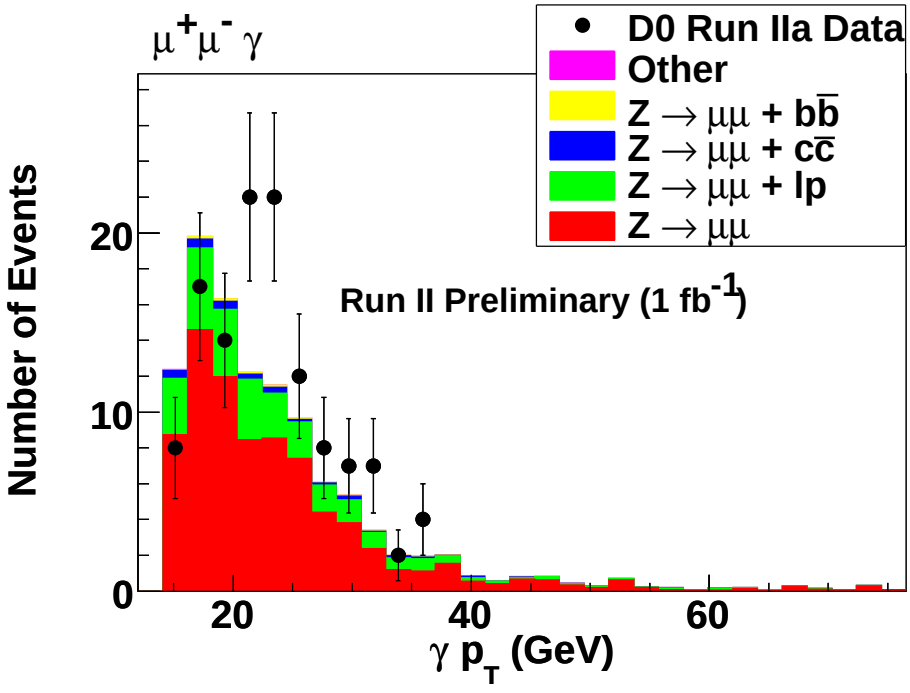


(b)

Figure 9.2: Figure 9.2(a) shows the excess of data in μ + 2 jets + $\not\!\!E_T$ to be focused on events with muons that have η values > 1.0. Figure 9.2(b) shows the $\Delta\phi$ distribution between a muon and the $\not\!\!E_T$, with the $\not\!\!E_T$ pointing opposite to a muon.



(a)



(b)

Figure 9.3: Two figures showing the p_T distributions of the photon.

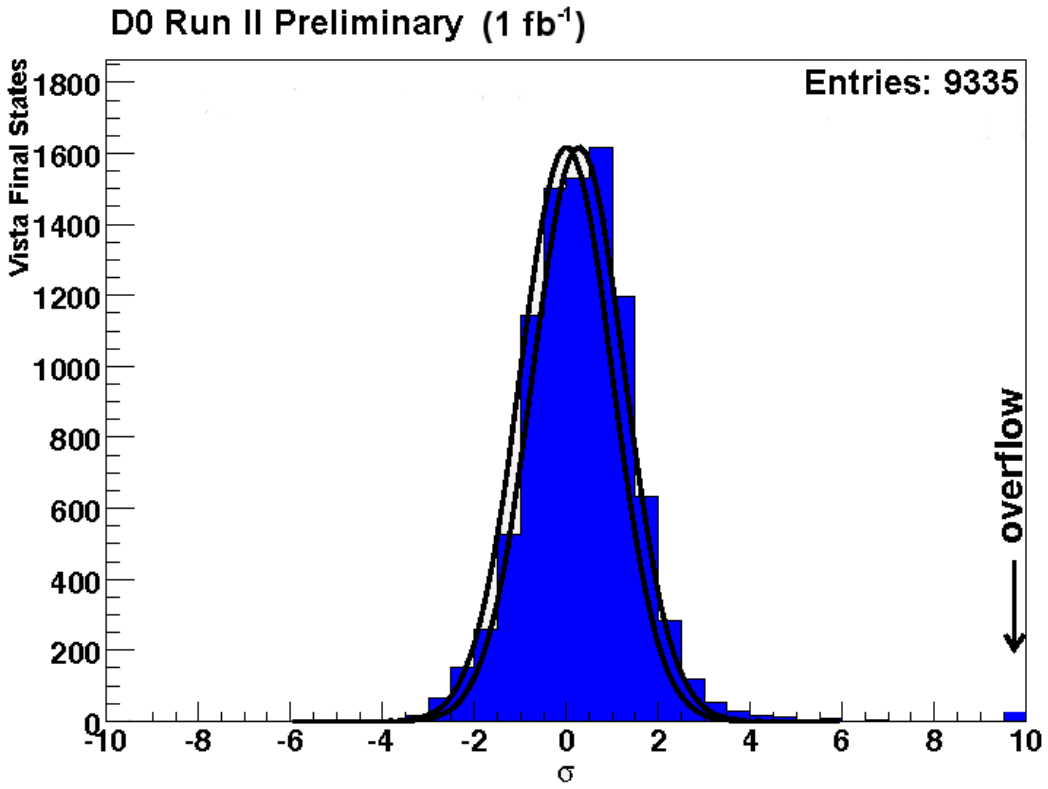
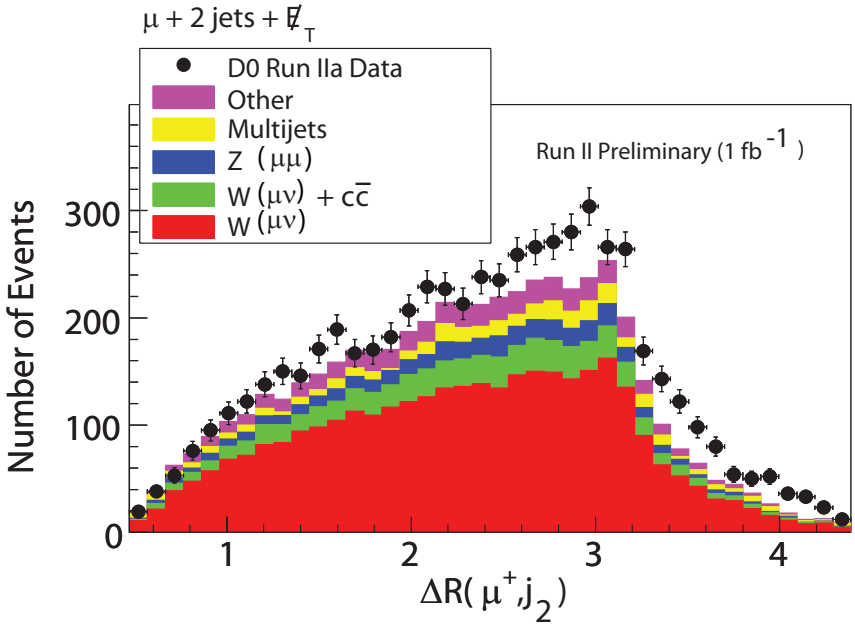
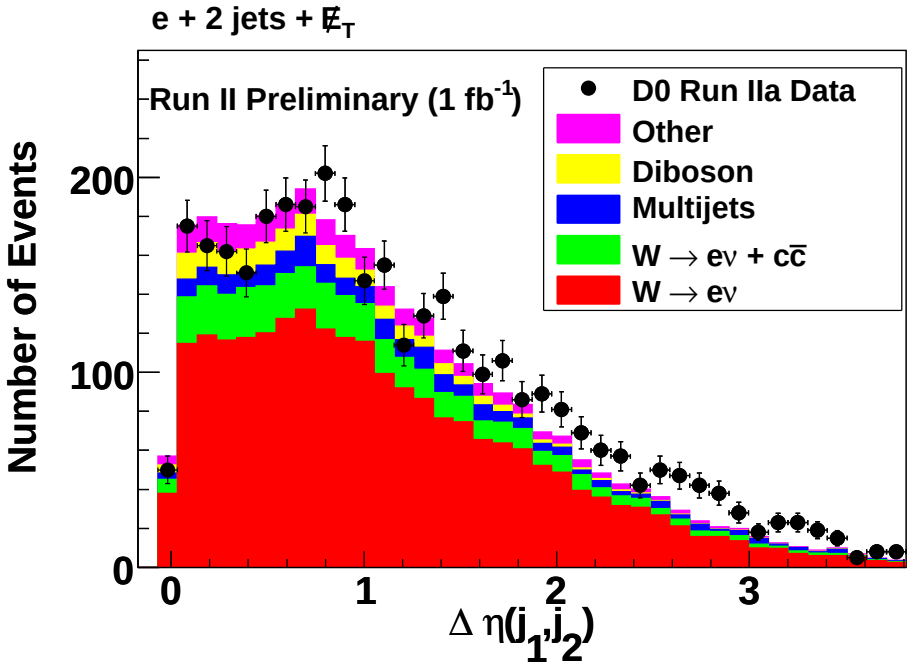


Figure 9.4: VISTA histogram σ distribution for 100% sample before accounting for the trials factor. Each curve is a Gaussian distribution. The curve that is shifted to lower values is centered at zero while the second curve is centered at the mean. The difference between the two curves approximates the average systematic uncertainty found in the plots.

correction. The majority of these are related to spatial distributions involving jets, low $\not\!\!E_T$ excesses in dilepton distributions and multijet-background-dominated τ distributions. All of these types of discrepancies are related to known oversimplifications in our modeling assumptions and would not be expected to severely affect the SLEUTH search for new physics in the high-p_T tails. Eight histogram shape discrepancies are shown in Figures 9.5(a) through 9.8(b). The full list of discrepant histograms is shown in Table 9.2.

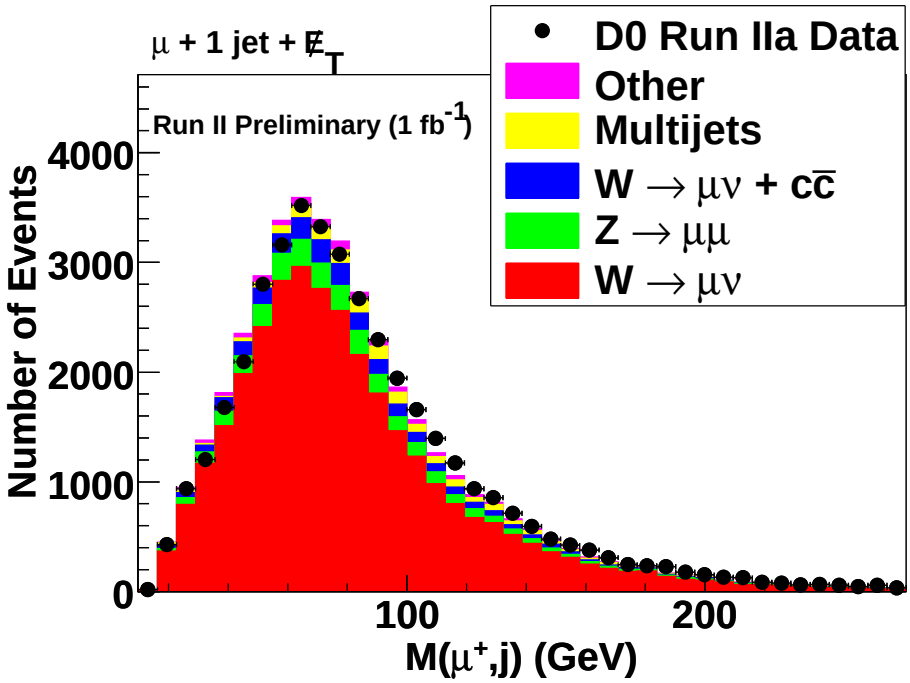


(a)

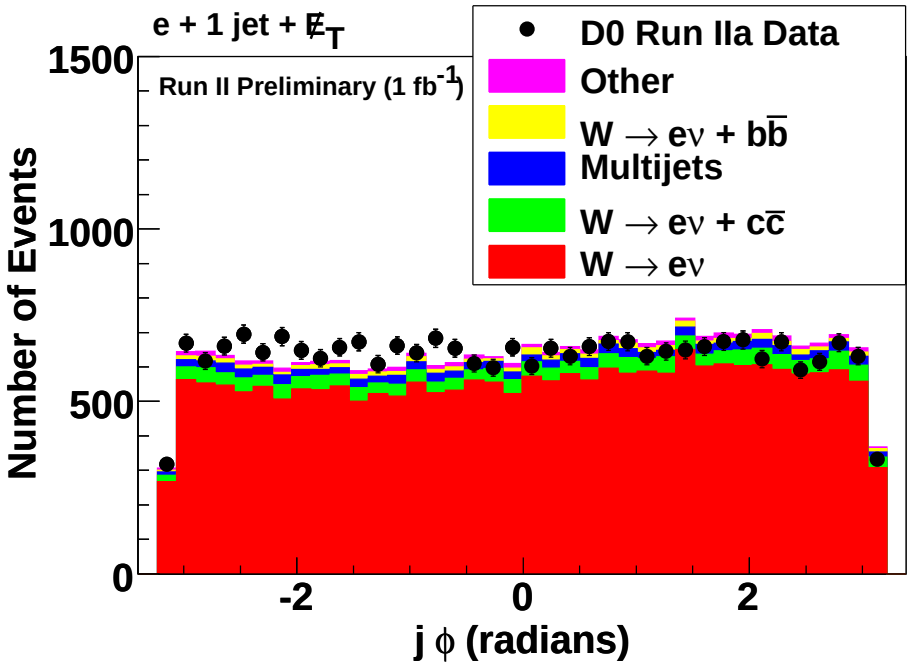


(b)

Figure 9.5: The plot 9.5(a) shows the $\Delta\mathcal{R}$ difference between the μ and trailing p_T jet. Figure 9.5(b) shows the $\Delta\eta$ distribution between the two jets in the e + 2 jets + $\not\!\!E_T$ final state.

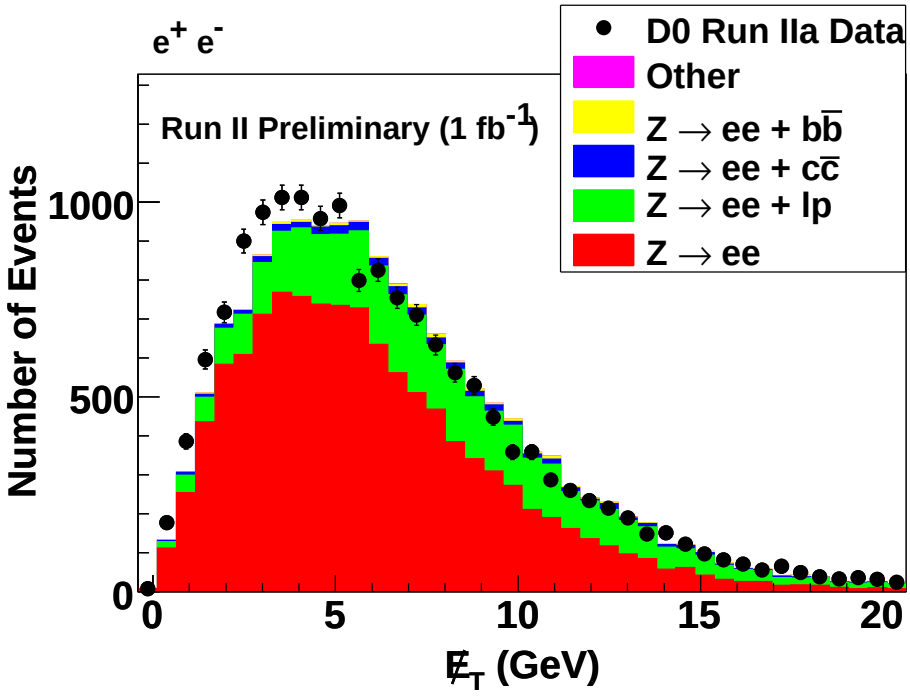


(a)

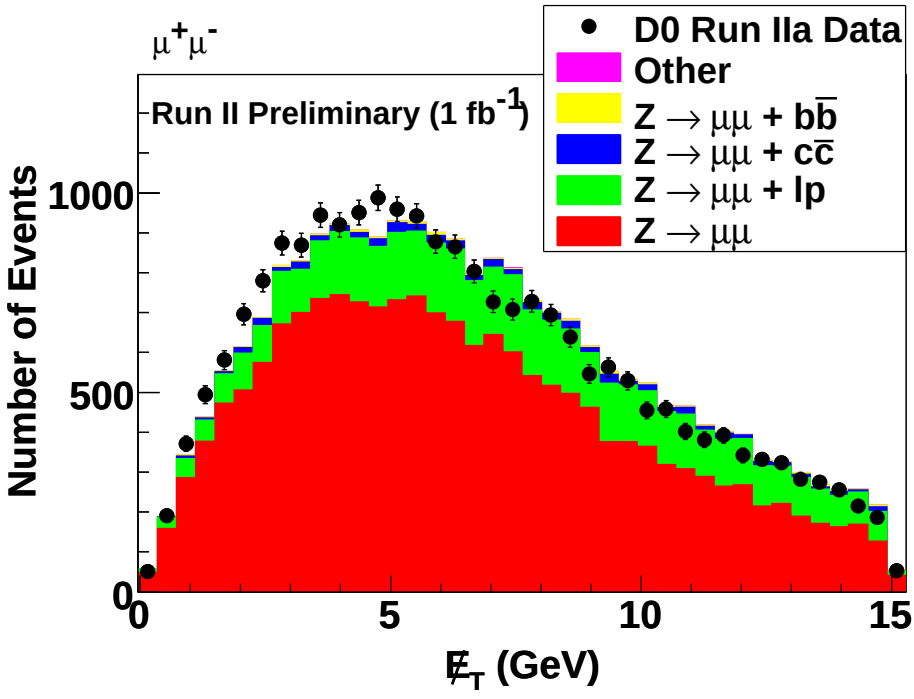


(b)

Figure 9.6: Figure 9.6(a) shows the invariant mass of the μ and the jet in a μ + jet + $\not\!\!E_T$ final state. Finally, 9.6(b) shows the ϕ distribution for the jet in the e + jet + $\not\!\!E_T$ final state. Each of these is tied to difficulties in spatial jet modeling.

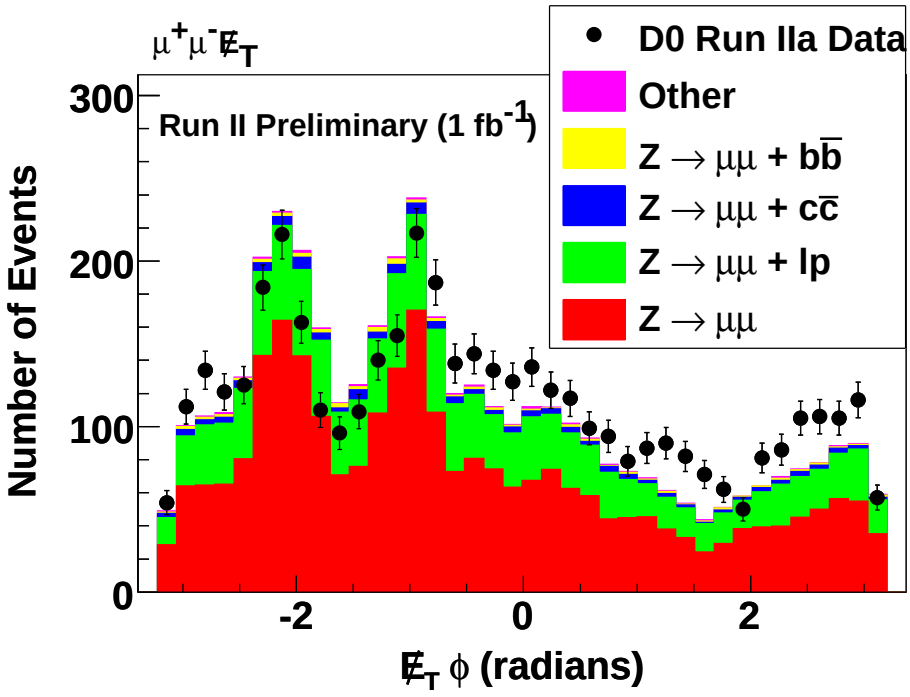


(a)

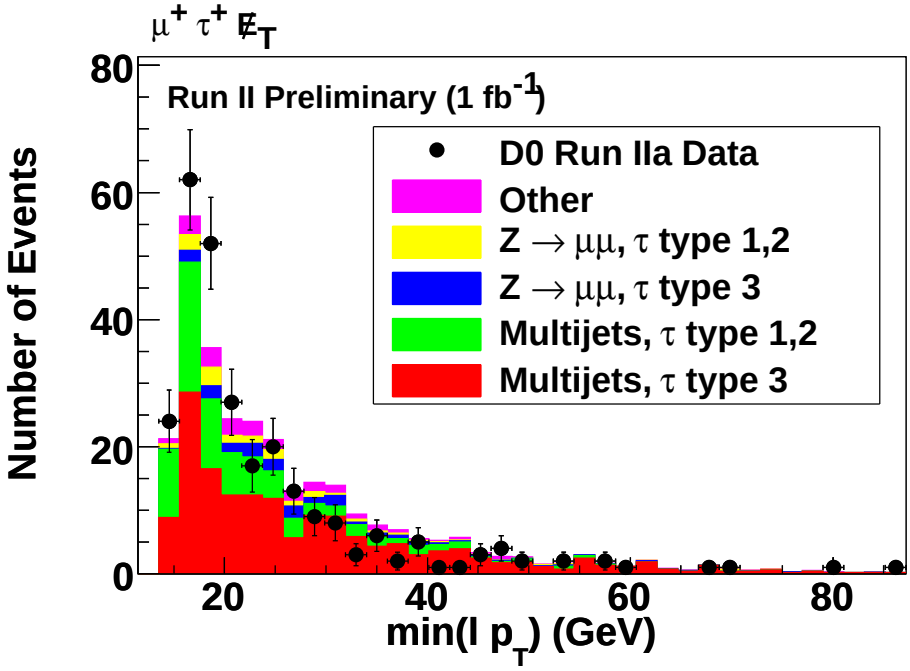


(b)

Figure 9.7: Plots 9.7(a) and 9.7(b) show the $\not\!\!E_T$ distribution in the opposite sign dielectron and dimuon final states. Both of these point to $\not\!\!E_T$ modeling issues in dilepton states.



(a)



(b)

Figure 9.8: Figure 9.8(a) shows the ϕ distribution of the $\not{E}_T$ in the dimuon state with large $\not{E}_T$, which also points to dilepton $\not{E}_T$ modeling issues. Finally, Figure 9.8(b) shows the minimum p_T of the μ and the τ for the same-sign $\mu\tau$ + $\not{E}_T$ final state which shows the difficulty in modeling the jet to τ misidentified background using loosened data cuts.

Table 9.2: The full list of VISTA shape discrepant histograms listed by VISTA final state.

VISTA FINAL STATE	Histogram	σ
μ + 2 jets + $\not\!\!E_T$	M(μ,j_2)	7.0
	M(μ,j_1,j_2)	4.6
	$\Delta\mathcal{R}(\mu,j_2)$	4.2
	M(W,j_2)	4.0
	$\not\!\!E_T$ ϕ	3.8
	$\Delta\eta(j_1,j_2)$	3.5
	M(j_1,j_2)	3.5
$\mu^{\pm}\mu^{\mp}\not\!\!E_T$	$\not\!\!E_T$ ϕ	3.1
$\mu^{\pm}\tau^{\pm}$	μ p_T	3.5
$\mu^{\pm}\tau^{\pm}\not\!\!E_T$	Min(ℓ p_T)	3.2
$\mu^{\pm}\mu^{\mp}$	$\not\!\!E_T$	3.4
μ + jet + $\not\!\!E_T$	M(W,j)	8.4
	j p_T	8.2
	M(μ,j)	7.2
	W p_T	5.7
	M_T(j,$\not\!\!E_T$)	5.3
	$\not\!\!E_T$ ϕ	5.3
	$\Delta\phi(\mu,\not\!\!E_T)$	4.8
	$\sum p_T$	3.3
e + 2 jets + $\not\!\!E_T$	$\Delta\eta$(j,j)	4.9
e + jet + $\not\!\!E_T$	j ϕ	3.3
$e^{\pm}e^{\mp}$	$\not\!\!E_T$	5.6
	$\sum p_T$	4.0

9.3 SLEUTH

All VISTA final states are input to SLEUTH, and the 180 final states are folded into 44 final states after applying global charge conjugation, rebinning in number of jets and using light lepton universality as described in Section 8.2.1. The several VISTA final states that show broad numerical excesses are found again with the SLEUTH algorithm as would be expected. One additional distribution crosses the discovery threshold of $\tilde{\mathcal{P}} < 0.001$, where $\tilde{\mathcal{P}}$ is the probability after all trials factors, described in detail in [91] and briefly in Appendix C. The final state that crosses the discovery threshold is $\mu^{\pm}$ + $e^{\mp}$ + $\not\!\!E_T$ as can be seen in Figure 9.9. Currently the evidence suggests that the muon tracking resolution is responsible for this discrepancy from the standard model. A large fraction of the events in the tail of the SLEUTH distribution have a muon with a very large p_T and large missing energy. With the present modeling of muon resolution, straight track events are underrepresented in the standard model background estimation. This state has 46 data events in the tail of this distribution compared to only 17 predicted by the Monte Carlo. A table of the top five SLEUTH final states that contain only leptons and jets is shown in Table 9.3. The known VISTA numerical excesses have been removed since this information is already known. All of these states are subject to the muon resolution issues discussed above. An example of another distribution expected to show the same issue is a single lepton with $\not\!\!E_T$. This is seen in Figure 9.10. A plot including all of the final state probabilities converted to units of σ can be seen in Figure 9.11. In this plot, e and μ states are combined according to light lepton universality. However, as we have seen, the systematic errors of the e and μ states differ. This distribution would be expected to improve if only electrons are searched because the electron energies are measured in the calorimeter. The corresponding electron-only plot is shown in Figure 9.12.

In the SLEUTH runs performed at CDF using a slightly different analysis strategy, the four most interesting observed final states were $\mu^{\pm}$ $e^{\pm}$, $\mu^{\pm}$ $e^{\pm}$ + 2 jets + $\not\!\!E_T$,

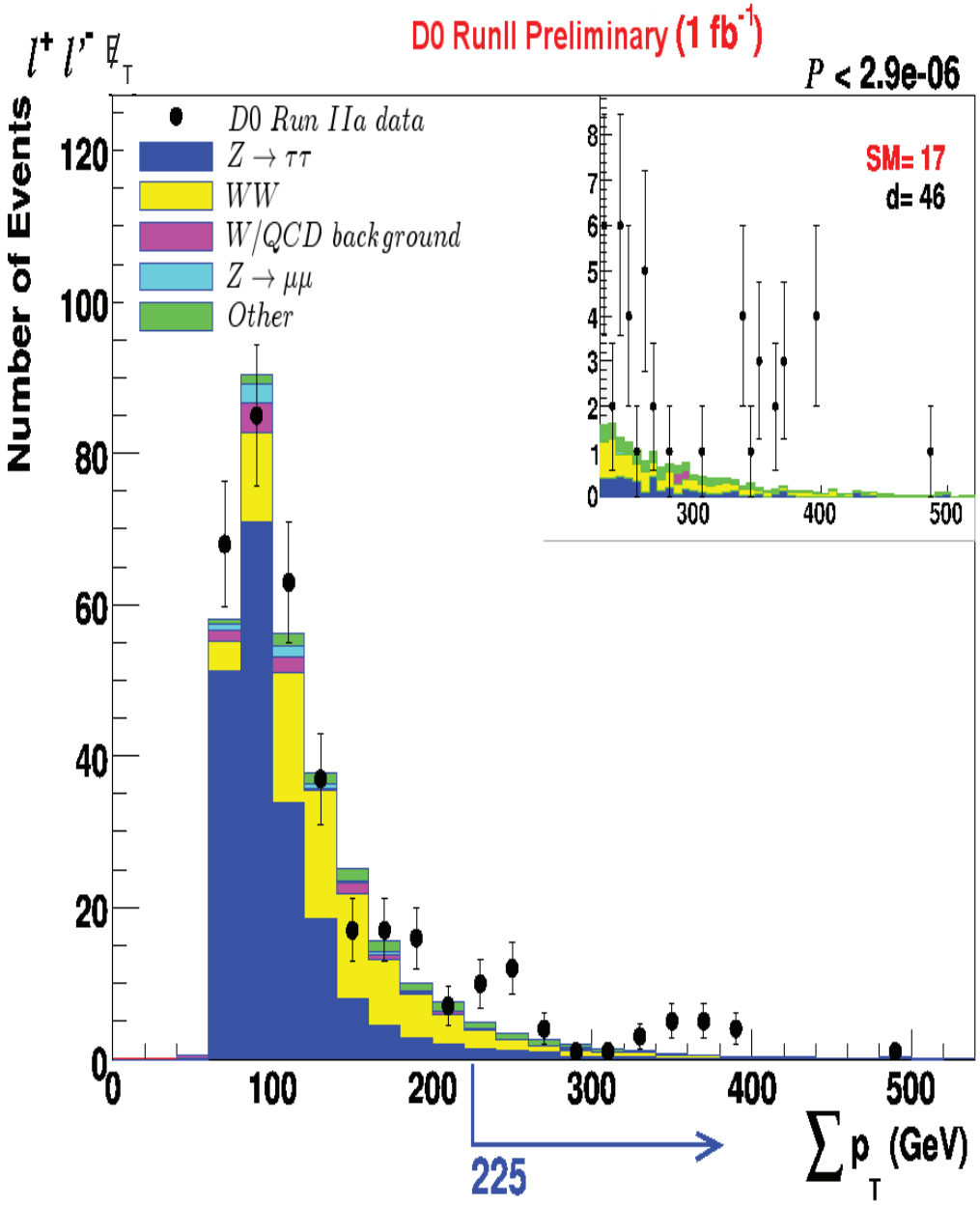


Figure 9.9: SLEUTH plot for opposite sign $\ell\ell^{'} + \not\!\!E_T$. The $\mathcal{P}$ value at the top right corner of the plot is the probability before final state trials factor.

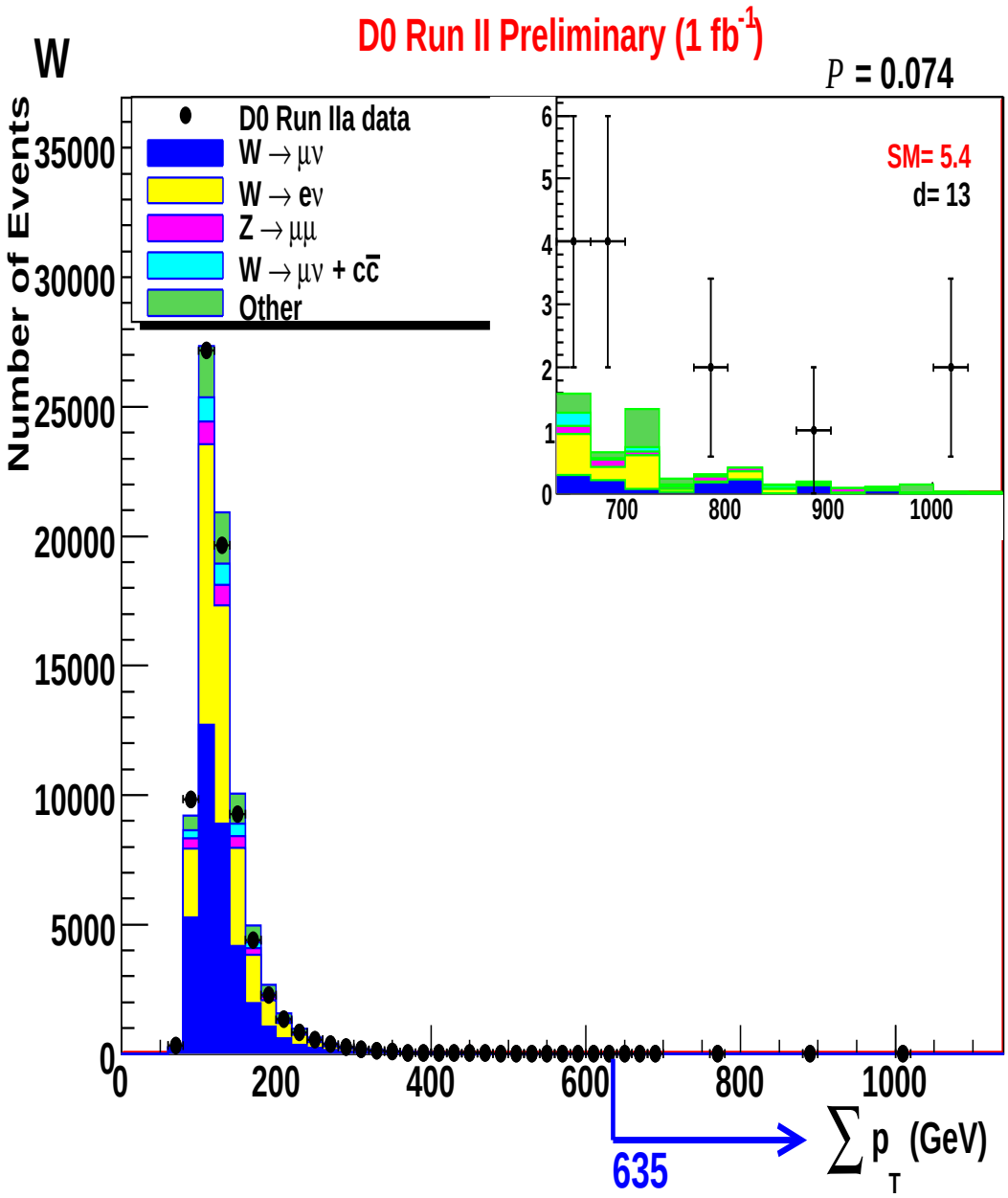


Figure 9.10: SLEUTH plot for $\ell + \not\!\!E_T$. The $\mathcal{P}$ value at the top right corner of the plot is the probability before final state trials factor. This plot shows the same issue in the tails of the distribution as Figure 9.9.

Table 9.3: The top five SLEUTH states with only leptons and jets. The value $\mathcal{P}$ represents the probability that the standard model background for an individual final state would have a fluctuation at any cut that would be more significant than what is seen in data. The variable $\tilde{\mathcal{P}}$ calculates the probability that one would observe a final state with $\mathcal{P}$ less than or equal to the one observed in data based on a statistical fluctuation.

Final State	$\mathcal{P}$	$\tilde{\mathcal{P}}^a$
$\ell^+\ell'^- + \not\!\!E_T$	2.9 E-6	0.00018
$\ell + \not\!\!E_T$	.00082	0.049
$\ell^+\ell'^-$	.0031	0.17
$\ell^+\tau^- + \not\!\!E_T$	0.006	0.31
$\ell^+\tau^+$	0.0066	0.33

[a] The value of $\tilde{\mathcal{P}}$ is not necessarily accurate below 0.001. The important check is whether the value drops below the threshold. Further discussion can be found in Appendix C and [91].

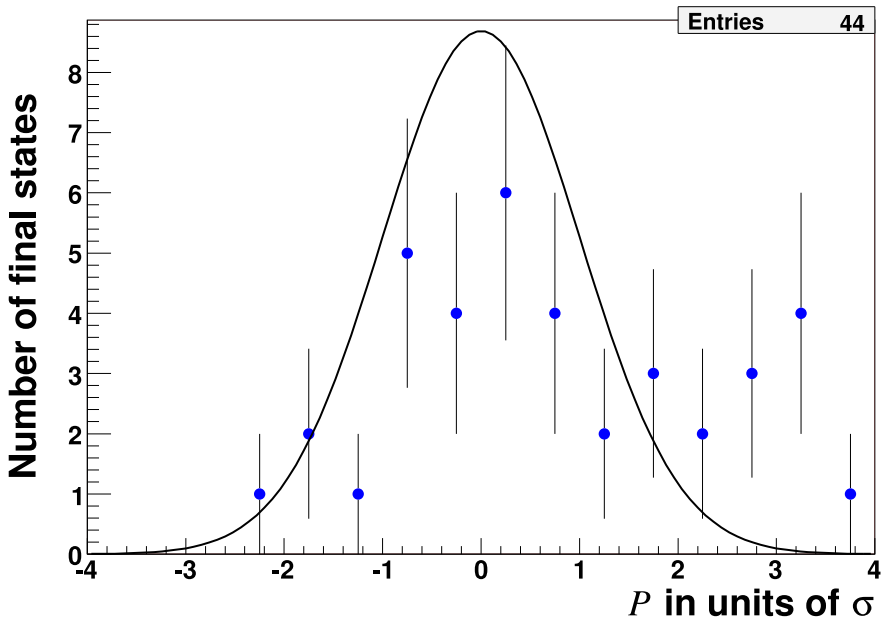


Figure 9.11: Distribution of final state SLEUTH probabilities converted into units of σ before inclusion of the final state trials factor.

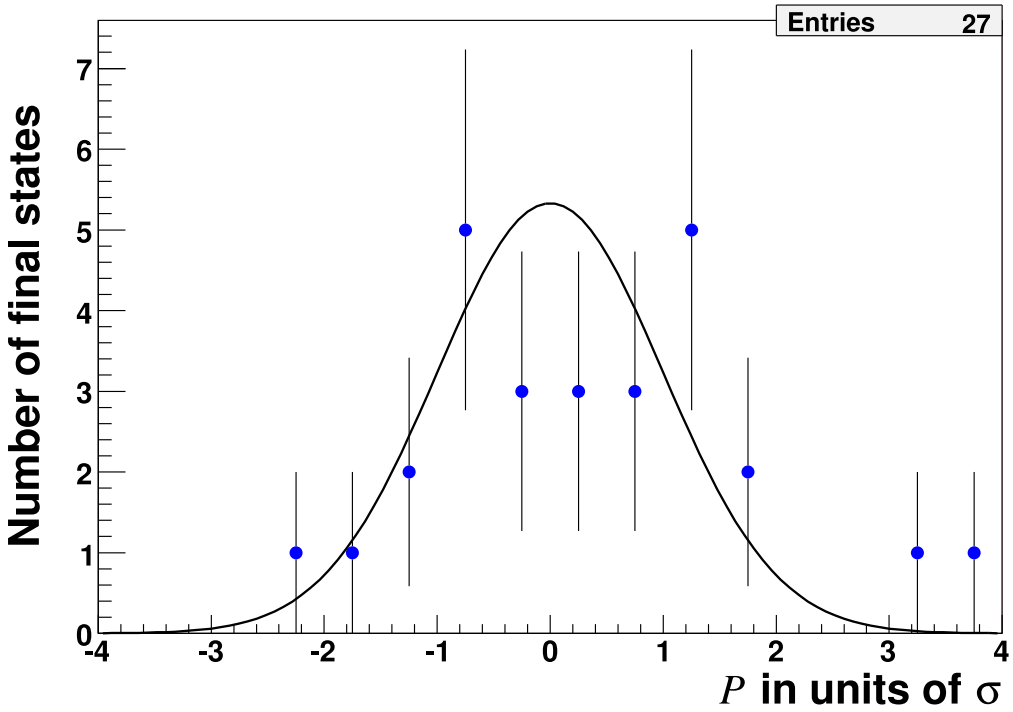


Figure 9.12: Electron-only distribution of final state SLEUTH probabilities converted into units of σ before inclusion of the final state trials factor. The two points in the tails show issues with jets misidentified as τ's.

$\mu^{\pm}\, e^{\pm} + \not\!\!E_T$ and $\ell^{\pm}\, \ell^{\mp}\, \ell^{'} + \not\!\!E_T$ with 2.0 fb^{-1} [28]. These states were also among the most discrepant observed by CDF for 927 pb^{-1} [91]. We show our results for these states in Figures 9.3, 9.3, 9.3 except for $\mu^{\pm}\, e^{\pm}$ + 2 jets + $\not\!\!E_T$ in which we have no data events. Figure 9.3 shows the similar final state where the muon and electron are of opposite sign rather than of the same sign where CDF sees the discrepancy. At DØ with 1.07 fb^{-1}, the $\mathcal{P}$ value is fairly low in Figures 9.3 and 9.3, but neither of these states are among the most discrepant.

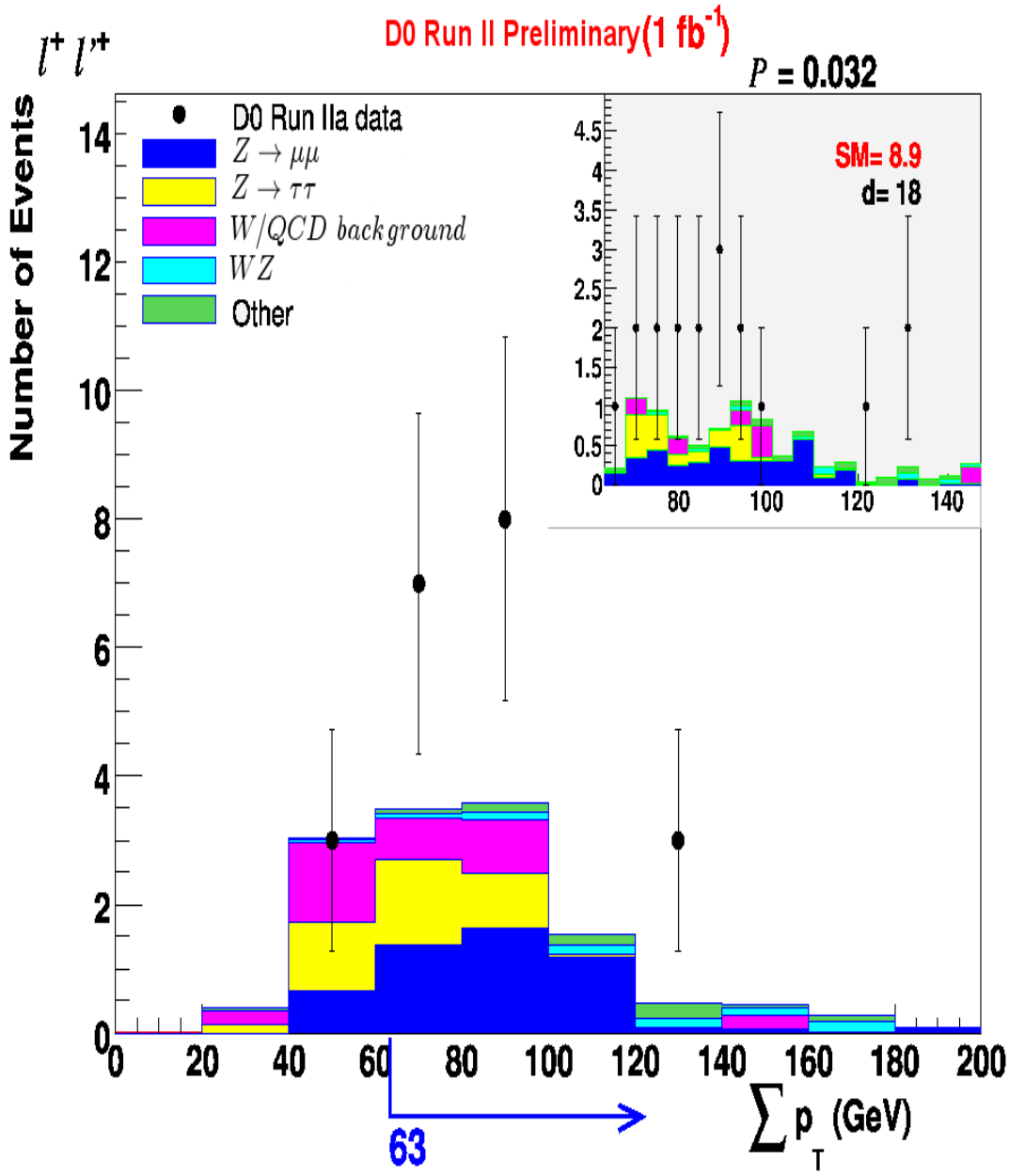


Figure 9.13: Check of most discrepant CDF plots from [28], same sign (SS) $\ell\ell'$. The $\mathcal{P}$ values at the top right corner of the plots are the probabilities before final state trials factors.

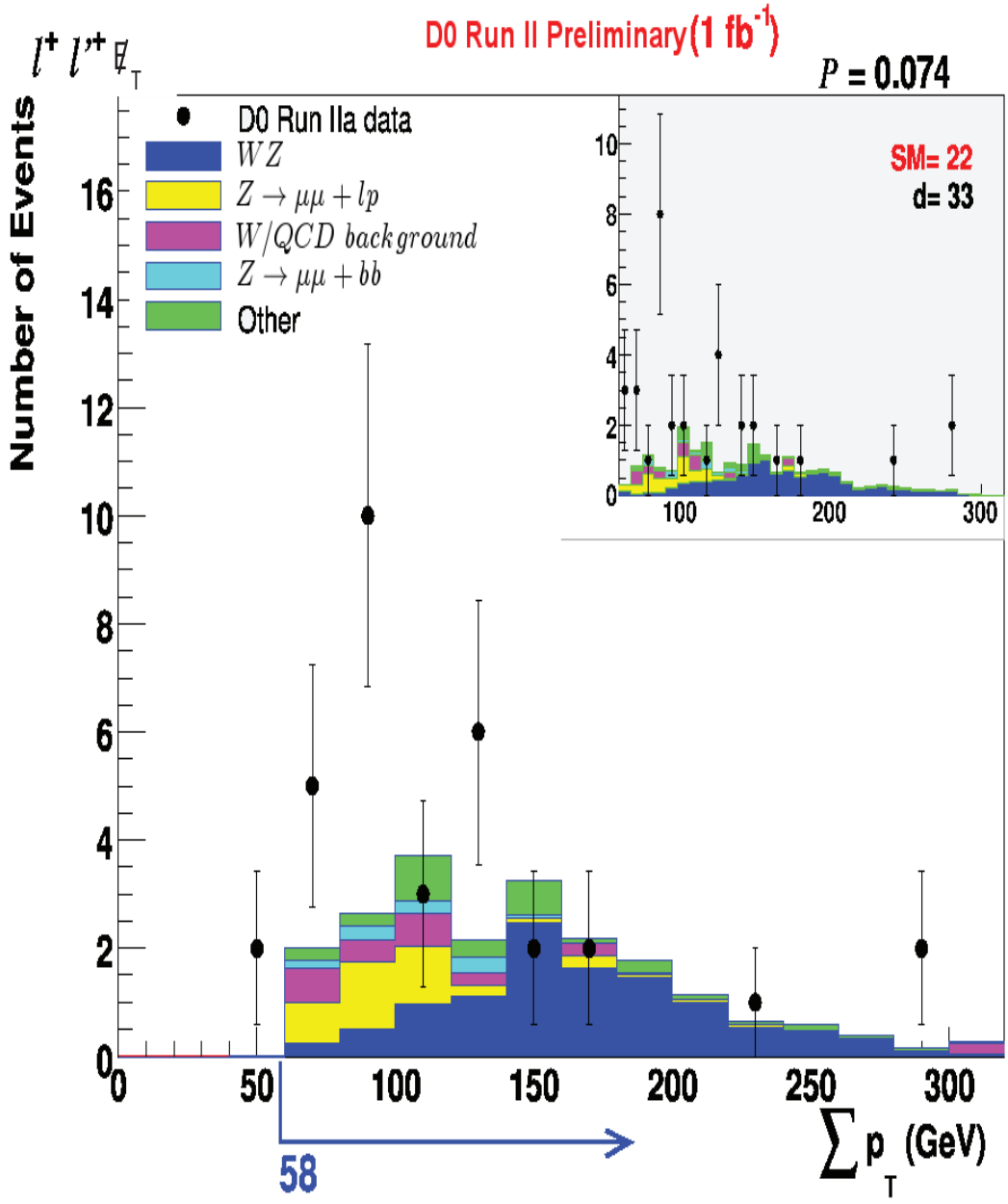


Figure 9.14: Check of most discrepant CDF plots from [28], same sign $\ell\ell' + \not{E}_T$. The $\mathcal{P}$ values at the top right corner of the plots are the probabilities before final state trials factors.

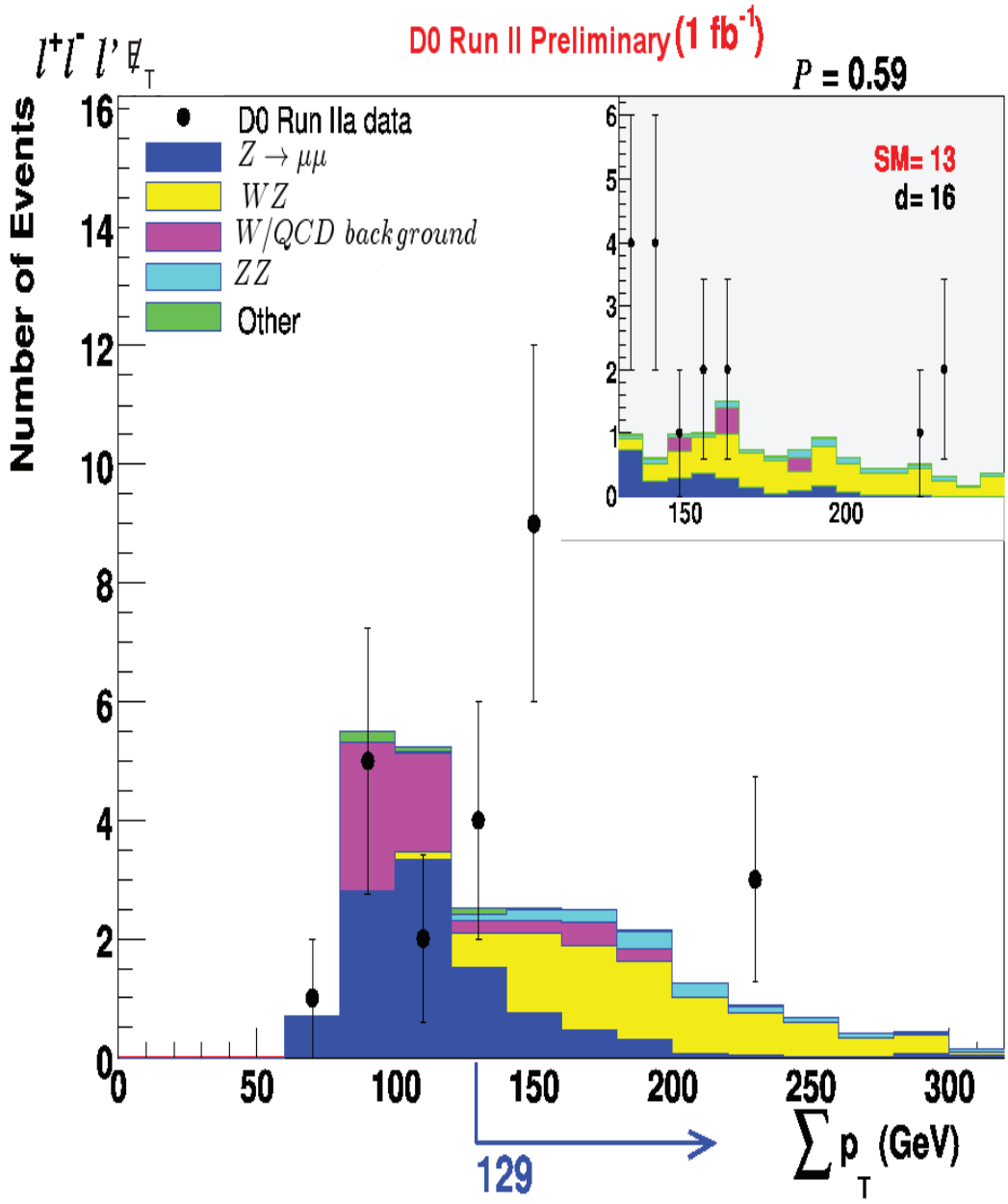


Figure 9.15: Check of most discrepant CDF plots from [28], $\ell^{\pm}\ell^{\mp}\ell' + \not{E}_T$. The $\mathcal{P}$ values at the top right corner of the plot is the probability before final state trials factors.

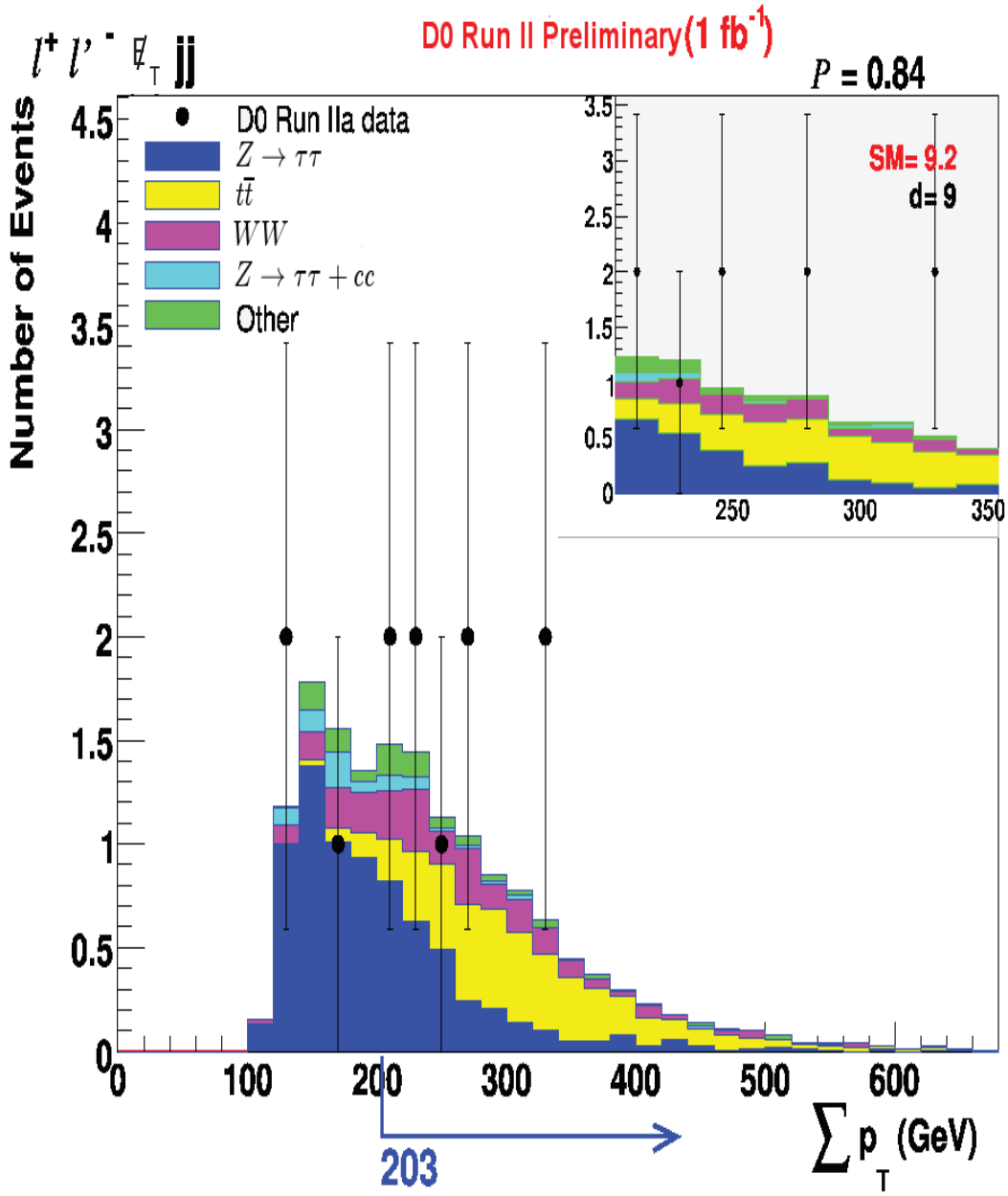


Figure 9.16: Since there are no data events in the for DØ in the descrepant CDF state, $\mu^{\pm}\ e^{\pm}$ + 2 jets + $\not\!\!E_T$, the distribution for $\mu^{\pm}\ e^{\mp}$ + 2 jets + $\not\!\!E_T$ is shown. The lack of data in 1 fb^{-1} shows that we do not see the same data excess in that final state.

Chapter 10

Conclusions

This analysis was an attempt to answer the basic question, "Do we see what we expect?". In our attempt to answer this, we performed a broad search for new physics over 1.07 fb^{-1} of data collected in Run II of the Fermilab Tevatron Collider at the DØ experiment. A total of 180 exclusive data final states and 9,335 relevant kinematic distributions were compared to the complete standard model background predictions using the VISTA algorithm. Only four out of 180 exclusive final states show a statistically significant discrepancy. Given the known modeling difficulties in all four final states, we refrain from attributing the observed discrepancies to new physics. A quasi-model independent search for new physics was also performed using the algorithm SLEUTH by looking at the regions of excess on the high-$\sum p_T$ tails of exclusive final states. Only $\mu^{\pm} + e^{\mp} + \not\!\!E_T$ surpasses the discovery threshold beyond the obvious excesses noticed in VISTA. This final state is potentially interesting for new physics processes. Several classes of theories, such as supersymmetry, can produce high momentum leptons with large missing energy, due to a non-interacting massive particle. The observed discrepancy in this analysis, however, strongly points to difficulties in modeling the muon p_T resolution. It is possible that there is a residual signal behind the known resolution issues, but we currently have no compelling case for this possibility. Further analysis of this same data set (subsequent to the work in

this thesis) also points to the likelihood that this is due to systematic underestimation of the muon resolution effects by the detector simulation.

While it is disappointing that we were unable to find clear signs of new physics in our data, the search accomplished two important tasks. First, we found that the vast majority of high-p_T data at DØ could be described through physics simulation of the standard model and the GEANT-based description of the detector. The ability to describe high-p_T standard model physics processes through Monte Carlo event generation requires a combination of calculations that can only be done with supercomputers, the integration and collation of information from physics experiments throughout the world, and a deep understanding of nature such that interactions where no calculation techniques are available can be estimated in exacting detail. This is then tied to a description of a 5,000 ton, $30' \times 30' \times 50'$ detector, often requiring accuracy at the level of microns. The multipurpose detector requires the integration of materials ranging from silicon to uranium. The tracking system, for instance, has 800,000 individual silicon strips and 70,000 scintillating fibers. With an average of 1.7 million proton/antiproton collisions each second, this analysis shows that this complicated system can provide a good description of the data (with a handful of well-motivated correction factors). Secondly, on the few areas where the detector description was less than perfect, such as the modeling of the curvature resolution of very straight tracks or the generator-level implementation of photonic radiation, this analysis has shined a strong light. These issues were not simply corrected away by looking at data outside the region of interest but were highlighted and brought to the attention of the collaboration. This information will help provide crucial insight as the DØ detector modeling and Monte Carlo generation is further improved, as well as point to areas of interest as future detector experiments are brought online.

In conclusion, the search for new physics tests our understanding of nature and the limits of technology. In searching for the answers to humanity's most fundamental

questions, even a null result provides profound insights.

Appendix A

Level 2 Global

The Level 2 triggering system was created to bridge the gap between triggering that could be done strictly with electronics (the Level 1 system) and the more detailed triggering using the full detector readout (the Level 3 system). Further detailed information on the Level 2 Global crate and the Level 2 triggering system in general can be found in [93], [20]. Much of the information in this section is adapted from these sources. The speed required for this intermediate region necessitates the use of pared down software combined with some firmware components. The original design parameters included a call for an input rate of 10 kHz, an output rate of 1 kHz with no more than 5% deadtime. As the triggering for Run II came together, it was found that the Level 1 system would have its output limited by the readout of the central tracking system. This kept the Level 1 output below a maximum of 2 kHz. This allowed Level 2 to operate with a rejection rate of around 50%. With a factor of two rejection rate, Level 2 can maintain a very high efficiency for physics objects while allowing the implementation of a more complex triggering scheme than originally imagined.

The Level 2 system consists of six processing crates that are connected to individual detector subsystems and the Level 1 trigger. Each of these crates runs an identical executable on a βeta processor [94] with individual configuration files designed for

each type of processing. The five preprocessors create basic physics objects from one particular subsystem. For example, the L2MUC processor (L2 central muon preprocessor) takes inputs from the scintillators and muon PDTs and creates basic muon objects from this information. Each of these preprocessor objects is formatted in a manner consistent with a common format for headers and trailers, and sent on to the central processing crate, L2 Global.

The L2 Global processor is responsible for making the L2 trigger decision. It takes all of the preprocessor objects, performs further processing, and determines whether any of the trigger terms included in the trigger list are met. If any of the triggers pass, the event is passed to L3 for further triggering. If all fail, then the event is rejected and triggering continues with the next event.

A.1 Data Flow

The flow of data across the Level 2 system can be seen in Figures A.1 and A.2. Definitions of each individual component can be found in the references. Inputs from the detector readout boards and Level 1 flow through the preprocessor crates and into Level 2 Global. The L2 Global crate layout can be seen in Figure A.3. The list of cards in the L2 Global VME crate is shown in Table A.1. Once each preprocessor crate creates its objects, the object data flows from each of the five preprocessor *βeta* [94] processing cards across the custom 128-bit wide, 100 ns Magic Bus [95] to the Magic Bus Transceiver [96]. Here the data is converted to a format compatible with Cypress Hotlinks cables [97] with a throughput of 16 MB/s. It is then sent to the input of the Magic Bus Transciever in the L2 Global crate. Next, it is sent across the Magic Bus in the Global crate to the Global *βeta* processor where L2 Global reads the lists of preprocessor objects and L1 framework information. The framework information includes the L1 decision mask: the list of which L1 terms passed and failed. L2 Global makes its decision based on the downloaded list of Level

2 triggers corresponding to the L1 terms that passed and using the input preprocessor objects. The list of preprocessor inputs to the L2 Global crate based on MBT source IDs is shown in Table A.2.

Photographs of the L2 Global crate can be seen in Figures A.4(a) and A.4(b).

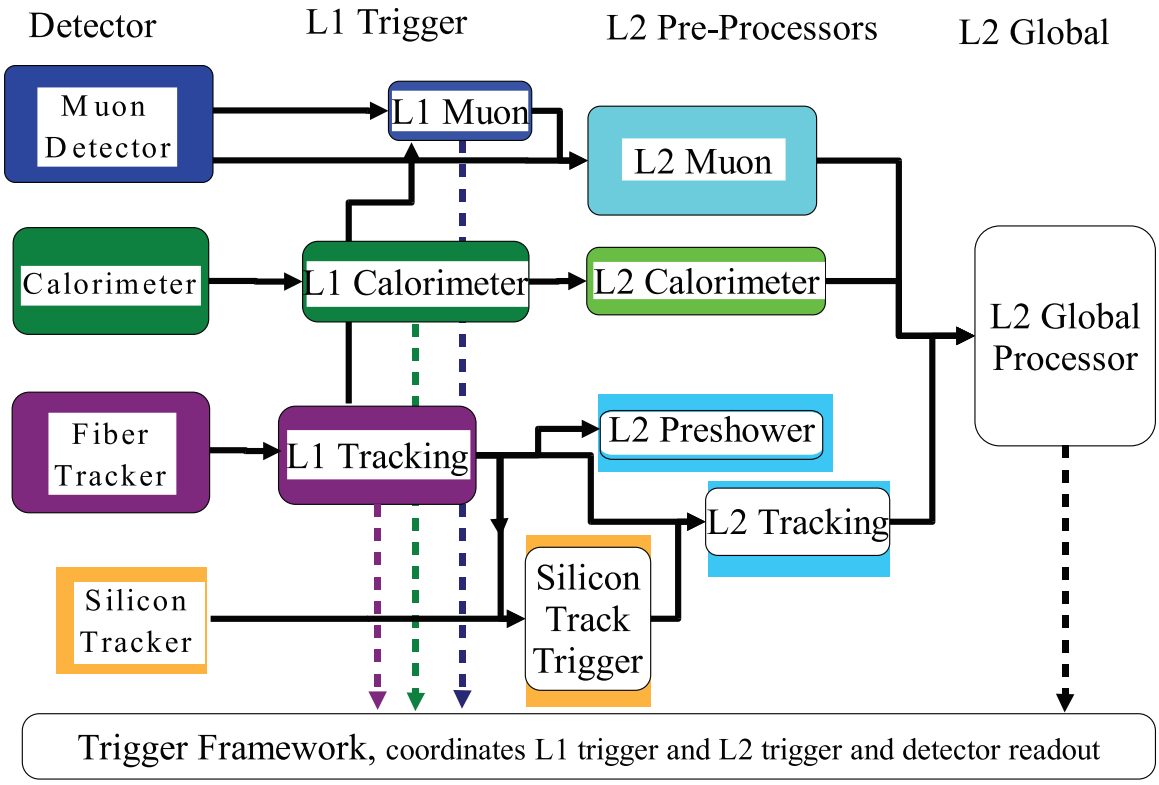


Figure A.1: Data flows from the front end detectors through the Level 1 and Level 2 trigger systems. The solid lines show the path of the detector data while the dotted lines show the path of the Level 1 and Level 2 triggers. The final Level 1 decision is determined by the trigger framework. The Level 2 system also sends the trigger decisions to the framework, but the Level 2 Global processor makes the final decisions on Level 2 event acceptance.

In all events, L2 Global prepares a header for L3 readout, but if the decision is to reject, the event is completely dropped. If L2 Global passes the event, then each of the Global objects that were used in passing any of the trigger conditions are tagged to save offline. This information is sent from the *βeta* processor across the VME [98] bus to the readout SBC (single board computer) [99] located in the crate. The Level 3 system uses this to draw events from Level 2 and into the Level 3 system. This

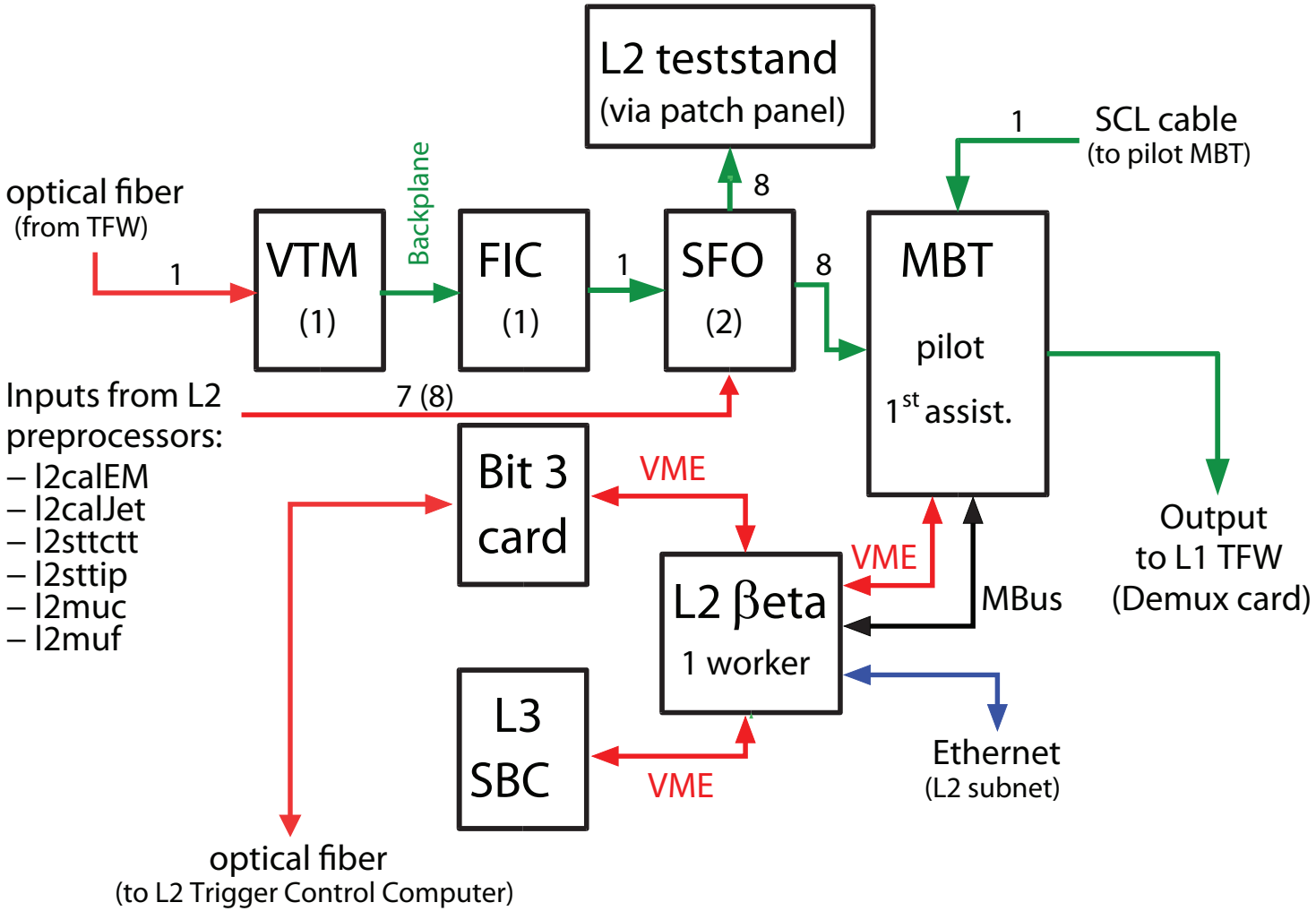


Figure A.2: Data flow within the L2 Global VME crate.

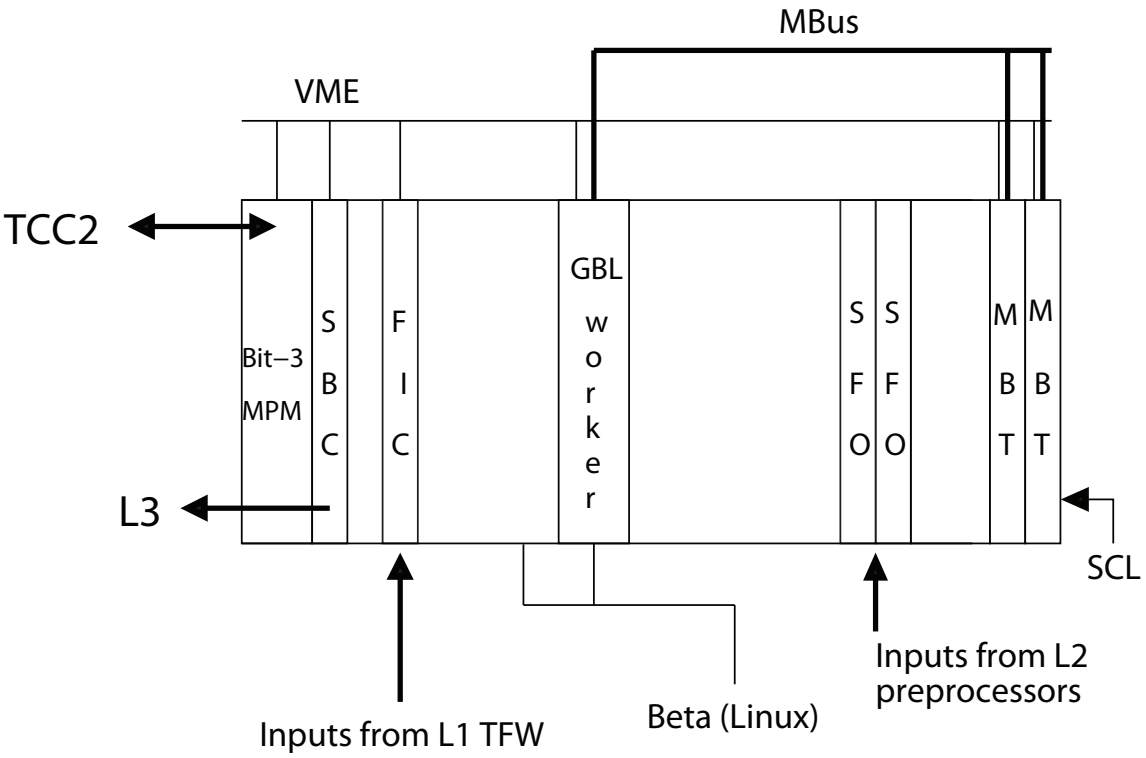


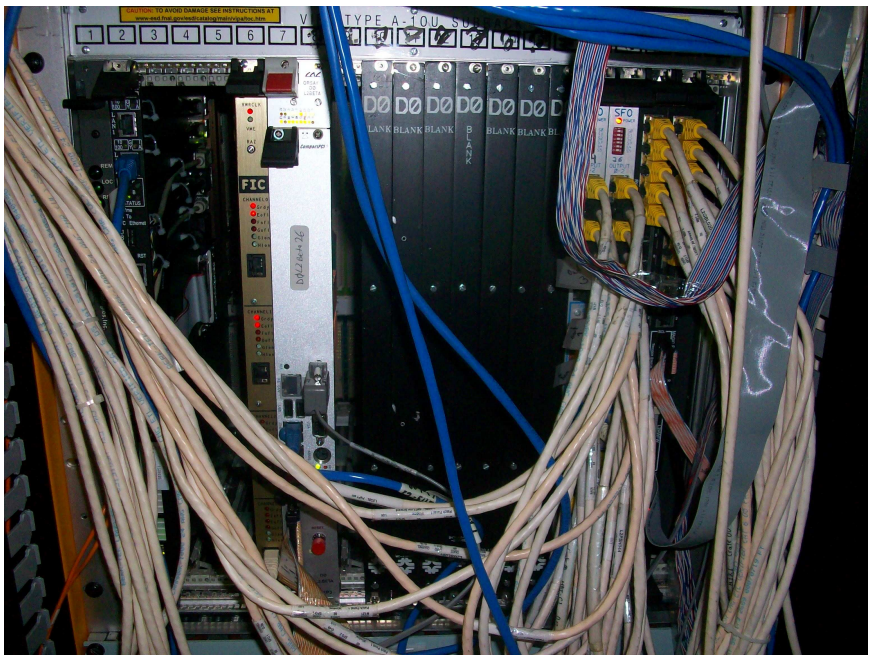
Figure A.3: Physical setup of the L2 Global cards within its VME crate.

object information is not used directly in the Level 3 trigger decision, but if the event passes, the objects are stored for possible offline analysis. A mask of all of the L2 decisions is also stored and used to determine which algorithms to run at Level 3.

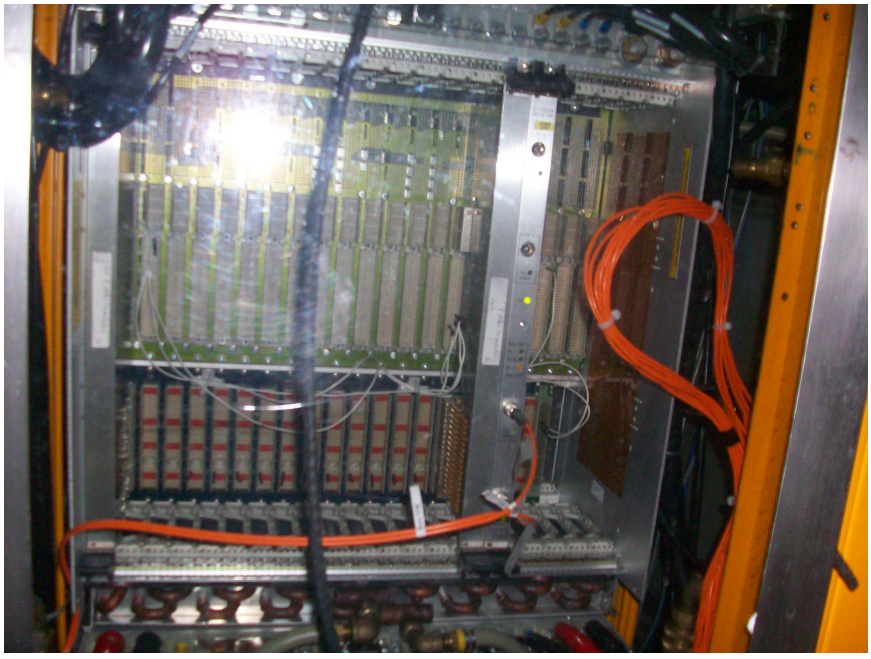
A.2 Trigger Configuration

As with the preprocessor crates, the configuration file that controls how L2 Global runs is stored on the Windows machine, DØ TCC2 (the trigger control computer). Level 2 Global differs from the preprocessor crates in that it must enter the decision state of the event loop and determine whether each event is accepted or rejected. The interpretation of each of the trigger conditions and the relation of each of these to the Level 1 decisions is a complex job exclusively done in the Global crate.

The trigger list is stored in the trigger database. The data acquisition coordination



(a)



(b)

Figure A.4: The Level 2 Global crate, front A.4(a) and back A.4(b) On the front, the visisble cards from left to right are the Bit 3 card, SBC, FIC, Beta, 2 SFOs and 2 MBTs. On the back, 2 MBUS terminators and the VTM for the L1 trigger framework input. The white jumpers shown in the photo of the back of the crate are needed for proper functioning of the readout SBCs. In order to run the L2 event loop at the teststand, the L3 handshaking must be faked.

Table A.1: The contents of the VME crate that houses the L2 Global processor. Each card is listed by the VME slot in which it resides.

Slot	**Card**
1	Bit 3 Multiport Memory
2	SBC
3-5	Spare
6	FIC input from L1 Trigger Framework
7-8	L2 Global Processor βeta
9-16	Spare
17	SFO
18	SFO
19	MBT
20	MBT
21	Spare

system (COOR) [100] retrieves this information from the database and sends it TCC2. From there, the triggers are sent to L2 Global. Each of these trigger conditions are considered only if the relevant Level 1 triggers have fired. There is a direct correspondence between a single Level 1 trigger and a group of Level 2 triggers. An example of a trigger used in this analysis is shown in Figure A.5.

A.2.1 Quick Overview of Relation Among Components of Level 2 Trigger Decision

- L1 Triggerbit- One of 128 possible trigger conditions at Level 1. Each of these bits is either 0 or 1 depending on whether the trigger passed.

TRIGGER Name/Version= **MUH5_LM15 / 2** , Use_Status= **used** , Current_Status= **obsolete** .
Created by Pompos on 2005-06-02 19:17:47
Description: **L1: wide region (CFT region) single muon trigger with tight scintillator and tight wire requirements and NOT Calorimeter unsuppressed readout. L2: pass events with at least one muon found with pT>5 GeV meeting MEDIUM quality(=2) requirements (no region requirement). L3: The trigger bit set to true if one muon is found with Pt>15 GeV.**

<table>
<tr><th>LEVEL</th><th>SCRIPT Name / Description</th></tr>
<tr><td>1</td><td>Level 1 SCRIPT Name/Version= mu1ptxbttx_ncu / 1
Description: Requires one muon in the forward region meeting tight pixel and tight wire requirements and NOT Calorimeter unsuppressed readout.</td></tr>
<tr><td>2</td><td>Level 2 GROUP Name/Version= MUON(0,5.,2,0,0,MUON5) / 1
Description: pass events with at least one muon found with pT>5 GeV meeting MEDIUM quality(=2) requirements (no region requirement).
<table>
<tr><th>LEVEL 2 SCRIPT Name / Description</th></tr>
<tr><td>Level 2 SCRIPT Name/Version= MUON(0,5.,2,0,0,MUON5) / 1
Description: pass events with at least one muon found with pT>5 GeV meeting MEDIUM quality(=2) requirements (no region requirement).</td></tr>
</table></td></tr>
<tr><td>3</td><td>Level 3 SCRIPT Name/Version= L3FMuon(MUON,1,0.,0.,2.5,15.,0.,LOOSE) / 2
Description: The trigger bit set to true if one muon is found with Pt>15 GeV.</td></tr>
</table>

Figure A.5: One of the triggers used in this analysis ("MUH5_LM15"). This is a screen shot taken from the trigger database for trigger list global_CMT-14.92 [29].

Table A.2: L2 Global data sources.

ID	**Source**
0	L1 Serial Command Link
1	L1 Hardware Framework
2	L2 Cal- EM
3	L2 Cal- Jets
4	L2 Cal- MET
5	L2 CTT- STT p_T
6	L2 CTT- STT ip
7	L2 MUC
8	L2 MUF
9	L2 PS- CPS (disabled)
10	L2 PS- FPS (disabled)
11	Spare
12	L2 CTT- CTT p_T (empty)
13-15	Spare

- Superscript- The superscript is a group of Level 2 trigger terms that are associated with a L1 triggerbit. There are 128 of these, one for each of the incoming L1 bits. If ANY of the scripts associated with a superscript passes, then the superscript (and the event) passes. This was used to easily expand the initial one-to-one correspondence between L1 and L2 triggers to accomodate several L2 triggers for each Level 1 trigger at the beginning of Run IIb.

- Script- This is the term that corresponds to a particular Level 2 trigger. Each script is made up of a filter and a number of objects that need to pass that filter. The parameters used in that filter are also specified in the script.

- Filter- The filter uses objects created by the tools or other filters to determine if they match the parameters set in the script. An example would be a minimum $\Delta\eta$ cut between a Global muon object from a muon filter and a Global jet object from a jet filter. The objects passing the muon and jet filters would have been determined from muon and jet tools described below.

- Tool- Level 2 tools are used to create Level 2 Global objects. These may be preprocessor objects directly, preprocessor objects with more complex properties or combinations of preprocessor objects. An example would be a preprocessor muon matched to an L2CTT (STT) track. Minimum requirements would be put on the muon p_T and a minimum distance would be needed to match the track to the muon.

- L2 Global Object- An L2 Global object is any object that is created by a global tool. For example, if an L2 jet tool is run, all of the preprocessor jet objects are looped over. Any of the objects that meet the criteria for the jet tool will be added to the list of global objects. These objects are then in turn used in the Level 2 filters.

- L2 Preprocessor Object- A preprocessor object is an object sent from one of the Level 2 preprocessors. These objects always contain only information that comes from a single preprocessor, which reflects just one part of the detector. For example, the calorimeter will send jet and EM preprocessor objects, L2MUC will send muon objects from the central muon system, etc.

- Tool Object List- Once a tool is run, it does not need to be run again. A tool will only be run if a particular filter requests it. If a different filter later requests the same tool, the objects passing that tool will not need to be recalculated. They will simply be read off the list of objects passing that particular tool. The object list is saved for the event and any object that contributed to the passing

of a trigger is saved and sent to Level 3.

- Filter Object List- Once one of the global objects that was part of the tool object list passes a filter, that object will be added to the filter object list. Once a filter is run one time, all of the objects that pass that particular filter are saved. If the filter were to be used again in a different script or as an input to a different filter, it would not need to be run again. The objects that passed that filter would simply be checked to see if they satisfy the higher level filter or script requirements. In the case of the script, only the number of objects would be further specified.

- Preprocessor Object List- As each stage of the decision is made, the preprocessor object associated with a particular Global object is saved. If a tool and filter pass as part of a passing trigger, all of the objects that could contribute to passing the trigger requirements are tagged and saved for offline analysis. When this happens, the preprocessor objects that are used in the creation of the global objects are also tagged and saved for offline analysis. In the end, the Global object and preprocessor object list are passed to Level 3.

A.3 Triggerbits, Superscripts, and Scripts

Which of the 128 Level 1 triggers that fire, determines which L2 processes will be run. The Level 1 trigger associated with the trigger above is shown in Figure A.6. Each of these corresponds to groups of Level 2 processes known as superscripts. There is an exact one-to-one correspondence between the L1 Trigger Bits and each L2 superscript. The list of which superscripts have passed is saved and sent to Level 3. Each superscript passes if ANY of an associated set of scripts passes. The superscript performs an OR of all of the scripts associated with it. Each of the scripts corresponds to a set of Level 2 triggers determined to most efficiently bring in physics objects while

limiting deadtime. The superscript associated with trigger "MUH5_LM15" is shown is Figure A.7.

The superscripts were not implemented to run directly as triggers. Their presence is based purely on the necessity to expand the old 128 possible trigger conditions to 1024 that are currently available [101]. Previously, only a single script could be run for each L1 trigger that fired. In order to expand this, another layer of complexity needed to be added, so that the superscripts now play the role that the individual scripts did previously. The trigger conditions are still applied at the script level, and the superscripts match individual L1 trigger decisions to the group of L2 trigger conditions that should be considered. An example script is shown in A.8.

Each script is made up of a particular filter (described in the next section) and the number of objects that need to pass the filter. As with the superscripts, the list of scripts that pass is saved and sent to Level 3 determining which Level 3 algorithms are run and for eventual offline analysis.

A.4 Filters and Tools

Filters are the conditions necessary for a particular script to pass. Each filter has a set of configurable parameters that are defined in the script. A filter may be used multiple times in different scripts with different object requirements. The filters set conditions on Global physics objects. Each Global physics object comes from preprocessor objects sent from the other Level 2 βetas, which can then be used directly, refined, or combined with other preprocessor objects. As an example of Global object creation, we can start with a preprocessor central muon. This can be refined by looking for overlap with a forward muon, and then checked for a match with an STT track. The new, more complex muon would then be added to the list of Global objects that can be used in the L2 filters. Each of these objects is created only when the need arises. If there are no filters that need this particular object, then the tool

Level 1 Term Name/Version= **MUO(1,ptx,B,T,T,X) / 1** , Use_Status= **used** , Current_Status= **current** , created by billl on 2004-05-03 16:07:46
Description: **requires one muon in the forward region meeting tight pixel and tight wire requirements.**
This Term is based on a Level 1 Object Name/Version= **MUO / 2.00**

Order	Parameter	Type	Value	Description	Default
1	Mult	int	**1**	one muon	-
2	Pt	string	**ptx**	no momentum selection	-
3	Eta	string	**B**	'between': |eta|<2.0 excluding |eta|<1.0	-
4	Scint	string	**T**	meeting tight pixel requirements	-
5	Wire	string	**T**	meeting tight PDT requirement	-
6	Option	string	**X**	no optional requirement	-

Figure A.6: The Level 1 trigger term associated with MUH5_LM15.

L2 Group Name/Version= MUON(0,5.,2,0,0,MUON5)/ 1 , Use_Status=used , Current_Status=current . Created by trigger_owner on 2006-08-30 09:50:56
Description: pass events with at least one muon found with pT>5 GeV meeting MEDIUM quality(=2) requirements (no region requirement).

Level 2 SCRIPT Name / Description
Level 2 Script Name/Version= MUON(0,5.,2,0,0,MUON5)/ 1 Description: pass events with at least one muon found with pT>5 GeV meeting MEDIUM quality(=2) requirements (no region requirement).

Figure A.7: The Level 2 superscript associated with MUH5_LM15. In the trigger database, the superscripts are known as L2 Groups.

L2 SCRIPT Name/Version=
MUON(0,5.,2,0,0,MUON5) / 1 , Use_Status= **used** ,
Current_Status= **current** . Created by Pompos on
2005-06-02 18:58:13
Description: **pass events with at least one muon found with pT>5 GeV meeting MEDIUM quality(=2) requirements (no region requirement).**

ORDER	Includes Level 2 Global TERM(s):	Preprocessor(s):
1	Term Name: **MUON(0,5.,2,0,0,muon5) / 1** , Minimum number required: **1**	missing !?!

Figure A.8: The single Level 2 script associated with MUH5_LM15. Since this trigger was used for Run IIa, there is only one script associated with the superscript. In Run IIb, more than one script are allowed for each superscript.

will never be run, saving processing time.

A full list of tools available to L2 Global is shown in Table A.3.

Table A.3: Full list of tools available to L2 Global with the configurable parameters.

Tool	**Parameter**	**Description**
Muon	L1PTTHRESH	Minimum L1 p_T
	REQUIRETRACK	Is match to track required?
	KINEFROMTRACK	Use kinematic track information rather than that of local muon stub?
	TRACKWINDOW	Maximum distance in *iphi* to match a track to the muon
	IPHI	

Continued on Next Page...

Table A.3 (cont'd)

Tool	Parameter	Description
EM	MINET	Minimum E_T
	MINNEIGHBOR	Threshold for which, if central cluster η neighbor is below, it will be turned into 2 separate EM tower objects
	ETACENET	
	'PHICENET	Same for central cluster ϕ neighbor
	'ETAFWNET	Same for forward cluster η neighbor
	'PHIFWNET	Same for forward cluster ϕ neighbor
	MINSINGLE TOWEREMFRAC	Value for which if EM fraction is greater, it will turn that cluster into two EM tower objects
	MINSINGLE TOWERET	In E_T for single tower EM object. Overrides MINET if MINNEIGHBORET is true.
	REQUIRETRACK	If set to 1, require track match to be found (0 is false).
	TRACKFILER	Filter used to define track that is to be used with track match.
	TRACKWINDOW	ϕ window to match track with EM cluster
	IPHI	
	REQUIRECPS	1 = require a CPS match (0 is false)
	CPSWINDOWIETA	η match window with CPS
	CPSWINDOWIPHI	ϕ match window with CPS
	MAXEM	Maximum number allowed to pass

Continued on Next Page...

Table A.3 (cont'd)

Tool	**Parameter**	**Description**
Commission		
InvMass	MININVMASS	Minimum invariant mass
	NFILTERS	Number of input files (1 or 2 depending on whether we are looking at two of the same type of objects)
	FILTER0	First filter
	FILTER1	Second filter
MET	REVERTEX	Find actual vertex rather than use default of zero
	FILTER0	Vertex filter
Jet	MINET	Minimum E_T
MJT	MINET	Minimum E_T of jets to be included in calculation
Tau	MINET	Minimum E_T
	MAXTAUS	Maximum number of taus
	MINRATIO	Minimum hadronic isolation fraction
	REQUIRETRACK	Require a track?
	TRACKWINDOW	How close in *iphi* to match tau with track?
	IPHI	
	TRACKFILTER	Filter used to select track

Continued on Next Page...

Table A.3 (cont'd)

Tool	**Parameter**	**Description**
BTag	MINET	Minimum p_T actually of tracks to be included in event b-tag calculation
	MINIPSIG	Minimum impact parameter significance for track
	MAXCHISQ	Maximum χ^2 value for track
Track	MINET	Minimum p_T of track
	TRACKSOURCE	Type of input track (now just p_T or ip ordered STT tracks)
	REQUIREL1ISO	Require L1 isolation confirmation
	REQUIREL1PS	Require L1 preshower confirmation
	L2ISOTYPE	Type of L2 isolation required. 0 = no requirement, 1 or 2 = require 1- or 3-prongs, 3 = require 1-prong
	MAXCHISQ	Maximum χ^2 allowed
	IPSIG	Minimum impact parameter significance
Vtx	MINET	Minimum p_T of tracks used in vertex finding
	MAXCHISQFIRST	Maximum χ^2 for tracks in first pass of vertex finding
	MAXCHISQSECOND	Maximum χ^2 in second pass

Continuing with the muon example, a filter may simply add a tighter transverse momentum cut to the muon object. It could also look at the η or ϕ separation between two muons, or a muon and a jet created by a different tool. The script containing

Level 2 Tool Term Name/Version= **muon5 / 1** , Use_Status= **used** , Current_Status= **current** , created by pompos on 2005-06-02 17:21:21
Description: **Find all muons.**
This Term is based on a Level 2 Tool Object Name/cvsPackage /Version= **MUON / l2gblmuon / 3**

Order	Parameter	Type	Value	PVTVersion	Default
1	MINOR_VERSION	int	**0**	-	-
2	L1PTTHRESH	int	0	-	0
3	TRACKWINDOWIPHI	int	5	-	5
4	REQUIRETRACK	int	0	-	0
5	TRACKFILTER	filter	PASS100	1	pass100
6	KINEFROMTRACK	int	0	-	0

Figure A.9: The Level 2 muon tool associated with MUH5_LM15. This is actually the only L2 muon tool used in this particular trigger list.

that filter then may require at least two of the muons. A total of 1024 total trigger conditions can be applied in L2 for each trigger list.

A full list of filters available is shown in Table A.4.

Table A.4: Full list of filters available to L2 Global with the configurable parameters.

Filter	**Parameter**	**Description**
Muon	MINET	Minimum E_T
	QUALITY	Minimum quality (based on number of hits)
	PROMPT	Minimum timing quality (based on scintillator times)
	SIGN	Required sign of muon to pass
	TOOL	Input tool
EM	EMFRAC	Minimum EM fraction
	ISOFRAC	Maximum isolation fraction
	MINET	Minimum E_T
	MAXEM	Maximum number allowed to pass
	MINLIKELIHOOD	Minimum value of the EM likelihood
	TOOL	Input tool
EMCalib	REQUIRENORTH	Require North fired
	REQUIRESOUTH	Require South fired
	REQUIRECENT	Require Central fired
	REQUIREANY	Require any fired

Continued on Next Page...

Table A.4 (cont'd)

Filter	**Parameter**	**Description**
BJet	BJETMIN	Sum of Jet p_Ts and Track *iphi*s
	JETFILTER	Jet filter
	TRACKFILTER	Track filter
Eta	FILTER	Filter used to choose object for *ieta* cut
	IETAMIN	Minimum *ieta* (for first region if NREGIONS is greater than 1)
	IETAMAX	Maximum *ieta* (for first region if NREGIONS is greater than 1)
	NREGIONS	Number of *ieta* regions to consider
	IETAMIN2	If more than one region, second min *ieta*
	IETAMAX2	If more than one region, second max *ieta*
	IETAMIN3	If more than two regions, third min *ieta*
	IETAMAX3	If more than two regions, third max *ieta*
	IETAMIN4	If four regions, fourth min *ieta*
	IETAMAX4	If four regions, fourth max *ieta*

Continued on Next Page...

Table A.4 (cont'd)

Filter	**Parameter**	**Description**
EtaPhiSep	NFILTERS	Number of input filters
	IETAMINSEP	Minimum η separation value
	IPHIMINSEP	Minimum ϕ separation value
	FILTER0	First input filter
	FILTER1	Second input filter
EtaSep	NFILTERS	Number of input filters
	IETAMINSEP	Minimum η separation value
	IETAMAXSEP	Maximum η separation value
	FILTER0	First input filter
	FILTER1	Possible second filter
FailAll		
HT	HTMIN	Minimum H_T
	NFILTERS	Number of input filters
	FILTER0	First input filter
	FILTER1	Second input filter
	FILTER2	Third input filter
InvMass	MININVMASS	Minimum invariant mass
	MAXINVMASS	Maximum invariant mass
	TOOL	Input tool
MET	MINET	Minimum $\not\!\!E_T$ required
	TOOL	Input tool
Phi	IPHIMIN	Minimum *iphi*
	IPHIMAX	Maximum *iphi*
	FILTER	Input filter

Continued on Next Page...

Table A.4 (cont'd)

Filter	**Parameter**	**Description**
PhiSep	NFILTERS	Number of input filters
	IPHIMINSEP	Minimum ϕ separation value
	IPHIMAXSEP	Maximum ϕ separation value
	FILTER0	First input filter
	FILTER1	Second input filter
PhiSepVeto	NFILTERS	Number of input filters
	IPHIMINSEP	Minimum ϕ separation value for veto
	IPHIMAXSEP	Maximum ϕ separation value for veto
	FILTER0	First input filter
	FILTER1	Second input filter
RSep	NFILTER	Number of input filters
	RMINSEP	Minimum $\Delta\mathcal{R}$ separation value
	FILTER0	First input filter
	FILTER1	Second input filter
RandomPass	PASSPERCENT	Percent of filters to pass
	TOOL	A commissioning tool
Spher	SPHERMIN	Minimum sphericity
	APLANMIN	Minimum acoplanarity
	NFILTERS	Number of filters
	FILTER0	First filter
	FILTER1	Second filter
	FILTER2	Third filter

Continued on Next Page...

Table A.4 (cont'd)

Filter	**Parameter**	**Description**
TimeDelay	DISTRIBUTION	Type of distribution (Delta, Gaussian, Exponential, Hyper-Exponential)
	MEANDELAY	Mean time delay
	MEANDELAY2	Second mean parameter used for hyper-exponential
	WIDTH	Width of Gaussian
	PROBABILITY	Probability of using first exponential in hyper-exponential with time delay, MEANDELAY. 1-PROBABILITY is probability of using exponential with mean MEANDELAY2
	TOOL	Input commissioning tool
TranMass	MINTRANSMASS	Minimum value of transverse mass
	FILTER0	First filter
	FILTER1	Second filter
Jet	MINET	Minimum E_T
	MAXJETS	Maximum number of jets allowed
	TOOL	Input tool
MJT	MINMJT	Minimum E_T
	TOOL	Input tool

Continued on Next Page...

Table A.4 (cont'd)

Filter	**Parameter**	**Description**
CalCalib	REQUIRENORTH	Require North fired
	REQUIRESOUTH	Require South fired
	REQUIRECENT	Require Central fired
	REQUIREANY	Require any fired
Tau	MINET	Minimum E_T
	MINRATIO	Minimum ratio of two highest energy hadronic towers to all jet towers
	TOOL	Input tool
BTag	MINGOOD	Minimum value of b-tagging parameter when just looking at best track
	MINALL	Minimum value of b-tagging parameter when looking at good tracks
	MINGOODTRACKS	Minimum number of good tracks in event
Track	MINET	Minimum p_T of track
	QUALITY	Minimum quality
	IP	Minimum impact parameter
	IPSIG	Minimum impact parameter significance
	TOOL	Input tool
Vtx	MAXVTXZ	Minimum MULM best cut value
	MINTRACKS	Minimum number of good tracks
	TOOL	Input tool

In the end, we have a list of tools which contain the Global objects for which

Level 2 Filter Term Name/Version= **MUON(0,5.,2,0,0,muon5) / 1** ,
Use_Status= **used** , Current_Status= **current** , created by pompos on
2005-06-02 17:35:04
Description: **Pass all muon tool candidates meeting quality='medium' requirements and has p_T>5 GeV.**
This Term is based on a Level 2 Filter Object Name/cvsPackage/Version=
MUON / l2gblmuon / 1

Order	Parameter	Type	Value	PVTVersion	Description	Default
1	MINOR_VERSION	int	**0**	-	-	-
2	MINET	float	**5.**	-	5 GeV	0.
3	QUALITY	int	**2**	-	medium quality	1
4	PROMPT	int	0	-	-	0
5	SIGN	int	0	-	-	0
6	TOOL	tool	muon5	1	-	-

Figure A.10: The main Level 2 filter associated with MUH5_LM15. There is also a PASS100 filter used as placeholder for a possible track requirement as seen in the muon tool.

filters have been run, and a list of filters that have been run for the scripts. Once these have been run once, they do not need to be run again for that event. All objects matching specific criteria will be saved in a final list. All objects that pass individual filter conditions will also be saved in a list. After all of these have been run, based on script requirements, the later scripts simply choose among these objects to see if they pass the necessary conditions.

A.5 L2 Global Packages

The L2 Global code is stored in several packages in CVS. Several packages are responsible for controlling input, output and decision making. The rest are used for filters and tools.

The code involved in packing and unpacking data, running tools and filters, reading L1 trigger masks, making decisions, and filling output is listed below. This also is where interpretation of possible errors coming from the L2 framework occurs.

- l2gblbase
- l2gblworker
- l2io

The rest of the code is made up of the individual tools and filters. These packages are listed below.

- l2gblem
- l2gblmuon
- l2gbltau
- l2gbljet

- l2gbltrack
- l2gblgeneric

All of these packages can use all of the input data. The separation by packages is for convenience rather than signaling partitions within the code.

A.6 Monitoring and Common Problems

The Level 2 Global crate typically runs well without interruption, but several monitoring tools are available to follow data taking and ensure the crate is running as expected.

A monitoring script called l2mon keeps track of global quantities from the trigger framework, as can be seen in Figure A.11. The trigger rates of each individual L2 trigger are also monitored as seen in Figure A.12. If the overall trigger rate jumps unexpectedly, the individual trigger rates can be checked to isolate the problem. These rates are plotted with a script called trigstripmon shown in Figure A.13. The L2 buffers are also monitored for each of the processors. The number of events that are sitting in buffers awaiting a Level 2 decision can also help with debugging a problem. If L2 Global stops issuing decisions, these buffers will quickly overflow. The part of the l2df program that shows this information is shown in Figure A.14.

As the Level 2 executable runs, configuration information and unusual running conditions are stored in a local log file. Significant errors are marked by a searchable term, "ELerror". The errors in L2 Global vary considerably from the preprocessors because of the additional coalescing of information and decision making that is only done in global.

Over 2008 and 2009, the L2 Global trigger ran into very few serious errors, but the log files were still large relative to the preprocessor crates, due to some errors occurring quite frequently. Four specific errors were found from 2008-2009 with three

TRGFR | L2 Dataflow | L2 Trigger | L2 HS | L2 Filter

L1 Trigger Monitor — Mon Jun 8 07:31:06 2009 D

Global | Spec Trigger | Geo Sector

L1 Accept (Hz)	598.406
L2 Accept (Hz)	375.048
L2 rejection (L1/L2)	1.596
L1 awaiting L2 (raw count)	0
L2 Fw Cycles (Hz)	598.172
L2 Fw Accept (Hz)	375.374
L2 Fw Reject (Hz)	222.797
L2 decisions enabled?	yes
Beam Xs in L2 Cycles (Hz)	2018484.244
Beam Xs with L2 Busies (Hz)	7586285.0
BXs waiting to advertize L2 (Hz)	7828.829
Luminosity Block Number	6774568
TRGFR paused by COOR?	no
Tot Alloc Sp Trgs	106
Tot Alloc Exp Grps	3
Tot Alloc Geo Sects	80

Figure A.11: Monitoring the global information from the trigger framework. This includes overall L2 accept rate and L2 rejection fraction.

Specific Trigger Display

Trig#	Trig Name	L1 accept (Hz)	av L1 acc. (Hz)	L2 accept (Hz)	av L2 accept (Hz)	L1/L2 ratio	L2 bit name
0	Afastz_ncu	2.768	2.465	2.768	2.458	1.0	1^min_bias_nim_NCU
1	ALiveBX_ncu	2.126	2.105	2.126	2.102	1.0	2^zero_bias_NCU
2	lMu2B_Download_NCU	0.0	0.0	0.0	0.0	reject all	L1MU2B_DWNL_NCU
3	L1CTT_download	0.0	0.0	0.0	0.0	reject all	4^L1CTT_DWNL_NCU
4	K_DOWNLOAD_NCU	0.0	0.0	0.0	0.0	reject all	5^L1CTK_DWNL_NCU
5	al2b_towers_term2_ncu	0.0	0.0	0.0	0.0	reject all	7^HCALCALIB7
6	pt5wttx_TTK(1,5.)_ncu	71.58	73.659	62.534	64.233	1.15	8^M5_TLM3_2JBID_V
7	al2b_towers_term0_ncu	0.0	0.0	0.0	0.0	reject all	16^PURPLE_HAZE
8	pt3wttx_TTK(1,3.)_ncu	0.0	0.0	0.0	0.0	reject all	17^M3_TMM3_VZ
9	CSWEM(1,19.,3.2)_ncu	86.954	87.028	49.483	51.138	1.7	21^E1_LH3ISH25
10	CSWEI(1,16.,3.2)_ncu	60.26	61.367	39.844	42.009	1.46	37^E2_LH3ISH25
11	SWEM(1,19.,3.2)_ncu^2	86.954	87.028	49.483	51.138	1.7	57^E1_LH2ISH24
12	CSWEI(1,16.,3.2)_ncu^2	60.26	61.367	39.844	42.009	1.46	58^E2_LH2ISH24
13	CSWEM(2,13.,3.2)_ncu	26.694	26.483	6.525	6.59	4.02	61^DE1_2L15SH15_L20
14	2)CSWEI(1,13.,3.2)_ncu	28.82	28.515	1.829	1.67	17.08	65^DE3_2L15SH15_L20
15	(E,1,10.,13.,W,X,X)_ncu	32.033	32.921	1.285	0.969	33.96	70^DE4_2L15SH15_L20
16	'K(E,1,6.7,7.,W,X,I)_ncu	31.885	32.154	6.13	5.566	5.78	76^DE5_2L15SH15_L20
17	K(E,1,8.,10.,W,X,X)_ncu	4.054	3.846	1.878	1.822	2.11	82^DE8_2L15SH15_L20
18	(E,1,10.,13.,W,X,X)_ncu	13.495	13.929	6.08	6.775	2.06	88^TE2_LH2ISH24
19	'K(E,1,8.,10.,W,C,I)_ncu	4.647	5.158	3.559	4.067	1.27	111^TE5_LH2ISH24
20	TK(E,2,3.,4.,W,C,I)_ncu	11.568	12.699	4.894	5.211	2.44	134^JP1_2_T10L10_L15
21	,I)_CSWEI(2,4,3.2)_ncu	2.818	3.059	0.445	0.475	6.44	138^JP2_2_T10L10_L15
22	2)CSWEI(1,10.,3.2)_ncu	11.716	11.031	5.685	5.226	2.11	142^EJT_ISHT15_M25
23	(E,1,10.,10.,W,X,X)_ncu	2.818	3.174	1.236	1.173	2.71	151^ET1_LH2ISH24
24	'K(E,1,6.7,7.,W,C,I)_ncu	2.521	2.697	1.335	1.27	2.12	169^ET2_LH2ISH24
25	K(E,1,6.7,7.,W,C,X)_ncu	1.483	1.377	0.692	0.546	2.52	187^ET3_LH2ISH24
26	(J,1,6.7,15.,W,X,X)_ncu	14.385	14.483	4.746	5.009	2.89	`ETAU2_SHT102TA10
27	TK(J,1,5.,10.,W,X,I)_ncu	9.64	10.637	4.152	4.675	2.28	`ETAU3_SHT102TA10

Figure A.12: More monitoring from l2mon program. This looks at each individual trigger and monitors the input, output and rejection information. This can be useful to isolate problems with trigger rates.

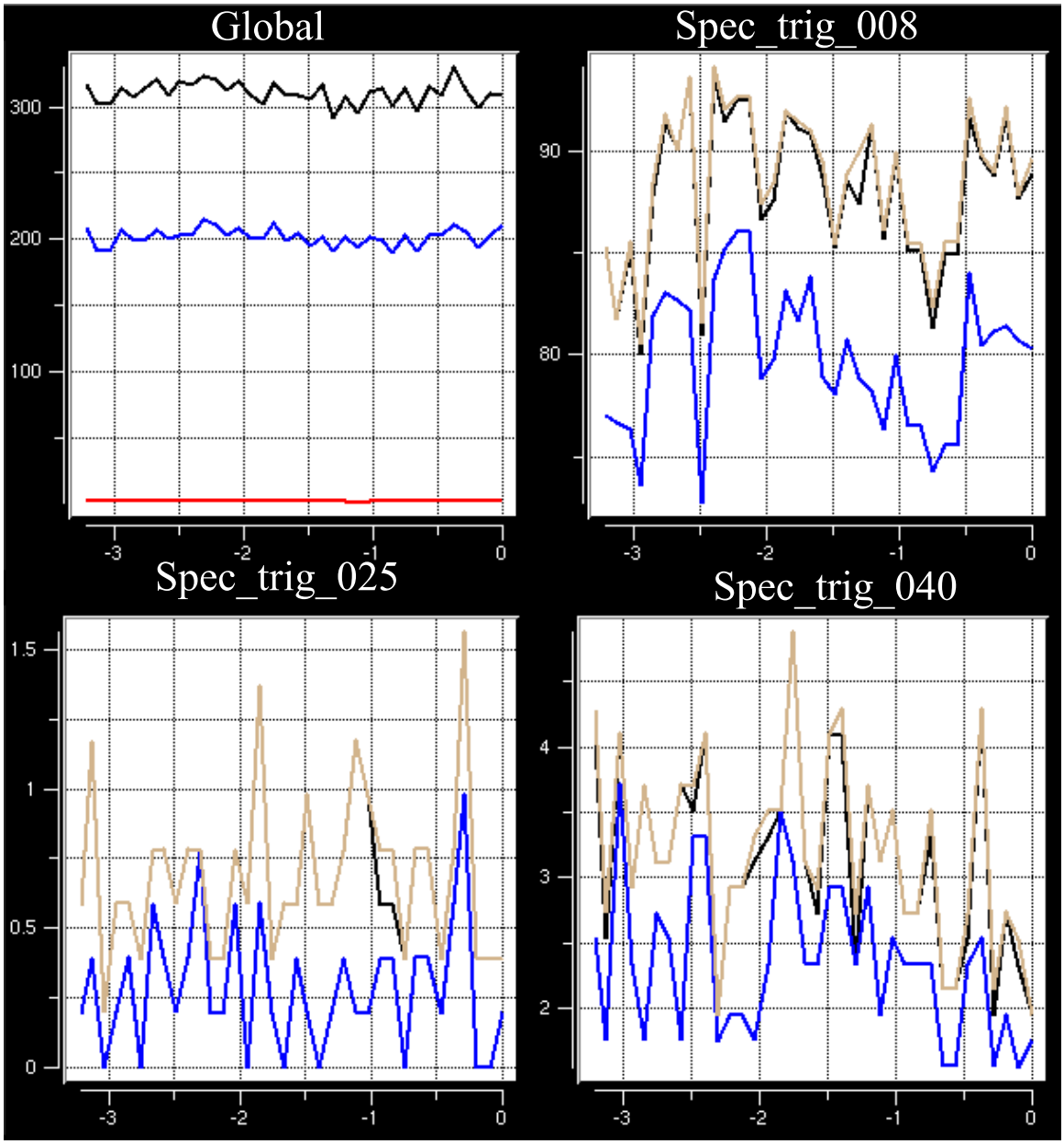


Figure A.13: The L2 monitoring program trigstripmon. Each trigger can be individually monitored with this program.

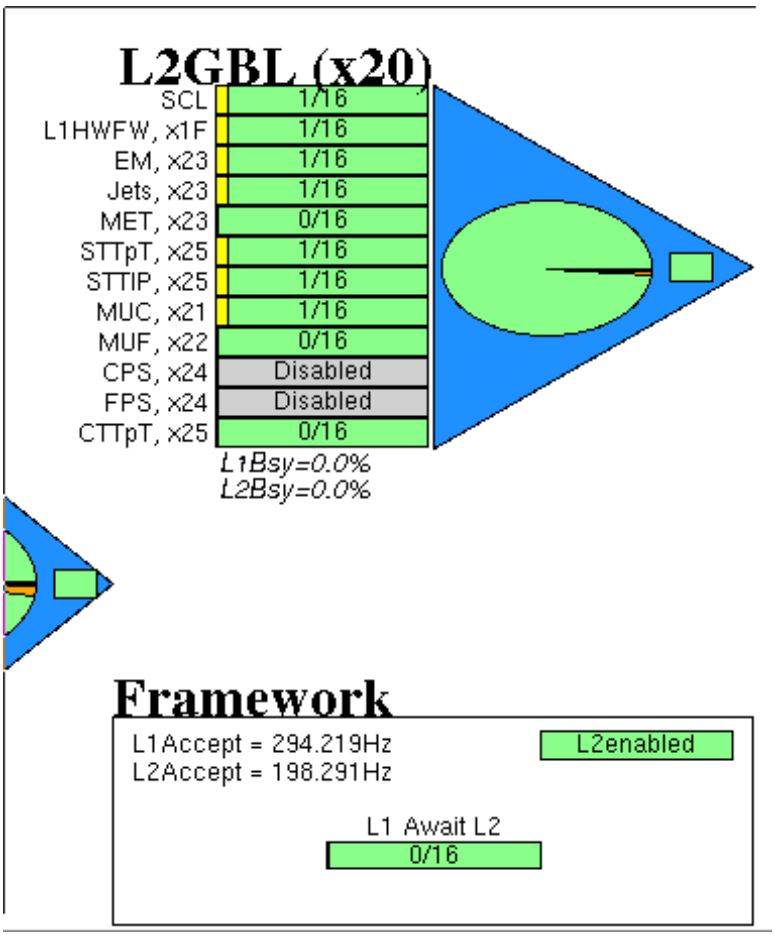


Figure A.14: The most common tool used in L2 monitoring. Shown are the parts of the GUI relevant to L2 Global.

of them occurring many times each day but without serious consequence, and one that is rarer but of significant interest. These are as follows:

- Script overflow error
- Undefined script error
- MBT channel overflow error
- L2 decision error

.

An example of each of these can be seen in the following messages taken from the Level 2 Global log files:

Mon Apr 27 00:12:36 CDT 2009: ——-ERROR LOG——-

Mon Apr 27 00:12:36 CDT 2009: message severity(Range is 0-14): 7(ELerror)

Mon Apr 27 00:12:36 CDT 2009: bunch: 128

Mon Apr 27 00:12:36 CDT 2009: rotation: 38008

Mon Apr 27 00:12:36 CDT 2009: message name: Script Overflow

Mon Apr 27 00:12:36 CDT 2009: message text:

Mon Apr 27 00:12:36 CDT 2009: Script.cpp(194): Object limit of 50 reached for L2 script 42, while in filter TRACKFILTER12

Sun Apr 26 12:21:59 CDT 2009: ——-ERROR LOG——-

Sun Apr 26 12:21:59 CDT 2009: message severity(Range is 0-14): 7(ELerror)

Sun Apr 26 12:21:59 CDT 2009: message name: Configuration

Sun Apr 26 12:21:59 CDT 2009: message text:

Sun Apr 26 12:21:59 CDT 2009: SuperScript.cpp(51): No subscript is defined. Breaking out of the loop (pass all).

Mon Apr 27 00:12:38 CDT 2009: ——-ERROR LOG——-

Mon Apr 27 00:12:38 CDT 2009: message severity(Range is 0-14): 7(ELerror)

Mon Apr 27 00:12:38 CDT 2009: bunch: 37

Mon Apr 27 00:12:38 CDT 2009: rotation: 56201

Mon Apr 27 00:12:38 CDT 2009: message name: Too many objects

Mon Apr 27 00:12:38 CDT 2009: message text:

Mon Apr 27 00:12:38 CDT 2009: FillableMBTChannel.hpp(186): Attempt to put too many objects into MBT channel with source ID 230 limit is 100 objects.

Sat Mar 21 14:59:36 CDT 2009: ——-ERROR LOG——-

Sat Mar 21 14:59:36 CDT 2009: message severity(Range is 0-14): 7(ELerror)

Sat Mar 21 14:59:36 CDT 2009: message name: L2 Decision Error

Sat Mar 21 14:59:36 CDT 2009: message text:

Sat Mar 21 14:59:36 CDT 2009: L2Decision.cpp(137): Buffer marked as pass but reject received

Sat Mar 21 14:59:36 CDT 2009: requesting SCL_INIT

The most common error found in the log files is the script overflow error. The maximum number of objects created with the global tools are set to fifty in the l2gblworker package. When an individual tool has more than fifty objects that pass the conditions set in the filters defined for that trigger list, the remaining objects will not be tagged to send to Level 3 for eventual offline analysis. After the maximum count is met, any other filter that uses these objects will be completely skipped in the object tagging portion of the code. This error has no effect on triggering. If a script passes because of an individual trigger, the result of that decision is appropriately saved. The effect of this error is simply that certain objects associated with one of the Level 2 triggers may not be available for eventual offline analysis. The event will be saved, but not all of the Level 2 objects that were involved in the event passing Level 2 will be available for further study. The frequency of this error increases with increasing luminosity as more and more objects satisfy trigger requirements. Two plots are shown in Figure A.15. The top plot shows the number of times the error occurred vs luminosity from September 2008 to June 2009. The bottom plot shows the upper plot normalized to one against the luminosity profile from the same time period.

During configuration, the undefined script error is frequently seen. As mentioned in Section A.3, each Level 1 trigger that passes is associated with a particular su-

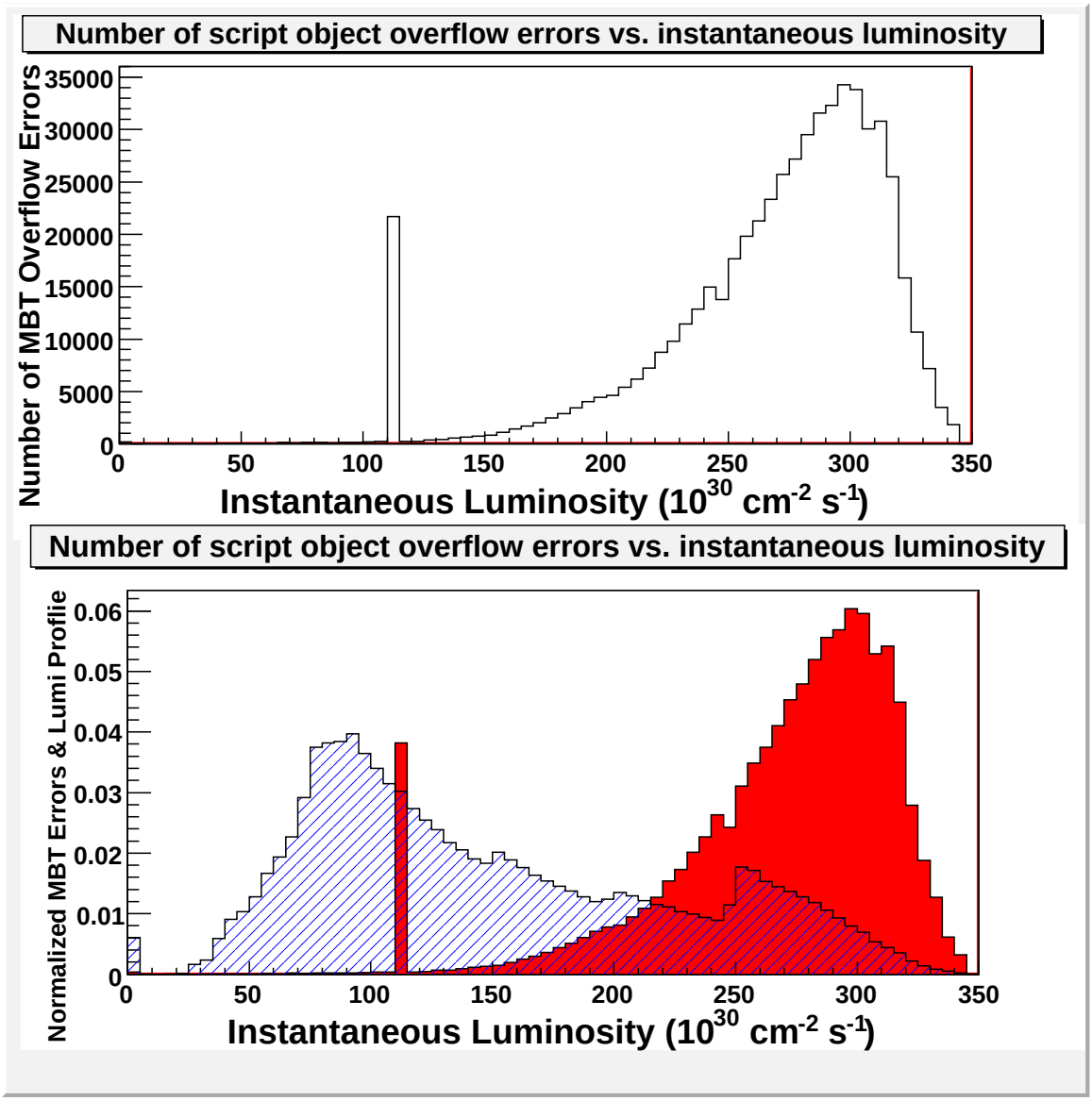


Figure A.15: The object overlow error shown as a function of luminosity in the top figure. The bottom shows the normalized number of occurances with respect to the overall luminosity profile. All data is from September 2008 - June 2009. The spike in the errors near $\mathcal{L} = 100 \cdot 10^{30}$ cm^{-2} s^{-1}, is due to an error in the luminosity fetching program.

perscript which is in turn associated with a number of scripts (or subscripts as they are called in the error message). If there are no subscripts defined for a particular superscript, then this error message is sent to the log file, and all events would automatically pass that particular trigger. Since any trigger would be sufficient to pass an event, if there were a condition in which a script was undefined, we would see no rejection at Level 2. Since this error occurs frequently, and Level 2 continues rejecting events, this error seems to be triggered outside of normal running conditions. Checking the timestamps associated with the errors confirms that this error does not happen during data taking. The C++ code associated with this error can be seen below:

```
// Initializes the SuperScript class using the parser
bool SuperScript::initialize(void) {
  doReset();                    // New run, reset everything,
including scripts
  char label[20];
  // Look for the super script's L1 bit number
  if(!isDefined("L1BIT")) {
      errlog(ERROR,"Configuration")
        << "Superscript configuration missing L1 bit number
assignment" << endmsg;
      return false;
  } else {
    _l1bit=getInteger("L1BIT");
  }
  // Loop over all the possible subscripts until an undefined one is
 found...
  for(int32 i=0;i<SuperScript::MAX_SUBSCRIPTS;++i) {
```

```
    // Construct the label that will be used to get the reference to
    // the attached subscript
    sprintf(label,"SLOT%d",i+1);
    // Check that the subscript is defined. If it is not then break
    // out of the loop immediately
    if(!isDefined(label)) {
      errlog(ERROR,"Configuration")
        << "No subscript is defined. Breaking out of the loop (pass
all)." << endmsg;
      break;
      // now using default: pass all
    }
```

One additional overflow error is also found frequently in the global log file. This one pertains to the number of objects that are to be sent to Level 3. This is also limited by object type. Even if the individual object count is not exceeded for a particular tool, each type of global object also has a maximum number that can be sent to Level 3. In the specific case that is seen in the log file, no particular track tool has more than 50 objects, but the sum of objects coming from all of the track tools exceeds one hundred. This is the cap for any tagged object to be sent to Level 3. The effects of this are similar to the script overflow error. There is no effect on triggering but a possible problem with offline trigger modeling. The translation of the MBT channels to L2 Global objects is shown in Table A.5. Two plots are shown in figure A.16 just as in the object overflow error. The top plot shows the number of times the error occurred vs luminosity from September 2008 to June 2009. The bottom plot shows the upper plot normalized to one against the luminosity profile from the same period also normalized.

The final error that has been observed recently is a potentially serious problem.

Table A.5: The translation of MBT channels to global objects sent to Level 3. Here the channel of interest is 256 which is translated as global track. The MBT channel overflow message is triggered when more than 100 of these objects are found.

Data ID	**Name**
217	Spher
218	Btag
219	TransMass
220	InvMass
221	Ht
222	MJT
223	MET
224	EMobj
225	Electron
226	Jet
227	Photon
228	Muon
229	Tau
230	Track
253	DFE_BOARD
254	CORRUPT
255	UNKNOWN

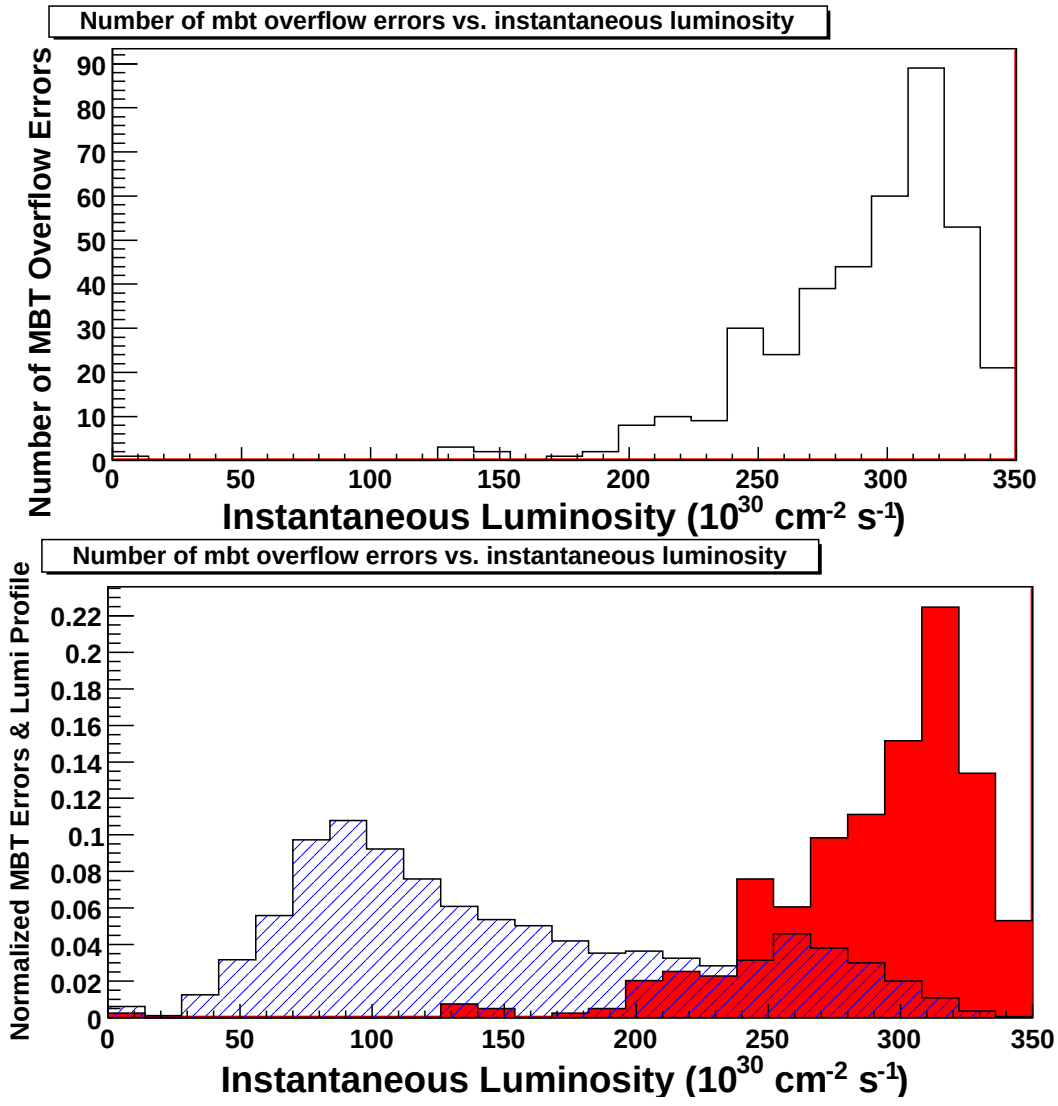


Figure A.16: The object overlow error shown as a function of lumi in the top figure. The bottom shows the normalized number of occurances with respect to the overall luminosity profile. All data is from September 2008 - June 2009.

It has not been seen since March 2009 but no explicit steps were taken to solve the problem, so it may come back. The L2 decision that is made in Global is sent to the trigger framework and returns a different result from what was sent. This decision error means that there is a problem in the path that the trigger mask travels and without the consistency check in Global, the wrong Level 2 decision would be issued. The decision errors cause the trigger system to reinitialize. These often occur in large bunches causing bad luminosity blocks which can seriously inhibit data taking. The C++ code where this check is made is shown in the code below. The route of the L2 trigger mask is shown in Figure A.17.

```
  // Check that the L2 decision matches if we did make a decision
  // based on this data, but only if the configuration flag is
set to make this check.
  if(buf->decision() && _checkDecision) {
    //IMPORTANT
    if(buf->pass() && l2scl->l2reject()) {
    //if(1) {
      // pulse bit #8 on ECL outputs if a decision errors is
 observed
      AlphaNode::tsi()->setScalerBit(8, true);
      AlphaNode::tsi()->setScalerBit(8, false);
      errlog(ERROR,"L2 Decision Error")
        << hex << "Buffer marked as pass but reject received \n"
        << "requesting SCL_INIT" << endmsg;
      errlog(INFO,"SCL Sync Info") << hex
<< "buf->bunch: " << buf->bunch()
        << " l2scl->bunch: " << l2scl->bunch()
        << " buf->rotation: " << buf->rotation()
```

```
<< " l2scl->rotation: " << l2scl->rotation()
   << endmsg;


        for (int i=0; i<NHISTORY; i++){
  errlog(INFO,"Decision History")
    << hex
    << "L1 "
    << _l1_history[i][3] << "\t" << _l1_history[i][2] <<
 "\t" << _l1_history[i][1] << "\t"<< _l1_history[i][0] << "\n"
            << "L2 "
    << _l2_history[i][3] << "\t" << _l2_history[i][2] <<
 "\t" << _l2_history[i][1] << "\t"<< _l2_history[i][0] << "\n"
    << endmsg;
}
//IMPORTANT
        throw RaiseL2Error();
```

A.7 Main Projects

After taking over the running and maintenance of Level 2 Global in late 2005, there are several projects in which I contributed personally.

A.7.1 Triggering on Events with b-Jets

The functionality to look for events with b-jets was added in late 2005. This trigger was developed by Sascha Caron to identify events that were likely to contain b-jets. A likelihood look-up table was created which took χ^2 and impact parameter significance to determine the likelihood for an individual track to have come from a b-quark. The

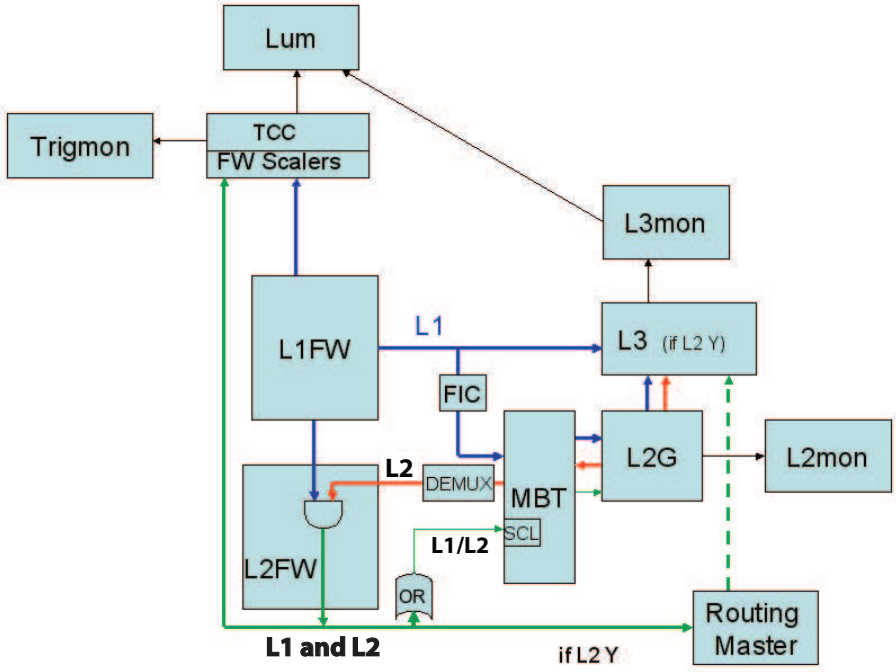


Figure A.17: The path of the L2 trigger decision through the online system. Thick lines correspond to full 128-bit triggers. The thin lines represent a single bit accept/reject decision. The blue lines are L1 trigger bits, red lines are L2 bits, and the green line is the AND of the L1 and L2 bits.

trigger then set a parameter to look at the maximum likelihood for any jet in the event or to look at the product of those values to determine if the event crossed a minimum threshold. The threshold would be set in the trigger along with the number of tracks that match minimum criteria associated with a secondary vertex.

A.7.2 Implementation of Electron Likelihood, Tau Objects, Sphericity, and Acoplanarity

The next update came in the summer of 2006. As we updated to a new trigger list associated with Run IIb, several users created tools that they integrated into the

Level 2 system. The electron likelihood was developed by Joseph Haley to better discriminate among between EM objects and jets. This was determined by looking at the electron energy deposition in the seed calorimeter cell, the highest neighboring cell, 5 X 5 blocks of cells and the corresponding information in the hadronic calorimeter.

Tau objects were introduced by Marco Verzocchi to distinguish between hadronic taus and the wider energy depositions typically associated with jets. A tau object was found by looking at the highest energy and highest neighboring hadronic calorimeter cells (similar to the electron in the EM calorimeter). These meaures of energy spread distinguish taus from light jets and comparing the EM and hadronic calorimeter depositions can distinguish between electrons and taus.

Finally, the sphericity and acoplanarity were introduced to trigger on all-jet and jet plus missing energy events that have meaningful physics. Many of the high energy jet events come from noise within the calorimeter. The sphericity and acoplanarity are measures of the event that can help determine if the energy deposition is spread through the calorimeter in a distribution that would be unusual to see in calorimeter noise.

A.7.3 L2 *βeta* Multiple Processing Boards Study

In order to reduce processing times, we conducted a study to see how practical it would be to add a further layer of parallelization in processing. This would allow the calorimeter preprocessor to have one board to focus on jets, another on EM objects, and one for the missing transverse energy. In the global crate, this would allow alternating events to be sent to different processors with the same setup. We explored the message passing system and modifications that would need to be made in the event loop to use multiple processors.

After several timing studies, it was determined that the processing time was not the limiting factor in preparing for an increase of instantaneous luminosity. The

processors that were potentially the most significant sources of deadtime were the L2CTT and L2PS. The L2PS is not used in triggering decisions, so if it becomes a significant source of deadtime, it can be turned off. The L2CTT is dominated by sending tracks to L2GBL. Parallelization of processing does not help with this process. After determining that this project would not be particularly helpful, it was decided to discontinue it.

Level 2 Global is the first place in the trigger system that each part of the detector can be considered together in making trigger decisions. This allows for more complex decisions than can be made at Level 1 by using relationships among objects throughout the detector. Additionally, Level 2 is the first part of the trigger that incorporates information from the silicon tracking detector. The Global processor drives the Level 2 system, making all of the Level 2 trigger decisions. This layer of triggering which did not exist during the first run of the Tevatron allows for high efficiency event selection while reducing the rate of incoming events to a level that can easily be consumed by the Level 3 system.

Appendix B

The High-p_T Data Format

A condensed data format was created to decrease the amount of time needed to run the analysis when small changes were made to the analysis strategy. The condensed format will only save basic object quantities such as object id, p_T, η, and ϕ. Additionally, up to four additional parameters are saved for each object depending on the object type. The information stored for a high-p_T object is shown in Figure B.1 and a summary of the various parameters is shown in Table B.1.

Events are selected to be part of the high-p_T skim if they satisfy one of the following selection criteria. These criteria for the skim are different and looser than those required for the MIS analysis fits and later VISTA and SLEUTH final states.

- μ $p_T > 20$ GeV
- e $E_T > 25$ GeV
- τ $E_T > 30$ GeV
- γ $E_T > 75$ GeV
- $\not\!\!E_T > 80$ GeV
- jet $E_T > 150$ GeV

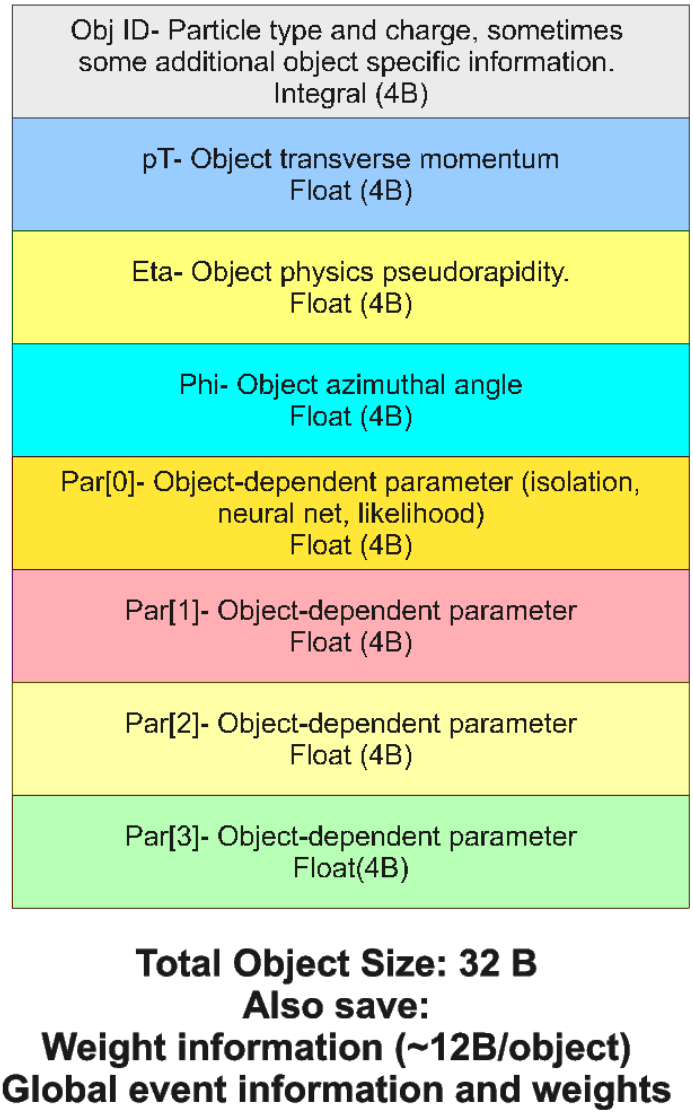


Figure B.1: The information stored in a high-p_T object is shown. Basic information is the same for all object types but four parameters are dependent upon the object type.

Table B.1: Additional parameters stored for each object in the high-p_T format.

Object	par[0]	par[1]	par[2]	par[3]
μ	Calorimeter Halo	Track Halo	Curvature Error	Track Hits
e	EM Fraction	Likelihood	8-variable H-matrix	Isolation
γ	NN4	NN5	X	X
τ	Output NN	EM NN or Tracks p_T	X	X
jet	b-tag NN	Negative tag NN	taggability SF	TRF / taggable RF

- Any two of μ, e, τ, γ with p_T or $E_T > 12$ GeV
- γ and $\not\!\!E_T$ both with $E_T > 30$ GeV

Since the reason behind the format is to use as little space as possible, additional threshold cuts are required for an object to be included in the information stored in the event. The focus of the analysis is on high-p_T objects, so basic selection criteria are also imposed on objects to be stored as part of the event. The following are the list of criteria for each individual object to be stored as part of the event.

- μ $p_T > 4$ GeV
- e $E_T > 10$ GeV
- τ $E_T > 10$ GeV
- γ $E_T > 15$ GeV

- *jet* $E_T > 15$ GeV

After the high-p_T skim the datasets used in the analysis are significantly reduced. This allows all of the data to be stored on local disks. Differences in the event sizes can be seen in Table B.2.

Table B.2: Storage comparison for some of the datasets used in this analysis comparing the standard DØ CAF tree format and the reduced high-p_T format after the high-p_T skim.

Sample	CAF Tree Size	High-p_T Size	CAF Events	High-p_T Events
EMinclusive	10.45 TB	5.5 GB	274M	19.9M
MUinclusive	8.83 TB	1.6 GB	267M	5.2M
W + 0lp	499 GB	2.2 GB	12.8M	4.2M
Drell-Yan $\mu\mu$ + 0lp, M=75-130	116 GB	1.1 GB	3.0M	1.7M
Diboson	230 GB	1.6 GB	3.7M	1.2M

Appendix C

Calculation of $\tilde{\mathcal{P}}$

The probability that a discrepancy seen in a given SLEUTH final state is due to a statistical fluctuation in the standard model background has been defined as $\mathcal{P}$. Once the minimum value of this probability $\mathcal{P}_{min}$ over all final states is found, an additional trials factor must be determined to account for the number of states that are checked. The value $\tilde{\mathcal{P}}$ represents the probability of seeing a final state as unlikely as the value of $\mathcal{P}_{min}$ based purely on the standard model background. This is determined by the formula

$$\tilde{\mathcal{P}} = 1 - \Pi_a(1 - \hat{p}_a), \tag{C.1}$$

where a represents all SLEUTH final states. The variable $\hat{p}_a$ is defined as the minimum of $\mathcal{P}_{min}$ and the probability of the total number of predicted events in a final state a to fluctuate up to three data events. Three events is found to be the minimum necessary to reasonably determine a value of $\tilde{\mathcal{P}}$ on the order of 0.001. A discussion of the determination of the minimum number of events can be found in [91].

Bibliography

[1] Gordon L. Kane. *Modern Elementary Particle Physics*. Westview Press, 1993.

[2] C. Amsler *et al.* Review of particle physics. *Phys. Lett.*, B667:1, 2008.

[3] Marion Arthaud, Frederic Deliot, Boris Tuchming, Viatcheslav Sharyy, and Didier Vilanova. Muon momentum oversmearing for p17 data. DØ Note 5444, DØ , 2007.

[4] Howard Haber. Higgs boson theory and phenomenology. The CERN-Fermilab Hadron Collider Physics Summer School, 2008. `http://projects.fnal.gov/hcpss/hcpss08/`.

[5] LEP Electroweak Working Group. The LEP electroweak working group website, 2009. `http://lepewwg.web.cern.ch/LEPEWWG/`.

[6] 2000. `http://www-visualmedia.fnal.gov/VMS_Site_2/index.shtml`.

[7] Joey Thompson. Introduction to colliding beams at fermilab. Fermilab Technical Report FERMILAB-TM-1909, Fermilab, 1994.

[8] Fermilab Accelerator Division. Accelerator concepts rookie book. Rookie book, Fermilab, 2009.

[9] Terry Asher. Main injector rookie book. Rookie book, Fermilab, 2003.

[10] Jim Morgan, Brian Drendel, and Stan Johnson. Antiproton rookie book. Rookie book, Fermilab, 2009.

[11] Sergei Nagaitsev. Introduction to electron cooling. `http://www-ecool.fnal.gov/default_files/ecool_introduction_files/v3_document.htm`.

[12] Darren Crawford. Tevatron rookie book. Rookie book, Fermilab, 2009.

[13] V.M. Abazov *et al.* (D0 Collaboration). The upgraded DØ detector. *Nucl. Inst. Methods*, **A**565:463, 2006.

[14] DØ Upgrade Collaboration. DØ silicon tracker technical design report. Technical report, D0, 1994.

[15] Dmitri Smirnov. Description of the central fiber tracker for DAQ shifters, 2007. Private website: `http://www-d0.hef.kun.nl//askArchive.php?base=agenda&categ=a09286&id=a09286s1t0/transparencies`.

[16] Lao, Miao, and Partridge. Luminosity monitor technical design report. DØ Note 3320, DØ , 1997.

[17] Alexander Khanov. HTF: histogramming method for finding tracks. the algorithm description. DØ Note 3778, DØ , 2000.

[18] A. Schwartzman and C. Tully. Primary vertex reconstruction by means of adaptive vertex fitting. DØ Note 4918, DØ , 2005.

[19] Daniel Duggan. *Measurements of the Differential Cross Sections for the Inclusive Production of a Photon and Heavy Flavor Jet*. PhD thesis, Florida State University, 2009.

[20] Adam Yurkewicz. *Search for Evidence of Supersymmetry in the Like-Sign Dimuon Channel at the DØ Experiment*. PhD thesis, Michigan State University, 2004.

[21] Philippe Gris. Electron smearing studies with Run IIa data. DØ Note 5400, DØ , 2007.

[22] Philippe Calfayan, Thomas Gadfort, Gavin Hesketh, Vincent Lesne, Mark Owen, Raimund Strohmer, Viatcheslav Sharyy, and Boris Tuchming. Muon identification certification in p17 data. DØ Note 5157, DØ , 2007.

[23] J. Hays, J. Mitrevski, C. Schwanenberger, and T. Toole. Single electron efficiencies in p17 data and Monte Carlo using d0correct from release p18.05.00. DØ Note 5105, DØ , 2006.

[24] Christophe Ochando and Jean-Francois Grivaz. SSR for p17. DØ Note 5609, DØ , 2008.

[25] Miruna Anastasoaie, Stephen Robinson, and Tim Scanlon. Performance of the NN b-tagging tool on p17 data. DØ Note 5213, DØ , 2007.

[26] Gavin Hesketh. Generator level reweighting of the inclusive w p_T distribution. DØ Note 5786, DØ , 2008.

[27] A. Evdokimov, J. Kraus, J. Linnemann, P. Mal, Md. Naimuddin, J. Piper, S. Protopopescu, and P. Renkel. Model independent search for new physics. DØ Note 5763, DØ , 2008.

[28] T. Aaltonen *et al.* Global search for new physics with 2.0 fb^{-1} at CDF. *Phys. Rev. D*, 79:011101, 2009.

[29] Mark Williams, 2009. `http://www-d0.fnal.gov/trigger_meister/private/www/tl_desc/global.html`.

[30] L. D. Landau and E. M. Lifshitz. *Mechanics*. Butterworth-Heinemann, third edition, 1976.

[31] Lewis H. Ryder. *Quantum Field Theory*. Cambridge University Press, second edition, 1996.

[32] Francis Halzen and Alan D. Martin. *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. Wiley, first edition, 1984.

[33] CERN and Fermilab. The CERN-Fermilab hadron collider physics summer school, 2008. `http://projects.fnal.gov/hcpss/hcpss08/`.

[34] Emmy Noether and M. A. Tavel. Invariant variation problems. 2005. Original paper published in 1918.

[35] Chen-Ning Yang and Robert L. Mills. Conservation of isotopic spin and isotopic gauge invariance. *Phys. Rev.*, 96:191–195, 1954.

[36] Peter W. Higgs. Broken symmetries, massless particles and gauge fields. *Phys. Lett.*, 12:132–133, 1964.

[37] Guido Altarelli. The standard model confronts the LHC. The CERN-Fermilab Hadron Collider Physics Summer School, 2008. `http://projects.fnal.gov/hcpss/hcpss08/`.

[38] R. Sekhar Chivukula. Strong dynamics. The CERN-Fermilab Hadron Collider Physics Summer School, 2008. `http://projects.fnal.gov/hcpss/hcpss08/`.

[39] Ian J. R. Aitchison. Supersymmetry and the MSSM: An elementary introduction. 2005.

[40] Ryan Hooper. *Search for Large Extra Dimensions and Z' Boson in the Dimuon Final State in $\sqrt{s} = 1.96$ TeV $p\overline{p}$ Collisions at DØ* . PhD thesis, Notre Dame University, 2004.

[41] V. M. Abazov *et al.* Search for Randall-Sundrum gravitons in dilepton and diphoton final states. *Physical Review Letters*, 95:091801, 2005. DØ Collaboration.

[42] Bruce Worthel. Linac rookie book. Rookie book, Fermilab, 2004.

[43] SLAC National Accelerator Laboratory, 2009. `http://www2.slac.stanford.edu/vvc/accelerators/klystron.html`.

[44] Bruce Worthel. Booster rookie book. Rookie book, Fermilab, 2009.

[45] K. Kleinknecht. *Detectors for Particle Physics*. Cambridge University Press, 2001.

[46] W. R. Leo. *Techniques for Nuclear and Particle Physics Experiments*. Springer, 1987.

[47] S. Abachi *et al.* The DØ Detector. *Nucl. Instrum. Meth.*, A338:185–253, 1994.

[48] V. M. Abazov *et al.* The muon system of the Run II DØ detector. *Nucl. Instrum. Meth.*, A552:372–398, 2005.

[49] Layer 0 conceptual design report. DØ Note 4415, DØ , 2003.

[50] Sergey Burdin. DØ silicon microstrip tracker. Fermilab Technical Report FERMILAB-PUB-05-515-E, Fermilab, 2005.

[51] Aurelio Juste. Description of the silicon microstrip tracking detector for DAQ shifters. `http://d0server1.fnal.gov/projects/Silicon/www/SMT_files/docs/smt_description.htm`.

[52] Herb Greenlee. Motion of a charged particle in a magnetic field. DØ Note 4180, DØ , 2003.

[53] Herb Greenlee. The DØ interaction propagator. DØ Note 4293, DØ , 2003.

[54] Guennadi Borissov. Ordering a chaos or.... technical details of AA tracking. All *DØ* Meeting, 2003. `http://www-d0.fnal.gov/global_tracking/talks/20030228/talk-adm-030228.ps`.

[55] Herb Greenlee. The DØ kalman track fit. DØ Note 4303, DØ , 2004.

[56] Ariel Schwartzman and Meenakshi Narain. Probabilistic primary vertex selection. DØ Note 4042, DØ , 2002.

[57] Yvonne Peters, Ariel Schwartzman, and Mike Strauss. Certification of the adaptive primary vertex in p17. DØ Note 5192, DØ , 2006.

[58] Ariel Schwartzman and Meenakshi Narain. Secondary vertex reconstruction using the kalman filter. DØ Note 3908, DØ , 2001.

[59] F. Fleuret. The DØ electron/photon analysis package emanalyze. DØ Note 3888, DØ , 2001.

[60] Ashish Kumar, Brajesh Choudhary, Joseph Kozminski, Robert Kehoe, Jonathon Hays, and Jan Stark. Electron likelihood study. DØ Note 4769, DØ , 2005.

[61] Joseph Haley. *First Evidence for WW and WZ Diboson Production with Semi-Leptonic Decays at a Hadron Collider*. PhD thesis, Princeton University, 2009.

[62] Oleksiy Atramentov, Andrew Askew, Dmitry Bandurin, Dan Duggan, Alexei Ferapontov, Yuri Gerstein, Yurii Maravin, and Greg Pawloski. Photon identification in p17 data. DØ Note 4976, DØ , 2006.

[63] D. Chakraborty, Y. Coadou, S. Duensing, C. Galea, Y. Gerstein, A-C. Le Bihan, C. Noding, and S. Protopopescu. Reconstruction of τ leptons in hadronic final states at DØ run 2. DØ Note 4210, DØ , 2003.

[64] H. Prosper. Some mathematical comments on feed-forward neural networks. DØ Note 1606, DØ , 1993.

[65] Emmanuel Busato and Bernard Andrieu. Jet algorithms in the DØ Run II software: Description and user's guide. DØ Note 4457, DØ , 2004.

[66] G. Bernardi, B. Olivier, B. Knuteson, and M. Strovink. NADA: A new event by event hot cell killer. DØ Note 3687, DØ , 1999.

[67] Jean-Roch Vlimant, Ursula Bassler, Gregorio Bernardi, and Sophie Trincaz-Duvoid. Technical description of the T42 algorithm for the calorimeter noise suppression. DØ Note 4146, DØ , 2003.

[68] JES Group. Jet energy scale calculation at DØ run II. DØ Note 5382, DØ , 2007.

[69] Tim Scanlon. *b-tagging and the search for neutral supersymmetric Higgs bosons at DØ* . PhD thesis, Imperial College, 2006.

[70] DØ Common Analysis Format Group. `http://www-d0.fnal.gov/Run2Physics/cs/index.html`.

[71] Z. Was. TAUOLA the library for [tau] lepton decay. *Nuclear Physics B - Proceedings Supplements*, 98(1-3):96 – 102, 2001.

[72] D. J. Lange. The EvtGen particle decay simulation package. *Nucl. Instrum. Meth.*, A462:152–155, 2001.

[73] T. Sjostrand *et al.* PYTHIA 6.2. *Computer Phys. Commun.*, **1**35:238, 2001.

[74] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. Nadolsky, and W. K. Tung. New generation of parton distributions with uncertainties from global QCD analysis. *JHEP0207*, 012, 2002.

[75] M. Mangano *et al.* Alpgen. *J. High Energy Phys.*, **0**7:001, 2003.

[76] S. Hoech *et al*, 2008. arxiv:hep-ph/0602031.

[77] Dag Gillberg. Heavy flavour removal and determination of weighting factors for ALPGEN W+jet Monte Carlo. DØ Note 5129, DØ , 2007.

[78] GEANT. CERN library long writeup. Technical Report W5013, CERN, 1994.

[79] Yuji Enari. Taggability studies update. b-ID Meeting, October 1, 2008, 2008. `http://www-d0.hef.kun.nl//askArchive.php?base=agenda&categ=a081708&id=a081708s1t16/transparencies`.

[80] Mansoora Shamim and Tim Bolton. Generator level reweighting of the p_T of the z boson. DØ Note 5565, DØ , 2008.

[81] V. M. Abazov *et al.* Measurement of the shape of the boson-transverse momentum distribution in $p\overline{p} \to z/\gamma^* \to e^+e^- + x$ events produced at $\sqrt{s} = 1.96$ tev. *Physical Review Letters*, 100(10):102002, 2008.

[82] Kirill Melnikov and Frank Petriello. Electroweak gauge boson production at hadron colliders through $\mathcal{O}(\alpha_s^2)$. *Physical Review D (Particles and Fields)*, 74(11):114017, 2006.

[83] Jeroen Hegeman. Luminosity determination and reweighting of Monte Carlo overlay luminosity for p17 hadronic top analyses. DØ Note 5561, DØ , 2008.

[84] Heidi Schellman. The longitudinal shape of the luminous region at DØ . DØ Note 5142, DØ , 2006.

[85] A. Evdokimov, J. Kraus, J. Linnemann, P. Mal, Md. Naimuddin, J. Piper, S. Protopopescu, and P. Renkel. Model independent search for new physics. DØ Note 5777, DØ , 2008. `http://www-d0.fnal.gov/Run2Physics/WWW/results/prelim/NP/N65/N65.pdf`.

[86] F. James. MINUIT: Function minimization and error analysis reference manual. Technical report, CERN Program Library Long Writeup, 1998.

[87] V. M. Abazov *et al.* Search for new physics in e mu x data at d0 using sleuth: A quasi-model-independent search strategy for new physics. *Phys. Rev. D*, 62:092004, 2000.

[88] V. M. Abazov *et al.* A quasi-model-independent search for new physics at large transverse momentum. *Phys. Rev. D*, 64:012004, 2001.

[89] V. M. Abazov *et al.* A quasi-model-independent search for new high p_t physics at dzero. *Phys. Rev. Let.*, 86:3712, 2001.

[90] A. Aktas *et al.* A general search for new phenomena in ep scattering at HERA. *Phys. Lett. B*, **602**(1-2):14–30, 2004.

[91] T. Aaltonen *et al.* Model-independent and quasi-model-independent search for new physics at CDF. *Phys. Rev. D*, 78:012002, 2008.

[92] Georgios Choudalakis. *Model Independent Search For New Physics At The Tevatron*. PhD thesis, Massachusetts Institute of Technology, 2008. `http://www.citebase.org/abstract?id=oai:arXiv.org:0805.3954`.

[93] Edmunds, Gross, Laurens, Linnemann, and Moore. L2 global TDR. DØ Note 3402, DØ , 1998.

[94] Hirosky, Baden, Burdin, Cros, Lavigne, Petroff, and Verdier. L2 beta design report. Internal Note v3.2.3, DØ , 2005.

[95] M. Campbell and C. Murphy, 1999. `http://home.fnal.gov/~ashmansk/doc/magicbus.ps.gz`.

[96] J. Linnemann, 2004. http://www.pa.msu.edu/hep/d0/l2/mbt.htm.

[97] Cypress Semiconductor. http://www.cypress.com/?rID=13675.

[98] http://www-esd.fnal.gov/esd/catalog/main/rittal/subrack_d0.htm.

[99] J. Backus Mayes, 2008. http://www-d0online.fnal.gov/www/groups/l3daq/doc/.

[100] S. Snyder. COOR. Internal Note v00-77-00, DØ , 2005.

[101] Hirosky, Kopal, Linnemann, Moore, and Yurkewicz. Design proposal to upgrade the level 2 trigger to more than 128 bits. D0 Note 4477, DØ , 2001.