Fermilab

FERMILAB-THESIS-2000-08

# A Quasi-Model-Independent Search for New High $p_T$ Physics at DØ

by

Bruce Owen Knuteson

B.A. (Rice University) 1997

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Physics

in the

GRADUATE DIVISION
of the
UNIVERSITY of CALIFORNIA at BERKELEY

Committee in charge:

Professor Mark Strovink, Chair
Professor Marjorie Shapiro
Professor Stephen Derenzo

Fall 2000

The dissertation of Bruce Owen Knuteson is approved:

Chair Date

Date

Date

University of California at Berkeley

Fall 2000

**A Quasi-Model-Independent Search for New High $p_T$ Physics at DØ**



by

Bruce Owen Knuteson

## Abstract

A Quasi-Model-Independent Search for New High $p_T$ Physics at DØ

by

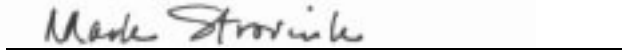
Bruce Owen Knuteson

Doctor of Philosophy in Physics

University of California at Berkeley

Professor Mark Strovink, Chair

We present a new quasi-model-independent strategy ("Sleuth") for searching for physics beyond the standard model. We define final states to be studied, and construct a rule that identifies a set of relevant variables for any particular final state. A novel algorithm searches for regions of excess in those variables and quantifies the significance of any detected excess. This strategy is applied to search for new high $p_T$ physics in $\approx 100$ pb$^{-1}$ of $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV collected by the DØ experiment during 1992–1996 at the Fermilab Tevatron. We systematically analyze many exclusive final states, and demonstrate sensitivity to a variety of models predicting new phenomena at the electroweak scale. No evidence of new high $p_T$ physics is observed.



Professor Mark Strovink
Dissertation Committee Chair

To my parents

# Contents

# List of Figures

# List of Tables

## Acknowledgements

I owe much to many. Mark Strovink has been an excellent mentor, providing consistently good advice and helpful criticisms. The patience and knowledge of Marvin Johnson and John Anderson, my instructors in the art of electronics, are greatly appreciated. I learned much from Greg Landsberg, the convener of the New Phenomena group at DØ, who together with Rich Partridge was an enormous help in putting together many of the data sets used in this thesis. It was a pleasure to work closely this past year with Dave Toback, my able and friendly fellow conspirator in the construction of Sleuth. The analysis presented here was checked carefully by an outstanding Editorial Board chaired by Hugh Montgomery. John Womersley helped to resolve a trademark dispute, and Sharon Hagopian suggested the renaming of "Sherlock" to "Sleuth." Tom Ferbel is at least partially responsible for my present writing style, and (together with Harrison Prosper) for a terrific summer school on the island of St. Croix. Ron Madaras, the LBNL DØ group leader, was responsible for many of the conveniences that allowed me to log long hours at the lab, and for enabling me to attend three summer schools and two conferences during the summer of 2000. Hannu Miettinen at Rice University first introduced me to DØ and the fascinating world of high energy physics. My parents, Lee and Chris Knuteson, shoulder most of the blame for my being who I am.

# Chapter 1

# Introduction

We begin in Chapter 2 by briefly reviewing the standard model and speculating what might lie beyond. Testing Nature at the smallest distances and largest energies currently accessible requires extremely sophisticated instrumentation; the DØ detector is described in Chapter 3. The standard model and I were born within a few years of each other, and DØ began to form as I was entering fourth grade; the credit for the content of Chapters 2 and 3 goes to an older generation of physicists.

The rest of this thesis, with the exception of the background estimates for the final states discussed in Secs. 4.4, 5.5, and 5.6, is original work. In Chapter 4 I introduce Sleuth, a new quasi-model-independent strategy for searching for physics beyond the standard model, and demonstrate its application to $e\mu X$ data. The analysis of over thirty additional final states is discussed in Chapter 5, with conclusions drawn in Chapter 6. Appendices include descriptions of other new and potentially useful analysis tools.

# Chapter 2

# Physics beyond the standard model

The fundamental interactions of nature are well described by the standard model of particle physics. The success of this model may perhaps be best appreciated by noting its outstanding agreement with nearly all experimental results published in the *Review of Particle Properties* [1], a telephone book sized tome summarizing the results of most particle physics experiments over the last half century. Nonetheless, there are reasons to believe that the standard model is incomplete, a low energy effective theory to a more fundamental theory at higher energies. The quest to reveal this underlying theory drives the current research in our field.

## 2.1 The standard model

The standard model is a field theory described by the Lagrangian

$$\begin{aligned}
\mathcal{L}_{\rm SM} &= -\tfrac{1}{4}W_{\mu\nu}W^{\mu\nu} - \tfrac{1}{4}B_{\mu\nu}B^{\mu\nu} - \tfrac{1}{4}G^a_{\mu\nu}G_a^{\mu\nu} \\
&\quad +\bar{L}\gamma^\mu(i\partial_\mu - \tfrac{1}{2}g\tau\cdot W_\mu - \tfrac{1}{2}g'YB_\mu)L \\
&\quad +\bar{R}\gamma^\mu(i\partial_\mu - \tfrac{1}{2}g'YB_\mu)R \\
&\quad -g''(\bar{q}\gamma^\mu T_a q)G^a_\mu \\
&\quad +\tfrac{1}{2}\,|(i\partial_\mu - \tfrac{1}{2}g\tau\cdot W_\mu - \tfrac{1}{2}g'YB_\mu)\phi|^2 - V(\phi) \\
&\quad -(G_1\bar{L}\phi R + G_2\bar{L}\phi_c R + h.c.).
\end{aligned} \tag{2.1}$$

The reader, though perhaps used to slightly different notation, will recognize the kinetic terms for the gauge fields in the first line of this expression, the electroweak couplings to the left- and right-handed components of the matter fields in the second and third lines, the strong coupling of quarks to gluons in the fourth, the Higgs kinetic and potential term in the fifth, and the fermion mass terms in the sixth and final line. The first four lines are well understood; the last two, which intimately involve the as yet unobserved Higgs field, beg the questions of how the electroweak and flavor symmetries are broken.

The standard model contains 21 free parameters, depending upon how one counts. Four of these parameters live in the CKM matrix; the masses of the six quarks, three charged leptons, and three neutrinos require another twelve parameters; four more appear in the values of the couplings $\alpha_{\rm EM}$, $G_F$, and $\alpha_S$, and the weak mixing angle $\sin^2\theta_W$; and the mass of the as yet undiscovered Higgs boson rounds out the set. Among the parameters that one could also include in this list are the so far undetermined elements of the leptonic analogue of the CKM matrix and a strong CP violating phase that is experimentally determined to

be so small ($\leq 10^{-9}$) that it very likely vanishes exactly by some symmetry principle not currently understood.

Saying that the standard model contains only 21 free parameters is actually a bit misleading. There are a host of "engineering numbers" that we measure that also help to define the standard model, such as particle fragmentation functions, parton distributions, and various hadron properties, that are calculable from the standard model in principle but not in practice. The gauge coupling of Quantum Chromodynamics, the theory of the strong interactions that confines quarks within hadrons, is sufficiently large at low energy scales that our standard perturbative tools are inapplicable. Although attempts are being made to perform nonperturbative QCD calculations by discretizing space-time on a lattice, this enormous computational effort has yet to provide the derivation of basic quantities (such as the proton mass) from the "21 free parameters" of the standard model. QCD is certainly a difficult theory to calculate within, but few doubt that it is correct.

Particle physics is unique among most scientific fields in being sufficiently mature that research efforts are directed toward solving a few prominent questions. Currently three outstanding problems are generally recognized in the field, and three types of experiments are hoping to shed light on these questions during this decade. The questions are:

- Can neutrino oscillations be understood by introducing a mass and mixing matrix analogous to the quark sector?
- What is the origin of CP violation?
- What is the cause of electroweak symmetry breaking (EWSB)?

Our focus will be on the last in this list.

## 2.2 Electroweak symmetry breaking

In the 1970's Glashow, Weinberg, and Salam unified the electromagnetic theory of Maxwell and the weak interaction of Fermi into the electroweak theory of the standard model, based on the gauge group $SU(2)_L \times U(1)$. This unification predicts the existence of three new gauge particles, the $W^+$, $W^-$, and $Z$ bosons, which were discovered at the S$p\bar{p}$S by the UA1 and UA2 experiments in 1983.

The $SU(2)_L \times U(1)$ symmetry that Glashow, Weinberg, and Salam postulated must be broken in order for the $W$ and $Z$ bosons to acquire masses of 82 and 91 GeV, respectively, while the photon remains massless. The mass of the $W$ and $Z$ bosons, in turn, are responsible for the apparant weakness (and short range) of the weak interaction. The requirement that the standard model Lagrangian remain invariant under $SU(2)_L$ prevents the insertion of explicit mass terms for the $W$ and $Z$ bosons, so we need some other mechanism for giving the $W$ and $Z$ particles mass.

The simplest such mechanism involves the introduction of a scalar field $\phi$, called the Higgs field. If this Higgs field obtains a vacuum expectation value (vev) $v$, the Lagrangian will contain terms of the form $v^2(W^{\pm})^2$ and $v^2(Z)^2$. This situation, in which a field obtains a vev and hides certain symmetries of the Lagrangian, is known as *spontaneous symmetry breaking*. This is to be contrasted with *explicit symmetry breaking*, which involves the introduction of symmetry-breaking terms by hand. The Higgs mechanism preserves the required $SU(2)_L \times U(1)$ symmetry, but hides it in the relationships among the coefficients of various terms.

A desirable feature of the Higgs mechanism is that it naturally accommodates the

observed tree-level relationship

$$\cos\theta_W = \frac{M_W}{M_Z}, \tag{2.2}$$

where $\theta_W$ is the weak mixing angle. The Higgs sector of the Lagrangian can be written

$$\mathcal{L}_{\text{Higgs}} = \frac{1}{2}(D_\mu\phi)^2 - \mu^2\phi^2 + \frac{\lambda^4}{4!}\phi^4; \tag{2.3}$$

this is just the fifth line of Eq. 2.1 with the kinetic terms contracted and the potential term expanded. Two free parameters appear in Eq. 2.3 — a mass term $\mu$, and a quartic coupling coefficient $\lambda$. These parameters may in turn be expressed in terms of the Higgs vev $v$ and the physical mass $m_h$. We know from precision electroweak measurements that $v = 246$ GeV, but $m_h$ is unknown.

## 2.3 New physics

Theoretical inconsistencies have historically pointed the way to new discoveries. The scattering of muon neutrinos on electrons, described in the Fermi theory of the weak interactions by the diagram shown in Fig. 2.1(a), diverges at high energies. This problem of unitarity is solved in the standard model by introducing the charged weak gauge bosons to mediate this interaction, as shown in Fig. 2.1(b). Similarly, the annihilation of $e^+e^-$ into $W^+W^-$ grows without bound at high energies if the process is mediated solely by $t$-channel $\nu_e$ exchange, as shown in Fig. 2.2(a). This high energy behavior is quenched with the introduction of the neutral weak gauge boson, which contributes the diagram in Fig. 2.2(b). The scattering of $W$ boson pairs, to which the diagrams in Fig. 2.3(a)–(c) contribute at tree level, violates unitarity at the TeV scale unless the Higgs boson (or other new physics)

exists to contribute diagrams such as those shown in Fig. 2.3(d)–(e). Since at these energies the standard model *sans* Higgs boson makes manifestly non-physical predictions, we are guaranteed to find either a standard-model-like Higgs boson or evidence of physics beyond the standard model at energies of $\approx 1$ TeV.

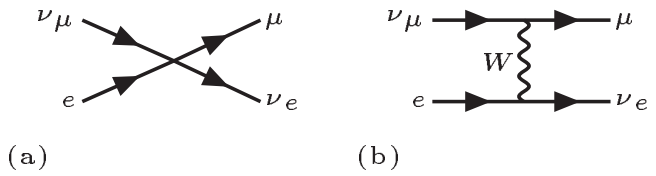



Figure 2.1: (a) The effective four-lepton vertex in the original Fermi theory of the weak interactions, which is sufficient to describe scattering at low energies. (b) The standard model diagram, involving the exchange of a $W$ boson, which accurately describes the process at higher energies.

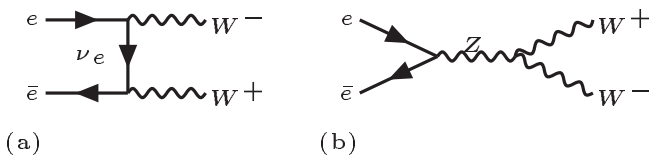



Figure 2.2: (a) $e^+e^-$ annihilation into $W^+W^-$ via $\nu_e$ exchange. The amplitude of this diagram grows as the square of the collision energy. (b) An additional diagram that contributes to this process once the $Z$ boson is postulated. Together these diagrams respect unitarity.

In fact, we are likely to find evidence of physics beyond the standard model at these energies even if a Higgs boson is observed. The Higgs in the standard model suffers from what is commonly known as the hierarchy problem, referring to the hierarchy between the Planck scale (at $10^{19}$ GeV) and the weak scale (at 1 TeV). The problem can be understood most simply by considering the 1-loop correction to the Higgs mass squared shown in Fig. 2.4. The amplitude for this diagram involves a momentum integral over a scalar loop, which diverges quadratically — or, if we impose a cutoff scale $\Lambda$, we find that the integral is proportional to

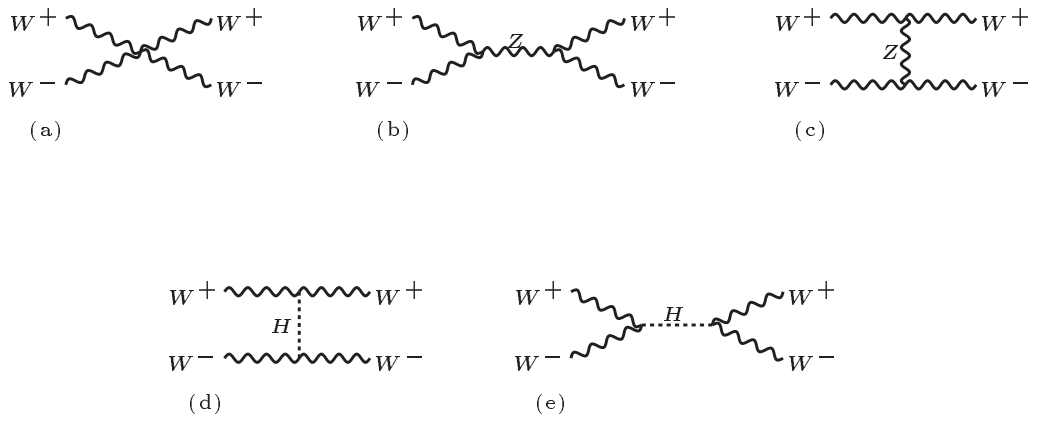



Figure 2.3: Diagrams contributing to $WW$ scattering in the standard model. If the Higgs boson in (d) and (e) does not exist, the $WW$ scattering cross section predicted by considering diagrams (a)–(c) grows without bound at large energies.

$\Lambda^2$. Incredible fine-tuning (at the level of one part in $10^{16}$) is necessary to bring the integral down to 1 TeV if it obtains corrections at the order of $M_{\text{Planck}}$. Avoiding this "unnatural" state of affairs requires the introduction of new physics at the electroweak scale.

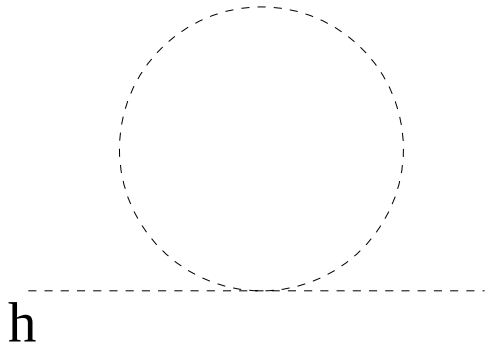



Figure 2.4: The simplest Feynman diagram giving a quadratically divergent correction to the Higgs mass squared.

There are currently three known solutions to this problem: we can invent a symmetry that allows the quadratic divergences to cancel, leaving only renormalizable, logarithmic divergences; we can assert that the problem does not exist because the Higgs is a composite

particle, so that there is no scalar in the theory to worry about; or we can argue that the upper limit of the integral should be $\approx 1$ TeV, rather than $10^{19}$ GeV. These three solutions to the hierarchy problem are obtained by positing the existence of supersymmetry [2], a new strong dynamics [3], and the existence of large extra dimensions [4], respectively.

**Supersymmetry** The "super" symmetry of supersymmetry is a broken symmetry between fermions and bosons that cancels the quadratic divergence in the quantum corrections to the Higgs mass. This symmetry requires that each standard model particle has an as yet undiscovered supersymmetric partner with identical gauge quantum numbers, but with intrinsic spin differing by 1/2. The spin-1 standard model gauge bosons ($g$, $\gamma$, $W^{\pm}$, $Z$) and the spin-0 Higgs boson thus have spin-$\frac{1}{2}$ counterparts, the *gauginos* ($\tilde{g}$, $\tilde{\chi}^{\pm}_{1,2}$, $\tilde{\chi}^{0}_{1,2,3,4}$), and the spin-$\frac{1}{2}$ standard model quarks and leptons are mirrored by spin-0 *squarks* ($\tilde{q}$) and *sleptons* ($\tilde{\ell}$). The existence of these particles is mandatory if supersymmetry is a correct theory. Any theory that does not extend this particle content is said to be a special case of the Minimal Supersymmetric Model (MSSM). In its full generality, the MSSM contains $\approx 100$ free parameters, most of which appear in the sector responsible for the breaking of the symmetry.

**Strong dynamics** Formation of a condensate of fermion pairs under the influence of a new strong force is the second candidate mechanism for electroweak symmetry breaking. The earliest models introduced new fermions (*technifermions*) that are charged under a scaled-up version of QCD (*technicolor*). Just as the breaking of the global chiral symmetry in QCD to isospin by the formation of a $q\bar{q}$ condensate leads to the existence of pseudo-

Goldstone bosons, the pions, this breaking in technicolor leads to the existence of *technipions* and a spectrum of technicolor states analogous to those observed in QCD. Three of these technicolor-charged Goldstone bosons are eaten by the Higgs mechanism, and become the longitudinal components of the $W$ and $Z$. With composite (rather than fundamental) scalars playing the role of the Higgs boson, the hierarchy problem disappears. Attempts have also been made to explain the origin of fermion masses with strong dynamics — *extended technicolor*, an embedding of technicolor, color, and flavor into a single gauge group, is one such example.

**Large extra dimensions** The third candidate cure for the hierarchy problem is to assert that the highest energy scale in the theory is $\approx 1$ TeV, rather than $M_{\text{Planck}}$. This might be the case if gravity deviates from its well-known $1/r^2$ behavior at small distances due to the presence of additional curled-up dimensions.

The "predictions" of models incorporating a supersymmetry or a new strong dynamics are many and varied, depending upon the particular values obtained for the many free parameters in the theories; theories incorporating large extra dimensions tend to be somewhat more predictive. Previous searches generally (and somewhat arbitrarily) fix many of the parameters, and results of the search are then expressed in terms of limits on the (few) parameters that are allowed to float. An obvious problem with this approach lies in the arbitrary fixing of many of the model's parameters. In addition, there are other predictions for new physics at the TeV scale that we would like to investigate at the same time: leptoquarks, contact interactions, excited leptons, a fourth generation of quarks, and addi-

tional heavy gauge bosons are just a few of the many possibilities. Theoretical reviews are available in Ref. [1]. We elaborate on the phenomenology of a few such potential signals in Sec. 5.3.

Regardless of which of these alternatives turns out to be correct (if any), we are confident that evidence of new physics must appear at the scale of 1 TeV. It was this energy scale and the typical electroweak cross sections at this scale that determined the design parameters (energy and luminosity) for the now-defunct Superconducting Super Collider (SSC) and the Large Hadron Collider (LHC) currently under construction at CERN. Whether we can find evidence of this new physics at the Tevatron remains to be seen.

# Chapter 3

# The DØ detector

## 3.1 DØ detector overview

DØ is a multipurpose detector designed to study QCD, electroweak physics, the top quark, and to search for physics beyond the standard model. In order to achieve such a comprehensive physics program, it is necessary to measure many properties of the collisions occurring within the detector.

The DØ detector [5] is a large (5,500 ton and four story tall) box, shown in Fig. 3.1; with the muon system removed, it assumes the form of a large cylinder. The beam pipe, an evacuated beryllium tube and the raceway for the colliding protons and anti-protons, runs horizontally along the axis of this cylinder. Going out in radius from this axis, one first encounters the tracking detectors that measure the tracks of charged particles produced in collisions. Outside the tracking detectors is the uranium and liquid argon calorimeter, which is used to measure the energies of electrons, photons, and hadrons. At a radius of $\approx 3.5$ m is a toroidal muon spectrometer that identifies and measures the momenta of muons, the only

particles able to penetrate the dense layers of the calorimeter.

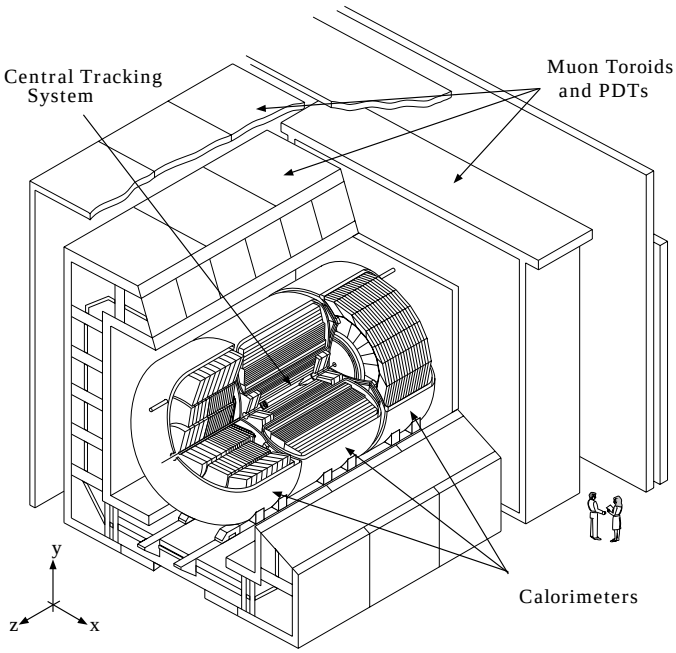



Figure 3.1: The DØ detector.

The geometry of the detector is dictated by the properties of the events that occur within it. Although ideally all collisions would occur at the same point in space, in reality the protons and anti-protons in the Tevatron collide over a "luminous region" that is generally taken to be a Gaussian centered at the center of the detector ($z = 0$) with a standard deviation of 25 cm. The center of momentum frame is the same as the laboratory reference frame in most $e^+e^-$ colliders, but in $p\bar{p}$ collisions one typically has interactions between two valence quarks that carry only a fraction of the total (anti-)proton's 0.9 TeV momentum. The event seen in the detector is therefore boosted in the laboratory frame, and the polar

angle, $\theta$, is not the most convenient angle to consider. A more appropriate measure of the polar angle is the pseudorapidity, $\eta$, defined as

$$\eta = -\ln \tan \frac{\theta}{2}, \tag{3.1}$$

which is additive under Lorentz boosts. The geometry of the tracking detectors and calorimeters is roughly cylindrical, with approximate azimuthal symmetry, but projecting back to the center of the detector. Since the longitudinal momentum of the interacting system is not known, the transverse momentum ($p_T$), defined as the magnitude of the momentum perpendicular to the beam direction, is usually considered. This unknown degree of freedom along the $z$ direction can be problematic when attempting to reconstruct the kinematics of events with missing transverse energy ($\not\!\!E_T$). Energy and momentum are conserved in these fundamental interactions, so the presence of neutrinos (or other particles without strong or electromagnetic interactions) can be inferred by the presence of substantial $\not\!\!E_T$. The DØ coordinate system is pictured in Fig. 3.2.

## 3.2 Tracking

The two beams circulating around the Tevatron spend most of their time confined in a stainless steel beam pipe of radius 3 cm. The beam pipe inside the DØ detector, however, is made of beryllium, a low-mass material, to reduce the probability that particles will scatter on the beam pipe. The detectors closest to the beam are tracking detectors, which are used to determine the trajectories of charged particles emerging from the collision. A secondary function of this system is to distinguish electrons from photons and neutral pions, and to identify photon conversions; another is to determine the location of the hard scattering.

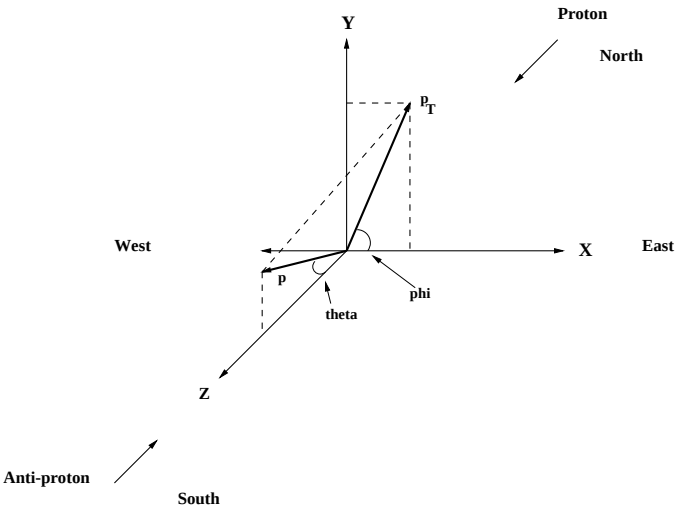



Figure 3.2: The DØ coordinate system.

As shown in Fig. 3.3, DØ's tracking detectors consist of four sub-detectors: a vertex drift chamber, a transition radiation detector, and central and forward drift chambers. DØ had no central magnetic field in Run I, and is therefore unable to measure track momenta.

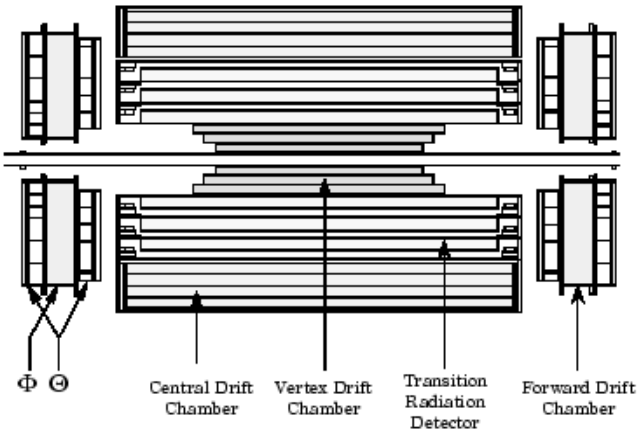



Figure 3.3: The DØ tracking system.

### 3.2.1 Vertex Drift Chamber

Immediately outside the beam pipe is the vertex drift chamber (VTX), designed for determining the location of the primary interaction in $z$ and the beam spot position in $xy$. The VTX extends from 3.7 cm to 16.2 cm in radius, and to $\pm 116$ cm in $z$. The chamber consists of three radial layers of drift cells supported by four carbon fiber tubes. The innermost layer has 16 sectors in azimuth; the outer second and third layers each have 32 sectors. Each sector contains 8 sense wires staggered by $\pm 100$ $\mu$m, helping us to avoid the left-right ambiguity that typically occurs in drift chambers. The sense wires are kept at a potential of +2.5 kV. A gas of 95% carbon dioxide ($CO_2$) and 5% ethane ($C_2H_6$) is used because of its low drift velocity (7 $\mu$m/ns) and low dispersion, which is needed in order to resolve tracks very close to the interaction region. Traces of $H_2O$ are added to the gas to reduce damage from radiation. The $xy$ position of the ionization is determined using the time it takes the electrons to drift toward the sense wires, and the position in $rz$ is determined using charge division. Unfortunately, due to the inherently coarse resolution obtained from charge division, the VTX is not particularly useful for determining the position in $rz$. The resolution of this detector is roughly 60 $\mu$m in $r\phi$ and 1.5 cm in $z$.

### 3.2.2 Transition Radiation Detector

A charged particle passing through different media must satisfy Maxwell's equations at the boundary of those media. A relativistic particle moving through a boundary emits transition radiation with energy proportional to the Lorentz factor $\gamma$ and concentrated primarily in a cone of angle $1/\gamma$ about the direction of motion. The transition radiation only

becomes appreciable (detectable) for particles that are sufficiently relativistic that $\gamma \gtrsim 10^3$, and electrons and positions are the only particles produced in Tevatron collisions that attain such high values of $\gamma$. The transition radiation detector (TRD) is useful for discriminating electrons and positrons from photons and $\pi^0$'s, which do not emit transition radiation, and which are difficult to distinguish from electrons in the calorimeter. The spectrum emitted by multi-GeV electrons peaks at roughly 8 keV, so most of the radiation is in the form of X-rays.

The TRD is built in three layers. Each layer consists of a stack of radiators — 393 layers of 18 $\mu$m thick polypropylene foils at a mean separation of 150 $\mu$m within dry nitrogen gas — followed by a proportional wire chamber. Xenon gas is used as the active medium in the wire chambers because of its large cross section for converting X-rays. The large number of foils is necessary since the probability for emitting radiation at any particular boundary much less than unity. All charged particles passing through the proportional chamber ionize some of the xenon atoms in their path, but the magnitude of the ionization signal for electrons is larger than that for other particles due to the transition radiation. (It will also be slightly larger for electrons since they are on the relativistic rise of the Bethe-Bloch energy loss curve, but this is a comparatively small effect.) The information from the TRD is combined with information on the track and the shower shape of the cluster of energy deposited in the calorimeter to construct a likelihood variable that enables us to distinguish electrons from photons and pions. This information is only available to us in the central region ($|\eta_{\rm det}| < 1.1$), due to the physical extent of the TRD.

### 3.2.3 Central Drift Chamber

The central drift chamber (CDC) lies between the TRD and the electromagnetic calorimeter. The tracks that are found in the CDC are used to obtain the $z$ position of the primary vertex. Knowing the $z$ position well is crucial for determining the transverse energies of the objects in the event, and consequently also the $\not\!\!E_T$ in the event. The amount of ionization along a CDC track also provides a hint as to whether the track was formed by an electron or a heavier particle, and this is used in electron identification.

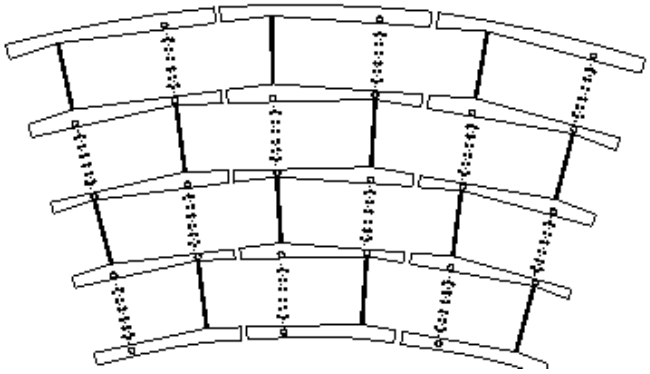


Figure 3.4: Central drift chambers.

The central drift chamber (a piece of which is shown in Fig. 3.4) is divided into four concentric cylinders, each of which lies between $\pm 90$ cm in $z$, corresponding to a reach in detector pseudorapidity of 1.2. The four cylinders lie between 49.5 cm and 74.5 cm in radius from the beam pipe. Each layer consists of 32 independent azimuthal sectors. The construction of the CDC is somewhat nonstandard in that each of these sectors is free-standing, so that each one can be replaced individually if needed. The cells of neighboring layers are offset by 1/2 cell (as shown in Fig. 3.4) in order to improve pattern recognition. Each cell

contains 7 anode sense wires staggered by $\pm 200$ $\mu$m in $\phi$ so that left-right ambiguities can be resolved. The cathode is on the boundary of the cell. The gas inside the cell consists of argon (92.5%), methane (4%), carbon dioxide (3%), and water (0.5%). There are two field-shaping wires for each sense wire in the cell, providing a uniform electric field of 620 V/cm and producing an electron drift velocity of 34 $\mu$m/ns. The measurement in $r\phi$ is obtained from the drift time of the electrons to the sense wires. The resolution from this method is between 150 and 200 $\mu$m, increasing slightly with radius. The $rz$ measurement is performed using two inductive delay lines in each cell. Each delay line consists of wire wrapped around a core of carbon fiber and epoxy, along which signals propagate at a speed of 2.4 $\mu$m/ns. An avalanche occurring near a sense wire induces a signal on the delay line; by reading out the delay line from both ends and measuring the time between signals, the $z$ position of the avalanche can be determined. The resolution in $z$ obtained in this manner is roughly 4 mm.

### 3.2.4 Forward Drift Chambers

The forward drift chambers (FDC's) extend our tracking coverage out to $|\eta_{\rm det}| \approx 3.0$. The FDC's are quite similar to the CDC's described above, using the same gas mixture and similar construction to shape the electric field inside the detector. Two identical forward drift chambers are used, one on either side of the interaction region and just beyond the ends of the cylinder formed by the VTX and CDC.

Figure 3.5 shows an exploded view of a forward drift chamber. Each FDC consists of three modules: one $\Phi$ module sandwiched between two $\Theta$ modules. The $\Phi$ module has sense wires directed radially from the beam, so as to measure $\phi$, while the $\Theta$ module has sense wires directed roughly in the $\phi$ direction, so as to measure $\theta$. The $\Phi$ module consists of 36 sectors

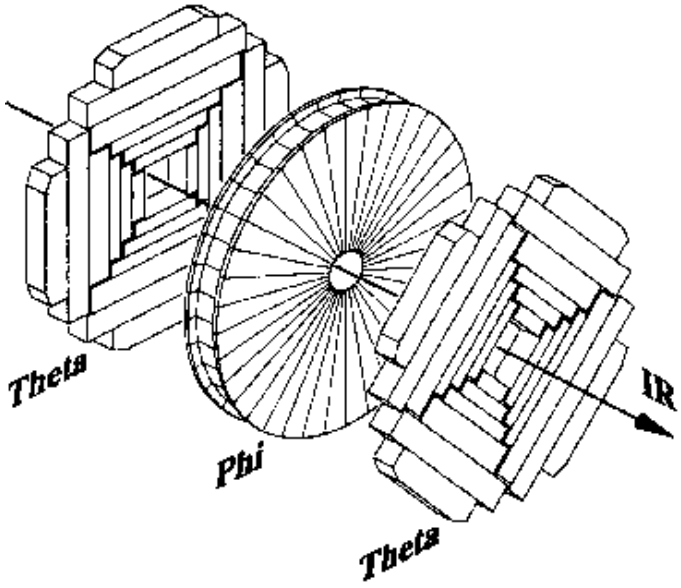



Figure 3.5: Forward drift chambers.

of 10 degrees each. Each sector contains sixteen 50 cm-long sense wires, arranged in a plane parallel to the beam axis. Each $\Theta$ module consists of four quadrants, each of which contains six rectangular cells, as shown in Fig. 3.5. Each cell contains 8 sense wires and one delay line. The delay line enables another $\phi$ measurement, which is helpful in pattern recognition. Since the construction is similar to the CDC, it is not surprising that the resolution is also similar to that of the CDC, being roughly 200 $\mu$m in the direction determined by the electron drift time, and 4 mm in the direction determined by measurements along the delay line.

## 3.3 Calorimeter

Since DØ lacked a central magnetic field in Run I, we rely heavily on the calorimeters (shown in Fig. 3.6) to measure the energies of electrons, photons, and hadrons. Assuming

that we know the interaction vertex from CDC tracks, we can also use the positions of the clusters in the calorimeter to determine the directions of the emerging energy flows (jets) in an event.

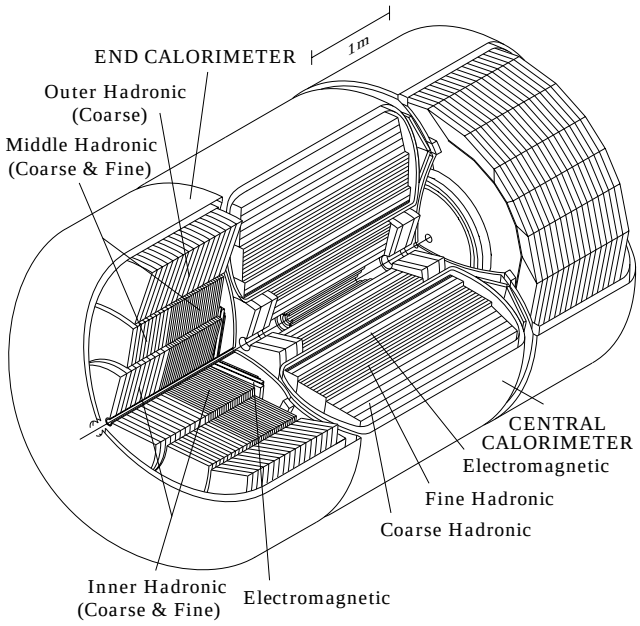



Figure 3.6: The DØ calorimeter.

Electromagnetic objects and hadronic objects lose their energies through different mechanisms, and DØ therefore has two different types of calorimeters. Both are uranium–liquid argon (U-LAr) sampling calorimeters, which use depleted uranium as the absorbing material and liquid argon as the active material. Sampling calorimeters infer total energy loss by sampling the energy loss in the active layers, which introduces a stochastic term in the resolution. Using a heterogeneous combination of active medium and heavy absorber has the advantage of allowing a more compact detector than if a homogeneous (non-sampling) calorimeter is used, however, and the cost of a sufficiently large homogeneous calorimeter is

generally not worth the gain in resolution. Due to the nuclear chemistry of uranium, the electromagnetic and hadronic parts of a shower produce similar signals. This is known as compensation; the relative response of the calorimeter of electrons to hadrons, known as $e/h$, is unity for a completely compensating calorimeter. Liquid argon is chosen as a sampling material because it is both stable and radiation hard. It has the disadvantage that it requires a cryogenic system, which places limitations on the geometry of the calorimeter.

The calorimeter comes in three sections: the central calorimeter (CC) covers the region $|\eta_{\rm det}| < 1$, and two end cap calorimeters (ECs) cover the regions $1.4 < |\eta_{\rm det}| < 4$. The CC and EC, which are similar in construction, are in turn divided radially into an electromagnetic (EM) calorimeter located just outside the central tracking detectors, a fine hadronic (FH) calorimeter located just beyond the EM calorimeter, and a coarse hadronic (CH) calorimeter, lying just outside the FH calorimeter. The calorimeters are divided into modules in $\phi$, each consisting of a series of depleted uranium absorber plates, followed by a 2.3 mm gap of liquid argon, followed by a multi-layer printed circuit board, followed by another 2.3 mm liquid argon gap, followed by another uranium absorber plate, and so on. An incident particle will interact with the absorber, producing a shower of particles that ionize the active liquid argon. The readout boards are covered with a resistive epoxy, and are held at high voltage with respect to the absorber plates, which are held at ground. The ionized electrons in the argon liquid then drift to the printed circuit board, where the charge is collected and read out. The absorber plates are not segmented in $\eta$, but the signal readout boards are segmented by cuts in the copper pads on the boards. Several readout boards at the same $\eta$ and $\phi$ are ganged together to produce a readout cell that combines signals

from several readout boards. The segmentation in $\Delta\eta \times \Delta\phi$ for most of the calorimeter is $0.1 \times 0.1$. The pads are cut such that the geometry of the calorimeter may be described as pseudo-projective. The term "pseudo-projective" indicates that while the absorber readout plates are parallel to the beam axis, the towers formed by the plates project back to the center of the detector, as shown in Fig. 3.7. The size of the towers vary with polar angle such that each cell spans $\Delta\eta = 0.1$. This is convenient for triggering and reconstruction.

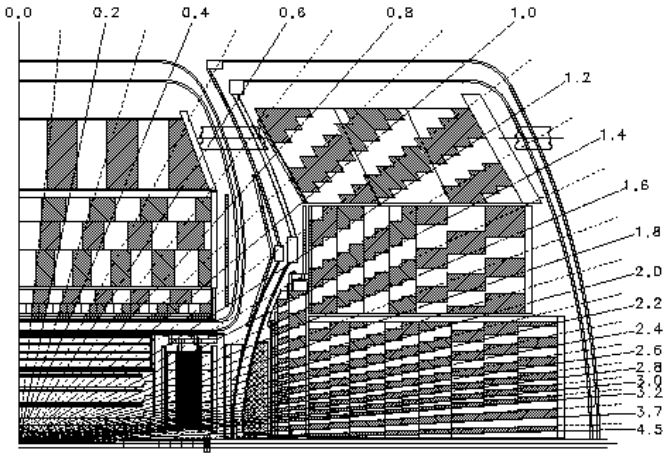



Figure 3.7: Side view of the DØ calorimeter.

The physics of electromagnetic showers is fairly well known, and is governed by basic QED processes. An energetic electron traversing a medium will lose most of its energy through bremsstrahlung. An energetic photon in turn loses most of its energy by pair production of an electron and a positron. Therefore an energetic electron traversing a material will radiate a photon, which will pair produce a positron and an electron, each of which will in turn radiate additional photons, and so on, building up an electromagnetic shower. The longitudinal extent of a shower is determined by a characteristic distance called the

radiation length ($X_0$), defined such that the average energy loss to bremsstrahlung is equal to the incident energy of the particle over the radiation length — i.e.,

$$\left\langle \frac{dE}{dx} \right\rangle_{\text{brem}} = \frac{E}{X_0}. \tag{3.2}$$

The characteristic scale for the width of the shower is the Moliére radius, which is typically about 2 cm. Expressions for the position of the shower maximum and the number of radiation lengths necessary to contain 99% of the energy deposited in the shower are well known. The Particle Data Group's *Review of Particle Properties* [1] contains an excellent review of the basic properties of electromagnetic showers.

The EM calorimeter is formed with 3 mm thick depleted uranium plates, segmented in 32 modules in phi and into four layers in radius. The four layers are known as EM1, EM2, EM3, and EM4, and contain 2, 2, 7, and 10 radiation lengths of material, respectively. Because some of the energy may be deposited in the first hadronic layer (FH1), the energy in this layer is added to the energy in the electromagnetic calorimeter when reconstructing the energies of electrons and photons. The maximum of the electromagnetic shower typically occurs in the third electromagnetic layer. In order to better see the shape of the cluster at shower max, the third electromagnetic layer is segmented more finely than the other three electromagnetic layers, with areas of $0.05 \times 0.05$ in $\eta\phi$. The use of shower shape information allows us to better reject backgrounds from $\pi^0$'s, whose two photons leave slightly broader clusters.

The physics governing hadronic showers is less well understood, due to the complication of nuclear processes. The characteristic depth of a hadronic shower in a medium is given by the nuclear interaction length ($\lambda$), which is inversely proportional to the nucle-

ar scattering cross section and to the atomic density of the absorbing material. Hadronic showers involve both electromagnetic and nuclear processes, and the signal that is obtained is different for the electromagnetic and hadronic components of the shower. This difference, known as the $e/\pi$ ratio, is close to unity in uranium.

The fine hadronic calorimeter is segmented into three layers radially. Each hadronic module contains 6 mm thick depleted uranium absorber plates (this time with an admixture of 1.7% niobium), and is segmented into 16 modules in $\phi$. The coarse hadronic calorimeter contains thicker 46.5 mm absorber plates made of copper, and is used for coarse energy measurement. The number of nuclear interaction lengths in the entire hadronic calorimeter range from $7.2\lambda$ at $\eta = 0$ to $10.3\lambda$ at the inside edge of the EC.

Because the entire calorimeter is encased in a cryostat, there are gaps in instrumentation between the CC and the EC. We instrument these regions with two different detectors in order to coarsely measure the energies of particles that pass in these directions. The first of these detectors are the "massless gaps," which contain readout boards in liquid argon, using the cryostat walls rather than uranium plates as the absorber. The second are the two inter-cryostat detectors (ICD's), one just inside the north EC and one just inside the south EC, each of which contains 384 scintillator tiles read out by photomultiplier tubes (PMT's). The tiles, like the calorimeters, are $0.1 \times 0.1$ in $\eta\phi$.

From test beam data we can determine the energy resolution of our calorimeter to electrons and to single pions. These were measured to be

$$\left.\frac{\sigma(E)}{E}\right|_e \approx \frac{15\%}{\sqrt{E}} \tag{3.3}$$

and

$$\left.\frac{\sigma(E)}{E}\right|_{\pi^{\pm}} \approx \frac{40\%}{\sqrt{E}}, \tag{3.4}$$

respectively. A more correct expression for the electron energy resolution also contains a constant term arising from nonuniformities in response at different positions in the detector. A more complete parameterization is given by

$$\delta E/E = 15\%/\sqrt{E} \oplus 0.3\%. \tag{3.5}$$

The $Z$ boson resonance in the dielectron invariant mass distribution as measured in the DØ detector is shown in Fig. 3.8 [6]; the 2.5 GeV intrinsic width of the $Z$ boson is comparable to the uncertainty in the invariant mass measurement. (The events in the shaded region of the histogram form the $Z$ final state, which we will meet again in Sec. 5.4.1.)

The energy resolution for jets is somewhat worse than the energy resolution for single pions, due to out of cone showering and noise from the underlying event. The jet energy resolution is determined by measuring the imbalance in transverse energy in events with exactly two jets, or in events with one photon and one jet. Figure 3.9 shows the distribution of $(p_T^{j1} - p_T^{j2})/(p_T^{j1} + p_T^{j2})$ for a sample of dijet events [7]. Similar studies with jets of different energy and in various regions of the detector lead to a jet energy resolution parameterized by

$$\delta E/E = 80\%/\sqrt{E}. \tag{3.6}$$

## 3.4 Muon system

The hadronic calorimeter is sufficiently thick that no strongly interacting particle is likely to penetrate its seven to ten nuclear interaction lengths. Neutrinos only interact

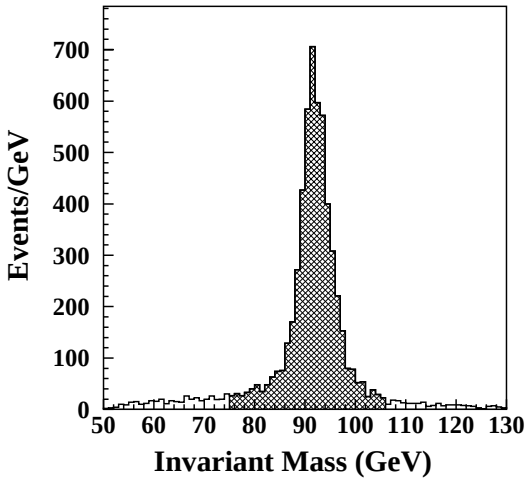



Figure 3.8: $Z$ boson resonance in the dielectron invariant mass distribution.

through the weak force, and are therefore invisible to our detector. Among the charged leptons, electrons deposit their energy by bremsstrahlung in the electromagnetic calorimeter, and taus typically decay within 1 mm of the interaction. Therefore the only (known, detectable) particle that makes it all the way through the calorimeter is the muon, as long as its energy is greater than some threshold ($\approx$ 3.5 GeV at $|\eta|= 0$, rising up to $\approx$ 5 GeV at larger $|\eta|$). Because the muon loses little energy in material, its momentum must be measured by the bending of the track in a magnetic field. This is accomplished with large toroidal magnets that surround the central detector and calorimeters.

The muon system, shown from the side in Fig. 3.10, is in two parts: WAMUS (Wide Angle MUon System), which covers the region $|\eta|< 1.7$, and SAMUS (Small Angle MUon

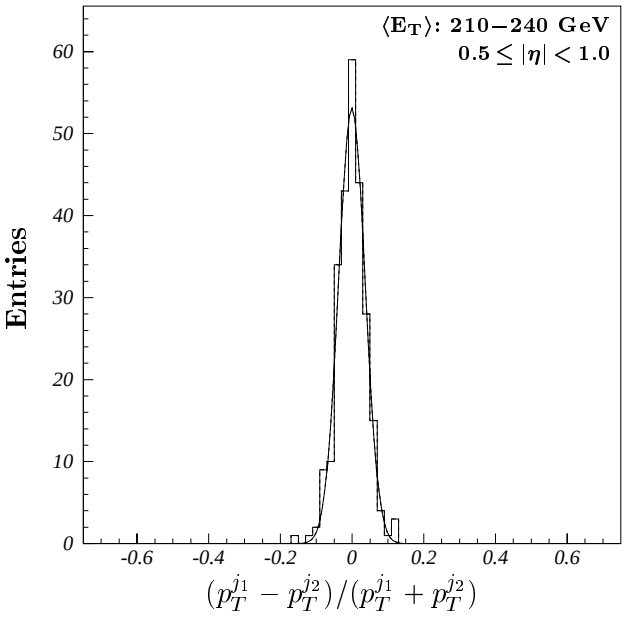



Figure 3.9: Dijet asymmetry distribution.

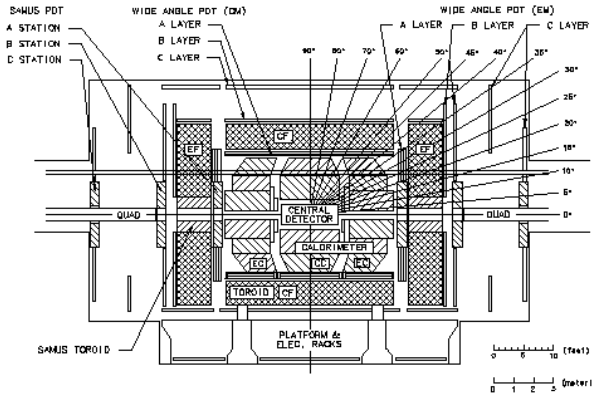



Figure 3.10: The DØ muon system.

System), which covers the region 2.5 $<|\eta|<$ 3.6. (The region 1.7 $<|\eta|<$ 2.5 is in principle covered by a combination of WAMUS and SAMUS, but the software required to reconstruct

tracks in this overlap region was never finalized in view of the high accidental background rates in this region.) WAMUS consists of three toroidal magnets: one CF (central Fe) toroid, which covers the region $-1.0 < \eta < 1.0$, and two EF (end Fe) toroids, which cover the regions $-1.7 < \eta < -1.0$ and $1.0 < \eta < 1.7$. The coverage of the CF is nearly 360 degrees, apart from two gaps underneath the detector required for support structures for the calorimeter.

The CF and EF toroids are roughly 1.1 m and 1.5 m thick, respectively. The wire coils around this iron carry a current of roughly 2500 A, producing an azimuthally-directed magnetic field that varies somewhat with position, reaching a peak magnitude of 2 T. A precise magnetic field map is used in the reconstruction of the muon momentum to account for the nonuniformity of the field.

The wide angle muon system contains three layers, labeled A, B, and C. Layer A is just inside the toroid, layer B is just outside the toroid, and layer C is roughly 1.4 m beyond layer B. Layer A consists of four sub-layers of proportional drift tubes (PDTs). The PDTs have a resolution of 1.6 mm along the wires, which are oriented in the direction of the magnetic field lines, and a resolution of 530 $\mu$m in the drift plane. The four sub-layers of PDTs in layer A measure the incident direction of the muon to 0.6 mrad, and its position to 100 $\mu$m. Layers B and C each contain 3 sub-layers of PDTs, and determine the direction and position of the outgoing muon to within 0.2 mrad and 170 $\mu$m, respectively. This accuracy results in a determination of the muon momentum $p$ that has a resolution parameterized by

$$\delta(1/p) = 0.18(p-2)/p^2 \oplus 0.003, \tag{3.7}$$

where $p$ is in units of GeV, and the $\oplus$ means addition in quadrature.

## 3.5 DAQ and Trigger

The peak luminosity at DØ in Run I reached $3 \times 10^{32}$ cm$^{-2}$s$^{-1}$. Since the total $p\bar{p}$ cross section is roughly 70 mb at $\sqrt{s} = 1.8$ TeV, the rate of interactions in the DØ detector is $\approx 1$ MHz. Collecting information about these interactions and deciding which are sufficiently interesting to record is the job of the data acquisition (DAQ) and trigger systems.

DØ employs a multilayer triggering system to select events of interest:

- The Level 0 trigger requires that two scintillator hodoscopes, mounted on the end calorimeters, detect the proton breakup that accompanies a hard interaction. The hodoscopes are also used to monitor instantaneous luminosity.

- Level 1 is implemented in hardware. Information from Level 0, summed energies in calorimeter towers, and hits in muon chambers are combined in a large and-or network to form 32 Level 1 triggers. The rate out of Level 1 is roughly 800 Hz.

- Level 1.5, also implemented in hardware, performs a crude clustering of electromagnetic calorimeter tower energies and basic track-finding with hits in the muon chambers, reducing the rate to 200 Hz.

- Further processing of the event is performed in software at Level 2, where the event rate is reduced to roughly 4 Hz.

Events passing Level 2 are then passed to the host computer, which copies the data to tape. The data are stored in a format based on Zebra [8], an extension of FORTRAN that allows dynamic memory allocation. This process is summarized in Fig. 3.11, which shows a block diagram for the DØ trigger and data acquisition system.

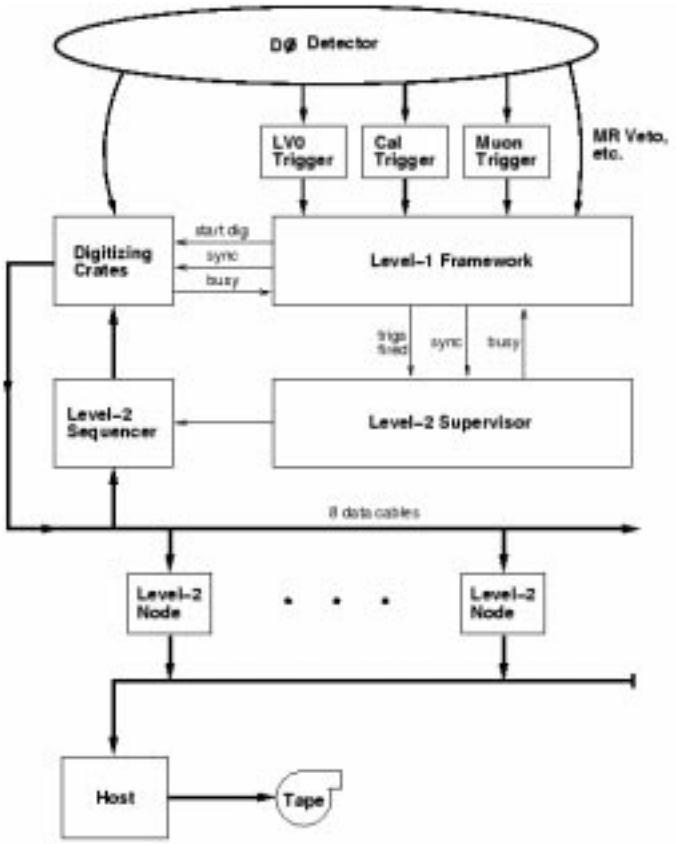



Figure 3.11: The DØ data acquisition system.

The primary triggers used to select the events analyzed in this thesis are provided in Table 3.1.

## 3.6 Event reconstruction

DØRECO, the DØ reconstruction program, takes the hits recorded in the detector — ionizations in drift chambers and drift tubes, and energy deposits in the calorimeters — and clusters them to form tracks and jets of energy flow. Tracks and jets passing carefully chosen criteria are interpreted as resulting from electrons, photons, muons, and quarks or gluons in

| Final states | Trigger name | Trigger requirements | $\int \mathcal{L}\, dt$ |
|---|---|---|---|
| $e\mu X$ | MU_ELE | $E_T^{\rm EM} > 7$, $p_T^{\mu} > 8$ | 108 pb$^{-1}$ |
| | ELE_JET_HIGH | $E_T^{\rm EM} > 15$, $\not\!\!E_T > 14$, $E_T^{j_{1,2}} > 10$ | 108 pb$^{-1}$ |
| | MU_JET_HIGH, MU_JET_CENT | $p_T^{\mu} > 10$, $E_T^{j} > 15$ | 108 pb$^{-1}$ |
| $e\not\!\!E_T\, 2j(nj)$ | ELE_HI | $E_T^{\rm EM} > 20$ | 11.2 pb$^{-1}$ |
| | EM1_EISTRKCC_MS | $E_T^{\rm EM} > 20$, $\not\!\!E_T > 15$ | 93.7 pb$^{-1}$ |
| | ELE_JET_HIGHA | $E_T^{\rm EM} > 17$, $\not\!\!E_T > 14$, $E_T^{j_{1,2}} > 10$ | 10.5 pb$^{-1}$ |
| $ee\, 2j(nj)$ | ELE_2_HI | $E_T^{\rm EM_{1,2}} > 10$ | 14.7 pb$^{-1}$ |
| | EM2_EIS_ESC(A) | $E_T^{\rm EM_1} > 20$, $E_T^{\rm EM_2} > 16$ | 108 pb$^{-1}$ |
| $\mu\not\!\!E_T\, 2j(nj)$ | MU_JET_* | $p_T^{\mu} > 10$, $E_T^{j} > 15$ | 94 pb$^{-1}$ |
| $\mu\mu\, 2j(nj)$ | MU_JET_* | $p_T^{\mu} > 10$, $E_T^{j} > 15$ | 94 pb$^{-1}$ |
| | DIMU_JET_CENT | $p_T^{\mu_{1,2}} > 3$, $E_T^{j} > 15$ | 94 pb$^{-1}$ |
| $(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$ | ELE_2_HI | $E_T^{\rm EM_{1,2}} > 10$ | 14.7 pb$^{-1}$ |
| | EM2_EIS_* | $E_T^{\rm EM_1} > 20$, $E_T^{\rm EM_2} > 16$ | 123 pb$^{-1}$ |
| | all $e\mu X$ triggers | | 108 pb$^{-1}$ |
| | all $\mu\mu(nj)$ triggers | | 94 pb$^{-1}$ |

Table 3.1: The primary triggers used to select events in each of the final states considered in this analysis. All energies are in units of GeV.

the final state, and any imbalance in detected transverse energy, suggesting the production of a particle that interacts only weakly, is noted. To first order, each 250 kilobyte event is thus summarized in the four-momenta of the few "physical" objects emerging from the hard scattering. The output of DØRECO is saved in two different formats: STA files, which store the raw data in addition to DØRECO output; and DST files, which store summaries of tracks from central detector and muon chambers, calorimeter clusters, and parameters for loosely-defined electron, photon, muon, and jet candidates, and $\not\!\!E_T$. Smaller versions of the DST files ($\mu$DSTs [9]) are then written with a subset of the information on the DSTs. Finally, $\mu$DSTs, which are still unwieldly for the purposes of most analyses, are used to construct an ntuple, which is easily manipulated with the Physics Analysis Workstation (PAW) [10].

### 3.6.1 Track-finding

Algorithms for clustering energy deposits in the calorimeter make use of transverse energies; computing transverse energies requires knowledge of the $z$ position of the primary vertex; knowledge of the primary vertex relies upon knowing the trajectories of tracks in the event. Track-finding is therefore the first task undertaken by DØRECO.

#### Central detector

Raw ADC information in the central detector is used first in $r\phi$, and then in $rz$. Considering for the moment only $r\phi$, the outermost hits in each of the four cylindrical layers that form the CDC are paired with the innermost hits in that layer within a section of $\phi$. Hits between each pair are added to form a *track segment* if they lie on the line defined by that pair. In order to allow for inefficiencies of the CDC, up to two sense wires without a hit are allowed along a track segment.

The track segments are combined into tracks traversing all four layers of the CDC by beginning with each segment in the outermost (fourth) layer, adding the best-fitting segment in the next-outermost (third) layer, then the best-fitting segment in the second layer, and finally the best-fitting segment in the innermost layer. Up to one layer may be skipped if collinear segments are found in the remaining three layers. After fitting tracks in the $r\phi$ view in this manner, delay line measurements are used to determine the $rz$ positions of the tracks. Tracks obtained in this way are determined to within roughly 2.5 mrad in $\phi$ and 28 mrad in $\theta$. Tracks in the FDC are found using a similar algorithm. Track-finding efficiencies are determined with $Z \to ee$ events to be $80 \pm 1\%$ in the CDC and $73.5 \pm 1\%$ in

the FDC.

Tracks from the CDC passing within 2.5 cm of the $z$ axis (the beam line) are used to determine the location of the hard scattering in $z$. A histogram is filled with the $z$ component of the point along each track closest to the $z$ axis. The primary vertex is estimated to lie at the global maximum of this histogram. Beam position measurements constrain the vertex to lie within 50 $\mu$m of the $z$ axis in $xy$.

**Muon chambers**

Track-finding in the muon chambers is complicated by the geometry of the detectors and by the bending of the muon in the field of the toroid magnet. Recall from Sec. 3.4 that the muon system consists of three layers of PDTs: one inside the magnet (layer A), and two outside (layers B and C). A muon passing through layer A will leave hits in up to four PDTs; a muon passing through layers B and C will leave hits in up to three PDTs in each layer. Track segments inside the magnet, requiring at least two hits, are defined in layer A; segments outside the magnet, requiring at least four hits, are defined using layers B and C together. All segments are required to point to less than 5 m from the center of the detector in order to reduce backgrounds from cosmic rays. Segments from layer A are combined with nearby segments from layers B and C to form muon candidates.

Although to first order the muon momentum is determined by the angle between the segments inside and outside of the magnet, a more accurate measurement of the momentum can be made by applying a global fit with 16 parameters: the $x$ and $y$ components of the event vertex, two parameters describing the quality of the match between the muon's track in the CDC and the trail of ionization left in the calorimeter, and twelve parameters defining

the CDC track and the track segments in layer A and in layers B and C.

Falsely identified muons arise from cosmic rays and electronic noise. Due to the thickness of the DØ calorimeter, the occurrence of hits in the muon chamber from the tail of a hadronic shower ("punch-through") is negligible. Variables used to reduce the backgrounds from cosmic rays and noise include:

- the quality of the track fit (IFW4);
- the fraction of hadronic calorimeter layers containing energy deposits along the path of the muon track;
- the fraction of energy deposited in the outermost layer of the hadronic calorimeter;
- $\int \vec{B} \cdot d\vec{l}$ along the muon trajectory, in order to reject tracks passing through the gap between the central and end toroids;
- the track impact parameter; and
- the difference between the time of muon detection and the beam crossing time.

In the first half of Run Ib, some muon chambers (particularly those in the EF and those near the Main Ring) were inefficient due to the buildup of polymers on the anode wires of the chambers. In the latter half of Run Ib, these polymers were removed with large electrical discharges. Slightly different selection criteria are therefore used in the two eras. Because accidentals contribute a large background in the forward regions, we require all muons to satisfy $|\eta| < 1.7$.

### 3.6.2 Clustering

Calorimeter ADC counts are converted to units of energy using calibration information from test beam runs with electron and pion beams of known energies, and extensive subsequent *in situ* calibration of both the electromagnetic and hadronic energy scales [11]. Uncertainties in the energy measurements of electrons, photons, and jets come not only from the intrinsic resolution of the DØ calorimeter, but also from the algorithms used to define these objects. In each case, a clustering algorithm is used to identify an object "candidate," and then properties of that object are queried in order to determine whether it passes more stringent identification cuts.

#### Electrons and photons

Electron and photon candidates are obtained using a nearest neighbor clustering algorithm. Starting with the most energetic tower in the EM calorimeter, neighboring towers with transverse energy greater than 50 MeV are added until all surrounding towers have transverse energy less than 50 MeV. This process is repeated with the most energetic tower in the electromagnetic calorimeter not already clustered, and so on until all electromagnetic towers are associated with a cluster. In order for a cluster to be identified as an electron or photon candidate, it must have at least 90% of its energy in the electromagnetic calorimeter, and at least 40% of its energy in a single tower. (Hadronic jets are typically much broader and deposit only $\approx 10\%$ of their energy in the electromagnetic calorimeter.) A shower centroid $\vec{x}_{\rm cog}$ is defined by

$$\vec{x}_{\rm cog} = \frac{\sum_i w_i \vec{x}_i}{\sum_i w_i}, \tag{3.8}$$

where the average is taken with weights of

$$w_i = \max(0, w_0 + \ln(E_i/E)), \tag{3.9}$$

and $w_0$ is a parameter chosen empirically to provide optimal resolution. Electron and photon candidates are then distinguished by whether the cluster has a CDC or an FDC track within a road of size $0.1 \times 0.1$ in $\eta\phi$ pointing from the primary vertex to the cluster.

Mistakenly identified electrons are due either to a jet containing an energetic $\pi^0$ decaying into two photons, combined with the overlap of a soft track from a charged hadron; or to a photon that converts into an $e^+e^-$ pair near the beam pipe. Five variables useful for rejecting these backgrounds have been identified:

- Jets that fragment into a leading $\pi^0$ often deposit energy around the candidate electron; a variable measuring the isolation of the electromagnetic energy is therefore useful for rejecting this background. "Isolation" is defined by

$$\text{Iso} = \frac{E(0.4) - EM(0.2)}{EM(0.2)}, \tag{3.10}$$

  where $EM(R)$ and $E(R)$ are the electromagnetic and total energies in a cone of radius $R$ around the shower centroid. In the $e\mu X$ final states, Iso $< 0.1$ is required for both electrons and photons. In the $W$+jets-like final states, Iso $< 0.15$ is required. In the $Z$+jets-like and $(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$ final states, Iso $< 0.15$ is required for electrons, and Iso $< 0.1$ is required for photons.

- The shapes of electromagnetic showers originating from two photons (from the decay of a neutral pion) or an electron-positron pair (from a photon conversion) are expected

to differ from showers originating from a single electromagnetic object. The fine segmentation of the DØ EM calorimeter is used to distinguish between these cases. The likeness of the shower profile to that of a single electromagnetic object is measured by a $\chi^2$-like quantity constructed from 41 variables: the fraction of energy deposited in the first, second, and fourth layers of the EM calorimeter; the fraction of energy in each cell in a $6 \times 6$ array around the center of the shower in the third EM layer; the logarithm of the total energy of the cluster; and the $z$ component of the position of the primary vertex. The means and covariances used in the construction of this $\chi^2$ are determined as a function of the pseudorapidity of the electron. Photons are required to have $\chi^2 < 100$ in all final states considered here.

- The quality of the match between an EM cluster produced by two photons and an accidental track should be worse than the quality of the match between an EM cluster produced by an electron and its associated track. A *track match significance* is therefore defined as

$$\sigma_{\rm trk} = \frac{\delta\phi}{\sigma_{\delta\phi}} \oplus \frac{\delta z}{\sigma_{\delta z}} \tag{3.11}$$

in the CC, and

$$\sigma_{\rm trk} = \frac{\delta\phi}{\sigma_{\delta\phi}} \oplus \frac{\delta r}{\sigma_{\delta r}} \tag{3.12}$$

in the EC, where $\delta\phi$ and $\delta z$ (or $\delta r$) are the distance from the track to the cluster, measured with errors $\sigma_{\delta\phi}$ and $\sigma_{\delta z}$ (or $\sigma_{\delta r}$).

- The energy loss in the CDC can also be used to discriminate backgrounds. After removing the Landau tail by ignoring the third of the CDC wires with largest signals,

the mean of signals on the remaining wires along the track is computed to find the average energy loss per unit distance $\langle dE/dx \rangle$.

- More transition radiation is expected from an electron-positron pair traversing the TRD than is expected from a single electron. The amount of transition radiation $\epsilon$ along a candidate electron track is used to distinguish photon conversions from true electrons.

These variables can be used independently, or combined into an "electron likelihood" $\mathcal{L}_e$, defined by

$$\mathcal{L}_e = \frac{p(\text{Iso}, \chi^2, \sigma_{\text{trk}}, \langle dE/dx \rangle, \epsilon \mid \text{background})}{p(\text{Iso}, \chi^2, \sigma_{\text{trk}}, \langle dE/dx \rangle, \epsilon \mid \text{electron})}, \tag{3.13}$$

where the densities are constructed from data samples and the correlations among variables are assumed to vanish. Electrons in the $W$+jets-like, $Z$+jets-like, and $(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$ final states are required to have $\mathcal{L}_e < 1.0$; those in the $e\mu X$ final states have $\mathcal{L}_e < 0.5$.

**Jets**

The identification of hadronic jets begins by "pre-clustering" the calorimeter, using the highest energy towers in the calorimeter as seeds. All towers within a cone of radius 0.5 in $\eta\phi$ are then added to each seed tower, and the energy-weighted center of the new jet is computed. A cone of radius 0.5 is then drawn about the new center, the jet is redefined, and a new center is computed. This process is repeated until an equilibrium state is reached. Overlapping jets are merged or split according to the following rule: if the transverse energy in the shared region is larger than half of the transverse energy of the softer jet, then the jets are merged into a single jet; otherwise, the jets are split, with each tower in the shared

region assigned to the jet with the nearest center. Only jets with $E_T > 8$ GeV are retained.

The kinematic criteria imposed upon electrons, photons, muons, jets, and missing transverse energy vary among the final states considered, and are described in the context of the analyses in Chapters 4 and 5. Additional details for the selection of events in the $e\mu X$ final states, the electron channel $W$+jets-like and $Z$+jets-like final states, and the muon channel $W$+jets-like and $Z$+jets-like final states are given in Refs. [12], [13], and [14], respectively. In all final states, events with activity in the Main Ring are rejected by demanding CAL_RECOVERY = 0 and GOOD_BEAM = 0.

### 3.6.3 Simulation

The effects of the reconstruction and selection just described on particular physics processes are simulated using a number of tools. The effects of the trigger requirements are studied using the package TRIGSIM, which consists of a simulation of the Level 1 (L1SIM) and Level 2 (L2SIM) triggers. L1SIM duplicates the Level 1 and-or network in software; L2SIM uses the same code as used in the Level 2 trigger. TRIGSIM provides a reasonable estimate of trigger efficiencies when combined with measurements using data sets collected with complementary triggers.

The reconstruction of Monte Carlo events uses DØGEANT, a package developed by DØ based upon GEANT [15], which simulates the interactions of particles as they traverse the DØ detector. DØGEANT can be run at varying levels of detail. At "plate level," all uranium plates and argon gaps in the DØ calorimeter are included in the specification of the

detector geometry, enabling a detailed simulation at the cost of intensive computation. An alternative model specifies the supports and individual modules of the detector, but treats the calorimeter as homogeneous blocks of a mixture of uranium, G10, and argon. A third approach uses a library of real showers, and chooses one of these showers at random rather than simulating each shower anew. The showers in all Monte Carlos that are indicated as "run through full DØGEANT" in Chapters 4 and 5 are taken from a shower library.

# Chapter 4

# Sleuth

In this chapter we introduce a new quasi-model-independent search strategy ("Sleuth") for the physics responsible for electroweak symmetry breaking. We define final states to be studied, and construct a rule that identifies a set of relevant variables for any particular final state. An algorithm searches for regions of excess in those variables and quantifies the significance of any detected excess. After demonstrating the sensitivity of the method, we apply it to the semi-inclusive channel $e\mu X$ collected in 108 pb$^{-1}$ of $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV at the DØ experiment during 1992–1996 at the Fermilab Tevatron. This chapter also appears as Ref. [16].

## 4.1 Introduction

It is generally recognized that the standard model, an extremely successful description of the fundamental particles and their interactions, must be incomplete. Although there is likely to be new physics beyond the current picture, the possibilities are sufficiently broad

that the first hint could appear in any of many different guises. This suggests the importance of performing searches that are as model-independent as possible.

The word "model" can connote varying degrees of generality. It can mean a particular model together with definite choices of parameters [e.g., mSUGRA [17] with specified $m_{1/2}$, $m_0$, $A_0$, $\tan\beta$, and sign($\mu$)]; it can mean a particular model with unspecified parameters (e.g., mSUGRA); it can mean a more general model (e.g., SUGRA); it can mean an even more general model (e.g., gravity-mediated supersymmetry); it can mean a class of general models (e.g., supersymmetry); or it can be a set of classes of general models (e.g., theories of electroweak symmetry breaking). As one ascends this hierarchy of generality, predictions of the "model" become less precise. While there have been many searches for phenomena predicted by models in the narrow sense, there have been relatively few searches for predictions of the more general kind.

In this chapter we describe an explicit prescription for searching for the physics responsible for stabilizing electroweak symmetry breaking, in a manner that relies only upon what we are sure we know about electroweak symmetry breaking: that its natural scale is on the order of the Higgs mass [18]. When we wish to emphasize the generality of the approach, we say that it is quasi-model-independent, where the "quasi" refers to the fact that the correct model of electroweak symmetry breaking should become manifest at the scale of several hundred GeV.

New sources of physics will in general lead to an excess over the expected background in some final state. A general signature for new physics is therefore a region of variable space in which the probability for the background to fluctuate up to or above the

number of observed events is small. Because the mass scale of electroweak symmetry breaking is larger than the mass scale of most standard model backgrounds, we expect this excess to populate regions of high transverse momentum ($p_T$). The method we will describe involves a systematic search for such excesses (although with a small modification it is equally applicable to searches for deficits). Although motivated by the problem of electroweak symmetry breaking, this method is generally sensitive to any new high $p_T$ physics.

An important benefit of a precise *a priori* algorithm of the type we construct is that it allows an *a posteriori* evaluation of the significance of a small excess, in addition to providing a recipe for searching for such an effect. The potential benefit of this feature can be seen by considering the two curious events seen by the CDF collaboration in their semi-inclusive $e\mu$ sample [19] and one event in the data sample we analyze in this chapter, which have prompted efforts to determine the probability that the standard model alone could produce such a result [20]. This is quite difficult to do *a posteriori*, as one is forced to somewhat arbitrarily decide what is meant by "such a result." The method we describe provides an unbiased and quantitative answer to such questions.

"Sleuth," a quasi-model-independent prescription for searching for high $p_T$ physics beyond the standard model, has two components:

- the definitions of physical objects and final states, and the variables relevant for each final state; and
- an algorithm that systematically hunts for an excess in the space of those variables, and quantifies the likelihood of any excess found.

We describe the prescription in Secs. 4.2 and 4.3. In Sec. 4.2 we define the physical objects

and final states, and we construct a rule for choosing variables relevant for any final state. In Sec. 4.3 we describe an algorithm that searches for a region of excess in a multidimensional space, and determines how unlikely it is that this excess arose simply from a statistical fluctuation, taking account of the fact that the search encompasses many regions of this space. This algorithm is especially useful when applied to a large number of final states. For a first application of Sleuth, we choose the semi-inclusive $e\mu$ data set ($e\mu X$) because it contains "known" signals (pair production of $W$ bosons and top quarks) that can be used to quantify the sensitivity of the algorithm to new physics, and because this final state is prominent in several models of physics beyond the standard model [21, 22]. In Sec. 4.4 we describe the data set and the expected backgrounds from the standard model and instrumental effects. In Sec. 4.5 we demonstrate the sensitivity of the method by ignoring the existence of top quark and $W$ boson pair production, and showing that the method can find these signals in the data. In Sec. 4.6 we apply the Sleuth algorithm to the $e\mu X$ data set assuming the known backgrounds, including $WW$ and $t\bar{t}$, and present the results of a search for new physics beyond the standard model.

## 4.2 Search strategy

Most recent searches for new physics have followed a well-defined set of steps: first selecting a model to be tested against the standard model, then finding a measurable prediction of this model that differs as much as possible from the prediction of the standard model, and finally comparing the predictions to data. This is clearly the procedure to follow for a small number of compelling candidate theories. Unfortunately, the resources required

to implement this procedure grow almost linearly with the number of theories. Although broadly speaking there are currently only three models with internally consistent methods of electroweak symmetry breaking — supersymmetry [2], strong dynamics [3], and theories incorporating large extra dimensions [4] — the number of specific models (and corresponding experimental signatures) is in the hundreds. Of these many specific models, at most one is a correct description of nature.

Another issue is that the results of searches for new physics can be unintentionally biased because the number of events under consideration is small, and the details of the analysis are often not specified before the data are examined. An *a priori* technique would permit a detailed study without fear of biasing the result.

We first specify the prescription in a form that should be applicable to any collider experiment sensitive to physics at the electroweak scale. We then provide aspects of the prescription that are specific to DØ. Other experiments wishing to use this prescription would specify similar details appropriate to their detectors.

### 4.2.1 General prescription

We begin by defining final states, and follow by motivating the variables we choose to consider for each of those final states. We assume that standard particle identification requirements, often detector-specific, have been agreed upon. The understanding of all backgrounds, through Monte Carlo programs and data, is crucial to this analysis, and requires great attention to detail. Standard methods for understanding backgrounds — comparing different Monte Carlos, normalizing background predictions to observation, obtaining instrumental backgrounds from related samples, demonstrating agreement in limited regions

of variable space, and calibrating against known physical quantities, among many others — are needed and used in this analysis as in any other. Uncertainties in backgrounds, which can limit the sensitivity of the search, are naturally folded into this approach.

### Final states

In this subsection we partition the data into final states. The specification is based on the notions of exclusive channels and standard particle identification.

**Exclusiveness.** Although analyses are frequently performed on inclusive samples, considering only exclusive final states has several advantages in the context of this approach:

- the presence of an extra object (electron, photon, muon, ...) in an event often qualitatively affects the probable interpretation of the event;
- the presence of an extra object often changes the variables that are chosen to characterize the final state; and
- using inclusive final states can lead to ambiguities when different channels are combined.

We choose to partition the data into exclusive categories.

**Particle identification.** We now specify the labeling of these exclusive final states. The general principle is that we label the event as completely as possible, as long as we have a high degree of confidence in the label. This leads naturally to an explicit prescription for labeling final states.

Most multipurpose experiments are able to identify electrons, muons, photons, and jets, and so we begin by considering a final state to be described by the number of isolated electrons, muons, photons, and jets observed in the event, and whether there is a significant imbalance in transverse momentum ($\not\!\!E_T$). We treat $\not\!\!E_T$ as an object in its own right, which must pass certain quality criteria. If $b$-tagging, $c$-tagging, or $\tau$-tagging is possible, then we can differentiate among jets arising from $b$ quarks, $c$ quarks, light quarks, and hadronic tau decays. If a magnetic field can be used to obtain the electric charge of a lepton, we split the charged leptons $\ell$ into $\ell^+$ and $\ell^-$ but consider final states that are related through global charge conjugation to be equivalent in $p\bar{p}$ or $e^+e^-$ (but not $pp$) collisions. Thus $e^+e^-\gamma$ is a different final state than $e^+e^+\gamma$, but $e^+e^+\gamma$ and $e^-e^-\gamma$ together make up a single final state. The definitions of these objects are logically specified for general use in all analyses, and we use these standard identification criteria to define our objects.

We can further specify a final state by identifying any $W$ or $Z$ bosons in the event. This has the effect (for example) of splitting the $eejj$, $\mu\mu jj$, and $\tau\tau jj$ final states into the $Zjj$, $eejj$, $\mu\mu jj$, and $\tau\tau jj$ channels, and splitting the $e\not\!\!E_T jj$, $\mu\not\!\!E_T jj$, and $\tau\not\!\!E_T jj$ final states into $Wjj$, $e\not\!\!E_T jj$, $\mu\not\!\!E_T jj$, and $\tau\not\!\!E_T jj$ channels.

We combine a $\ell^+\ell^-$ pair into a $Z$ if their invariant mass $M_{\ell^+\ell^-}$ falls within a $Z$ boson mass window ($82 \leq M_{\ell^+\ell^-} \leq 100$ GeV for DØ data) and the event contains neither significant $\not\!\!E_T$ nor a third charged lepton. If the event contains exactly one photon in addition to a $\ell^+\ell^-$ pair, and contains neither significant $\not\!\!E_T$ nor a third charged lepton, and if $M_{\ell^+\ell^-}$ does not fall within the $Z$ boson mass window, but $M_{\ell^+\ell^-\gamma}$ does, then the $\ell^+\ell^-\gamma$ triplet becomes a $Z$ boson. If the experiment is not capable of distinguishing between $\ell^+$ and $\ell^-$

and the event contains exactly two $\ell$'s, they are assumed to have opposite charge. A lepton and $\not\!\!E_T$ become a $W$ boson if the transverse mass $M^T_{\ell \not\!\!E_T}$ is within a $W$ boson mass window ($30 \leq M^T_{\ell \not\!\!E_T} \leq 110$ GeV for DØ data) and the event contains no second charged lepton. Because the $W$ boson mass window is so much wider than the $Z$ boson mass window, we make no attempt to identify radiative $W$ boson decays.

We do not identify top quarks, gluons, nor $W$ or $Z$ bosons from hadronic decays because we would have little confidence in such a label. Since the predicted cross sections for new physics are comparable to those for the production of detectable $ZZ$, $WZ$, and $WW$ final states, we also elect not to identify these final states.

**Choice of final states to study.** Because it is not realistic to specify backgrounds for all possible exclusive final states, choosing prospective final states is an important issue. Theories of physics beyond the standard model make such wide-ranging predictions that neglect of any particular final state purely on theoretical grounds would seem unwise. Focusing on final states in which the data themselves suggest something interesting can be done without fear of bias if all final states and variables for those final states are defined prior to examining the data. Choosing variables is the subject of the next section.

### Variables

We construct a mapping from each final state to a list of key variables for that final state using a simple, well-motivated, and short set of rules. The rules, which are summarized in Table 4.1, are obtained through the following reasoning:

- There is strong reason to believe that the physics responsible for electroweak symmetry breaking occurs at the scale of the mass of the Higgs boson, or on the order of a few hundred GeV. Any new massive particles associated with this physics can therefore be expected to decay into objects with large transverse momenta in the final state.

- Many models of electroweak symmetry breaking predict final states with large missing transverse energy. This arises in a large class of $R$-parity conserving supersymmetric theories containing a neutral, stable, lightest supersymmetric particle; in theories with "large" extra dimensions containing a Kaluza-Klein tower of gravitons that escape into the multidimensional "bulk space" [4]; and more generally from neutrinos produced in electroweak boson decay. If the final state contains significant $\not\!\!E_T$, then $\not\!\!E_T$ is included in the list of promising variables. We do not use $\not\!\!E_T$ that is reconstructed as a $W$ boson decay product, following the prescription for $W$ and $Z$ boson identification outlined above.

- If the final state contains one or more leptons we use the summed scalar transverse momenta $\sum p_T^\ell$, where the sum is over all leptons whose identity can be determined and whose momenta can be accurately measured. Leptons that are reconstructed as $W$ or $Z$ boson decay products are not included in this sum, again following the prescription for $W$ and $Z$ boson identification outlined above. We combine the momenta of $e$, $\mu$, and $\tau$ leptons because these objects are expected to have comparable transverse momenta on the basis of lepton universality in the standard model and the negligible values of lepton masses.

- Similarly, photons and $W$ and $Z$ bosons are most likely to signal the presence of new phenomena when they are produced at high transverse momentum. Since the expected transverse momenta of the electroweak gauge bosons are comparable, we use the variable $\sum p_T^{\gamma/W/Z}$, where the scalar sum is over all electroweak gauge bosons in the event, for final states with one or more of them identified.

- For events with one jet in the final state, the transverse energy of that jet is an important variable. For events with two or more jets in the final state, previous analyses have made use of the sum of the transverse energies of all but the leading jet [23]. The reason for excluding the energy of the leading jet from this sum is that while a hard jet is often obtained from QCD radiation, hard second and third radiative jets are relatively much less likely. We therefore choose the variable $\sum' p_T^j$ to describe the jets in the final state, where $\sum' p_T^j$ denotes $p_T^{j_1}$ if the final state contains only one jet, and $\sum_{i=2}^{n} p_T^{j_i}$ if the final state contains two or more jets. Since QCD dijets are a large background in all-jets final states, $\sum' p_T^j$ refers instead to $\sum_{i=3}^{n} p_T^{j_i}$ for final states containing $n$ jets and nothing else, where $n \geq 3$.

When there are exactly two objects in an event (e.g., one $Z$ boson and one jet), their $p_T$ values are expected to be nearly equal, and we therefore use the average $p_T$ of the two objects. When there is only one object in an event (e.g., a single $W$ boson), we use no variables, and simply perform a counting experiment.

Other variables that can help pick out specific signatures can also be defined. Although variables such as invariant mass, angular separation between particular final state objects, and variables that characterize event topologies may be useful in testing a particular

model, these variables tend to be less powerful in a general search. Appendix A.1 contains a more detailed discussion of this point. In the interest of keeping the list of variables as general, well-motivated, powerful, and short as possible, we elect to stop with those given in Table 4.1. We expect evidence for new physics to appear in the high tails of the $\not\!\!E_T$, $\sum p_T^{\ell}$, $\sum p_T^{\gamma/W/Z}$, and $\sum' p_T^{j}$ distributions.

| If the final state includes | then consider the variable |
|---|---|
| $\not\!\!E_T$ | $\not\!\!E_T$ |
| one or more charged leptons | $\sum p_T^{\ell}$ |
| one or more electroweak bosons | $\sum p_T^{\gamma/W/Z}$ |
| one or more jets | $\sum' p_T^{j}$ |

Table 4.1: A quasi-model-independently motivated list of interesting variables for any final state. The set of variables to consider for any particular final state is the union of the variables in the second column for each row that pertains to that final state. Here $\ell$ denotes $e$, $\mu$, or $\tau$. The notation $\sum' p_T^{j}$ is shorthand for $p_T^{j_1}$ if the final state contains only one jet, $\sum_{i=2}^{n} p_T^{j_i}$ if the final state contains $n \geq 2$ jets, and $\sum_{i=3}^{n} p_T^{j_i}$ if the final state contains $n$ jets and nothing else, with $n \geq 3$. Leptons and missing transverse energy that are reconstructed as decay products of $W$ or $Z$ bosons are not considered separately in the left-hand column.

### 4.2.2 Search strategy: DØ Run I

The general search strategy just outlined is applicable to any collider experiment searching for the physics responsible for electroweak symmetry breaking. Any particular experiment that wishes to use this strategy needs to specify object and variable definitions that reflect the capabilities of the detector. This section serves this function for the DØ detector [5] in its 1992–1996 run (Run I) at the Fermilab Tevatron. Details in this subsection supersede those in the more general section above.

**Object definitions**

The particle identification algorithms used here for electrons, muons, jets, and photons are similar to those used in many published DØ analyses. We summarize them here.

**Electrons.** DØ had no central magnetic field in Run I; therefore, there is no way to distinguish between electrons and positrons. Electron candidates with transverse energy greater than 15 GeV, within the fiducial region of $|\eta|< 1.1$ or $1.5 <|\eta|< 2.5$ (where $\eta = -\ln\tan(\theta/2)$, with $\theta$ the polar angle with respect to the colliding proton's direction), and satisfying standard electron identification and isolation requirements as defined in Ref. [24] are accepted.

**Muons.** We do not distinguish between positively and negatively charged muons in this analysis. We accept muons with transverse momentum greater than 15 GeV and $|\eta|< 1.7$ that satisfy standard muon identification and isolation requirements [24].

**$\not\!\!E_T$.** The missing transverse energy, $\not\!\!E_T$, is the energy required to balance the measured energy in the event. In the calorimeter, we calculate

$$\not\!\!E_T^{\rm cal} = |\sum_i E_i \sin\theta_i (\cos\phi_i\,\hat{x} + \sin\phi_i\,\hat{y})|, \tag{4.1}$$

where $i$ runs over all calorimeter cells, $E_i$ is the energy deposited in the $i^{th}$ cell, and $\phi_i$ is the azimuthal and $\theta_i$ the polar angle of the center of the $i^{th}$ cell, measured with respect to the event vertex.

An event is defined to contain a $\not\!\!E_T$ "object" only if we are confident that there

is significant missing transverse energy. Events that do not contain muons are said to contain $\not\!\!E_T$ if $\not\!\!E_T^{\rm cal} > 15$ GeV. Using track deflection in magnetized steel toroids, the muon momentum resolution in Run I is

$$\delta(1/p) = 0.18(p-2)/p^2 \oplus 0.003, \tag{4.2}$$

where $p$ is in units of GeV, and the $\oplus$ means addition in quadrature. This is significantly coarser than the electromagnetic and jet energy resolutions, parameterized by

$$\delta E/E = 15\%/\sqrt{E} \oplus 0.3\% \tag{4.3}$$

and

$$\delta E/E = 80\%/\sqrt{E}, \tag{4.4}$$

respectively. Events that contain exactly one muon are deemed to contain $\not\!\!E_T$ on the basis of muon number conservation rather than on the basis of the muon momentum measurement. We do not identify a $\not\!\!E_T$ object in events that contain two or more muons.

**Jets.** Jets are reconstructed in the calorimeter using a fixed-size cone algorithm, with a cone size of $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = 0.5$ [25]. We require jets to have $E_T > 15$ GeV and $|\eta| < 2.5$. We make no attempt to distinguish among light quarks, gluons, charm quarks, bottom quarks, and hadronic tau decays.

**Photons.** Isolated photons that pass standard identification requirements [26], have transverse energy greater than 15 GeV, and are in the fiducial region $|\eta| < 1.1$ or $1.5 < |\eta| < 2.5$ are labeled photon objects.

**$W$ bosons.** Following the general prescription described above, an electron (as defined above) and $\not\!\!E_T$ become a $W$ boson if their transverse mass is within the $W$ boson mass window ($30 \leq M^T_{\ell \not\!\!E_T} \leq 110$ GeV), and the event contains no second charged lepton. Because the muon momentum measurement is coarse, we do not use a transverse mass window for muons. From Sec. 4.2.2, any event containing a single muon is said to also contain $\not\!\!E_T$; thus any event containing a muon and no second charged lepton is said to contain a $W$ boson.

**$Z$ bosons.** We use the rules in the previous section for combining an $ee$ pair or $ee\gamma$ triplet into a $Z$ boson. We do not attempt to reconstruct a $Z$ boson in events containing three or more charged leptons. For events containing two muons and no third charged lepton, we fit the event to the hypothesis that the two muons are decay products of a $Z$ boson and that there is no $\not\!\!E_T$ in the event. If the fit is acceptable, the two muons are considered to be a $Z$ boson.

**Variables**

The variables provided in the general prescription above also need minor revision to be appropriate for the DØ experiment.

**$\sum p_T^\ell$.** We do not attempt to identify $\tau$ leptons, and the momentum resolution for muons is coarse. For events that contain no leptons other than muons, we define $\sum p_T^\ell = \sum p_T^\mu$. For events that contain one or more electrons, we define $\sum p_T^\ell = \sum p_T^e$. This is identical to the general definition provided above except for events containing both one or more electrons and one or more muons. In this case, we have decided to define $\sum p_T^\ell$ as the sum of the momenta of the electrons only, rather than combining the well-measured electron momenta

with the poorly-measured muon momenta.

**$\not\!\!E_T$.** $\not\!\!E_T$ is defined by $\not\!\!E_T = \not\!\!E_T^{\rm cal}$, where $\not\!\!E_T^{\rm cal}$ is the missing transverse energy as summed in the calorimeter. This sum includes the $p_T$ of electrons, but only a negligible fraction of the $p_T$ of muons.

**$\sum p_T^{\gamma/W/Z}$.** We use the definition of $\sum p_T^{\gamma/W/Z}$ provided in the general prescription: the sum is over all electroweak gauge bosons in the event, for final states with one or more of them. We note that if a $W$ boson is formed from a $\mu$ and $\not\!\!E_T$, then $p_T^W = \not\!\!E_T^{\rm cal}$.

## 4.3 Sleuth algorithm

Given a data sample, its final state, and a set of variables appropriate to that final state, we now describe the algorithm that determines the most interesting region in those variables and quantifies the degree of interest.

### 4.3.1 Overview

Central to the algorithm is the notion of a "region" ($R$). A region can be regarded simply as a volume in the variable space defined by Table 4.1, satisfying certain special properties to be discussed in Sec. 4.3.2. The region contains $N$ data points and an expected number of background events $\hat{b}_R$. We can consequently compute the weighted probability $p_N^R$, defined in Sec. 4.3.3, that the background in the region fluctuates up to or beyond the observed number of events. If this probability is small, we flag the region as potentially interesting.

In any reasonably-sized data set, there will always be regions in which the probability for $b_R$ to fluctuate up to or above the observed number of events is small. The relevant issue is how often this can happen in an ensemble of hypothetical similar experiments (hse's). This question can be answered by performing these hypothetical similar experiments; i.e., by generating random events drawn from the background distribution, finding the least probable region, and repeating this many times. The fraction of hypothetical similar experiments that yields a probability as low as the one observed in the data provides the appropriate measure of the degree of interest.

Although the details of the algorithm are complex, the interface is straightforward. What is needed is a data sample, a set of events for each background process $i$, and the number of background events $\hat{b}_i \pm \delta\hat{b}_i$ from each background process expected in the data sample. The output gives the region of greatest excess and the fraction of hypothetical similar experiments that would yield such an excess.

The algorithm consists of seven steps:

1. Define regions $R$ about any chosen set of $N = 1, \ldots, N_{\text{data}}$ data points in the sample of $N_{\text{data}}$ data points.

2. Estimate the background $\hat{b}_R$ expected within these $R$.

3. Calculate the weighted probabilities $p_N^R$ that $b_R$ can fluctuate to $\geq N$.

4. For each $N$, determine the $R$ for which $p_N^R$ is minimum. Define $p_N = \min_R (p_N^R)$.

5. Determine the fraction $P_N$ of hypothetical similar experiments in which the $p_N(\text{hse})$ is smaller than the observed $p_N(\text{data})$.

6. Determine the $N$ for which $P_N$ is minimized. Define $P = \min_N (P_N)$.

7. Determine the fraction $\mathcal{P}$ of hypothetical similar experiments in which the $P(\text{hse})$ is smaller than the observed $P(\text{data})$.

Our notation is such that a lowercase $p$ represents a probability, while an uppercase $P$ or $\mathcal{P}$ represents the fraction of hypothetical similar experiments that would yield a less probable outcome. The symbol representing the minimization of $p_N^R$ over $R$, $p_N$ over $N$, or $P_N$ over $N$ is written without the superscript or subscript representing the varied property (i.e., $p_N$, $p$, or $P$, respectively). The rest of this section discusses these steps in greater detail.

### 4.3.2 Steps 1 and 2: Regions

When there are events that do not appear to follow some expected distribution, such as the event at $x = 61$ in Fig. 4.1, we often attempt to estimate the probability that the event is consistent with coming from that distribution. This is generally done by choosing some region around the event (or an accumulation of events), integrating the background within that region, and computing the probability that the expected number of events in that region could have fluctuated up to or beyond the observed number.

Of course, the calculated probability depends on how the region containing the events is chosen. If the region about the event is infinitesimal, then the expected number of background events in the region (and therefore this probability) can be made arbitrarily small. A possible approach in one dimension is to define the region to be the interval bounded below by the point halfway between the interesting event and its nearest neighbor, and bounded above by infinity. For the case shown in Fig. 4.1, this region would be roughly

the interval $(46, \infty)$.

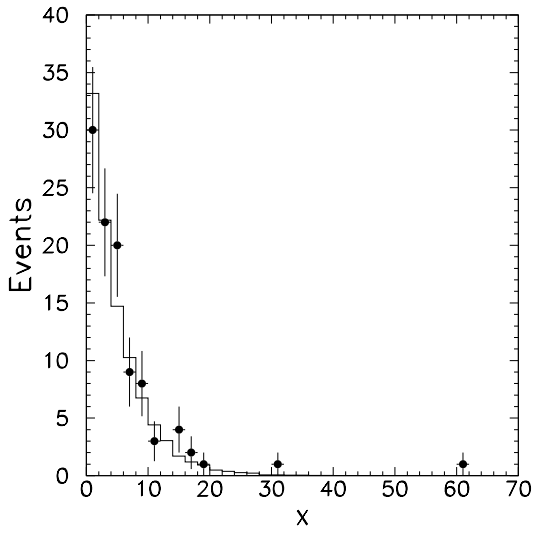



Figure 4.1: Example of a data set with a potentially anomalous point. The solid histogram is the expected distribution, and the points with error bars are the data. The bulk of the data is well described by the background prediction, but the point located at $x = 61$ appears out of place.

Such a prescription breaks down in two or more dimensions, and it is not entirely satisfactory even in one dimension. In particular, it is not clear how to proceed if the excess occurs somewhere other than at the tail end of a distribution, or how to generalize the interval to a well-defined contour in several dimensions. As we will see, there are significant advantages to having a precise definition of a region about a potentially interesting set of data points. This is provided in Sec. 4.3.2, after we specify the variable space itself.

**Variable transformation**

Unfortunately, the region that we choose about the point on the tail of Fig. 4.1 changes if the variable is some function of $x$, rather than $x$ itself. If the region about each

data point is to be the subspace that is closer to that point than to any other one in the sample, it would therefore be wise to minimize any dependence of the selection on the shape of the background distribution. For a background distributed uniformly between 0 and 1 (or, in $d$ dimensions, uniform within the unit "box" $[0,1]^d$), it is reasonable to define the region associated with an event as the variable subspace closer to that event than to any other event in the sample. If the background is not already uniform within the unit box, we transform the variables so that it becomes uniform. The details of this transformation are provided in Appendix A.2.

With the background distribution trivialized, the rest of the analysis can be performed within the unit box without worrying about the background shape. A considerable simplification is therefore achieved through this transformation. The task of determining the expected background within each region, which would have required a Monte Carlo integration of the background distribution over the region, reduces to the problem of determining the volume of each region. The problem is now completely specified by the transformed coordinates of the data points, the total number of expected background events $\hat{b}$, and its uncertainty $\delta\hat{b}$.

**Voronoi diagrams**

Having defined the variable space by requiring a uniform background distribution, we can now define more precisely what is meant by a region. Figure 4.2 shows a 2-dimensional variable space $V$ containing seven data points in a unit square. For any $v \in V$, we say that $v$ *belongs to* the data point $D_i$ if $|v - D_i| < |v - D_j|$ for all $j \neq i$; that is, $v$ belongs to $D_i$ if $v$ is closer to $D_i$ than to any other data point. In Fig. 4.2(a), for example, any $v$ lying within

the variable subspace defined by the pentagon in the upper right-hand corner belongs to the data point located at $(0.9, 0.8)$. The set of points in $V$ that do not belong to any data point [those points on the lines in Fig. 4.2(a)] has zero measure and may be ignored.

We define a *region* around a set of data points in a variable space $V$ to be the set of all points in $V$ that are closer to one of the data points in that set than to any data points outside that set. A region around a single data point is the union of all points in $V$ that belong to that data point, and is called a 1-region. A region about a set of $N$ data points is the union of all points in $V$ that belong to any one of the data points, and is called an $N$-region; an example of a 2-region is shown as the shaded area in Fig. 4.2(b). $N_{\rm data}$ data points thus partition $V$ into $N_{\rm data}$ 1-regions. Two data points are said to be neighbors if their 1-regions share a border – the points at $(0.75, 0.9)$ and $(0.9, 0.8)$ in Fig. 4.2, for example, are neighbors. A diagram such as Fig. 4.2(a), showing a set of data points and their regions, is known as a *Voronoi diagram.* We use a program called HULL [27] for this computation.

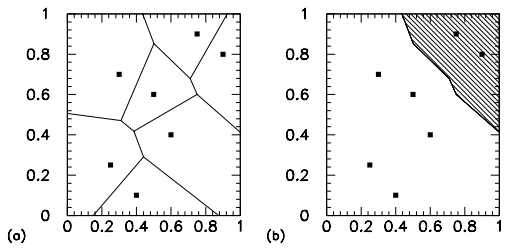



Figure 4.2: A Voronoi diagram. (a) The seven data points are shown as black dots; the lines partition the space into seven regions, with one region belonging to each data point. (b) An example of a 2-region.

**Region criteria**

The explicit definition of a region that we have just provided reduces the number of contours we can draw in the variable space from infinite to a mere $2^{N_{\rm data}} - 1$, since any region either contains all of the points belonging to the $i^{th}$ data event or it contains none of them. In fact, because many of these regions have a shape that makes them implausible as "discovery regions" in which new physics might be concentrated, the number of possible regions may be reduced further. For example, the region in Fig. 4.2 containing only the lower-leftmost and the upper-rightmost data points is unlikely to be a discovery region, whereas the region shown in Fig. 4.2(b) containing the two upper-rightmost data points is more likely (depending upon the nature of the variables).

We can now impose whatever criteria we wish upon the regions that we allow Sleuth to consider. In general we will want to impose several criteria, and in this case we write the net criterion $c_R = c_R^1 c_R^2 \ldots$ as a product of the individual criteria, where $c_R^i$ is to be read "the extent to which the region $R$ satisfies the criterion $c^i$." The quantities $c_R^i$ take on values in the interval $[0,1]$, where $c_R^i \to 0$ if $R$ badly fails $c^i$, and $c_R^i \to 1$ if $R$ easily satisfies $c^i$.

Consider as an example $c = \textit{AntiCornerSphere}$, a simple criterion that we have elected to impose on the regions in the $e\mu X$ sample. Loosely speaking, a region $R$ will satisfy this criterion ($c_R \to 1$) if all of the data points inside the region are farther from the origin than all of the data points outside the region. This situation is shown, for example, in Fig. 4.2(b). For every event $i$ in the data set, denote by $r_i$ the distance of the point in the unit box to the origin, let $r'$ be $r$ transformed so that the background is uniform in $r'$ over

the interval $[0,1]$, and let $r'_i$ be the values $r_i$ so transformed. Then define

$$c_R = \begin{cases} 0 & , \left(\frac{1}{2} + \frac{r'^{\text{in}}_{\text{min}} - r'^{\text{out}}_{\text{max}}}{\xi}\right) < 0 \\ \left(\frac{1}{2} + \frac{r'^{\text{in}}_{\text{min}} - r'^{\text{out}}_{\text{max}}}{\xi}\right) & , 0 \leq \left(\frac{1}{2} + \frac{r'^{\text{in}}_{\text{min}} - r'^{\text{out}}_{\text{max}}}{\xi}\right) \leq 1 \\ 1 & , 1 < \left(\frac{1}{2} + \frac{r'^{\text{in}}_{\text{min}} - r'^{\text{out}}_{\text{max}}}{\xi}\right) \end{cases} \tag{4.5}$$

where $r'^{\text{in}}_{\text{min}} = \min_{i \in R}(r'_i)$, $r'^{\text{out}}_{\text{max}} = \max_{i \notin R}(r'_i)$, and $\xi = 1/(4N_{\text{data}})$ is an average separation distance between data points in the variable $r'$.

Notice that in the limit of vanishing $\xi$, the criterion $c$ becomes a boolean operator, returning "true" when all of the data points inside the region are farther from the origin than all of the data points outside the region, and "false" otherwise. In fact, many possible criteria have a scale $\xi$ and reduce to boolean operators when $\xi$ vanishes. This scale has been introduced to ensure continuity of the final result under small changes in the background estimate. In this spirit, the "extent to which $R$ satisfies the criterion $c$" has an alternative interpretation as the "fraction of the time $R$ satisfies the criterion $c$," where the average is taken over an ensemble of slightly perturbed background estimates and $\xi$ is taken to vanish, so that "satisfies" makes sense. We will use $c_R$ in the next section to define an initial measure of the degree to which $R$ is interesting.

We have considered several other criteria that could be imposed upon any potential discovery region to ensure that the region is "reasonably shaped" and "in a believable location." We discuss a few of these criteria in Appendix A.3.

### 4.3.3 Step 3: Probabilities and uncertainties

Now that we have specified the notion of a region, we can define a quantitative measure of the "degree of interest" of a region.

**Probabilities**

Since we are looking for regions of excess, the appropriate measure of the degree of interest is a slight modification of the probability of background fluctuating up to or above the observed number of events. For an $N$-region $R$ in which $\hat{b}_R$ background events are expected and $\hat{b}_R$ is precisely known, this probability is

$$\sum_{i=N}^{\infty} \frac{e^{-\hat{b}_R}(\hat{b}_R)^i}{i!}. \tag{4.6}$$

We use this to define the weighted probability

$$p_N^R = \left( \sum_{i=N}^{\infty} \frac{e^{-\hat{b}_R}(\hat{b}_R)^i}{i!} \right) c_R + (1 - c_R), \tag{4.7}$$

which one can also think of as an "average probability," where the average is taken over the ensemble of slightly perturbed background estimates referred to above. By construction, this quantity has all of the properties we need: it reduces to the probability in Eq. 4.6 in the limit that $R$ easily satisfies the region criteria, it saturates at unity in the limit that $R$ badly fails the region criteria, and it exhibits continuous behavior under small perturbations in the background estimate between these two extremes.

**Systematic uncertainties**

The expected number of events from each background process has a systematic uncertainty that must be taken into account. There may also be an uncertainty in the shape of a particular background distribution — for example, the tail of a distribution may have a larger systematic uncertainty than the mode.

The background distribution comprises one or more contributing background processes. For each background process we know the number of expected events and the sys-

tematic uncertainty on this number, and we have a set of Monte Carlo points that tell us what that background process looks like in the variables of interest. A typical situation is sketched in Fig. 4.3.

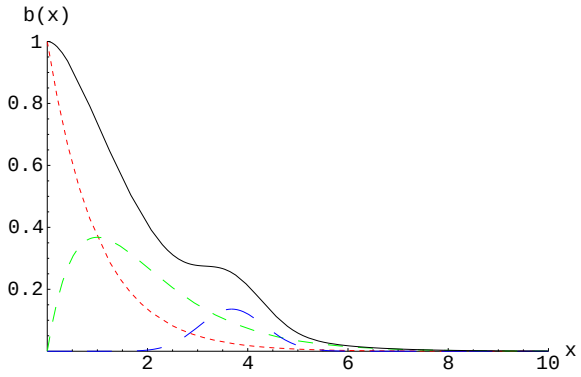



Figure 4.3: An example of a one-dimensional background distribution with three sources. The normalized shapes of the individual background processes are shown as the dashed lines; the solid line is their sum. Typically, the normalizations for the background processes have separate systematic errors. These errors can change the shape of the total background curve in addition to its overall normalization. For example, if the long-dashed curve has a large systematic error, then the solid curve will be known less precisely in the region $(3, 5)$ than in the region $(0, 3)$ where the other two backgrounds dominate.

The multivariate transformation described in Sec. 4.3.2 is obtained assuming that the number of events expected from each background process is known precisely. This fixes each event's position in the unit box, its neighbors, and the volume of the surrounding region. The systematic uncertainty $\delta\hat{b}_R$ on the number of background events in a given region is computed by combining the systematic uncertainties for each individual background process. Eq. 4.7 then generalizes to

$$p_N^R = c_R \int_0^\infty \sum_{i=N}^{\infty} \frac{e^{-b} b^i}{i!} \frac{1}{\sqrt{2\pi}(\delta\hat{b}_R)} \exp\left(-\frac{(b-\hat{b}_R)^2}{2(\delta\hat{b}_R)^2}\right) db \; + \; (1 - c_R), \tag{4.8}$$

which is seen to reduce to Eq. 4.7 in the limit $\delta\hat{b}_R \to 0$.

This formulation provides a way to take account of systematic uncertainties on the

shapes of distributions, as well. For example, if there is a larger systematic uncertainty on the tail of a distribution, then the background process can be broken into two components, one describing the bulk of the distribution and one describing the tail, and a larger systematic uncertainty assigned to the piece that describes the tail. Correlations among the various components may also be assigned.

We vary the number of events generated in the hypothetical similar experiments according to the systematic and statistical uncertainties. The systematic errors are accounted for by pulling a vector of the "true" number of expected background events $\vec{b}$ from the distribution

$$p(\vec{b}) = \frac{1}{\sqrt{2\pi\ |\Sigma|}} \exp\left(-\frac{1}{2}(b_i - \hat{b}_i)\Sigma_{ij}^{-1}(b_j - \hat{b}_j)\right), \tag{4.9}$$

where $\hat{b}_i$ is the number of expected background events from process $i$, as before, and $b_i$ is the $i^{th}$ component of $\vec{b}$. We have introduced a covariance matrix $\Sigma$, which is diagonal with components $\Sigma_{ii} = (\delta\hat{b}_i)^2$ in the limit that the systematic uncertainties on the different background processes are uncorrelated, and we assume summation on repeated indices in Eq. 4.9. The statistical uncertainties in turn are allowed for by choosing the number of events $N_i$ from each background process $i$ from the Poisson distribution

$$P(N_i) = \frac{e^{-b_i} b_i^{N_i}}{N_i!}, \tag{4.10}$$

where $b_i$ is the $i^{th}$ component of the vector $\vec{b}$ just determined.

### 4.3.4 Step 4: Exploration of regions

Knowing how to calculate $p_N^R$ for a specific $N$-region $R$ allows us to determine which of two $N$-regions is more interesting. Specifically, an $N$-region $R_1$ is more interesting than

another $N$-region $R_2$ if $p_N^{R_1} < p_N^{R_2}$. This allows us to compare regions of the same size (the same $N$), although, as we will see, it does not allow us to compare regions of different size.

Step 4 of the algorithm involves finding the most interesting $N$-region for each fixed $N$ between 1 and $N_{\text{data}}$. This most interesting $N$-region is the one that minimizes $p_N^R$, and these $p_N = \min_R(p_N^R)$ are needed for the next step in the algorithm.

Even for modestly sized problems (say, two dimensions with on the order of 100 data points), there are far too many regions to consider an exhaustive search. We therefore use a heuristic to find the most interesting region. We imagine the region under consideration to be an amoeba moving within the unit box. At each step in the search the amoeba either expands or contracts according to certain rules, and along the way we keep track of the most interesting $N$-region so far found, for each $N$. The detailed rules for this heuristic are provided in Appendix A.4.

### 4.3.5 Steps 5 and 6: Hypothetical similar experiments, Part I

At this point in the algorithm the original events have been reduced to $N_{\text{data}}$ values, each between 0 and 1: the $p_N$ ($N = 1, \ldots, N_{\text{data}}$) corresponding to the most interesting $N$-regions satisfying the imposed criteria. To find the *most* interesting of these, we need a way of comparing regions of different size (different $N$). An $N_1$-region $R_{N_1}$ with $p_{N_1}^{\text{data}}$ is more interesting than an $N_2$-region $R_{N_2}$ with $p_{N_2}^{\text{data}}$ if the fraction of hypothetical similar experiments in which $p_{N_1}^{\text{hse}} < p_{N_1}^{\text{data}}$ is less than the fraction of hypothetical similar experiments in which $p_{N_2}^{\text{hse}} < p_{N_2}^{\text{data}}$.

To make this comparison, we generate $N_{\text{hse}^1}$ hypothetical similar experiments. Generating a hypothetical similar experiment involves pulling a random integer from Eq. 4.10

for each background process $i$, sampling this number of events from the multidimensional background density $b(\vec{x})$, and then transforming these events into the unit box.

For each hse we compute a list of $p_N$, exactly as for the data set. Each of the $N_{\mathrm{hse}^1}$ hypothetical similar experiments consequently yields a list of $p_N$. For each $N$, we now compare the $p_N$ we obtained in the data ($p_N^{\mathrm{data}}$) with the $p_N$'s we obtained in the hse's ($p_N^{\mathrm{hse}_i^1}$, where $i = 1, \ldots, N_{\mathrm{hse}^1}$). From these values we calculate $P_N$, the fraction of hse's with $p_N^{\mathrm{hse}^1} < p_N^{\mathrm{data}}$:

$$P_N = \frac{1}{N_{\mathrm{hse}^1}} \sum_{i=1}^{N_{\mathrm{hse}^1}} \Theta\left(p_N^{\mathrm{data}} - p_N^{\mathrm{hse}_i^1}\right), \tag{4.11}$$

where $\Theta(x) = 0$ for $x < 0$, and $\Theta(x) = 1$ for $x \geq 0$.

The most interesting region in the sample is then the region for which $P_N$ is smallest. We define $P = P_{N_{\mathrm{min}}}$, where $P_{N_{\mathrm{min}}}$ is the smallest of the $P_N$.

### 4.3.6 Step 7: Hypothetical similar experiments, Part II

A question that remains to be answered is what fraction $\mathcal{P}$ of hypothetical similar experiments would yield a $P$ less than the $P$ obtained in the data. We calculate $\mathcal{P}$ by running a second set of $N_{\mathrm{hse}^2}$ hypothetical similar experiments, generated as described in the previous section. (We have written $\mathrm{hse}^1$ above to refer to the first set of hypothetical similar experiments, used to determine the $P_N$, given a list of $p_N$; we write $\mathrm{hse}^2$ to refer to this second set of hypothetical similar experiments, used to determine $\mathcal{P}$ from $P$.) A second, independent set of hse's is required to calculate an unbiased value for $\mathcal{P}$. The quantity $\mathcal{P}$ is then given by

$$\mathcal{P} = \frac{1}{N_{\mathrm{hse}^2}} \sum_{i=1}^{N_{\mathrm{hse}^2}} \Theta\left(P^{\mathrm{data}} - P^{\mathrm{hse}_i^2}\right). \tag{4.12}$$

This is the final measure of the degree of interest of the most interesting region. Note that $\mathcal{P}$ is a number between 0 and 1, that small values of $\mathcal{P}$ indicate a sample containing an interesting region, that large values of $\mathcal{P}$ indicate a sample containing no interesting region, and that $\mathcal{P}$ can be described as the fraction of hypothetical similar experiments that yield a more interesting result than is observed in the data. $\mathcal{P}$ can be translated into units of standard deviations ($\mathcal{P}_{[\sigma]}$) by solving the unit conversion equation

$$\mathcal{P} = \frac{1}{\sqrt{2\pi}} \int_{\mathcal{P}_{[\sigma]}}^{\infty} e^{-t^2/2}\, dt \tag{4.13}$$

for $\mathcal{P}_{[\sigma]}$.

### 4.3.7 Interpretation of results

In a general search for new phenomena, Sleuth will be applied to $N_{\rm fs}$ different final states, resulting in $N_{\rm fs}$ different values for $\mathcal{P}$. The final step in the procedure is the combination of these results. If no $\mathcal{P}$ value is smaller than $\approx 0.01$ then a null result has been obtained, as no significant signal for new physics has been identified in the data.

If one or more of the $\mathcal{P}$ values *is* particularly low, then we can surmise that the region(s) of excess corresponds either to a poorly modeled background or to possible evidence of new physics. The algorithm has pointed out a region of excess ($\mathcal{R}$) and has quantified its significance ($\mathcal{P}$). The next step is to interpret this result.

Two issues related to this interpretation are combining results from many final states, and confirming a Sleuth discovery.

**Combining the results of many final states**

If one looks at many final states, one expects eventually to see a fairly small $\mathcal{P}$, even if there really is no new physics in the data. We therefore define a quantity $\tilde{\mathcal{P}}$ to be the fraction of hypothetical similar *experimental runs*[1] that yield a $\mathcal{P}$ that is smaller than the smallest $\mathcal{P}$ observed in the data. Explicitly, given $N_{\rm fs}$ final states, with $\hat{b}_i$ background events expected in each, and $\mathcal{P}_i$ calculated for each one, $\tilde{\mathcal{P}}$ is given to good approximation by[2]

$$\tilde{\mathcal{P}} = 1 - \prod_{i=1}^{N_{\rm fs}} \sum_{j=0}^{n_i-1} \frac{e^{-\hat{b}_i}\hat{b}_i^j}{j!}, \tag{4.14}$$

where $n_i$ is the smallest integer satisfying

$$\sum_{j=n_i}^{\infty} \frac{e^{-\hat{b}_i}\hat{b}_i^j}{j!} \leq \mathcal{P}_{\rm min} = \min_i \mathcal{P}_i. \tag{4.15}$$

**Confirmation**

An independent confirmation is desirable for any potential discovery, especially for an excess revealed by a data-driven search. Such confirmation may come from an independent experiment, from the same experiment in a different but related final state, from an independent confirmation of the background estimate, or from the same experiment in the same final state using independent data. In the last of these cases, a first sample can be presented to Sleuth to uncover any hints of new physics, and the remaining sample can be

[1] In the phrase "hypothetical similar experiment," "experiment" refers to the analysis of a single final state. We use "experimental runs" in a similar way to refer to the analysis of a number of different final states. Thus a hypothetical similar experimental run consists of $N_{\rm fs}$ different hypothetical similar experiments, one for each final state analyzed.

[2] Note that the naive expression $\tilde{\mathcal{P}} = 1 - (1 - \mathcal{P}_{\rm min})^{N_{\rm fs}}$ is not correct, since this requires $\tilde{\mathcal{P}} \to 1$ for $N_{\rm fs} \to \infty$, and there are indeed an infinite number of final states to examine. The resolution of this paradox hinges on the fact that only an integral number of events can be observed in each final state, and therefore final states with $\hat{b}_i \ll 1$ contribute very little to the value of $\tilde{\mathcal{P}}$. This is correctly accounted for in the formulation given in Eq. 4.14.

subjected to a standard analysis in the region suggested by Sleuth. An excess in this region in the second sample helps to confirm a discrepancy between data and background. If we see hints of new physics in the Run I data, for example, we will be able to predict where new physics might show itself in the upcoming run of the Fermilab Tevatron, Run II.

## 4.4 The $e\mu X$ data set

As mentioned in Sec. 4.1, we have applied the Sleuth method to DØ data containing one or more electrons and one or more muons. We use a data set corresponding to 108.3±5.7 $\text{pb}^{-1}$ of integrated luminosity, collected between 1992 and 1996 at the Fermilab Tevatron with the DØ detector. The data set and basic selection criteria are identical to those used in the published $t\bar{t}$ cross section analysis for the dilepton channels [24]. Specifically, we apply global cleanup cuts and select events containing

- one or more high $p_T$ ($p_T > 15$ GeV) isolated electrons, and
- one or more high $p_T$ ($p_T > 15$ GeV) isolated muons,

with object definitions given in Sec. 4.2.2.

The dominant standard model and instrumental backgrounds to this data set are

- top quark pair production with $t \to Wb$, and with both $W$ bosons decaying leptonically, one to $e\nu$ (or to $\tau\nu \to e\nu\nu\nu$) and one to $\mu\nu$ (or to $\tau\nu \to \mu\nu\nu\nu$),
- $W$ boson pair production with both $W$ bosons decaying leptonically, one to $e\nu$ (or to $\tau\nu \to e\nu\nu\nu$) and one to $\mu\nu$ (or to $\tau\nu \to \mu\nu\nu\nu$),
- $Z/\gamma^* \to \tau\tau \to e\mu\nu\nu\nu\nu$, and

- instrumental ("fakes"): $W$ production with the $W$ boson decaying to $\mu\nu$ and a radiated jet or photon being mistaken for an electron, or $b\bar{b}/c\bar{c}$ production with one heavy quark producing an isolated muon and the other a false electron [25].

A sample of 100,000 $t\bar{t} \rightarrow$ dilepton events was generated using HERWIG [28], and a $WW$ sample of equal size was generated using PYTHIA [29]. We generated $\gamma^* \rightarrow \tau\tau \rightarrow e\mu\nu\nu\nu\nu$ (Drell-Yan) events using PYTHIA and $Z \rightarrow \tau\tau \rightarrow e\mu\nu\nu\nu\nu$ events using ISAJET [30]. The Drell-Yan cross section is normalized as in Ref. [31]. The cross section for $Z \rightarrow \tau\tau$ is taken to be equal to the published DØ $Z \rightarrow ee$ cross section [32]; the top quark production cross section is taken from Ref. [33]; and the $WW$ cross section is taken from Ref. [34]. The $t\bar{t}$, $WW$, and $Z/\gamma^*$ Monte Carlo events all were processed through GEANT [15] and the DØ reconstruction software. The number and distributions of events containing fake electrons are taken from data, using a sample of events satisfying "bad" electron identification criteria [35].

We break $e\mu X$ into exclusive data sets, and determine which variables to consider in each set using the prescription given in Sec. 4.2. The exclusive final states within $e\mu X$ that are populated with events in the data are listed in Table 4.2. The number of events expected for the various samples and data sets in the populated final states are given in Table 4.3; the number of expected background events in all unpopulated final states in which the number of expected background events is $> 0.001$ are listed in Table 4.4. The dominant sources of systematic error are given in Table 4.5.

| Final State | Variables |
|---|---|
| $e\mu\not\!\!E_T$ | $p_T^e$, $\not\!\!E_T$ |
| $e\mu\not\!\!E_T j$ | $p_T^e$, $\not\!\!E_T$, $p_T^j$ |
| $e\mu\not\!\!E_T jj$ | $p_T^e$, $\not\!\!E_T$, $p_T^{j_2}$ |
| $e\mu\not\!\!E_T jjj$ | $p_T^e$, $\not\!\!E_T$, $p_T^{j_2} + p_T^{j_3}$ |

Table 4.2: The exclusive final states within $e\mu X$ for which events are seen in the data and the variables used for each of these final states. The variables are selected using the prescription described in Sec. 4.2. Although all final states contain "$e\mu\not\!\!E_T$," no missing transverse energy cut has been applied explicitly; $\not\!\!E_T$ is inferred from the presence of the muon, following Sec. 4.2.2.

| Data set | Fakes | $Z \to \tau\tau$ | $\gamma^* \to \tau\tau$ | $WW$ | $t\bar{t}$ | Total |
|---|---|---|---|---|---|---|
| $e\mu\not\!\!E_T$ | 18.4±1.4 | 25.6±6.5 | 0.5±0.2 | 3.9±1.0 | 0.011±0.003 | 48.5±7.6 |
| $e\mu\not\!\!E_T j$ | 8.7±1.0 | 3.0 ± 0.8 | 0.1±0.03 | 1.1±0.3 | 0.4±0.1 | 13.2±1.5 |
| $e\mu\not\!\!E_T jj$ | 2.7±0.6 | 0.5±0.2 | 0.012±0.006 | 0.18±0.05 | 1.8±0.5 | 5.2±0.8 |
| $e\mu\not\!\!E_T jjj$ | 0.4±0.2 | 0.07±0.05 | 0.005±0.004 | 0.032±0.009 | 0.7±0.2 | 1.3±0.3 |
| $e\mu X$ | 30.2±1.8 | 29.2±4.5 | 0.7±0.1 | 5.2±0.8 | 3.1±0.5 | 68.3±5.7 |

Table 4.3: The number of expected background events for the populated final states within $e\mu X$. The errors on $e\mu X$ are smaller than on the sum of the individual background contributions obtained from Monte Carlo because of an uncertainty on the number of extra jets arising from initial and final state radiation in the exclusive channels.

| Final State | Background expected |
|---|---|
| $e\mu\not\!\!E_T jjjj$ | 0.3 ± 0.15 |
| $ee\mu\not\!\!E_T$ | 0.10 ± 0.05 |
| $e\mu\mu$ | 0.04 ± 0.02 |
| $e\mu\not\!\!E_T\gamma$ | 0.06 ± 0.03 |

Table 4.4: The number of expected background events for the unpopulated final states within $e\mu X$. The expected number of events in final states with additional jets is obtained from those listed in the table by dividing by five for each jet. These are all rough estimates, and a large systematic error has been assigned accordingly. Since no events are seen in any of these final states, the background estimates shown here are used solely in the calculation of $\tilde{\mathcal{P}}$ for all $e\mu X$ channels.

## 4.5 Sensitivity

We choose to consider the $e\mu X$ final state first because it contains backgrounds of mass scale comparable to that expected of the physics responsible for electroweak symmetry breaking. Top quark pair production ($q\bar{q} \to t\bar{t} \to W^+W^-b\bar{b}$) and $W$ boson pair production

| Source | Error |
|---|---|
| Trigger and lepton identification efficiencies | 12% |
| $P(j \rightarrow$"$e$") | 7% |
| Multiple Interactions | 7% |
| Luminosity | 5.3% |
| $\sigma(t\bar{t} \rightarrow e\mu X)$ | 12% |
| $\sigma(Z \rightarrow \tau\tau \rightarrow e\mu X)$ | 10% |
| $\sigma(WW \rightarrow e\mu X)$ | 10% |
| $\sigma(\gamma^* \rightarrow \tau\tau \rightarrow e\mu X)$ | 17% |
| Jet modeling | 20% |

Table 4.5: Sources of systematic uncertainty on the number of expected background events in the final states $e\mu\not\!\!E_T$, $e\mu\not\!\!E_T j$, $e\mu\not\!\!E_T jj$, and $e\mu\not\!\!E_T jjj$. $P(j \rightarrow$"$e$") denotes the probability that a jet will be reconstructed as an electron. "Jet modeling" includes systematic uncertainties in jet production in PYTHIA and HERWIG in addition to jet identification and energy scale uncertainties.

are excellent examples of the type of physics that we would expect the algorithm to find.

Before examining the data, we decided to impose the requirements of AntiCornerSphere and Isolation (see Appendix A.3) on the regions that Sleuth is allowed to consider. The reason for this choice is that, in addition to allowing only "reasonable" regions, it allows the search to be parameterized essentially by a single variable — the distance between each region and the lower left-hand corner of the unit box. We felt this would aid the interpretation of the results from this initial application of the method.

We test the sensitivity in two phases, keeping in mind that nothing in the algorithm has been "tuned" to finding $WW$ and $t\bar{t}$ in this sample. We first consider the background to comprise fakes and $Z/\gamma^* \rightarrow \tau\tau$ only, to see if we can "discover" either $WW$ or $t\bar{t}$. We then consider the background to comprise fakes, $Z/\gamma^* \rightarrow \tau\tau$, and $WW$, to see whether we can "discover" $t\bar{t}$. We apply the full search strategy and algorithm in both cases, first (in this section) on an ensemble of mock samples, and then (in Sec. 4.6) on the data.

### 4.5.1 Search for $WW$ and $t\bar{t}$ in mock samples

In this section we provide results from Sleuth for the case in which $Z/\gamma^* \to \tau\tau$ and fakes are included in the background estimates and the signal from $WW$ and $t\bar{t}$ is "unknown." We apply the prescription to the exclusive $e\mu X$ final states listed in Table 4.2.

Figure 4.4 shows distributions of $\mathcal{P}$ for mock samples containing only $Z/\gamma^* \to \tau\tau$ and fakes, where the mock events are pulled randomly from their parent distributions and the numbers of events are allowed to vary within systematic and statistical errors. The distributions are uniform in the interval $[0, 1]$, as expected, becoming appropriately discretized in the low statistics limit. (When the number of expected background events $\hat{b} \lesssim 1$, as in Fig. 4.4(d), it can happen that zero or one events are observed. If zero events are observed then $\mathcal{P} = 1$, since all hypothetical similar experiments yield a result as interesting or more interesting than an empty sample. If one event is observed then there is only one region for Sleuth to consider, and $\mathcal{P}$ is simply the probability for $\hat{b} \pm \delta\hat{b}$ to fluctuate up to exactly one event. In Fig. 4.4(d), for example, the spike at $\mathcal{P} = 1$ contains 62% of the mock experiments, since this is the probability for $0.5 \pm 0.2$ to fluctuate to zero events; the second spike is located at $\mathcal{P} = 0.38$ and contains 28% of the mock experiments, since this is the probability for $0.5 \pm 0.2$ to fluctuate to exactly one event. Similar but less pronounced behavior is seen in Fig. 4.4(c).) Figure 4.5 shows distributions of $\mathcal{P}$ when the mock samples contain $WW$ and $t\bar{t}$ in addition to the background in Fig. 4.4. Again, the number of events from each process is allowed to vary within statistical and systematic error. Figure 4.5 shows that we can indeed find $t\bar{t}$ and/or $WW$ much of the time. Figure 4.6 shows $\tilde{\mathcal{P}}$ computed for these samples. In over 50% of these samples we find $\tilde{\mathcal{P}}_{[\sigma]}$ to correspond to more than two

standard deviations.

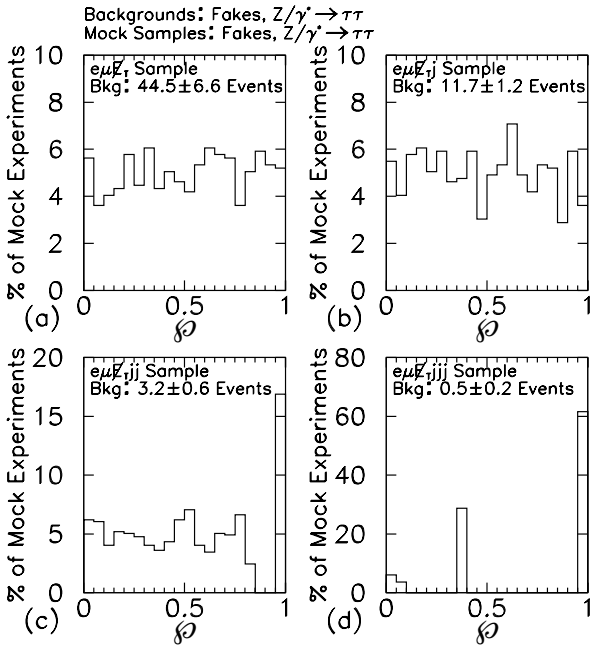



Figure 4.4: Distributions of $\mathcal{P}$ for the four exclusive final states (a) $e\mu\not{E}_T$, (b) $e\mu\not{E}_T j$, (c) $e\mu\not{E}_T jj$, and (d) $e\mu\not{E}_T jjj$. The background includes only $Z/\gamma^* \to \tau\tau$ and fakes, and the mock samples making up these distributions also contain only these two sources. As expected, $\mathcal{P}$ is uniform in the interval $[0, 1]$ for those final states in which the expected number of background events $\hat{b} \gg 1$, and shows discrete behavior for $\hat{b} \lesssim 1$.

### 4.5.2 Search for $t\bar{t}$ in mock samples

In this section we provide results for the case in which $Z/\gamma^* \to \tau\tau$, fakes, and $WW$ are all included in the background estimate, and $t\bar{t}$ is the "unknown" signal. We again apply the prescription to the exclusive final states listed in Table 4.2.

Figure 4.7 shows distributions of $\mathcal{P}$ for mock samples containing $Z/\gamma^* \to \tau\tau$, fakes, and $WW$, where the mock events are pulled randomly from their parent distributions, and the numbers of events are allowed to vary within systematic and statistical errors. As found in the

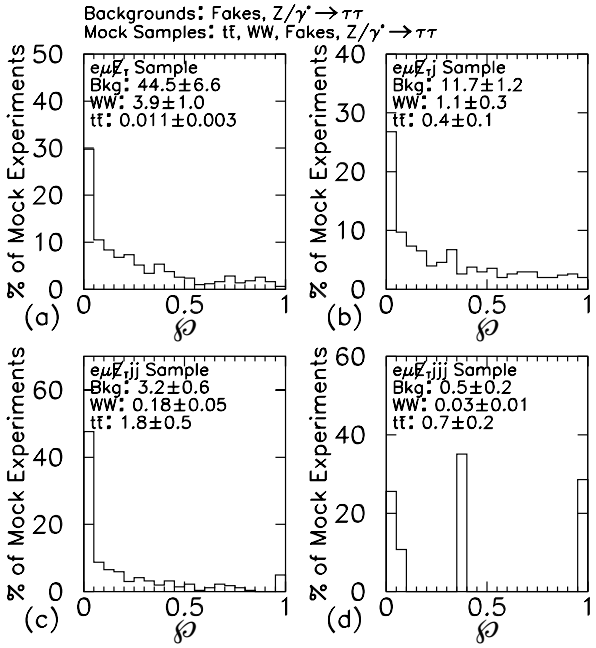



Figure 4.5: Distributions of $\mathcal{P}$ for the four exclusive final states (a) $e\mu\not{E}_T$, (b) $e\mu\not{E}_T j$, (c) $e\mu\not{E}_T jj$, and (d) $e\mu\not{E}_T jjj$. The background includes only $Z/\gamma^* \to \tau\tau$ and fakes. The mock samples for these distributions contain $WW$ and $t\bar{t}$ in addition to $Z/\gamma^* \to \tau\tau$ and fakes. The extent to which these distributions peak at small $\mathcal{P}$ can be taken as a measure of Sleuth's ability to find $WW$ or $t\bar{t}$ if we had no knowledge of either final state. The presence of $WW$ in $e\mu\not{E}_T$ causes the trend toward small values in (a); the presence of $t\bar{t}$ causes the trend toward small values in (c) and (d); and a combination of $WW$ and $t\bar{t}$ causes the signal seen in (b).

previous section, the distributions are uniform in the interval $[0, 1]$, becoming appropriately discretized when the expected number of background events becomes $\lesssim 1$. Figure 4.8 shows distributions of $\mathcal{P}$ when the mock samples contain $t\bar{t}$ in addition to $Z/\gamma^* \to \tau\tau$, fakes, and $WW$. Again, the number of events from each process is allowed to vary within statistical and systematic errors. The distributions in Figs. 4.8(c) and (d) show that we can indeed find $t\bar{t}$ much of the time. Figure 4.9 shows that the distribution of $\tilde{\mathcal{P}}_{[\sigma]}$ is approximately a Gaussian centered at zero of width unity for the case where the background and data both

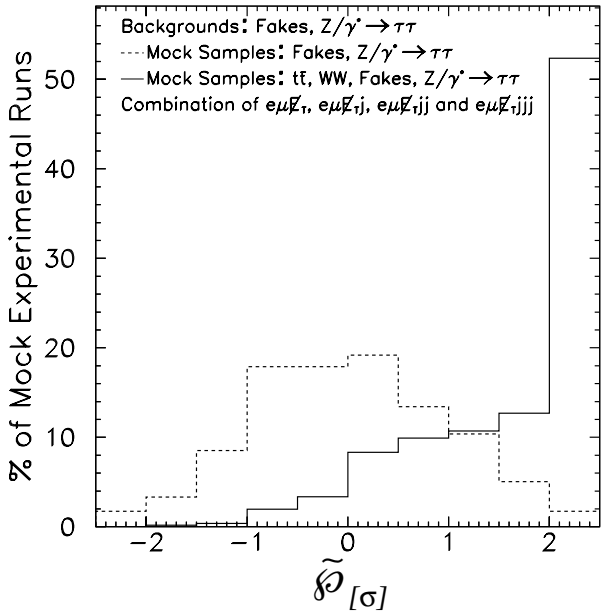



Figure 4.6: Distribution of $\tilde{\mathcal{P}}_{[\sigma]}$ from combining the four exclusive final states $e\mu\not\!\!E_T$, $e\mu\not\!\!E_T j$, $e\mu\not\!\!E_T jj$, and $e\mu\not\!\!E_T jjj$. The background includes only $Z/\gamma^* \to \tau\tau$ and fakes. The mock samples making up the distribution shown as the solid line contain $WW$ and $t\bar{t}$ in addition to $Z/\gamma^* \to \tau\tau$ and fakes, and correspond to Fig. 4.5; the mock samples making up the distribution shown as the dashed line contain only $Z/\gamma^* \to \tau\tau$ and fakes, and correspond to Fig. 4.4. All samples with $\tilde{\mathcal{P}}_{[\sigma]} > 2.0$ appear in the rightmost bin. The fact that $\tilde{\mathcal{P}}_{[\sigma]} > 2.0$ in 50% of the mock samples can be taken as a measure of Sleuth's sensitivity to finding $WW$ and $t\bar{t}$ if we had no knowledge of the existence of the top quark or the possibility of $W$ boson pair production.

contain $Z/\gamma^* \to \tau\tau$, fakes, and $WW$ production, and is peaked in the bin above 2.0 for the same background when the data include $t\bar{t}$.

### 4.5.3 New high $p_T$ physics

We have shown in Secs. 4.5.1 and 4.5.2 that the Sleuth prescription and algorithm correctly finds nothing when there is nothing to be found, while exhibiting sensitivity to the expected presence of $WW$ and $t\bar{t}$ in the $e\mu X$ sample. Sleuth's performance on this "typical"

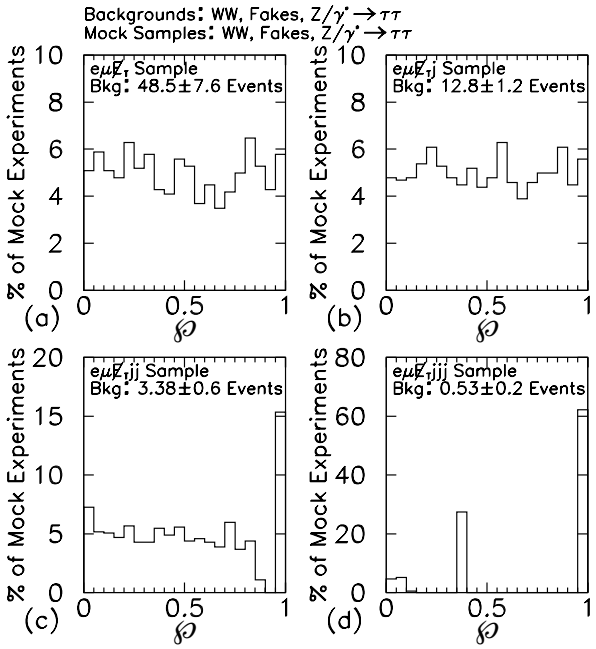



Figure 4.7: Distributions of $\mathcal{P}$ for the four exclusive final states (a) $e\mu\not{E}_T$, (b) $e\mu\not{E}_T j$, (c) $e\mu\not{E}_T jj$, and (d) $e\mu\not{E}_T jjj$. The background includes $Z/\gamma^* \to \tau\tau$, fakes, and $WW$, and the mock samples making up these distributions also contain these three sources. As expected, $\mathcal{P}$ is uniform in the interval $[0,1]$ for those final states in which the expected number of background events $\hat{b} \gg 1$, and shows discrete behavior when $\hat{b} \lesssim 1$.

new physics signal is encouraging, and may be taken as some measure of the sensitivity of this method to the great variety of new high $p_T$ physics that it has been designed to find. Making a more general claim regarding Sleuth's sensitivity to the presence of new physics is difficult, since the sensitivity obviously varies with the characteristics of each candidate theory.

That being said, we can provide a rough estimate of Sleuth's sensitivity to new high $p_T$ physics with the following argument. We have seen that we are sensitive to $WW$ and $t\bar{t}$ pair production in a data sample corresponding to an integrated luminosity of $\approx 100$ pb$^{-1}$. These events tend to fall in the region $p_T^e > 40$ GeV, $\not{E}_T > 40$ GeV, and $\sum' p_T^j > 40$ GeV (if

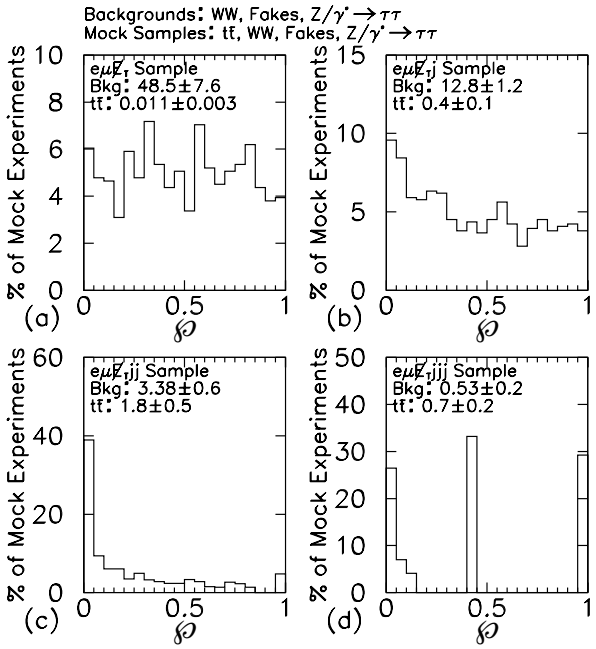



Figure 4.8: Distributions of $\mathcal{P}$ for the four exclusive final states (a) $e\mu\not{E}_T$, (b) $e\mu\not{E}_T j$, (c) $e\mu\not{E}_T jj$, and (d) $e\mu\not{E}_T jjj$. The background includes $Z/\gamma^* \rightarrow \tau\tau$, fakes, and $WW$. The mock samples for these distributions contain $t\bar{t}$ in addition to $Z/\gamma^* \rightarrow \tau\tau$, fakes, and $WW$. The extent to which these distributions peak at small $\mathcal{P}$ can be taken as a measure of Sleuth's sensitivity to finding $t\bar{t}$ if we had no knowledge of the top quark's existence or characteristics. Note that $\mathcal{P}$ is flat in $e\mu\not{E}_T$, where the expected number of top quark events is negligible, peaks slightly toward small values in $e\mu\not{E}_T j$, and shows a marked low peak in $e\mu\not{E}_T jj$ and $e\mu\not{E}_T jjj$.

there are any jets at all). The probability that any true $e\mu X$ event produced will make it into the final sample is about 15% due to the absence of complete hermeticity of the DØ detector, inefficiencies in the detection of electrons and muons, and kinematic acceptance. We can therefore state that we are as sensitive to new high $p_T$ physics as we were to the roughly eight $WW$ and $t\bar{t}$ events in our mock samples if the new physics is distributed relative to all standard model backgrounds as $WW$ and $t\bar{t}$ are distributed relative to backgrounds from $Z/\gamma^* \rightarrow \tau\tau$ and fakes alone, and if its production cross section $\times$ branching ratio into this

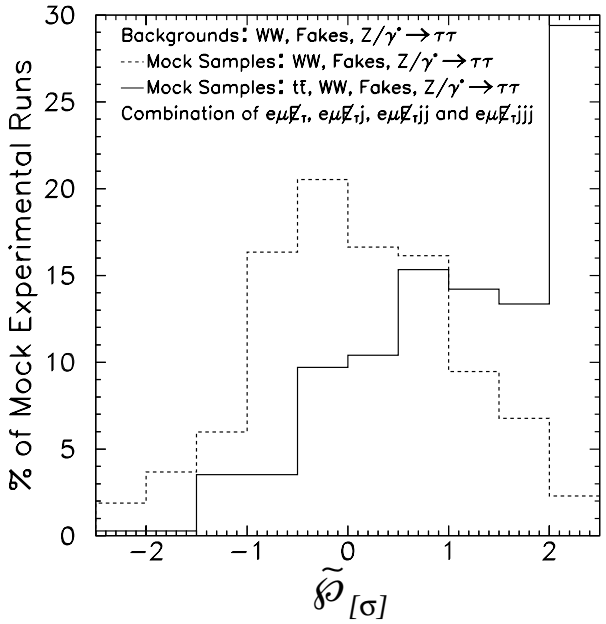



Figure 4.9: Distribution of $\tilde{\mathcal{P}}_{[\sigma]}$ from combining the four exclusive final states $e\mu\not\!\!E_T$, $e\mu\not\!\!E_T j$, $e\mu\not\!\!E_T jj$, and $e\mu\not\!\!E_T jjj$. The background includes $Z/\gamma^* \to \tau\tau$, fakes, and $WW$. The mock samples making up the distribution shown as the solid line contain $t\bar{t}$ in addition to $Z/\gamma^* \to \tau\tau$, fakes, and $WW$, corresponding to Fig. 4.8; the mock samples making up the distribution shown as the dashed line contain only $Z/\gamma^* \to \tau\tau$, fakes, and $WW$, and correspond to Fig. 4.7. All samples with $\tilde{\mathcal{P}}_{[\sigma]} > 2.0$ appear in the rightmost bin. The fact that $\tilde{\mathcal{P}}_{[\sigma]} > 2.0$ in over 25% of the mock samples can be taken as a measure of Sleuth's sensitivity to finding $t\bar{t}$ if we had no knowledge of the top quark's existence or characteristics.

final state is $\gtrsim 8/(0.15 \times 100\ \text{pb}^{-1}) \approx 600$ fb. Readers who are interested in a possible signal with a different relative distribution, or who prefer a more rigorous definition of "sensitivity," should adjust this cross section accordingly.

## 4.6 Results

In the previous section we studied what can be expected when Sleuth is applied to $e\mu X$ mock samples. In this section we confront Sleuth with data. We observe 39 events in

the $e\mu\not\!\!E_T$ final state, 13 events in $e\mu\not\!\!E_T j$, 5 events in $e\mu\not\!\!E_T jj$, and a single event in $e\mu\not\!\!E_T jjj$, in good agreement with the expected background in Table 4.3. We proceed by first removing both $WW$ and $t\bar{t}$ from the background estimates, and next by removing only $t\bar{t}$, to search for evidence of these processes in the data. Finally, we include all standard model processes in the background estimates and search for evidence of new physics.

### 4.6.1 Search for $WW$ and $t\bar{t}$ in data

The results of applying Sleuth to DØ data with only $Z/\gamma^* \to \tau\tau$ and fakes in the background estimate are shown in Table 4.6 and Fig. 4.10. Sleuth finds indications of an excess in the $e\mu\not\!\!E_T$ and $e\mu\not\!\!E_T jj$ states, presumably reflecting the presence of $WW$ and $t\bar{t}$, respectively. The results for the $e\mu\not\!\!E_T j$ and $e\mu\not\!\!E_T jjj$ final states are consistent with the results in Fig. 4.5. Defining $r'$ as the distance of the data point from $(0,0,0)$ in the unit box (transformed so that the background is distributed uniformly in the interval $[0,1]$), the top candidate events from DØ's recent analysis [36] are the three events with largest $r'$ in the $e\mu\not\!\!E_T jj$ sample and the single event in the $e\mu\not\!\!E_T jjj$ sample, shown in Fig. 4.10. The presence of the $WW$ signal can be inferred from the events designated interesting in the $e\mu\not\!\!E_T$ final state.

| Data set | $\mathcal{P}$ |
|---|---|
| $e\mu\not\!\!E_T$ | 0.008 |
| $e\mu\not\!\!E_T j$ | 0.34 |
| $e\mu\not\!\!E_T jj$ | 0.01 |
| $e\mu\not\!\!E_T jjj$ | 0.38 |
| $\tilde{\mathcal{P}}$ | 0.03 |

Table 4.6: Summary of results on the $e\mu\not\!\!E_T$, $e\mu\not\!\!E_T j$, $e\mu\not\!\!E_T jj$, and $e\mu\not\!\!E_T jjj$ channels when $WW$ and $t\bar{t}$ are not included in the background. Sleuth identifies a region of excess in the $e\mu\not\!\!E_T$ and $e\mu\not\!\!E_T jj$ final states, presumably indicating the presence of $WW$ and $t\bar{t}$ in the data. In units of standard deviation, $\tilde{\mathcal{P}}_{[\sigma]} = 1.9$.

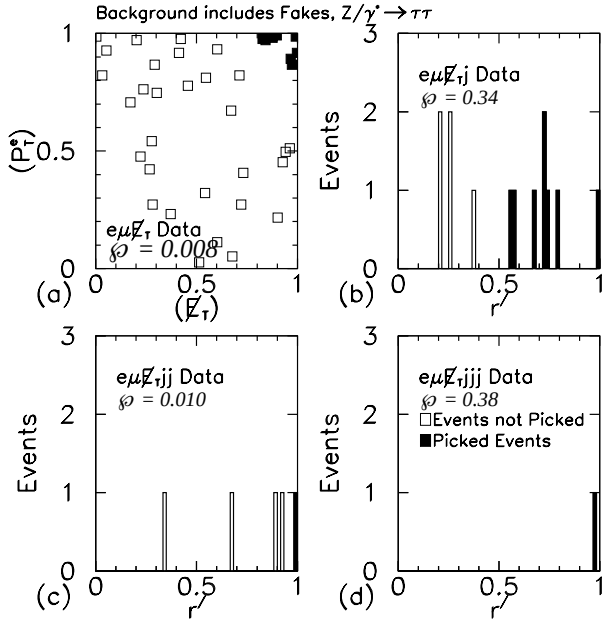



Figure 4.10: Positions of data points following the transformation of the background from fake and $Z/\gamma^*$ sources in the space of variables in Table 4.1 to a uniform distribution in the unit box. The darkened points define the region Sleuth found most interesting. The axes of the unit box in (a) are suggestively labeled $(p_T^e)$ and $(\not\!\!E_T)$; each is a function of both $p_T^e$ and $\not\!\!E_T$, but $(p_T^e)$ depends more strongly on $p_T^e$, while $(\not\!\!E_T)$ more closely tracks $\not\!\!E_T$. $r'$ is the distance of the data point from $(0,0,0)$ (the "lower left-hand corner" of the unit box), transformed so that the background is distributed uniformly in the interval $[0,1]$. The interesting regions in the $e\mu\not\!\!E_T$ and $e\mu\not\!\!E_T jj$ samples presumably indicate the presence of $WW$ signal in $e\mu\not\!\!E_T$ and of $t\bar{t}$ signal in $e\mu\not\!\!E_T jj$. We find $\tilde{\mathcal{P}} = 0.03$ ($\tilde{\mathcal{P}}_{[\sigma]} = 1.9$).

### 4.6.2 Search for $t\bar{t}$ in data

The results of applying Sleuth to the data with $Z/\gamma^* \to \tau\tau$, fakes, and $WW$ included in the background estimate are shown in Table 4.7 and Fig. 4.11. Sleuth finds an indication of excess in the $e\mu\not\!\!E_T jj$ events, presumably indicating the presence of $t\bar{t}$. The results for the $e\mu\not\!\!E_T$, $e\mu\not\!\!E_T j$, and $e\mu\not\!\!E_T jjj$ final states are consistent with the results in Fig. 4.8. The $t\bar{t}$ candidates from DØ's recent analysis [36] are the three events with largest

$r'$ in the $e\mu\not\!\!E_T jj$ sample, and the single event in the $e\mu\not\!\!E_T jjj$ sample, shown in Fig. 4.11.

| Data set | $\mathcal{P}$ |
|---|---|
| $e\mu\not\!\!E_T$ | 0.16 |
| $e\mu\not\!\!E_T j$ | 0.45 |
| $e\mu\not\!\!E_T jj$ | 0.03 |
| $e\mu\not\!\!E_T jjj$ | 0.41 |
| $\tilde{\mathcal{P}}$ | 0.11 |

Table 4.7: Summary of results on the $e\mu\not\!\!E_T$, $e\mu\not\!\!E_T j$, $e\mu\not\!\!E_T jj$, and $e\mu\not\!\!E_T jjj$ channels when $t\bar{t}$ production is not included in the background. Sleuth identifies a region of excess in the $e\mu\not\!\!E_T jj$ final state, presumably indicating the presence of $t\bar{t}$ in the data. In units of standard deviation, $\tilde{\mathcal{P}}_{[\sigma]} = 1.2$.

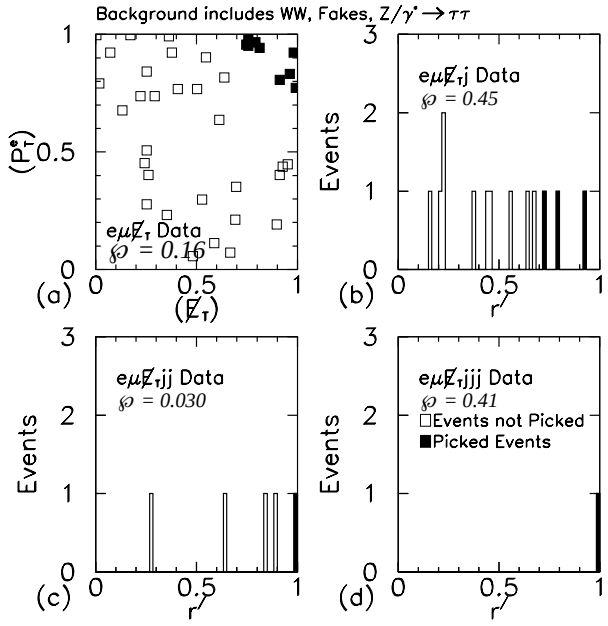



Figure 4.11: Positions of data points following the transformation of the background from the three sources $Z/\gamma^* \to \tau\tau$, fakes, and $WW$ in the space of variables in Table 4.1 to a uniform distribution in the unit box. The darkened points define the region Sleuth found most interesting. The interesting region in the $e\mu\not\!\!E_T jj$ sample presumably indicates the presence of $t\bar{t}$. We find $\tilde{\mathcal{P}} = 0.11$ ($\tilde{\mathcal{P}}_{[\sigma]} = 1.2$).

A comparison of this result with one obtained using a dedicated top quark search illustrates an important difference between Sleuth's result and the result from a dedicated

search. DØ announced its discovery of the top quark [37] in 1995 with 50 $\mathrm{pb}^{-1}$ of integrated luminosity upon observing 17 events with an expected background of $3.8 \pm 0.6$ events, a $4.6\sigma$ "effect," in the combined dilepton and single-lepton decay channels. In the $e\mu$ channel alone, two events were seen with an expected background of $0.12 \pm 0.03$ events. The probability of $0.12 \pm 0.03$ events fluctuating up to or above two events is 0.007, corresponding to a $2.5\sigma$ "effect." In a subsequent measurement of the top quark cross section [24], three candidate events were seen with an expected background of $0.21 \pm 0.16$, an excess corresponding to a $2.75\sigma$ "effect." Using Sleuth, we find $\mathcal{P} = 0.03$ in the $e\mu \not\!\!E_T jj$ sample, a $1.9\sigma$ "effect," when complete ignorance of the top quark is feigned. When we take into account the fact that we have also searched in all of the final states listed in Table 4.3, we find $\tilde{\mathcal{P}} = 0.11$, a $1.2\sigma$ "effect." The difference between the $2.75\sigma$ "effect" seen with a dedicated top quark search and the $1.2\sigma$ "effect" that Sleuth reports in $e\mu X$ lies partially in the fact that Sleuth is not optimized for $t\bar{t}$; and partially in the careful accounting of the many new physics signatures that Sleuth considered in addition to $t\bar{t}$ production, and the correspondingly many new physics signals that Sleuth might have discovered.

### 4.6.3 Search for physics beyond the standard model

In this section we present Sleuth's results for the case in which all standard model and instrumental backgrounds are considered in the background estimate: $Z/\gamma^* \to \tau\tau$, fakes, $WW$, and $t\bar{t}$. The results are shown in Table 4.8 and Fig. 4.12. We observe excellent agreement with the standard model. We conclude that these data contain no evidence of new physics at high $p_T$, and calculate that a fraction $\tilde{\mathcal{P}} = 0.72$ of hypothetical similar experimental runs would produce a more significant excess than any observed in these data.

Recall that we are sensitive to new high $p_T$ physics with production cross section $\times$ branching ratio into this final state as described in Sec. 4.5.3.

| Data set | $\mathcal{P}$ |
|---|---|
| $e\mu\not\!\!E_T$ | 0.14 |
| $e\mu\not\!\!E_T j$ | 0.45 |
| $e\mu\not\!\!E_T jj$ | 0.31 |
| $e\mu\not\!\!E_T jjj$ | 0.71 |
| $\tilde{\mathcal{P}}$ | 0.72 |

Table 4.8: Summary of results on all final states within $e\mu X$ when all standard model backgrounds are included. The unpopulated final states (listed in Table 4.4) have $\mathcal{P} = 1.0$; these final states are included in the calculation of $\tilde{\mathcal{P}}$. We observe no evidence for the presence of new high $p_T$ physics.

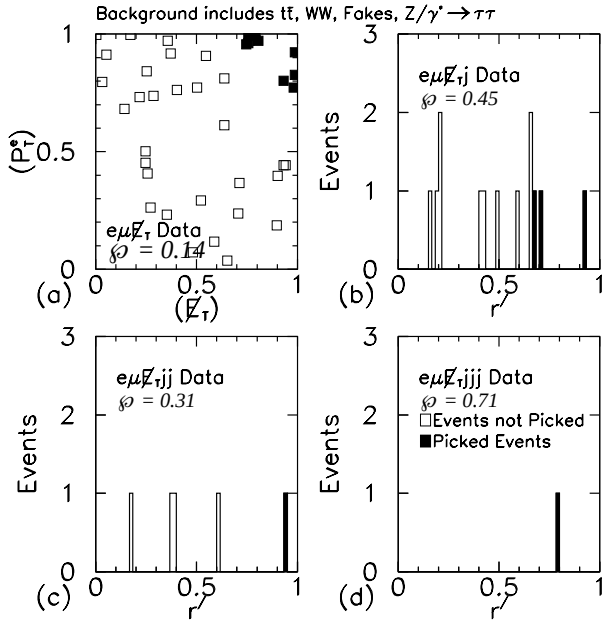



Figure 4.12: Positions of the data points following the transformation of the background from $Z/\gamma^* \to \tau\tau$, fakes, $WW$, and $t\bar{t}$ sources in the space of variables in Table 4.1 to a uniform distribution in the unit box. The darkened points define the region that Sleuth chose. We find $\tilde{\mathcal{P}} = 0.72$, and distributions that are all roughly uniform and consistent with background. No evidence for new high $p_T$ physics is observed.

## 4.7 Conclusions

We have developed a quasi-model-independent technique for searching for the physics responsible for stabilizing electroweak symmetry breaking. Our prescription involves the definition of final states and the construction of a rule that identifies a set of relevant variables for any particular final state. An algorithm (Sleuth) systematically searches for regions of excess in those variables, and quantifies the significance of any observed excess. This technique is sufficiently *a priori* that it allows an *ex post facto*, quantitative measure of the degree to which curious events are interesting. After demonstrating the sensitivity of the method, we have applied it to the set of events in the semi-inclusive channel $e\mu X$. Removing $WW$ and $t\bar{t}$ from the calculated background, we find indications of these signals in the data. Including these background channels, we find that these data contain no evidence of new physics at high $p_T$. A fraction $\tilde{\mathcal{P}} = 0.72$ of hypothetical similar experimental runs would produce a more significant excess than any observed in these data.

# Chapter 5

# Further analysis of DØ data

In this chapter we apply Sleuth to many exclusive final states in $\approx 100$ pb$^{-1}$ of $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV collected by the DØ experiment during 1992–1996 at the Fermilab Tevatron. Along the way we demonstrate sensitivity to a variety of models predicting new phenomena at the electroweak scale. This chapter also appears as Ref. [38].

## 5.1 Introduction

The standard model is an impressive theory, accurately predicting, or at least accommodating, the results of nearly all particle physics experiments to date. It is generally accepted, however, that there is good reason to believe that hints of new physics are likely to appear at or around the energy scale of 1 TeV.

Electroweak symmetry is broken in the standard model when a scalar field (the Higgs field) acquires a vacuum expectation value. Since the quantum corrections to the renormalized mass squared of a scalar field grow as the square of the heaviest energy scale in

the theory (naively the Planck scale, of order $10^{19}$ GeV), and since the mass of the standard model Higgs boson is of the order of a few hundred GeV, a fine-tuning at the level of one part in $10^{16}$ appears to be required to keep the Higgs mass at the electroweak scale.

Two of the most popular solutions to this hierarchy problem are supersymmetry [2] and strong dynamics [3]. In their most general form these classes of models are capable of "predicting" any of many different signatures, depending upon the values that are chosen for the model's parameters. Previous searches for these signals have fought to strike a balance between the simultaneous desires to assume as little as possible about the signal and yet achieve "optimal sensitivity" to more specific signals. These are necessarily contradictory objectives.

Many new phenomena have been predicted in addition to those resulting from these proposed solutions to the hierarchy problem. Among them are leptoquarks, proposed in an attempt to explain the relationship between quarks and leptons in the standard model and appearing in many grand unified theories; composite quarks and leptons, in case the "fundamental" particles of the standard model turn out not to be fundamental at scales $\lesssim 10^{-18}$ meters; a fourth generation of quarks or leptons; excited quarks and leptons, in analogy to the excited states of hadrons observed at much lower energies; new heavy gauge bosons, arising from additional gauge symmetries in models extending the $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ of the standard model; and many others. Of course, Nature may have other ideas. The CDF and DØ collaborations have performed many searches on the data collected during Run I of the Fermilab Tevatron, but have we looked in all the right places?

Figure 5.1 diagrams the final states that are populated (i.e., that contain events) in

the DØ Run I data. In this article we undertake a systematic and quasi-model-independent analysis of many of these exclusive final states, in the hope of finding some evidence for physics beyond the standard model.

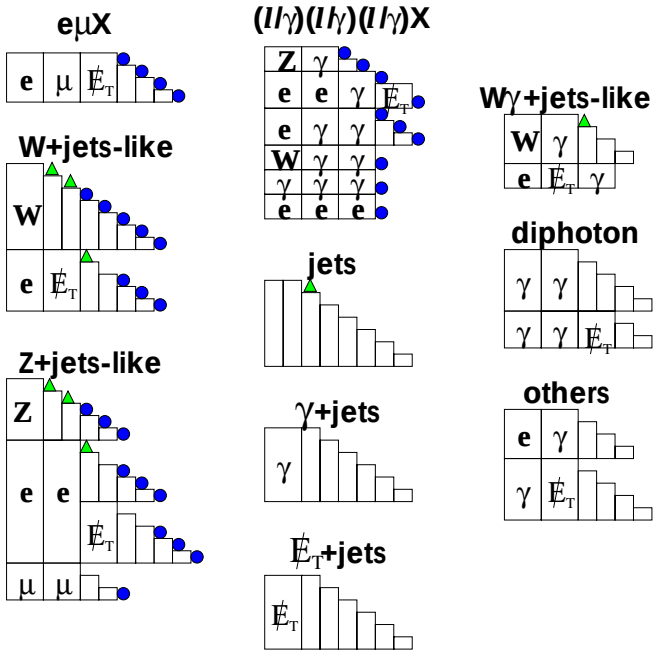



Figure 5.1: A diagram showing the final states populated in DØ data in Run I. Each row in a given column represents the final state defined by the objects in that row; to reduce clutter, jets are represented by an empty rectangle, rather than by a rectangle containing "$j$." Reading down the left column are the final states $e\mu\not\!\!E_T$, $e\mu\not\!\!E_T j$, $e\mu\not\!\!E_T\, 2j$, $e\mu\not\!\!E_T\, 3j$, $W$, $Wj$, $W\, 2j$, and so on. Rows with triangles (e.g., $W$ and $Wj$) indicate final states analyzed previously by DØ in a manner similar to the strategy we use here, but without using Sleuth; rows with filled circles indicate final states analyzed with Sleuth. The remaining rows show populated final states not discussed in this article.

In Refs. [16] we introduced a quasi-model-independent search strategy ("Sleuth"), designed to systematically search for new high $p_T$ physics at any collider experiment sensitive to physics at the electroweak scale, and applied it to all events in the DØ data containing one or more electrons and one or more muons ($e\mu X$). Considering again Fig. 5.1, we see

that the number of final states within $e\mu X$ is a small fraction of the total number of final states populated by the DØ Run I data. If there is indeed a signal in the data, our chances of finding it grow proportionally to the number of final states considered.

In this article we present a systematic analysis of thirty-two of these final states — those marked with a filled circle in Fig. 5.1. A large number of unpopulated final states with additional objects are analyzed implicitly; e.g., $ee\mu\not\!\!E_T$ and $e\mu\not\!\!E_T\gamma$ are among a host of unpopulated final states analyzed within the context of $e\mu X$.

The notation we use to label final states may require explanation. Electrons and muons are confidently identified with the DØ detector on an event-by-event basis, but taus are not; $\ell$ and the word "lepton" will therefore denote an electron ($e$) or a muon ($\mu$) in this article. We use the composite symbol ($\ell/\gamma$) to denote an electron, muon, or photon. $X$ will denote zero or more objects, and ($nj$) will denote zero or more jets. Any inclusive final state [i.e., any state whose label includes the symbol $X$ or ($nj$)] will refer to the physics objects actually reconstructed in the detector. Thus $ee\,2j(nj)$ denotes the set of all events with two electrons and two or more jets. Any exclusive final state is defined according to the rules in Sec. 4.2. For example, since these rules include a prescription for identifying a $Z$ boson from two charged leptons of the same flavor, we use $ee\,2j$ to denote the set of all events with two electrons and two jets having $m_{ee}$ substantially different from $M_Z$, while events with two electrons and two jets having $m_{ee} \approx M_Z$ fall within the final state $Z\,2j$.

We begin in Sec. 5.2 by providing a brief review of the Sleuth search strategy and algorithm, and describing a slight change from the method advanced in Ref. [16]. In Sec. 5.4 we discuss eight final states already analyzed by DØ in a manner similar to Sleuth, and

motivate the final states to be considered in this article. In Sec. 5.5 we describe the analysis of the $W$+jets-like final states — events containing a single lepton, missing transverse energy ($\not\!\!E_T$), and two or more jets. In Sec. 5.6 we present the analysis of the $Z$+jets-like final states — events containing two leptons and two or more jets. In Sec. 5.7 we analyze the final states containing several objects, at least three of which are either an electron, muon, or photon $[(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X]$. In Sec. 5.8 we present the combined results of all of these final states. Section 5.9 contains our conclusions.

## 5.2 New region criteria

Use of Sleuth requires the specification of criteria that define the regions that Sleuth is allowed to consider. In the analysis of $e\mu X$ we imposed two criteria: *AntiCornerSphere* ($c^A$), which restricts the allowed region to be defined by those data points greater than a distance $r$ from the origin of the unit box, where $r$ is allowed to vary; and *Isolation* ($c^I$), which requires that there exist no data points outside the region that are closer than $\xi$ to any data point inside the region, where $\xi = 1/(4N_{\rm data}^{1/d})$ is a characteristic distance between the $N_{\rm data}$ data points in the $d$-dimensional unit box.

In this article we use *Hyperplanes* ($c^H$), a criterion defined but not used in Ref. [16]. Hyperplanes is less restrictive than AntiCornerSphere, in the sense that any region satisfying AntiCornerSphere will also satisfy Hyperplanes. Hyperplanes has the advantage of allowing regions that lie in the high tails of only a subset of the variables considered. A region $R$ in a $d$-dimensional unit box is said to satisfy Hyperplanes if, for each data point $p$ inside $R$, one can draw a $(d-1)$-dimensional hyperplane through $p$ such that all data points on the side

of the hyperplane containing the point $\vec{1}$ (the "upper right-hand corner of the unit box") are inside $R$. An example of a region satisfying Hyperplanes is shown in Fig. 5.2.

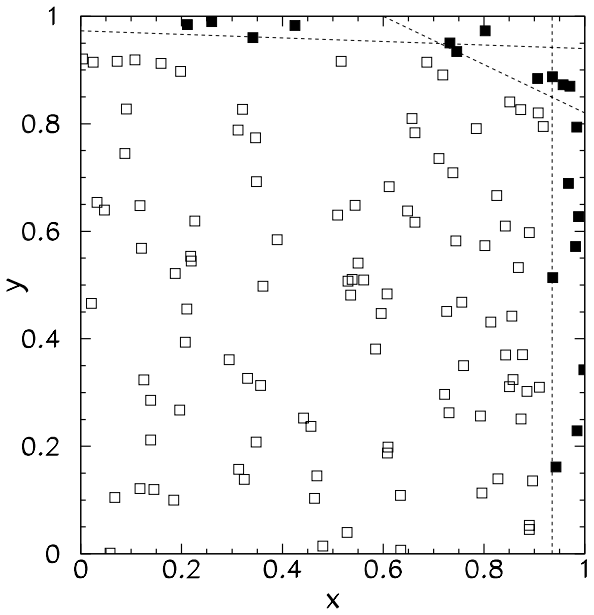



Figure 5.2: An example of a region satisfying Hyperplanes. The boundary of the figure is the unit box; open squares represent data points outside the region $R$; filled squares represent data points inside the region $R$. The three dashed lines indicate hyperplanes $h_i$ (which are lines in this two-dimensional case) that can be drawn through the points at $(x, y)_i = (0.34, 0.96)$, $(0.74, 0.95)$, and $(0.935, 0.515)$ with the property that all of the data points "up and to the right" of $h_i$ are inside $R$.

We continue this boolean criterion to the unit interval $[0, 1]$ in order to ensure the continuity of the final result under small changes in the background estimate. For each data point $i$ inside the candidate region $R$ and each hyperplane $h_i$ through $i$, we define $d_{jh_i}$ to be the distance between a data point $j$ lying outside $R$ and the hyperplane $h_i$. This quantity is taken to be positive if $j$ and the point $\vec{1}$ are on the same side of $h_i$, and negative otherwise.

Letting

$$\theta(x) = \begin{cases} 0 & ,x < 0 \\ x & ,0 \leq x \leq 1 \\ 1 & ,1 < x \end{cases} , \tag{5.1}$$

we define

$$c_R^H = \prod_{i \in R} \theta(1 - \min_{h_i} \max_{j \notin R} d_{jh_i}/\xi). \tag{5.2}$$

Note that $c_R^H$ reduces to the boolean operator of the preceding paragraph in the limit $\xi \to 0$.

We also impose the criterion *Connectivity* ($c^C$) to ensure connected regions, and the criterion *ReasonableSize* ($c^R$) to limit the size of the regions we consider to that expected for a typical signal and to reduce the computational cost of finding the most interesting region. A region $R$ is said to satisfy Connectivity if, given any two points $a$ and $b$ within $R$, there exists a list of points $p_1 = a, p_2, \ldots, p_{n-1}, p_n = b$ such that all the $p_i$ are in $R$, and the 1-region about $p_{i+1}$ shares a border with the 1-region about $p_i$. A region is said to satisfy ReasonableSize if it contains fewer than 50 data points. These criteria are summarized in Table 5.1.

In Ref. [16] we demonstrated Sleuth's ability to find $t\bar{t}$ in the $e\mu X$ final states using the criteria $c^A c^I$. Figure 5.3 shows that the combination $c^H c^C c^R$ (solid) performs similarly to these criteria (dashed) in this test.

## 5.3 Examples of signals that might appear

In this section we provide a few examples of signals that might be discovered in the course of this analysis. This discussion is provided to give the reader a taste of the many

| Symbol | Name | A region satisfies this criterion if |
|---|---|---|
| $c^A$ | AntiCornerSphere | One can draw a sphere centered on the origin of the unit box containing all data events outside the region and no data events inside the region. |
| $c^I$ | Isolation | There exist no data points outside the region that are closer than $\xi$ to any data point inside the region. |
| $c^H$ | Hyperplanes | For each data point $p$ inside $R$, one can draw a $(d-1)$-dimensional hyperplane through $p$ such that all data points on the side of the hyperplane containing the point $\vec{1}$ are inside $R$. |
| $c^C$ | Connectivity | Given any two points $a$ and $b$ within the region, there exists a list of points $p_1 = a, p_2, \ldots, p_{n-1}, p_n = b$ such that all the $p_i$ are in the region and $p_{i+1}$ is a neighbor of $p_i$. |
| $c^R$ | ReasonableSize | The region contains fewer than 50 data points. |

Table 5.1: Summary of the region criteria imposed in our previous analysis of $e\mu X$ (above middle line) and those imposed in the analyses described in this article (below middle line). $\xi = 1/(4N_{\rm data}^{1/d})$ is a characteristic distance between the $N_{\rm data}$ data points in the $d$-dimensional unit box.

processes that might appear in the final states to be analyzed, and is by no means intended to be complete. The possibility that the correct answer is "none of the following" is one of the strongest motivations for pursuing a quasi-model-independent search.

### 5.3.1 $e\mu X$

In supersymmetric models (denoting the supersymmetric particles as in Ref. [2]), the process $q\bar{q} \to Z/\gamma^* \to \tilde{\chi}_1^{\pm}\tilde{\chi}_1^{\mp} \to e\mu\nu\nu\tilde{\chi}_1^0\tilde{\chi}_1^0$ can produce events appearing in the $e\mu\not{E}_T$ final state. More generally, any process involving the production of two charginos has the potential for producing a final state containing an electron, a muon, and $\not{E}_T$. This final state may also be reached through the leptonic decays of two taus, obtained (for example)

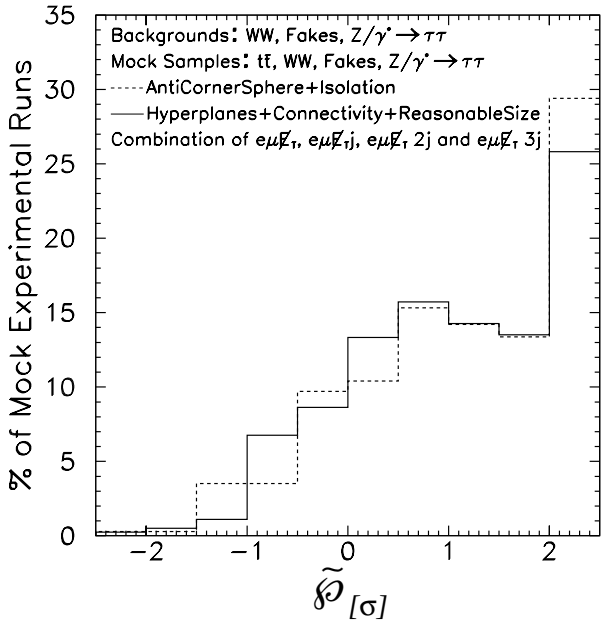



Figure 5.3: Distribution of $\tilde{\mathcal{P}}_{[\sigma]}$ in an ensemble of mock experimental runs on the four exclusive final states $e\mu\not{E}_T$, $e\mu\not{E}_T j$, $e\mu\not{E}_T\, 2j$, and $e\mu\not{E}_T\, 3j$. The background includes $Z/\gamma^* \rightarrow \tau\tau$, fakes, and $WW$. The mock samples making up the distributions contain $t\bar{t}$ in addition to $Z/\gamma^* \rightarrow \tau\tau$, fakes, and $WW$.

from the production of two $\tilde{\tau}$ particles that each decay to $\tau\tilde{\chi}_1^0$, or from the production of a heavy $Z$-like object that couples strongly to the third generation. An anomalous correction to the standard model $WW\gamma$ vertex or anomalies involving the top quark could also appear in these final states.

### 5.3.2 Final states already considered

A sampling of the types of new physics that might appear in a few of the final states described in Sec. 5.4.1 is provided here.

$2j$. The dijet final state could contain hints of a massive object (such as an additional neutral gauge boson) produced through $q\bar{q}$ annihilation and decaying back into $q\bar{q}$. It

could also contain indications that quarks are in fact composite objects, interacting through terms in an effective Lagrangian of the form $\frac{c}{\Lambda^2} q\bar{q}q'\bar{q}'$, where $\Lambda \gtrsim 1$ TeV is a compositeness scale and $c$ is a constant of order unity.

$e\not{E}_T$. Models containing symmetry groups larger than the $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ group of the standard model often contain an additional SU(2) group, suggesting the existence of a heavy $W$-like gauge boson ($W'$) that would decay into the $e\not{E}_T$ final state, with the transverse mass of the electron and neutrino greater than that expected for the standard model $W$. Production of $\tilde{\ell}\tilde{\nu}$ decaying to $\ell\tilde{\chi}_1^0\nu\tilde{\chi}_1^0$ could also produce events in this final state, as could production of $\tilde{\chi}_1^\pm\tilde{\chi}_2^0$ decaying to $\ell\nu\tilde{\chi}_1^0\nu\nu\tilde{\chi}_1^0$.

$ee$. If both quarks and leptons are composite objects, there will be four-fermion contact terms of the form $\frac{c}{\Lambda^2} q\bar{q}\ell^+\ell^-$ in addition to the $\frac{c}{\Lambda^2} q\bar{q}q'\bar{q}'$ terms postulated in the discussion of the $2j$ final state above. Such an interaction would produce events with large transverse momentum, opposite-sign leptons, and should appear in the $ee$ and $\mu\mu$ final states. Some models that employ a strong dynamics to break electroweak symmetry predict the existence of composite "techni-"particles, such as the $\omega_T$, $\rho_T$, and $\pi_T$, that are analogous to the composite $\omega$, $\rho$, and $\pi$ mesons that arise from confinement in QCD. The technirho ($\rho_T$) and techniomega ($\omega_T$), if produced, will decay into an $\ell^+\ell^-$ pair if their preferred decay mode to technipions ($\pi_T$) is kinematically forbidden. Such events will appear as a bump in the tail of the $ee$ invariant mass distribution and as an excess in the tail of the electron $p_T$ distribution. Models containing symmetry groups larger than that of the standard model typically contain a heavy neutral boson (generically called a $Z'$) in addition to the $W'$ boson described above. If this $Z'$ boson couples to leptons, the process $q\bar{q} \to Z' \to \ell\ell$ could produce

a signature similar to that expected from the decay of a $\rho_T$ or $\omega_T$.

### 5.3.3 $W$+jets-like final states

A variety of new signals have been predicted that would manifest themselves in the $W$+jets-like final states — those final states containing events with a single lepton, missing transverse energy, and zero or more jets. A plethora of supersymmetric signatures could appear in these states. A chargino and neutralino, produced from $q\bar{q}$ through an $s$-channel $W$ boson, can proceed to decay as $\tilde{\chi}_1^{\pm} \to \ell\nu\tilde{\chi}_1^0$ and $\tilde{\chi}_2^0 \to q\bar{q}\tilde{\chi}_1^0$, leaving an event that will be partitioned into either the $e\not{E}_T\,2j$ or $W\,2j$ final state. Pair production of top squarks, with $\tilde{t} \to b\tilde{\chi}_1^{\pm}$ and subsequent decays of the charginos to $e\nu\tilde{\chi}_1^0$ and $qq'\tilde{\chi}_1^0$, will produce events likely to fall into the $e\not{E}_T\,4j$ or $W\,4j$ final states. Depending upon the particular model, even gluino decays can give rise to leptons. Events with gluinos that are pair-produced and decay, one into $qq'\tilde{\chi}_1^{\pm}$ and the other into $q\bar{q}\tilde{\chi}_1^0$, can also find themselves in the $e\not{E}_T\,4j$ or $W\,4j$ final state. Other possible decays of the supersymmetric spectrum allow many more signals that might populate these final states.

The decay of a $\rho_T^+$, produced by $q\bar{q}$ annihilation, can produce a $W^+$ boson and a $\pi_T^0$, which in turn may decay to $b\bar{b}$ or $gg$. Such an event should appear in the high tails of the $p_T^W$ and $\sum' p_T^j$ distributions in our analysis of the $W\,2j$ final state if the technipion is sufficiently massive. The same final state may also be reached by the process $q\bar{q} \to \rho_T^0 \to W^-\pi_T^+ \to \ell^-\nu c\bar{b}$. A neutral color-octet technirho ($\rho_{T8}^0$) produced by $q\bar{q}$ annihilation can decay to two technipions carrying both color and lepton quantum numbers ($\pi_{\rm LQ}$), each of which in turn decays preferentially into a massive quark and a massive lepton. If the technipion is heavier than the top quark then the decay $\pi_{\rm LQ} \to t\tau$ or $t\nu_\tau$ is kinematically

allowed. Appropriate decays of the $W$ bosons from the two top quarks leave the event containing one high transverse momentum lepton, substantial $\not\!\!E_T$, and several energetic jets.

The standard model contains three generations of quarks and leptons, but there appears to be no fundamental reason that Nature should choose to stop at three. A massive charge $-1/3$ fourth-generation quark ($b'$), which could be pair-produced at the Tevatron, would be apt to decay weakly into a $W$ boson and a top quark. Events in which one of the four $W$ bosons then decays leptonically will result in a final state containing one lepton, substantial missing transverse energy, and many jets.

Leptoquarks, a consequence of many theories that attempt to explain the peculiar symmetry between quarks and leptons in the standard model, could also be pair-produced at the Tevatron. If their branching ratio to charged leptons $\beta = 0.5$ then the pair will decay to $\ell\nu q\bar{q}$ 50% of the time, resulting in events that will be classified either as $e\not\!\!E_T\, 2j$ or $W\, 2j$.

Models invoking two Higgs doublets predict a charged Higgs that may appear in occasional decays of the top quark. In such models a top quark pair, produced by $q\bar{q}$ or $gg$ annihilation, can decay into $H^+bW^-\bar{b}$. Depending upon the mass of the charged Higgs particle, it may decay into $W^+b\bar{b}$, $c\bar{s}$, or $\tau^+\nu$. Appropriate decay of the $W$ boson(s) in the event will result in the event populating one of the $W\, 2j(nj)$ final states. Other predictions abound.

### 5.3.4 $Z$+jets-like final states

Just as in the $W$+jets-like final states, there are a host of theoretical possibilities for new physics in the $Z$+jets-like final states. Although some of these processes involve the production of two same-flavor, opposite-sign leptons via the production of a standard model

$Z$ boson, many others involve particles that decay to leptons of different flavor, or with the same charge. These different possibilities typically are partitioned into different final states according to our prescription: events that contain leptons of different flavor (those within $e\mu X$) are considered in Sec. 4.4; events containing leptons of similar charge (e.g., an $e^+e^+\,2j$ event) would in principle be partitioned into different final states than events containing leptons of opposite charge (e.g., an $e^+e^-\,2j$ event) if DØ distinguished electron charge; and events in which the leptons have an invariant mass consistent with the hypothesis that they are the decay products of a $Z$ boson are partitioned into different final states than those with a dilepton invariant mass outside the $Z$ boson mass window.

Models containing supersymmetry and imposing conservation of $R$-parity predict signatures containing substantial missing transverse energy. Such events might therefore populate the $ee\not\!\!E_T\,2j(nj)$ or $\mu\mu\not\!\!E_T\,2j(nj)$ channels. Final state leptons may be obtained in supersymmetric models from the decays of neutralinos (which can produce two same-flavor, oppositely-charged leptons), or charginos or sleptons (which decay into a single charged lepton and missing transverse energy). The process $q\bar{q}' \to W^* \to \tilde{\chi}_1^\pm\tilde{\chi}_2^0$, with subsequent decay of the chargino to $qq'\tilde{\chi}_1^0$ and the neutralino to $\ell^+\ell^-\tilde{\chi}_1^0$, results in an event with two same-flavor, opposite-sign leptons, two jets, and missing transverse energy, and would appear in our $ee\not\!\!E_T\,2j$ or $\mu\mu\,2j$ final states. Events in which gluinos are pair-produced and decay via $\tilde{g} \to qq'\tilde{\chi}_1^\pm$ will appear in the $ee\not\!\!E_T\,4j$ and $\mu\mu\,4j$ final states when the gaugino decays to $\ell\nu\tilde{\chi}_1^0$. Pair production of scalar top quarks ($q\bar{q}/gg \to g \to \tilde{t}\tilde{t}^*$) that decay via $\tilde{t} \to b\tilde{\chi}_1^\pm$ and $\tilde{\chi}_1^\pm \to \ell\nu\tilde{\chi}_1^0$ again produce events that populate the $ee\not\!\!E_T\,2j$ and $\mu\mu\,2j$ final states, in addition to the $e\mu\not\!\!E_T\,2j$ final states already considered. If $R$-parity is violated, then

supersymmetric signals could populate final states without missing transverse energy. Pair production of gluinos decaying to $\bar{c}\tilde{c}_L$ could produce events that land in the $ee\,4j$ final state if the $R$-parity-violating decay $\tilde{c}_L \to e^+ d$ is allowed.

Color-octet models predict the existence of a color-octet technirho, which can decay to $\pi_{\rm LQ}\pi_{\rm LQ}$. These technipions decay preferentially to massive particles, like the color-singlet $\pi_T$, but their decay products will carry both color and lepton quantum numbers. Events in which each $\pi_{\rm LQ}$ decays to a $b$ quark and a $\tau$ lepton will populate $ee\not\!\!E_T\,2j$ and $\mu\mu\,2j$ final states, among others. Leptoquarks motivated by grand unified theories could be pair-produced at the Tevatron via $q\bar{q} \to Z/\gamma^* \to LQ\overline{LQ}$, and might populate the final states $ee\,2j$ and $\mu\mu\,2j$. Again, other examples abound.

### 5.3.5 $(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$

There are few standard model processes that produce events in which the sum of the numbers of electrons, muons, and photons is $\geq 3$. The $(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$ final states are therefore quite clean, and the presence of even a few events in any of these states could provide a strong indication of new physics.

Supersymmetric models predict a variety of possible signatures in these states. Those models in which $R$-parity is conserved produce events with missing transverse energy in addition to three $(\ell/\gamma)$ objects. Models in which the lightest neutralino ($\tilde{\chi}_1^0$) is the lightest supersymmetric particle (LSP) usually produce final states without photons. This case occurs for many models in which the supersymmetry is broken in a hidden sector and communicated to the visible sector through gravitational forces (gravity-mediated supersymmetry breaking). Models in which the gravitino ($\tilde{G}$) is the LSP often produce final

states with photons from the decay of the next-to-lightest supersymmetric particle via (for example) $\tilde{\chi}_1^0 \to \gamma\tilde{G}$. This case, in turn, obtains for many models in which the breaking of the supersymmetry is mediated by gauge fields (gauge-mediated supersymmetry breaking). For example, the production of a chargino and neutralino through $q\bar{q}$ annihilation into a virtual $W$ boson can produce events in these final states through the decays $\tilde{\chi}_1^\pm \to \ell\nu\tilde{\chi}_1^0$ and $\tilde{\chi}_2^0 \to \ell\ell\tilde{\chi}_1^0$ if the lightest neutralino is the LSP, or through the decays $\tilde{\chi}_1^\pm \to e\nu\tilde{\chi}_1^0$, $\tilde{\chi}_2^0 \to q\bar{q}\tilde{\chi}_1^0$, and $\tilde{\chi}_1^0 \to \gamma\tilde{G}$ if the gravitino is the LSP.

Charginos can be pair-produced in the reaction $q\bar{q} \to Z/\gamma^* \to \tilde{\chi}_1^\pm\tilde{\chi}_1^\mp$. If they decay to $e\nu\tilde{\chi}_2^0$, and if $\tilde{\chi}_2^0$ in turn decays to $\gamma\tilde{\chi}_1^0$, these events will populate the final state $ee\gamma\gamma\not{E}_T$. The production of slepton pairs can also result in events falling into the final state $ee\gamma\gamma\not{E}_T$, since a typical decay of a selectron in a model with gravity-mediated supersymmetry breaking is $\tilde{e} \to e\tilde{\chi}_2^0$, with $\tilde{\chi}_2^0 \to \gamma\tilde{\chi}_1^0$. If a pair of sufficiently massive sleptons are produced, each can decay into the corresponding standard model lepton and the second-lightest neutralino ($\tilde{\chi}_2^0$), which in turn could decay into $\ell\ell\tilde{\chi}_1^0$. A similar production of $\tilde{\ell}\tilde{\nu}$ can easily lead to a final state with one fewer charged lepton, through the decay chain $\tilde{\nu} \to \ell\tilde{\chi}_1^\pm$, and $\tilde{\chi}_1^\pm \to \ell\nu\tilde{\chi}_1^0$. The standard model backgrounds to such events, containing five or more charged leptons and substantial missing transverse energy, are vanishingly small. Events with four charged leptons and substantial $\not{E}_T$ could result from the decay of a $\tilde{\chi}_2^0\tilde{\chi}_2^0$ pair, in which each $\tilde{\chi}_2^0$ decays to $\ell\ell\tilde{\chi}_1^0$. Even pair production of gluinos, each decaying to $q\bar{q}\tilde{\chi}_2^0$, with one neutralino decaying to $ee\tilde{\chi}_1^0$ and the other to $\gamma\tilde{\chi}_1^0$, could produce events in these final states. With this particular decay, such events would appear in the final state $ee\gamma\,2j$.

If leptons exist in excited states several hundred GeV above their ground state, just

as hadrons exist in excited states at energy scales a thousand times smaller, they could be produced in the process $q\bar{q} \to Z/\gamma^* \to \ell^*\ell^*$ or $q\bar{q}' \to W^* \to \ell^*\nu^*$. The excited leptons can decay by emitting a photon, so that $\ell^* \to \ell\gamma$ and $\nu^* \to \nu\gamma$. Such events would populate the $\ell\ell\gamma\gamma$ and $\ell\not{E}_T\gamma\gamma$ final states. If the technirho exists and is sufficiently massive, it can decay to $WZ$. Roughly one time in fifty both the $W$ and $Z$ bosons will decay to leptons, producing a $\ell^+\ell^-\ell'\not{E}_T$ event. More generally, any process producing anomalous triboson couplings will affect the $(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$ final states, and (as we show in Sec. 5.7.3) our method is likely to be sensitive to such a signal.

## 5.4 Charted and uncharted territory

The DØ experiment [5] began collecting data at $\sqrt{s} = 1.8$ TeV in 1992, and completed its first series of runs in 1996. These data have been carefully scrutinized by the DØ Collaboration. Nonetheless, the incredible richness of these data, which probe fundamental physics at the highest energy scales currently achievable, allows for the possibility that something there may yet remain undiscovered.

### 5.4.1 Final states already considered by DØ

Some portions of these data have been more comprehensively scrutinized than others. In particular, there are eight final states — those marked with triangles in Fig. 5.1 — that DØ has already analyzed in a manner similar to the Sleuth prescription.

In final states that contain only a single object (such as a $W$ or $Z$ boson), there are no non-trivial momentum variables to consider, and the Sleuth search strategy reduces

in this case to a counting experiment. In final states containing exactly two objects (such as $ee$, $Zj$, or $W\gamma$), the single momentum variable available to us is the average (scalar) transverse momentum of the two objects, assuming that both are sufficiently central. DØ has analyzed eight final states in these limiting cases. These analyses do not precisely follow the Sleuth prescription — they were performed before Sleuth was created — so $\mathcal{P}$ is not calculated for these final states. Nonetheless, they are sufficiently close to our prescription (and therefore sufficiently quasi-model-independent) that we briefly review them here, both for completeness and in order to motivate the final states that we treat in Secs. 5.5–5.7. Examples of the types of new physics that could be expected to appear in a few of these final states are provided in Appendix 5.3.2.

$2j$. DØ has performed an analysis of the dijet mass spectrum [39] and angular distribution [40] in a search for quark compositeness. We note that the dijet mass and the polar angle of the jet axis (in the center-of-mass frame of the system) together completely characterize these events, and that two central jets with large invariant mass also have large average $p_T$. No compelling evidence of an excess at large jet transverse momentum is seen in either case.

$W$. The Sleuth-defined $W$ final state contains all events with either: one muon and no second charged lepton; or one electron, significant missing transverse energy, and transverse mass $30 < m_T^{e\nu} < 110$ GeV. The Sleuth prescription reduces to a cross section measurement in this case. DØ has measured the inclusive $W$ boson cross section [32], and finds it to be in good agreement with the standard model prediction.

$e\not\!\!E_T$. Events that contain one electron, no second charged lepton, substantial $\not\!\!E_T$, and have transverse mass $m_T^{e\nu} > 110$ GeV belong to the $e\not\!\!E_T$ final state. This final state contains two objects (the electron and the missing transverse energy), so we consider the average object $p_T$, which is approximately equal in this case to $m_T^{e\nu}/2$. DØ has performed a search for right-handed $W$ bosons and heavy $W'$ bosons in 79 pb$^{-1}$ of data [41], looking for an excess in the tail of the transverse mass distribution. No such excess is observed.

$Wj$. In the two-object final state $Wj$, the average transverse momentum of the two objects is essentially $p_T^W$, the transverse momentum of the $W$ boson. DØ has measured the $W$ boson $p_T$ distribution [42], and finds good agreement with the standard model.

$W\gamma$. Similarly, the transverse momentum distribution of the photon in $W\gamma X$ events has been analyzed by DØ in a measurement of the $WW\gamma$ gauge boson coupling parameters [43]. No excess at large $p_T^\gamma$ is observed. (The Sleuth prescription for defining final states is less well satisfied in DØ's corresponding measurement of $p_T^\gamma$ in $Z\gamma X$ events [44].)

$Z$. As in the case of the $W$ final state, our prescription reduces to a counting experiment in the $Z$ final state. DØ has published a measurement of the inclusive $Z$ boson cross section [32], and finds it to be in good agreement with the standard model prediction.

$ee$. Events containing two electrons and nothing else fall into the final state $ee$ if the invariant mass $m_{ee}$ is outside the $Z$ boson mass window of $(82, 100)$ GeV. The single variable we consider in this two-object final state is the average scalar transverse momentum of the two electrons, which is simply related to the invariant mass $m_{ee}$ for sufficiently central elec-

trons. DØ has analyzed the high mass Drell-Yan cross section in a search for indications of quark-lepton compositeness with the full data set [45], and has analyzed the $ee$ invariant mass distribution in the context of a search for additional neutral gauge bosons in a subset of those data [46]. No discrepancy between the data and expected background is observed.

**$Zj$.** In the two-object final state $Zj$, the average transverse momentum of the two objects is essentially the transverse momentum of the $Z$ boson. DØ's published measurement of the $Z$ boson $p_T$ distribution [47] is in good agreement with the standard model prediction.

### 5.4.2 Final states considered in this article

The decision as to which of the remaining final states should be subjected to a Sleuth analysis was made on the basis of our ability to estimate the standard model and instrumental backgrounds in each final state, and the extent to which a systematic analysis for new physics is lacking in each final state. The final states we chose to analyze arranged themselves into four "classes": $e\mu X$, $W$+jets-like final states, $Z$+jets-like final states, and $(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$. The first of these classes has been analyzed in Ref. [16] and summarized in Sec. 4.4. A systematic Sleuth analysis of the remaining three classes of final states is the subject of the next three sections.

## 5.5 $W$+jets-like final states

In this section we analyze the $W$+jets-like final states — events containing a single lepton, missing transverse energy, and two or more jets. In Sec. 5.5.1 we describe the $e\not\!\!E_T\,2j(nj)$ and $\mu\not\!\!E_T\,2j(nj)$ data sets and background estimates, and in Sec. 5.5.2 we present

the results. After this, we feign ignorance of the heaviest quark in the standard model and check the sensitivity of our method to top quark pair production in Sec. 5.5.3. A few of the many signals that might appear in these final states are described in Sec. 5.3.3.

### 5.5.1 Data sets and background estimates

$e\not\!\!E_T\, 2j(nj)$

The $e\not\!\!E_T\, 2j(nj)$ data set [48] comprises $115 \pm 6$ pb$^{-1}$ of collider data, collected with triggers that require the presence of an electromagnetic object, with or without jets and missing transverse energy. Offline event selection requires: one electron with transverse energy $p_T^e > 20$ GeV and pseudorapidity $|\eta_{\rm det}| < 1.1$ or $1.5 < |\eta_{\rm det}| < 2.5$ [49]; $\not\!\!E_T > 30$ GeV; and two or more jets with $p_T^j > 20$ GeV and $|\eta_{\rm det}| < 2.5$. Effects of jet energy mismeasurement are reduced by requiring the $\not\!\!E_T$ vector to be separated from the jets by $\Delta\phi > 0.25$ radians if $\not\!\!E_T < 120$ GeV. To reduce background from a class of events in which a fake electron's energy is overestimated, leading to spurious $\not\!\!E_T$, we reject events with $p_T^W < 40$ GeV. Events containing isolated muons appear in a sample analyzed previously with this method ($e\mu X$), and are not considered here.

The dominant standard model and instrumental backgrounds to the $e\not\!\!E_T\, 2j(nj)$ final states are from:

- $W$ + jets production, with $W \to e\nu$;
- multijet production, with mismeasured $\not\!\!E_T$ and one jet faking an electron; and
- $t\bar{t}$ pair production, with $t \to Wb$ and with at least one $W$ boson decaying to an electron or to a tau that in turn decays to an electron.

The $W$+jets background is simulated using VECBOS [50], with HERWIG [28] used for fragmenting the partons. The background from multijet events containing a jet that is misidentified as an electron, and with $\not\!\!E_T$ arising from the mismeasurement of jet energies, is modeled using multijet data. The probability for a jet to be misidentified as an electron is estimated [51] to be $(3.50 \pm 0.35) \times 10^{-4}$. The background from $t\bar{t}$ decays into an electron plus two or more jets is simulated using HERWIG with a top quark mass of 170 GeV. All Monte Carlo event samples are processed through the DØ detector simulation based on the GEANT [15] package.

We estimate the number of $t\bar{t}$ events in the $W$+jets-like final states to be $18 \pm 6$ using the measured $t\bar{t}$ production cross section of $5.5 \pm 1.8$ pb [24]. The multijet background is estimated to be $21 \pm 7$ events, using a sample of events with three or more jets with $\not\!\!E_T > 30$ GeV. This is done by multiplying the fake probability by the number of ways the events satisfy the selection criteria with one of the jets passing the electron $p_T$ and $\eta$ requirements. After the estimated numbers of $t\bar{t}$ and multijet background events are subtracted, the number of events with transverse mass of the electron and neutrino ($m_T^{e\nu}$) below 110 GeV is used to obtain an absolute normalization for the $W$+jets background.

Following the Sleuth prescription, we combine the electron and missing transverse energy into a $W$ boson if $30 < m_T^{e\nu} < 110$ GeV, and reject events with $m_T^{e\nu} < 30$ GeV. The expected numbers of background events for the exclusive final states within this $e\not\!\!E_T\, 2j(nj)$ sample are provided in Table 5.2.

| Final State | $W$+jets | QCD fakes | $t\bar{t}$ | Total | Data |
|---|---|---|---|---|---|
| $e\not\!\!E_T\,2j$ | $6.7 \pm 1.4$ | $3.3 \pm 0.9$ | $1.7 \pm 0.6$ | $11.6 \pm 1.7$ | 7 |
| $e\not\!\!E_T\,3j$ | $1.0 \pm 0.4$ | $0.48 \pm 0.22$ | $1.0 \pm 0.4$ | $2.5 \pm 0.6$ | 5 |
| $e\not\!\!E_T\,4j$ | $0.15 \pm 0.11$ | $0.38 \pm 0.19$ | $0.26 \pm 0.09$ | $0.80 \pm 0.24$ | 2 |
| $W(\rightarrow e\not\!\!E_T)\,2j$ | $334 \pm 51$ | $12.0 \pm 2.6$ | $4.0 \pm 1.4$ | $350 \pm 51$ | 387 |
| $W(\rightarrow e\not\!\!E_T)\,3j$ | $57 \pm 9$ | $3.4 \pm 0.9$ | $6.0 \pm 2.1$ | $66 \pm 9$ | 56 |
| $W(\rightarrow e\not\!\!E_T)\,4j$ | $5.9 \pm 1.3$ | $1.1 \pm 0.4$ | $3.9 \pm 1.4$ | $10.9 \pm 1.9$ | 11 |
| $W(\rightarrow e\not\!\!E_T)\,5j$ | $0.8 \pm 0.3$ | $0.19 \pm 0.12$ | $0.73 \pm 0.26$ | $1.8 \pm 0.4$ | 1 |
| $W(\rightarrow e\not\!\!E_T)\,6j$ | $0.12 \pm 0.06$ | $0.030 \pm 0.015$ | $0.10 \pm 0.04$ | $0.25 \pm 0.07$ | 1 |

Table 5.2: Expected backgrounds to the $e\not\!\!E_T\,2j(nj)$ final states. The final states labeled "$W(\rightarrow e\not\!\!E_T)$" have $m_T^{e\nu} < 110$ GeV; the final states labeled "$e\not\!\!E_T$" have $m_T^{e\nu} > 110$ GeV. We have extrapolated our background estimates to final states with five or more jets. Berends scaling and the data in this table suggest that a factor of $\approx 7$ in cross section is the price to be paid for an additional radiated jet with transverse energy above 20 GeV.

$\mu\not\!\!E_T\,2j(nj)$

The $\mu\not\!\!E_T\,2j(nj)$ data set [52] corresponds to $94 \pm 5$ pb$^{-1}$ of integrated luminosity. The initial sample is composed of events passing any of several muon + jets triggers requiring a muon with $p_T^{\mu} > 5$ GeV within $|\eta_{\rm det}| < 1.7$ and one or more jets with $p_T^{j} > 8$ GeV and $|\eta_{\rm det}| < 2.5$. Using standard jet and muon identification criteria, we define a final sample containing one muon with $p_T > 25$ GeV and $|\eta_{\rm det}| < 0.95$, two or more jets with $p_T^{j} > 15$ GeV and $|\eta_{\rm det}| < 2.0$ and with the most energetic jet within $|\eta_{\rm det}| < 1.5$, and missing transverse energy $\not\!\!E_T > 30$ GeV. Because an energetic muon's momentum is not well measured in the detector, we are unable to separate "$W$-like" events from "non-$W$-like" events using the transverse mass, as we have done above in the electron channel. The muon and missing transverse energy are therefore always combined into a $W$ boson.

The dominant standard model and instrumental backgrounds to these final states are from:

- $W$ + jets production with $W \rightarrow \mu\nu$;

- $Z$ + jets production with $Z \to \mu\mu$, where one of the muons is not detected;
- $WW$ pair production with one $W$ boson decaying to a muon or to a tau that in turn decays to a muon; and
- $t\bar{t}$ pair production with $t \to Wb$ and with at least one $W$ boson decaying to a muon or to a tau that in turn decays to a muon.

Samples of $W$ + jets and $Z$ + jets events are generated using VECBOS, employing HERWIG for parton fragmentation. Background due to $WW$ pair production is simulated with PYTHIA [29]. Background from $t\bar{t}$ pair production is simulated using HERWIG with a top quark mass of 170 GeV. All Monte Carlo samples are again processed through a detector simulation program based on the GEANT package.

The expected backgrounds for the exclusive final states within $\mu\not\!\!E_T\, 2j(nj)$ are listed in Table 5.3. These $W(\to \mu\not\!\!E_T)\, 2j(nj)$ final states are combined with the $W(\to e\not\!\!E_T)\, 2j(nj)$ final states described in Sec. 5.5.1 to form the $W\, 2j(nj)$ final states treated in Sec. 5.5.1. For consistency in this combination, we also require $p_T^W > 40$ GeV for the $W(\to \mu\nu)\, 2j(nj)$ final states.

| Final State | $W$+jets | $Z$+jets | $WW$ | $t\bar{t}$ | Total | Data |
|---|---|---|---|---|---|---|
| $W(\to \mu\not\!\!E_T)\, 2j$ | $48 \pm 15$ | $1.6 \pm 0.4$ | $0.5 \pm 0.3$ | $0.42 \pm 0.14$ | $50 \pm 15$ | 54 |
| $W(\to \mu\not\!\!E_T)\, 3j$ | $10 \pm 3$ | $0.27 \pm 0.08$ | $0.41 \pm 0.26$ | $0.58 \pm 0.20$ | $11 \pm 3$ | 11 |
| $W(\to \mu\not\!\!E_T)\, 4j$ | $2.8 \pm 1.3$ | $0.022 \pm 0.011$ | – | $0.61 \pm 0.21$ | $3.5 \pm 1.3$ | 4 |

Table 5.3: Expected backgrounds for the $W(\to \mu\not\!\!E_T)\, 2j(nj)$ final states.

$W\,2j(nj)$

Combining the results in Tables 5.2 and 5.3 gives the expected backgrounds for the $W\,2j(nj)$ final states shown in Table 5.4. We note the good agreement in all final states between the total number of background events expected and the number of data events observed. This of course is due in part to the method of normalizing the $W$+jets background. The agreement in the final states containing additional jets is also quite good. A more detailed comparison between data and background in the more heavily populated final states ($W\,2j$, $W\,3j$, and $W\,4j$) is provided in Figs. 5.4–5.6.

| Final State | Total | Data |
|---|---|---|
| $W\,2j$ | $400 \pm 53$ | 441 |
| $W\,3j$ | $77 \pm 10$ | 67 |
| $W\,4j$ | $14.3 \pm 2.3$ | 15 |
| $W\,5j$ | $1.8 \pm 0.4$ | 1 |
| $W\,6j$ | $0.25 \pm 0.07$ | 1 |

Table 5.4: Expected backgrounds to the $W\,2j(nj)$ final states.

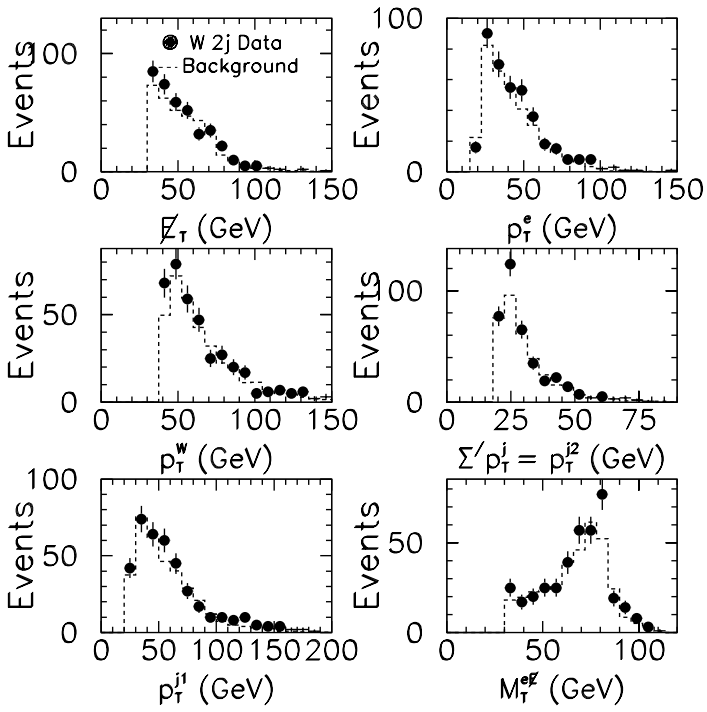



Figure 5.4: Comparison of background to data for $W\,2j$.

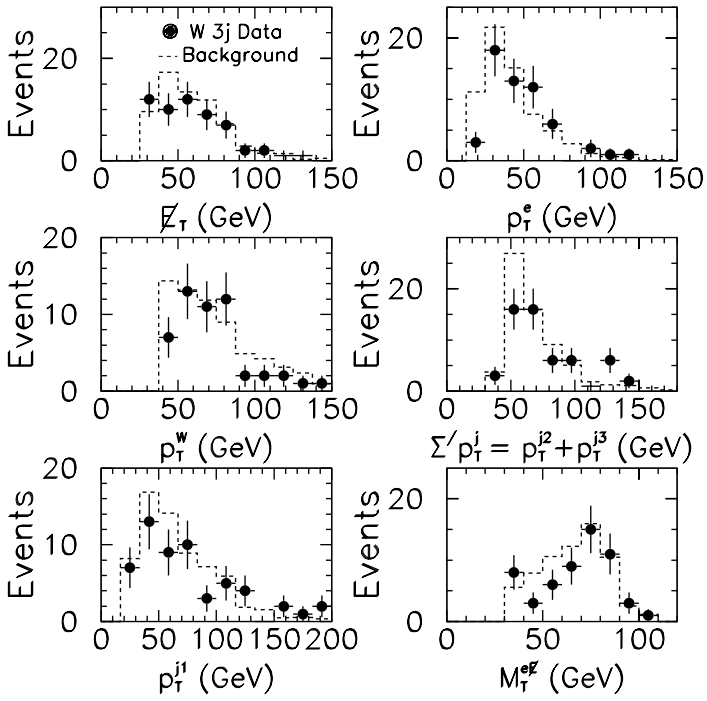



Figure 5.5: Comparison of background to data for $W\,3j$.

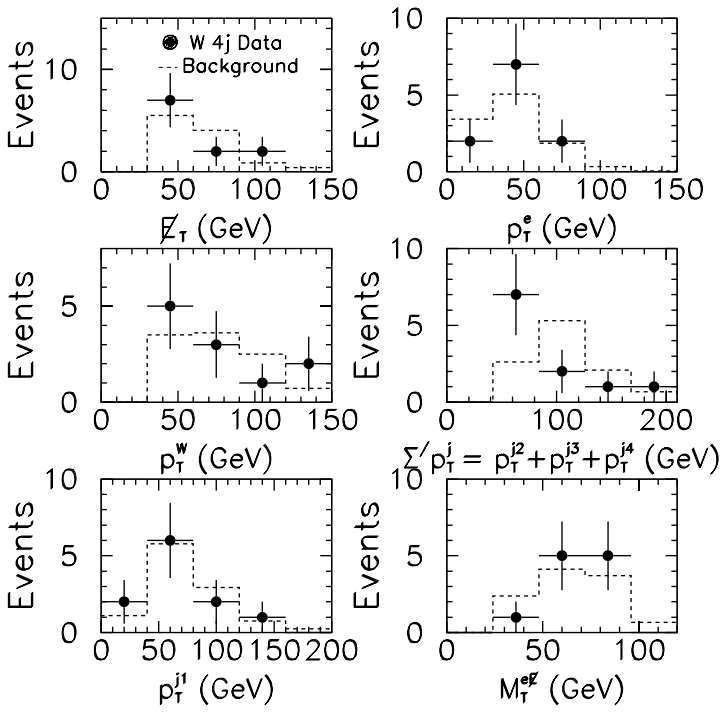



Figure 5.6: Comparison of background to data for $W\,4j$.

Monte Carlo programs suitable for estimating backgrounds to final states with many additional jets are not readily available. It has been observed that the rate of a process may be related to the rate of the process with an additional radiated jet by a multiplicative factor of 1/4–1/7, depending upon the $p_T$ and $\eta$ thresholds used to define a jet — this phenomenological law is known as Berends scaling [50]. We estimate that this factor is $\approx 1/5$ for jets with $|\eta_{\rm det}| < 2.5$ and $p_T > 15$ GeV, and that it is $\approx 1/7$ for jets with $|\eta_{\rm det}| < 2.5$ and $p_T > 20$ GeV. This will be used to estimate particular background contributions to final states in which the expected background is $\lesssim 1$.

### 5.5.2 Results

The results of applying Sleuth to the $e\not\!\!E_T\, 2j(nj)$ and $W\, 2j(nj)$ data sets are summarized in Table 5.5 and in Figs. 5.7 and 5.8. Recall from Sec. 4.3.2 that the positions of the data points within the unit box are determined by the background distribution, which defines the transformation from the original variable space, in addition to the location of the data points in that original space. We observe quite good agreement with the standard model in the $W$+jets-like final states.

| Data set | $\mathcal{P}$ |
|---|---|
| $e\not\!\!E_T\, 2j$ | 0.76 |
| $e\not\!\!E_T\, 3j$ | 0.17 |
| $e\not\!\!E_T\, 4j$ | 0.13 |
| $W\, 2j$ | 0.29 |
| $W\, 3j$ | 0.23 |
| $W\, 4j$ | 0.53 |
| $W\, 5j$ | 0.81 |
| $W\, 6j$ | 0.22 |

Table 5.5: Summary of results on $e\not\!\!E_T\, 2j(nj)$ and $W\, 2j(nj)$.

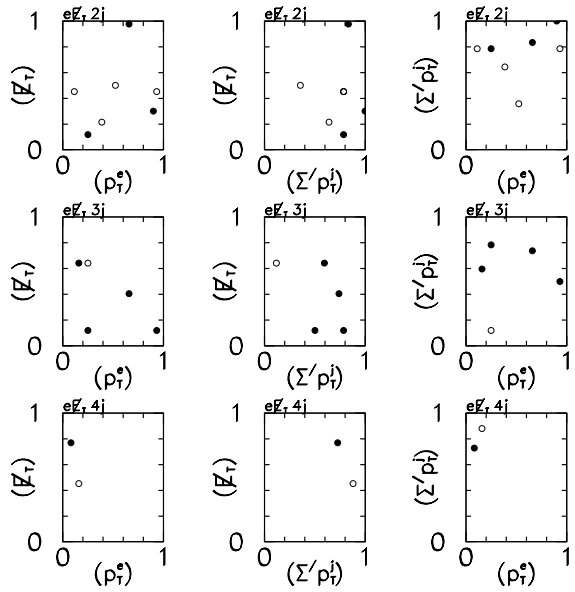



Figure 5.7: The positions of the transformed data points in the final states $e\not\!\!E_T\,2j$, $e\not\!\!E_T\,3j$, and $e\not\!\!E_T\,4j$. The data points inside the region chosen by Sleuth are shown as filled circles; those outside the region are shown as open circles. For these final states the variables $p_T^e$, $\not\!\!E_T$, and $\sum' p_T^j$ are considered, and the unit box is in this case a unit cube. The two-dimensional views shown here are the projections of that cube onto three orthogonal faces.

### 5.5.3 Sensitivity check: $t\bar{t} \to \ell$+jets

In this section we check Sleuth's sensitivity to $t\bar{t}$ in the final states $W\,3j$, $W\,4j$, $W\,5j$, and $W\,6j$. After briefly putting this signal into context, we test Sleuth's ability to find $t\bar{t}$ in the data, and then in an ensemble of mock experiments.

In 1997 DØ published a measurement of the top quark production cross section [24] based on events in the dilepton, $\ell$+jets, $\ell$+jets($/\mu$), and "$e\nu$" channels, where "$/\mu$" indicates that one or more of the jets contains a muon, and hence is likely to be the product of a $b$ quark. 19 events with no $b$-quark tag are observed in $\ell$+jets (nine events in the electron channel, and ten events in the muon channel) with an expected background of $8.7 \pm 1.7$. An

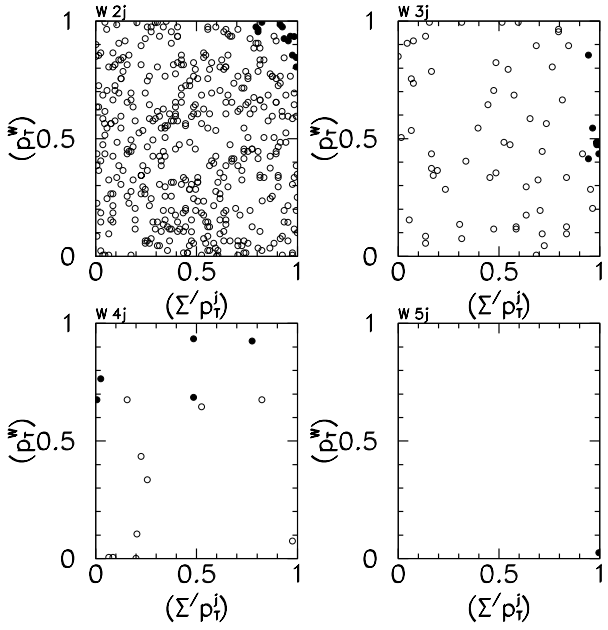



Figure 5.8: The positions of the transformed data points in the final states $W\,2j$, $W\,3j$, $W\,4j$, and $W\,5j$. The data points inside the region chosen by Sleuth are shown as filled circles; those outside the region are shown as open circles. The single event in the $W\,5j$ final state is in the lower right-hand corner of the unit square, having $\sum' p_T^j = 300$ GeV.

additional eleven events are observed with a $b$-quark tag (five events in the electron channel, and six events in the muon channel) with an expected background of $2.5 \pm 0.5$ events. Three or more jets with $p_T > 15$ GeV are required in both cases. The number of events observed in all four channels is 39 with an expected background of $13 \pm 2.2$ events. The probability for $13 \pm 2.2$ to fluctuate up to or above 39 is $6 \times 10^{-7}$, or 4.8 standard deviations. In the $\ell$+jets channel alone, the probability that $8.7 \pm 1.7$ fluctuates to the 19 events observed is 0.005, corresponding to a "significance" of $2.6\sigma$.

Figures 5.9(a) and 5.9(c) show where $t\bar{t}$ Monte Carlo events fall in the unit box in the final states $W\,3j$ and $W\,4j$. The distribution of these events is quite diffuse in the case

of $W\,3j$, since $t\bar{t}$ is similar to the background in the variables $p_T^W$ and $\sum' p_T^j$ in this channel. In the $W\,4j$ final state $t\bar{t}$ tends to populate regions of large $\sum' p_T^j$, but the signal is nearly indistinguishable from background in the variable $p_T^W$. A check of Sleuth's ability to find $t\bar{t}$ in the $W\,3j(nj)$ final states tests how well Sleuth performs when the signal shows up in a subset of the variables we choose to consider.

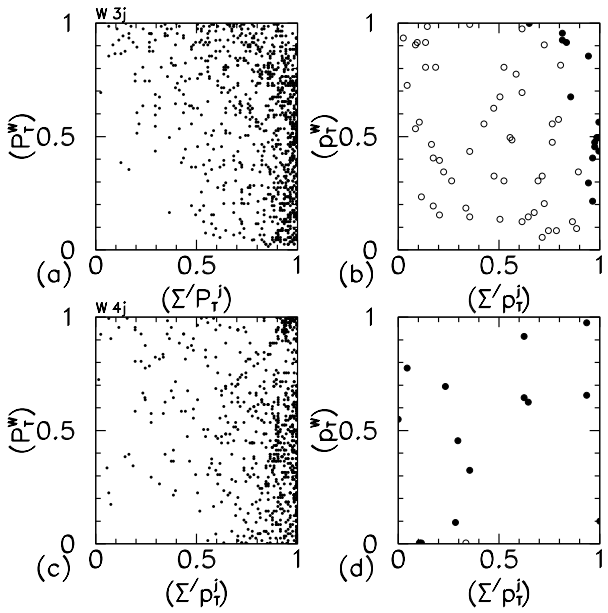



Figure 5.9: Scatter plot of where $t\bar{t}$ Monte Carlo events fall in the unit box in the final states $W\,3j$ (a) and $W\,4j$ (c). Although top quark events appear in the high tails of $\sum' p_T^j$, the variable $p_T^W$ is not particularly discriminating. The locations of the data points are shown in (b) and (d). The backgrounds are taken to include all standard model processes except top quark pair production.

Figures 5.9(b) and 5.9(d) show DØ data in the final states $W\,3j$ and $W\,4j$, when $t\bar{t}$ is not included in the background estimate. Notice that the region chosen by Sleuth in the $W\,3j$ final state in Fig. 5.9(b) is very similar to the region populated by $t\bar{t}$ in Fig. 5.9(a). In the $W\,4j$ final state (d), the region chosen by Sleuth is nearly the entire unit box. Comparison

with Fig. 5.8 shows how the absence of $t\bar{t}$ in the background estimate in this figure affects the transformation from the original variable space into the unit box. Applying Sleuth to these data while continuing to feign ignorance of $t\bar{t}$, we find $\mathcal{P}_{W\,3j} = 0.12$, $\mathcal{P}_{W\,4j} = 0.18$, $\mathcal{P}_{W\,5j} = 0.37$, and $\mathcal{P}_{W\,6j} = 0.09$. Upon combining these results, we find $\mathcal{P}_{\rm min} = \min(\mathcal{P}_{W\,3j}, \mathcal{P}_{W\,4j}, \mathcal{P}_{W\,5j}, \mathcal{P}_{W\,6j}) = 0.09$ $(1.3\sigma)$.

Figure 5.10 shows a histogram of $\mathcal{P}_{\rm min}$ for a sample of mock experimental runs in which the backgrounds include $W$+jets and QCD events, and the mock samples include $t\bar{t}$ in addition to the expected background. The number of background and $t\bar{t}$ events in the mock samples are allowed to vary according to statistical and systematic errors. Note that since four final states are considered, the distribution of $\mathcal{P}_{\rm min}$ for an ensemble of experiments including background only has a median of $\approx 1\sigma$. We see that Sleuth is able to find indications of the presence of $t\bar{t}$ in these final states, returning $\mathcal{P}_{{\rm min}\,[\sigma]} > 3$ in 30% of an ensemble of mock experimental runs containing $t\bar{t}$ events, compared to only 0.5% of an ensemble of mock experimental runs containing background only.

We conclude from this sensitivity check that Sleuth would not have been able to "discover" $t\bar{t}$ in the DØ $W$+jets data, but that in 30% of an ensemble of mock experimental runs Sleuth would have found $\mathcal{P}_{{\rm min}\,[\sigma]} > 3$.

## 5.6 $Z$+jets-like final states

In this section we analyze the $Z$+jets-like final states. We first describe the data sets and background estimates for the dielectron+jets channels, and we then discuss the dimuon+jets channels. After presenting our results, we check the sensitivity of our method

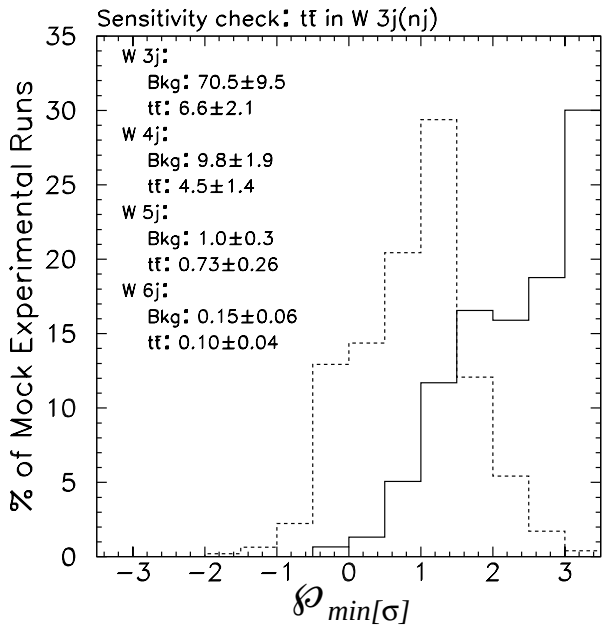



Figure 5.10: Histogram of $\mathcal{P}_{\rm min} = \min(\mathcal{P}_{W\,3j}, \mathcal{P}_{W\,4j}, \mathcal{P}_{W\,5j}, \mathcal{P}_{W\,6j})$ for an ensemble of mock experimental runs in which the backgrounds include $W$+jets and QCD events, and the mock samples include (solid) / do not include (dashed) $t\bar{t}$ in addition to the expected background. All experimental runs with $\mathcal{P}_{\rm min} > 3\sigma$ are in the rightmost bin.

to the presence of first generation scalar leptoquarks. Section 5.3.4 describes signals that might appear in these final states.

### 5.6.1 Data sets and background estimates

$ee\,2j(nj)$

The $ee\,2j(nj)$ data set [51], corresponding to an integrated luminosity of $123 \pm 7$ pb$^{-1}$, is collected with triggers requiring the presence of two electromagnetic objects. Offline event selection requires two electrons passing standard identification criteria with transverse momenta $p_T^e > 20$ GeV and pseudorapidity $|\eta_{\rm det}| < 1.1$ or $1.5 < |\eta_{\rm det}| < 2.5$, and

two or more jets with $p_T^j > 20$ GeV and $|\eta_{\rm det}| < 2.5$. At least one electron is required to have a matching track in the central tracking detectors and to satisfy ionization requirements in the tracking chambers and transition radiation detector. For these data the trigger energy threshold forces a transverse momentum cut of 20 GeV, rather than the Sleuth-preferred requirement of 15 GeV. We cut on a likelihood described in Appendix B.1 in order to correctly identify any events with significant missing transverse energy. Electron pairs are combined into a $Z$ boson if $82 < m_{ee} < 100$ GeV, unless the event contains significant $\not\!\!E_T$ (in which case it falls within $ee\not\!\!E_T X$, discussed in this section) or a third charged lepton (in which case it falls within $(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$, discussed in Sec. 5.7).

The dominant standard model and instrumental backgrounds to this data set are:

- Drell-Yan + jets production, with $Z/\gamma^* \to ee$;
- QCD multijets, with two jets faking electrons; and
- $t\bar{t}$ pair production with $t \to Wb$ and with each $W$ boson decaying to an electron or to a tau that in turn decays to an electron.

Monte Carlo samples for the Drell-Yan events are generated using ISAJET [30]. The Drell-Yan cross section normalization is fixed by comparing the Monte Carlo events with $Z$ + $\geq 2$ jets data in the $Z$ boson region. Top quark events are generated using HERWIG at a top quark mass of 170 GeV with all dilepton final states included. The DØ measured $t\bar{t}$ production cross section of $5.5 \pm 1.8$ pb at a top quark mass of 173.3 GeV was used [24]. The multijet background is estimated from a sample of events with four or more jets in which the probability for two jets or photons to be misidentified as electrons is weighted by the number of jets in the event that passed the electron $p_T$ and $\eta$ requirements. This misidentification

probability is calculated from a sample of events with three jets to be $(3.50 \pm 0.35) \times 10^{-4}$ for an electron with a reconstructed track and $(1.25 \pm 0.13) \times 10^{-3}$ for an electron without a reconstructed track. The uncertainties in these probabilities reflect a slight dependence on the jet $p_T$ and $\eta$. The expected backgrounds for the exclusive final states within $ee\,2j(nj)$ are listed in Table 5.6.

| Final State | $Z/\gamma^*$+jets | QCD fakes | Total | Data |
|---|---|---|---|---|
| $ee\,2j$ | $20 \pm 4$ | $12.2 \pm 1.8$ | $32 \pm 4$ | 32 |
| $ee\,3j$ | $2.6 \pm 0.6$ | $1.85 \pm 0.28$ | $4.5 \pm 0.6$ | 4 |
| $ee\,4j$ | $0.40 \pm 0.20$ | $0.24 \pm 0.04$ | $0.64 \pm 0.20$ | 3 |
| $ee\not\!\!E_T\,2j$ | $3.7 \pm 0.8$ | – | $3.7 \pm 0.8$ | 2 |
| $ee\not\!\!E_T\,3j$ | $0.45 \pm 0.13$ | – | $0.45 \pm 0.13$ | 1 |
| $ee\not\!\!E_T\,4j$ | $0.061 \pm 0.028$ | – | $0.061 \pm 0.028$ | 1 |
| $Z(\to ee)\,2j$ | $94 \pm 19$ | $1.88 \pm 0.28$ | $96 \pm 19$ | 82 |
| $Z(\to ee)\,3j$ | $12.7 \pm 2.7$ | $0.27 \pm 0.04$ | $13.0 \pm 2.7$ | 11 |
| $Z(\to ee)\,4j$ | $1.8 \pm 0.5$ | $0.034 \pm 0.006$ | $1.8 \pm 0.5$ | 1 |
| $Z(\to ee)\,5j$ | $0.26 \pm 0.10$ | $0.0025 \pm 0.0009$ | $0.26 \pm 0.10$ | 0 |

Table 5.6: Expected backgrounds to the $ee\,2j(nj)$, $ee\not\!\!E_T\,2j(nj)$, and $Z(\to ee)\,2j(nj)$ final states.

### $\mu\mu\,2j(nj)$

The $\mu\mu\,2j(nj)$ data set [53] corresponds to $94 \pm 5$ pb$^{-1}$ of integrated luminosity. The initial sample is composed of events passing any of several muon + jets triggers requiring a muon with $p_T^\mu > 5$ GeV within $|\eta_{\rm det}| < 1.7$ and one or more jets with $p_T^j > 8$ GeV and $|\eta_{\rm det}| < 2.5$. Using standard jet and muon identification criteria, we define a final sample containing two or more muons with $p_T > 20$ GeV and $|\eta_{\rm det}| < 1.7$ and at least one muon in the central detector ($|\eta_{\rm det}| < 1.0$), and two or more jets with $p_T^j > 20$ GeV and $|\eta_{\rm det}| < 2.5$.

We combine a $\mu\mu$ pair into a $Z$ boson if the muon momenta can be varied within their resolutions such that $m_{\mu\mu} \approx M_Z$ and the missing transverse energy becomes negligible.

More specifically, we combine a muon pair into a $Z$ boson if

$$\chi = \min_{a,b} \quad \left( \frac{1/a - 1/p^{\mu 1}}{\delta(1/p^{\mu 1})} \oplus \frac{1/b - 1/p^{\mu 2}}{\delta(1/p^{\mu 2})} \oplus \right.$$
$$\left. \frac{m_{ab} - M_Z}{\Gamma_Z} \oplus \frac{\not{E}_{T\,ab}}{\delta(\not{E}_T)} \right) < 20, \tag{5.3}$$

where $\delta(1/p) = 0.18(p-2)/p^2 \oplus 0.003$ is the uncertainty in the reciprocal of the muon momentum; $\delta(\not{E}_T) = 0.7$ GeV$\sqrt{\sum p_T^j / \text{GeV}}$ is the error on the missing transverse energy measured in the calorimeter; $m_{ab}$ and $\not{E}_{T\,ab}$ are the muon pair invariant mass and missing transverse energy, computed taking the muons to have scalar momenta $a$ and $b$; $M_Z$ and , $_Z$ are the mass and width of the $Z$ boson; and $\oplus$ means addition in quadrature. The cut of $\chi < 20$ is chosen so that $Z(\rightarrow \mu\mu)$ is not the dominant background to the $\mu\mu\, 2j(nj)$ final states.

The most significant standard model and instrumental backgrounds to this data set are

- $Z$ + jets production with $Z \rightarrow \mu\mu$,
- $WW$ pair production with each $W$ boson decaying to a muon or to a tau that in turn decays to a muon, and
- $t\bar{t}$ pair production with $t \rightarrow Wb$ and with each $W$ boson decaying to a muon or to a tau that in turn decays to a muon.

A sample of $Z$ + jets events was generated using VECBOS, employing HERWIG for parton fragmentation. Background due to $WW$ pair production is simulated with PYTHIA. Background from $t\bar{t}$ pair production is simulated using HERWIG with a top quark mass of

170 GeV. All Monte Carlo samples are processed through a detector simulation program based on the GEANT package.

The expected backgrounds for the exclusive final states within $\mu\mu\,2j(nj)$ are listed in Table 5.7. The $Z(\to \mu\mu)\,2j(nj)$ final states are combined with the $Z(\to ee)\,2j(nj)$ final states described in Sec. 5.6.1 to form the $Z\,2j(nj)$ final states treated in Sec. 5.6.1.

| Final State | $Z$+jets | $WW$ | $t\bar{t}$ | Total | Data |
|---|---|---|---|---|---|
| $\mu\mu\,2j$ | $0.112 \pm 0.029$ | $0.25 \pm 0.13$ | $0.14 \pm 0.05$ | $0.50 \pm 0.15$ | 2 |
| $\mu\mu\,3j$ | $0.007 \pm 0.004$ | $0.06 \pm 0.04$ | $0.065 \pm 0.025$ | $0.13 \pm 0.05$ | 0 |
| $Z(\to \mu\mu)\,2j$ | $2.2 \pm 0.4$ | – | $0.050 \pm 0.020$ | $2.3 \pm 0.4$ | 3 |
| $Z(\to \mu\mu)\,3j$ | $0.24 \pm 0.05$ | – | $0.018 \pm 0.009$ | $0.26 \pm 0.06$ | 1 |

Table 5.7: Expected backgrounds to the $Z(\to \mu\mu)\,2j(nj)$ and $\mu\mu\,2j(nj)$ final states.

$Z\,2j(nj)$

Combining the results in Tables 5.6 and 5.7 gives the expected backgrounds for the $Z\,2j(nj)$ final states, shown in Table 5.8. The number of dimuon events in these tables is significantly smaller than the number of dielectron events due to especially tight identification requirements on the muons.

$Z/\gamma^*$ is the dominant background to nearly all final states discussed in this section, although other sources of background contribute significantly when the dilepton mass is outside the $Z$ boson mass window. The agreement between the total number of events expected and the number observed in the data is quite good, even for final states with several jets. While any analysis of $Z$+jets-like states will need to rely to some degree on an accurate $Z/\gamma^*$+jets Monte Carlo, having a reliable estimate of the jet distributions in such events is especially important when exclusive final states are considered. We anticipate that

this will become increasingly important in the next Tevatron run. Differential agreement between data and the expected background may be seen by considering a comparison of various kinematic quantities in Figs. 5.11 and 5.12.

| Final State | Total | Data |
|---|---|---|
| $Z\,2j$ | $98 \pm 19$ | 85 |
| $Z\,3j$ | $13.2 \pm 2.7$ | 12 |
| $Z\,4j$ | $1.9 \pm 0.5$ | 1 |
| $Z\,5j$ | $0.26 \pm 0.10$ | 0 |

Table 5.8: Expected backgrounds to the $Z\,2j(nj)$ final states.

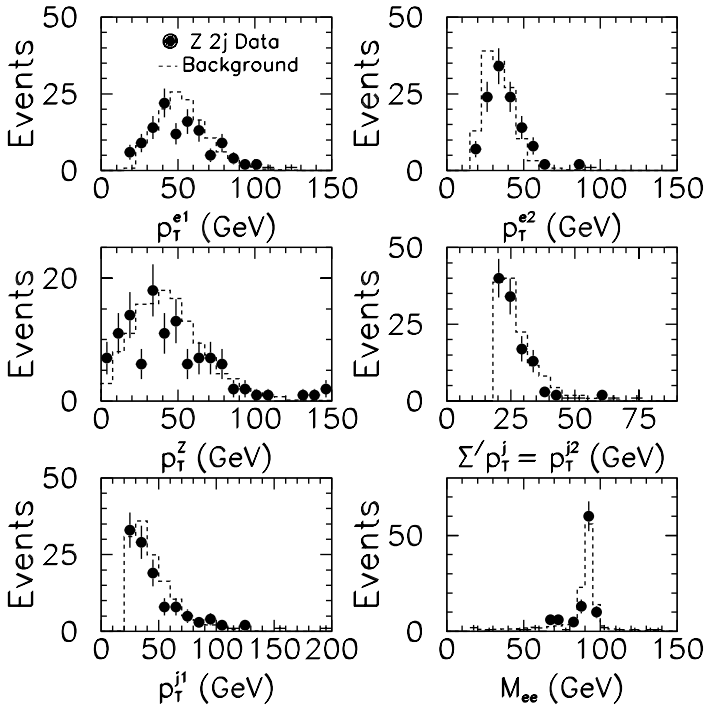



Figure 5.11: Comparison of background to data for $Z\,2j$.

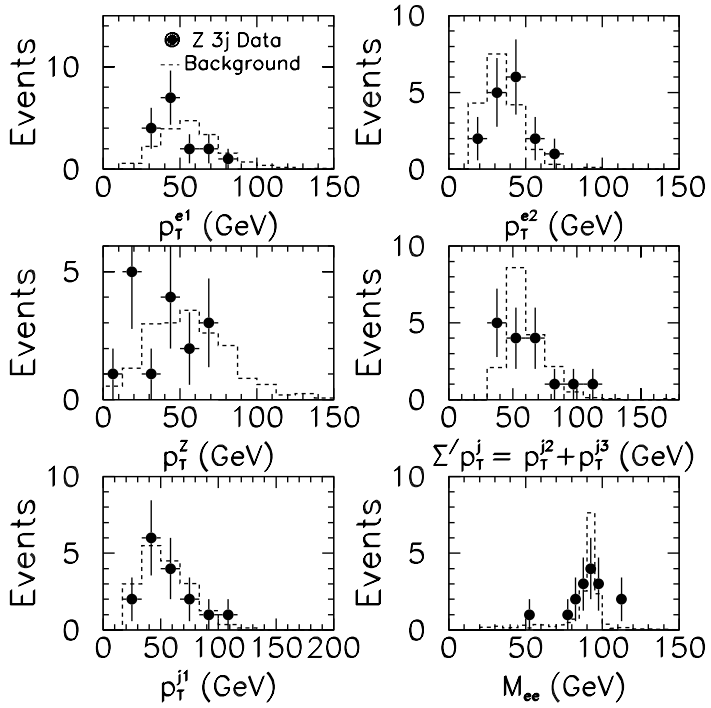



Figure 5.12: Comparison of background to data for $Z\,3j$.

### 5.6.2 Results

The results of applying Sleuth to the $Z\,2j(nj)$ and $\ell\ell\,2j(nj)$ data sets are summarized in Table 5.9 and Figs. 5.13 and 5.14. Figure 5.13 shows the location of the data within the unit box for those final states in which the two leptons are not combined into a $Z$ boson, while Fig. 5.14 displays the data for those final states in which a $Z$ boson has been identified. Large $\mathcal{P}$'s are found for most final states, as expected. The smallest $\mathcal{P}$'s in this class of final states are observed in the $ee\,4j$ and $ee\not\!\!E_T\,4j$ final states. Although the number of events is small, it is interesting to compare the number of events observed in the $Z$ + 2, 3, and 4 jet final states (showing good agreement with expected backgrounds) with the

number of events observed in the $ee$ + 2, 3, and 4 jet and $ee\not\!\!E_T$ + 2, 3, and 4 jet final states. There is a small but statistically insignificant excess in final states with four jets — we find in Sec. 5.8 that we expect to find at least one $\mathcal{P} \lesssim 0.04$ in the analysis of so many final states. Additionally, one of the three $ee\,4j$ events has an $ee$ invariant mass barely outside the $Z$ boson mass window. The kinematics of the events in the $ee\,4j$ and $ee\not\!\!E_T\,4j$ final states are provided in Appendix B.2.

| Data set | $\mathcal{P}$ |
|---|---|
| $ee\,2j$ | 0.72 |
| $ee\,3j$ | 0.61 |
| $ee\,4j$ | 0.04 |
| $ee\not\!\!E_T\,2j$ | 0.68 |
| $ee\not\!\!E_T\,3j$ | 0.36 |
| $ee\not\!\!E_T\,4j$ | 0.06 |
| $\mu\mu\,2j$ | 0.08 |
| $\mu\mu\,3j$ | 1.00 |
| $Z\,2j$ | 0.52 |
| $Z\,3j$ | 0.71 |
| $Z\,4j$ | 0.83 |
| $Z\,5j$ | 1.00 |

Table 5.9: Summary of results on the $Z$+jets-like final states.

### 5.6.3 Sensitivity check: leptoquarks

As a sensitivity check in the $Z$+jets-like final states we consider a scalar, first generation leptoquark [54] of mass $m_{LQ} = 170$ GeV, and assume a branching fraction to charged leptons of $\beta = 1.0$. The cross section for the process $q\bar{q} \to LQ\overline{LQ}$ with these parameters is 0.54 pb. The overall efficiency for this type of event is $24 \pm 4\%$ [51], including trigger and object requirement efficiencies and geometric and kinematic acceptances. If such a leptoquark were to exist, we would expect $11.2 \pm 1.5$ events of signal in the inclusive

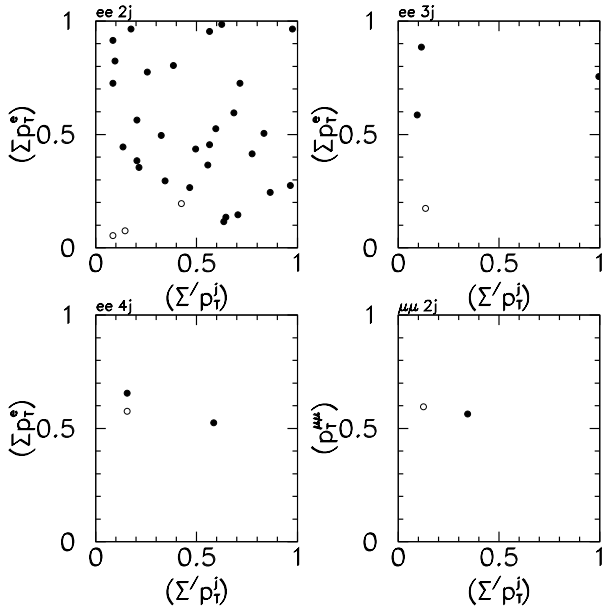



Figure 5.13: The positions of the transformed data points in the final states $ee\,2j$, $ee\,3j$, $ee\,4j$, and $\mu\mu\,2j$. The data points inside the region chosen by Sleuth are shown as filled circles; those outside the region are shown as open circles.

sample $ee\,2jX$, of which $5.9 \pm 0.8$ events would fall in the exclusive final state $ee\,2j$, on a background of $32 \pm 4$ events. Figure 5.15 shows the result of Sleuth applied to an ensemble of mock experiments in this final state. We see that Sleuth finds $\mathcal{P}$ larger than 3.5 standard deviations in over 80% of these mock samples.

## 5.7 $(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$

In this section we analyze the $(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$ final states. After describing the data sets and background estimates, we provide the results obtained by applying Sleuth to these channels. We conclude the section with a sensitivity check $[X' \to (\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X]$ that is more general in nature than those provided for the $e\mu X$, $W$+jets-like, and $Z$+jets-like

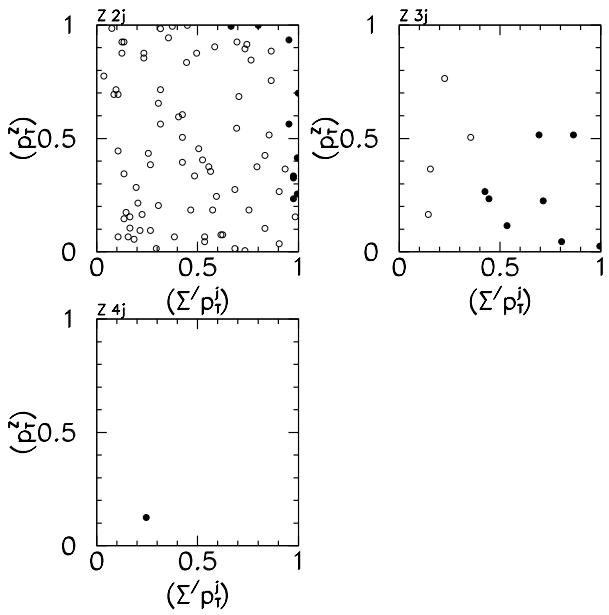



Figure 5.14: The positions of the transformed data points in the final states $Z\,2j$, $Z\,3j$, and $Z\,4j$. The data points inside the region chosen by Sleuth are shown as filled circles; those outside the region are shown as open circles.

final states above. Examples of a few of the many signals that might appear in these final states are provided in Section 5.3.5.

### 5.7.1 Data sets and background estimates

The $(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$ data set corresponds to an integrated luminosity of $123 \pm 7$ pb$^{-1}$. Global cleanup cuts are imposed as above. In this section we strictly adhere to standard particle identification criteria. All objects (electrons, photons, muons, and jets) are required to have transverse momentum $\geq$ 15 GeV, to be isolated, to be within the fiducial volume of the detector, and to be central. For electrons and photons the fiducial requirement is $|\eta_{\rm det}| < 1.1$ or $1.5 < |\eta_{\rm det}| < 2.5$; for muons it is $|\eta_{\rm det}| < 1.7$. For the case

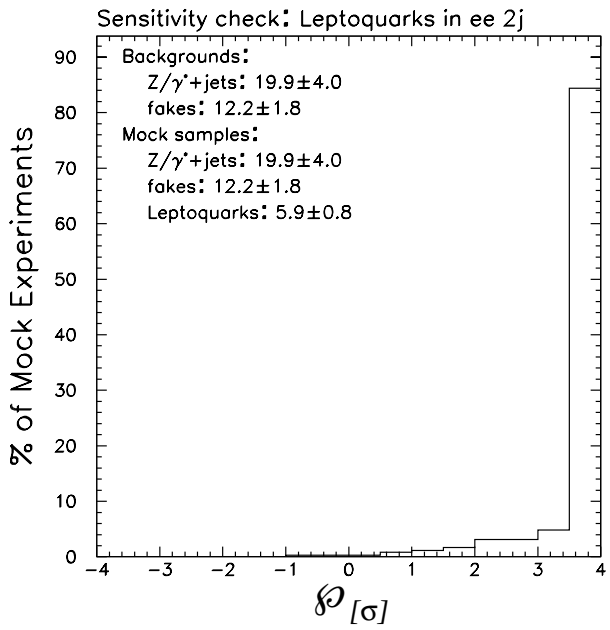



Figure 5.15: Histogram of $\mathcal{P}$ for an ensemble of mock experiments in which the backgrounds include $Z/\gamma^*$+jets and QCD fakes, and the mock samples include leptoquark pair production (with an assumed leptoquark mass of 170 GeV and $\beta = 1$) in addition to the expected background. All samples with $\mathcal{P} > 3.5\sigma$ are in the rightmost bin. Sleuth finds $\mathcal{P}$ larger than 3.5 standard deviations in over 80% of these mock samples.

of hadronic jets our centrality requirement of $|\eta| < 2.5$ is more stringent than the fiducial requirement of $|\eta_{\rm det}| \lesssim 4$. We require electrons, photons, and muons to be separated by at least 0.4 in $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$. $\not\!\!E_T$ is identified as an object if its magnitude is larger than 15 GeV. The selection of events is facilitated by use of the database described in Ref. [55].

We make frequent use of the (mis)identification probabilities determined for these identification criteria, which are summarized in Table 5.10.

| | $e$ | $\gamma$ |
|---|---|---|
| $e$ | $0.61 \pm 0.04$ [48] | $0.28 \pm 0.03$ [56] |
| $\gamma$ | $0.16 \pm 0.016$ [56] | $0.73 \pm 0.012$ [56] |
| $j$ | $0.00035 \pm 0.000035$ [48] | $0.00125 \pm 0.00013$ [48] |

Table 5.10: (Mis)identification probabilities. The number at (row $i$, column $j$) is the probability that the object labeling row $i$ will be reconstructed as the object labeling column $j$.

$ee\gamma X$

The dominant background to $ee\gamma X$ is the standard model process $Z/\gamma^*(\rightarrow ee)\gamma$. We use a matrix element Monte Carlo [57] to estimate this background. The $p\bar{p} \rightarrow Z/\gamma^*(\rightarrow ee)\gamma$ cross section, multiplied by our kinematic and geometric acceptance, is $0.50 \pm 0.05$ pb. From Table 5.10, the probability for two true electrons and one true photon to be reconstructed as two electrons and one photon is 0.33. From these numbers we estimate the expected background from this process into the $ee\gamma X$ final states to be $14.3 \pm 2.9$ events. Of these, $7.6 \pm 1.5$ events have $m_{ee} < 82$ GeV or $m_{ee} > 100$ GeV, and $82 < m_{ee\gamma} < 100$ GeV. Following the prescription in Sec. 4.2, such events are placed in the $Z$ final state, and are not considered in this section.

A smaller background in these final states is $Z$+jets production, with the jet faking a photon. From Ref. [47], we expect $1100 \pm 200$ $Z(\rightarrow ee)$+1 jet events in our data; the probability that this jet will fake a photon is given in Table 5.10. We therefore expect $0.99 \pm 0.27$ events of background in $Z\gamma$ from this source, and $0.13 \pm 0.04$ events (roughly 10% of the number expected in $Z\gamma$, determined by PYTHIA) in $ee\gamma$.

The dominant background to the $ee\gamma\not\!\!E_T$ final state comes from $W(\rightarrow e\nu)Z(\rightarrow ee)$, in which one of the three electrons is reconstructed as a photon. The $WZ$ production cross section in the standard model is calculated to be 2.5 pb [58]; DØ's geometric acceptance

for these events is determined using PYTHIA. Using the (mis)identification probabilities in Table 5.10, we estimate the contribution from standard model $WZ$ production to this final state to be $0.23 \pm 0.10$ events.

The numbers of expected background events in final states with additional jets are obtained by multiplying by a factor of 1/5 for each additional jet. The number of events expected in each final state, together with the number of events observed in the data, is given in Table 5.11. We find good agreement between the expected background and the numbers of events observed in the data.

| Final State | $Z\gamma$ | $Zj$ | $WZ$ | Total | Data |
|---|---|---|---|---|---|
| $Z\gamma$ | $3.3 \pm 0.7$ | $0.99 \pm 0.27$ | – | $4.3 \pm 0.7$ | 3 |
| $ee\gamma$ | $2.1 \pm 0.4$ | $0.13 \pm 0.04$ | – | $2.2 \pm 0.4$ | 1 |
| $Z\gamma j$ | $0.80 \pm 0.30$ | $0.23 \pm 0.06$ | – | $1.03 \pm 0.31$ | 1 |
| $ee\gamma j$ | $0.50 \pm 0.25$ | $0.033 \pm 0.009$ | – | $0.53 \pm 0.25$ | 0 |
| $ee\gamma \not\!\!E_T$ | $0.010 \pm 0.005$ | $0.024 \pm 0.007$ | $0.23 \pm 0.10$ | $0.26 \pm 0.10$ | 1 |

Table 5.11: Expected backgrounds for the $ee\gamma X$ final states.

$\mu\mu\gamma X$

The dominant background to the $\mu\mu\gamma X$ final states is standard model $Z/\gamma^*(\rightarrow \mu\mu)\gamma$. The matrix element Monte Carlo used to estimate the backgrounds to $ee\gamma X$ is also used for this final state. The normalization is determined by multiplying the number of expected $Z/\gamma^*(\rightarrow ee)\gamma$ events by the square of the ratio of efficiency $\times$ acceptance for muons and electrons. For muons, the efficiency $\times$ acceptance is roughly $0.5 \times 0.5$; for electrons, the number is approximately $0.6 \times 0.8$. The number of expected events in $\mu\mu\gamma$ is thus $3.9 \pm 0.9$. No events are seen in this final state. The probability of seeing zero events when $3.9 \pm 0.9$

are expected is 2.8%.

$e\gamma\gamma X$

The dominant background to $e\gamma\gamma X$ is the standard model process $Z/\gamma^*(\rightarrow ee)\gamma$, where one of the electrons is reconstructed as a photon. From Table 5.10 and the $Z(\rightarrow ee)\gamma$ estimate in Sec. 5.7.1, we determine the number of expected events in the $e\gamma\gamma$ final state to be $10.7 \pm 2.1$ events. Twelve $e\gamma\gamma X$ events are seen in the data, appearing in the final states shown in Table 5.12. We model the $e\gamma\gamma$ backgrounds with the Monte Carlo used for the $ee\gamma X$ final states above.

Three of the events in the $e\gamma\gamma j$ final state have $m_{e\gamma_1\gamma_2} = 95.8$ GeV, $m_{e\gamma_1\gamma_2} = 85.9$ GeV, and $m_{e\gamma_1} = 97.9$ GeV, respectively, and are consistent with $Z\gamma$ production with a radiated jet. The invariant masses of the objects in the fourth event all lie substantially outside the $Z$ boson mass window. Lacking an adequate $Z(\rightarrow ee)\gamma j$ Monte Carlo, we simply calculate the probability that the expected background fluctuates up to or above the observed number of events in this final state. The single event in the $e\gamma\gamma\, 2j$ final state has $m_{e\gamma_1\gamma_2} = 92.4$ GeV; this appears to be a $Z$ boson produced in association with two jets.

One event in this sample contains significant $\not\!\!E_T$ in addition to one electron and two photons. In this event $m_{e\gamma_1} = 95.9$ GeV, but the missing transverse energy in the event is large, and directly opposite the electron in $\phi$. The transverse mass $m_T^{e\nu} = 71.9$ GeV, so this event falls in the $W\gamma\gamma$ final state. The dominant background to this final state is $W(\rightarrow e\nu)Z(\rightarrow ee)$, in which two electrons are reconstructed as photons; the number of such events expected in this final state is determined to be $0.11 \pm 0.05$. $W(\rightarrow e\nu)\gamma\gamma$ is a slightly smaller but comparable background to this final state, which we estimate using a

matrix element Monte Carlo [59]. The total cross section for $W(\to e\nu)\gamma\gamma$ with all three detected objects in the fiducial region of the detector and $\not\!\!E_T > 15$ GeV is determined to be $0.77 \pm 0.08$ fb. The number of $W(\to e\nu)\gamma\gamma$ events in our data is therefore expected to be $0.026 \pm 0.010$. Backgrounds from $W\gamma j$ and $W\,2j$, where the jets fake photons, are comparable but smaller. This event will be combined in the next section with any events containing one muon and two photons to form the $W\gamma\gamma$ final state.

| Final state | Bkg | Data |
|---|---|---|
| $e\gamma\gamma$ | $10.7 \pm 2.1$ | 6 |
| $W(\to e\nu)\gamma\gamma$ | $0.14 \pm 0.05$ | 1 |
| $e\gamma\gamma j$ | $2.3 \pm 0.7$ | 4 |
| $e\gamma\gamma\,2j$ | $0.37 \pm 0.15$ | 1 |

Table 5.12: Population of final states within $e\gamma\gamma X$.

$\mu\not\!\!E_T\gamma\gamma X$

The dominant backgrounds to the $\mu\not\!\!E_T\gamma\gamma X$ final states, like those from the $e\not\!\!E_T\gamma\gamma X$ final states, come from $WZ$ and from a $W$ boson produced in association with two photons. The number of expected events from $WZ$ is determined as above to be $0.05 \pm 0.02$. The background from standard model $W\gamma\gamma$ is estimated by multiplying the number of expected $W(\to e\nu)\gamma\gamma$ events above by the ratio of efficiency $\times$ acceptance for electrons and muons.

Adding the number of events expected from $W(\to e\nu)\gamma\gamma$ to the number of events expected from $W(\to \mu\nu)\gamma\gamma$, we find the total number of expected background events in the $W\gamma\gamma$ final state to be $0.21 \pm 0.08$. No events are seen in the muon channel, so the only event in this final state is the event in the electron channel described above.

$\gamma\gamma\gamma X$

The dominant background to $\gamma\gamma\gamma$ is the standard model process $Z/\gamma^*(\rightarrow ee)\gamma$, where both of the electrons are reconstructed as photons. Taking the probability of an electron faking a photon from Table 5.10 and using the number of $Z/\gamma^*(\rightarrow ee)\gamma$ events determined above, we find the number of expected events in this final state from this process to be $2.5 \pm 0.5$ events. The contributions from $3j$, $\gamma\, 2j$, and $\gamma\gamma j$ are smaller by an order of magnitude.

Two events are seen in the data, both in the final state $\gamma\gamma\gamma$. One of these events has a three-body invariant mass $m_{\gamma\gamma\gamma} = 100.4$ GeV, consistent with the expectation that it is truly a $Z\gamma$ event. The other has a three-body invariant mass $m_{\gamma\gamma\gamma} = 153$ GeV, but two photons may be chosen whose two-body invariant mass is $m_{\gamma\gamma} = 90.3$ GeV. This event also appears to fit the $Z\gamma$ hypothesis.

$eeeX$

The dominant background to the final state $eee$ is again $Z/\gamma^*(\rightarrow ee)\gamma$, where this time the photon is reconstructed as an electron. The cross section quoted above for $Z/\gamma^*(\rightarrow ee)\gamma$, folded with the (mis)identification probabilities from Table 5.10, predicts $2.6 \pm 1.0$ events expected in the final state $eee$. One event is seen in the data. The $eee$ invariant mass in this event is 87.6 GeV, consistent with the standard model process $Z/\gamma^*(\rightarrow ee)\gamma$, where the photon is reconstructed as an electron.

$\mu\mu\mu X$

The dominant background to $\mu\mu\mu$ is standard model $WZ$ production. We use the $WZ$ production cross section above and take our efficiency $\times$ acceptance for picking up all three muons in the event to be roughly $(0.5 \times 0.5)^3 = 0.02$. The total number of expected background events in $\mu\mu\mu$ from $WZ$ production is thus $0.020 \pm 0.010$ events. Zero events are seen in the data.

The only populated final states within $\gamma\gamma\gamma X$, $eeeX$, and $\mu\mu\mu X$ are $\gamma\gamma\gamma$ and $eee$; these are summarized in Table 5.13.

| Final state | Bkg | Data |
|---|---|---|
| $\gamma\gamma\gamma$ | $2.5 \pm 0.5$ | 2 |
| $eee$ | $2.6 \pm 1.0$ | 1 |

Table 5.13: Population of final states with three like objects.

### 5.7.2 Results

Having estimated the backgrounds to each of these final states, we proceed to apply Sleuth to the data. Large $\mathcal{P}$'s are determined for all final states, indicating no hints of new physics within $(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$. Table 5.14 summarizes the results. We note that *all* final states within $(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$ have been analyzed, including (for example) $ee\gamma\gamma\not{E}_T$ and $\mu\mu\gamma\gamma\, 2j$. All final states within $(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$ not listed in Table 5.14 are unpopulated, and have $\mathcal{P} = 1.00$.

| Data set | $\mathcal{P}$ |
|---|---|
| $\gamma\gamma\gamma$ | 0.41 |
| $eee$ | 0.89 |
| $Z\gamma$ | 0.84 |
| $Z\gamma j$ | 0.63 |
| $ee\gamma$ | 0.88 |
| $ee\gamma \not\!\!E_T$ | 0.23 |
| $e\gamma\gamma$ | 0.66 |
| $e\gamma\gamma j$ | 0.21 |
| $e\gamma\gamma\, 2j$ | 0.30 |
| $W\gamma\gamma$ | 0.18 |

Table 5.14: Summary of results on the $(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$ final states.

### 5.7.3 Sensitivity check: $X' \to (\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$

The backgrounds to the $(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$ final states are sufficiently small that a signal present even at the level of one or two events can be significant. Due to the variety of final states treated in this section and the many processes that could produce signals in one or more of these final states, our sensitivity check for this section is the general process $X' \to (\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$, rather than a specific process such as $p\bar{p} \to \tilde{\chi}_2^0\tilde{\chi}_1^\pm \to \ell\ell\ell' \not\!\!E_T$. We (pessimistically) take the kinematics of the final state particles to be identical to the kinematics of the standard model background. In reality the final state objects in the signal are expected to have significantly larger momenta than those in the backgrounds, and the calculated $\mathcal{P}$ will be correspondingly smaller. With this minimal assumption about the kinematics of the signal, the details of the Sleuth algorithm are irrelevant, and $\mathcal{P}$ is given on average by the probability that the background fluctuates up to or above the number of expected background events plus the number of expected signal events.

The quantity $\tilde{\mathcal{P}}$ obtained by combining the $\mathcal{P}$'s calculated in all final states is a very different measure of "significance" than the measure familiar to most high energy physicists.

The fact that a "significance" of five standard deviations is unofficially but generally accepted as the threshold for a discovery results from a rough collective accounting of the number of different places such an effect could appear. We can better understand this accounting by first noting that five standard deviations corresponds to a (one-sided) probability of $3 \times 10^{-7}$. We then estimate that there are *at least* $5 \times 10^3$ distinct regions in the many variable spaces that are considered in a multipurpose experiment such as DØ in which one could realistically claim to see a signal. A probability of $1.5 \times 10^{-3}$, in turn, corresponds to three standard deviations. We can therefore understand the desire for a "$5\sigma$ effect" in our field to really be a desire for a "$3\sigma$ effect" (one time in one thousand), after a rigorous accounting for the number of places that such an effect might appear.

One of the advantages of Sleuth is that this rigorous accounting is explicitly performed. The final output of Sleuth takes the form of single number, $\tilde{\mathcal{P}}$, which is "the fraction of hypothetical similar experimental runs in which you would see something as interesting as what you actually saw in the data." The discussion in the preceding paragraph suggests that finding $\tilde{\mathcal{P}} \geq 3\sigma$ is as improbable (if not more so) as finding a "$5\sigma$ effect."

The number of final states that we consider, together with the number of background events expected in each, defines the mapping between $\mathcal{P}_{\rm min}$ (the smallest $\mathcal{P}$ found in any final state) and $\tilde{\mathcal{P}}$. For the final states that we have considered in this article, this mapping is shown in Fig. 5.16. We see that finding $\tilde{\mathcal{P}} \geq 3\sigma$ requires finding $\mathcal{P} \geq 4.2\sigma$ in some final state.

Let $N_Y$ be the smallest integer for which the probability that the background in the final state $Y$ fluctuates up to or above the expected background $\hat{b}$ plus $N_Y$ is $\leq 1.5 \times 10^{-5}$

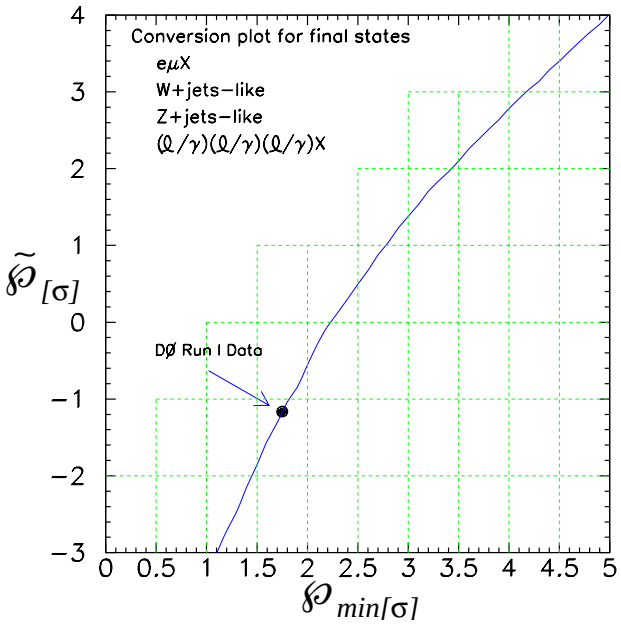



Figure 5.16: Correspondence between $\mathcal{P}_{\rm min}$ and $\tilde{\mathcal{P}}$ for the final states we have considered.

($4.2\sigma$). This is the number of events which, if observed in $Y$, would correspond to a discovery. This number can be related to the most probable cross section $\sigma_\vartheta$ of the new process $\vartheta$ into the final state $Y$ through

$$\sigma_\vartheta = \frac{N_Y}{a_\vartheta \epsilon_Y \mathcal{L}}, \tag{5.4}$$

where $a_\vartheta$ are the appropriate kinematic and geometric acceptance factors for the process $\vartheta$ and the DØ detector, $\epsilon_Y$ is the probability that the objects in the true final state $Y$ will be correctly reconstructed (which can be determined using Table 5.10), and $\mathcal{L} \approx 85$ pb$^{-1}$ is the effective luminosity of the DØ data after application of global cleanup cuts. The numbers $N_Y$ for some of the final states within $(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$ are given in Table 5.15. (These final states are all unpopulated in the DØ data.) Even with our pessimistic assumptions, using the

Sleuth strategy but setting aside the sophisticated Sleuth algorithm, we see that a discovery could have been made had even a few signal events populated one of these channels.

| Final State | $\hat{b}$ | $N$ |
|---|---|---|
| $ee\gamma j\not{E}_T$ | $0.059 \pm 0.020$ | 4 |
| $ee\gamma\, 2j$ | $0.10 \pm 0.05$ | 4 |
| $Z\gamma\, 2j$ | $0.13 \pm 0.05$ | 5 |
| $Z\gamma\, 3j$ | $0.025 \pm 0.010$ | 3 |
| $Z\gamma\, 4j$ | $0.0049 \pm 0.0020$ | 3 |
| $ee\mu\not{E}_T$ | $0.10 \pm 0.05$ | 4 |
| $e\mu\mu$ | $0.040 \pm 0.020$ | 4 |
| $\mu\mu\mu$ | $0.020 \pm 0.010$ | 3 |
| $W\gamma\gamma$ | $0.21 \pm 0.08$ | 5 |

Table 5.15: The number of signal events $N$ required in some of the final states within $(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$ in order to find $\tilde{\mathcal{P}} \geq 3\sigma$ (see the discussion in the text). This number is pessimistic, as it assumes that the signal is distributed identically to the backgrounds in the variables of interest. Most tenable models predict events containing final state objects that are significantly more energetic than the backgrounds, and in this case $N$ decreases accordingly.

## 5.8 Summary

Table 5.16 summarizes the values of $\mathcal{P}$ obtained for all populated final states analyzed in this article. Taking into account the many final states (both populated and unpopulated) that have been considered in this analysis, we find $\tilde{\mathcal{P}} = 0.89$ $(-1.23\sigma)$. Figure 5.17 shows a histogram of the $\mathcal{P}$'s computed for the populated final states analyzed in this article, together with the distribution expected from a simulation of many mock experimental runs. Good agreement is observed.

Although no statistically significant indications of new physics are observed in this analysis, some final states appear to hold greater promise than others. The smallest $\mathcal{P}$'s (0.04 and 0.06) are found in the final states $ee\, 4j$ and $ee\not{E}_T\, 4j$. The kinematics of the events

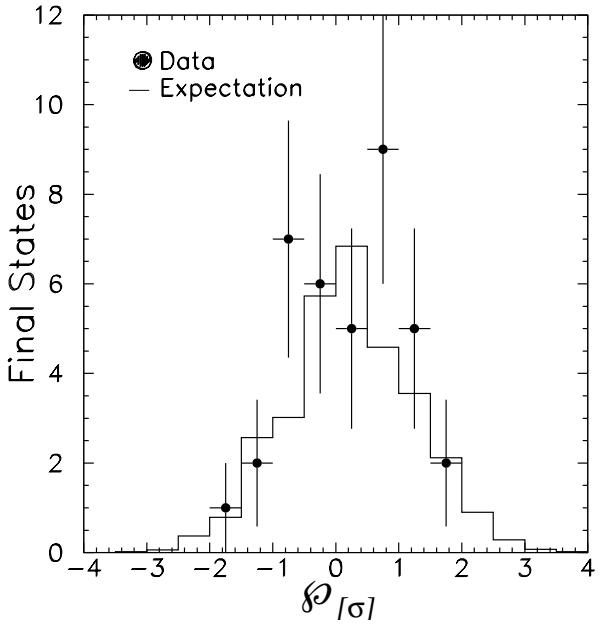



Figure 5.17: Histogram of the $\mathcal{P}$'s computed for the populated final states considered in this article. The distribution agrees well with the expectation.

in these final states are provided in Appendix B.2.

It is very difficult to quantify the sensitivity of Sleuth to arbitrary new physics, since the sensitivity necessarily depends on the characteristics of that new physics. We have provided examples of Sleuth's performance on "typical," particular signatures. This function is served by the sensitivity checks provided at the end of each of Secs. 5.5–5.7. In the analysis of the $e\mu X$ data in Ref. [16], our signal was first $WW$ and $t\bar{t}$ together, and then only $t\bar{t}$. This was a difficult signal to find, for although both $WW$ and $t\bar{t}$ cluster in the upper right-hand corner of the unit box, as desired, we expect only 3.9 $WW$ events in $e\mu\not\!\!E_T$ (with a background of 45.6 events), and 1.8 $t\bar{t}$ events in $e\mu\not\!\!E_T\,2j$ (with a background of 3.4 events). We were able to consistently find indications of the presence of $WW$ and $t\bar{t}$ in an ensemble

of mock experiments, but we would not have been sufficiently sensitive to claim a discovery.

In the $W$+jets-like final states we again chose $t\bar{t}$ for our sensitivity check. This was both a natural sequel to the sensitivity check in $e\mu X$ and a test of Sleuth's performance when the signal populates the high tails of only a subset of the variables considered. We find $\mathcal{P}_{\rm min} > 3\sigma$ in 30% of an ensemble of mock experimental runs containing $t\bar{t}$ events on the final states $W\,3j$, $W\,4j$, $W\,5j$, and $W\,6j$, compared with only 0.5% of an ensemble of mock experimental runs containing background only.

In the $Z$+jets-like final states we considered a leptoquark signal. This is in many ways an ideal signature — a relatively large number of events (about six) are predicted, and the signal appears in the high tails of both variables under consideration. Sleuth finds $\mathcal{P} > 3.5\sigma$ in over 80% of the mock experiments performed.

Finally, in the final states $(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$ we introduced the mapping between $\mathcal{P}_{\rm min}$ and $\tilde{\mathcal{P}}$ and briefly discussed its interpretation. The generic sensitivity check we considered $[X' \to (\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X]$ demonstrates the advantages of considering exclusive final states. While the other sensitivity checks rely heavily upon the Sleuth algorithm, this check shows that a careful and systematic definition of final states by itself can lead to a discovery with only a few events.

## 5.9 Conclusions

We have applied the Sleuth algorithm to search for new high $p_T$ physics in data spanning over thirty-two exclusive final states collected by the DØ experiment during Run I of the Fermilab Tevatron. A quasi-model-independent, systematic search of these data has

produced no evidence of physics beyond the standard model.

| Data set | $\mathcal{P}$ |
|---|---|
| $e\mu X$ | |
| $e\mu \not\!\!E_T$ | 0.14 $(+1.08\sigma)$ |
| $e\mu \not\!\!E_T j$ | 0.45 $(+0.13\sigma)$ |
| $e\mu \not\!\!E_T\, 2j$ | 0.31 $(+0.50\sigma)$ |
| $e\mu \not\!\!E_T\, 3j$ | 0.71 $(-0.55\sigma)$ |
| $W$+jets-like | |
| $W\, 2j$ | 0.29 $(+0.55\sigma)$ |
| $W\, 3j$ | 0.23 $(+0.74\sigma)$ |
| $W\, 4j$ | 0.53 $(-0.08\sigma)$ |
| $W\, 5j$ | 0.81 $(-0.88\sigma)$ |
| $W\, 6j$ | 0.22 $(+0.77\sigma)$ |
| $e \not\!\!E_T\, 2j$ | 0.76 $(-0.71\sigma)$ |
| $e \not\!\!E_T\, 3j$ | 0.17 $(+0.95\sigma)$ |
| $e \not\!\!E_T\, 4j$ | 0.13 $(+1.13\sigma)$ |
| $Z$+jets-like | |
| $Z\, 2j$ | 0.52 $(-0.05\sigma)$ |
| $Z\, 3j$ | 0.71 $(-0.55\sigma)$ |
| $Z\, 4j$ | 0.83 $(-0.95\sigma)$ |
| $ee\, 2j$ | 0.72 $(-0.58\sigma)$ |
| $ee\, 3j$ | 0.61 $(-0.28\sigma)$ |
| $ee\, 4j$ | 0.04 $(+1.75\sigma)$ |
| $ee \not\!\!E_T\, 2j$ | 0.68 $(-0.47\sigma)$ |
| $ee \not\!\!E_T\, 3j$ | 0.36 $(+0.36\sigma)$ |
| $ee \not\!\!E_T\, 4j$ | 0.06 $(+1.55\sigma)$ |
| $\mu\mu\, 2j$ | 0.08 $(+1.41\sigma)$ |
| $(\ell/\gamma)(\ell/\gamma)(\ell/\gamma)X$ | |
| $eee$ | 0.89 $(-1.23\sigma)$ |
| $Z\gamma$ | 0.84 $(-0.99\sigma)$ |
| $Z\gamma j$ | 0.63 $(-0.33\sigma)$ |
| $ee\gamma$ | 0.88 $(-1.17\sigma)$ |
| $ee\gamma \not\!\!E_T$ | 0.23 $(+0.74\sigma)$ |
| $e\gamma\gamma$ | 0.66 $(-0.41\sigma)$ |
| $e\gamma\gamma j$ | 0.21 $(+0.81\sigma)$ |
| $e\gamma\gamma\, 2j$ | 0.30 $(+0.52\sigma)$ |
| $W\gamma\gamma$ | 0.18 $(+0.92\sigma)$ |
| $\gamma\gamma\gamma$ | 0.41 $(+0.23\sigma)$ |
| $\tilde{\mathcal{P}}$ | 0.89 $(-1.23\sigma)$ |

Table 5.16: Summary of results for populated final states. The most interesting final state is found to be $ee\, 4j$, with $\mathcal{P} = 0.04$. Upon taking into account the many final states we have considered using the curve in Fig. 5.16, we find $\tilde{\mathcal{P}} = 0.89$. The values of $\mathcal{P}$ obtained in these final states are histogrammed in Fig. 5.17, and compared to the distribution we expect from an ensemble of mock experimental runs. No evidence for new high $p_T$ physics is observed in these data.

# Chapter 6

# Conclusions

We have tackled the important question of whether DØ's Run I data contains any evidence of physics beyond the standard model, and have come up empty-handed. While the result is disappointing, we are now able to make a fairly sweeping statement about the presence of new high $p_T$ physics in the Run I DØ data.

DØ and CDF will begin to collect data again in the spring of 2001 (Run II) with improved detectors and an accelerator providing increased luminosity and slightly higher collision energies. The analysis tools that we have developed here may prove to be very useful for discovering the presence of new physics in this upcoming run.

# Bibliography

[1] D. E. Groom *et al.*, The European Physical Journal C **15**, 1 (2000).

[2] See for example H. Haber and G. Kane, Phys. Rept. **117**, 75 (1985); H. Nilles, Phys. Rept. **110**, 1 (1984); P. Nath *et al.*, *Applied* $N = 1$ *Supergravity*, (World Scientific 1984); X. Tata, The Standard Model and Beyond, ed. J. Kim (World Scientific, 1991).

[3] S. Weinberg, Phys. Rev. D **13**, 974 (1976); ibid. **19**, 1277 (1979); L. Susskind, Phys. Rev. D **20**, 2619 (1979); S. Dimopoulos and L. Susskind, Nuc. Phys. B **155**, 237 (1979); E. Eichten and K. Lane, Phys. Lett. B **90**, 125 (1980).

[4] N. Arkani-Hamed, S. Dimopoulos, and G. Dvali, Phys. Lett. B **429**, 263 (1998).

[5] DØ Collaboration, S. Abachi *et al.*, Nucl. Instr. and Methods **A338**, 185 (1994).

[6] G. Gomez, Ph.D. thesis, University of Maryland, 1999 (unpublished).

[7] L. Babukhadia, Ph.D. thesis, University of Arizona, 1999 (unpublished).

[8] M. Goosens *et al.*, CERN Library Program No. Q100 (1991).

[9] I. Adam *et al.*, DØ internal note 2327 (1994).

[10] R. Brun *et al.*, CERN Program Library Long Writeup Q121, 1989.

[11] R. Kehoe, DØ internal notes 2052 (1994) and 2908 (1996).

[12] J. Cochran, Ph.D. Thesis, State University of New York at Stony Brook, 1993; H. Singh, Ph.D. thesis, University of California at Riverside, 1999. D. Toback *et al.*, DØ internal note 3700 (1999).

[13] P. Bhat *et al.*, DØ internal notes 3208 and 3308 (1997).

[14] D. Karmgard, Ph.D. Thesis, Florida State University, 1999; D. Karmgard, DØ internal notes 3545 and 3518 (1999).

[15] R. Brun and F. Carminati, CERN Program Library Long Writeup W5013, 1993.

[16] DØ Collaboration, B. Abbott *et al.*, Phys. Rev. D **62**, 92004 (2000).

[17] L. Hall, J. Lykken and S. Weinberg, Phys. Rev. D **27**, 2359 (1983).

[18] J. Gunion, H. E. Haber, G. L. Kane, and S. Dawson, *The Higgs Hunter's Guide* (Addison-Wesley, Redwood City, 1990).

[19] M. Hohlmann, Ph. D. thesis, University of Chicago (1997).

[20] R. M. Barnett and L. J. Hall, Phys. Rev. Lett. **77**, 3506 (1996).

[21] P. Nath and R. Arnowitt, Mod. Phys. Lett. A **2**, 331 (1987); R. Barbieri *et al.*, Nucl. Phys. B **367**, 28 (1991); H. Baer and X. Tata, Phys. Rev. D **47**, 2739 (1993); J. Lopez *et al.*, Phys. Rev. D **48**, 2062 (1993).

[22] G. Valencia and S. Willenbrock, Phys. Rev. D **50**, 6843 (1994).

[23] DØ Collaboration, S. Abachi *et al.*, Phys. Rev. Lett. **79**, 1197 (1997).

[24] DØ Collaboration, S. Abachi *et al.*, Phys. Rev. Lett. **79**, 1203 (1997).

[25] DØ Collaboration, S. Abachi *et al.*, Phys. Rev. D **52**, 4877 (1995).

[26] DØ Collaboration, S. Abachi *et al.*, Phys. Rev. Lett. **78**, 2070 (1997).

[27] K. Clarkson, http://cm.bell-labs.com/netlib/voronoi/hull.html (1996).

[28] G. Marchesini *et al.*, hep-ph/9607393, 1996; G. Marchesini *et al.*, Comp. Phys. Comm. **67**, 465 (1992). We used v5.7.

[29] T. Sjöstrand, Comp. Phys. Comm. **82**, 74 (1994). We used v5.7.

[30] F. Paige and S. Protopopescu, BNL Report No. 38304 1986. We used v7.22 with CTEQ2L parton distribution functions.

[31] J. McKinley, Ph.D. thesis, Michigan State University, 1996.

[32] DØ Collaboration, B. Abbott *et al.*, Phys. Rev. D **61**, 072001 (2000).

[33] E. Laenen, J. Smith and W. van Neerven, Phys. Lett. B **321**, 254 (1994).

[34] J. Ohnemus, Phys. Rev. D **44**, 1403 (1991).

[35] DØ Collaboration, S. Abachi *et al.*, Phys. Rev. D **58**, 052001 (1998).

[36] H. Singh, Ph.D. thesis, University of California at Riverside, 1999.

[37] DØ Collaboration, S. Abachi *et al.*, Phys. Rev. Lett. **74**, 2632 (1995).

[38] DØ Collaboration, B. Abbott *et al.*, in preparation.

[39] DØ Collaboration, B. Abbott *et al.*, Phys. Rev. Lett. **82**, 2457 (1999).

[40] DØ Collaboration, B. Abbott *et al.*, Phys. Rev. Lett. **80**, 666 (1998).

[41] DØ Collaboration, S. Abachi *et al.*, Phys. Rev. Lett. **76**, 3271 (1996).

[42] DØ Collaboration, B. Abbott *et al.*, submitted to Phys. Rev. Lett., hep-ex/0010026 (2000).

[43] DØ Collaboration, S. Abachi *et al.*, Phys. Rev. Lett. **78**, 3634 (1997).

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

[45] DØ Collaboration, B. Abbott *et al.*, Phys. Rev. Lett. **82**, 4769 (1999)

[46] DØ Collaboration, S. Abachi *et al.*, Phys. Lett. B **385**, 471 (1996).

[47] DØ Collaboration, B. Abbott *et al.*, Phys. Rev. D **61**, 032004 (2000); DØ Collaboration, B. Abbott *et al.*, Phys. Rev. Lett. **84**, 2792 (2000).

[48] DØ Collaboration, B. Abbott *et al.*, Phys. Rev. Lett. **80**, 2051 (1998).

[49] We distinguish between "physics $\eta$," defined as $\eta = -\ln\tan(\theta/2)$ where $\theta$ is the polar angle as measured from the interaction point, and "detector $\eta$," defined as $\eta_d = -\ln\tan(\theta_d/2)$ where $\theta_d$ is the polar angle as measured from the center of the detector.

[50] F. A. Berends *et al.*, Nucl. Phys. B **357**, 32 (1991); we used v3.0.

[51] DØ Collaboration, B. Abbott *et al.*, Phys. Rev. Lett. **79**, 4321 (1997).

[52] DØ Collaboration, B. Abbott *et al.*, Phys. Rev. Lett. **83**, 2896 (1999).

[53] DØ Collaboration, B. Abbott *et al.*, Phys. Rev. Lett. **84**, 2088 (2000).

[54] J. C. Pati and A. Salam, Phys. Rev. D **19**, 275 (1974); H. Georgi and S. Glashow, Phys. Rev. Lett. **32**, 438 (1974); also, see J. L. Hewett and T. G. Rizzo, Phys. Rep. **183**, 193 (1989) and references therein.

[55] M. Bowen, G. Landsberg, and R. Partridge, in *Proceedings of the International Conference on Computing in High Energy Physics (CHEP 2000)*, edited by M. Mazzucato, (INFN, Padova, 2000), p. 478.

[56] DØ Collaboration, B. Abbott *et al.*, Phys. Rev. Lett. **81**, 524 (1998).

[57] U. Baur and E. L. Berger, Phys. Rev. D **47**, 4889 (1993).

[58] DØ Collaboration, S. Abachi *et al.*, Phys. Rev. Lett. **77**, 3303 (1996).

[59] U. Baur and T. Stelzer, Phys. Rev. D **61**, 073007 (2000).

[60] L. Holmström, S. R. Sain, and H. E. Miettinen, Comp. Phys. Commun. **88**, 195 (1995).

[61] D. Scott. *Multivariate Density Estimation.* John Wiley & Sons, 1992.

[62] M. Wand and M. Jones. *Kernel Smoothing.* Chapman & Hall, 1995.

[63] G. DiLoreto and F. Nang. The Triple Differential Dijet Cross Section at $\sqrt{s} = 1.8$ TeV. http://msupc3.fnal.gov/gian/results/triple/triple_note_8-14-98.ps, 1998.

[64] DØ Collaboration, S. Abachi *et al.*, Phys. Rev. D **56**, 6742 (1997).

[65] I am grateful to Bob Olivier (LPNHE, Paris) for his help with Figs. D.4–D.6.

[66] A.L. Lyon, Ph.D. Thesis, University of Maryland, 1997.

[67] Marvin Johnson and Joey Thompson, private communication (1998).

# Appendix A

# Sleuth details

## A.1 Further comments on variables

We have excluded a number of "standard" variables from the list in Table 4.1 for various reasons: some are helpful for specific models but not helpful in general; some are partially redundant with variables already on the list; some we have omitted because we felt they were less well-motivated than the variables on the list, and we wish to keep the list of variables short. Two of the perhaps most significant omissions are invariant masses and topological variables.

- Invariant masses: If a particle of mass $m$ is produced and its decay products are known, then the invariant mass of those decay products is an obvious variable to consider. $M_{\ell\nu}^T$ and $M_{\ell^+\ell^-}$ are used in this spirit to identify $W$ and $Z$ bosons, respectively, as described in Sec. 4.2. Unfortunately, a non-standard-model particle's decay products are generally not known, both because the particle itself is not known and because of

final state combinatorics, and resolution effects can wash out a mass peak unless one knows where to look. Invariant masses turn out to be remarkably ineffective for the type of general search we wish to perform. For example, a natural invariant mass to consider in $e\mu\not{E}_T jj$ is the invariant mass of the two jets ($m_{jj}$); since top quark events do not cluster in this variable, they would not be discovered by its use. A search for any *particular* new particle with known decay products is best done with a dedicated analysis. For these reasons the list of variables in Table 4.1 does not include invariant masses.

- Shape variables: Thrust, sphericity, aplanarity, centrality, and other topological variables often prove to be good choices for model-specific searches, but new physics could appear in a variety of topologies. Many of the processes that could show up in these variables already populate the tails of the variables in Table 4.1. If a shape variable is included, the choice of that particular variable must be justified. We choose not to use topological variables, but we do require physics objects to be central (e.g., $|\eta_j| < 2.5$), to similar effect.

## A.2 Transformation of variables

The details of the variable transformation are most easily understood in one dimension, and for this we can consider again Fig. 4.1. It is easy to show that if the background distribution is described by the curve $b(x) = \frac{1}{5}e^{-x/5}$ and we let $y = 1 - e^{-x/5}$, then $y$ is distributed uniformly between 0 and 1. The situation is more complicated when the background is given to us as a set of Monte Carlo points that cannot be described by a simple

parameterization, and it is further complicated when these points live in several dimensions.

There is a unique solution to this problem in one dimension, but an infinity of solutions in two or more dimensions. Not all of these solutions are equally reasonable, however — there are two additional properties that the solution should have.

- Axes should map to axes. If the data live in a three-dimensional space in the octant with all coordinates positive, for example, then it is natural to map the coordinate axes to the axes of the box.

- Points that are near each other should map to points that are near each other, subject to the constraint that the resulting background probability distribution be flat within the unit box.

This somewhat abstract and not entirely well-posed problem is helped by considering an analogous physical problem:

> The height of the sand in a $d$-dimensional unit sandbox is given by the function $b(\vec{x})$, where $\vec{x}$ is a $d$-component vector. (The counting of dimensions is such that a physical sandbox has $d = 2$.) We take the $d$-dimensional lid of the sandbox and squash the sand flat. The result of this squashing is that a sand grain at position $\vec{x}$ has moved to a new position $\vec{y}$, and the new function $b'(\vec{y})$ describing the height of the sand is a constant. Given the function $b(\vec{x})$, determine the mapping $\vec{x} \to \vec{y}$.

For this analogy to help, the background first needs to be put "in the sandbox." Each of the background events must also have the same weight (the reason for this will become clear shortly). The background probability density is therefore estimated in the original variables using Probability Density Estimation [60], and $M$ events are sampled from this distribution.

These $M$ events are then put "into the sandbox" by transforming each variable (individually) into the interval $[0, 1]$. The new variable is given by

$$x_j \to x'_j = \frac{1}{M} \int_{-\infty}^{x_j} \sum_{i=1}^{M} \frac{1}{\sqrt{2\pi}\sigma_j h} \exp\left( -\frac{(t - \mu_{ij})^2}{2\sigma_j^2 h^2} \right) dt, \tag{A.1}$$

where $\mu_{ij}$ is the value of the $j^{th}$ variable for the $i^{th}$ background event, $\sigma_j$ is the standard deviation of the distribution in the $j^{th}$ variable, and $h = M^{-\frac{1}{d+4}}$, where $d$ is the dimensionality of the space.

The next step is to take these $M$ events and map each of them to a point on a uniform grid within the box. The previous paragraph defines a mapping from the original variables into the unit sandbox; this step defines a mapping from a lumpy distribution in the sandbox to a flat distribution. The mapping is continued to the entire space by interpolating between the sampled background events.

The mapping to the grid is done by first assigning each sampled background point to an arbitrary grid point. Each background point $i$ is some distance $d_{ij}$ away from the grid point $j$ with which it is paired. We then loop over pairs of background points $i$ and $i'$, which are associated with grid points $j$ and $j'$, and swap the associations (associate $i$ with $j'$ and $i'$ with $j$) if $\max(d_{ij}, d_{i'j'}) > \max(d_{i'j}, d_{ij'})$. This looping and swapping is continued until an equilibrium state is reached.

## A.3 Region criteria

In Sec. 4.3.2 we introduced the formal notion of *region criteria* — properties that we require a region to have for it to be considered by Sleuth. The two criteria that we have decided to impose in the analysis of the $e\mu X$ data are *Isolation* and *AntiCornerSphere*.

**Isolation** We want the region to include events that are very close to it. We define $\xi = \frac{1}{4} N_{\text{data}}^{-\frac{1}{d}}$ as a measure of the mean distance between data points in their transformed coordinates, and call a region *isolated* if there exist no data points outside the region that are closer than $\xi$ to a data point inside the region. We generalize this boolean criterion to the interval $[0, 1]$ by defining

$$c_R^{Isolation} = \min\left(1, \frac{\min|(\vec{x})^{\text{in}} - (\vec{x})^{\text{out}}|}{2\xi}\right), \tag{A.2}$$

where the minimum is taken over all pairwise combinations of data points with $(\vec{x})^{\text{in}}$ inside $R$ and $(\vec{x})^{\text{out}}$ outside $R$.

**AntiCornerSphere** One must be able to draw a sphere centered on the origin of the unit box containing all data events outside the region and no data events inside the region. This is useful if the signal is expected to lie in the upper right-hand corner of the unit box. We generalize this boolean criterion to the interval $[0, 1]$ as described in Sec. 4.3.2.

A number of other potentially useful region criteria may be imagined. Among those that we have considered are *Connectivity*, *Convexity*, *Peg*, and *Hyperplanes*. Although we present only the boolean forms of these criteria here, they may be generalized to the interval $[0, 1]$ by introducing the scale $\xi$ in the same spirit as above.

**Connectivity** We generally expect a discovery region to be one connected subspace in the variables we use, rather than several disconnected subspaces. Although one can posit cases in which the signal region is not connected (perhaps signal appears in the two regions $\eta > 2$ and $\eta < -2$), one should be able to easily avoid this with an appropriate choice of variables.

(In this example, we should use $|\eta|$ rather than $\eta$.) We defined the concept of neighboring data points in the discussion of regions in Sec. 4.3.2. A *connected region* is defined to be a region in which given any two points $a$ and $b$ within the region, there exists a list of points $p_1 = a, p_2, \ldots, p_{n-1}, p_n = b$ such that all the $p_i$ are in the region and $p_{i+1}$ is a neighbor of $p_i$.

**Convexity** We define a *non-convex* region as a region defined by a set of $N$ data points $P$, such that there exists a data point $\hat{\vec{p}}$ not within $P$ satisfying

$$\sum_{i=1}^{N} \vec{p}_i \lambda_i = \hat{\vec{p}} \tag{A.3}$$

$$\sum_{i} \lambda_i = 1 \tag{A.4}$$

$$\lambda_i \geq 0 \quad \forall\, i, \tag{A.5}$$

for suitably chosen $\lambda_i$, where $\vec{p}_i$ are the points within $P$. A convex region is then any region that is not non-convex; intuitively, a convex region is one that is "roundish," without protrusions or intrusions.

**Peg** We may want to consider only regions that live on the high tails of a distribution. More generally, we may want to only consider regions that contain one or more of $n$ specific points in variable space. Call this set of points $\tilde{x}_i$, where $i = 1, \ldots, n$. We transform these points exactly as we transformed the data in Sec. 4.3.2 to obtain a set of points $\tilde{y}_i$ that live in the unit box. A region $R$ is said to be *pegged* to these points if there exists at least one $i \in 1, \ldots, n$ such that the closest data point to $\tilde{y}_i$ lies within $R$.

**Hyperplanes** Connectivity and Convexity are criteria that require the region to be "reasonably-shaped," while Peg is designed to ensure that the region is "in a believable location." It is

possible, and may at times be desirable, to impose a criterion that judges both shape and location simultaneously. A region $R$ in a $d$-dimensional unit box is said to satisfy *Hyperplanes* if, for each data point $p$ inside $R$, one can draw a $(d-1)$-dimensional hyperplane through $p$ such that all data points on the side of the hyperplane containing the point $\vec{1}$ (the "upper right-hand corner of the unit box") are inside $R$.

More complicated region criteria may be built from combinations and variations of these and other basic elements.

## A.4 Search heuristic details

The heuristic Sleuth uses to search for the region of greatest excess may usefully be visualized as a set of rules for an amoeba to move within the unit box. We monitor the amoeba's progress by maintaining a list of the most interesting region of size $N$ (one for each $N$) that the amoeba has visited so far. At each state, the amoeba is the region under consideration, and the rules tell us what region to consider next.

The initial location and size of the amoeba is determined by the following rules for *seeding*:

1. If we have not yet searched this data set at all, the starting amoeba fills the entire box.

2. Otherwise, the amoeba starts out as the region around a single random point that has not yet inhabited a "small" region that we have considered so far. We consider a region $R$ to be small if adding or removing an individual point can have a sizeable effect on the $p_N^R$; in practice, a region is small if $N \lesssim 20$.

3. If there is no point that has not yet inhabited a small region that we have considered so far, the search is complete.

At each stage, the amoeba either *grows* or *shrinks*. It begins by attempting to grow. The rules for growth are:

1. Allow the amoeba to encompass a neighboring data point. Force it to encompass any other data points necessary to make the expanded amoeba satisfy all criteria. Check to see whether the $p_N^R$ of the expanded amoeba is less than the $p_N^R$ of the region on the list of the same size. If so, the amoeba has successfully grown, the list of the most interesting regions is updated, and the amoeba tries to grow again. If not, the amoeba shrinks back to its former size and repeats the same process using a different neighboring data point.

2. If the amoeba has tried all neighboring data points and has not successfully grown, it shrinks.

The rules for shrinking are:

1. Force the amoeba to relinquish the data point that owns the most background, subject to the requirement that the resulting shrunken amoeba be consistent with the criteria.

2. If the amoeba has shrunk out of existence or can shrink no further, we kill this amoeba and reseed.

The result of this process is a list of regions of length $N_{\text{data}}$ (one region for each $N$), such that the $N^{th}$ region in the list is the most interesting region of size $N$ found in the data set.

# Appendix B

# Analysis details

## B.1 $\not\!\!E_T$ significance

We determine the significance of any missing transverse energy in an event in the $Z$+jets-like final states by computing a probability density $p(\not\!\!E_T)$. This is a true probability density in the sense that, for a given event, the probability that the actual missing transverse energy in that event is between $\not\!\!E_T$ and $\not\!\!E_T + \delta\not\!\!E_T$ is given by $p(\not\!\!E_T)\delta\not\!\!E_T$. This density is computed with a Monte Carlo calculation. For each data event we generate an ensemble of events similar to the original but with the energies of the objects smeared according to their resolutions. Jets are smeared with a Gaussian of width $\sigma = 80\%\sqrt{E}$, and electrons are smeared with a Gaussian of width $\sigma = 20\%\sqrt{E}$ (a slight inflation of the measured resolution of $15\%\sqrt{E}$), where $E$ is the energy of the object in GeV. The component of the missing transverse energy $\not\!\!E_{T_a}$ along the direction of the original $\not\!\!E_T$ is recalculated for each smeared event, and the values that are obtained are histogrammed. The histogram is then smoothed,

and the likelihood

$$\mathcal{L}_{\not\!\!E_T} = \frac{p(\not\!\!E_{Ta})_{\max}}{p(\not\!\!E_{Ta} = 0)} \tag{B.1}$$

is calculated. Studies have shown that a cut of $\log_{10} \mathcal{L}_{\not\!\!E_T} > 3$ does an excellent job of retaining events with true $\not\!\!E_T$ while rejecting QCD background.

## B.2 Kinematics of interesting events

Table B.1 provides information about the events in the most interesting final states seen in the course of this analysis. Invariant masses of objects in these events are given in Table B.2.

| run:event | object | $p_T$ (GeV) | $\phi$ | $\eta$ |
|---|---|---|---|---|
| $ee\,4j$ | | | | |
| 85918:12437 | $e$ | 58.0 | 0.74 | −0.42 |
| | $e$ | 37.9 | 0.30 | −1.51 |
| | $j$ | 89.0 | 3.94 | −0.10 |
| | $j$ | 26.0 | 4.20 | −0.98 |
| | $j$ | 21.3 | 2.55 | −1.25 |
| | $j$ | 21.2 | 2.07 | 0.77 |
| 90278:31411 | $e$ | 53.1 | 4.15 | 0.00 |
| | $e$ | 33.6 | 0.28 | −1.85 |
| | $j$ | 80.2 | 0.78 | 1.24 |
| | $j$ | 39.9 | 4.46 | 1.81 |
| | $j$ | 34.0 | 2.94 | −1.55 |
| | $j$ | 24.2 | 2.92 | 0.05 |
| 92746:25962 | $e$ | 64.6 | 1.99 | 0.99 |
| | $e$ | 40.6 | 5.72 | 0.55 |
| | $j$ | 26.8 | 3.84 | −2.13 |
| | $j$ | 25.6 | 4.83 | 0.49 |
| | $j$ | 20.0 | 5.73 | −1.12 |
| | $j$ | 21.5 | 1.86 | 2.62 |
| $ee\not{E}_T\,4j$ | | | | |
| 89815:17253 | $e$ | 87.7 | 5.93 | 1.00 |
| | $e$ | 22.5 | 4.19 | 1.33 |
| | $\not{E}_T$ | 59.8 | 0.97 | – |
| | $j$ | 69.8 | 2.42 | −1.33 |
| | $j$ | 53.1 | 2.88 | 0.36 |
| | $j$ | 52.2 | 4.27 | −1.30 |
| | $j$ | 25.4 | 5.81 | −0.18 |

Table B.1: Kinematic properties of the most interesting events seen in this analysis.

| run:event | $m_{ee}$ | $m_T^{e\not{E}_T}$ | | $m_{4j}$ |
|---|---|---|---|---|
| $ee\,4j$ | | | | |
| 85918:12437 | 57.4 | | | 149 |
| 90278:31411 | 119.5 | | | 342 |
| 92746:25962 | 100.6 | | | 323 |
| $ee\not{E}_T\,4j$ | | | | |
| 89815:17253 | 69.4 | 89.0 | 73.3 | 239 |

Table B.2: Invariant masses (in units of GeV) of objects in the most interesting events seen in this analysis.

# Appendix C

# $\vec{\alpha}PDE$

This appendix introduces a new multivariate analysis technique for parameter estimation. We describe how to estimate probability densities nonparametrically [60], and then show how to obtain posterior distributions for the parameters of interest by a direct manipulation of these densities. We explain how to apply this procedure to an ensemble of events containing background.

Past DØ analyses have shown that multivariate methods frequently prove to be more powerful than univariate methods, and a non-parametric method has the advantage of not requiring one to make assumptions about the forms of probability distributions. We hope that those who complain about the "black-box" quality of neural networks will welcome the straightforward manipulation of probability densities used in this method, and the intuitive graphical interpretation that results. Because we construct and manipulate the probability densities directly, obtaining any additional statistical information — confidence intervals, for example — is a straightforward exercise.

We begin in Sec. C.1 by setting up a typical analysis problem that might be faced in Run II. In Sec. C.2 we give our recipe for solving this type of problem. Along the way we will describe the salient features of our method and discuss some of its possible advantages. A few useful details are discussed in Sec. C.3. We conclude in Sec. C.4 with thoughts on the application of this method to future discovery analyses.

## C.1 The problem

With the notable exception of the study of CP-violation in the B system, the next decade of collider physics will be focusing mainly on new particle searches. For heavy, exotic particles such as the $Z'$, leptoquarks, or even the standard model Higgs, a peak in an invariant mass distribution provides a very convincing signature. These searches motivated the development of this method, and they provide the motivation for the toy problem we consider in this paper: measuring the mass of the standard model Higgs, where the Higgs decays to $b\bar{b}$.

We will assume that two variables $\vec{x} = (x, y)$ have already been identified as good variables for this analysis. This pair might be the transverse energies of two energetic $b$-tagged jets; it might be the invariant mass of the two leading jets and a measure of the sphericity of the event; it might be the scalar sum of all jet transverse energies and the output of a neural network built with event-shape variables. To provide a check for systematic errors, the variables might be the instantaneous luminosity and the probability of having found an incorrect primary vertex. The point is simply that no special assumptions are made about the variables $\vec{x}$.

Nothing would prevent us from applying this same technique to determine the coupling of the Higgs to $b\bar{b}$, or to determine the setting of any other dial that theorists can turn to tweak predicted distributions. Given the plethora of supersymmetric dials, one can imagine the utility of having an analysis method that works for all of them (at least in principle). For now, however, the problem at hand is to concoct a recipe for determining the expected Higgs mass $\hat{m}$ given a data event with variable values $\vec{x}$.

## C.2 The recipe

Our goal was to construct a method that performs as well as (or better than) such popular algorithms as neural networks, but to keep the method simple enough that it reads like a recipe. The recipe we have constructed is given below.

### C.2.1 Specify $p(m)$

This method has its roots in Bayesian statistics, and as a result it has the advantage (disadvantage) of enabling (requiring) us to specify a function $p(m|\mathcal{I})$ that represents the *a priori* probability that nature has chosen the Higgs mass to be $m$. $\mathcal{I}$ here is used in standard Bayesian notation to represent all assumptions implicit in our specification of this *a priori* probability; $p(m|\mathcal{I})$ is known as the *prior distribution* for $m$. The basic assumptions contained in $\mathcal{I}$ will not change, so we drop it from here on, writing simply $p(m)$.

A natural choice, used when one strongly believes that the true mass must lie somewhere between $a$ and $b$ but has no reason to prefer any value within that range over any other, is to take $p(m) = \frac{1}{b-a}$ for $a < m < b$, and 0 elsewhere. Alternatively, one can

specify the prior to accommodate previous experimental results, which can be a useful way to combine analyses into a single definitive answer.

### C.2.2 Generate Monte Carlo events

In order to demonstrate sensitivity to the Higgs mass in this analysis, we need to know what Higgs events look like for assumed values of the Higgs mass. We therefore generate some large number $N$ of Monte Carlo events with Higgs masses $m$ pulled from the distribution $p(m)$ specified above. That is, the probability that we generate an event with a Higgs mass between $m$ and $m + \delta m$ is $p(m)\,\delta m$.[1] For each Monte Carlo event we calculate the two variables $\vec{x} = (x, y)$.

What we have done is very simple: a histogram in $(x, y, m)$ filled with the events we have generated approximates the *joint density* $p(x, y, m)$. This function has the property that, given an event in which a Higgs was produced and decayed to $b\bar{b}$, the probability that the Higgs mass was between $m$ and $m + \delta m$, the first variable between $x$ and $x + \delta x$, and the second variable between $y$ and $y + \delta y$, is simply $p(x, y, m)\,\delta x\,\delta y\,\delta m$.

### C.2.3 Construct a training array $T$

Each of the $N$ Monte Carlo events we generated above is characterized by three numbers: the value of $x$, the value of $y$, and the Higgs mass $m$. We label these $N$ Monte Carlo events with the index $i$ ($i = 1, .., N$), and label the three numbers corresponding to the $i^{\text{th}}$ event by $x_i$, $y_i$, and $m_i$. For this $i^{\text{th}}$ Monte Carlo event, we can then define the *event*

[1] Sec. C.3.2 provides an alternative to this approach, which may be necessary if practical constraints make generating this random sample of events difficult.

*vector*

$$\vec{v}_i = (\vec{x}_i, m_i), \tag{C.1}$$

and proceed to define the *training array* $T$ for the entire set of Monte Carlo events to be the matrix given by

$$T_{ij} = (\vec{v}_i)_j. \tag{C.2}$$

Here and below $i$ ranges from 1 to $N$ and indexes the Monte Carlo events, and $j$ goes from 1 to 3 and indexes the components of the event vector $\vec{v}$.

### C.2.4 Calculate the covariance matrix

Having defined the event vector $\vec{v}$, we can calculate the *mean event vector*

$$\langle \vec{v} \rangle = \frac{1}{N} \sum_{i=1}^{N} \vec{v}_i, \tag{C.3}$$

and can then construct the *training covariance matrix*,

$$\Sigma_{kl} = \frac{1}{N} \sum_{i=1}^{N} ((\vec{v}_i)_k - \langle \vec{v} \rangle_k)((\vec{v}_i)_l - \langle \vec{v} \rangle_l), \tag{C.4}$$

in the standard way. $\Sigma$ is a 3 by 3 symmetric matrix, $\Sigma_{21} = \mathrm{Cov}(x, y)$, and so on.

### C.2.5 Estimate the joint density $p(\vec{v})$

In Sec. C.2.2 we imagined filling a three-dimensional histogram in $\vec{v}$ with Monte Carlo events, and we recognized that the resulting histogram represents an estimation of a probability density. A well-known technique in multivariate statistics involves estimating a probability density not by filling a histogram, but rather by summing "bumps" of probability placed at each data point. A favorite "bump" of many statisticians is the multivariate Gaussian:

$$K(\vec{v}) = \frac{1}{(\sqrt{2\pi}h)^3\sqrt{\det(\Sigma)}} \exp\left(\frac{-\vec{v}^T\Sigma^{-1}\vec{v}}{2h^2}\right), \tag{C.5}$$

where $K$ is for *kernel* (the technical name for "bump"). The vector $\vec{v}$ is the same three-component vector defined above, and $\Sigma^{-1}$ is the matrix inverse of the training array covariance matrix $\Sigma$ defined above. The variable $h$ that we have introduced is known in the language of density estimation as a *smoothing parameter*; it controls the width of the bumps we place around each data point. Theoretical arguments suggest an optimal choice $h^* \approx N^{-1/(d+4)}$ for the smoothing parameter as a function of the number of data points $N$ and the dimensionality $d$ of the variable space.[2] For $N = 10^5$ and $d = 3$, the optimal choice for $h$ is roughly 0.20.

We may now obtain an estimate of the joint probability density $p(\vec{v})$ by simply summing kernels centered about each of the $N$ data points $\vec{v}_i$, so that

$$p(\vec{v}) = \frac{1}{N}\sum_{i=1}^{N} K(\vec{v} - \vec{v}_i). \tag{C.6}$$

### C.2.6 Compute $\hat{m}$

A simple application of Bayes' theorem now tells us everything we could possibly want to know. The most natural function to consider is the *posterior density* for $m$ – the probability that the Higgs mass is $m$ given that you have observed an event with variable values $\vec{x}$. Now the probability of obtaining both $\vec{x}$ and $m$ is equal to the probability of

[2] This expression for $h^*$ depends on assumptions about the probability density that we have not made explicit, and it is intended only to provide a ballpark figure. For the reader interested in the theory behind this expression, we suggest [61] and [62]. In practice, this parameter may be optimized for any set of Monte Carlo events by constructing and minimizing some appropriate error estimate $\chi(h)$.

obtaining $\vec{x}$ multiplied by the probability of obtaining $m$ given that you have obtained $\vec{x}$:

$$p(\vec{x}, m) = p(\vec{x})p(m|\vec{x}), \tag{C.7}$$

and the probability of obtaining $\vec{x}$ is given by integrating the probability of obtaining both $\vec{x}$ and $m$ over all values of $m$:

$$p(\vec{x}) = \int p(\vec{x}, m')\, dm'. \tag{C.8}$$

Putting these two equations together we derive the formula for the posterior density $p(m|\vec{x})$, which is the natural function to consider for analysis purposes, in terms of the the joint density $p(\vec{x}, m)$, which is the natural function to construct. This formula is simply

$$p(m|\vec{x}) = \frac{p(\vec{x}, m)}{\int p(\vec{x}, m')\, dm'}. \tag{C.9}$$

Our best estimate $\hat{m}$ for the mass of a single event is then the value of the mass that maximizes this posterior density, and the equation

$$p(\hat{m}|\vec{x}) = \max_{m} p(m|\vec{x}) \tag{C.10}$$

may be solved numerically for $\hat{m}$. Note that since the denominator of Eq. C.9 is independent of $m$, maximizing the posterior density $p(m|\vec{x})$ is equivalent to maximizing the joint density $p(\vec{x}, m)$, which we construct explicitly.

Sec. C.3.3 discusses other possible ways to use this posterior density.

## C.3 Details

In putting the discussion above into practice, the following details may prove useful.

### C.3.1 The general multivariate case

For pedagogical reasons we chose to introduce this method through a specific example: determining the mass $m$ of the Higgs boson from the two measured quantities $x$ and $y$. The price we have paid for this is that the formulae we have developed are specific to that example. In this section we provide the formulae for the general case.

In the general case, each Monte Carlo event will be characterized by $d$ numbers; of these, $d_1$ will be known variables (corresponding to $x$ and $y$ in the example given above), and $d_2$ will be parameters to be determined (corresponding to $m$ in that example). We denote the known variables by the $d_1$-dimensional vector $\vec{x}$, and the parameters to be determined by the $d_2$-dimensional vector $\vec{\alpha}$. The $d$-dimensional vector $\vec{v} = (\vec{x}, \vec{\alpha})$, and $d = d_1 + d_2$.

The $i^{\text{th}}$ Monte Carlo event is described by the event vector

$$\vec{v}_i = (\vec{x}_i, \vec{\alpha}_i), \tag{C.11}$$

and the entire Monte Carlo sample is described by the training array

$$T_{ij} = (\vec{v}_i)_j, \tag{C.12}$$

where now $j$ ranges from 1 to $d$. The mean event vector is

$$\langle \vec{v} \rangle = \frac{1}{N} \sum_{i=1}^{N} \vec{v}_i, \tag{C.13}$$

the training covariance matrix is

$$\Sigma_{kl} = \frac{1}{N} \sum_{i=1}^{N} ((\vec{v}_i)_k - \langle \vec{v} \rangle_k)((\vec{v}_i)_l - \langle \vec{v} \rangle_l), \tag{C.14}$$

as before, and the general multivariate Gaussian is given by

$$K(\vec{v}) = \frac{1}{(\sqrt{2\pi}h)^d \sqrt{\det(\Sigma)}} \exp\left(\frac{-\vec{v}^T \Sigma^{-1} \vec{v}}{2h^2}\right). \tag{C.15}$$

Finally, in Eqs. C.9 and C.10, $m$ should be replaced by the vector $\vec{\alpha}$.

### C.3.2 Alternative to generating a random sample of Monte Carlo events

In this appendix we describe a modification to the above procedure that is necessary if practical constraints prevent the generation of events with Higgs masses $m$, pulled from the distribution $p(m)$. We assume instead the possibility of generating events at $q$ different Higgs masses $M_j$, where $j = 1, .., q$.

Only two changes need to be made in the first five steps of the recipe (Secs. C.2.1 to C.2.5). First, Monte Carlo events should be generated at the $q$ different masses $M_j$ instead of as described in Sec. C.2.2. Second, the function calculated in Eq. C.6, which may no longer be interpreted as a joint density, should be re-labeled. For lack of a better alternative, call it $\xi(\vec{v})$.

We now add a step $5\frac{1}{2}$ between Secs. C.2.5 and C.2.6. This function $\xi(\vec{v})$ is clearly not an appropriate density function — if events have been generated assuming five different masses $M_j$, a graph of $\xi(\vec{v})$ might appear as shown in Fig. C.1. We see that the density has ridges along the values of $m$ for which events have been generated, with corresponding valleys in the regions between these values.

We can generate an appropriately rescaled probability density $p(\vec{x}, m)$ by multiplying $\xi(\vec{v})$ by a normalizing $m$-dependent factor $s(m)$:

$$p(\vec{x}, m) = \xi(\vec{x}, m)s(m). \tag{C.16}$$

This normalizing factor will correct for the fact that we have introduced valleys into our

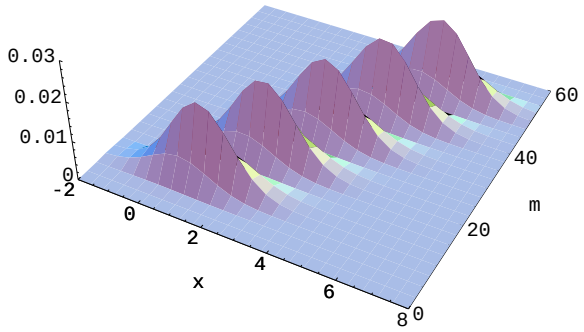



Figure C.1: A sample function $\xi(x, m)$ that might be constructed from Monte Carlo events at masses $M = 10$, 20, 30, 40, and 50. Notice the ridges in this function, due to the fact that it is constructed from events at specific masses.

density by only generating events at specific masses $M_j$. The requirement that

$$\int p(\vec{x}, m)\, d\vec{x} = p(m) \tag{C.17}$$

determines this normalizing factor uniquely. The desired joint probability density $p(\vec{v})$ is then given by

$$p(\vec{v}) = \frac{\xi(\vec{v})p(m)}{\int d\vec{x}'\, \xi(\vec{x}', m)}, \tag{C.18}$$

and the final step (Sec. C.2.6) is exactly as before. The rescaled density of Fig. C.1 is shown in Fig. C.2.

We mention briefly a useful shortcut when calculating integrals such as that appearing in the denominator of Eq. C.18. Multidimensional integrals are difficult to calculate in general, but this integral may be handled analytically provided kernels of the Gaussian form (Eq. C.15) are used. Assume as in Sec. C.3.1 that the vector of known variables $\vec{x}$ is of $d_1$ dimensions, that the vector of unknown variables $\vec{\alpha}$ is of $d_2$ dimensions, and that the

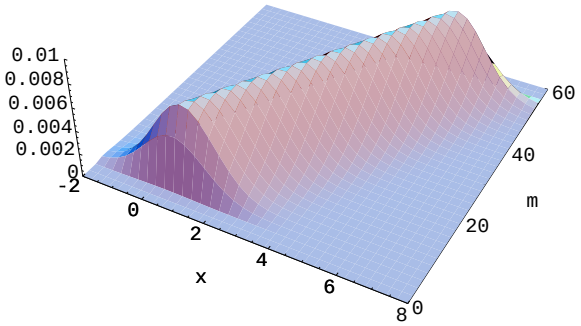



Figure C.2: The density $p(x, m)$ formed by rescaling the function $\xi(x, m)$ shown in Fig. C.1. Notice how this rescaling corrects for the fact that only events at specific masses were used in the construction of $\xi(x, m)$.

Monte Carlo has a covariance matrix $\Sigma$. Then the relevant formula is

$$\int K(\vec{x}, \vec{\alpha})\, d\vec{x} = \frac{1}{(\sqrt{2\pi}h)^{d_2}\sqrt{\det(\Sigma')}} \exp\left(\frac{-\vec{\alpha}^T \Sigma'^{-1}\vec{\alpha}}{2h^2}\right), \tag{C.19}$$

where $\Sigma'$ is the $d_2$ by $d_2$ submatrix of $\Sigma$ formed by retaining elements with row and column numbers larger than $d_1$.

### C.3.3 Background events

In the text we considered the problem of determining the mass $m$ for one candidate event. In any practical analysis we will have $n$ such events, and of those we will generally expect some fraction $b$ to be background events – events that do not contain a Higgs boson at all. In this section we show how to apply this method to a full-blown analysis.

We begin by generating signal Monte Carlo events and background Monte Carlo events, and constructing the signal and background probability densities $p_s(\vec{x}, \vec{\alpha})$ and $p_b(\vec{x}, \vec{\alpha})$ as described in Secs. C.2.1 to C.2.5. From a careful analysis of background efficiencies we

determine the probability $p(b)$ that a fraction $b$ of our events are background events. We label the $n$ data events by $\vec{x}_1, .., \vec{x}_n$.

The goal is to compute the posterior density $p(\vec{\alpha}|\vec{x}_1, .., \vec{x}_n)$; once this is obtained, we can compute any statistical quantities desired. Since the observations $\vec{x}_1, .., \vec{x}_n$ are assumed to be independent, the joint probability density $p(\vec{x}_1, .., \vec{x}_n, \vec{\alpha}, b)$ factors into a product:

$$p(\vec{x}_1, .., \vec{x}_n, \vec{\alpha}, b) = p(\vec{\alpha})p(b) \prod_{i=1}^{n} p(\vec{x}_i|\vec{\alpha}, b). \tag{C.20}$$

The probability density for one data event $p(\vec{x}, \vec{\alpha}, b)$ may be written in terms of the signal and background probability densities as

$$p(\vec{x}, \vec{\alpha}, b) = (1 - b)\, p_s(\vec{x}, \vec{\alpha})p(b) + b\, p_b(\vec{x}, \vec{\alpha})p(b). \tag{C.21}$$

Dividing Eq. C.21 by $p(\vec{\alpha})p(b)$, one gets the term $p(\vec{x}_i|\vec{\alpha}, b)$ needed in the product of (C.20). One then integrates out the *nuisance parameter* $b$ in Eq. C.20 and divides by $\prod_{i=1}^{n} p(\vec{x}_i)$ to obtain the desired $p(\vec{\alpha}|\vec{x}_1, .., \vec{x}_n)$.

We could now maximize this quantity with respect to $\vec{\alpha}$ to determine the most likely values of the parameters as we did in the Higgs example, perhaps estimating an error from the width of the peak. If desired, we could compute the expectation value

$$\langle \vec{\alpha} \rangle = \int \vec{\alpha}\, p(\vec{\alpha}|\vec{x}_1, .., \vec{x}_n)\, d\vec{\alpha}, \tag{C.22}$$

and along with it the covariance

$$\sigma_{rs} = \int (\alpha_r - \langle \vec{\alpha} \rangle_r)(\alpha_s - \langle \vec{\alpha} \rangle_s)\, p(\vec{\alpha}|\vec{x}_1, .., \vec{x}_n)\, d\vec{\alpha}. \tag{C.23}$$

One of the many great advantages to dealing explicitly with probability densities is the ability to use them to easily calculate quantities such as these. Another advantage is the

possibility of using a graphing package to plot these densities and view them directly. This method, even when applied in several dimensions, lends itself to a very intuitive, graphical interpretation.

## C.4 Summary

We hope that the analysis method described here will prove to be a generally useful tool. If the Monte Carlo distributions do not lend themselves to an obvious parameterization, a non-parametric approach such as the one utilized here may be especially attractive. If one wishes to use several correlated variables simultaneously in a fit, this method provides a kludgeless way to do so. If one wishes to fit for two or three unknown variables simultaneously, this method again provides a kludgeless way to do so. If one wishes to assess possible systematic errors or to study correlations among the event variables, one can use this method to fit for variables and then compare them with their known values. If one wishes to obtain confidence intervals or other statistical data from this analysis, one can simply manipulate the probability densities that have been generated along the way. The method itself is extraordinarily simple: we simply compute the value of $m$ that maximizes the posterior density $p(m|\vec{x})$. We anticipate that this method will be useful for future discovery searches for these reasons.

# Appendix D

# $\not\!\!E_T$ significance

The missing transverse energy resolution of the DØ detector is determined by the convolution of a number of effects. Choosing an incorrect vertex, mismeasuring the energy of a shower in the calorimeter, and failing to identify the presence of a "hot" calorimeter cell are among the most prominent. In a large ensemble of events the Central Limit Theorem ensures that we will be able to speak of a "$\not\!\!E_T$ resolution"; i.e., the $x$- and $y$-components of $\not\!\!E_T$ will closely approximate Gaussians centered at zero with some width $\delta\not\!\!E_T/\sqrt{2}$. If the ensemble is binned in one or more quantities, then the dependence of $\delta\not\!\!E_T$ on those quantities may be determined. The quantity most typically used is the scalar summed transverse energy of the event ($\sum E_T$), leading one to conclude that $\delta\not\!\!E_T \approx \sqrt{\sum E_T}$, when energies are expressed in units of GeV. Given this function for the $\not\!\!E_T$ resolution, a quantity often referred to as the "$\not\!\!E_T$ significance" is often defined by $\not\!\!E_T/\delta\not\!\!E_T$. This $\not\!\!E_T$ significance, being a property of each individual event, is a quantity that one can use to preferentially select events with true $\not\!\!E_T$. This common definition, a function only of the scalar $\not\!\!E_T$ and scalar summed transverse

energy, completely ignores the topology of the event.

A physicist given a display of an event with substantial $\not\!\!E_T$ is likely to argue that the $\not\!\!E_T$ is or is not significant by seeing whether the event can be "smeared" (i.e., modified within reason) in such a way that the $\not\!\!E_T$ vanishes. In this spirit we present a new method for determining whether an event contains true $\not\!\!E_T$. The "smearing" is made quantitative by introducing the notion of $p(\not\!\!E_T)$, the probability density for the missing transverse energy in an event. This is a true probability density in that, for a given event, the probability that the actual missing transverse energy in the event lies between $\not\!\!E_T$ and $\not\!\!E_T+\delta\not\!\!E_T$ is given by $p(\not\!\!E_T)\delta\not\!\!E_T$. This density is used to form a likelihood $\mathcal{L}_{\not\!\!E_T}$ that may be interpreted as the significance of the $\not\!\!E_T$ in the event.

## D.1 Probability densities

### D.1.1 Uncertainty in the vertex: $p(z)$

The probability density for the $z$ vertex position is constructed from tracks found in the event. In a typical event in Run I, $n_v$ candidate vertices $i$ are determined with position $z = \mu_i$ and resolution $\delta z = \sigma_i$. The probability $p(i)$ that the $i^{\text{th}}$ vertex is the correct one may be estimated crudely in Run I as the fraction of tracks associated with the vertex; in Run II we expect more sophisticated algorithms to be employed. The density may then be defined as

$$p(z) = \sum_{i=1}^{n_v} p(i)\, g(\mu_i, \sigma_i; z), \tag{D.1}$$

where

$$g(\mu, \sigma; z) \equiv \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-(z-\mu)^2}{2\sigma^2}\right) \tag{D.2}$$

is a Gaussian kernel centered at the position $\mu$ with width $\sigma$. We might expect $p(z)$ to look something like the graph shown in Fig. D.1.

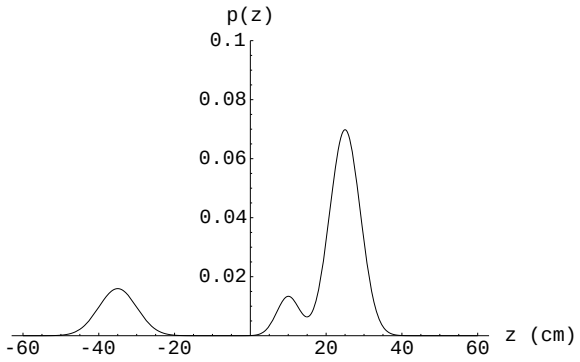



Figure D.1: Plot of a probability density $p(z)$. This particular event has a "main" vertex at $z \approx 25$ cm and a second vertex at $z \approx -35$ cm, in addition to two tracks that appear to originate from $z \approx 10$ cm.

### D.1.2 Uncertainty in measured energies: $p(E)$

The remaining probability densities that we will need are the densities $p(E)$ for the energies $E$ of the physics objects in the event. Explicitly, we want to construct $p(E)$ such that the probability that the final state object in question has an energy between $E$ and $E+\delta E$ is $p(E)\delta E$.

#### Jets

To first order the uncertainty in the calculated energy $\mu$ for a jet may be handled by smearing this value across a Gaussian of width

$$\sigma = 80\%\sqrt{\mu}, \tag{D.3}$$

so that

$$p(E) = g(\mu, \sigma; E). \tag{D.4}$$

At higher order the energy resolution also depends upon $\eta_{\rm det}$. In this case we enlarge the dependence of the resolution and consider $\sigma(\mu, \eta_{\rm det})$, which depends upon both the energy deposited and the location in the calorimeter. We use the results of measurements made in the context of the analysis of the triple differential dijet cross section. [63]

**Electrons**

Similarly, the energies of any electrons in the event having measured energy $\mu$ are smeared with a Gaussian of width

$$\sigma = 20\%\sqrt{\mu}. \tag{D.5}$$

**Muons**

The rather coarse muon energy resolution in Run I will be much improved with the presence of a central magnetic field in Run II. The Run I muon resolution is typically described in terms of a Gaussian in the reciprocal of the muon energy. A muon with measured energy $\mu$ has

$$p(E) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-(1/E - 1/\mu)^2}{2\sigma^2}\right), \tag{D.6}$$

where $\sigma = 0.18(\mu - 2)/\mu^2 \oplus 0.003$, the $\oplus$ means add in quadrature, and the difference between the muon momentum and energy has been ignored. [64]

**Hot cells**

Hot cells are accounted for in a similarly probabilistic fashion. We calculate the probability that a given cell is anomalously hot (with some method not specified here), and remove the cell from the event with this probability.

**Unclustered energy**

Some of the energy in an event remains unassociated with physics objects, and is called *unclustered energy*. We treat this unclustered energy as a "physics object" in its own right, and smear its energy according to Sec. D.1.2, assuming that the unclustered energy is primarily hadronic.

## D.2 Uncertainty in the quantity of interest: $p(\not\!\!E_T)$

Once we have determined $p(z)$ and the $p(E)$, we can use these to calculate the probability density $p(\not\!\!E_T)$ for the missing transverse energy by a Monte Carlo. First, a value of $z$ for the primary vertex is chosen randomly from the distribution $p(z)$. Next, values for the energies of the objects in the event are randomly chosen from the distributions $p(E)$. The $\not\!\!E_T$ in the event is computed using this vertex and these energies. If we do this many times and histogram the computed missing transverse energies, the probability density $p(\not\!\!E_T)$ is obtained. Graphs of what this might look like for a dijet event and for a W event are shown in Figs. D.2 and D.3, respectively.

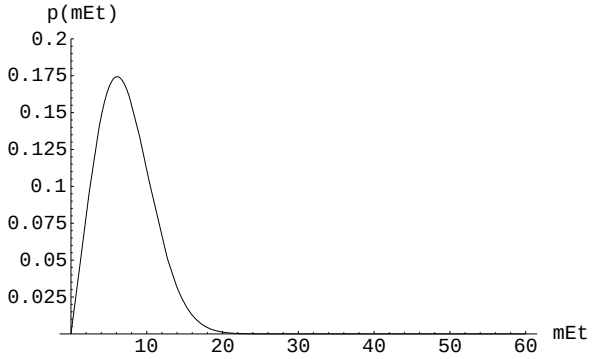



Figure D.2: Graph of a sample probability density for the missing transverse energy for a dijet event. The corresponding two-dimensional probability density in the $\not\!\!E_{Tx}$–$\not\!\!E_{Ty}$ plane is actually peaked at the origin; a Jacobian zero at the origin appears when one considers the one-dimensional scalar $\not\!\!E_T$.

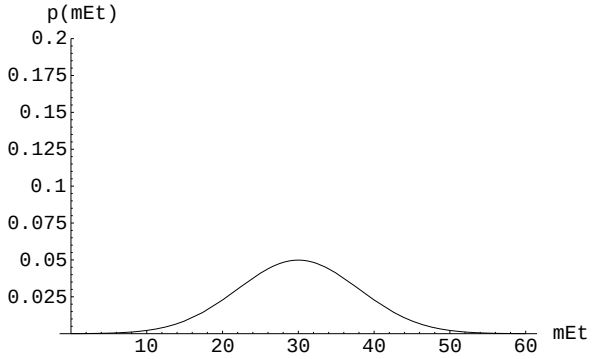



Figure D.3: Graph of a sample probability density for the missing transverse energy for a W event.

## D.3 Significance: $\mathcal{L}_{\not\!\!E_T}$

Using the projection $\not\!\!E_{Ta}$ of the missing transverse energy along the axis $a$ defined by the measured missing transverse energy in order to avoid the problems associated with the Jacobian zero at the origin, we define a likelihood by

$$\mathcal{L}_{\not\!\!E_T} = \log_{10} \frac{p(\not\!\!E_{Ta})_{\max}}{p(\not\!\!E_{Ta} = 0)}. \tag{D.7}$$

## D.4 Measures of performance

### D.4.1 Testing $p(\not\!\!E_T)$ for correctness

The probability that the missing transverse energy is between $\not\!\!E_T$ and $\not\!\!E_T + \delta\not\!\!E_T$ will be $p(\not\!\!E_T)\delta\not\!\!E_T$ if and only if the construction of $p(\not\!\!E_T)$ is correct. Whether this is so can be checked for the components $\not\!\!E_{Tx}$ and $\not\!\!E_{Ty}$ individually using W+jets Monte Carlo, where the W decays to leptons. Each event $i$ has some known parton-level $\not\!\!E_{Ti}$ corresponding to the neutrino in the event. For each of these events we calculate the quantity

$$\xi_i = \int_{-\infty}^{\not\!\!E_{Txi}} p_i(\not\!\!E_{Tx})\, d\not\!\!E_{Tx}, \tag{D.8}$$

where $p_i(\not\!\!E_{Tx})$ is the $\not\!\!E_{Tx}$ probability density calculated for the $i^{\text{th}}$ event, and $\not\!\!E_{Txi}$ is the true missing transverse energy in the Monte Carlo event. $p(\not\!\!E_T)$ is an accurate estimate of the probability density for the missing transverse energy if and only if $\xi$ has a uniform distribution between 0 and 1. A similar check may of course be made for $\not\!\!E_{Ty}$.

### D.4.2 Testing $\mathcal{L}_{\not\!\!E_T}$ for usefulness

The goal of this effort is to use the function $p(\not\!\!E_T)$ to distinguish between signal events with large true missing transverse energy and background events with zero true missing transverse energy. To obtain a quantitative measure of "usefulness" we look at two samples, one with true missing transverse energy, and one without. The first is a sample of $W+ \geq 2$ jets Monte Carlo events, and the second is a sample of QCD events containing $\geq$ 3 jets collected with the `jet_85` trigger, in which any missing transverse energy presumably arises from detector effects. We require the events in both samples to have $30 < \not\!\!E_T < 40$.

Figure D.4 [65] shows the type of topological cut made in a DØ search for squarks and gluinos in the $\not\!\!E_T\, 3j$ channel [66] — all events outside the solid lines are removed by this cut. We find it convenient for purposes of comparison to introduce a variable $s$, defined such that an event lying on the solid line in Fig. D.4 has $s = 1$ and an event lying on the boundary of the plot has $s = 0$; $s$ is a measure of how far an event is from the boundary and the lower right-hand corner of the plot.

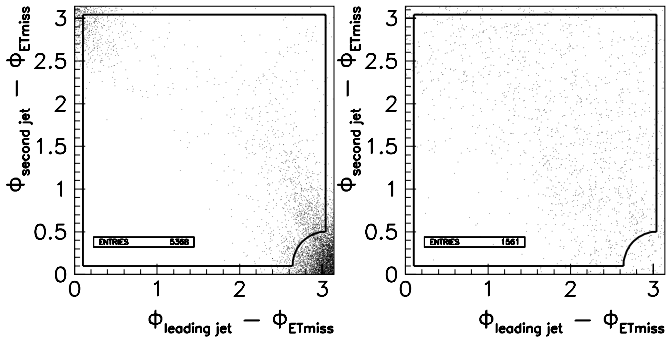



Figure D.4: A standard topological cut to reject QCD background — events outside the solid lines are discarded. (Left) QCD events with $\geq 3$ jets and $30 < \not\!\!E_T < 40$ GeV, collected with the `jet_85` trigger. (Right) $W$+jets Monte Carlo events with $\geq 2$ jets and $30 < \not\!\!E_T < 40$ GeV. This topological cut clearly eliminates a greater fraction of QCD events than $W$+jets events, but can one do better?

Figure D.5 shows distributions of the topological variable $s$ and the likelihood variable $\mathcal{L}$ for the $W$+jets and QCD samples. We find that the standard cut of $s > 1$ retains 83% of the $W$+jets Monte Carlo events, and we tune our cut on the likelihood in order to achieve the same efficiency on this sample. A cut of $\mathcal{L} > 3$ retains 82% of the $W$+jets events. Applying the same cuts to our QCD sample, we find that a cut of $s > 1$ keeps 29% of these unwanted events, while a cut of $\mathcal{L} > 3$ retains only 5.4% of these events. That is, for the same efficiency on events with true missing transverse energy, our likelihood cut appears to

reject a factor of five more background events than does the topological cut that has been used in previous jets+$\not\!\!E_T$ analyses.

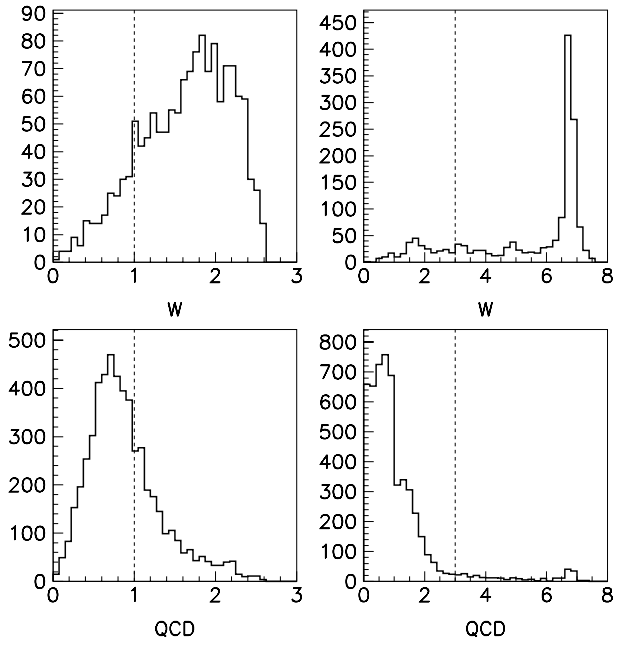



Figure D.5: A comparison of the topological variable $s$ and the likelihood $\mathcal{L}$ for $W$+jets Monte Carlo events and QCD data events. The $W$+jets Monte Carlo events are required to have $\geq 2$ jets and $30 < \not\!\!E_T < 40$ GeV; the QCD events, collected from `jet_85`, are required to have $\geq 3$ jets and $30 < \not\!\!E_T < 40$ GeV. (Top left) The topological variable $s$ for the $W$+jets Monte Carlo events. A cut of $s > 1$ keeps 83% of this sample. (Top right) The likelihood variable $\mathcal{L}$ for these events. A cut of $\mathcal{L} > 3$ keeps 82% of this sample. (Bottom left) The topological variable $s$ for the QCD data events. A cut of $s > 1$ keeps 29% of this sample. (Bottom right) The likelihood variable $\mathcal{L}$ for these events. A cut of $\mathcal{L} > 3$ retains only 5.4% of these unwanted events, less than one fifth of the number retained using the standard topological variable $s$ for the same $W$+jets efficiency.

Figure D.6 shows the distributions in $s$ and $\mathcal{L}$ for a $W$+jets data sample, selected by requiring one good electron, $\geq 2$ jets, $50 < M_T^{e\nu} < 90$ GeV, and $\not\!\!E_T > 30$ GeV. We see that the cuts $s > 1$ and $\mathcal{L} > 3$ perform comparably on real data, as well as on Monte Carlo.

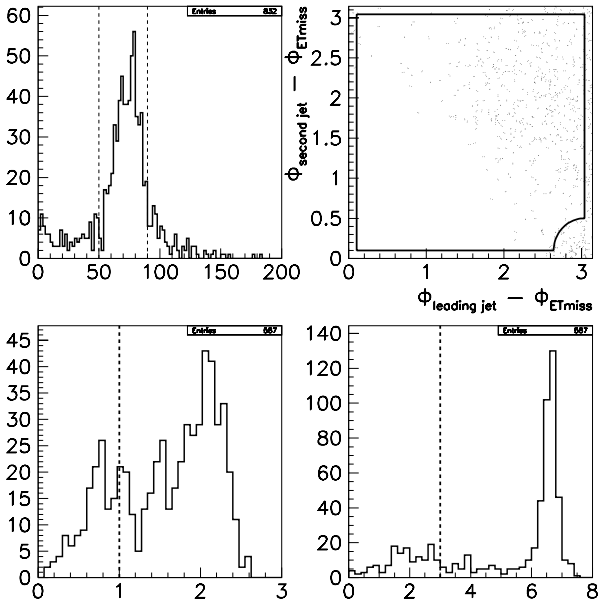



Figure D.6: A comparison of the topological variable $s$ and the likelihood $\mathcal{L}$ for $W$+jets data events. (Top left) The transverse mass distribution for the $W+\geq 2$ jets data sample; we keep events with $50 < M_T^{e\nu} < 90$ GeV. (Top right) Scatter plot of these events in the variable space considered in previous DØ analyses. (Bottom left) The topological variable $s$ for these events. A cut of $s > 1$ keeps 74% of the events. (Bottom right) The likelihood variable $\mathcal{L}$ for these events. A cut of $\mathcal{L} > 3$ keeps 76% of these events.

## D.5 Summary

We have described a method for quantifying the "significance" of $\not\!\!E_T$ on an event-by-event basis. Central to the method is the construction of the probability density $p(\not\!\!E_T)$, which quantifies the intuitive "smearing" that physicists perform in order to argue that an event does or does not contain true missing transverse energy. The components (determination of $p(z)$, $p(E)$, etc.) of the method presented here can easily be extended to include additional information provided by more sophisticated algorithms in Run II.

# Appendix E

# Measure of Sensitivity

In this appendix we describe a useful measure of sensitivity for analyses that set limits on parameters of a theory. Because the results of unsuccessful searches in particle physics are frequently expressed in terms of a lower limit on the mass of the undiscovered particle, this is the case that we consider. Extensions of this measure of sensitivity to other types of parameters, or the case of multiple parameters, is straightforward.

## E.1 Motivation

A physicist testing a specific hypothesis is usually interested only in a subset of the data collected in the experiment. She is therefore faced with the problem of choosing among many possible sets of cuts to reduce backgrounds while retaining the signal she hopes to find. In order to avoid introducing bias when setting these cuts, it is helpful to have a rigorous criterion for preferring one set of cuts over another. Such a criterion is equivalent to a mapping $S$ from the set $\mathcal{C}$ of all possibles cuts to the real numbers. $S$ should have the

property that for any two sets of cuts $C$ and $C'$, we agree to say that the set of cuts $C$ is better than the set of cuts $C'$ if and only if $S(C) > S(C')$.

In the DØ first generation scalar leptoquark analysis [51], the function $S$ was chosen to be of the form

$$S(C) = \begin{cases} \hat{s}, & \hat{b} \leq 0.4 \\ 0, & \text{otherwise,} \end{cases} \tag{E.1}$$

where $\hat{s}$ is the expected number of signal events passing the set of cuts $C$, and $\hat{b}$ is the expected number of background events passing the set of cuts $C$. The rationale for this choice of $S$ is that for $\hat{b} < 0.4$ the probability that 0 data events pass the cuts $C$ is about 66%, and a very good limit may be set if in fact 0 data events pass those cuts. The choice of $\hat{b} < 0.4$ (rather than $\hat{b} < .2$ or $\hat{b} < .7$, for example), is in some sense determined by how much one wants to "gamble" that zero events will be observed in the data. Unfortunately, this means that the final specification of the mapping $S$ is subject to the habits of the individual performing the analysis, and thus does not meet the goal of generality prescribed above.

## E.2 Weighted gambling

A more explicit alternative to the suggestion for $S$ in Eq. E.1 may be obtained by pursuing this notion of gambling. We would like to choose a set of cuts that maximizes the expected lower limit on the mass (usually at the 95% confidence level). This expected limit may be written

$$\hat{M}^{95\%}(C) = \sum_{k=0}^{\infty} P(k,\hat{b}) M^{95\%}(k,\hat{b},\hat{s}(M)), \tag{E.2}$$

where $P(k,\hat{b})$ denotes the probability that $k$ data events pass the set of cuts $C$, and $M^{95\%}(k,\hat{b},\hat{s}(M))$ is the mass limit that will be obtained if $k$ data events pass these cuts. $\hat{M}^{95\%}(C)$ has been written to emphasize the dependence of $\hat{M}^{95\%}$ upon the choice of cuts $C$.

If Poisson statistics are assumed, $P(k,\hat{b})$ is simply given by

$$P(k,\hat{b}) = \frac{e^{-\hat{b}}\hat{b}^k}{k!}. \tag{E.3}$$

The second piece of the summand in Eq. E.2, $M^{95\%}$, depends upon the assumed number of data events $k$, the number of expected background events $\hat{b}$, and the number of expected signal events $\hat{s}(M)$ passing our set of cuts $C$. The parameters $\hat{b}$ and $\hat{s}(M)$ are functions of our cuts, and $\hat{s}(M)$ is explicitly written as a function of the mass $M$.

The upper 95% confidence level on the cross section may be found by numerically computing the value $N$ with the property that the probability for the expected background $\hat{b}$ to fluctuate up to or above $N$ is 5%. Define $s = N - \hat{b}$, and re-write the standard formula

$$\sigma^{95\%}(M) = \frac{(N-\hat{b})\sigma_{th}(M)}{\hat{s}(M)} \tag{E.4}$$

in the form

$$\frac{\sigma^{95\%}(M)}{\sigma_{th}(M)} = \frac{N-\hat{b}}{\hat{s}(M)}. \tag{E.5}$$

The mass $M$ at which $\sigma^{95\%}(M)$ and $\sigma_{th}(M)$ intersect is determined by

$$\frac{\sigma^{95\%}(M)}{\sigma_{th}(M)} = 1; \tag{E.6}$$

the numerical solution of

$$\hat{s}(M) = N - \hat{b} \tag{E.7}$$

for $M$ thus provides the desired $M^{95\%} = M^{95\%}(k, \hat{b}, \hat{s}(M))$ referred to above.

## E.3 Incorporating errors

We have assumed in Sec. E.2 that the quantities $\hat{b}$ and $\hat{s}(M)$ are known precisely for the set of cuts $\mathcal{C}$. As long as the errors on these quantities are small, this is a reasonable approximation. If these errors are not small, they should be taken into account.

The prescription for incorporating the errors on $\hat{b}$ and $\hat{s}(M)$ into the calculation of sensitivity is straightforward. The error on the expected background is described by a probability distribution $p(b)$ with mean $\hat{b}$. Similarly, for a given $M$, the error on the expected number of signal events is given by a probability distribution $p(s(M))$ with mean $\hat{s}(M)$. The obvious generalization of Eq. E.2 is now

$$\hat{M}^{95\%}(C) = \int_0^\infty db\, p(b) \int_0^\infty ds(M)\, p(s(M)) \sum_{k=0}^{\infty} P(k,b) M^{95\%}(k,b,s(M)). \tag{E.8}$$

Note that this reduces to Eq. E.2 when $p(b) = \delta(b - \hat{b})$ and $p(s(M)) = \delta(s(M) - \hat{s}(M))$, as expected.

## E.4 Application to $LQ\overline{LQ} \to eejj$

This measure of sensitivity has been applied to DØ's search for leptoquarks decaying to two electrons and two energetic jets. The final cut in that analysis was on a variable $S_T$,

| $S_T$ | $\mathcal{S}$ |
|---|---|
| 200 | 45.13 |
| 225 | 176.85 |
| 250 | 206.62 |
| 275 | 212.48 |
| 300 | 216.44 |
| 325 | 219.49 |
| 350 | 220.30 |
| 375 | 219.62 |
| 400 | 217.45 |

Table E.1: Computed limit setting sensitivity $\mathcal{S}$ for given values of the cut parameter $S_T$.

defined as the scalar sum of the transverse energies of all jets in an event. The computed value of the limit setting sensitivity $\mathcal{S}$ as a function of this cut variable is given in Table 1.

It can be seen from this table that the sensitivity is maximized for $S_T \approx 350$, and that with this cut a mass limit of roughly 220 GeV is expected. In the actual analysis this was in fact the cut that was found with the "maximize signal at an expected background of 0.4" prescription. Zero events were seen in the data, enabling a mass limit of 225 GeV to be set. This limit is slightly larger than the expected limit of 220 GeV quoted in the table above; we might say that DØ was 5 GeV "lucky" in this case.

## E.5 Summary

We have constructed a well-defined measure of sensitivity appropriate for analyses that attempt to set a lower limit on the mass of an undiscovered particle. This measure of sensitivity is simply the expected limit with a set of cuts $C$; explicitly,

$$S = \hat{M}^{95\%} = \sum_{k=0}^{\infty} P(k, \hat{b}) M^{95\%}(k, \hat{b}, \hat{s}(M)). \tag{E.9}$$

$P(k, \hat{b})$ is given in closed form in Eq. E.3, and an algorithm has been given in the form of Eqs. E.4–E.7 to compute $M^{95\%}(k, \hat{b}, \hat{s}(M))$. Errors in the expected background, signal efficiency, luminosity, and theoretical cross section may be incorporated into the calculation simply by replacing Eq. E.2 with Eq. E.8. This sensitivity was applied to the search for leptoquarks decaying to two electrons and two jets, and reasonable results were obtained. We believe that the sensitivity measure described here is both simple and powerful, and encourage its use in any future analysis whose final result will be expressed as a limit on one or more parameters of a theory.

# Appendix F

# Test beam simulator

Ideally, the DØ Central Fiber Tracker (CFT) readout electronics would be exercised in a test beam before the beginning of Run II. However, the time and cost of such an effort is prohibitive. This appendix describes recently developed hardware, designed to simulate scintillation signals from the CFT fibers, for use in a system-level test of the CFT readout.

## F.1 General information

### F.1.1 Trigger test stand

The trigger test stand is designed to test the Visible Light Photon Counters (VLPCs) and their associated readout electronics. Photons from scintillating fibers in the Central Fiber Tracker (CFT), Central Preshower Detector (CPS), and Forward Preshower Detector (FPS) will be channeled from the detectors to the racks housing readout electronics through clear fibers. The photons, produced in the scintillating fibers and then shifted in wavelength with wavelength shifting fibers (WSFs), will have wavelengths of 525 nm as they travel

through the clear fibers. In order to convert these photons into an electrical signal, a technology is needed that is sensitive to small numbers of photons, has low noise, and operates in the visible spectrum. VLPCs, developed by Rockwell in collaboration with Fermilab, have these characteristics.

The preferred method for testing the readout electronics is to use a test beam, in which the entire system is put into a secondary beam from the Tevatron and the electronics tested *in situ*. Since resources are limited and the scheduled roll-in date (March 1, 2001) is quickly approaching, we have decided instead to simulate a test beam. The actual test beam would have been used to produce pulses of light to be read out by the VLPCs from the passage of charged particles through the scintillating fibers. Our simulated beam will produce pulses of light directly, using light emitting diodes (LEDs). We would like to test the readout electronics by providing input as similar as possible to the signals they will actually see when collecting data.

Consider a set of 512 readout channels corresponding to one of the eighty sectors of the CFT. A typical signal might contain a charged particle track through one of these sectors together with a number of hits due to noise. In order to mimic this input signal, it is necessary that we be able to control the number of photons emitted by each LED, and that we be able to control each LED channel independently. In addition to providing a check for crosstalk and other problems in the readout system, this flexibility will allow us to produce real Monte Carlo "events" to test the track-finding trigger. The trigger test stand is being designed with the intention of testing 512 channels simultaneously. We will have four identical VME Test Beam Generator (VTBG) boards, each of which will control 128

independent channels.

The trigger test stand will consist of several functional blocks, shown in Fig. F.1. The user interacts with the PC, which is used to download trigger patterns onto the VTBG through VMEbus. Once the pulsing sequence is initiated, analog electrical signals from the VTBG are converted into analog light signals on the VTBG-A. These are transmitted over fiber-optic cable to the VLPCs, which are housed in cassettes kept at approximately six degrees Kelvin. Here the number of photons incident on each VLPC is converted back into an electrical signal, which is then processed by the readout system, consisting of the Front End, the Sequencer, and the VME Readout Board (VRB). Final trigger patterns are then read back into the PC, where they are compared with the originals.

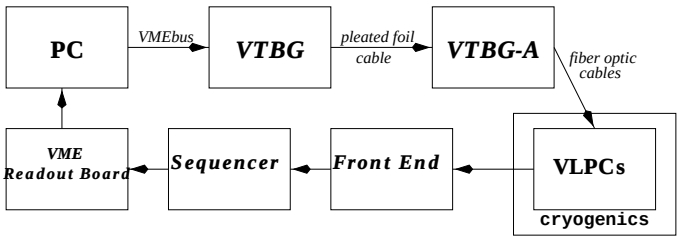



Figure F.1: Block diagram for the system designed to test the VLPCs and CFT readout electronics.

### F.1.2 VTBG requirements

Each VTBG board must have the following functionality:

1. It must accept and store in local memory $2^{13} = 8192$ events. An event is defined as a vector of $2^7 = 128$ signals, where a signal is defined as an 8-bit word (corresponding to a decimal integer between 0 and 255) defining the relative number of photons to be emitted by that particular LED. Thus a signal with value 255 corresponds to the

maximum LED output, and a signal with value 0 corresponds to no light emitted from the LED. The total number of 16-bit memory locations (one memory location for every two signals) on the board will thus be $2^{13} \times 2^7/2 = 2^{19}$.

2. It must talk with a PC through VMEbus. It must accept 16-bit reads and writes to each of its $2^{19}$ memory locations. After all events are loaded onto local memory, it must accept a signal from the PC to initiate the pulsing sequence. It must also accept a signal from the PC to halt a pulsing sequence in progress.

3. It must pulse every 132 ns in sync with the global 7 MHz clock on the test stand. The global 7 MHz clock determines the timing of the simulated beam crossing. The 7 MHz clock signal will be high for 10 ns during the simulated beam crossing. It will be low for the remaining 122 ns of its 132 ns cycle.

4. It must utilize a 53 MHz clock, global to the test stand, for its internal logic.

5. Its output at each crossing must consist of 128 analog signals (one event), with each analog signal's amplitude determined by its digital signal, an 8-bit word. The pulse width will be 19 ns. There should be no output during the remaining 113 ns between crossings.

6. The analog signal must be converted into light output. This will be accomplished on a daughterboard, the VME Test Beam Generator – Analog (VTBG-A) board connected to the VTBG by a 30-foot pleated foil cable. The tailoring of the analog signal for the LEDs is performed on this daughterboard so that the VTBG may be used to control other outputs if the need arises. By modularizing the design in this way, a new

daughterboard can be used to convert the VTBG signals into whatever form is desired.

7. The board must physically sit inside a VME crate, which puts modest constraints on its size.

8. The cost of this project (four printed circuit boards plus one spare) may not exceed the allocated budget of $50,000. The final board must be completed by the date set for the beginning of the test stand: March 15, 1999.

## F.2 Theory of operation and operating modes

### F.2.1 VME memory map

| A23–A21 | A20 | A19–A14 | A13–A01 | Action |
|---|---|---|---|---|
| n | X | X | X | None |
| y | 0 | <chip> | <address> | The 6 bits in <chip> select one of the 64 static RAMs on the board. The 13 bits in <address> select one of the 8192 memory addresses on that chip. |
| y | 1 | X | X | Start/stop pulsing |

Table F.1: The VTBG's VME memory map. (y,n) denotes (does, does not) match DIP switch settings on the board. The three DIP switches are used to determine the board's location in VME address space.

Table F.1 summarizes how to access the memory of the VTBG through VME.

### F.2.2 Loading memory

The VTBG will respond only to standard 24-bit VME addressing and 16-bit data transfers. It requires that the address modifier (AM5-0) bits be set to any of $3E (Standard Supervisory Program), $3D (Standard Supervisory Data), $3A (Standard Non-Privileged

Program), or $39 (Standard Non-Privileged Data). Address bits A23-21 are used to select the VTBG boards from other boards in the VME crate. The address of each VTBG is set by three dip switches on the board; if A23-21 do not match this dip switch setting, the VTBG will not respond. Address bit A20 functions as the high-level control bit. For memory reads and writes, this bit must be low. The board contains $2^6 = 64$ high-speed (15 ns) dual-port static RAMs from Cypress. Each memory chip contains $2^{13} = 8192$ memory locations, and each memory location holds 16 bits of data. We store exactly two digital signals ($2 \times 8$ bits of data) in each memory location. Address bits A19-14 are used to select one of the 64 memory chips on the board. Address bits A13-A01 are used to select a particular memory location within that chip. If WRITE* (from VME) is low, the data presented on the data transfer bus will be written to that memory location. If WRITE* is high, the data in the memory location will be presented to the data transfer bus for a VME read. The details of the implementation of this design are shown on the board schematic page labeled InterfaceVME.

### F.2.3 Pulsing sequence

The pulsing sequence is enabled by accessing the pulsing control memory address. This address is accessed with A20 high. Addresses A19-01 are not checked in this case. Upon receiving this address, a signal is sent to a local Altera FPGA, which controls the logic for the pulsing sequence. To halt a pulsing sequence in progress, one accesses the pulsing control memory address again.

Since video DACs, which come three to a package, are used on this board, the 128 channels are broken into 21 "hexachannels," each of which contains six individual channels, plus two extra channels in order to bring the total number of channels to 128. Each

hexachannel consists of three memory chips and two DACs. Since each event contains 128 analog signals, each of the 64 memory chips on the board must contain 16 bits for each event. That is, the data in one 16-bit memory location will be read out of each memory chip during each 132 ns.

The 132 ns (7 MHz) crossing cycle is broken into 7 stages by the synchronous 53 MHz clock. The output DACs have a one-stage pipeline; new data is latched into the pipeline, and the old data appears on the output, on the rising edge of the 53 MHz clock. During 6 of the 7 stages in each cycle, memory address #0000 is presented to the DAC input; we require that this memory address contain data 00 on each memory chip. During the $7^{\text{th}}$ stage in the cycle, memory address $\#j$ is presented to the DAC input, assuming that this is the $j^{\text{th}}$ crossing (mod 8192) since the start of the pulsing sequence. Upon reaching event #8191 (which occurs roughly one millisecond after the start of the sequence), the pulser loops back to event #0001 and begins the count again.

The output pulses must be synchronized to the global 7 MHz clock, which determines the simulated beam crossing. Many factors, including the rise time of the DACs, the length of cable between the VTBG and the VTBG-A, the circuit design of the daughterboard, and the rise time of the LEDs, will shift our output signal relative to this clock. The timing of the output signal may be fine-tuned within the controlling FPGA. A ByteBlaster circuit incorporated on the board enables the reprogramming of the FPGA through a 25-pin connection to a PC.

### F.2.4 LEDs

We are using high speed GaAlAs red LEDs from Optek, which emit light within the sensitivity region of the VLPCs. The circuitry used to drive the LEDs has been prototyped. Using each LED in the feedback loop of a high speed op amp from Burr Brown, driven by a TMC3503 triple video DAC, we see light pulses with rise times of less than 5 ns. The total number of emitted photons from one LED at full scale within an 20 ns time window is measured to be roughly one million. The interface between the LEDs and the clear fibers will consist of a block of black Plexiglas placed over the LEDs and mounted on the VTBG-A. Each LED will feed into a single fiber, and each fiber will receive light from only one LED. Typically, the number of photons emitted by an LED over the 19 ns pulse width will greatly exceed the number of photons we want to enter the clear fiber. A neutral density filter will be placed at the interface of the LEDs and fibers to limit the intensity. This scheme was successfully employed in a previous project at DØ several years ago. [67]

### F.2.5 Diagnostic features

The control board allows the user at the PC to write directly to memory, and to subsequently read out exactly what was written. This should enable the user to diagnose any problems that relate to the loading of the static RAMs. A header will be placed on one set of 16 traces leading from a memory chip to its corresponding DAC for diagnostic purposes. This will allow us to detect a possible error in the digital output of the pulsing sequence, before conversion to an analog signal.

### F.2.6 Calibration

Many factors (including temperature variation on the board, varying thickness of the neutral density filter, and variance among DAC outputs) may cause a considerable variance in luminosity among the LEDs we use. Since we will need to control the luminosity of the LEDs fairly precisely, it is crucial that we have some reasonable system of calibration. Calibration is performed using the entire test stand. We pulse each LED with the same signals and read back, through the VRB, the signal detected. These signals are then compared to compute an offset signal for each channel. The software on the PC then modifies each digital signal in memory by this offset signal, and pulses each LED with the recalculated signal. This time, the detected signals should be of the same magnitude.

Of course, the different signals may be the fault of a piece of the electronics in the test stand. If so we want to know about it, and we do not want to simply calibrate it away. We can determine whether the LEDs or a piece of electronics is causing the difference by manually switching which LEDs pulse into a given bundle of fibers. A problem that is the fault of the LED will remain with the corresponding LED; a problem that is the fault of a particular channel will remain with the corresponding fiber.

## F.3 Schematics, equations, and data sheets

A design for the VTBG, consisting of 11 hierarchical schematics, has been drawn using Orcad Capture 7.20. These schematics are archived at DØ with document number 3823-113-ED-330231. A design for the VTBGA, consisting of 3 hierarchical schematics, has also been drawn using Orcad Capture 7.20. These schematics are archived with document

number 3823-113-ED-330234. The VTBG is a standard 9U board; the VTBG-A is a 6U by 280mm board. The VTBG contains two CPLDs from Altera: one for controlling the pulsing sequence and handling logic within the interface to VME, and one spare to accommodate additional logic if needed.

## F.4 Final product

The first prototype VTBG and VTBG-A boards met specifications with minimal changes. Four VTBG and four VTBG-A boards were then produced, each of which also met specifications, allowing the simultaneous testing of 512 channels with one spare board. The boards were produced on time and within budget.