Measurement of the Top Quark Mass at DØ

A Dissertation Presented

by

Scott Stuart Snyder

to

The Graduate School

in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

in

Physics

State University of New York
at
Stony Brook

May 1995

State University of New York
at Stony Brook

The Graduate School

Scott Stuart Snyder

We, the dissertation committee for the above candidate for the Doctor of Philosophy degree, hereby recommend acceptance of the dissertation.

---

dissertation director
Professor Paul Grannis
Department of Physics

---

chairman of defense
Professor Michael Rijssenbeek
Department of Physics

---

committee member
Professor John Smith
Department of Physics

---

outside member
Professor C. H. Sah
Department of Mathematics

This dissertation is accepted by the Graduate School.

Graduate School

Abstract of the Dissertation
Measurement of the Top Quark Mass at DØ

by

Scott Stuart Snyder

Doctor of Philosophy

in

Physics

State University of New York at Stony Brook

1995

The DØ experiment has recently reported the discovery of the standard model top quark in proton-antiproton collisions with a center of mass energy of 1.8 TeV, based on an integrated luminosity of approximately $50\,\mathrm{pb}^{-1}$ accumulated during the period 1992–1995. This work describes a measurement of the mass of the top using the lepton + jets channels of this data. The result is $m_t = 199^{+19}_{-21}(\mathrm{stat.})^{+14}_{-21}(\mathrm{syst.})\,\mathrm{GeV}/c^2$.

To my parents.

Contents

List of Figures

List of Tables

Acknowledgements

First of all, I would like to thank all the members of the DØ collaboration both for their hard work building and operating the detector and for putting up with me for the past five years.

Special thanks go to my advisor, Paul Grannis, for all the kindness and wisdom he has given to me. I find it hard to imagine a better advisor.

Thanks to Tom Ferbel, Myungyun Pang, Harrison Prosper, Mark Strovink, and others in the top mass group and the top group as a whole for many valuable suggestions.

Thanks to the many friends I've made here, who have helped me in innumerable ways: my housemates, Jim Cochran and Srini Rajagopolan; Paul Rubinov; Joey and Rita Thompson; Dhiman Chakraborty, Chris Klopfenstein, and the rest of the old Stony Brook crowd; and Meena Narain. Thanks to Paul Russo and John Featherly, who ably kept the computers running.

Thanks to Paul Rubinov, Jim Cochran, and Sailesh Chopra for reading early drafts of pieces of this thesis.

Thanks to the Free Software Foundation and others for a wide variety of hackable software tools.

Thanks to all my teachers, especially John Kolena, who taught astronomy and modern physics during my senior year of high school. It is primarily due to him that I ended up in this crazy field rather than something like computer science.

Finally, thanks to my parents for everything.

Chapter 1

Introduction

With the recent announcements of the discovery of the top quark by the DØ and CDF collaborations [1, 2], interest now turns towards improving the measurements of its properties, such as its production cross section and mass. This thesis presents a detailed account of the mass measurement published in [1].

The general plan is as follows. Chapter 2 summarizes the role of the top quark in the Standard Model, and discusses aspects of its production and decay. Chapter 3 discusses the experimental apparatus used for this analysis, the DØ detector. Chapter 4 describes the algorithms used to identify final-state objects such as leptons and jets, while Chapter 5 summarizes the analyses of [1] which provide the main evidence for the existence of top, and identifies the samples of candidate events which will be used for mass fitting. Chapter 6 discusses the mass fitting algorithm itself and explores various Monte Carlo studies which have been done on the algorithm, and Chapter 7 explains the maximum likelihood analysis used to extract the final mass value. Chapter 8

applies the techniques from Chapters 6 and 7 to the data samples identified in Chapter 5. Chapter 9 summarizes the results and discusses the prospects for future improvements. Finally, two appendices (A and B) present some mathematical details of the constrained fit method and the detailed kinematic parameters of the candidate events.

Chapter 2

The Top Quark in the Standard Model

This chapter discusses the role of the top quark in the standard model, including the characteristics of its production and decays. A short review of past experimental results is also included.

2.1 A Brief Tour of the Standard Model

The *standard model* of particle physics is a description of nature at very small distance scales, typically scales smaller than that of an atomic nucleus ($\approx 10^{-15}$ m). To the extent that the predictions of the theory can be calculated, the standard model provides a good description of all known phenomena in this regime.

Mathematically, the standard model is a theory of interacting quantum fields. Excitations in these fields correspond to particles, and each separate field corresponds to a different type (or *flavor*) of particle. See Table 2.1 for a listing of the particle types of the standard model. These can be divided into

	symbol	name	mass (MeV/c^2)	charge (e)
Quarks	d	down	≈ 10	$-1/3$
(spin = 1/2)	u	up	≈ 5	2/3
	s	strange	≈ 200	$-1/3$
	c	charm	≈ 1500	2/3
	b	bottom	≈ 4500	$-1/3$
	t	top	$\approx 175 - 200$ GeV/c^2	2/3
Leptons	e	electron	0.511	-1
(spin=1/2)	ν_e	electron neutrino	< 7 eV	0
	μ	muon	105.7	-1
	ν_μ	muon neutrino	< 0.27	0
	τ	tau	1777	-1
	ν_τ[a]	tau neutrino	< 31	0
Gauge bosons	γ	photon	0	0
(spin = 1)	W	W	80.2 GeV/c^2	1
	Z	Z	91.2 GeV/c^2	0
	g	gluon	0	0
Higgs	H[b]	Higgs	?	?

Table 2.1: Particles of the Standard Model [3, pp. 1191–1195]

[a] Not yet definitively observed.

[b] Not yet observerved.

three major groups: quarks, leptons, and gauge bosons. The quarks and leptons are all spin-$\frac{1}{2}$ particles, and thus obey the Pauli exclusion principle; they make up what is usually thought of as 'matter'. Both the quarks and leptons are grouped into three *generations* of two particles each. The corresponding particles in each generation have similar properties, except for their masses, which increase with each successive generation. All 'normal' matter (protons, neutrons, and electrons) is composed of particles from the first generation. Particles in higher generations can be produced in high-energy interactions (such as when cosmic rays hit the upper atmosphere), but they are unstable and ultimately decay into first generation particles or photons.

Each generation of leptons consists of one charged particle (the electron, muon, and tau) and an associated uncharged particle (the neutrinos). Experimentally, masses of the neutrinos are constrained to be quite small; the standard model assumes that they are zero. The charged leptons interact electromagnetically, but the neutrinos are affected only by the weak interaction (see below). This implies that for most purposes, neutrinos cannot be detected directly. Their presence must be inferred from an imbalance in the total measured momentum.

Quarks have two major features which qualitatively separate them from leptons. First, they have fractional electric charge — either 1/3 or 2/3 the charge of an electron. Second, they are affected by the strong force, which binds quarks together inside nuclei, and is described in more detail below.

The particles of the third major class, the *gauge bosons*, are responsible for the interactions between particles. The equations of the standard model

couple the fields of each gauge boson with fields of all the particles which feel that particular force. Interactions between two particles thus involve two couplings of the particles to the gauge boson; this can be viewed as a process in which the two particles exchange a virtual gauge boson.

Electromagnetism ('quantum electrodynamics' or 'QED'), for example, is mediated by the photon, which couples to particles which have electric charge. An additional feature of electromagnetism is the fact that the coupling strength is not constant: it increases as the energy involved in the interaction increases. This is called a *running coupling*, and is a general feature of interactions in the standard model.

The weak interaction is mediated by the W and Z bosons. Unlike the photon, which is massless, the W and Z are quite heavy, with masses close to 100 GeV/c^2. This implies that unlike electromagnetism, the weak force will operate only at rather small distance scales. One of the major features of the standard model is the fact that it treats the weak force and electromagnetism in a unified manner; these two forces are often referred to collectively as the 'electroweak' force.

The strong force ('quantum chromodynamics' or 'QCD') is mediated by gluons. Gluons couple to objects which possess 'color' charge, which are the quarks plus the gluons themselves. A color charge has three possible values, conventionally called 'red,' 'green,' and 'blue' (for quarks; antiquarks come in 'anti-red,' 'anti-green,' and 'anti-blue'). As is the case for the electromagnetic interaction, the value of the strong coupling runs. However, the direction of the effect is different: as the energy of the interaction increases, the strength

of the coupling gets smaller. This has the desirable consequence that at the high energies typical of modern high-energy experiments ($E \gtrsim 10\,\mathrm{GeV}$), quarks behave nearly like free particles ('*asymptotic freedom*'), and the behavior of the strong force can be calculated using the same sort of perturbative techniques as are used for electromagnetism. However, at lower energies (such as would be typical of quarks bound in a nucleon) the coupling strength becomes large enough that perturbation theory breaks down. The behavior of such systems is presently not calculable from first principles.

The fact that the strength of the strong interaction increases as the energy of the interaction decreases, or equivalently, as the distance scale of the interaction increases, also ensures that at distance scales larger than a nucleon, quarks always appear in bound states. It is thought that these bound states are always arranged so that the color charges exactly cancel and have integral electric charge (either a quark and and its antiquark with the opposite color, or a mixture of all three colors).

In order to pull a quark out of a bound state such as a nucleon, one must expend sufficient energy to create a new quark-antiquark pair, one of which will pair with the removed quark, and the other of which will take the place of the removed quark. This means that if a quark is produced or knocked out of a nucleus in some interaction, it will rapidly 'clothe' itself with other quarks which bind together to form a collection of composite particles. (This process is usually called *fragmentation.*) Experimentally, what one 'sees' is not a single quark or gluon, but a collimated *jet* of many hadronic particles moving along directions close to that of the original quark.

The remaining force is gravity, and the particle which mediates it has been named the 'graviton'. The bad news is that at present there is no workable theory of quantum gravity. The good news is that, since gravity is so much weaker than the other three forces, it is completely ignorable in almost all experiments of interest in high-energy physics.

The remaining ingredient of the standard model is the Higgs boson. The standard method of introducing a new interaction into models like the standard model (by demanding a gauge symmetry) requires that the associated gauge bosons be massless. This is a problem for the case of the weak force, since the W and Z bosons must be quite massive in order to explain the observed low-energy behavior. The Higgs mechanism is a way around this problem; it introduces a new scalar particle which interacts with the W and Z in exactly the right way so that they acquire mass. The quarks and leptons can also acquire masses through this mechanism. If this description is correct, the Higgs should appear as a real, observable particle. To date, however, it has not been directly observed.

Within the standard model, the top mass, the Higgs mass, and the ratio of the W and Z masses are interrelated, as shown in Figure 2.1. Although the dependence on the Higgs mass is weak (logarithmic), it is apparent that a sufficiently accurate measurement of the top mass can constrain the allowable range of Higgs masses.

For more information, there are many good texts which one can consult; for example, [5, 6, 7, 8].

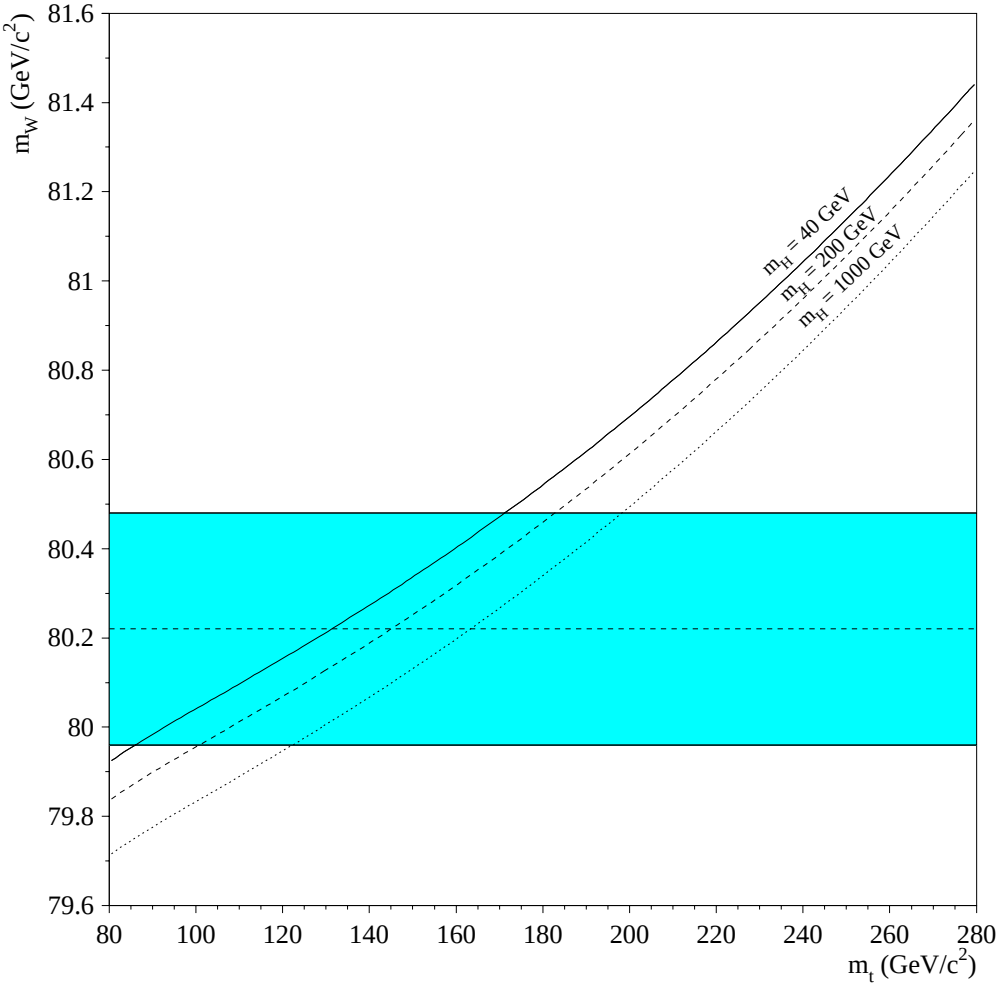

Figure 2.1: The Standard Model relation between top mass and W mass, for different values of the Higgs mass. [4] Calculated using the LEP Z mass of $91.187\,\mathrm{GeV}/c^2$. The shaded band indicates a W mass of $m_W = 80.22 \pm 0.26\,\mathrm{GeV}/c^2$ (from [3, p. 1191]).

2.2 Why Top Must Exist

As of last year, the existence of five quark flavors (d, u, s, c, and b) had been firmly established. There is also now good evidence for the existence of a sixth quark (t). However, even in the absence of direct evidence for top, there would be good reason to believe that it exists, as otherwise the standard model would not be consistent [5, chapter 25] [9].

2.2.1 Anomalies

In the standard electroweak model, if one looks at triangle diagrams of the form illustrated in Figure 2.2, one finds that they diverge in a way that cannot be removed by the usual renormalization procedures [6, 7]. However, each flavor of fermion which can 'run around' inside the loop makes a separate contribution to this process. It turns out that if one adds up the contributions from each fermion in a generation, then the divergences will exactly cancel, provided that the sum of electric charge for these fermions is zero [6, pp. 137–139] [7, pp. 376–384]. (Note that each quark gets counted three times, once for each color state.) Thus, the presence of top would ensure the cancellation of these anomalies. However, while this is the simplest way to eliminate this problem, it may not be the only way, so this argument in itself is not conclusive [9].

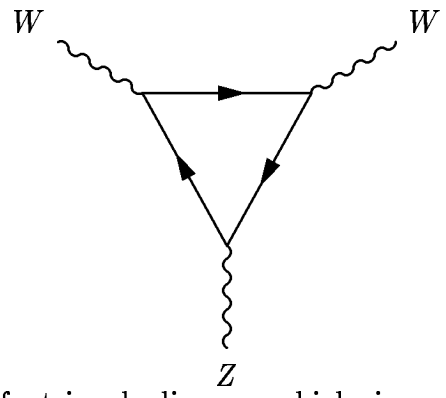

Figure 2.2: Example of a triangle diagram which gives rise to a chiral anomaly.

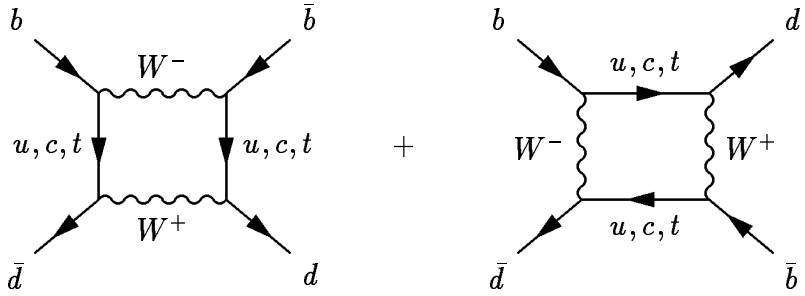

Figure 2.3: $B^0 - \bar{B}^0$ mixing box diagrams.

2.2.2 $B^0 - \bar{B}^0$ Mixing

The B^0 and $\bar{B}^0$ mesons can mix with each other through box diagrams involving internal quark lines, as shown in Figure 2.3. Within the standard model, the top contribution dominates; in fact, a fairly heavy top ($m_t \gtrsim 45\,\mathrm{GeV}$) is required to explain the observed level of $B^0 - \bar{B}^0$ mixing [10]. However, it is possible to build models in which other states can supply the needed contribution [9], so this argument is also, on its own, not conclusive.

2.2.3 Forward-Backward Asymmetry in $e^+e^- \to b\bar{b}$

In the context of the electroweak standard model, particles are grouped into $SU(2)$ spin multiplets. Each particle (or more precisely, each helicity state of each particle) thus has a spin quantum number called the *weak isospin* (denoted T_3) which is integral if there are an odd number of particles in the multiplet and half-integral otherwise. The strength of the weak interaction depends in part on the value of T_3, so T_3 can be measured under some conditions. This has been done for the reaction $e^+e^- \to b\bar{b}$, which can be mediated by either a photon or a Z^0. The interference between these two processes gives rise to an asymmetric angular distribution for b production; the amount of this asymmetry depends on the weak isospin of the left-handed b-quark T^b_{3L}. Experimentally, the data [11, 12, 13] [5, p. 268] give $T^b_{3L} = -0.504^{+0.018}_{-0.011}$, implying that the left-handed b-quark is a member of an isospin doublet. Its partner is, by definition, the top quark.

2.2.4 Bottom Decays

Suppose that top did not exist and both helicity states of b were $SU(2)$ singlets. Then the b would not interact with W's and could not decay via the usual weak processes. But the b is observed to decay; the only way the standard model could be made to accommodate this is to postulate that the b mixes with a lighter quark through some mechanism; the lighter quark, being in a $SU(2)$ doublet, could then decay normally through a virtual W (see Figure 2.4). However, if this were to be the case, then the corresponding process involving

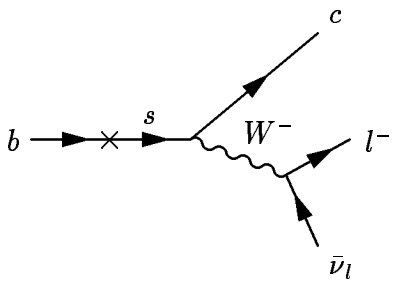

Figure 2.4: Hypothetical b-decay occuring via quark mixing.

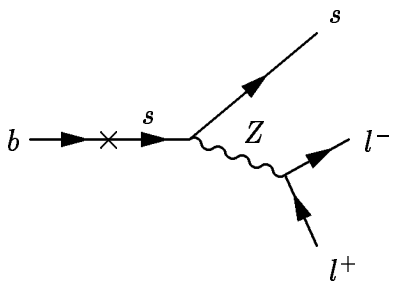

Figure 2.5: Additional b-decay which would occur if the process in Figure 2.4 was allowed.

a Z (shown in Figure 2.5) would also be present, with a cross section of at least 12% that of the first process [14]. However, the experimental upper limit for this ratio is several orders of magnitude less than this value [15, 16]. This is a further indication that the bottom must be in a $SU(2)$ doublet with the top.

2.3 Top Production and Decay

At the Tevatron, the dominant mechanisms for top production are expected to be the pair production processes $q\bar{q} \rightarrow t\bar{t}$ and $gg \rightarrow t\bar{t}$ [17]. The lowest order diagrams for these processes are illustrated in Figure 2.6. Of these two, the $q\bar{q} \rightarrow t\bar{t}$ process dominates over $gg \rightarrow t\bar{t}$ for the top masses of interest [17].

The top production cross section has been computed in perturbative QCD for both the α_s^2 leading order (LO) and α_s^3 next-to-leading order (NLO) terms [18, 21]. In a regime where perturbation theory is valid, the NLO contribution should be small compared to the LO terms. However, for top

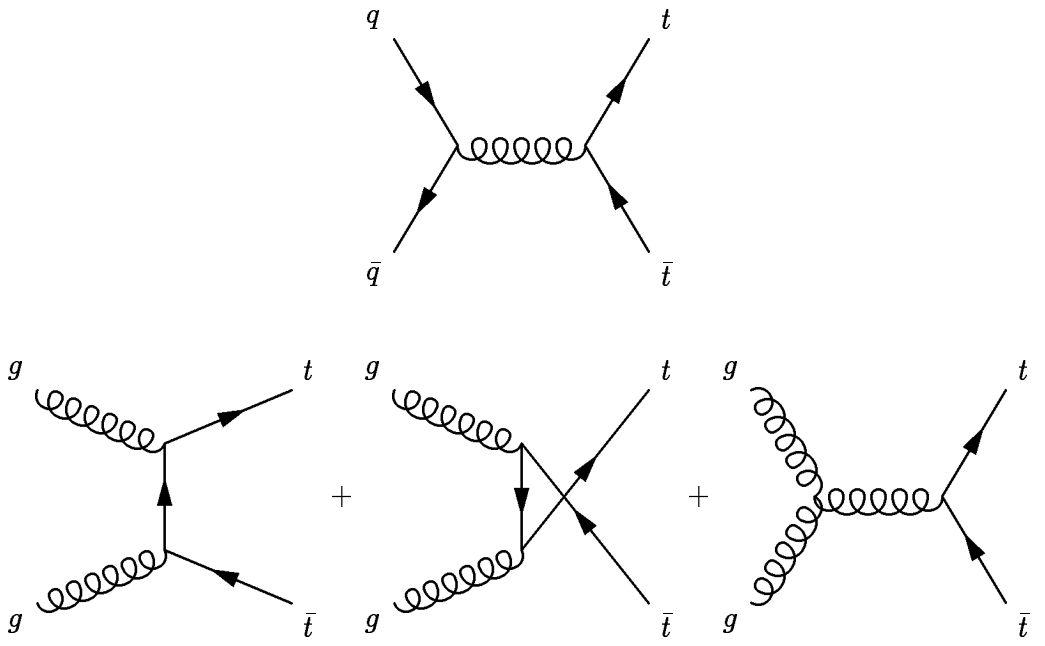

Figure 2.6: Lowest order $t\bar{t}$ production processes.

production at the Tevatron, the NLO contributions are worryingly large: for the gg process, the NLO contribution is about 70% of the LO terms [17]. (The situation is better for the $q\bar{q}$ process, where the size of the NLO contribution is about 20% that of LO.) The major contributor to the large difference between LO and NLO is found to be processes involving the emission of soft initial state gluons. Fortunately, it is possible, through a technique called *resummation*, to calculate the sums of the dominant logarithms from these processes to all orders in perturbation theory. This calculation has recently been carried out by Laenen, Smith, and van Neerven; the results are summarized in [17]. A plot of the resulting total cross section for top pair production is shown in Figure 2.7. This result depends on two unknown scale factors. The first is the usual renormalization scale μ. The second is required by the fact that the gluon series being resummed eventually diverges due to nonperturbative effects when

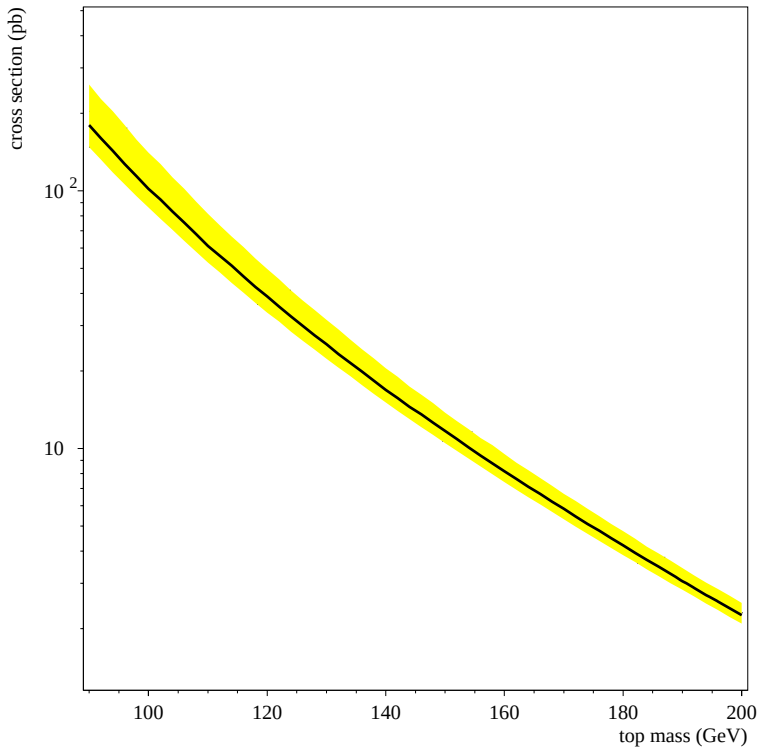

Figure 2.7: The $t\bar{t}$ production cross section at $\sqrt{s} = 1.8$ TeV. This is the NLO calculation with soft gluon resummation from [17]. The shaded region shows the upper and lower limits of the quoted error bars.

α_s becomes large. The solution is simply to stop the resummation at a specific scale μ_0, which should be bounded roughly by the QCD scale Λ on the low end and by m_t on the high end. The value chosen for this parameter may be different for the $q\bar{q}$ and gg channels. The scales chosen for the central value in the plot are $\mu = m_t$, $\mu_0(q\bar{q}) = 0.1m_t$, and $\mu_0(gg) = 0.25m_t$. For the upper limit, the nonperturbative scales are lowered to $\mu_0(q\bar{q}) = 0.05m_t$ and $\mu_0(gg) = 0.2m_t$. The lower limit was determined by taking the soft gluon series out to only one additional term ($O(\alpha_s^4)$) rather than summing the entire series [23].

	$W \to e\nu_e$ (1/9)	$W \to \mu\nu_\mu$ (1/9)	$W \to \tau\nu_\tau$ (1/9)	$W \to q\bar{q}$ (2/3)
$W \to e\nu_e$ (1/9)	1/81	1/81	1/81	2/27
$W \to \mu\nu_\mu$ (1/9)	1/81	1/81	1/81	2/27
$W \to \tau\nu_\tau$ (1/9)	1/81	1/81	1/81	2/27
$W \to q\bar{q}$ (2/3)	2/27	2/27	2/27	4/9

Table 2.2: Possible decay modes for a $t\bar{t}$ pair.

Within the standard model, a top with $m_t > m_W + m_b$ will almost always decay into a (real) W and a b [9]. (The presently known values for the elements of the CKM mixing matrix [3, p. 1315] imply that the branching ratio for $t \to W + s$ is less than about 0.2%; $t \to W + d$ should be smaller still.) The b will form a jet, while the W will decay into either a lepton-neutrino pair or a quark-antiquark pair. To a good approximation, each possible final state of the W is equally probable; however, one must remember to count each quark flavor three times, since quarks come in three colors. Thus, the probability for the W to decay into each of the three lepton flavors is about 1/9, and the probability for it to decay into each of the two available quark final states is about 1/3.

Since there are two tops in each event, and since the W's decay independently, the events may be classified based on how the W's decay as follows (see Table 2.2).

- Events in which both W's decay leptonically are called *dilepton* events. Since taus are difficult to identify, the particular dilepton channels which have received the most attention are the ee, $\mu\mu$, and $e\mu$ channels. These

final states are expected to have small backgrounds (especially $e\mu$). However, as can be seen from Table 2.2, they also have a rather small branching ratio, with all three of these channels comprising only about $4/81 \approx 4.9\%$ of top decays. Dilepton events also have the drawback of containing two unobserved neutrinos in the final state, which prevents one from completely reconstructing the decay.

- Events in which one W decays leptonically and the other decays into quarks are called *semileptonic* or *lepton + jets* events. The particular channels which have been examined are the e+jets and μ+jets channels. Compared to the dilepton channels, the lepton + jets channels have a much larger cross section — the branching ratio for each lepton + jets channel is about $4/27 \approx 15\%$. Since the final state contains only one neutrino, there is sufficient information to completely reconstruct the decay (once a particular jet assignment is assumed). The drawback to these channels, however, is a large background from single-W production with associated jets. This final state is examined in more detail in the next section.

- Events in which both W's decay into quarks are called *all-jets* events. This channel boasts the largest cross section. Unfortunately, that is more than countered by a huge background from QCD multijet processes.

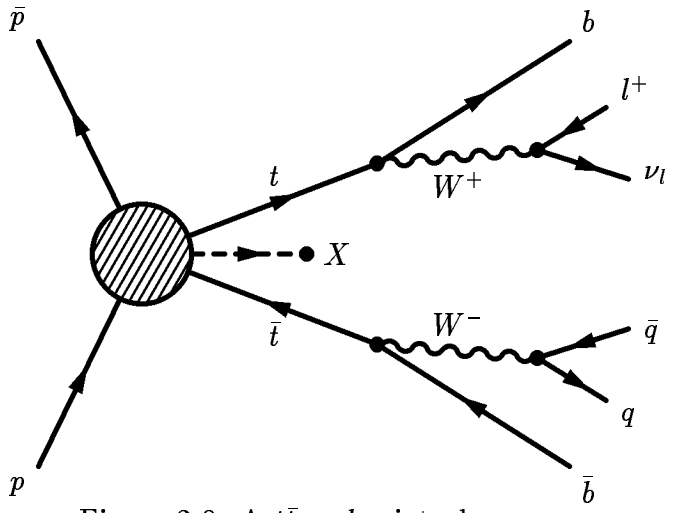

Figure 2.8: A $t\bar{t} \rightarrow l +$ jets decay

2.4 Lepton + Jets Decays

The decay modes used in this analysis are the lepton + jets modes. As illustrated in Figure 2.8, this final state contains a lepton, a neutrino, two b-jets, and two additional jets from the hadronic W decay. In addition, it is not unlikely that there will be still more jets arising from QCD radiation from either one of the incoming partons or one of the final state quarks.

The neutrino can not be detected directly; however, its presence can be inferred from a momentum imbalance in the plane transverse to the beam. (Note that one cannot apply momentum balance arguments in the direction along the beam, since the spectator quarks in the $p\bar{p}$ collision will always carry off a large amount of energy, most of which escapes down the beam pipe.) The experimental signature thus consists of one high-p_T isolated lepton, a substantial imbalance in total transverse momentum (indicating a neutrino), and several energetic jets.

There are two major backgrounds to this channel. The first is single-W

production with additional radiated jets. The second is QCD multijet production, where one jet is misidentified as an electron, and, in addition, the missing E_T is substantially mismeasured. The former can be reduced by cutting on the number and energies of the jets as well as on various topological quantities; and the latter can be reduced by tightening the electron identification cuts and by raising the missing E_T threshold.

It is also worth noting that each top decay will contain two b-jets. One might then ask if b-jets can be distinguished from other types of jets (this is called *tagging*). This would be useful for separating top decays from background since the W + jets and QCD backgrounds do not contain many heavy quarks. It would also help in the problem of properly assigning jets to parent partons within a top decay. One way of doing this tagging is to exploit the fact that the B-meson lifetime is large enough (about 1.5 ps [3, p. 1207]) that the b-decay vertex can sometimes be separated from the primary event vertex. (A B with an energy of ≈ 20 GeV will have $\gamma \approx 4$, so it can cover a distance of about 2 mm in the lab frame in a 1.5 ps proper time.) The precision required of this measurement usually demands a silicon vertex detector, and is difficult under any conditions in the crowded environment of a top decay.

Another way of identifying b-jets is to notice that a b decay will have a muon in the final state about 22% of the time. Thus, one can identify b-jets by looking for muons embedded in jets. (One can in principle do the same thing for electrons; however, it is much more difficult to identify a nonisolated electron.)

In the lepton + jets final states, if one knows the proper assignment of

jets, one can completely reconstruct the decay, even though one component of the neutrino is not measured. In fact, the problem is overconstrained by two equations. Let's stop and enumerate the variables and constraints in this type of event.

Referring back to Figure 2.8, one can identify 13 particles: p, $\bar{p}$; t, b, W^+, l^+, ν_l; $\bar{t}$, $\bar{b}$, W^-, q, $\bar{q}$; and the pseudoparticle X. Each particle has four kinematic variables, so the total number of variables is $13 \times 4 = 52$.

Now, what do we know? First of all, we directly measure the energy and direction of most of the final-state particles. Also assume that the masses of these particles are known.

b, $\bar{b}$, q, $\bar{q}$, l ($\times 3$ components each):	15
$+$ one mass constraint for each:	5
$+$ the two components of $\not\!\!E_T$:	2

We also know the initial four-vectors of the p and $\bar{p}$:

p, $\bar{p}$ initial conditions ($\times 4$ components each):	8

We have total four-momentum conservation at each of the five internal vertices:

5 vertices ($\times 4$ components each):	20

We know the masses of the W's and the ν:

W^+, W^-, and ν masses:	3

Finally, we know that the two tops have the same mass:

$m_t = m_{\bar{t}}$	1

Adding these up, we get

Total	54

Thus, we have 52 variables and 54 constraints; i.e., the system is overconstrained by two.

If one solves out all trivial constraints, one is left with the 17 measured variables (three variables each for the lepton and the four jets, and two components of $\not\!\!E_T$) and one unmeasured variable (the z-component of the neutrino momentum). There are then three constraints left: the two W-mass constraints, and the $m_t = m_{\bar{t}}$ constraint.

2.5 Experimental Searches for Top

2.5.1 Mass Limits

Although top has now been observed, it is still interesting to review the limits which have been placed on its mass. The best experimental lower limit for the top mass is given by DØ [24, 25] and is $m_t > 128\,\mathrm{GeV}/c^2$ with a 95% confidence limit. This search assumes standard model decays of top; if it were to have other decay modes (such as $t \to b + H^+$), it could have been missed by this search. However, a measurement of the total width of W decays gives a model-independent limit, since any decays of the form $W \to t + \bar{b}$ would contribute to this width. Measurements of $\Gamma(W)$ at DØ and CDF yield a model-independent limit of $m_t > 62\,\mathrm{GeV}/c^2$ (95% CL) [26, 27].

Strictly speaking, the analysis that yields the $m_t > 128\,\mathrm{GeV}/c^2$ limit also assumes that $m_t > m_W + m_b$. However, the range from $40\,\mathrm{GeV}/c^2 < m_t <$

91 GeV/c^2 is excluded (at 95% CL) by earlier CDF searches [28, 29]. In addition, the range $m_t < 45.8\,\mathrm{GeV}/c^2$ is excluded (95% CL) by searches in e^+e^- collisions at SLC and LEP [30, 31].

2.5.2 Standard Model Predictions

Since the top mass is correlated with other electroweak observables, it is possible to make a prediction for the top mass using the high-precision electroweak measurements which have been made at LEP and elsewhere (assuming, of course, the validity of the standard model). A recent such fit [32] gives $m_t = 178^{+11\,+18}_{-11\,-19}\,\mathrm{GeV}/c^2$, where the central value assumes a Higgs mass of $m_H = 300\,\mathrm{GeV}/c^2$. The first set of error bars results from propagating the errors on the measured quantities which are input to the fit, and the second set of error bars is the result of varying m_H from $60\,\mathrm{GeV}/c^2$ to $1000\,\mathrm{GeV}/c^2$.

2.5.3 Observations

In 1994, following the 1992–1993 '1A' collider run, the CDF collaboration announced evidence for top production, with a cross section of $13.9^{+6.1}_{-4.8}\,\mathrm{pb}$ and a mass of $174 \pm 10^{+13}_{-12}\,\mathrm{GeV}/c^2$ [33, 34]. The statistical significance of the result, however, was too small to claim a discovery. The DØ collaboration also reported an excess of observed events over background with a cross section of $8.2 \pm 5.1\,\mathrm{pb}$ [35, 25], but with a still smaller statistical significance.

In March 1995, part way through the '1B' collider run, both experiments claimed observation of a top signal, DØ with a cross section of $6.4 \pm 2.2\,\mathrm{pb}$ and

a mass of $199^{+19}_{-21} \pm 22\,\mathrm{GeV}/c^2$ [1], and CDF with a cross section of $6.8^{+3.6}_{-2.4}\,\mathrm{pb}$ and a mass of $176 \pm 13\,\mathrm{GeV}/c^2$ [2].

The remainder of this thesis will describe the DØ result, especially the mass measurement, in detail.

Chapter 3

Experimental Apparatus

This chapter contains a description of the DØ detector. The 'official' reference for the detector as a whole is [36]. The reader should consider a reference to [36] implied in almost everything in the remainder of this chapter.

3.1 Coordinate Systems

In what follows, a right-handed coordinate system will be used, with the positive z-axis aligned along the beam in the direction of the protons and the positive y-axis pointing up. Cylindrical (r, ϕ, z) coordinates are sometimes used, as are spherical (r, ϕ, θ) coordinates. The angular variables are defined so that $\phi = \pi/2$ is parallel to the positive y-axis, and $\theta = 0$ is coincident with the positive z-axis. Instead of θ, it is often convenient to use the pseudorapidity η defined as

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

The pseudorapidity approximates the true rapidity,

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z}, \tag{3.2}$$

in the limit that $m \ll E$ (where m is the invariant mass $m^2 = E^2 - p^2$).

It is also often convenient to use instead of momentum, the 'transverse' momentum, which is the momentum vector projected onto a plane perpendicular to the beam axis:

$$p_T = p \sin \theta. \tag{3.3}$$

This is particularly useful due to the fact that in a $p\bar{p}$ collision, the momenta along the beam of the colliding partons are not known (since many of the products of the collision escape down the beam pipe). However, their transverse momenta are very small compared to their momenta along the beam, so one can apply momentum conservation in the transverse plane.

One can also define a 'transverse energy' by

$$E_T = E \sin \theta. \tag{3.4}$$

When treated as a vector, the direction of E_T should be taken to be the same as the p_T vector.

3.2 Of Luminosities and Cross Sections

A reaction rate measurement is usually expressed as a *cross section*, the interaction probability per unit flux. The particle flux is also called the *luminosity*; for a collider, it is proportional to the square of the number of particles

passing through a unit area per unit time. Luminosities are typically measured in units of $\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, while cross sections are often expressed in *barns*, where $1\,\mathrm{barn} = 10^{-24}\,\mathrm{cm}^2$. The luminosity at the Tevatron during run 1A was typically in the range $1 - 10 \times 10^{30}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. The reaction rate R, the luminosity $\mathcal{L}$, and the cross section σ are related by

$$R = \sigma \mathcal{L}. \tag{3.5}$$

It follows that the number of events of a specific type expected after running an experiment for a period of time is found by integrating the luminosity with respect to time:

$$N = \sigma \int \mathcal{L}\, dt. \tag{3.6}$$

The quantity $\int \mathcal{L}\, dt$ is called *integrated luminosity*; the total integrated luminosity for run 1A is $13.5\,\mathrm{pb}^{-1} \pm 12\%$. The luminosity is determined by measuring the rate of inelastic $p\bar{p}$ scatterings, which have a known observable cross section of 48.2 mb at DØ (after detection efficiencies are accounted for) [37, 38]. (This was derived by averaging measurements from E710 and CDF.)

3.3 The Beam

The DØ detector is located at the Fermilab Tevatron [39, 40, 41, 42], presently the world's highest-energy hadron collider, with a center-of-mass energy of 1800 GeV. A schematic of the accelerator complex is shown in Figure 3.1.

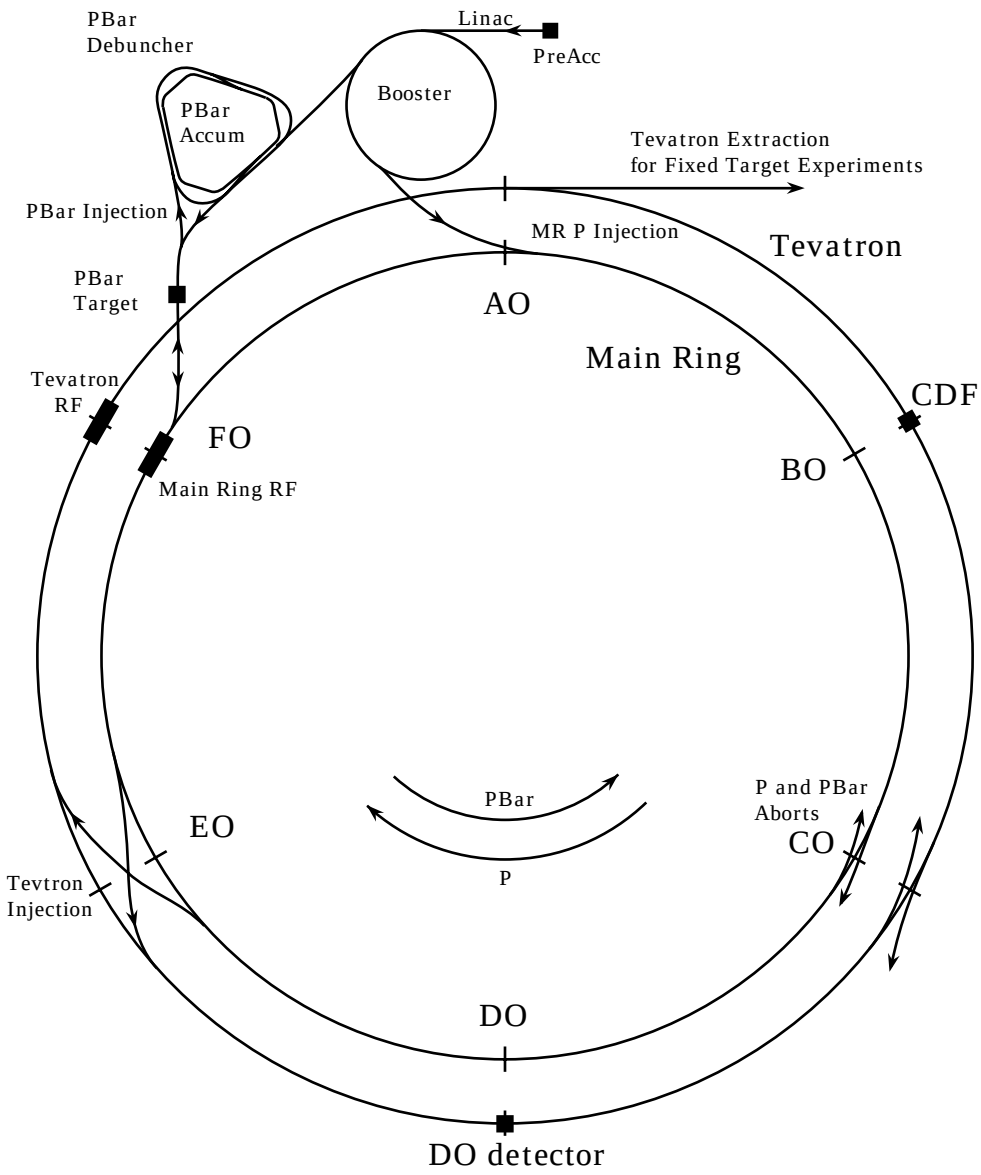

Figure 3.1: Schematic of the Fermilab accelerator complex (not to scale). [43, p. 112]

The Tevatron is a proton storage ring, composed of superconducting magnets. The ring is filled with bunches of protons and antiprotons, which circulate in opposite directions. At the B0 and D0 experimental areas, these beams are made to collide with each other. The process of filling the ring is quite complicated; a summary of the major steps is given below, but the reader should consult the cited references for more details.

The beams originate in the preaccelerator. There, H^- ions are formed and accelerated to 750 keV by an electrostatic Cockroft-Walton accelerator. The preaccelerator operates in a pulsed mode with a frequency of 15 Hz. The ions are bunched and transported to the start of the Linac. The Linac is a 150 m long linear accelerator, which boosts the energy of the ions to 200 MeV[1]. After emerging from the Linac, the ions are passed through a carbon foil which strips off the electrons, leaving bare protons. The protons are then injected into the Booster, a 151 m diameter synchrotron. (A synchrotron is a device which confines charged particles in a closed orbit using bending magnets. RF cavities can be used to increase the energy of the stored particles; when this is done, the field of the bending magnets must also be increased in a synchronous manner in order to keep the particles in the same orbit.) One of the interesting features of the Booster is its rapid cycle rate of 15 Hz. To achieve this, the magnets are combined with capacitor banks to form LC circuits which resonate at 15 Hz. The Booster accelerates the protons to an energy of 8 GeV. The protons are then injected into the Main Ring, a large (1000 m radius) synchrotron

[1]For run 1B, the Linac has been upgraded to an energy of 400 MeV.

composed of conventional magnets. The Main Ring lies mostly in a plane, except at the B0 and D0 experimental areas where it is bent into overpasses to allow room for the detectors (the separation between the Main Ring and the Tevatron is 19 feet at B0 and 89.2 inches at D0). Protons in the Main Ring can be used to make antiprotons (see below), or they can be accelerated to 150 GeV and injected into the Tevatron.

The Tevatron is a proton synchrotron made from superconducting magnets [40, 42]. It lies just below the Main Ring in the accelerator tunnel, and has a maximum beam energy of 900 GeV. (Upgrades to the cryogenic system are expected to raise this to 1000 GeV.) The Tevatron can be operated in one of two major modes. In fixed-target mode, the Tevatron is filled with protons which are accelerated and then extracted and directed towards numerous experimental areas. This cycle repeats with a frequency of about once per minute. In collider mode, the Tevatron is filled with six bunches of protons and six bunches of antiprotons, traveling in opposite directions. The beams are accelerated to the maximum energy of 900 GeV each and allowed to collide at the B0 and D0 experimental areas. (At other points where the beams would collide, they are kept apart by electrostatic separators). The beams are typically kept colliding for about 20 hours, after which the machine is emptied and refilled with new batches of protons and antiprotons.

The remaining major part of the accelerator complex is the antiproton source [41, 44], which is used to produce and store antiprotons for use in the collider. While collisions are occurring in the Tevatron, the Main Ring continually runs antiproton production cycles at a rate of one every 2.4 s.

Protons are accelerated to 120 GeV and extracted onto a nickel target. Each of these collisions produces a spray of nuclear debris, which includes some antiprotons. Immediately following the target is a lithium lens, a cylindrical piece of lithium through which a large (0.5 MA) current is passed. This generates an azimuthal magnetic field which acts to focus negatively-charged particles passing through it. Following the lens is a bending magnet which selects negatively-charged particles with energies of 8 GeV and transports them to the Debuncher. The Debuncher is a storage ring in which antiprotons are first 'debunched' (rotated in phase space from a configuration with a small time spread and large momentum spread to one with a large time spread but small momentum spread) and then stochastically 'cooled' to further reduce the momentum spread. Stochastic cooling [44, 45] operates by measuring the trajectory of collections of particles relative to the desired orbit. From this information, a correction signal is derived which is passed across the ring to kicker electrodes which apply a force on the particles to move them back towards the desired orbit. The effect on any single particle is very small due to the incoherent contribution of all the other particles near it in the beam, but when repeated over a large number of turns, the effect becomes significant. The antiprotons are kept in the Debuncher until just before the next pulse arrives, about 2.4 s later. They are then transferred to the Accumulator, another storage ring which lies inside the Debuncher. There, cooling continues for several hours, and eventually the antiprotons settle into a dense core near the inner radius of the Accumulator. When enough have accumulated to fill the Tevatron (typically on the order of $50 - 150 \times 10^{10}$), they are extracted

Accelerator radius	1000 m
Maximum beam energy	900 GeV
Injection energy	150 GeV
Peak luminosity	$\approx 10 \times 10^{30}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$
Number of bunches	6 p, 6 $\bar{p}$
Intensity per bunch	$\approx 100 \times 10^{9} p$, $\approx 50 \times 10^{9} \bar{p}$
Crossing angle	0°
Bunch length	50 cm
Transverse beam radius	43 μm
Energy spread	0.15×10^{-3}
RF frequency	53 MHz
$\bar{p}$ stacking rate	$\approx 3.5 \times 10^{10}$/hour
Beam crossing frequency	290 kHz
Period between crossings	3.5 μs

Table 3.1: Run 1A Tevatron Parameters. [46, p. 17][47, ch. 2][43, app. A]

from the Accumulator, accelerated to 150 GeV in the Main Ring, and injected in bunches into the Tevatron.

Some of the major parameters of the Tevatron for run 1A are given in Table 3.1. A more detailed introduction to the accelerator may be found in [43, appendix A].

3.4 DØ Overview

DØ is a large, multipurpose detector for studying $p\bar{p}$ collisions which has been operating at the Fermilab Tevatron since 1992. The design was optimized for the study of high-p_T physics and high mass states, and stresses the

identification and measurement of electrons and muons, the measurement of the direction and total energy of high-p_T jets, and the determination of missing transverse energy. Little emphasis is placed on identifying and tracking individual particles within jets, as the details of hadronization are not relevant to the underlying hard scattering.

Detectors for colliding beam experiments are composed of many different particle-detection devices, each with their own strengths and weaknesses. The general layout, however, is dictated by the physics of how particles interact with matter. Closest to the interaction point are the *tracking* detectors, which are devices designed to measure the three-dimensional trajectories of particles passing through them. Often, the tracking detectors are immersed in a magnetic field; this permits a determination of the momentum of the charged particles via a measurement of their bending radius. Surrounding the tracking detectors is typically a *calorimeter*; this is a device which measures the energy of particles which hit it. A calorimeter should be 'thick' so that it will absorb all the energy of incident particles; conversely, tracking detectors should contain as little material as possible so as to minimize multiple scattering and losses prior to the calorimeter. A calorimeter is typically made thick enough to stop all known particles except for muons and neutrinos. Muons are identified by the use of tracking chambers outside the calorimeter; any charged tracks originating from the interaction point and penetrating the calorimeter are likely to be muons. Neutrinos are not detected directly; their presence is inferred from an imbalance in the total detected momentum perpendicular to the beam.

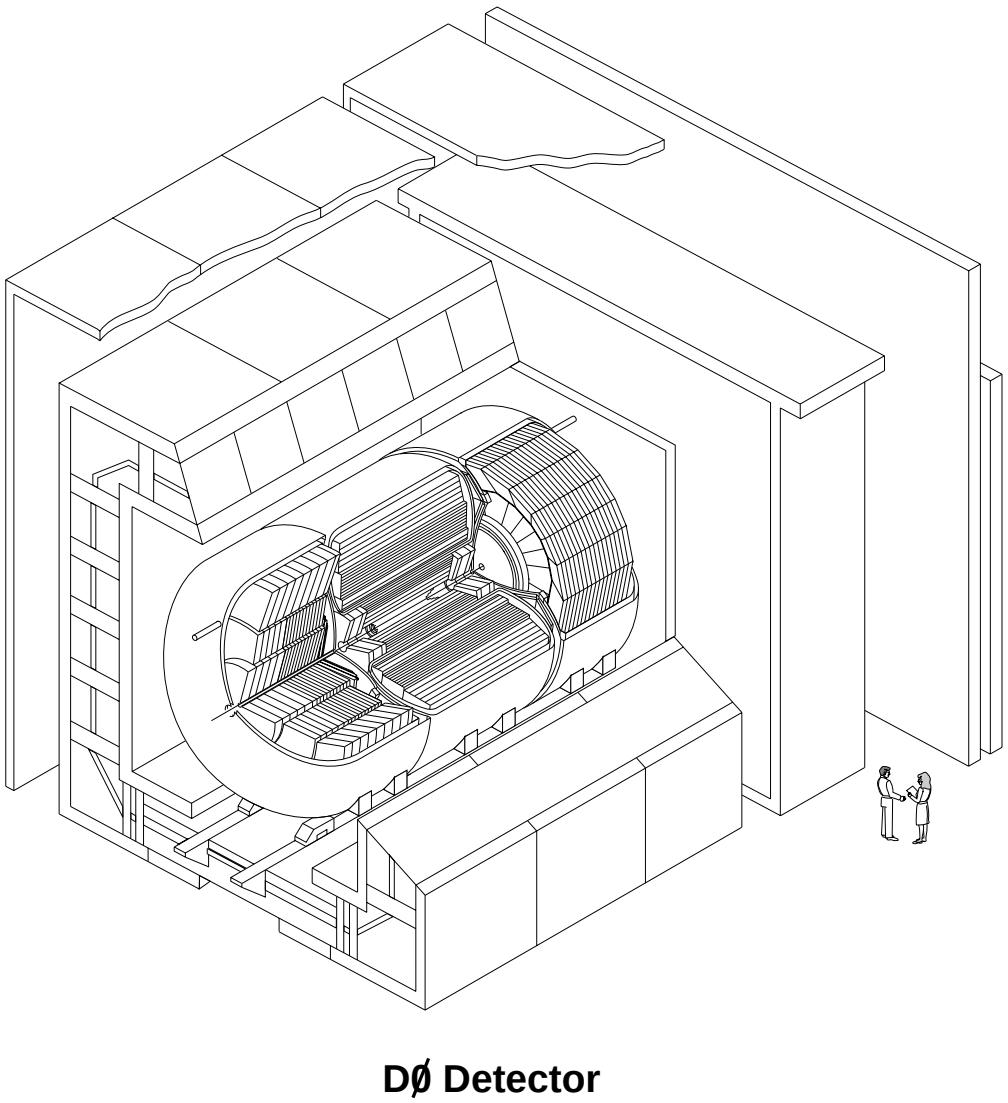

Figure 3.2: Cutaway view of the DØ detector. [36]

The DØ detector, illustrated in Figure 3.2, follows the general plan outlined above. Surrounding the beam pipe are a set of tracking detectors. At DØ, however, there is no central magnetic field. Part of the reason for this is that the momentum resolution in a magnetic tracker is roughly proportional to the momentum, while the energy resolution of a calorimeter is roughly proportional to $1/\sqrt{E}$. Thus, for high-p_T objects, a calorimeter will yield a better momentum measurement than a tracking detector. A magnetic tracker would also tend to sweep low-momentum charged particles out of jets, degrading the calorimetric energy measurement. In addition, the material of the solenoid is absent. The absence of a tracking momentum measurement, however, implies the need for a very good calorimeter. DØ achieves this using a liquid argon sampling calorimeter made mostly from depleted uranium. To identify muons, an additional set of tracking chambers is installed surrounding the calorimeter. To provide a measurement of the muon momentum, magnetized iron toroids are placed between the first two muon tracking layers.

The detector is quite large; the entire assembly is about 13 m high $\times$ 11 m wide $\times$ 17 m long with a total weight of about 5500 tons [43, p. 26]. As indicated in Figure 3.2, the Tevatron beam pipe passes through the center of the detector. Also visible is the Main Ring beam pipe, which passes through the upper portion of the calorimeter, 89.2 inches above the Tevatron beam pipe. Not shown in the figure is the detector support platform, on which the entire assembly rests. This platform is mounted on rollers so that the entire detector may be rolled from the assembly area to the collision hall. In addition, the platform provides rack space for detector electronics and other

support services.

The use of clocked electronics, such as digitizers, is kept to a minimum on the detector platform in order to reduce electronic noise. Analog signals are led from the platform, out of the collision hall, and into the Moving Counting House (MCH), a three-story structure containing the digitization electronics, the level-1 trigger, high-voltage supplies, and other services. In order to keep the cable runs short, the MCH is also mounted on rollers and follows the detector as it moves into and out of the collision hall. The detector data cables and other communications lines are led out of the MCH into the second floor of the assembly building (sometimes called the fixed counting house, or FCH).

Some of the individual elements making up the DØ detector are described below.

3.5 Central Detector

The main goal of the central tracking chambers is to reconstruct the three-dimensional trajectories of charged particles which pass through them. This information can be used to decide if an electromagnetic shower in the calorimeter was caused by an electron or by a γ/π^0. By also measuring the dE/dx of a track, one can determine if the track was actually caused by several closely spaced charged particles, such as would result from a $\gamma \rightarrow e^+e^-$ conversion. The central trackers are also responsible for making a precise measurement of

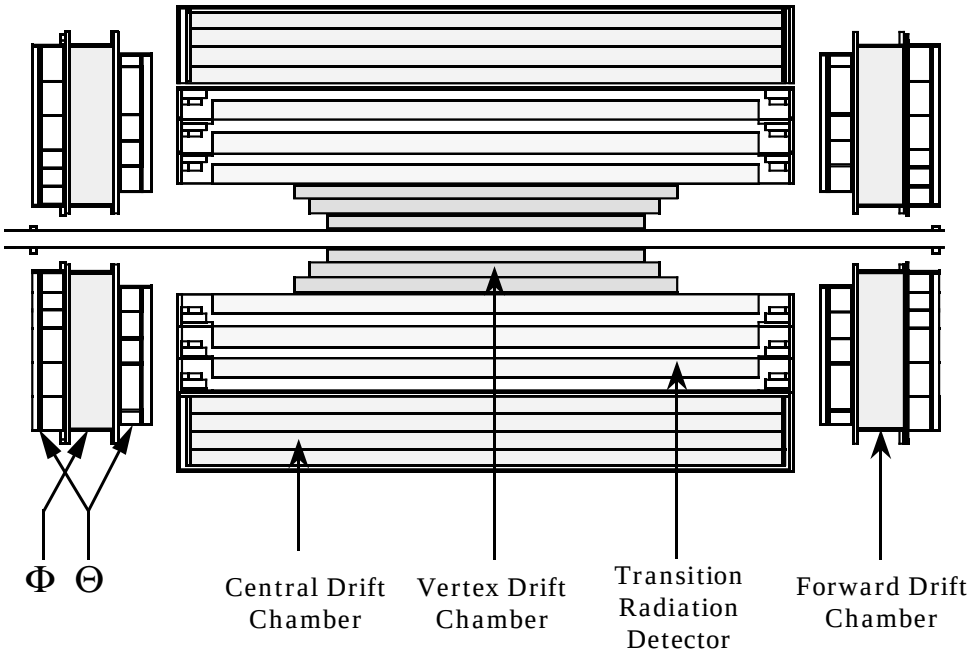

Figure 3.3: Side view of the DØ central detector. [36]

the location of the interaction vertices for each event. Precise position measurements are also useful for calibrating the calorimeter position measurements and for improving the accuracy of muon momentum measurements.

The central detector consists of four major subsystems (see Figure 3.3). Immediately outside the beryllium beam pipe is the vertex drift chamber (VTX). Surrounding that is the transition radiation detector (TRD) for electron identification. Outside that is the cylindrical central drift chamber (CDC) capped on both ends by the two disk-shaped forward drift chambers (FDC). The entire assembly is contained within a nonmagnetic cylindrical volume of radius $r = 75\,\mathrm{cm}$ and length $l = 270\,\mathrm{cm}$.

3.5.1 Drift Chamber Principles

When a charged particle passes through a gas, it will interact electromagnetically with nearby atomic electrons, resulting in the creation of electron/ion pairs along the path of the particle. The number of such pairs created depends on the energy of the particle and the type of gas, but for a typical gas at STP and a particle with unit charge, the mean number of electron/ion pairs formed will be on the order of 100/cm [48, p. 22].

If an electric field is applied, the electrons will start to drift through the gas towards the positive electrode, undergoing repeated collisions with the gas molecules. (The ions also drift in the opposite direction. However, their drift speed is much less than that of the electrons, so they can be ignored in this discussion.) If the electric field near the anode is strong enough, an electron can acquire enough energy between collisions to knock an additional electron free from a gas molecule. This additional electron can then go on to ionize more gas molecules; in this way, an avalanche is formed in which the number of electrons increases exponentially. When this avalanche reaches the positive electrode, it gives rise to a measurable current, the size of which is proportional to the original number of ions created. The ratio between the final number of electrons collected and the initial number deposited is called the *gas gain*, and for practical detectors is typically on the order of 10^4–10^6.

The large electric field needed to obtain gas amplification is usually obtained by forming the anode from a very thin (20–100 μm) wire. An electron sitting in the gas far away from the anode will see a much smaller electric field,

and will drift towards the anode with a velocity roughly proportional to the field. When it gets close to the anode, the electric field will start to rapidly increase, and the electron will initiate an avalanche.

The fact that an electron drifts with a predictable speed over most of the distance to the anode implies that one can turn a measurement of the time an electron took to drift to the anode into a measurement of the distance of the original source particle from the anode. A device designed for this type of measurement is called a *drift chamber*. In order to obtain a linear relationship between distance and time, it is necessary that the electric field be made as constant as possible over as large a volume as possible. Typically, additional field-shaping electrodes will be inserted into a drift chamber in order to make the field more uniform. One is also aided by the fact that the relationship between electron drift velocity and electric field tends to flatten out for sufficiently large electric fields. It is thus highly desirable to operate a drift chamber in this *saturation* region.

For further discussion of drift chambers and their application to high-energy physics, consult [49, 50, 51].

3.5.2 Vertex Chamber

The vertex chamber (VTX) [52, 53, 54] is the innermost of the tracking chambers. It can be used to accurately determine event vertex positions and complements the other tracking chambers by identifying conversions which occur in the TRD. It consists of four carbon fiber cylinders enclosing three

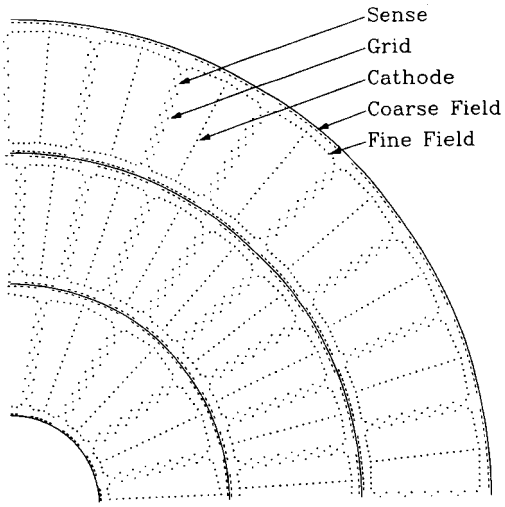

Figure 3.4: End view of one quadrant of the VTX chamber. [36]

concentric layers and occupies the region 3.7 cm $\leq r \leq$ 16.2 cm. The inner layer has a length of 97 cm, with each successive layer being about 10 cm longer.

An end view of the VTX is shown in Figure 3.4. The geometry of the VTX is called a *jet* geometry because the sense wires are arranged in planes which are parallel to the paths of particles emerging from the interaction region. The inner layer of the VTX is divided into sixteen cells; the outer two layers contain 32 cells each. Between layers, the cells are staggered in ϕ to avoid dead regions and also to aid in resolving left-right ambiguities when linking together track segments in different layers. Within each cell, there are eight sense wires, which are staggered relative to each other by 100 μm in order to further resolve left-right ambiguities. Each cell also contains a set of aluminum traces mounted on the carbon fiber tubes plus an additional set of wires for field shaping. The $r\phi$ position of a hit is determined from the drift time and

Active radius	3.7 cm – 16.2 cm
Number of layers	3
Active lengths of each layer	96.6 cm, 106.6 cm, 116.8 cm
Number of cells	16 inner layer; 32 in outer two
Number of sense wires	8 per cell, 640 total
Sense wire separation	4.57 mm radially with 100 μm stagger
Sense wire specifications	25 μm NiCoTin, 80 g tension
Sense wire voltage	+2.5 kV
Field wire specifications	152 μm Au-plated Al, 360 g tension
Gas composition	CO_2 95%, ethane 5%, H_2O 0.5%
Gas pressure	1 atm
Average drift field	1 kV/cm
Drift velocity	$\approx 7.3\,\mu$m/ns
Maximum drift distance	1.6 cm
Gas gain	4×10^4
Position resolution	$r\phi \approx 60\,\mu$m, $z \approx 1.5$ cm

Table 3.2: Vertex Chamber Parameters. [47, p. 38][43, p. 34][53]

the wire hit. The z position is determined using a technique called charge division: the resistive sense wire is read out at both ends and treated as a voltage divider. Unfortunately, this method requires that the sense wire pulses be well separated and that the cell occupancy be low, a condition which is rare in a high-multiplicity environment such as DØ. The VTX is not operated at saturation, so extra care is required with high voltage to obtain a reliable time-distance relation.

Further parameters of the VTX are given in Table 3.2.

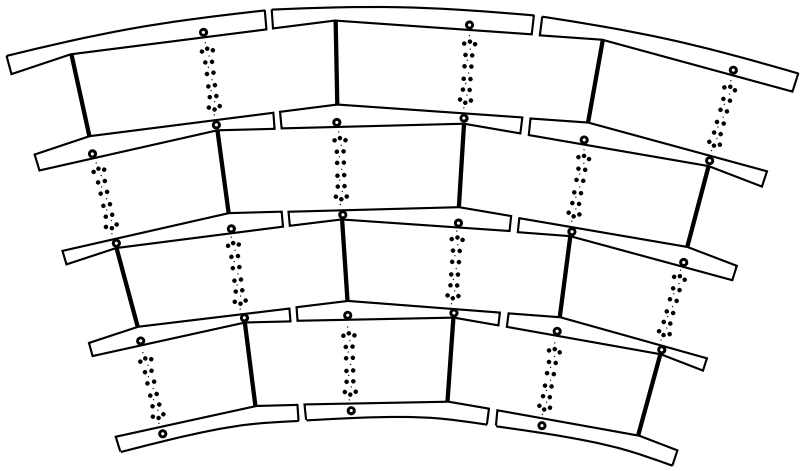

Figure 3.5: End view of 3 of 32 CDC modules. [36]

3.5.3 Central Drift Chamber

The central drift chamber (CDC) [48, 53, 55, 56] is the outermost tracker in the central region (covering a pseudorapidity range of about $|\eta| \leq 1.2$). Like the vertex chamber, it has a jet geometry. It consists of four layers of cells occupying the region $49.5\,\text{cm} \leq r \leq 74.5\,\text{cm}$, with a length of 184 cm. An end view of a portion of the CDC is shown in Figure 3.5.

The CDC is constructed from 32 separate identical modules, which are arranged in a cylindrical ring. The cylinder is contained inside of an aluminum tube, with a carbon fiber tube sealing the inner radius. The ends of the modules are capped with G10 walls; the chamber as a whole is capped by aluminum endplates which hold the wire tension. The module walls are constructed from Rohacell foam laminated with Kevlar and wrapped with Kapton. Within each cell, there are 7 sense wires, staggered by 200 μm relative to each other to help resolve left-right ambiguities. Each cell also has a set of field-shaping wires,

Active radius	51.8 cm – 71.9 cm
Number of layers	4
Active length	179.4 cm, 106.6 cm, 116.8 cm
Number of cells per layer	32
Number of sense wires	7 per cell, 896 total
Sense wire separation	6.0 mm radially with 200 μm stagger
Sense wire specifications	30 μm Au-plated W, 110 g tension
Sense wire voltage	+1.45 kV (inner SW) – +1.58 kV (outer SW)
Number of delay lines	2 per cell, 256 total
Delay line velocity	2.35 mm/ns
Field wire specifications	125 μm Au-plated CuBe, 670 g tension
Gas composition	Ar 93%, CH_4 4%, CO_2 3%, H_2O 0.5%
Gas pressure	1 atm
Average drift field	620 V/cm
Drift velocity	$\approx$ 34 μm/ns
Maximum drift distance	7 cm
Gas gain	2×10^4 (inner SW) – 6×10^4 (outer SW)
Position resolution	$r\phi \approx 180\,\mu$m, $z \approx 2.9$ mm (best case)

Table 3.3: Central Drift Chamber Parameters. [47, p. 41][43, p. 36][53]

plus a set of resistive strips silk-screened onto the Kapton cell surface. The $r\phi$ position of a hit is determined via the drift time and the wire hit. The z position of a hit is measured using inductive delay lines embedded in the module walls in the sense wire plane. When an avalanche occurs near an outer sense wire, a pulse is induced in the nearby delay line. By comparing the arrival times of the pulse at both ends, the z position can be determined.

Further parameters of the CDC are given in Table 3.3. See [48, 56] for a detailed description of the construction and testing of the chamber.

3.5.4 Forward Drift Chambers

The forward drift chambers (FDCs) [46, 53, 55, 57] extend the outer tracking coverage from where the CDC leaves off down to an angle of $\theta \approx 5^\circ$ ($\eta \approx 3.1$). There are two sets of chambers, one located at each end of the CDC. Figure 3.6 shows an exploded view of one of the FDCs. Each FDC consists of three layers of chambers: two Θ layers sandwiching one Φ layer. The Φ layer is a single chamber divided into 36 azimuthal drift cells, each containing sixteen radial sense wires arranged in a plane containing the beam line. The two Θ cells consist of four separate quadrant modules, each of which is composed of six rectangular drift cells at increasing radii. Each cell contains eight sense wires oriented in a plane parallel to the z-axis and normal to the radial direction. The inner three cells are half cells in which the wires are placed at the edge of the cell; thus the electrons can drift in only one direction. Each Θ cell also contains a delay line of the same type as in the CDC to measure the position along the length of the cell. There are no delay lines in the Φ chamber. The two Θ chambers are rotated relative to each other by an angle of $\pi/4$. The construction of the FDCs is similar to that of the CDC, except that the FDCs use 25 μm aluminum etchings instead of the CDC's epoxy strips.

Further parameters of the FDC are given in Table 3.4. See [46, 57, 58] for additional details on the construction and testing of these chambers.

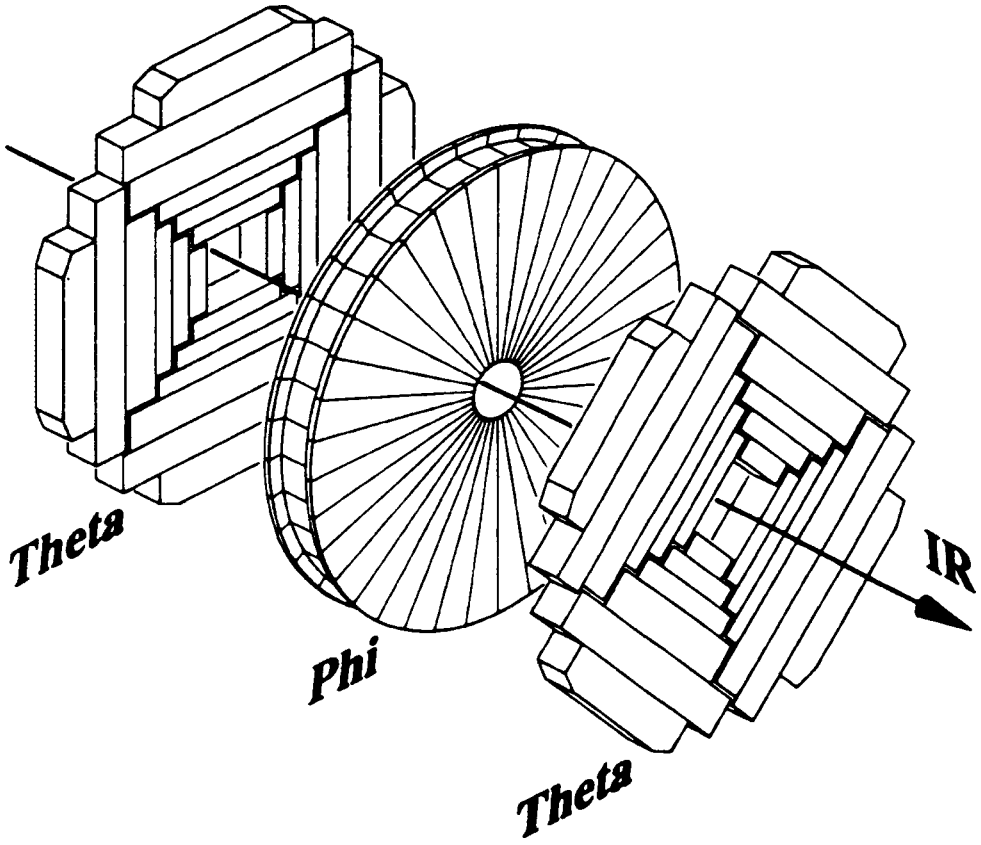

Figure 3.6: Exploded view of one of the FDCs. [36]

	Θ modules	Φ modules
Active radial extent	11 cm – 62 cm	11 cm – 61.3 cm
Active z extent	104.8 – 111.2 cm 128.8 – 135.2 cm	113.0 – 127.0 cm
Number of cells per layer	4 quadrants of 6 cells	32
Number of sense wires	8 per cell, 384/FDC	16 per cell, 576/FDC
Sense wire separation	8.0 mm radially with 200 μm stagger	
Sense wire specifications	30 μm Au-plated W, 50 – 100 g tension	
Sense wire voltage	+1.55 kV	+1.66 kV
Number of delay lines	1 per cell, 48/FDC	—
Delay line velocity	2.35 mm/ns	—
Field wire specifications	163 μm Au-plated Al, 100 – 150 g tension	
Gas composition	Ar 93%, CH_4 4%, CO_2 3%, H_2O 0.5%	
Gas pressure	1 atm	1 atm
Average drift field	1.0 kV/cm	1.0 kV/cm
Drift velocity	40 μm/ns	37 μm/ns
Maximum drift distance	5.3 cm	5.3 cm
Gas gain	2.3×10^4 (inner SW) 5.3×10^4 (outer SW)	3.6×10^4
Position resolution (drift)	$\approx$ 300 μm	$\approx$ 200 μm

Table 3.4: Forward Drift Chamber Parameters. [47, p. 45][43, p. 38][53]

3.5.5 Transition Radiation Detector

When a charged particle crosses between two materials with different dielectric constants, it radiates in the forward direction [50, p.136]. The intensity of this radiation is proportional to the relativistic gamma factor, $\gamma = E/(mc^2)$, and is concentrated on a cone with a half opening angle of $1/\gamma$. These characteristics suggest that one can use this effect to discriminate between particles which have similar energies but different masses. The radiation emitted from a particle crossing a single boundary is not readily detectable, so one must stack a large number of boundaries in order to give a reasonable signal. For highly relativistic particles, the radiation will be in the X-ray range.

The DØ transition radiation detector (TRD) [59, 60] is a device designed to distinguish electrons from heavier particles using transition radiation. It occupies the radial space between the VTX and the CDC and consists of three radial layers. Each layer has a radiator consisting of 393 layers of 18 μm polypropylene foil with a mean separation of 150 μm. The gaps between the foils are filled with dry nitrogen. Surrounding each radiator is a cylindrical xenon-filled drift chamber to detect the emitted X-rays. At the Tevatron, electrons are the only charged particles likely to be produced with sufficient energy to produce detectable transition radiation. See [60, 61] for further information about the performance of the TRD.

3.5.6 Central Detector Readout

For the first stage of signal processing, the signals from the chamber wires are led into preamplifiers mounted on the ends of the chambers themselves. From there, the signals are fed into analog pulse shaping cards located on the platform underneath the detector. Finally the signals are sent to flash analog-to-digital converters (FADCs) located in the moving counting house. There the signals are sampled and digitized at a rate of 106 MHz, starting at the beam crossing. If the event is not accepted by the level-1 trigger (see Section 3.8.1), the data are overwritten by the next crossing. Otherwise, the data are compressed by eliminating the flat portions of the signal between the pulses ('zero suppression') and sent on to the level-2 trigger. See [46, pp. 41–45] and [62] for more information on the CD electronics.

3.6 Calorimetry

Conceptually, a calorimeter is a device which stops particles in an absorber and measures the amount of deposited kinetic energy. Since DØ has no central magnetic field, calorimetry is the only available method for measuring the energy of most types of particles. Thus, good calorimetry is a vitally important part of the detector. In-depth discussions of calorimetry in high-energy physics may be found in [50, 63, 64].

3.6.1 Calorimetry Principles

When a high-energy electron ($\gg$ 10 MeV) passes through a material with a high atomic number, the primary mechanism by which it loses energy is through Bremsstrahlung, in which a charged particle interacts with the Coulomb field around a nucleus and emits an energetic photon. A high-energy photon, on the other hand, will interact predominately via pair production, in which a photon converts into an electron-positron pair in the vicinity of a nucleus. The particles emitted in these interactions can themselves undergo Bremsstrahlung or pair production. Thus, an energetic electron or photon passing through a dense material will result in a shower of secondary electrons, positrons, and photons. This process is called an *electromagnetic shower*. The shower will continue to develop until all the secondaries have sufficiently low energies that other energy loss mechanisms (mostly ionization) become important. The rate at which an incident particle loses energy is a constant of the material, and is usually specified as a *radiation length* X_0:

$$\frac{dE}{E} = -\frac{dx}{X_0}. \tag{3.7}$$

As an example of a typical value, the radiation length for uranium is about 3.2 mm [3, page 1241].

Hadronic particles also cause showers, but they are qualitatively different from the electromagnetic showers caused by electrons and photons. Hadrons lose energy primarily through inelastic collisions with atomic nuclei. These collisions produce secondary hadrons, which can in turn undergo inelastic collisions. This process is called a *hadronic shower*, and it continues until all

particles have either been stopped by ionization losses or absorbed by nuclear processes. The scale for the size of these showers is given by the nuclear interaction length λ for the material. For uranium, $\lambda \approx 10.5\,\mathrm{cm}$ [3, page 1241]. Thus, hadronic showers are much more extended in space than electromagnetic showers of similar energy.

The showering process converts a single high-energy particle into many low-energy particles, all of which have about the same energy. The next step is to measure these low-energy particles. One approach is to make the absorber also function as an active material, such as a scintillator. Some typical materials for this are NaI, BGO ($B_4Ge_3O_{12}$), and lead glass. Such *homogeneous* calorimeters offer the best energy resolution; however, they are often not practical for large, high-energy detectors.

An alternate approach is to interleave layers of a dense, inert absorber with layers of a material which is sensitive to particles passing through it (such as a scintillator or some medium sensitive to ionization). This approach yields what is called a *sampling* calorimeter. Since most of the energy is absorbed in the inert material, only a portion of the incident energy can be detected. This fraction is called the *sampling fraction*. Since this is a statistical process, sampling degrades the achievable energy resolution.

Note that there is no *a priori* reason why the response of a calorimeter (i.e., the ratio of the measured signal to the energy of an incident particle) should be the same for electromagnetic and hadronic showers. In fact, the response will tend to be smaller for hadronic showers [65], since ν's and μ's produced by π and K decays will escape from the detector. The energy spent

breaking up nuclei will also be invisible. This is quantified by the e/π *ratio*, the ratio of the calorimeter responses to electrons and pions. It is highly desirable that this number be close to 1. The reason is this: a hadronic shower will include not only hadronic components but also electromagnetic components deriving from γ's from π^0 and η decays. The fraction of a hadron's energy which is deposited as electromagnetic showers can undergo large variations from shower to shower, but if the e/π ratio is 1, these fluctuations will not affect the energy resolution. A calorimeter with $e/\pi \approx 1$ is called a *compensating* calorimeter.

Even if a calorimeter is perfectly compensating, however, there are a large number of additional effects which will tend to degrade the resolution. Since the showering and sampling processes are statistical in nature, one would expect the fractional error in the calorimeter signal to scale as $1/\sqrt{N_{ion}}$, where N_{ion} is the number of ionization electrons liberated. Since this quantity should be proportional to the total energy of the incident particle, one would expect the resolution to scale as $1/\sqrt{E}$. Some other sources of energy fluctuations in the DØ calorimeter include [47, p. 51]:

- Leakage of energy out of the calorimeter.
- Variations in high voltage, absorber thickness and spacing, electronics gain, LAr temperature, or LAr purity.
- Noise due to the natural radioactivity of the uranium plates.
- Electronic noise.

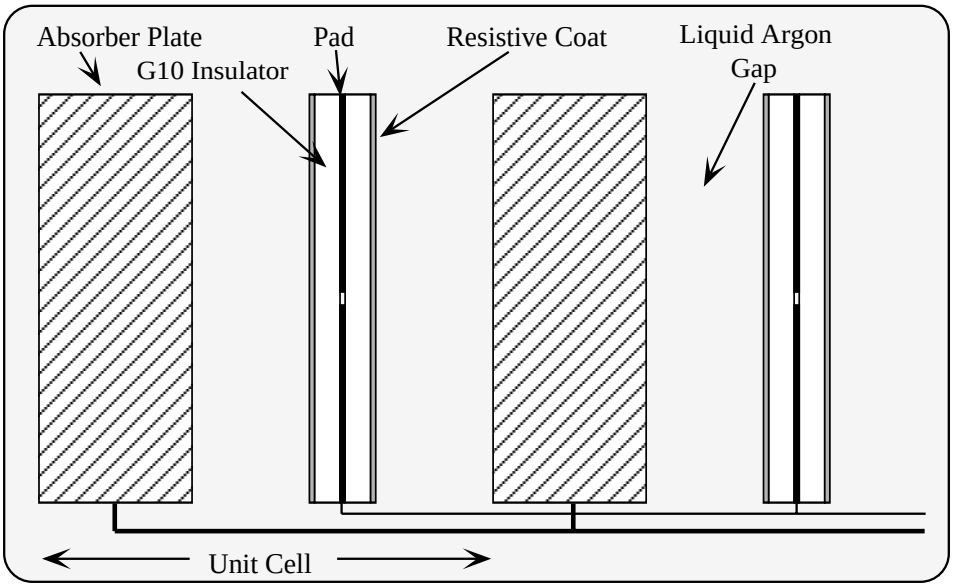

Figure 3.7: Schematic view of a calorimeter cell. [43, p. 44]

3.6.2 Calorimeter Geometry

The DØ calorimeter is a sampling calorimeter, with liquid argon (LAr) as the ionization medium. The primary absorber material is depleted uranium, with copper and stainless steel used in the outer regions. Since uranium is very dense, the calorimeter is relatively compact.

The calorimeter is divided into a large number of modules, each of which consists of a stack of interleaved absorber plates and signal boards. Figure 3.7 shows a cross-sectional schematic view of a section of this stack. The absorber plates are separated from the signal boards by a LAr-filled gap of 2.3 mm. The signal boards consist of a copper pad sandwiched between two 0.5 mm thick pieces of G10. The outer surfaces of these boards are coated with a resistive epoxy coating. During operation, the absorber plates are grounded, while a

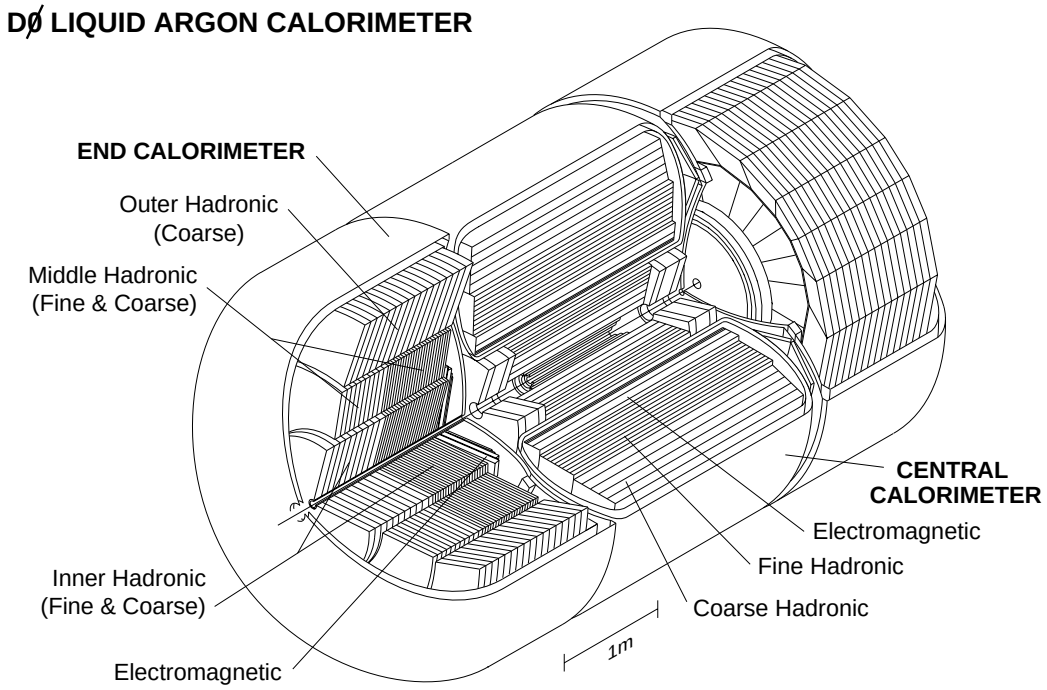

Figure 3.8: The DØ calorimeter. [36]

positive voltage of 2.0 – 2.5 kV is applied to the resistive coatings. As a shower develops in the calorimeter, charged particles crossing the LAr gap leave a trail of ionization. The liberated electrons are collected on the signal board after a drift time on the order of 450 ns, and induce a signal on the copper pad via capacitive coupling. In order to measure the transverse positions of showers, the readout pads are subdivided into smaller cells. The corresponding cells in adjacent signal boards are ganged together in depth to form *readout cells*; the details of this ganging vary from module to module.

An overview of the calorimeter is shown in Figure 3.8. Since it uses LAr, it must sit inside of a cryostat in order to keep the argon cold. In order to facilitate assembly and to allow access to the central detectors, the calorimeter

is divided into three major assemblies, each in its own cryostat: the central calorimeter (CC) and the two end calorimeters (EC).

The central calorimeter provides coverage out to a pseudorapidity of about 1.2. It is roughly toroidal, and consists of three concentric layers of modules. The inner layer consists of 32 electromagnetic (EM) modules, which are thick enough to contain most electromagnetic showers. The middle layer consists of 16 fine hadronic (FH) modules, which measure the showers due to hadronic particles. The final layer consists of 16 coarse hadronic (CH) modules, which measure any leakage of energy out of the FH layer, and which also serve to reduce any leakage out of the back of the calorimeter into the muon system ('punchthrough'). The parameters of the CC modules are given in Table 3.5.

The two sections of the end calorimeter provide coverage on each side of the CC from a pseudorapidity of about 1.3 out to about 4. The EC is composed of three concentric layers of modules. Like the CC, the modules are divided into electromagnetic and fine and coarse hadronic types; however, the geometry is rather different. The center of the EC consists of a disc-shaped electromagnetic module, backed by the cylindrical fine and coarse inner hadronic modules. Arranged in a ring around this central core are the fine and coarse middle hadronic modules, and around them is a final ring of coarse outer hadronic modules. The parameters of the EC modules are given in Table 3.6.

In both the CC and EC, the area in η, ϕ space covered by a typical readout cell is 0.1×0.1. However, in the third layer of the EM modules, where electromagnetic showers typically deposit the bulk of their energy, the readout

Module type	EM	FH	CH
Rapidity coverage	± 1.2	± 1.0	± 0.6
Number of modules	32	16	16
Absorber [a]	DU	DU-Nb	Cu
Absorber thickness (mm)	3	6	46.5
Argon gap (mm)	2.3	2.3	2.3
Number of signal boards	21	50	9
Number of readout layers	4	3	1
Cells per readout layer	2,2,7,10	20,16,14	9
Total radiation lengths (X_0) [b]	20.5	96.0	32.9
Total nuclear absorption lengths (λ) [b]	0.76	3.2	3.2
Sampling fraction (%)	11.79	6.79	1.45
Total readout cells	10,368	3000	1224

Table 3.5: Central Calorimeter Parameters. [47, p. 55] [43, p. 49] [46, p. 26] [66, p. 22]

[a] Depleted uranium (DU), depleted uranium with 1.7% niobium (DU-Nb), or copper (Cu).

[b] At $\eta = 0$.

Module type	EM	IFH	ICH	MFH	MCH	OH
Rapidity range	1.3–3.7	1.6–4.5	2.0–4.5	1.0–1.7	1.3–1.9	0.7–1.4
No. of modules	1	1	1	16	16	16
Absorber [a]	DU	DU-Nb	SS	DU-Nb	SS	SS
Thickness (mm)	4	6	6	6	46.5	46.5
LAr gap (mm)	2.3	2.1	2.1	2.2	2.2	2.2
Signal boards	18	64	12	60	14	24
R.O. layers	4	4	1	4	1	3
Cells/r.o. layer	2,2,6,8	16	14	15	12	8
Total rad. len.	20.5	121.8	32.8	115.5	37.9	65.1
Total abs. len.	0.95	4.9	3.6	4.0	4.1	7.0
Samp. frac. (%)	11.9	5.7	1.5	6.7	1.6	1.6
Tot. r.o. cells	7488	4288	928	1472	384 + 64 + 896 [b]	

Table 3.6: End Calorimeter Parameters. [47, p. 57] [43, p. 49] [46, p. 29] [66, p. 25]

[a] Depleted uranium (DU), depleted uranium with 1.7% niobium (DU-Nb), or stainless steel (SS).

[b] MCH and OH cells are summed together at $|\eta| = 1.4$

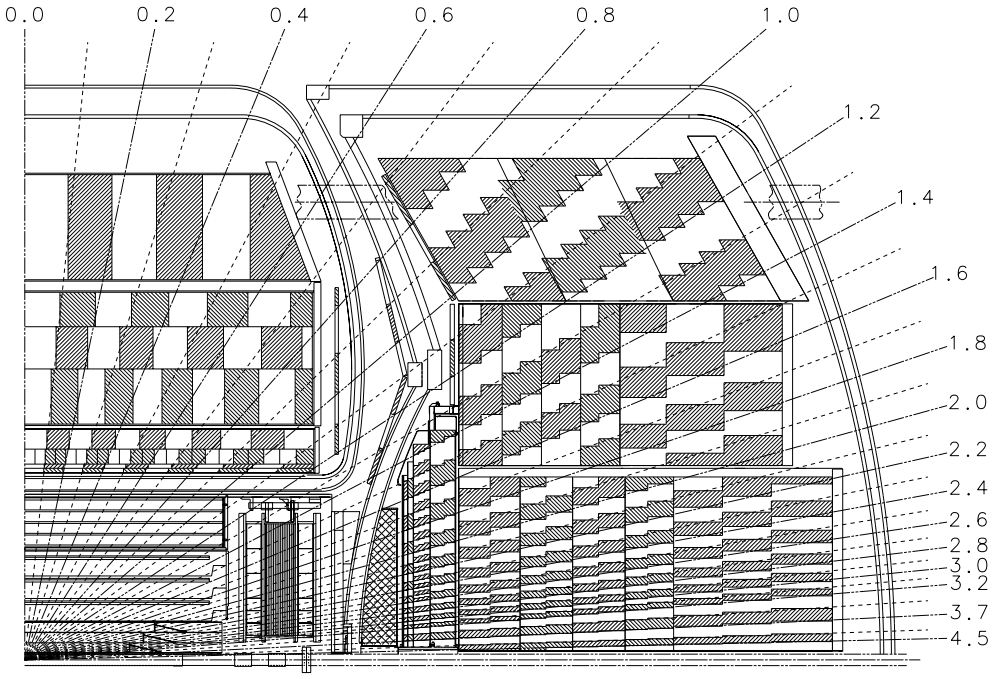

Figure 3.9: Side view of the calorimeters. [36]

cells have areas of 0.05×0.05. In addition, cells with $|\eta| > 3.2$ have a ϕ size of 0.2 and are somewhat larger in η as well. See Figure 3.9 and 3.10 for an illustration of the calorimeter segmentation.

If one examines Figure 3.9, it is apparent that in the transition region between the CC and the EC ($0.8 < |\eta| < 1.4$) there is a relatively large amount of uninstrumented material. This is primarily due to the cryostat walls and the support structures for the calorimeter modules. Two additional devices are used in this region to recover some of the energy deposited in this dead material. The first of these is the massless gaps (MG). These are simply rings of two signal boards mounted on the end plates of the CCFH, ECMH, and ECOH modules. The second device is the intercryostat detector (ICD). This

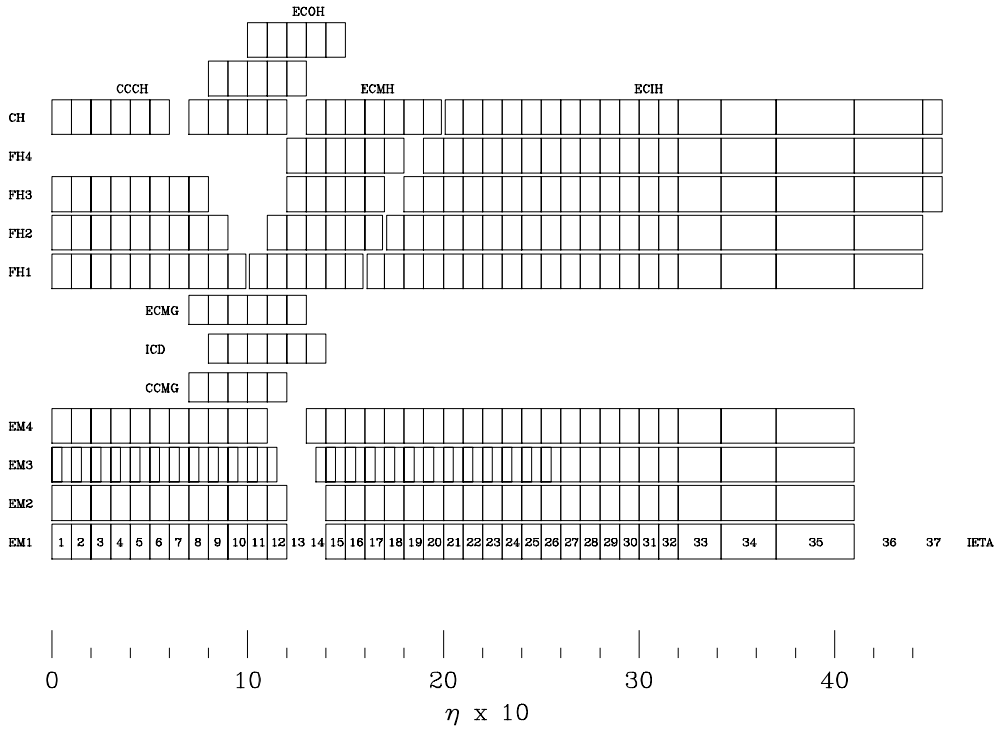

Figure 3.10: Layout of calorimeter channels in depth and η. [67]

is a ring of scintillation counters mounted on the exterior of the EC cryostats. Both the MG and the ICD have the standard segmentation of 0.1×0.1 in η, ϕ space.

3.6.3 Calorimeter Readout

The signals induced on the readout pads are pulses with widths on the order of 450 ns [43, p. 48]. These signals are led out though four ports in the cryostats to charge sensitive preamplifiers mounted on top of the cryostats. From the preamplifiers, the signals are led to base line subtractor (BLS) modules located in the platform below the detector. The BLS modules perform analog shaping and split the signal into two paths.

The first path is used for triggering. The signals from all the fine hadronic cells within a 0.2×0.2 tower are summed; the signals from the electromagnetic cells are similarly summed. These signals then form the input to the level-1 calorimeter trigger (see Section 3.8.1).

The second path is used for the data readout. The incoming signal is sampled just before the beam crossing and again 2.2 μs later. The difference between these two samples is a DC voltage which is proportional to the total collected charge. This difference is sent to the ADCs where, if the event is accepted by the level-1 trigger, the signals are digitized and sent on to the level-2 trigger.

3.6.4 Calorimeter Performance

The response of the calorimeter modules to single electrons and pions has been studied extensively in test beams [36, sect. 3.7], [68, 69, 70, 71]. The response for both is found to be linear to within 0.5%.

The resolution is parameterized as

$$\left(\frac{\sigma}{E}\right)^2 = C^2 + \frac{S^2}{E} + \frac{N^2}{E^2}, \tag{3.8}$$

where the constants C, S, and N represent calibration errors, sampling fluctuations, and noise contributions, respectively. For electrons, the measured resolutions are

$$C = 0.003 \pm 0.002, \quad S = 0.157 \pm 0.005(\mathrm{GeV})^{\frac{1}{2}}, \quad N \approx 0.140\,\mathrm{GeV} \tag{3.9}$$

and for pions,

$$C = 0.032 \pm 0.004, \quad S = 0.41 \pm 0.04(\mathrm{GeV})^{\frac{1}{2}}, \quad N \approx 1.28\,\mathrm{GeV}. \tag{3.10}$$

The position resolution for electrons is found to be about 0.8–1.2 mm, varying approximately as $1/\sqrt{E}$.

The e/π ratio of the calorimeter falls from about 1.11 at 10 GeV to about 1.04 at 150 GeV.

3.7 Muon System

Although muons decay into electrons, their lifetime of 2.2 μs is so large compared with the scale of the detector that for all practical purposes a relativistic muon may be regarded as stable. Muons do not interact strongly, and

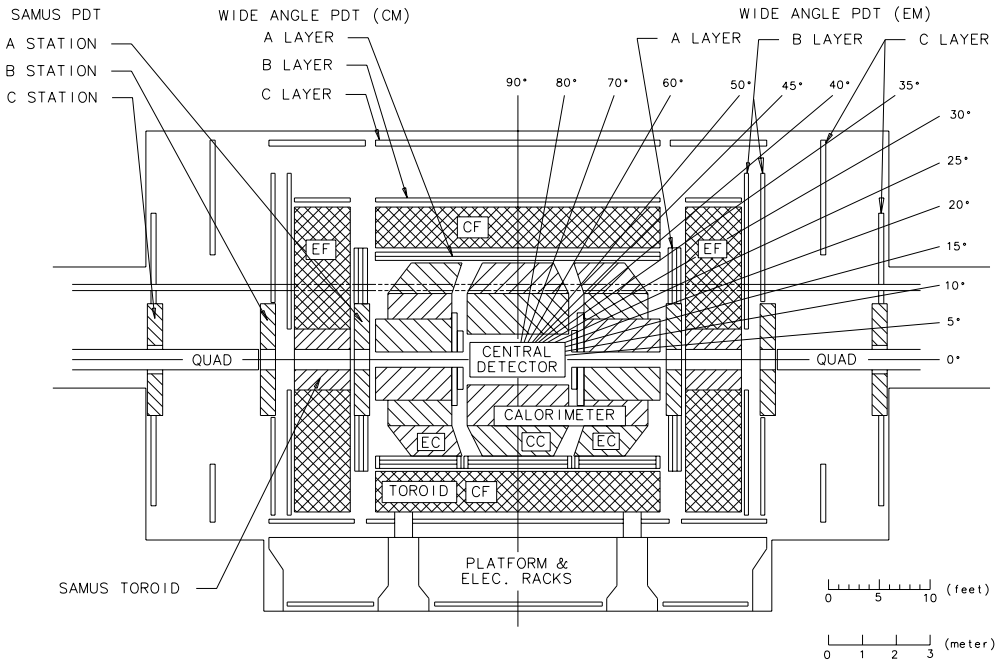

Figure 3.11: Side elevation of the muon system. [36]

moreover, their mass is too large ($\approx 200m_e$) to readily initiate an electromagnetic shower at Tevatron energies. Thus, any charged particle which makes it out of the calorimeter is likely to be a muon.

The DØ muon system [72] consists of five magnetized iron toroids which are surrounded by three layers of proportional drift tubes (PDTs). See Figure 3.11. The PDTs measure the trajectory of muons before and after they traverse the magnetized iron; thus a measurement of the muon momentum can be made. The five magnets are the CF (Central Fe), covering the angular range $|\eta| < 1$, the two EFs (End Fe), covering the angular range from $|\eta| = 1$ out to about $|\eta| = 2.5$, and the two SAMUS (Small Angle MUon System) magnets, covering the range from about $|\eta| = 2.5$ out to about $|\eta| = 3.6$. The CF and the two EFs together are referred to as the Wide Angle Muon

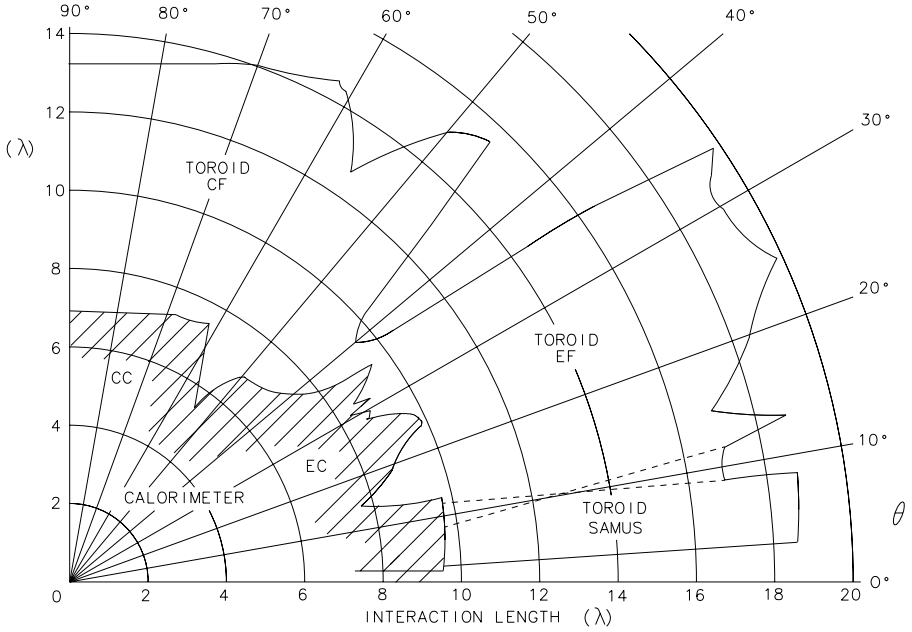

Figure 3.12: Number of nuclear interaction lengths as a function of polar angle. [36]

System, or WAMUS. Each section has one layer of drift tubes (the A layer) just inside the magnet, a second layer (B layer) just outside the magnet, and a third layer (C layer) 1–3 m further out. Due to the necessity of supporting the calorimeter, some regions underneath the detector are not covered by the full compliment of three drift tube layers.

The more material which must be traversed prior to entering the muon system, the smaller the background due to hadronic punchthrough will be. Figure 3.12 shows how the thickness of the detector in nuclear interaction lengths varies with polar angle. At $\eta = 0$ the minimum muon momentum required to make it through both the calorimeter and iron is about 3.5 GeV/c [47, p. 59]. At higher η, this rises to about 5 GeV/c.

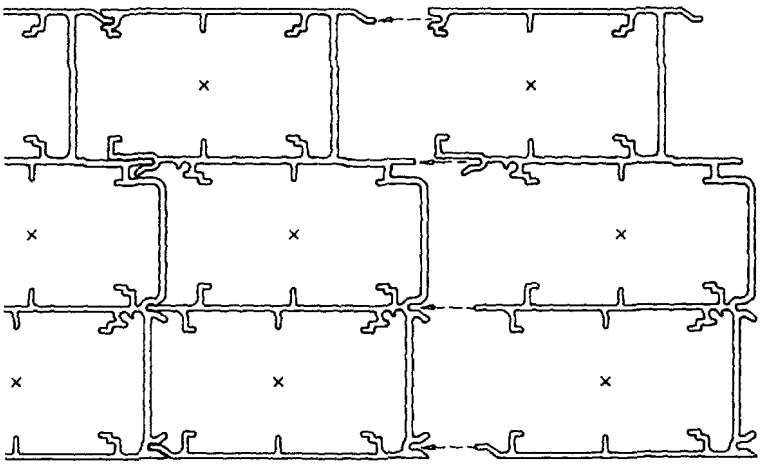

Figure 3.13: End view of muon drift tubes. [36]

3.7.1 WAMUS System

The PDTs making up the WAMUS system are rectangular in cross section, with one sense wire per drift cell. They are formed out of aluminum extrusions into rectangular modules containing either four (A layer) or three (B and C layers) planes of drift tubes (see Figure 3.13). There are 164 modules, each containing between 14 and 24 drift tubes, the lengths of which are between 191 and 579 cm. The tubes are oriented roughly parallel to the direction of the magnetic field in the iron toroid. In this manner, the deflection due to the magnet is measured by the drift time. The distance of a hit from the sense wire can be determined with a resolution of about 0.5 mm.

In order to facilitate access to the chamber electronics, all readout is done from one end of the drift tubes. To accomplish this, tubes are jumpered together in pairs on one end. The front end electronics measures the arrival times of pulses at the end of each wire, as well as the time difference between

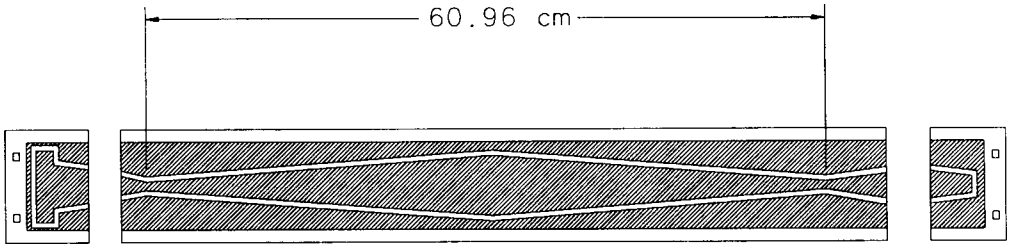

Figure 3.14: Muon cathode pad. [36]

pulses arriving at the ends of each jumpered pair of sense wires. Using this time difference, a crude measure of the position of a hit along the wire can be found (within 10 – 20 cm). A more precise determination of the hit position is made by inserting at the top and bottom of each tube a set of vernier cathode pads (see Figure 3.14). These pads are insulating Glasteel coated with copper cladding and forming a diamond pattern which repeats with a period of $\approx$ 60 cm. The cladding is separated into an inner region and an outer region. Electron avalanches on the sense wire will induce pulses on the cathode pads. The ratio of the charge deposited on the inner and outer pads can be used to localize the hit to within about 3 mm modulo the half-repeat period of about 30 cm. This ambiguity is resolved using the coarse time difference measurement. The cathode pads in adjacent layers of PDTs are staggered by about 1/6 of a repeat period relative to each other to further resolve ambiguities which occur near the corners of the diamond pattern. Each tube also has a single bit output which is set if there were any hits on the cathode pad in that tube (the *pad latch*). These bits form the input to the muon trigger (see Section 3.8.1).

To measure the momentum of a muon, its trajectory before and after

passing through the magnet must be known. Tracks through the B and C layers give the trajectory after the magnet, while tracks though the A layer give the trajectory before the magnet. If possible, the A layer tracks are matched to tracks in the central detector and to minimum ionizing traces in the calorimeter in order to improve the direction measurement. The primary interaction vertex found by the central detector may also be used to define the incoming track. The momentum (P) resolution is most easily formulated in terms of the inverse momentum $k = 1/P$ as

$$\left(\frac{\delta k}{k}\right)^2 = (0.18)^2 + \left(\frac{0.01}{k\,\mathrm{GeV}}\right)^2 . \tag{3.11}$$

Further parameters of the muon system are given in Table 3.7. For further details, see [36, 72, 73].

3.7.2 SAMUS System

Due to the high occupancies in the forward region, the SAMUS system uses smaller drift tubes. It is composed of three stations each consisting of three drift tube planes. Each plane is composed of two subplanes, offset by half a tube diameter. The three layers are rotated with respect to each other. The drift tubes themselves are constructed from stainless steel tubes with a 3 cm diameter, each containing a single sense wire with a 50 μm diameter. Some additional parameters of the SAMUS system are given in Table 3.7. For further details, see [36, 72, 74].

	WAMUS	SAMUS
Rapidity coverage	$\|\eta\| < 1.7$	$1.7 \leq \|\eta\| \leq 3.6$
Magnetic field	2 T	2 T
Nuclear interaction lens.	≈ 13.4	≈ 18.7
Number of modules	164	6
Number of drift cells	11,386	5308
Sense wire specifications	50 μm Au-plated W, 300 g tension	50 μm Au-plated W, 208 g tension
Maximum sagitta	0.6 mm	2.4 mm
Sense wire voltage	+4.56 kV	+4.0 kV
Cathode pad voltage	+2.3 kV	—
Gas composition	Ar 93%, CF_4 5%, CO_2 5%	CF_4 90%, CH_4 10%
Bend view resolution	$\pm$0.53 mm	$\pm$0.35 mm
Non-bend view resolution	$\pm$0.3 mm	$\pm$0.35 mm
Average drift velocity	6.5 cm/μs	9.7 cm/μs
Maximum drift distance	5 cm	1.45 cm

Table 3.7: Muon System Parameters. [36, 72][47, p. 61]

3.8 Triggering and Readout

At the Tevatron, beam crossings occur at the interaction region at a rate of about 290 kHz. At a luminosity of $5 \times 10^{30}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, an inelastic collision will occur in about 3/4 of these crossings [36, sect. 5.1]. However, the processes which are of the greatest interest are much rarer. Because it is not feasible to record and process data from every crossing, there must be some mechanism to select out the small fraction of interesting events for permanent storage. This process is called *triggering*.

The overall layout of the DØ trigger system is shown in Figure 3.15. It can be conceptually divided into two hierarchical pieces: level-1 and level-2. The level-1 trigger is a collection of dedicated hardware processors which operate on a coarse subset of the event data. Most level-1 trigger decisions can be made within the 3.5 μs interval between beam crossings, permitting operation without deadtime. However, some triggers, called level-1.5 triggers, may require more time. The goal of the level-1 trigger is to reduce the event rate from the beam crossing frequency of 290 kHz to a rate of 200 – 300 Hz.

Once an event has been accepted by level-1, the complete event is digitized and the data transferred to one of 48 level-2 nodes. These are general-purpose computers which process events in parallel. They perform a fast reconstruction of the event, and can use general software filters to decide whether or not an event should be kept. If level-2 passes an event, it is transferred to the host system, where it is permanently recorded on magnetic tape.

The components of the trigger system are further described below.

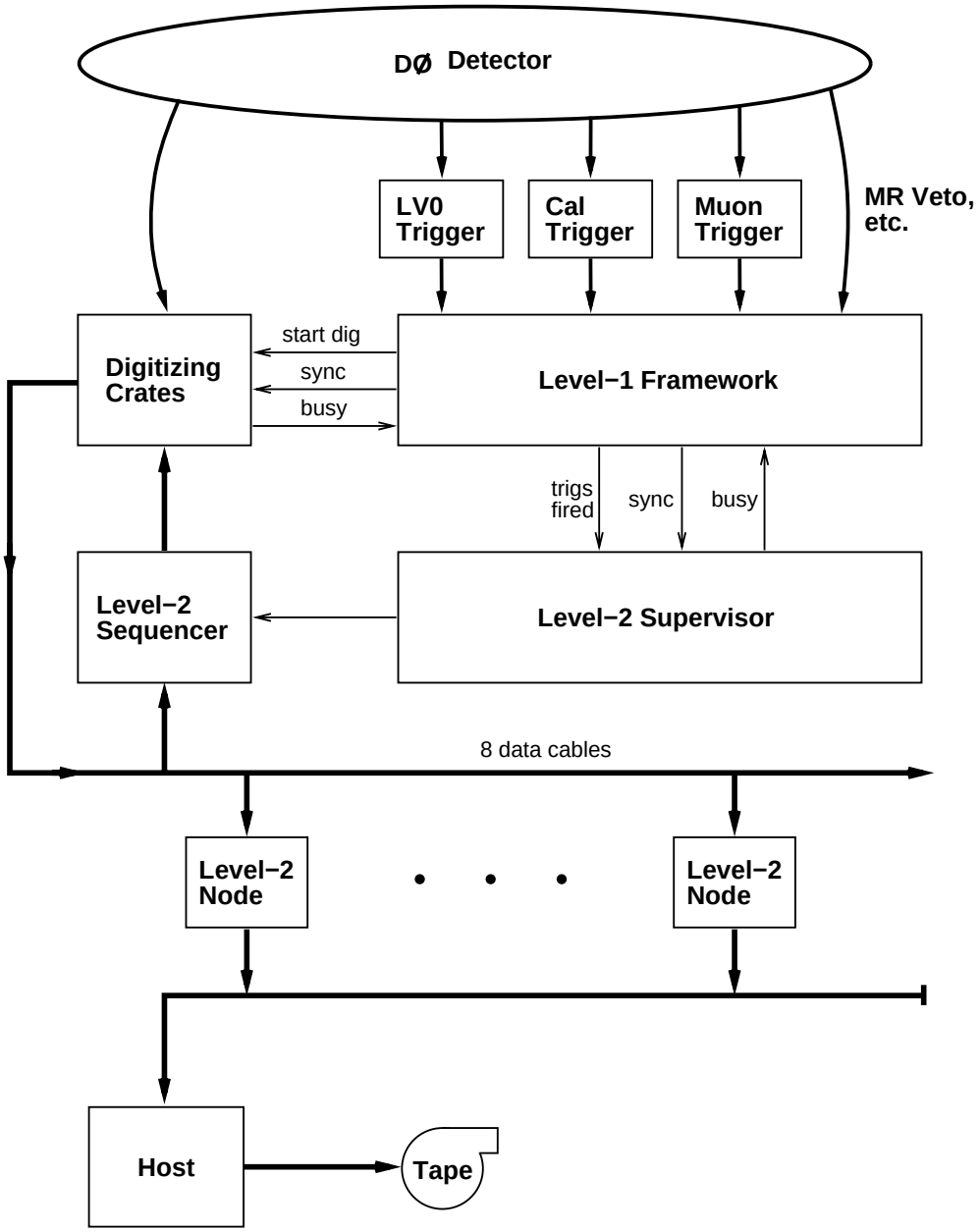

Figure 3.15: Overall layout of the DØ trigger system.

3.8.1 Level 1

At the heart of the trigger system is the level-1 trigger framework [75, 76, 77]. This is a special-purpose hardware processor which is responsible for combining the results of individual level-1 components into a set of global decisions, commanding the readout of the digitization crates, and interfacing with level-2. The primary input to the framework consists of 256 *trigger terms*. Each of these is a single bit, indicating that some specific requirement is met for the present event (such as 'at least two muons', or '$\not\!\!E_T > 20\,\mathrm{GeV}$'). Most of these inputs come from detector-specific level-1 processors, but some come directly from sources such as scintillators or from accelerator timing signals.

The 256 trigger terms are reduced to a set of 32 level-1 *trigger bits*, also known as *specific triggers*, by an and/or network. Each trigger bit can be configured to require that certain trigger terms be set and that certain other terms not be set. Each trigger bit also has a programmable prescale; setting the prescale to some value N means that the trigger will only actually fire once in every N times that its trigger term conditions are satisfied.

The digitizing hardware is located in 86 front-end VME crates in the moving counting house. For purposes of readout, these crates are grouped into 32 *geographic sectors*. For each specific trigger bit, the framework has a list of geographic sections to be read out. When a trigger fires, the framework commands the appropriate set of sectors to begin digitizing. The crates are double buffered, so they can start digitizing an event while a previous one is still being transferred to level-2. Each geographic sector also sends a busy

signal back to the framework when any crate in it is unable to accept another trigger. This will inhibit any further triggers involving that geographic sector.

At the same time that the framework commands the front ends to begin digitizing, it sends a signal to the level-2 system containing a mask of the specific triggers which fired. The level-2 supervisor then manages the transfer of the event data from the digitizing crates to a level-2 node, as described below. The level-2 supervisor also sends a busy signal to the framework when there are no free nodes to accept an event.

If a level-1.5 decision is required, the framework starts the digitization cycle as usual, but delays notifying level-2. When level-1.5 processing is complete, the framework either sends the event on to level-2 or aborts the digitization cycle, depending on the result.

The trigger framework also maintains numerous counters to monitor system dead time, the luminosity seen by each trigger, and other performance measures.

The trigger framework is controlled by a dedicated Vaxstation 4000/60, called the trigger control computer (TCC). The TCC does not directly participate in trigger processing, but instead is responsible for programming the level-1 hardware at the beginnings and ends of runs, and for monitoring the performance of the system. The TCC is the interface through which the rest of the data acquisition system talks to the trigger framework.

The sections below describe some of the major inputs to the trigger framework.

Level 0

The level-0 system [78, 79] is designed to detect events containing an inelastic scattering and to provide a fast measurement of the location of the event vertex. It consists of two separate detectors located at each end of the central detector between the FDC and the EC. Each detector consists of two layers of rectangular scintillator bars, read out with photomultipliers. Due to the rectangular geometry of the scintillators, the coverage is not uniform in ϕ. There is nearly complete coverage in the range $2.2 < |\eta| < 3.9$ and partial coverage extending out to $1.9 < |\eta| < 4.3$. An inelastic $p\bar{p}$ collision will typically include a large amount of activity in the far forward regions (from the spectator quarks); thus, one looks for a coincidence between the signals from the two scintillator arrays. The level-0 trigger is $> 99\%$ efficient for nondiffractive inelastic collisions. By comparing the arrival times of the signals from the two arrays, the approximate position of the interaction vertex may be found. A fast vertex determination with a resolution of $\pm 15\,\mathrm{cm}$ is available within 800 ns after the collision. A more accurate determination with a resolution of $\pm 3.5\,\mathrm{cm}$ is available within $2.1\,\mu\mathrm{s}$ [43, p. 59]. The vertex position is available as several level-1 trigger terms; it can also be used in level-2 processing. The level-0 system can also identify events which are likely to contain multiple interactions.

Main Ring Vetoes

During normal operation of the Tevatron, the Main Ring is used to produce antiprotons, with a cycle period of 2.4 s [43, p. 129]. Since the Main Ring passes through the DØ detector, losses from the Main Ring will show up in the detector and must be rejected. The largest losses occur when beam is injected to the Main Ring, and again 0.3 s later when the beam passes through transition[2]. These losses are dealt with by vetoing on the `MRBS_LOSS` trigger term. This term is asserted as a possible veto during a 0.4 s window starting at injection, continuing through transition, and allowing time for the calorimeter and muon high voltage to recover from the large losses. This results in a dead time of about $0.4/2.4 \approx 17\%$ [80].

Even after injection is complete, however, it is still possible to have observable losses whenever a Main Ring bunch passes through the detector. To

[2]Consider a bunch of non-relativistic particles traveling in a circular orbit. The particles with a larger than average momentum will also have a larger than average velocity and will pull ahead of the rest of the bunch. In order to keep the bunch from blowing up longitudinally, one must therefore arrange for particles near the front of the bunch to be decelerated relative to the rest of the bunch, and for those near the tail to be accelerated (again, relative to the rest of the bunch). For highly relativistic particles, however, the situation is different. In this regime, the velocity of a particle is nearly constant (at c) regardless of its momentum. However, the path length is not constant: a particle with larger than average momentum will have a larger than average bending radius and will thus fall behind the rest of the bunch. So in this situation, one must accelerate the head of the bunch more than the tail. The point in the acceleration cycle at which the switch between these two descriptions occurs is called *transition*; the energy at which it occurs depends both on the mass of the particles being accelerated and the size of the accelerator ring. Properly rearranging the accelerating fields when passing through transition is tricky, and accelerators often experience extra losses at that point.

guard against this, an additional trigger term is available called the *microblank* term. This term is asserted for a particular beam crossing if a Main Ring bunch is present in the detector within ± 800 ns of the crossing. By vetoing on this term, one can reject events which may be contaminated by Main Ring losses. This adds an additional $\approx 8\%$ deadtime [80].

Calorimeter Trigger

The level-1 calorimeter trigger [76, 77], located along with the trigger framework on the first floor of the moving counting house, is responsible for making trigger decisions based on calorimeter information. It is a pure level-1 trigger, and thus must be able to return a decision within $2.4\,\mu$s [81]. The inputs derive from the trigger pickoffs in the calorimeter BLS cards, which sum cells into towers of size 0.2×0.2 in $\eta - \phi$ out to a pseudorapidity of 4 (see Section 3.6.3). Separate inputs are provided for cells originating from EM modules and FH modules, 1280 of each (the CH calorimeter modules are not used for the level-1 trigger). At the input of the calorimeter trigger, all the inputs are simultaneously flash digitized; all subsequent calculations are entirely digital. The trigger calculates a number of global sums of calorimeter inputs. These sums are:

- The total electromagnetic energy, $E(\mathrm{em}) = \sum_i E_i(\mathrm{em})$.
- The total hadronic energy, $E(\mathrm{had}) = \sum_i E_i(\mathrm{had})$.
- The total scalar sum of electromagnetic transverse energy, $E_T(\mathrm{em}) = \sum_i E_i(\mathrm{em}) \sin\theta_i$.

- The total scalar sum of hadronic transverse energy,
  $E_T(\text{had}) = \sum_i E_i(\text{had}) \sin\theta_i$.

- The total transverse energy, $E_T(\text{tot}) = E_T(\text{em}) + E_T(\text{had})$.

- The missing transverse energy, $\not\!\!E_T = \sqrt{E_x^2 + E_y^2}$, where

$$E_x = \sum_i \left(E_i(\text{em}) + E_i(\text{had})\right) \sin\theta_i \cos\phi_i \tag{3.12}$$

  and

$$E_y = \sum_i \left(E_i(\text{em}) + E_i(\text{had})\right) \sin\theta_i \sin\phi_i. \tag{3.13}$$

These quantities are then compared with a set of programmable thresholds. Each of these comparisons yields a trigger term which is input to the trigger framework. This allows one to specify trigger requirements such as 'total scalar E_T above 140 GeV', or 'missing E_T above 20 GeV'.

Each individual trigger tower also has a set of associated thresholds. The EM energy in each tower is compared with four programmable thresholds; the hadronic energy is also compared with a separate set of four thresholds. For each tower and pair of thresholds, a bit is set if the EM energy is greater than, and the hadronic energy less than, their respective thresholds. In addition, the sum of the EM and hadronic energy is compared with another set of four thresholds and a separate bit set for each tower for which this threshold is exceeded. A count is then made of the number of towers for which each bit is set and the result compared to another set of thresholds. For each threshold that is exceeded, a trigger term is asserted. This allows one to specify trigger

requirements such as 'one EM tower above 5 GeV in the central region' or 'two towers with total energy above 10 GeV anywhere in the calorimeter'.

Muon Trigger

The inputs to the level-1 muon trigger [82] are the pad latch outputs from the muon system (see Section 3.7.1). There is one bit from each tube, for a total of about 16,700 bits. The level-1 trigger divides the muon detector into five regions: CF, EF-north, EF-south, SAMUS-north, and SAMUS-south. Within each region, the trigger looks for patterns of hits which are consistent with a muon emitted from the nominal interaction vertex. In most regions, hits are required in all three layers of muon drift tubes; however, in some CF areas without three layer coverage, tracks are allowed to pass with hits in only two layers. The level-1 trigger counts the number of track candidates in each region and compares the result with a set of preset thresholds. If any thresholds are exceeded, the appropriate trigger term is asserted to the level-1 framework.

The pure level-1 muon trigger is not capable of distinguishing between muons of different momenta because it matches tracks between layers very coarsely. There is also a level-1.5 muon trigger which is capable of imposing a p_T cut on a muon by requiring a finer match between layers. However, due to the combinatoric problems of doing this matching, especially in the busy SAMUS regions, this computation often takes longer than the 2.4 μs allotted for level-1 trigger term decisions. (Decision times typically range from 1 to 5 μs in the WAMUS regions, but can take up to 100 μs in the busy SAMUS

region.) Thus, enabling the level-1.5 trigger may cause extra dead time.

3.8.2 Level 2

The level-2 trigger [83, 84, 85] is essentially a large farm of general-purpose processors which run software filters using the complete data for an event. Before this can happen, however, the data must be collected together in a single place.

The digitizing hardware is located in a large number of VME crates. Each crate also contains a card called a VME buffer/driver (VBD). These cards are connected together along one of eight high-speed (40 MB/s) data cables, each of which is a loop originating and terminating at a special sequencer card. When the crate has finished digitizing, the VBD copies the data into an internal buffer. It then waits to receive a readout token from the sequencer before sending the data out over the data cable.

Each level-2 node is a Vaxstation 4000/60 (some of which have been upgraded to 4000/90's for run 1B), running the VaxELN realtime operating system. The bus of each node is extended out to a VME crate. Each crate contains four dual multiport memory (MPM) boards which are connected to the eight data cables. Each level-2 node also contains a VBD which is used to transmit events up to the host system.

The process of transferring an event from the digitizing crates to a level-2 node is controlled by the level-2 supervisor. When the supervisor receives a trigger from level-1, it picks an idle level-2 node and enables its MPMs to

receive data from the cables. The supervisor then causes the sequencers to start circulating tokens on the data cables. These tokens are used to arbitrate access to the data cable: a VBD may only transmit data while it is in possession of the readout token. When a VBD receives the token and has data to send, it grabs the token and dumps its data out on the cable. Any MPMs on that cable which have been enabled will receive and store this data. The VBD follows its data with another token, which has been modified to indicate that that crate has been read out. Once the complete event has been received by the level-2 node, it is converted to Zebra format [86] by adding appropriate headers to the data received from each cable. The node can then start doing filter processing.

Since the data for an event is initially distributed among many front-end crates, and since each crate may have several events buffered, it is conceivable that a readout error may cause pieces of different events to get mixed together during the readout process. To prevent this, the framework stamps each event with a unique number. This number is sent along with the trigger mask to the level-2 system. The lower four bits, called *sync bits*, are also distributed to each crate along with the start-digitize commands. Level-2 puts the sync bits into the tokens it circulates, and a VBD will only read out if the sync bits in the token match the sync bits associated with the data that that crate has buffered.

The level-2 node software [87] is modular and is composed of a framework, which is responsible for getting events into and out of the node and for communicating with the rest of the system, and a set of *tools*, which do the actual

filtering. Each tool can take a set of parameters describing the cuts to be made. For example, the parameters used by the jet tool include, among others, the number of jets to require, the minimum E_T of the jets, the cone size of the jets, etc. Each tool then returns a boolean value telling whether or not the current event passed the cuts. Lists of tools with specific values for the parameters are collected together into *filter scripts*. Each script is associated with one level-1 trigger bit. During normal operation, the framework looks at the mask of level-1 bits for each event. It calls in sequence every script associated with the level-1 bits which were set. For every script, every tool in the script is called; if all the tools in the script pass, then the script as a whole passes. For each script which passes, a bit is set in a 128-bit mask of *filter bits*. If any filter bits are set, the event gets sent to the host, as described below.

3.9 Host Processes

The DØ host cluster is a cluster of Vax computers running the VMS operating system. For most of run 1A, it consisted of three machines (an 8700, a 6410, and a 6620) networked with DEC's 'Cluster Interconnect', three HSC disk controllers, a collection of RA-series disks, and a group of satellite Vaxstations. Dedicated X-windows terminals were also used to increase the number of available displays.

The online software is generally structured as a set of detached server processes which respond to requests from users, to requests from other servers, or to events taking place in the experiment. Most communication between

processes is done using a locally-developed package called ITC (Inter-Task Communication) [88], which provides a convenient layer over the VMS DECnet and mailbox communication services.

The online processes can be broadly classed into four groups: run control, data logging, downloading, and monitoring.

3.9.1 Run Control

At the heart of the system is the run coordination program, COOR [89]. This program is responsible for three principal tasks:

- Allocating hardware and software resources to individual users of the system.
- Arranging for hardware and trigger systems to be programmed according to the requests of users of the system.
- Stepping the other parts of the system, such as the trigger framework, level-2 trigger, and the data logger, through the steps needed to begin or end a run.

Users typically interact with COOR through a program called Taker [90]. Through Taker, users choose a *trigger configuration* [91] which they want to use. A trigger configuration is a set of text files which describe to COOR how to configure the online system for a particular purpose. A configuration file will contain descriptions for one or more level-1 trigger bits (see Section 3.8.1). For each bit, it specifies:

- The level-1 programming: what trigger terms are required and what thresholds to set for the individual trigger processors.
- The list of geographical sections to be read out, and any special mode settings for the digitization hardware.
- A list of level-2 filter scripts to be run for events which pass the level-1 trigger.
- For each level-2 script, a list of recording streams to which events passing the filter should be written.

When it receives a trigger configuration request, COOR examines its internal model of the online system to determine if the new request would conflict with any previous requests. If not, COOR issues the appropriate commands to configure the system for the new request and adjusts its internal model accordingly. Since a configuration can be set up to read out only a portion of the detector, it is possible to have multiple users using the system simultaneously, each reading out a different subdetector. This is useful for calibration and testing tasks.

The configuration files which COOR reads are rather cumbersome to work with for complex trigger configurations. To ease the construction and maintenance of these files, a utility called 'trigparse' has been written which takes as input a compact, physics-oriented specification of the triggers and produces as output the full set of configuration files required by COOR [92]. In order to allow reconstruction of past configurations and to track changes, all changes

to configuration files are logged using DEC CMS.

3.9.2 Data Logging

Once a level-2 node has passed an event, it transmits it to the host cluster, where it is received by the data logger [93], a program running on one of the host computers. This program and others associated with it are responsible for receiving raw data from the level-2 system and copying it to magnetic tape, while performing all necessary bookkeeping tasks. Data may also be sent to the online DAQ pool for online monitoring and to the express line. The express line is a collection of dedicated Vaxstations used to immediately reconstruct a subset of the data for rapid analysis. The overall flow of data in the host system is illustrated in Figure 3.16.

The process of reading out an event through the logger starts when a level-2 node has passed an event and wants to send it to the host. The node copies the event to the VBD (VME Buffer/Driver) card in its VME crate and tells the VBD to read out.

The VBDs are connected along a readout cable, similar to the cables used to convey data into the level-2 nodes. The sequencer for this cable is controlled by D0SNTY, a Vaxstation 4000/60 running VaxELN. While D0SNTY waits for an event to arrive, it circulates tokens around the readout cable. If the token reaches a VBD which has data to send, that VBD grabs the token and sends its event data along the readout cable. D0SNTY then ceases to circulate tokens until it is ready for another event. (Note that this implies

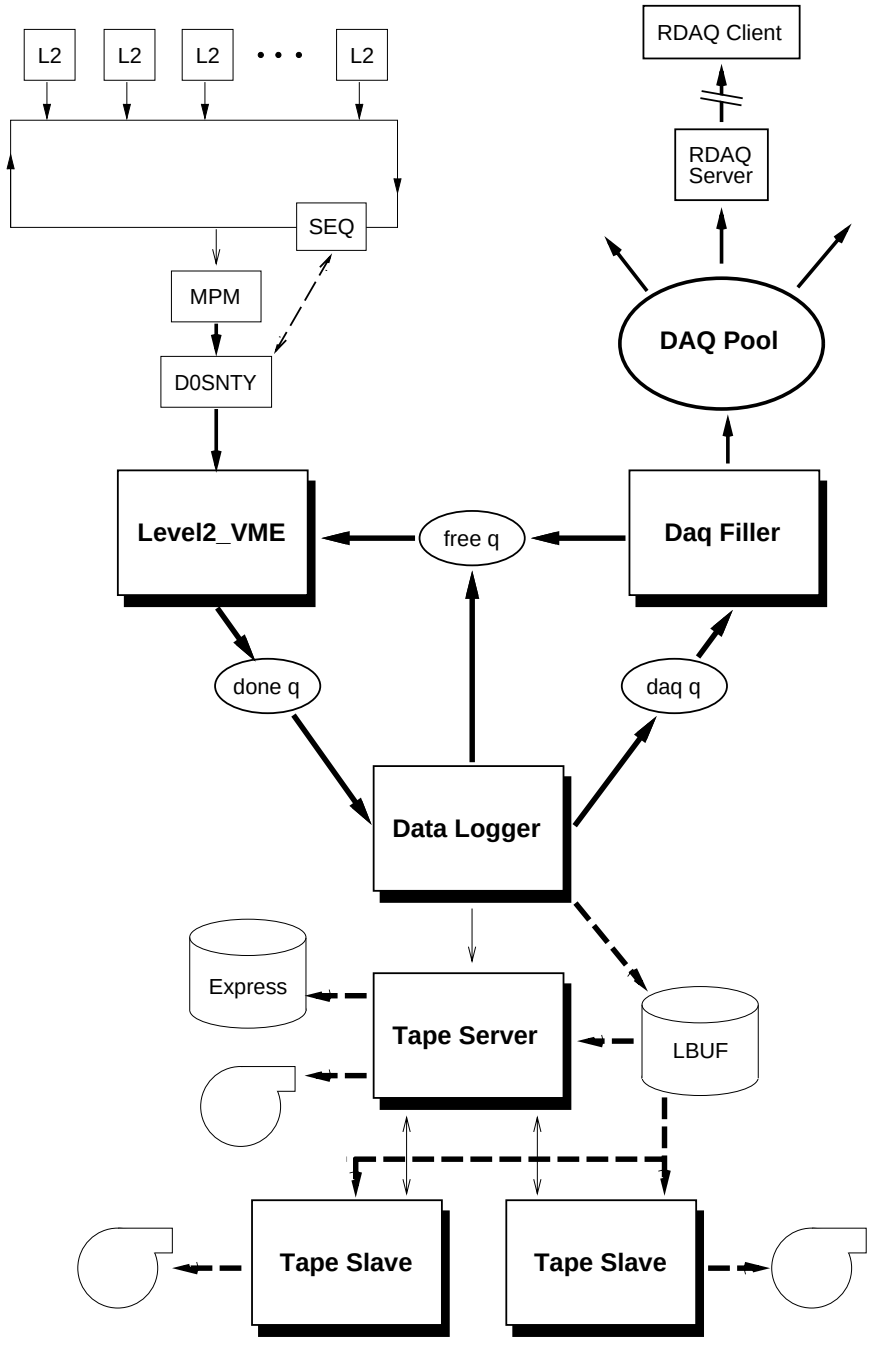

Figure 3.16: The DØ data logger. [93]

that the level-2 nodes at the start of the cable are read out preferentially to those further down the cable.)

As the event is read out over the cable, it is stored in a multiport memory (MPM) board in the D0SNTY VME crate. When the readout is complete, D0SNTY copies the event from the MPM to a prearranged location in VME memory which is mapped to memory in the host VAX through a VME to XMI (Vax bus) interface (DWMVA). Once this copy is complete, D0SNTY notifies the host by sending an interrupt. When the host is ready for another event, it notifies D0SNTY and tells it where in VME memory the next event should be written.

On the host computer, there are three processes which comprise the core of the logger. These are called `Level2_VME`, `Data Logger`, and `DAQ Filler`. Between them, they maintain a small pool of event buffers. These buffers are passed among these processes through three queues, the 'free queue', 'done queue', and 'daq queue'.

The `Level2_VME` program is responsible for interfacing with the level-2 readout system. It obtains a free buffer from the free queue, and arranges for that buffer to be mapped into VME address space. It tells D0SNTY to read out into this buffer, and waits for the interrupt signalling the completion of the transfer. It then tells `Data Logger` that a new event is available by placing the buffer in the done queue, and goes to find another free buffer. Once an event is in an event buffer, it does not need to be copied again within the logger.

Note that `Level2_VME` can be replaced with another program for diagnostic or testing purposes. For example, there is a version which reads raw data

from a file instead of receiving it from level-2. There is also a version which implements the old DECnet protocol for receiving data from level-2 (which can still be used if there are problems with the primary data path).

When `Data Logger` gets an event from `Level2_VME`, it verifies that the event is intact by computing its checksum and by validating the Zebra structure. If the event appears damaged, it is thrown away. (Damaged events can optionally be dumped to disk for later analysis.) If the event is good, the logger proceeds to process the event by storing the symbolic names of all the triggers and filters which fired into the event's TSUM bank. The logger then examines the mask of filters which the event passed, and on the basis of this decides to which output streams this event should be written. For each stream, there is a file present on one of several buffer disks to which events in that stream are written. Before writing an event, the logger stamps it with the proper run number and output event number[3], and performs other bookkeeping tasks. The disk files which are written have a fixed maximum length, which is usually set to 250 MB. When a file reaches this size, it is closed and a new partition is opened. A command is then sent to the tape server to copy the old file to tape, and information about the file is recorded in a database. There is a fixed limit for any stream of 99 parts per run. When this limit is reached, the logger sends a command to the run coordination program COOR

[3]For run 1A, events were numbered consecutively *in each stream.* This meant that if an event was written to multiple streams, it would in general have a different number in each. For run 1B, the numbering scheme was changed so that events are numbered consecutively within a run; an event then has the same number across all streams.

instructing it to stop the run.

After being written to disk, events are sent to `DAQ Filler` by placing the buffer in the daq queue. (Actually, this is only done if `DAQ Filler` is idle, in order to prevent hanging the entire system if a problem should develop with the DAQ pool.) `DAQ Filler` copies the events into the DAQ pool (a.k.a. the *Global Shared Common*) [94]. This is a pool of memory on the host which can hold a collection of events. `DAQ Filler` inserts events at the start of this list (it is a *producer*), and other processes (*consumers*) can copy events starting from the end of the list. When the pool gets full, old events are dropped from the end of the list to make room for new events. The events in the pool are tagged with the level-2 filter mask, so that events can be selectively read from the pool on the basis of what triggers and filters they passed. This mechanism is used to feed data to a set of programs used to monitor the quality of the data as it is being recorded ('examines').

Strictly speaking, the DAQ pool is local to a single node (since it is implemented using shared memory). However, it is possible to read events from the pool from another node by using the RDAQ (Remote DAQ) services. A program using RDAQ communicates over DECnet with a server process running on the node on which the DAQ pool resides. This server copies events out of the pool and ships them over the network to its client.

To return to the data written to the buffer disk: After receiving a command from the logger to copy a file to tape, the tape server assigns it to a tape, records this assignment in the tape database, and queues the file to be copied to that tape. If the file is in an express stream (signified by the stream name

starting with `EXP_`) the tape server copies the file to one of the expressline disks. Files in a global monitor stream (starting with `GMS_`) are also copied to the expressline disks, but are not written to tape.

The tape server is capable of writing data to any tape drive which is served to the cluster. However, using a tape drive which is not local to the node on which the server is running incurs a penalty of extra network I/O. In addition, the current VMS implementation of tape serving is not very robust, and has been linked to a number of system crashes. To solve these problems, tape copies should be performed by a process running on the node to which the tape drive is attached. This is accomplished by running a tape slave process on such nodes. Before the tape server attempts to mount a tape on a remote node, it attempts to connect to a tape slave running on that node. If it is successful, the tape server instructs the slave to mount the tape and handles all further copies to that tape by sending commands to the slave. If the server is unable to connect to a slave, it does the mount and copy itself.

After a file has been successfully copied to tape, it is removed from the buffer disk. In addition, a small text file (the *drool* file) is written to a particular directory in order to communicate information about the file to the offline production group. If there is an error while copying a file to tape, the tape is ejected, and that file and all other files which were queued for that tape are requeued to another tape.

3.9.3 Downloading

Most of the programmable hardware at DØ resides in one of a large number of VME crates. Through the use of a device called a *vertical interconnect*, the address spaces of several VME crates are mapped into another VME bus. This master crate will typically contain a dedicated 68000-based microcontroller, which is linked with other such controllers by an experiment-wide token ring. By sending it the appropriate messages, the microcontroller can be made to read from or write to any addresses in the VME crates which it controls. Communication between devices on the token ring and the host cluster is handled by three *gateways*, dedicated microvax computers which are connected both to the token ring and to the local online cluster network. Programs running on the host cluster talk to the gateways (and thus to any devices on the token ring) via a protocol called CDAQ (Control and Data Acquisition).

One of the primary users of CDAQ is a process called `COMM_TKR`. This process receives hardware configuration requests from COOR and is responsible for turning high-level, functional requests (such as 'turn on the pulser in such-and-such crate') into low-level, hardware-oriented requests (such as 'write 0x00080000 into VME address 0x00801240 in the front end at token-ring address 747'). The information needed to perform this conversion comes from the *hardware database*. `COMM_TKR` is also responsible for downloading calibration constants to the hardware. This data is kept in a number of DBL3 [95] databases. These databases are populated by periodically taking special calibration runs using a program called Calib.

3.9.4 Monitoring

Another set of processes is responsible for monitoring the health of the experiment. Chief among these is the alarm server, which provides a centralized collection point for alarm messages and other significant occurrences. Besides handling downloads, the microcontrollers on the token ring continually scan sets of environmental sensors, such as thermometers, pressure and humidity sensors, and readbacks from power supplies. If any of these readings fall outside of a preset tolerance, the microcontroller sends an alarm message to the alarm server. Any program running on the host cluster can also send error messages through the alarm system. Some alarms are deemed fatal; these indicate conditions which could compromise the quality of the data being taken. When a fatal alarm occurs, any physics run in progress is automatically paused until the condition is corrected.

Also used for monitoring incoming data are the 'examine' processes. These are typically run on the satellite workstations. They receive events from the data stream via RDAQ, run some subset of the full DØ event reconstruction on the data, and produce a set of histograms for each run which can help to identify potential hardware problems. The online event display also uses this mechanism.

3.10 Offline Data Processing

During run 1A, events were primarily written to two streams. The ALL stream contained all events sent to the host and was directly written to tape. The EXPRESS stream contained only events from a small subset of filters which was of the most interest for rapid analysis (mostly top and W/Z triggers). Besides being written to tape, EXPRESS stream events were copied to the express line, where they were reconstructed and rapidly made available for analysis. For run 1A, data was recorded at a rate of 2 – 3 Hz, with about 10% of the events going to the express line.

Every day, the raw data tapes which were written by the online system were collected and transported to the Feynman computing center. There, the raw data was reconstructed on a farm of Silicon Graphics machines [96]. The product of the reconstruction is two sets of files. STAs contain the raw data of the event augmented with the results of the reconstruction, and are about 600-1000 kbytes/event. DSTs contain only the reconstruction results for high-level objects, such as electrons, muons, etc. They are about 15 kbytes/event. Most analysis starts with the DSTs. STAs are usually required only when one wants to re-reconstruct an event, or to examine an event in detail with the event display.

The reconstructed events were further split into a set of offline streams. This was done on the basis of filter procedures supplied by the various physics groups, which were designed to select out the events of most interest. The filtered event streams were then made available to users on the DØ file server.

Chapter 4

Reconstruction and Particle ID

The raw event data which comes from the detector is given in terms of quantities such as digitized counts in a calorimeter cell, counts per time bin for a tracking chamber wire, and so on. However, these quantities in themselves are not very interesting. The patterns of ionization in the calorimeter and tracking chambers are presumably due to particles originating from a $p\bar{p}$ collision which interact within the detector; what one would like to know are the kinematic parameters of these physical objects. The process of turning the raw detector data into descriptions of objects such as leptons and jets is called *reconstruction*, and is carried out by a computer program called DØRECO.

4.1 The Reconstruction Program

The reconstruction process can be divided into three major phases:

- *Hit finding*, during which the raw data is unpacked and converted into 'hits' (i.e., energy deposits in calorimeter cells, or pulses on tracking

chamber wires) of definite energy and spatial location.

- *Tracking and clustering*, during which hits which are close together spatially are joined together to produce 'clusters' in the calorimeter and 'tracks' in the tracking chambers.

- *Particle identification*, during which information from all parts of the detector is combined to produce a collection of objects which are candidates for being jets, electrons, or muons. The criteria used for identifying these candidates are deliberately made quite loose so that they have high efficiency, but there will also be a large background. When performing an analysis, one will typically make much tighter cuts on the reconstructed objects. Both the reconstruction cuts and the tighter cuts used in the top analysis are detailed in the particle ID sections below.

Hit finding for the tracking chambers starts by unpacking the raw digitized data of charge versus time and identifying individual pulses by looking for leading and trailing edges. Each pulse is integrated to find the total deposited charge (used to calculate dE/dx). The time of arrival of the pulse (after corrections for channel-to-channel variations) is used to determine the position of the pulse: the time required to drift to the sense wire gives the distance of the hit from the sense wire, and the arrival time of the pulse from the delay line gives its location along the sense wire. Due to left/right ambiguities, there may be two possibilities for the location of a hit; both of the possibilities are used as input to the tracking phase. Due to the stagger in the sense wires, usually only the correct solutions will yield a good track.

In central detector tracking, the object is to identify groups of hits which lie along a straight line. Tracking is first done for each individual layer of the detector to produce track *segments*. Segments are then matched between the layers of each detector to form *tracks*. Finally, tracks are matched between the vertex chamber, the TRD, and the outer tracking chambers (CDC and FDC). For further details about the central detector hit finding and tracking, see [46, 48, 56, 57].

For the calorimeter, hit finding consists primarily of converting the energy deposited in each cell from digitized counts to GeV. This conversion ultimately comes from test beam measurements, in which the response of calorimeter modules to beams of known energy was measured [97, ch. 9]. Additional corrections are made for cell-by-cell variations in the electronics gain and pedestal values. These corrections are measured periodically during periods in which there are no collisions and are stored in a database, which the reconstruction program can later access.

Following cell unpacking, the cell energies are converted to transverse energy values using the position of the primary interaction vertex, as determined by the central tracking chambers. Cells with the same η and ϕ coordinates are summed together in the electromagnetic and hadronic calorimeters to produce *towers*. These towers are the input to the jet and electron clustering algorithms, described in the particle ID sections below.

The processing of the data from the muon system is similar in spirit to that done for the central detectors but quite different in detail, due to the differences in geometry and in the nature of the front-end electronics.

4.2 Vertex Finding

In order to calculate the transverse momentum or energy of a particle from the definitions (3.3) and (3.4), we must know θ, its polar angle in the lab frame. But for a particle which is detected only in the calorimeter, the only information we have about its direction is the location where it hits the calorimeter[1]. In order to extract an angle from this, we must also know one other point along its trajectory. This can be done by finding the location of the hard $p\bar{p}$ collision from which this particle presumably originated. This point is called the *interaction vertex* (or just the vertex of the event).

The x and y positions of the vertex can be known quite well simply due to the fact that the cross-section of the beam is made as small as possible in these dimensions, in order to maximize the luminosity. The typical cross-section of the beam was about $50\,\mu$m at a location of about 3–4 mm from the center of the detector with a drift over the length of a data run of less than $50\,\mu$m [98, pp. 14–16]. Thus, the (x, y) position of the vertex can be taken as a constant, and for many purposes can be set to $(0, 0)$ (the geometrical center of the detector).

The z-coordinate of the vertex, however, is less well constrained. Each bunch of particles in the Tevatron has some extent along the beam direction, and the resulting width of the z-distribution of interaction vertices in the detector is about 30 cm. Thus, it is necessary to measure the z-position of

[1] Actually, one could get an idea of the direction of the incident particle by looking at how the location of the shower varies with depth in the calorimeter. Something like this is done for isolated muons, but not for any other type of particle.

the vertex for each event individually. This is done using tracks found in the CDC [98, pp. 12–13] [66, pp. 39–41]:

- Take the tracks found in the CDC, and project them back towards the center of the detector.
- For each track, calculate the impact parameter — the minimum distance between the track and the z-axis of the detector. Throw out all tracks with an impact parameter larger than some cutoff. (This eliminates low-momentum tracks which have undergone a large amount of multiple scattering.)
- Project each track into the (r, z) plane, and compute the intersection with the z-axis. Histogram the z-positions of the intersections.
- Fit a gaussian around the peak of the resulting distribution. The mean is the estimate of the z-position of the vertex. The outlying regions of the histogram are also searched for any secondary peaks.

This procedure yields a resolution for the vertex z-coordinate of about 1–2 cm. Multiple vertices can typically be separated if they are at least 7 cm apart [66, pp. 40–41].

4.3 Electron Identification

Electrons are identified as localized deposits of energy in the electromagnetic calorimeter with an associated central detector track pointing back to the interaction vertex.

4.3.1 Candidate Construction

To identify electron candidates, the reconstruction program takes the following steps:

- Clusters are formed from calorimeter towers using a 'nearest neighbor' algorithm [99] [47, pp. 71–72]. Starting with the highest-E_T tower, adjacent towers are added to the cluster provided that they are above an E_T threshold, and that the cluster is not too big.

- A cluster is required to have at least 90% of its energy in the electromagnetic calorimeter, and at least 40% of the energy must be contained in a single tower.

- The centroid of the cluster is computed using the cells in the third electromagnetic layer. If $\vec{x}_i$ is the position of the center of cell i and E_i is the amount of energy deposited in that cell, then the centroid is the log-weighted center-of-gravity [100, pp. 83–85] [101]

$$\vec{x}_{\mathrm{COG}} = \frac{\sum_i w_i \vec{x}_i}{\sum_i w_i}. \tag{4.1}$$

  The weights w_i are

$$w_i = \max\left(0, w_0 + \ln\left(\frac{E_i}{\sum_j E_j}\right)\right). \tag{4.2}$$

  The parameter w_0 is chosen to optimize the position resolution, and the sums are over all EM3 cells in the cluster. The position resolution achieved is about 1.5–2 mm.

- Finally, the reconstruction program searches for a central detector track pointing from the interaction vertex to the calorimeter cluster within a "road" of $\Delta\eta = \pm 0.1$, $\Delta\phi = \pm 0.1$. If such a track is found, the cluster is identified as an electron candidate; otherwise, it becomes a photon candidate.

4.3.2 Selection Cuts

The cuts used in forming electron candidates are purposely quite loose; users are expected to apply further cuts which can be tailored to the particular analysis being carried out. There are numerous additional variables available for recognizing electrons; those used in this analysis [25] [100, pp. 81–87] are:

- Covariance matrix χ^2: The primary tool used for quantifying the information contained in the shape of the electromagnetic shower is the 'H-matrix' analysis [102, 103, 104]. Suppose one has a set of N observations of events of a given type, where each observation consists of M variables: $\mathbf{x}^i = (x_1^i, \ldots, x_M^i)$. One can form the covariance matrix from the outer products

$$\mathbf{V} = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{x}^i - \bar{\mathbf{x}})^T(\mathbf{x}^i - \bar{\mathbf{x}}), \tag{4.3}$$

where $\bar{\mathbf{x}}$ is the mean value of the N measurements:

$$\bar{\mathbf{x}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}^i. \tag{4.4}$$

The 'H-matrix' is then the inverse of this covariance matrix:

$$\mathbf{H} = \mathbf{V}^{-1}. \tag{4.5}$$

For any subsequent measurement $\mathbf{y}$ one can define a χ^2 which describes how likely it is that $\mathbf{y}$ came from the same sample as the $\mathbf{x}$'s:

$$\chi^2 = (\mathbf{y} - \bar{\mathbf{x}})\mathbf{H}(\mathbf{y} - \bar{\mathbf{x}})^T. \tag{4.6}$$

For the problem of electron identification, the events $\mathbf{x}$ used to build the H-matrix are Monte Carlo electron events. A total of 41 observables are used, consisting of the fractional energies in layers 1, 2, and 4 of the EM calorimeter, the fractional energies in each cell of a 6×6 array in the third EM layer (centered on the most energetic tower in the cluster), the z-position of the interaction vertex, and the logarithm of the total cluster energy. A separate matrix is built for each ring of calorimeter cells with the same $|\eta|$ coordinate. This analysis requires $\chi^2 < 100$ (41 degrees of freedom). Figures 4.1(a),(b) show the χ^2 parameter for electron and background samples.

- Isolation: The decay electron from a W should not be close to any other objects in the event. To quantify this, define the *isolation fraction.* Let $E_{\rm TOT}(0.4)$ be the energy deposited in all calorimeter cells within a cone of radius $R < 0.4$ around the electron direction (where $R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$). Similarly, let $E_{\rm EM}(0.2)$ be the energy deposited in the electromagnetic calorimeter within a cone of radius $R < 0.2$. Then

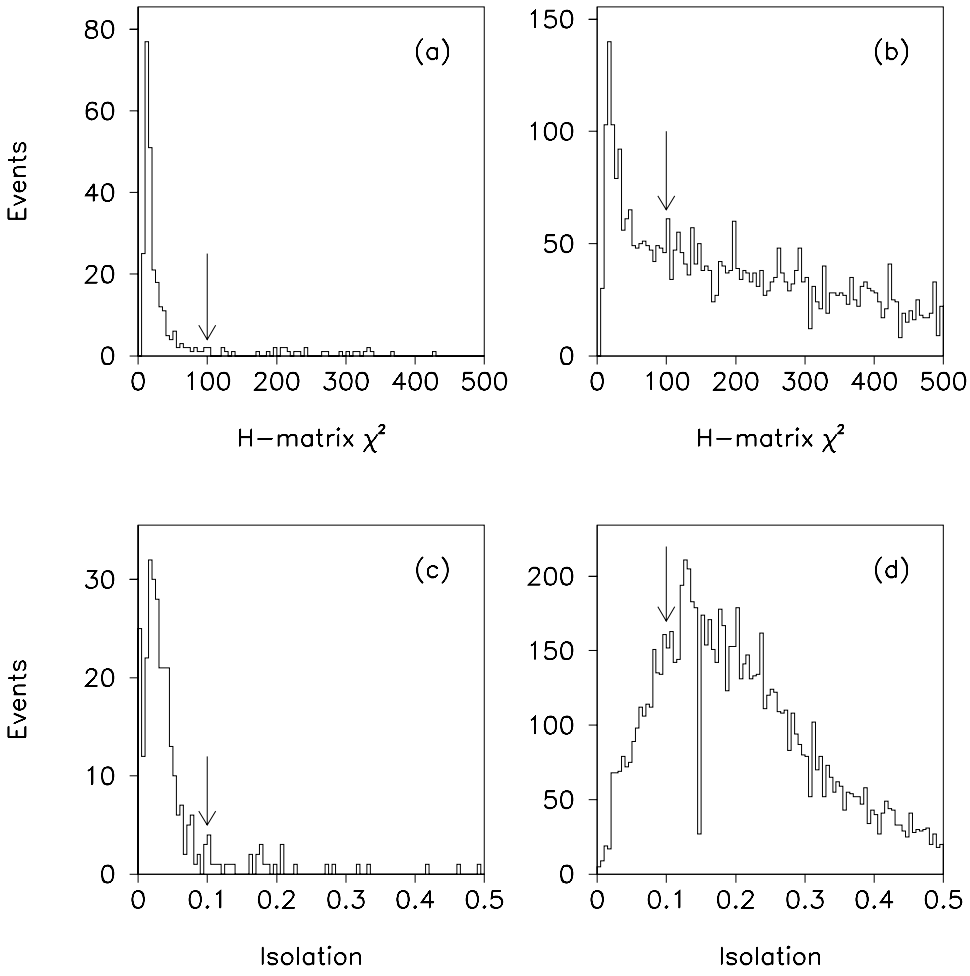

Figure 4.1: Covariance matrix χ^2 and isolation parameter. (a), (c) are from electrons in $Z \rightarrow ee$ events, and (b), (d) are from EM clusters in inclusive jet data. [25]

the isolation fraction $f_{\rm iso}$ is

$$f_{\rm iso} \equiv \frac{E_{\rm TOT}(0.4) - E_{\rm EM}(0.2)}{E_{\rm EM}(0.2)}. \tag{4.7}$$

This analysis requires $f_{\rm iso} < 0.1$. Figures 4.1(c),(d) show $f_{\rm iso}$ for electron and background samples.

- Track match significance: A significant source of background to electrons is photons, either produced directly or by the decay of π^0 and η mesons. Such photons do not create tracks in the central detector, but might appear to do so if some charged particle is nearby. This background can be reduced by requiring that the track point accurately at the centroid of the calorimeter cluster. To quantify this, define the *track match significance* [25]:

$$S = \sqrt{\left(\frac{\Delta\phi}{\sigma_{\Delta\phi}}\right)^2 + \left(\frac{\Delta z}{\sigma_{\Delta z}}\right)^2}, \tag{4.8}$$

where $\Delta\phi$, Δz are the coordinate differences between the cluster centroid and the point at which the track hits the calorimeter, and $\sigma_{\Delta\phi}$, $\sigma_{\Delta z}$ are the corresponding measurement resolutions. (This form is appropriate for the central calorimeter. In the end region, r replaces z.) This analysis requires $S < 5$. Figures 4.2(a),(b) show S for electron and background samples.

- Track ionization: Since DØ has no central magnetic field, e^+e^- pairs resulting from photon conversions do not diverge very far from each other, and are often reconstructed as a single track. However, the energy deposition per unit length (dE/dx) will be twice that of a single electron

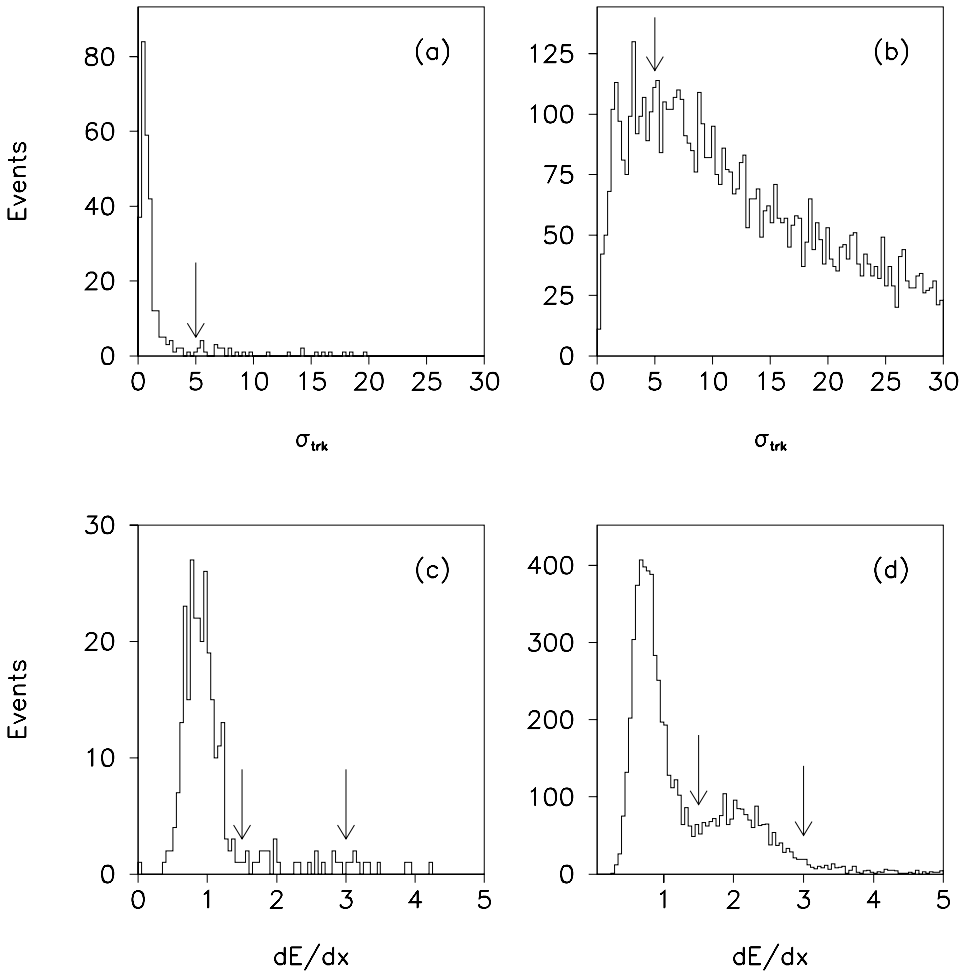

Figure 4.2: Track match significance S and ionization dE/dx (in the CDC). (a), (c) are from electrons in $Z \rightarrow ee$ events, and (b), (d) are from EM clusters in inclusive jet data. [25]

(which is called one 'MIP', for 'minimum ionizing particle'). Thus, the background due to conversions can be reduced by cutting out the region around 2 MIPs. For tracks in the CDC, the excluded region is $1.5 \leq dE/dx \leq 3.0$ and for tracks in the FDC, it is $1.3 \leq dE/dx \leq 2.7$. Figures 4.2(c),(d) show the ionization for signal and background samples.

In addition to these quality cuts, the analysis makes the kinematic cuts $E_T > 20\,\mathrm{GeV}$ and $|\eta^{\mathrm{det}}| < 2$. The 'detector-eta' η^{det} is the η which would be calculated with an interaction vertex fixed at the origin.

A summary of the electron selection cuts is given in Table 4.1. The efficiency for identifying electrons with these cuts is about 72% in the CC and 43% in the EC [25]. For the mass fitting Monte Carlo studies, where rejection of background is not an issue, a slightly looser set of cuts is used. These cuts are indicated in the 'relaxed cuts' column.

4.3.3 Electron Energy Corrections

The absolute energy scale of the calorimeters was originally set using test beam calibration data. However, due to differences in conditions between the test beam setup and the DØ installation, this calibration is slightly low. One useful point of reference is the mass of the Z in $Z \rightarrow e^+ + e^-$ events, which has been measured very accurately by the LEP experiments [3, p. 1356]. For this analysis, the measured electron energies are scaled up so that the mass peak in $Z \rightarrow e^+ + e^-$ matches the LEP measurement [100, p. 87]. This correction is about 5% in the central calorimeter, and 1–2% in the end calorimeters.

	Standard cuts	Relaxed cuts
Electromagnetic fraction	$F_{\rm EM} > 0.9$	$F_{\rm EM} > 0.9$
Cluster χ^2	$\chi^2 < 100$	$\chi^2 < 100$, in CC $\chi^2 < 200$, in EC
Isolation	$f_{\rm iso} < 0.1$	$f_{\rm iso} < 0.1$
Track match significance	$S < 5$	$S < 10$
Track ionization	CDC: $dE/dx < 1.5$ or $dE/dx > 3.0$ FDC: $dE/dx < 1.3$ or $dE/dx > 2.7$	—
Detector η	$\lvert\eta^{\rm det}\rvert < 2$	
Transverse Energy	$E_T > 20\,{\rm GeV}$	

Table 4.1: Summary of electron ID cuts.

4.4 Muon Identification

Muons are identified as tracks in the muon chambers which point back at the interaction vertex. There are two major backgrounds to contend with: cosmic ray muons, and leakage out of the backs of hadronic showers. This latter background is only important in the transition regions between the central and end calorimeters where there is not so much material.

4.4.1 Candidate Construction

Analogous to the central detector reconstruction, muon reconstruction proceeds by converting the raw hit and time information into three-dimensional hit positions, and then attempting to find hits which lie on straight lines

pointing towards the interaction vertex [105]. For muons, the situation is complicated by the fact that there is a magnetic field between the first and second layers of drift tubes. Thus, tracking is done separately for segments before and after the magnet. The segments are matched, and a measurement of the muon momentum can be made by measuring by how much the track bends while passing through the magnet.

The muon system by itself does not give a very good measurement of the muon momentum. However, the momentum resolution can be improved if the muon track can be associated with a track in the central detector and with an interaction vertex. This allows the trajectory of the muon before the magnet to be determined more accurately. (The process is referred to as *global fitting* [106].)

The momentum measured by the muon system is, of course, its momentum *after* it has passed through the calorimeter. However, a muon will typically lose several GeV of energy in the calorimeter. This energy loss is estimated through a Monte Carlo calculation and added to the measured muon momentum (for the dilepton channels, the measured energy loss is used [107]). The fact that a real muon will deposit a small, but nonzero, amount of energy in the calorimeter will also be useful for background rejection.

4.4.2 Selection Cuts

Muons are required to be entirely contained in the WAMUS system, with $\eta < 1.7$. (For run 1B data, muons are required to be contained in the

central muon (CF) system, due to efficiency problems in the end chambers caused by chamber aging.) For muons from W decay, this analysis requires $p_T > 15\,\mathrm{GeV}/c$. In addition, just as for electrons, there are a number of quality cuts made on candidate muon tracks [25] [43, pp. 70–83]:

- A-stubs: The latest version of the muon reconstruction can form muon tracks consisting of only hits in the innermost (A) layer. These 'A-stubs' are excluded from further consideration.

- Impact Parameters: Two impact parameter cuts are used to require that the muon tracks point towards the interaction vertex, and thus reject cosmic ray backgrounds.

  The *non-bend impact parameter* is defined by projecting the muon track into the x, y plane (it does not get bent in this plane), extrapolating the track formed by the B and C layers towards the center of the detector, and calculating the impact parameter between this extrapolated track and the interaction vertex. This impact parameter is required to be less than 40 cm.

  The *bend-view impact parameter* is calculated by projecting the track into the plane in which the muon bends and calculating the impact parameter of this projection. This parameter is required to be less than 25 cm.

- Cosmic Ray Veto (MUCTAG): A cosmic ray muon which penetrates the entire detector will leave hits in the muon chambers on both sides of the

interaction region. Therefore, a track in the central region ($\eta^{\rm det} < 1.0$) is rejected if there is another track or an excessive number of hits located back-to-back in both ϕ and η [108].

- Track Timing: When trying to find the location of a track using a drift chamber, one needs to know the time at which the particle went through the chamber (the 'T_0'). Normally, this is deduced from the time at which the beams cross. However, this will not be correct in general for cosmic rays, which are not synchronized with the accelerator. Thus, for a cosmic ray, the quality of the track can often be improved by taking a T_0 different from the beam crossing. This provides another way to reject cosmic rays: the track χ^2 is minimized with respect to the T_0 and the result compared with the nominal T_0 of the beam crossing. If the difference is larger than 100 ns, the track is rejected.

- Hit Multiplicity: A muon track will typically have hits in 7–10 drift tubes, depending on the region of the detector. High-p_T muon tracks in the end regions are required to have at least 5 hits (no explicit cut is made in the central region).

- Muon Quality (IFW4): For each track, the muon reconstruction code makes a set of cuts on the number of modules hit, impact parameters, and hit residuals. The number of cuts which the track fails is called 'IFW4'; it is required to be no more than one.

- Calorimeter Confirmation: A muon passing through the detector will typically deposit between 1 and 3 GeV of energy in the calorimeter. The energy contained in all cells within a one cell radius surrounding the muon track is summed and is required to be at least 1 GeV. If there is no CD track matched with the muon, the cut is set higher, at 1.5 GeV.

- Path Length Through Iron Toroids: Muons which pass through the thin region of the iron toroid (in the range $0.8 < |\eta| < 1.0$) are poorly measured, and may be contaminated by a background of leakage from the tail ends of hadronic showers. To reject such tracks, the integrated magnetic field $\int B \cdot dl$ is required to be greater than 1.83 Tm.

- Isolation: The analysis requires that the distance in R between the muon and the nearest jet be greater than 0.5.

The above cuts define high-p_T isolated muons. It is also of interest to look for soft muons from semileptonic decays of b quarks. These muons are selected using requirements which are somewhat looser than those outlined above. In addition, a tag muon is required to be within $R < 0.5$ of a jet. A summary of the muon ID cuts is given in Table 4.2. The efficiency for identifying muons with these cuts is about $\sim 41\%$ [25].

4.5 Jet Reconstruction

When a quark or gluon leaves the site of a hard scattering, it cannot remain free, but instead *hadronizes*, or *fragments* into a collection of colorless

	Isolated muon	Tag muon
Momentum	$p_T > 15\,\mathrm{GeV}$	$p_T > 4\,\mathrm{GeV}$
Pseudorapidity	$\lvert\eta\rvert < 1.7$ [a]	$\lvert\eta\rvert < 1.7$ [a]
Cosmic veto	yes	—
A-stub removal	yes	yes
Impact Parameters	BVIP $< 25\,\mathrm{cm}$ NBIP $< 40\,\mathrm{cm}$	—
Floating T_0	$\Delta T_0 < 100\,\mathrm{ns}$	—
Minimum Hits	$N_{\mathrm{hits}} \geq 5$, in EF	$N_{\mathrm{hits}} \geq 4$
Track Quality	IFW4 ≤ 1	IFW4 ≤ 1, for $\lvert\eta\rvert < 1$ IFW4 $= 0$, for $\lvert\eta\rvert > 1$
Calorimeter Confirmation	$> 0.5\,\mathrm{GeV}$, w/ CD track $> 1.5\,\mathrm{GeV}$, otherwise	$> 1.5\,\mathrm{GeV}$
Path Length	$\int B \cdot dl > 1.83\,\mathrm{Tm}$	—
Isolation	$\Delta R(\mu,\mathrm{jet}) \geq 0.5$	$\Delta R(\mu,\mathrm{jet}) < 0.5$

Table 4.2: Summary of muon ID cuts.

[a] Restricted to CF region ($\lvert\eta\rvert \lesssim 0.7$) for run 1B.

hadronic particles. This collection will typically lie in a cone around the direction of motion of the original parton, and will show up in a calorimeter as a cluster of energy. This is called a *jet*.

If this sounds somewhat ill-defined, that's because it is. To a much greater extent than for electrons and muons, the results one gets from a jet reconstruction depend on exactly how a jet is defined.

The jet definition most commonly used in a $p\bar{p}$ environment is the 'cone algorithm', in which jets are taken to be the energy inside of cones of a fixed radius (in η–ϕ space). This definition was used by UA1 [109, 110], UA2, and CDF, and is also used by most DØ analyses; a description of this algorithm is given below. Another method which has been explored is the 'nearest neighbor' algorithm, which employs a clustering method similar to that used for electrons. In principle, it should yield a better resolution than the cone algorithm, but at present, it is less well understood.

For further information about jets at DØ, see [66, ch. 3], [111, ch. 4], [112].

4.5.1 Cone Jet Algorithm

The following is a description of the cone jet algorithm as used at DØ [113] [112, ch. 4] [100, pp. 88-90]:

- Preclustering: The calorimeter towers (see Section 4.1) are first sorted in E_T, and a set of 'seed' clusters are formed. Starting with the highest-E_T tower which has not yet been assigned to a precluster, the precluster

is formed from all contiguous towers within $|\Delta\eta| < 0.3$, $|\Delta\phi| < 0.3$ with $E_T > 1\,\mathrm{GeV}$. Preclustering continues until all towers with $E_T > 1\,\mathrm{GeV}$ have been assigned to a seed cluster. For each precluster, the E_T-weighted centroid defines the axis of the corresponding jet candidate.

- Cone Clustering: A new cluster is defined around the trial axis including all calorimeter cells within a fixed distance in η, ϕ space. The centroid of this new cluster is computed, which defines a new jet axis. This process is then repeated until it stabilizes.

- Merging and Splitting: Once the cone clustering has completed, some cells may turn out to have been assigned to more than one jet. If two jets share some cells, the fraction of the total energy which is shared between them is examined. If the fraction is greater than 50%, the two jets are merged together, and the jet axis recalculated from the centroid of all the cells in the merged jet. Otherwise, the jets are split, and each shared cell is assigned to the closest jet.

- E_T Cut: A jet is required to have a total E_T above a threshold, which is usually set to 8 GeV.

- Jet E_T Definition: For this analysis, the E_T of a jet is defined by

$$E_T = \sqrt{E_x^2 + E_y^2}, \tag{4.9}$$

where E_x and E_y are the sums of the components of the individual cell energies:

$$E_x = \sum_i E_x^i$$
$$E_y = \sum_i E_y^i. \tag{4.10}$$

For event selection, this analysis will use jets with a radius of $R = 0.5$. Kinematic fitting, however, will use narrower jets, with radius $R = 0.3$.

4.5.2 Jet Corrections

Ideally, one would like the measured jet energy to give back the energy of the original parton which formed the jet. However, there are systematic biases in jet measurements which need corrections. In addition to having to determine the energy scale (just as for electrons), there are several other effects which become important due to the extended, multiparticle nature of jets:

- Many of the particles in even a high-E_T jet will be fairly soft ($\lesssim 2\,\mathrm{GeV}$). However, the response of the calorimeter becomes nonlinear in this region, so simply summing the calorimeter responses to each particle will not give the correct total energy.

- Since the hadronic shower is an extended object, some portion of the shower may extend beyond the jet cone. In addition, some of the particles radiated by the initial parton may fall outside of the cone.

- A jet will pick up some extra energy due to the underlying ('spectator') event, as well as noise due to the natural radioactivity of the uranium

absorber. Both of these sources affect jets much more than electrons due to the fact that hadronic showers are much larger than electromagnetic showers.

- The zero-suppression used in the calorimeter readout can also give rise to a systematic shift in energy which scales with the shower size.

The method adopted to obtain the jet corrections at DØ is called the Missing E_T Projection Fraction (MPF) method. It was originally used by the CDF collaboration [114]; its application at DØ is described in [111, pp. 54–63] [100, pp. 90-94] [112, 115, 116]. Briefly, the idea is to look at events which consist of one jet with a very high electromagnetic content (it is required to pass the reconstruction cuts for a photon), a second jet lying opposite in ϕ, and no other objects in the event. The EM jet is corrected using the electron corrections. There should be no energetic neutrinos in these events, so any missing transverse energy ($\not\!\!E_T$) remaining in the event can then be attributed to a mismeasurement of the hadronic jet. By projecting the $\not\!\!E_T$ along the jet axis, the needed correction for the jet can be derived. This is averaged over similar jets in the sample to produce a correction which is a function of jet E_T, η, and electromagnetic content. Out-of-cone showering, underlying event, and noise effects were determined in separate studies using Monte Carlo and minimum-bias event samples. The resulting corrections are shown for two different η values in Figure 4.3.

Note that the MPF corrections are not entirely sufficient to return to the parton level. While the effects of showers leaking out of the jet cone is

accounted for, energy lost due to QCD radiation outside of the jet cone is not. Corrections for this effect are made later, before doing kinematic fitting.

4.6 Missing Energy Reconstruction

From momentum conservation and the fact that the colliding proton and antiproton have nearly opposite momenta, it follows that the total vector sum of the momenta of all final-state particles in the event must be zero. However, one cannot usefully apply total momentum conservation in the direction along the beam, since many particles will escape detection by going down the beam pipe. But the very fact that they do escape implies that they have very small transverse momenta; thus, one can apply momentum conservation in the plane perpendicular to the beam. If the sum of the transverse momenta of the detected particles is significantly different from zero, the discrepancy is attributed to one or more neutrinos which escaped detection, and which must have transverse momentum opposite the total detected transverse momentum.

4.6.1 Definition of Missing Transverse Energy

Each cell in the calorimeter is given a four-vector, with an energy equal to the measured energy in the cell, a direction pointing from the interaction vertex to the center of the cell, and a mass of zero. The transverse components of these vectors are summed over all the calorimeter cells (including the ICD).

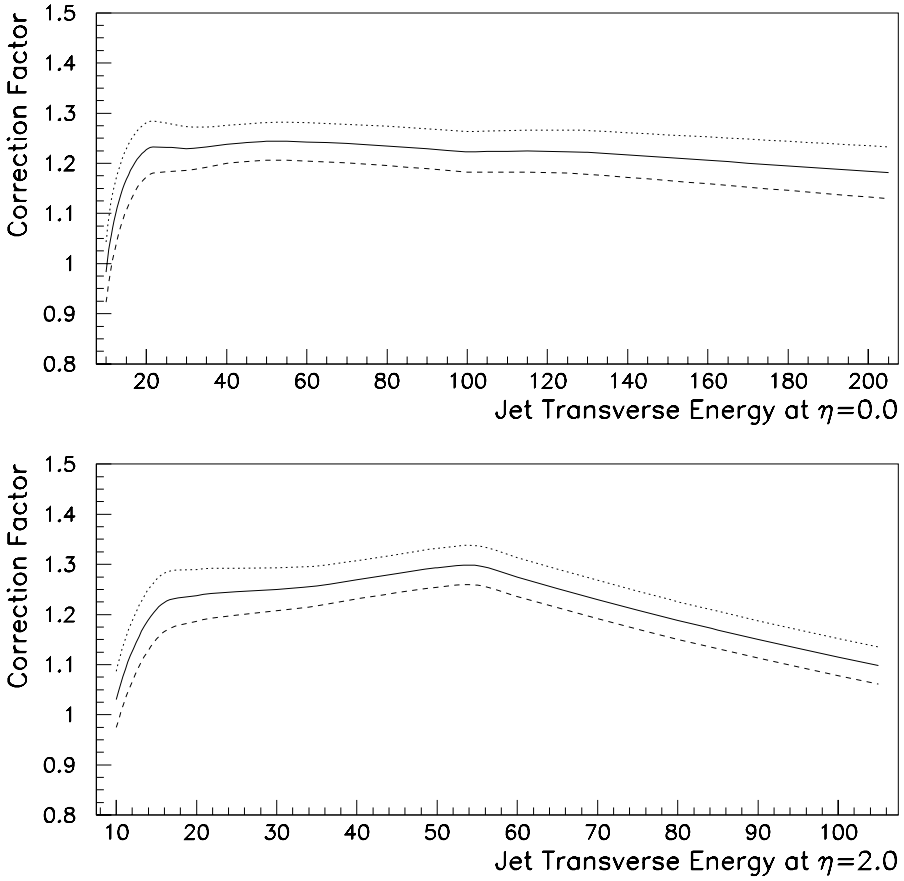

Figure 4.3: Energy scale corrections for $R = 0.5$ cone jets. Shown for both the central region ($|\eta| = 0.0$) and the forward region ($|\eta| = 2.0$). [25]

The negation of this vector is then the 'calorimeter missing-E_T', or $\vec{\not\!\!E}_T^{\rm cal}$:

$$\vec{\not\!\!E}_T^{\rm cal} = -\sum_i \vec{E}_T^i. \tag{4.11}$$

This does not yet take into account muons, which (for high-p_T muons) only deposit a small portion of their energy in the calorimeter. The 'total missing E_T', or just $\vec{\not\!\!E}_T$, is then $\vec{\not\!\!E}_T^{\rm cal}$ with the transverse momenta of all good muon tracks subtracted.

4.6.2 Corrections

Since the $\not\!\!E_T^{\rm cal}$ is the sum over everything in the calorimeter, it follows that if any object in the calorimeter is mismeasured, then $\not\!\!E_T^{\rm cal}$ will be mismeasured by exactly the same amount. So whenever corrections are made to electrons and jets, the corresponding correction must be made to $\not\!\!E_T^{\rm cal}$ [117]. This is straightforward to do: as each object is corrected, one simply adds the uncorrected object $\vec{E}_T$ to $\vec{\not\!\!E}_T^{\rm cal}$ and then subtracts the corrected object $\vec{E}_T$.

Chapter 5

Selection Cuts and Background Calculation

This chapter summarizes the so-called 'counting' analysis, which furnishes the main evidence for the existence of top. The idea is to construct a set of criteria, or 'cuts', which should preferentially select the top signal over the background processes. One then counts the number of events in the data which pass the cuts and compares it to the number of events which one would expect from the backgrounds. Any statistically significant excess could be evidence for a top signal. The results of this analysis, in the form of a set of data events which pass the cuts and the expected background in the sample, also form the input to the mass analysis.

The results here correspond to those presented in [1]. More detailed information about the techniques used may be found in [25] and also in [43, 47, 100, 118, 119].

The data used are the complete run 1A sample which was recorded from July, 1992 through June, 1993 plus the data from run 1B recorded from December, 1993 through January, 1995. The total integrated luminosity is about

$13.5\,\mathrm{pb}^{-1}$ from run 1A and $37\,\mathrm{pb}^{-1}$ from run 1B.

5.1 Event Simulation

In order to evaluate the detailed response of the detector to a potential top signal, Monte Carlo methods must be used to construct samples of simulated top events. In addition, while the background calculation is based as much as possible on real data, it is still necessary to use simulated data to investigate the details of the event shapes. This section lists the computer programs used to perform the event simulations for this analysis. For further information, see [25] [47, ch. 5] [100, ch. 5].

5.1.1 Signal Simulation

To model the expected $t\bar{t}$ signal, the ISAJET Monte Carlo [120] is used. When cross sections are needed, the ISAJET results are rescaled to match the calculation of [17] (see Section 2.3). The output of ISAJET is then passed through the standard (showerlibrary) DØGEANT detector simulation [121, 122, 123, 124], trigger simulator [125, 126], and reconstruction program (see Chapter 4).

To evaluate systematic errors due to the event generator, the results from ISAJET are compared with those from the HERWIG Monte Carlo [127]. The differences between the two are usually small.

5.1.2 W + jets Background Simulation

The VECBOS Monte Carlo program [128] is used to model the W + jets and Z + jets backgrounds. This program incorporates the exact leading order matrix elements for W, Z production with up to four additional final-state partons. However, it does not include higher order and fragmentation effects. To estimate these effects, the output from VECBOS is fed into ISAJET starting at the QCD evolution phase [25] [100, pp. 107–117]. As before, the result is then passed through the detector simulation and reconstruction.

The settings of the VECBOS parameters were

- $E_T > 10\,\text{GeV}$ for all final-state partons.
- $\Delta R > 0.5$ for every pair of jets.
- CTEQ1M structure functions.
- $\langle Q^2 \rangle = m_W^2$ for the dynamical scale.

5.2 Summary of Dilepton Channels

Since the dilepton channels are not used for mass fitting, only a brief summary of these results will be presented. For more details about these channels, see [25, 47, 118, 119].

The signature for the dilepton channels is two high-p_T leptons, two jets, and missing E_T. The precise values used for the cuts are listed in Table 5.1. The $\not\!\!E_T$ cut is not made for the $\mu\mu$ channel, due to the relatively poor muon

	ee + jets	$e\mu$ + jets	$\mu\mu$ + jets
Leptons	$E_T(e) > 20\,\text{GeV}$ $\lvert\eta(e)\rvert < 2.5$	$E_T(e) > 15\,\text{GeV}$ $\lvert\eta(e)\rvert < 2.5$ $p_T(\mu) > 12\,\text{GeV}/c$ $\lvert\eta(\mu)\rvert < 1.7^a$	$p_T(\mu) > 15\,\text{GeV}/c$ $\lvert\eta(\mu)\rvert < 1.7^a$
Missing E_T	$\not\!\!E_T > 25\,\text{GeV}$	$\not\!\!E_T > 10\,\text{GeV}$ $\not\!\!E_T^{\text{cal}} > 20\,\text{GeV}$	—
Jets	2 jets with $E_T > 15\,\text{GeV}$ and $\lvert\eta\rvert < 2$		
Additional cuts	$H_T > 120\,\text{GeV}$ $\not\!\!E_T > 40\,\text{GeV}$ if $79 < m_{ee} < 103$ (GeV/c^2)	$H_T > 120\,\text{GeV}$ $\Delta R(e,\mu) > 0.25$	$H_T > 100\,\text{GeV}$ $\chi^2 < 0.01$

Table 5.1: Summary of kinematic selection cuts for the dilepton channels. [1]

[a] Restricted to CF for run 1B.

momentum resolution. Instead, a constrained fit is made to the hypothesis $Z \to \mu\mu$, and a cut made on the χ^2 of the fit. For all channels, an additional cut is made on the quantity H_T, which is defined as the scalar sum of the transverse energies of the jets plus the transverse energy of the leading electron. (For the $\mu\mu$ channel, only the jets are used.)

There are three dilepton events which pass all the cuts. The major backgrounds are due to Z and continuum Drell-Yan production ($Z, \gamma^* \to ee, \mu\mu, \tau\tau$), vector boson pairs ($WW$, WZ), heavy flavor ($b\bar{b}$ and $c\bar{c}$), and jets misidentified as electrons; the methods used to calculate them are the same as described in [25, 35] (also see [47]). The total expected background for the dilepton channels is 0.65 ± 0.15 events. Table 5.2 summarizes the dilepton results.

5.3 Lepton + Jets Channels

The signature for the lepton + jets channels is an isolated high-p_T lepton, missing E_T and several (3–4) jets (see Section 2.4). The major backgrounds to these channels are W + jet production and QCD multijets where one jet is misidentified as a lepton. Even after requiring the presence of all the above objects, the background is still many times larger than the expected signal, especially for for higher top masses.

Due to this large background, additional cuts must be made to improve the signal/background ratio. Two complimentary approaches are used. The first involves making cuts on global quantities which describe aspects of the

m_t (GeV/c^2)		$e\mu$ + jets	ee + jets	$\mu\mu$ + jets	All dilepton
	$\varepsilon \times \mathcal{B}(\%)$	0.17 ± 0.02	0.11 ± 0.02	0.06 ± 0.01	
140	$\langle N \rangle$	1.36 ± 0.21	1.04 ± 0.19	0.46 ± 0.08	2.86 ± 0.29
	$\varepsilon \times \mathcal{B}(\%)$	0.24 ± 0.02	0.15 ± 0.02	0.09 ± 0.02	
160	$\langle N \rangle$	0.94 ± 0.13	0.69 ± 0.12	0.34 ± 0.07	1.97 ± 0.19
	$\varepsilon \times \mathcal{B}(\%)$	0.28 ± 0.02	0.17 ± 0.02	0.10 ± 0.02	
180	$\langle N \rangle$	0.57 ± 0.07	0.40 ± 0.07	0.19 ± 0.04	1.16 ± 0.11
	$\varepsilon \times \mathcal{B}(\%)$	0.31 ± 0.02	0.20 ± 0.03	0.11 ± 0.02	
200	$\langle N \rangle$	0.34 ± 0.04	0.25 ± 0.05	0.11 ± 0.02	0.70 ± 0.07
Background		0.12 ± 0.03	0.28 ± 0.14	0.25 ± 0.04	0.65 ± 0.15
$\int \mathcal{L}dt$ (pb^{-1})		47.9 ± 5.7	55.7 ± 6.7	44.2 ± 5.3	
Data		2	0	1	3

Table 5.2: Summary of results from dilepton channels, showing the efficiency $\times$ branching ratio ($\varepsilon \times \mathcal{B}$) and the expected number of top events ($\langle N \rangle$). The central value of the cross section of [17] is used. Also given are the expected background, integrated luminosity, and the number of observed events in each channel. From [1].

event shape. This is sometimes called the 'topological' analysis. The second approach is to look for a nonisolated low-p_T muon, such as would be produced by a semileptonic b-decay. This is a good cut for separating the signal from background, since each $t\bar{t}$ decay will have two b-jets, but the backgrounds should contain relatively little heavy flavor. This is called the 'tagging' analysis. The cuts used to define these two analyses are constructed so that they do not overlap either with each other or with the dilepton channels, so that all the channels may be combined at the end.

For each analysis, two sets of cuts are defined. The 'standard' cuts were tuned (on Monte Carlo samples) to optimize the expected significance for a heavy ($\gtrsim 160\,\mathrm{GeV}$) top. The 'loose' cuts back off on the total energy cuts. This provides a useful consistency check, as well as avoiding extra biases in mass fitting.

5.3.1 Topological Analysis

The selection cuts for the topological analysis require a high-p_T lepton, a large missing E_T, and at least four jets. To ensure orthogonality with the tagged channels, events are rejected if they contain a μ-tag.

Two additional variables are used to define the selection cuts. The first is H_T, which is defined as the scalar sum of the transverse energies of all jets which pass the selection cuts:

$$H_T \equiv \sum_i |E_T(\mathrm{jet}_i)|. \tag{5.1}$$

A large H_T is a signature of the decay of a massive object.

The second variable used is the aplanarity $\mathcal{A}$ [8, p. 280] [129]. This is defined in terms of the normalized three-momentum tensor constructed from the selected jets:

$$M^{ab} \equiv \sum_i p_i^a p_i^b / \sum_i p_i^2. \tag{5.2}$$

The aplanarity is then

$$\mathcal{A} \equiv \frac{3}{2} \times (\text{smallest eigenvalue of M}). \tag{5.3}$$

The aplanarity takes on its maximum value of 0.5 for a spherical event. For a planar or linear event, it is zero. Top events tend to be more spherical than events due to radiative QCD background processes [100, pp. 135–138].

The triggers used for these channels varied somewhat over the course of the two runs. However, the general strategy was to require a relatively high-E_T central lepton candidate, a central jet in addition to the lepton, and, for the $e+$jets channel, some missing E_T in the calorimeter. Some additional cuts are made to reject events which are contaminated by leakage from the Main Ring or by instrumental effects ('hot cells') in the calorimeter. The selection cuts used are summarized in Table 5.3.

The results of this selection are shown in Figure 5.1 as scatter plots in the $\mathcal{A} - H_T$ plane for the data, 180 GeV ISAJET top Monte Carlo, and the expected backgrounds. A total of 8 events survive the standard cuts and 23 events survive the loose cuts.

The total normalization of the background is derived entirely from data. However, the VECBOS Monte Carlo is employed to find the shapes of the $\mathcal{A}$ and H_T distributions for the $W+$jets background.

	e + jets	μ + jets
Trigger	$E_T(e) > 15\,\text{GeV}$ $\lvert\eta(e)\rvert < 2.5$ 2 jets with $E_T(j) > 10\,\text{GeV}$ $\lvert\eta(j)\rvert < 2.5$ $\not\!\!E_T^{\text{cal}} > 13\,\text{GeV}$	$p_T(\mu) > 10\,\text{GeV}$ μ track in WAMUS 1 jet with $E_T(j) > 15\,\text{GeV}$ $\lvert\eta(j)\rvert < 2.5$
Lepton	$E_T(e) > 20\,\text{GeV}$ $\lvert\eta^{\text{det}}(e)\rvert < 2$	$p_T(\mu) > 15\,\text{GeV}/c$ $\lvert\eta(\mu)\rvert < 1.7^a$
Missing E_T	$\not\!\!E_T^{\text{cal}} > 25\,\text{GeV}$	$\not\!\!E_T^{\text{cal}} > 20\,\text{GeV}$ $\not\!\!E_T > 20\,\text{GeV}$
Jets	4 jets with $E_T > 15\,\text{GeV}$ and $\lvert\eta\rvert < 2$	
Loose cuts	$\mathcal{A} > 0.03$	
Standard cuts	$\mathcal{A} > 0.05$ $H_T > 200\,\text{GeV}$	
	No soft μ tag	

[a] Restricted to CF for run 1B.

Table 5.3: Summary of kinematic selection cuts for the untagged lepton + jets channels. [1]

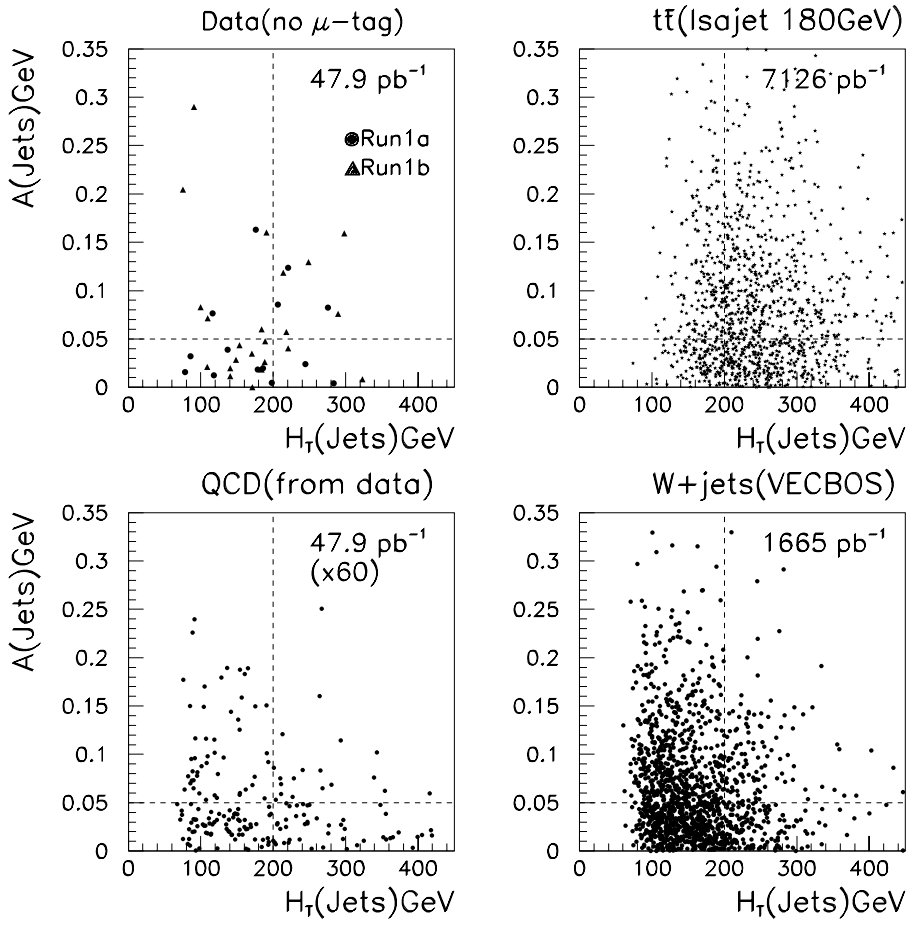

Figure 5.1: Results from the lepton + jets selection, for data, 180 GeV ISAJET top Monte Carlo, QCD fakes, and VECBOS W + jet Monte Carlo. [130]

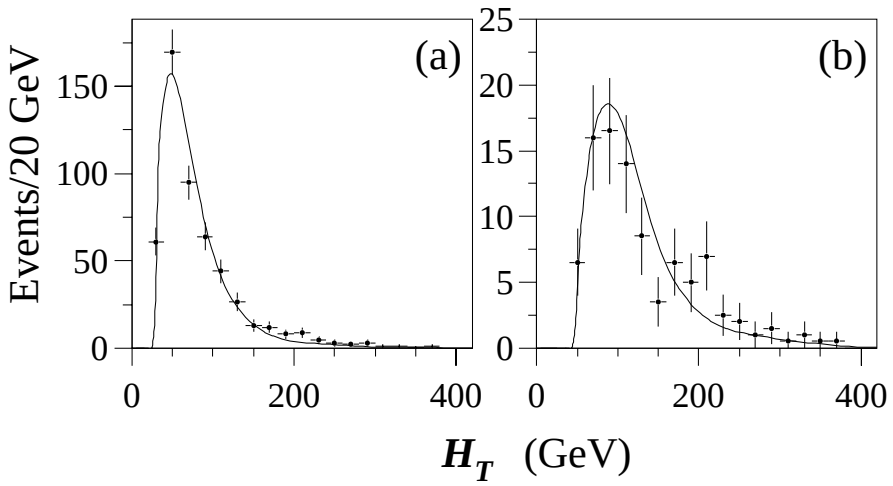

Figure 5.2: Observed H_T distributions (points) compared to the distributions expected from background (line) for $\not\!\!E_T > 25\,\mathrm{GeV}$ for (a) $e+ \geq 2$ jets and (b) $e+ \geq 3$ jets. [1]

The shape of the non-W QCD multijet background is estimated by using events which contain a 'bad' lepton candidate and several jets. For the $e +$ jets channel, the sample consists of multijet events which contain a highly electromagnetic jet which fails the final electron-ID cuts. This jet is then treated as an electron for the remainder of the analysis. For the $\mu +$ jets channel, the sample used consists of events containing muons which fail the isolation cut and are thus embedded in a jet.

These background models can be tested by looking at event selections which are dominated by background. Figure 5.2 compares the H_T distributions for the data and the calculated backgrounds for $e + 2$ jet and $e + 3$ jet events; good agreement is seen. (Here, and in what follows, jet multiplicities are *inclusive*; i.e., $e + 2$ jets means 'two or more jets.')

The calculation of the background normalization is based on the fact that

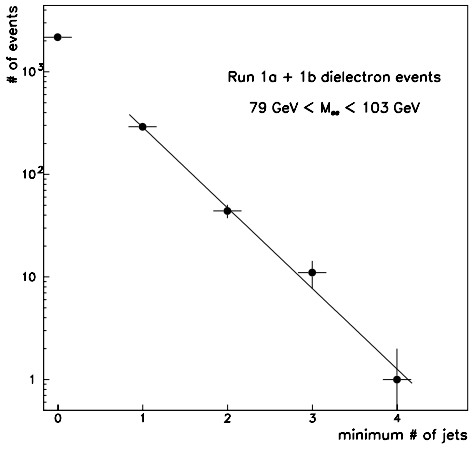

Figure 5.3: Jet multiplicity spectrum for $(Z \rightarrow ee) +$ jets events. [130]

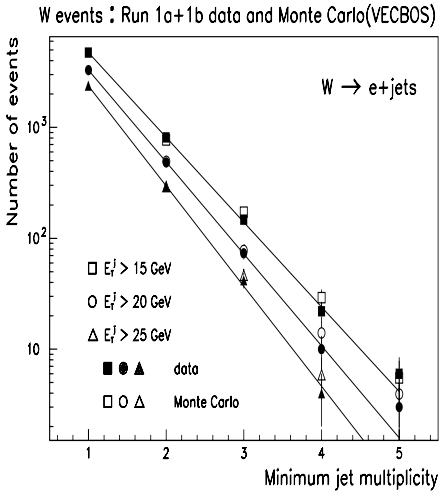

Figure 5.4: Jet multiplicity spectrum for $(W \rightarrow e\nu) +$ jets events. [130]

the background jet multiplicity spectrum for radiative QCD processes falls nearly exponentially; i.e.,

$$\frac{\text{Number of } l+ \geq n \text{ jet events}}{\text{Number of } l+ \geq (n-1) \text{ jet events}} \approx \text{const.} \tag{5.4}$$

Roughly speaking, each additional jet adds an extra factor of the strong coupling α_s to the cross section. This was suggested on theoretical grounds in [131], and has been shown to work well empirically for small n ($n \leq 5$) for $W +$ jets, $Z +$ jets, QCD multijet, and VECBOS $W +$ jets samples. This is illustrated for Z and W events in Figures 5.3 and 5.4.

For the $e+$jets channel, the calculation starts by estimating the amount of non-W background (dominantly QCD multijets) present in the sample. This is done by taking the data sample before the lepton-ID and $\not\!\!E_T$ cuts and dividing it up into four sets corresponding to the four possible outcomes for the two cuts. Label the set containing events which passed both cuts 1 and the other

three 2, 3, and 4. Then the number of observed events in each set N_i can be written as a sum of W and QCD contributions:

$$N_i = N_i^W + N_i^Q. \tag{5.5}$$

The distribution of W events among the four sets is determined using the VECBOS Monte Carlo. This determines the ratios

$$A_{ij} = \frac{N_i^W}{N_j^W}, \tag{5.6}$$

and gives the three additional equations

$$N_i = A_{i1} N_1^W + N_i^Q, \quad \text{for } i = 2, 3, 4. \tag{5.7}$$

One additional constraint is needed in order to obtain a unique solution. It is assumed that for QCD events, the lepton-ID and $\not\!\!E_T$ cuts are independent; i.e.,

$$\frac{N^Q_{\text{good e, good } \not\!\!E_T}}{N^Q_{\text{bad e, good } \not\!\!E_T}} = \frac{N^Q_{\text{good e, bad } \not\!\!E_T}}{N^Q_{\text{bad e, bad } \not\!\!E_T}} \tag{5.8}$$

This system of equations can be solved iteratively to yield N_1^Q and N_1^W. This is repeated for each jet multiplicity, and the resulting estimates for the QCD background are subtracted from the data. The result for the e + jets channel is plotted in Figure 5.4.

The e+jets data now consist of a QCD-subtracted event count for each jet multiplicity. Change notation slightly and denote these points by N_i, where i is the jet multiplicity. The data are now considered to be the sum of W + jets and top contributions, and are fit to the form

$$N_i = \alpha^{i-1} N_W + f_i N_t. \tag{5.9}$$

Here, N_W is the number of $W+ \geq 1$ jet events, N_t is the total number of top events, f_i is the fraction of top events expected at each multiplicity, and α is the ratio from the $W+$jets scaling law

$$\alpha = \frac{\text{Number of } W+ \geq n \text{ jet events}}{\text{Number of } W+ \geq (n-1) \text{ jet events}}. \tag{5.10}$$

The top fractions f_i are determined using Monte Carlo, and N_W, N_t, and α are obtained from the fit. This is then evaluated to yield the expected number of $W+$jets events with four or more jets.

The final step is to apply the $\mathcal{A}$ and H_T cuts. The efficiencies for these cuts are found separately for the $W+$jets and QCD backgrounds using the VECBOS Monte Carlo and the QCD multijet samples, respectively. The background estimates are then multiplied by these efficiencies.

The estimation of the $\mu+$jets background is similar, except that the QCD background is not subtracted off before the fit to the scaling law. Instead, the scaling law is used to obtain the total background, including both $W+$jets and QCD multijets. The amount of non-W background contributing to the final background number is estimated using the nonisolated muon sample.

The results for the $e+$jets and $\mu+$jets channels are summarized in Table 5.4.

5.3.2 Tagging Analysis

The basis of the tagging analysis is to require the presence of a nonisolated muon, such as would be produced by the semileptonic decay of a b-jet via the

		Standard Cuts			Loose cuts		
m_t (GeV/c^2)		$e+$jets	$\mu+$jets	All $l+$jets	$e+$jets	$\mu+$jets	All $l+$jets
	$\varepsilon \times \mathcal{B}(\%)$	0.50 ± 0.10	0.33 ± 0.08		1.7 ± 0.3	1.07 ± 0.32	
140	$\langle N \rangle$	4.05 ± 0.94	2.47 ± 0.68	6.52 ± 1.16	16.95 ± 4.04	8.00 ± 2.57	24.95 ± 4.79
	$\varepsilon \times \mathcal{B}(\%)$	0.80 ± 0.10	0.57 ± 0.13		1.8 ± 0.3	1.23 ± 0.31	
160	$\langle N \rangle$	3.13 ± 0.54	2.04 ± 0.53	5.17 ± 0.76	7.08 ± 1.69	4.43 ± 1.22	11.51 ± 2.08
	$\varepsilon \times \mathcal{B}(\%)$	1.20 ± 0.30	0.76 ± 0.17		1.5 ± 0.3	1.35 ± 0.32	
180	$\langle N \rangle$	2.42 ± 0.67	1.41 ± 0.36	3.83 ± 0.76	4.30 ± 1.03	2.51 ± 0.67	6.81 ± 1.23
	$\varepsilon \times \mathcal{B}(\%)$	1.70 ± 0.20	0.96 ± 0.21		2.5 ± 0.4	1.44 ± 0.32	
200	$\langle N \rangle$	1.84 ± 0.31	0.95 ± 0.24	2.79 ± 0.39	2.80 ± 0.67	1.44 ± 0.37	4.24 ± 0.77
Backgnd: QCD		0.85 ± 0.32	0.61 ± 0.26	1.46 ± 0.41	7.52 ± 2.90	4.62 ± 1.78	12.14 ± 3.40
$W+$jets		0.37 ± 0.17	0.10 ± 0.11	0.47 ± 0.20	2.33 ± 0.81	1.25 ± 0.92	3.58 ± 1.23
Total		1.22 ± 0.42	0.71 ± 0.28	1.93 ± 0.50	9.85 ± 3.43	5.87 ± 2.00	15.73 ± 3.97
$\int \mathcal{L}\,dt$ (pb^{-1})		47.9 ± 5.7	44.2 ± 5.3		47.9 ± 5.7	44.2 ± 5.3	
Data		5	3	8	13	10	23

Table 5.4: Summary of results from $l+$jets channels, showing the efficiency $\times$ branching ratio ($\varepsilon \times \mathcal{B}$) and the expected number of top events ($\langle N \rangle$). The central value of the cross section of [17] is used. Also given are the expected background, integrated luminosity, and the number of observed events in each channel. From [1, 132, 133].

processes

$$b \to c\mu\nu, \quad b \to c \to s\mu\nu. \tag{5.11}$$

The probability that a $t\bar{t}$ event contains at least one such muon is about 44%. These muons can be detected at DØ with an efficiency of about 45%, so the probability for observing a tag muon in a $t\bar{t}$ event is about 20%. By contrast, the fraction of $W + \geq 3$ jet events with a muon passing the tagging cuts is expected to be about 2% [25]. Thus, the tagging requirement offers about an order of magnitude background reduction.

Due to the effectiveness of the tagging cut, the kinematic cuts are relaxed somewhat relative to the untagged analysis. The minimum number of jets is reduced from four to three (although the jet E_T cut is increased from 15 GeV

	$e + \text{jets}/\mu$	$\mu + \text{jets}/\mu$
Trigger	Same as untagged channels (Table 5.3)	
Lepton	$E_T(e) > 20\,\text{GeV}$ $\lvert\eta^{\text{det}}(e)\rvert < 2$	$p_T(\mu) > 15\,\text{GeV}/c$ $\lvert\eta(\mu)\rvert < 1.7^a$
Missing E_T	$\not\!\!E_T^{\text{cal}} > 20\,\text{GeV}$ $\not\!\!E_T^{\text{cal}} > 35\,\text{GeV}$ if $\Delta\phi(\not\!\!E_T^{\text{cal}}, \mu) < 25°$	$\not\!\!E_T^{\text{cal}} > 20\,\text{GeV}$ $\not\!\!E_T > 20\,\text{GeV}$ $\Delta\phi(\not\!\!E_T, \mu) < 170°{}^b$ $\lvert\Delta\phi(\not\!\!E_T, \mu) - 90°\rvert/90°$ $< \not\!\!E_T/(45\,\text{GeV})^b$
$Z \to \mu\mu$ fit	—	$\chi^2 < 0.01$
Tag μ	$p_T > 4\,\text{GeV}$ and $\Delta R(\text{tag } \mu, j) < 0.5$	
Jets	3 jets with $E_T > 20\,\text{GeV}$ and $\lvert\eta\rvert < 2$	
Loose cuts	—	
Standard cuts	$H_T > 140\,\text{GeV}$	

[a] Restricted to CF for run 1B.

[b] The highest-p_T muon is used for these two cuts.

Table 5.5: Summary of kinematic selection cuts for the tagged lepton + jets channels. [1]

to 20 GeV), the H_T cut is significantly reduced, and the aplanarity cut is removed. A summary of the selection cuts for the tagging analysis is given in Table 5.5. A total of 6 events survive, for both the loose and tight cuts.

The motivation for the complicated $\not\!\!E_T$ cuts in these selections is to obtain good rejection of the QCD multijet background. Figures 5.5 and 5.6 show the results of the $\not\!\!E_T$ selections for samples of simulated signal and background events for the $e + \text{jets}/\mu$ and $\mu + \text{jets}/\mu$ channels, respectively.

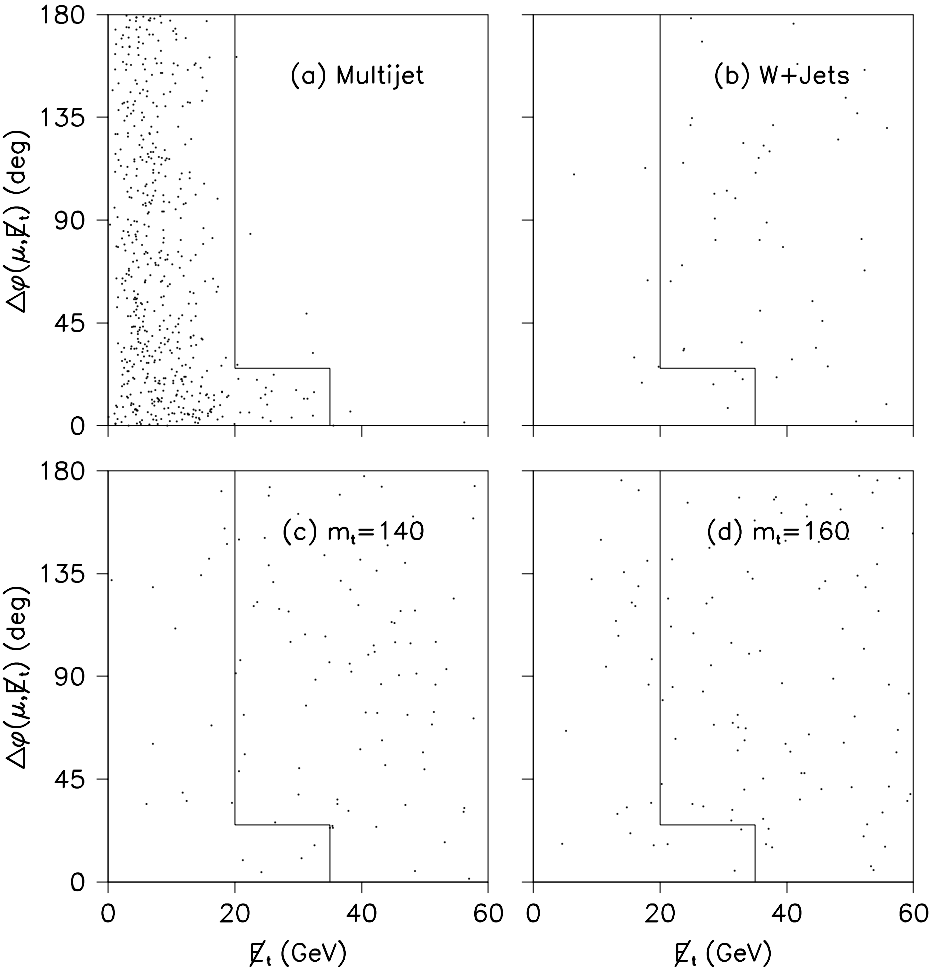

Figure 5.5: Distributions of $\Delta\phi(\not\!\!E_T^{\rm cal}, \mu)$ vs. $\not\!\!E_T^{\rm cal}$ for $e+$jets/μ for (a) Multijet background sample; (b) W + jets sample (VECBOS Monte Carlo); (c) Top Monte Carlo (ISAJET, $m_t = 140\,\mathrm{GeV}/c^2$); (d) Top Monte Carlo (ISAJET, $m_t = 160\,\mathrm{GeV}/c^2$). The contour shows the cut values. [25]

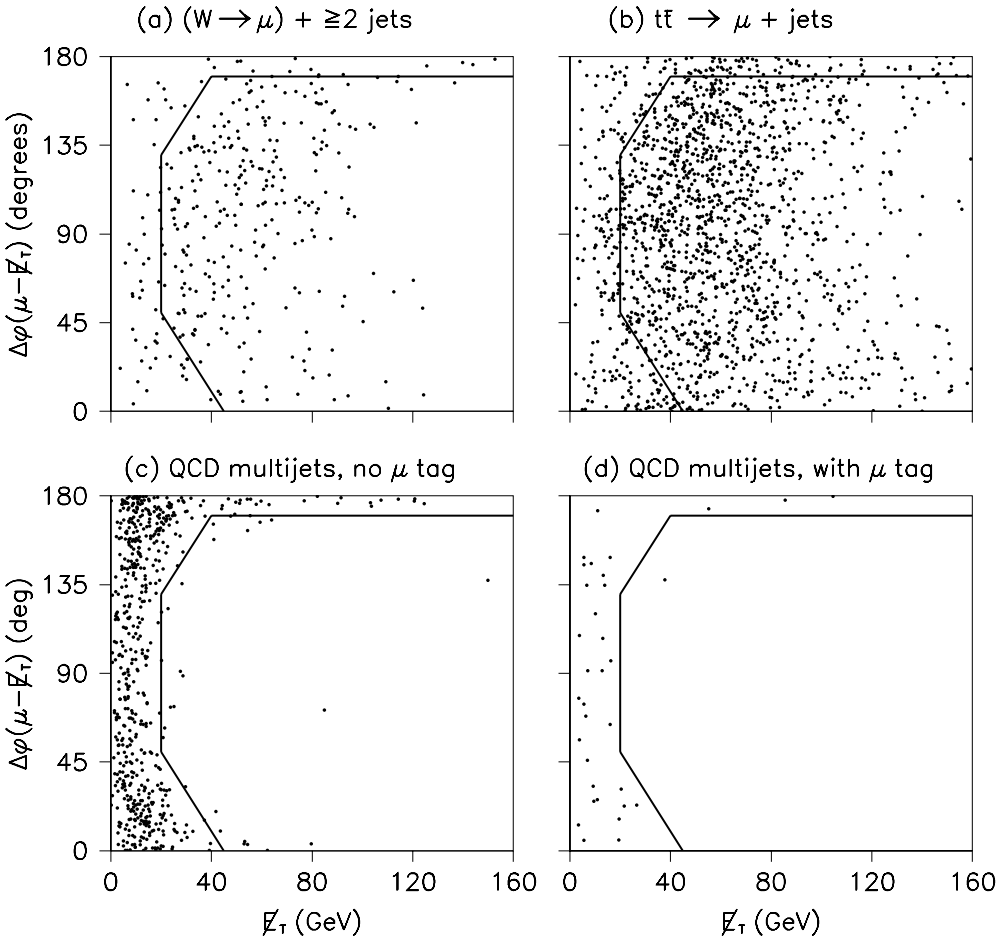

Figure 5.6: The correlation between $\not\!\!E_T$ and the azimuthal angle between $\not\!\!E_T$ and the highest-p_T muon for μ + jets/μ events for (a) VECBOS W + jets Monte Carlo ($80\,\mathrm{pb}^{-1}$); (b) ISAJET $t\bar{t}$ Monte Carlo ($m_t = 160\,\mathrm{GeV}/c^2$, $3240\,\mathrm{pb}^{-1}$); (c) Events with a nonisolated high-p_T muon with no μ-tag; (d) Events with a nonisolated high-p_T muon with an additional tagging muon. [25]

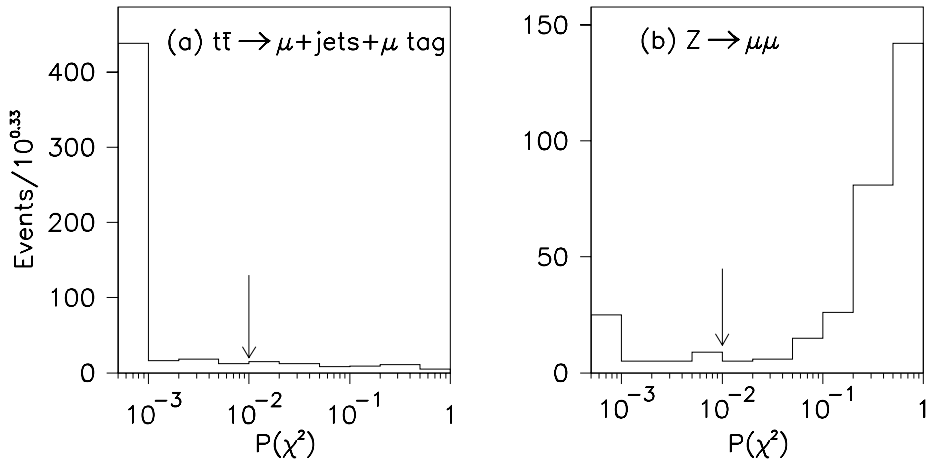

Figure 5.7: The χ^2 probability distribution for (a) Monte Carlo $t\bar{t} \rightarrow \mu$+jets+μ−tag events ($m_t = 160\ \mathrm{GeV}/c^2$) and (b) Monte Carlo ($Z \rightarrow \mu\mu$)+3 jet events. [25]

Like the dimuon channel, the μ + jets/μ channel rejects $Z \rightarrow \mu\mu$ background by performing a kinematic fit to that hypothesis. Events which are found to be compatible with the Z hypothesis are rejected. Figure 5.7 shows the distribution of the fit χ^2 for simulated signal and background samples.

The major ingredient going into the background calculation is the probability of tagging a jet in the background samples. Figure 5.8 shows the tagging rate per event as measured in bad electron, QCD 5-jet, and VECBOS W + jets samples. (A 'bad' electron is a highly-electromagnetic cluster which satisfies the electron trigger requirements, but fails the offline selection cuts.) Note that the per-event tagging rate is proportional to the number of jets in the event, and that the tagging rates for the three samples are consistent with each other. This supports the hypothesis that it is sensible to measure a per-jet tagging rate for these samples. The resulting tagging rate per jet is shown in

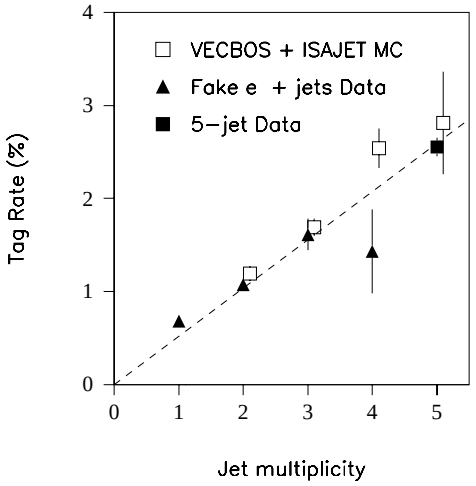

Figure 5.8: Fraction of events containing a muon as a function of jet multiplicity. [25]

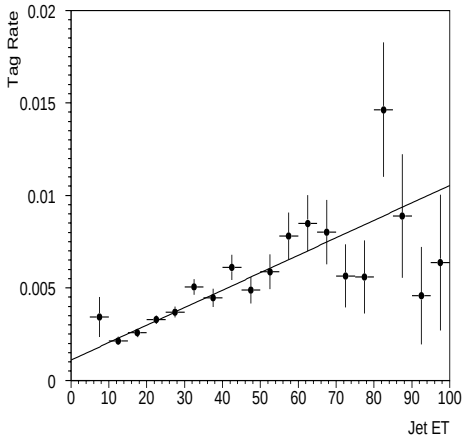

Figure 5.9: Tagging probability per jet as a function of jet E_T in fake $e+$ jets events. [130]

Figure 5.9 as a function of jet E_T.

The two major backgrounds to the e + jets/μ channel are, as in the untagged case, W + jets and QCD multijets. The W + jets background is estimated as follows. First, a sample is selected consisting of all $e + \geq 3$ jet events which satisfy the event selection, but without regard to the presence of a tag. Next, the amount of non-W background in that sample is estimated. This is done by looking at the ratio of the number of fake electrons which pass the electron ID cuts to the number of electromagnetic clusters which pass the loose reconstruction cuts but which fail the final electron ID cuts ('bad' electrons; see Section 4.3). This ratio is measured using single-electron triggers to be about 3% [134]. This is then multiplied by the number of 3 jet events with a bad electron to obtain the fake electron contribution to the $e + \geq 3$ jet sample. The untagged e + jet sample is then used to estimate the tagged W + jet

background by taking each jet in the sample, multiplying it by the tagging probability appropriate to the jet E_T, and summing up the resulting probabilities. This is then scaled down by the estimated fraction of real W + jet events present in the sample. The QCD multijet background is estimated separately by counting the number of 3 jet events which have a μ tag but which also have a bad electron candidate and multiplying this by the ratio mentioned previously.

For the μ + jets/μ channel, there are three major backgrounds. The W + jets and QCD multijet backgrounds are present as before, but there is an additional background from $Z \rightarrow \mu\mu$. The amount of this Z background expected to pass the χ^2 cut and all other kinematic cuts is estimated using VECBOS Monte Carlo. The W + jet background is estimated in a manner similar to that used for the e + jets/μ channel: a sample of events satisfying the kinematic cuts is selected, but without imposing the μ tag requirement. The transverse energies of the jets in this sample are then folded together with the measured jet tagging probabilities to obtain the estimate for the background after the tagging requirement. An additional correction is made at the end to take into account the fact that since the tagged events have two muons, they are more likely to pass the muon triggers than events with only one muon. Finally, the QCD multijet background was studied with a sample of events which satisfy all selection cuts except for the tagging requirement and for which the high-p_T muon was nonisolated (see Figure 5.6). The QCD background is then the product of the following three factors:

- The number of events in the nonisolated sample which pass the analysis cuts (with the exception of the μ-tag requirement), normalized to match the luminosity of the data sample.

- The probability for an event to have a soft muon tag. This is determined by convoluting the jet E_T spectrum in the nonisolated sample with the measured tag probability function. (The jet in which the nonisolated muon is embedded is excluded from this calculation.)

- The probability for a μ + jets system to fluctuate to give an isolated muon. This is measured using the ratio of the number of n-jet isolated muon events to the number of $(n+1)$-jet nonisolated muon events.

Figure 5.10 summarizes the observed data and the calculated background. Note that there is an excess of the data over the expected background at the higher multiplicity points, while for the 1 jet point, the calculated background agrees well with the observed data. (Note that the jet multiplicity is inclusive, so any deficit which is present at one multiplicity will also be present at all lower multiplicities). The final results for the tagged channels are summarized for both the standard and loose cuts in Table 5.6.

5.4 Significance and Cross Section

In summary, there are 17 events which pass the standard cuts, with an expected background of 3.79 ± 0.55 events. One can define the 'significance' of this result as the probability of seeing at least 17 events assuming that only the

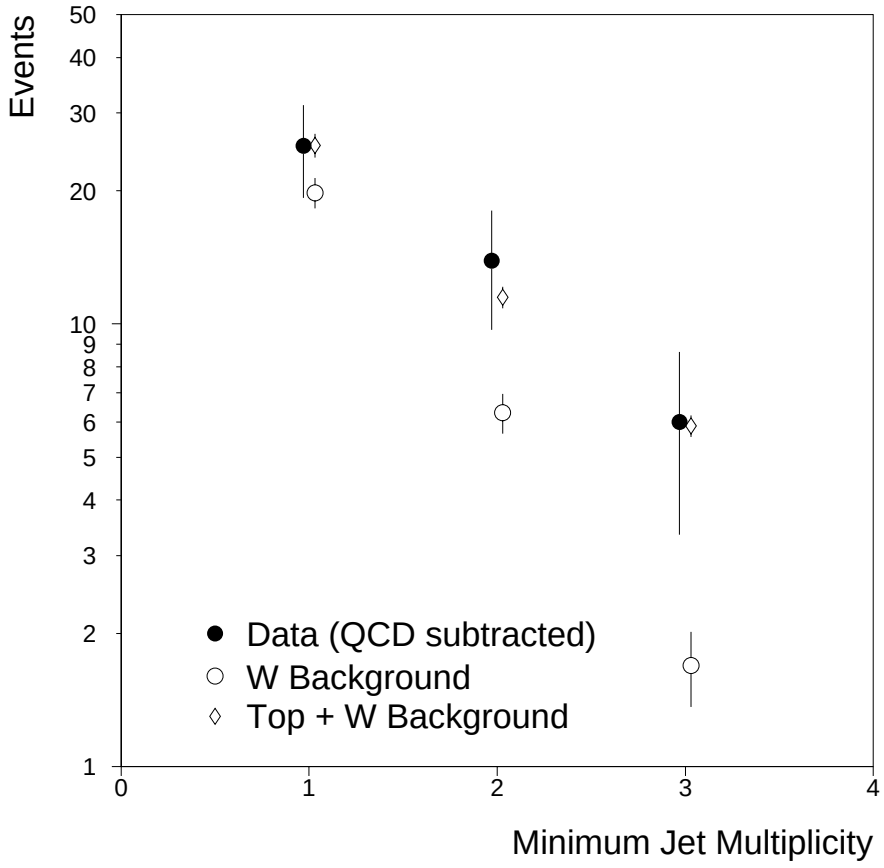

Figure 5.10: Number of events in the μ-tag channels as a function of inclusive jet multiplicity. Plotted for the observed data (with the calculated QCD background subtracted), calculated W + jets background, and the sum of the background and the expectation for a 200 GeV/c^2 top. [130]

		Standard Cuts			Loose cuts		
m_t (GeV/c^2)		$e+\text{jets}/\mu$	$\mu+\text{jets}/\mu$	All tagged	$e+\text{jets}/\mu$	$\mu+\text{jets}/\mu$	All tagged
	$\varepsilon \times \mathcal{B}(\%)$	0.36 ± 0.07	0.20 ± 0.05		0.58 ± 0.10	0.26 ± 0.06	
140	$\langle N \rangle$	2.93 ± 0.68	1.48 ± 0.42	4.41 ± 0.80	4.70 ± 0.98	1.92 ± 0.51	6.62 ± 1.10
	$\varepsilon \times \mathcal{B}(\%)$	0.50 ± 0.08	0.25 ± 0.06		0.75 ± 0.13	0.31 ± 0.07	
160	$\langle N \rangle$	1.95 ± 0.39	0.92 ± 0.24	2.87 ± 0.46	2.93 ± 0.62	1.10 ± 0.28	4.03 ± 0.68
	$\varepsilon \times \mathcal{B}(\%)$	0.56 ± 0.09	0.35 ± 0.08		0.69 ± 0.11	0.39 ± 0.09	
180	$\langle N \rangle$	1.14 ± 0.22	0.64 ± 0.16	1.78 ± 0.27	1.39 ± 0.28	0.73 ± 0.18	2.12 ± 0.33
	$\varepsilon \times \mathcal{B}(\%)$	0.74 ± 0.11	0.41 ± 0.08		0.90 ± 0.14	0.43 ± 0.09	
200	$\langle N \rangle$	0.81 ± 0.16	0.41 ± 0.10	1.22 ± 0.19	0.97 ± 0.19	0.43 ± 0.11	1.40 ± 0.22
Backgnd: QCD		0.50 ± 0.08	0.08 ± 0.03	0.58 ± 0.09	0.87 ± 0.16	0.12 ± 0.03	0.99 ± 0.16
$W+\text{jets}$		0.35 ± 0.12	0.28 ± 0.08	0.63 ± 0.14	0.60 ± 0.13	0.57 ± 0.14	1.17 ± 0.19
Total		0.85 ± 0.14	0.36 ± 0.08	1.21 ± 0.16	1.47 ± 0.21	0.69 ± 0.14	2.16 ± 0.25
$\int \mathcal{L}\, dt$ (pb^{-1})		47.9 ± 5.7	44.2 ± 5.3		47.9 ± 5.7	44.2 ± 5.3	
Data		3	3	6	3	3	6

Table 5.6: Summary of results from $l+\text{jets}/\mu$ channels, showing the efficiency $\times$ branching ratio ($\varepsilon \times \mathcal{B}$) and the expected number of top events ($\langle N \rangle$). The central value of the cross section of [17] is used. Also given are the expected background, integrated luminosity, and the number of observed events in each channel. From [1, 134, 135].

background is present. If the probability distribution for the expected number of background events is taken to be a gaussian with mean B and width σ_B, then the significance can be written

$$\text{significance} = \sum_{n=N_{\text{obs}}}^{\infty} \int_0^{\infty} q(N_{\text{obs}}, \mu)\, g(\mu; B, \sigma_B)\, d\mu, \tag{5.12}$$

where N_{obs} is the number of observed events, $q(N, \mu)$ is the Poisson distribution

$$q(N, \mu) = e^{-\mu} \mu^N / N!, \tag{5.13}$$

and $g(x; B, \sigma_B)$ is the gaussian distribution

$$g(x; B, \sigma_B) = \frac{1}{\sqrt{2\pi}\sigma_B} e^{-(x-B)^2/2\sigma_B^2}. \tag{5.14}$$

The result of this calculation for the tight cuts is about 2×10^{-6} (equivalent to 4.6 standard errors for a gaussian distribution).

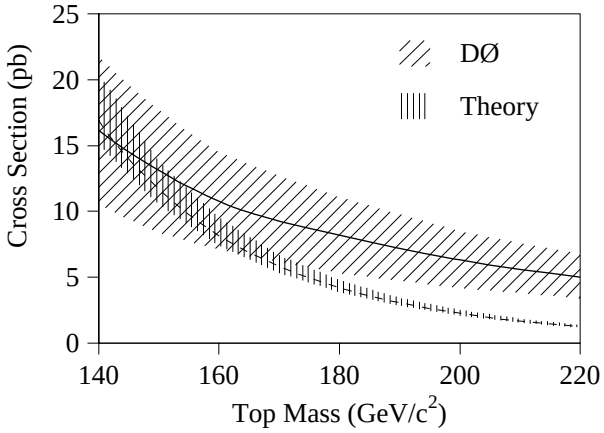

Figure 5.11: DØ measured $t\bar{t}$ production cross section (solid line with 1σ error band) as a function of assumed top mass. Also shown is the theoretical cross section curve of [17] (dashed line).

Given the numbers for efficiency $\times$ branching ratio ($\varepsilon\mathcal{B}$), integrated luminosity ($L = \int \mathcal{L}\, dt$), and background B for each channel, one can compute the total cross section for top production using

$$\sigma_{t\bar{t}} = \frac{\sum_{i\in\text{channels}}(N_i - B_i)}{\sum_{i\in\text{channels}} \varepsilon_i \mathcal{B}_i L_i}. \tag{5.15}$$

Since the signal efficiency depends on the top mass, the measured cross section will as well. The result is plotted as a function of top mass in Figure 5.11, along with the theoretical expectation. At $m_t = 200\,\text{GeV}/c^2$, the cross section is $6.3 \pm 2.2\,\text{pb}$.

Table 5.7 summarizes the significance and cross section for all of the different analyses.

For mass fitting, the dilepton channels are not (yet) used. In that case, there are, with the standard cuts, 14 observed events with an expected background of 3.14 ± 0.52. This yields a significance of 2×10^{-5}. With the loose

Channel	Cuts	Nevts	Background	Signif.	$\sigma_{t\bar{t}}$ (pb)
Dilepton	Loose	4	2.66 ± 0.40	0.28	4.4 ± 6.8
	Standard	3	0.65 ± 0.15	0.03	7.5 ± 5.7
$l +$ jets	Loose	23	15.7 ± 3.1	0.09	4.0 ± 3.2
	Standard	8	1.9 ± 0.5	0.002	4.9 ± 2.5
$l +$ jets/μ	Loose	6	2.2 ± 0.3	0.03	6.3 ± 4.2
	Standard	6	1.2 ± 0.2	0.002	8.9 ± 4.8
Total	Loose	33	20.6 ± 3.2	0.023	4.5 ± 2.5
	Standard	17	3.8 ± 0.6	2×10^{-6}	6.3 ± 2.2

Table 5.7: Summary of number of observed events, expected background, significance, and top cross section (for $m_t = 200\,\mathrm{GeV}/c^2$) for each analysis.

cuts, there are a total of 29 observed events, with an expected background of 17.9 ± 4.0, giving a significance of 4×10^{-2}. These two samples form the input to the mass analysis in Chapter 8.

Chapter 6

Mass Fitting

The goal of this chapter is to develop the tools needed to reconstruct the top mass in a $t\bar{t} \rightarrow l + \text{jets}$ decay. Unfortunately, there is not a unique prescription for accomplishing this; further, there are relatively large, mass-dependent biases which are difficult to eliminate. Therefore, the quantity which will be constructed is best thought of not as *the* top mass for an event, but rather as merely an observable which is correlated with the top mass. In order to distinguish the two, the result of this procedure will be referred to as the 'fitted' mass. The final stage of the analysis (discussed in Chapter 7) will go from fitted masses to the final top mass.

Further information on mass fitting at DØ may be found in [98, 136].

6.1 Introduction

Here is a summary of the steps used in constructing the fitted mass. The points raised here will be expanded on later in this chapter.

- The inputs to the analysis are the kinematic parameters of the lepton, jets, and missing E_T in the event.

- To reduce the confusion caused when multiple partons get merged together in a single jet, the relatively narrow cone size of $R = 0.3$ is used. As mentioned in Section 4.5.2, the standard MPF jet corrections do not fully account for out-of-cone radiation, so an additional correction needs to be made (see Section 6.2.1). Jets with a muon tag receive an extra correction to account for the energy carried off by the muon and neutrino.

- Jets are required to have $|\eta| < 2.5$ and $E_T > 15\,\text{GeV}$ (after all corrections). There must be at least four such jets in an event in order to start the fit; events with fewer jets are discarded. Only the four jets with the highest E_T are used in the fit.

- If the correspondence between jets and partons were known, one could simply sum them up in the appropriate combinations to obtain the masses of the two tops. But in general, this is not known. Instead, all twelve possible assignments of the four jets are tried and only the 'best' solutions are retained.

- But how to quantify 'best'? Recall from Section 2.4 that the system is overconstrained by two equations. One can thus define a χ^2 which tells how far the kinematic parameters of the event must be pulled from their measured values in order for the constraints to be satisfied. The 'best' permutations are then the ones with the lowest χ^2 values. The results

of this kinematic fit can also be used to define an improved value for the top mass.

- Finally, the fitted mass of an event is defined as the χ^2-weighted average of (up to) the best three jet permutations which satisfy a loose χ^2 cut of $\chi^2 < 7$. If the event contains a tagged jet, a jet permutation is considered only if it assigns the tagged jet as a b jet.

6.2 Final Jet Corrections

6.2.1 Out-Of-Cone Correction

Following the MPF jet corrections (see Section 4.5.2), an additional correction is applied. The intent of this correction is to account for radiation outside of the jet cone; it attempts to make the jet energies match those of the original partons. It was derived by looking at an ISAJET $t\bar{t}$ Monte Carlo sample and comparing the energies of the MPF-corrected reconstructed jets with the energies of the partons which most nearly match those jets. (The ISAJET `ISAQ` partons were used, i.e., after QCD evolution but before fragmentation.) The correction is a simple linear form [98]:

$$E_T(\text{corr}) = 5.556 + 1.087 E_T(\text{MPF}) \tag{6.1}$$

(for $R = 0.3$ cone jets).

Figure 6.1 shows the reconstructed mass of the hadronic W in ISAJET top events, after the full detector simulation and reconstruction, both before and

after the out-of-cone corrections. (ISAJET parton-ID information was used to find the jets resulting from the decay of the W. All such jets satisfying $E_T > 15\,\mathrm{GeV}$ and $|\eta| < 2$ are used, not just those in the highest four E_T jets.) It can readily be seen the the correction moves the distribution much closer to the true W mass.

As an additional check, one can look at $(Z \to ee) + 1$ jet events. In these events, the Z is well-measured, and should be balanced in the transverse plane by the single jet. Figure 6.2 shows the difference between the electron pair and jet transverse energies for data and Monte Carlo both before and after the out-of-cone correction. One sees that there is good agreement between the data and the Monte Carlo, and also that the corrections significantly reduce the transverse momentum imbalance between the electron pair and the jet.

The Z data are also used to estimate the systematic error on the jet energy scale, including the out-of-cone corrections. A clean sample of $(Z \to ee)+1$ jet events is selected. Events with more than one ($R = 0.3$ cone) jet are excluded from the sample (with no E_T cut on the jet). Define the quantity f as

$$f = \frac{E_T(\mathrm{jet}) - E_T(ee)}{E_T(ee)}. \tag{6.2}$$

An upper limit on the energy scale error is then estimated by computing the absolute value of the mean of f and its error, and adding these two numbers in quadrature. The results are as follows [137].

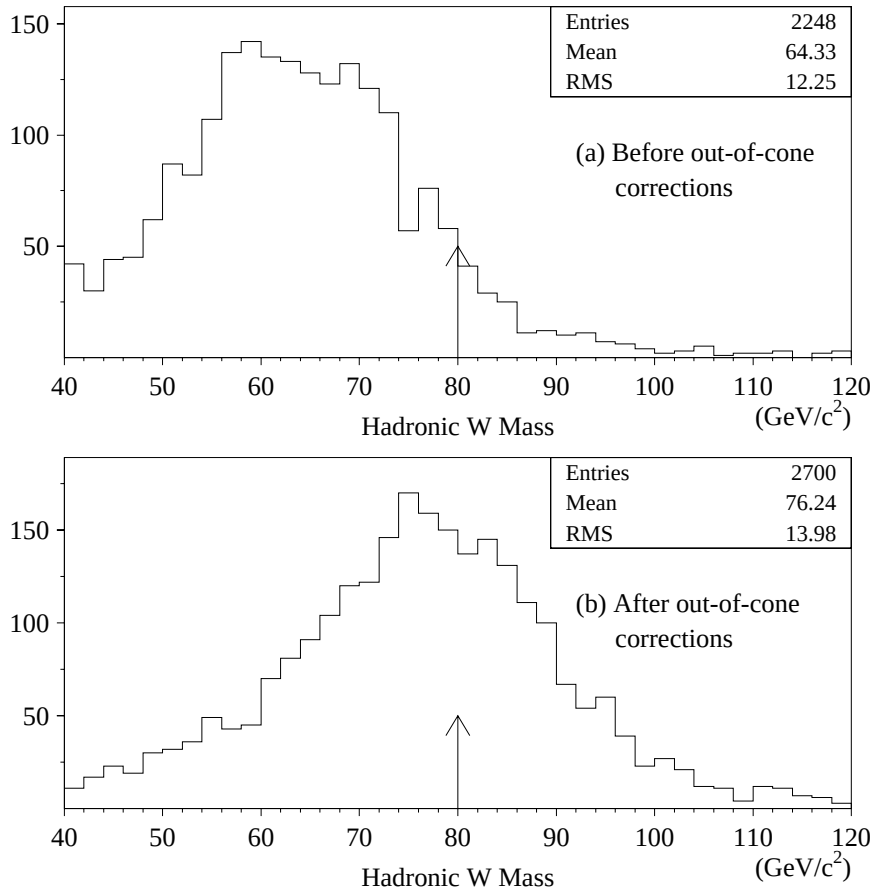

Figure 6.1: Reconstructed hadronic W mass, (a) before and (b) after out-of-cone corrections (using $R = 0.3$ cone jets). From ISAJET 180 GeV/c^2 $t\bar{t}$ Monte Carlo, using all jets which were actually from W decay and with $E_T > 15$ GeV and $|\eta| < 2$.

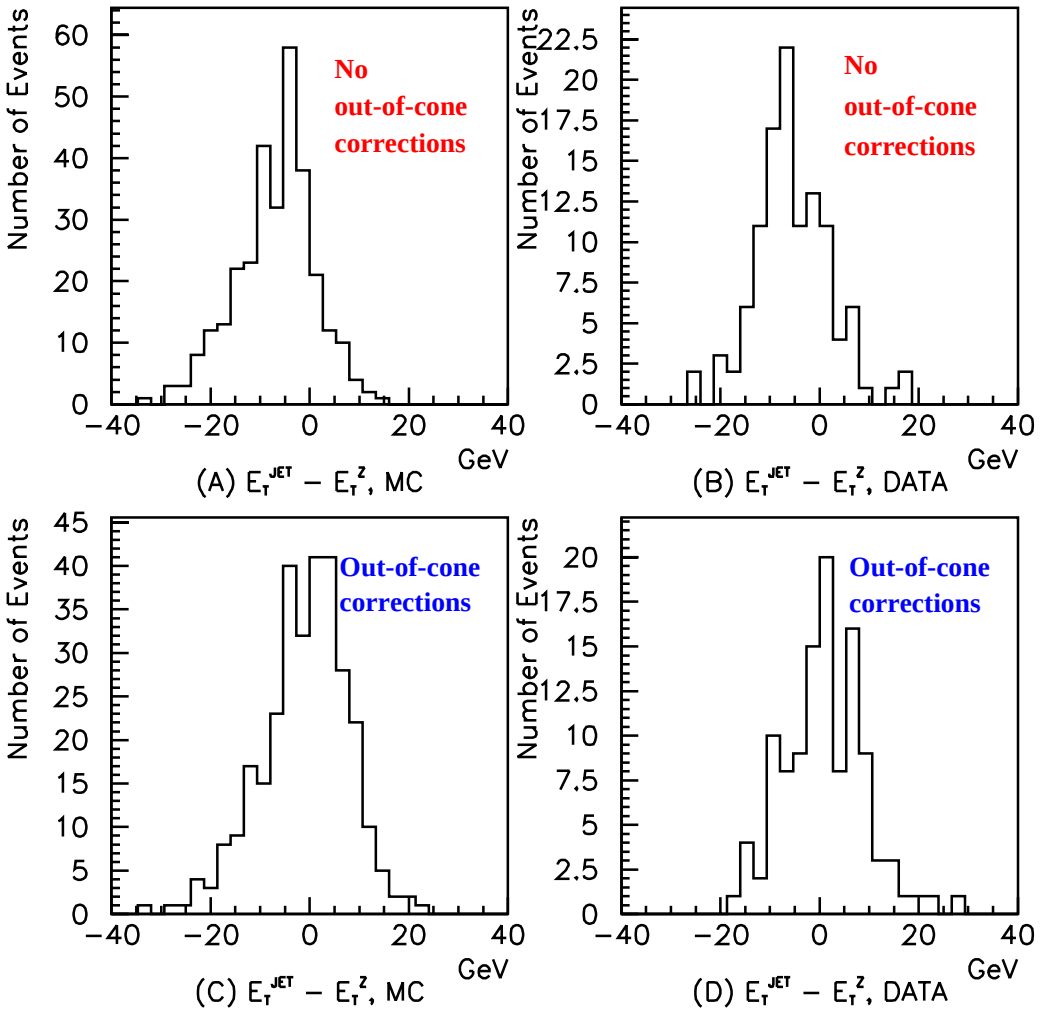

Figure 6.2: Difference between the transverse momenta of the electron pair and the jet in $(Z \rightarrow ee) + 1$ jet events. For Monte Carlo and data, with and without out-of-cone corrections. [98, p. 61]

$E_T(Z)$	$\lvert\langle f\rangle\rvert$	Error on $\lvert\langle f\rangle\rvert$	Estimate of jet scale error
10–20 GeV	0.3	0.06	30%
20–30 GeV	0.004	0.05	5.0%
30–40 GeV	0.06	0.05	7.8%
40–50 GeV	0.023	0.08	8.6%
$>$ 50 GeV	0.018	0.07	7.0%

The lowest energy band is somewhat anomalous, but most jets in top events have higher transverse energies. Thus, an upper limit on the error on jet energy scale is estimated to be about 10%.

For further details about the out-of-cone corrections, see [98, pp. 55–61].

6.2.2 Tagged Jet Correction

For jets which have a muon tag, there is one further correction which should be made. These jets presumably came from a b which underwent a semileptonic decay into a muon and a neutrino. To get back the momentum of the original b, the momenta of both the muon and the neutrino should be added into the jet. Of course, the momentum of the neutrino is not measured, but as Figure 6.3 shows, adding twice the muon momentum to the jet is a reasonable approximation.

6.3 Fitting Algorithm

Now that the jet corrections are sorted out, the kinematic fitting algorithm itself can be examined.

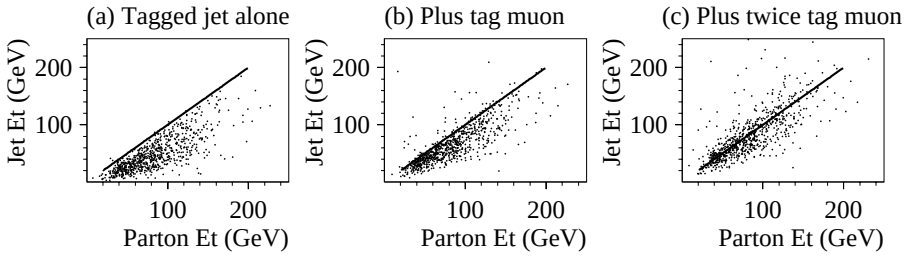

Figure 6.3: Comparison of transverse energies of reconstructed tagged jets (vertical axis) to matched partons (horizontal axis). (a) No tagging correction. (b) Tag muon momentum added to the tagged jet. (c) Twice the tag muon momentum added to the tagged jet. From ISAJET 180 GeV/c^2 $t\bar{t}$ Monte Carlo.

6.3.1 Fit Variables

The first thing to do is to define the variables which will be used in the fit. It is convenient to choose them so that they are as uncorrelated as possible.

For the lepton and each jet, there are three measured variables: its energy (or momentum), and its direction (which is a two dimensional quantity). The variables which will be used to describe an electron or jet are its momentum p, azimuthal angle ϕ, and pseudorapidity η. Muons are described in a similar manner, except that instead of p, the inverse momentum $k = 1/p$ is used, since the errors are more nearly gaussian in that variable. The masses of the jets are fixed to zero, except for the jets which are assigned as b-jets, which are given masses of 5 GeV/c^2. (This is done by taking the momentum components of the jet as the measurement, and changing the energy component of the four-vector to set the desired mass. Since these masses are small compared to the typical jet energies (> 15 GeV), the exact prescription used doesn't make

much difference for the final result. For another method of dealing with jet masses, see [98].)

To represent the transverse momentum of the neutrino, one could use the x and y components of the missing E_T. However, that is not ideal, since the missing E_T is highly correlated with all the other energy measurements: a fluctuation in jet measurement will show up as an exactly corresponding fluctuation in the missing E_T. A better quantity is k_T, which is defined as

$$\vec{k_T} = \vec{E}_T(\text{lep}) + \vec{\not{E}}_T + \sum_{i \in \text{jets}} \vec{E}_T(\text{jet}_i). \tag{6.3}$$

The sum over jets here includes only jets assumed to be part of the final state. This variable can be interpreted as the transverse momentum of the $t\bar{t}$ system (it is the negation of the 'baby jet' of [98]). The x and y components of $\vec{k_T}$ are used as fit variables.

There is one remaining variable, which is taken to be the z-component of the neutrino momentum. This gives a total of eighteen variables (for four jets).

Each measured variable still needs to have an error attached to it; this will be discussed later. All the errors are assumed to be uncorrelated (i.e., a diagonal error matrix).

6.3.2 Jet Permutations

As mentioned above, a fit is tried for each possible jet permutation. For four jets, there are twelve such permutations: there are twenty-four permutations of four objects, but the two jets which are assigned to the hadronic W

can be interchanged without changing the solution, so only twelve of them are distinct.

Before attempting a fit, loose cuts are made to reject permutations which are likely to yield poor fits. The hadronic W mass, before fitting, is required to be within the range $40\,\mathrm{GeV}/c^2 < m_W < 120\,\mathrm{GeV}/c^2$. In addition, permutations which have a large value for the smallest solution for the z-component of the neutrino momentum (see Section 6.3.3 and Appendix B) are rejected.

For future reference, consider the problem of jet permutations for more than four jets. There are two possibilities for dealing with an extra jet: it can be considered initial state radiation (ISR) and dropped from the problem, or it can be considered to have been radiated from one of the final state partons (final-state radiation, or FSR), in which case it should be merged with the appropriate jet. Due to interference terms in the production and decay matrix elements, this separation into ISR and FSR is only an approximation, but it works reasonably well for radiation which is large enough to be seen as a separate jet [138].

In combinatorial terms, the problem can be stated as follows. Find all distinct ways of tagging N objects with the labels

- b_l, for the leptonic-side b,
- b_h, for the hadronic-side b,
- w, for the decay products of the hadronic W, and
- i, for initial state radiation,

subject to the conditions that there be at least one b_l, one b_h, and two w's. The number of such permutations grows extremely rapidly with N; the first few values are as follows:

$$\begin{aligned} N &= 4 \quad 12 \\ N &= 5 \quad 140 \\ N &= 6 \quad 1020 \\ N &= 7 \quad 5992 \end{aligned}$$

This combinatorial explosion is one of the main motivations for restricting the fit to just the top four jets.

6.3.3 Constrained Fit

The constrained fit itself can now be described. Form the fit variables into a one-dimensional vector $\mathbf{x}$, and define a χ^2 by

$$\chi^2 = (\mathbf{x} - \mathbf{x}^m)^T \mathbf{G} (\mathbf{x} - \mathbf{x}^m), \tag{6.4}$$

where $\mathbf{x}$ is the vector of fit variables, $\mathbf{x}^m$ is the vector of the measured values of those variables, and $\mathbf{G}$ is the inverse error matrix. This quantity should be minimized subject to the constraints

$$m_{W\mathrm{lep}} = 80.2\,\mathrm{GeV}/c^2 \tag{6.5}$$

$$m_{W\mathrm{had}} = 80.2\,\mathrm{GeV}/c^2 \tag{6.6}$$

$$m_t = m_{\bar{t}}. \tag{6.7}$$

Were these constraints linear in the fit variables, the problem would be exactly solvable using the method of Lagrange multipliers and simple linear algebra.

Unfortunately, that is not the case. Instead, an iterative technique is used. Starting at the point given by the measured values of the variables, The constraint equations are expanded in a power series around this point to linearize them, and the minimization is solved with these linearized constraints. This result defines the starting point for the next minimization step, and the iteration continues until the constraints are satisfied and the χ^2 stops changing. Further details on the fitting method may be found in Appendix A.

There is one additional wrinkle to this procedure. To start the above procedure, one must specify an initial value for *all* variables — not just the measured ones. In particular, a starting value is needed for the z-component of the neutrino momentum, p_z^ν. This is done by choosing it such that the two tops have equal mass. This yields the following quadratic equation for p_z^ν:

$$\left((p_z^c)^2 - (E^c)^2\right)(p_z^\nu)^2 + \alpha p_z^c p_z^\nu - (E^c p_T{}^\nu)^2 + \alpha^2/4 = 0, \tag{6.8}$$

where

$$\alpha = m_t^2 - m_c^2 + 2\vec{p_T}{}^\nu \cdot \vec{p_T}{}^c, \tag{6.9}$$

and the four-vector c is the sum of the four-vectors of the lepton and b-jet. If the solutions are complex, the real part is used. Otherwise, there will be two real solutions. Both are tried, and the fit with the lowest χ^2 is retained.

Note that since p_z^ν does not enter into the χ^2 (since its measurement error is effectively infinite), the only effect it can have on the final answer is to influence which local minimum the fit will find, if there happen to be more than one. Only in about 10–20% of top events do two distinct neutrino solutions yield different fit results.

6.3.4 Solution Averaging

The procedure so far gives a set of (χ^2, fitted mass) pairs, one for each jet permutation which was successfully fit. However, it would be most convenient to be able to represent the result of the fit by a single number. One way of doing this would be to pick the solution which has the best χ^2. This has the drawback, however, that this value can undergo large, discontinuous changes for small changes in the input parameters, as the χ^2 values change and different permutations come to have the smallest χ^2. On the other hand, the result should not depend strongly on fits which have a relatively poor χ^2. An additional observation is that while there is at most one 'correct' jet permutation in an event (and it rarely turns out to have the best χ^2), there are several additional 'semicorrect' permutations (where the jets are assigned to the proper tops) which contain nearly as much mass information as the correct permutation.

Therefore, a compromise is to take the weighted average of the best few solutions, where the weight is taken to be $e^{-\chi^2/2}$. For this analysis, a loose χ^2 cut is imposed ($\chi^2 < 7$), and the weighted average of the fitted masses for the three best solutions is used (or however many good solutions there are, if this number is less than three). If the event has a tagged jet, the only permutations used in the average are those in which the tagged jet was used as a b.

6.4 Parton-Level Tests

The first tests of the fitter will use parton-level results from the ISAJET Monte Carlo [120]. Here, 'parton-level' means the results from ISAJET following the QCD evolution phase, but before fragmentation (the '`ISAQ`' particles). No detector simulation is included. To (crudely) model the effects of the jet-finding algorithm, partons are grouped together inside of cones of radius $R = 0.45$. The resulting objects, called 'pjets' [139], are then treated as jets.

Unless otherwise stated, the data were smeared with gaussians with the widths given in the following table.

$\frac{\sigma E(e)}{E(e)} = \frac{0.18\sqrt{\text{GeV}}}{\sqrt{E}}$	$\frac{\sigma E(\text{jet})}{E(\text{jet})} = \frac{1.19\sqrt{\text{GeV}}}{\sqrt{E}}$
$\sigma\phi(e) = 0.0060$ rad	$\sigma\phi(\text{jet}) = 0.035$ rad
$\sigma\eta(e) = 0.0072$	$\sigma\eta(\text{jet}) = 0.040$
$\sigma k_{Tx} = 12$ GeV	$\sigma k_{Ty} = 12$ GeV

(6.10)

These same numbers are used for the object resolutions in the constrained fit.

The parton-level studies are restricted to the e + jets channel. The kinematic cuts applied are the same as used for the data analysis:

$$\begin{aligned} &E_T(e) > 20\,\text{GeV} \qquad |\eta(e)| < 2 \\ &E_T(\text{jet}) > 15\,\text{GeV} \qquad |\eta(\text{jet})| < 2 \\ &\not\!\!E_T > 25\,\text{GeV}. \end{aligned} \tag{6.11}$$

The four jets with the largest E_T within $|\eta| < 2.5$ are selected for use in the fit. For these parton-level studies, a χ^2 cut of $\chi^2 < 20$ is required of all fits.

To start with, examine Figure 6.4(a). This shows the result of fitting an unsmeared $t\bar{t}$ Monte Carlo sample with $m_t = 180\,\text{GeV}/c^2$. The unshaded

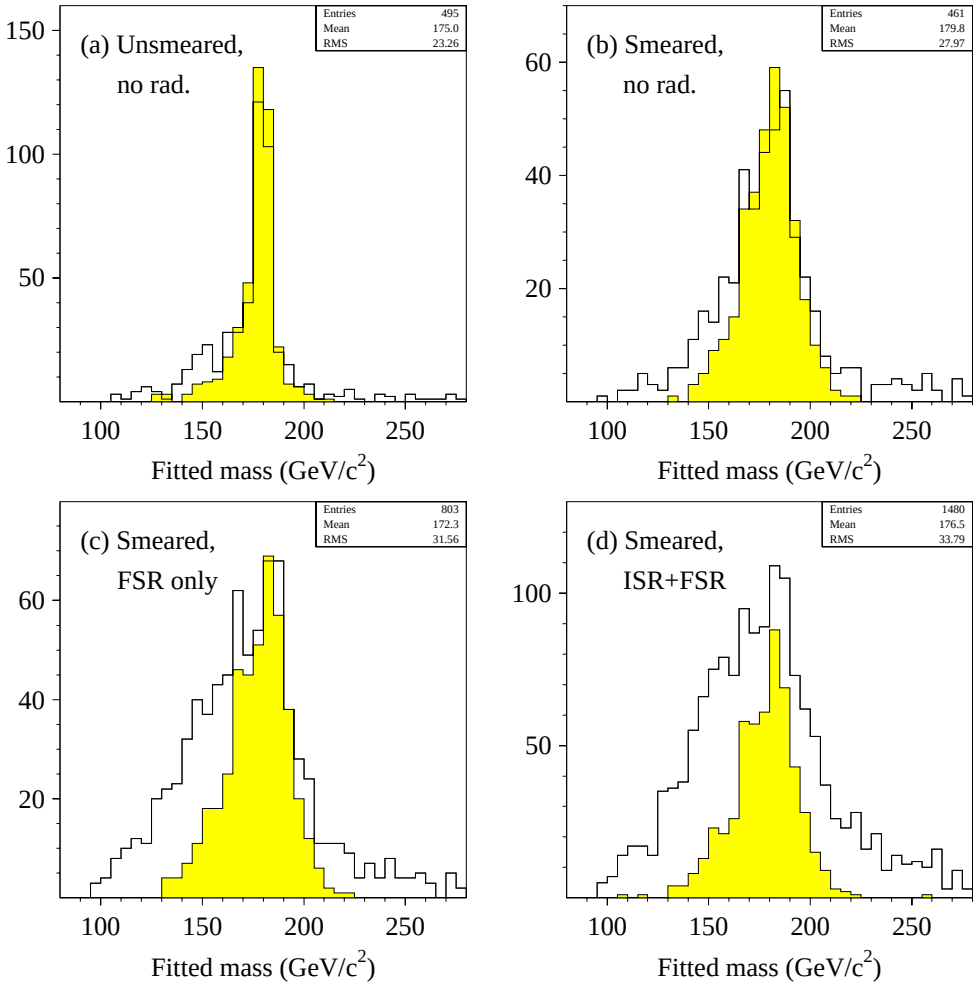

Figure 6.4: Results of fits with a $t\bar{t} \to e+$jets parton-level ISAJET Monte Carlo with $m_t = 180\,\mathrm{GeV}/c^2$, for different radiation conditions. The unshaded plot is the jet permutation with the smallest χ^2, and the shaded plot is the correct jet permutation. (a) Unsmeared, no radiation (see text). (b) Smeared, no radiation. (c) Smeared, FSR only. (d) Smeared, ISR and FSR. For all plots, the fit χ^2 was required to be less than 20. Statistics reflect the unshaded plot.

histogram is the jet permutation with the smallest χ^2 and the shaded histogram is the correct jet permutation. (Note that for any given event, the result from the correct permutation will usually be different than that from the best permutation. Thus, although the events appearing in the shaded plots are a subset of those in the unshaded plots, the shaded plots do not necessarily lie completely beneath the unshaded ones.) For this plot, the sample was further filtered to remove events containing significant radiation, either initial-state (ISR) or final state (FSR). More precisely, events were required to have exactly four jets which correspond to the four primary partons. The distributions peak in nearly the right place, but there is a small tail on the low end caused by residual radiation. Another factor which contributes to the width of these plots is the fact that the W masses are generated from a Breit-Wigner distribution, while the fit always forces the W masses to a constant value.

Continuing, Figure 6.4(b) shows what happens when smearing is turned on. In Figure 6.4(c), events are allowed to have significant final state radiation, but events with significant initial state radiation are still filtered out. This increases the low-end tail, since energy gets lost when one picks just the top four jets in an event with FSR. Finally, in Figure 6.4(d), there is no filtering out of radiation. This inclusion of ISR increases the tail at the high end, but also brings the means of the distributions back towards the input mass.

The χ^2 distribution for this sample (with full radiation) is shown in Figure 6.5, for the first three lowest χ^2 fits and the correct permutation. The distribution of correct fits peaks at zero, as it should, while those from the

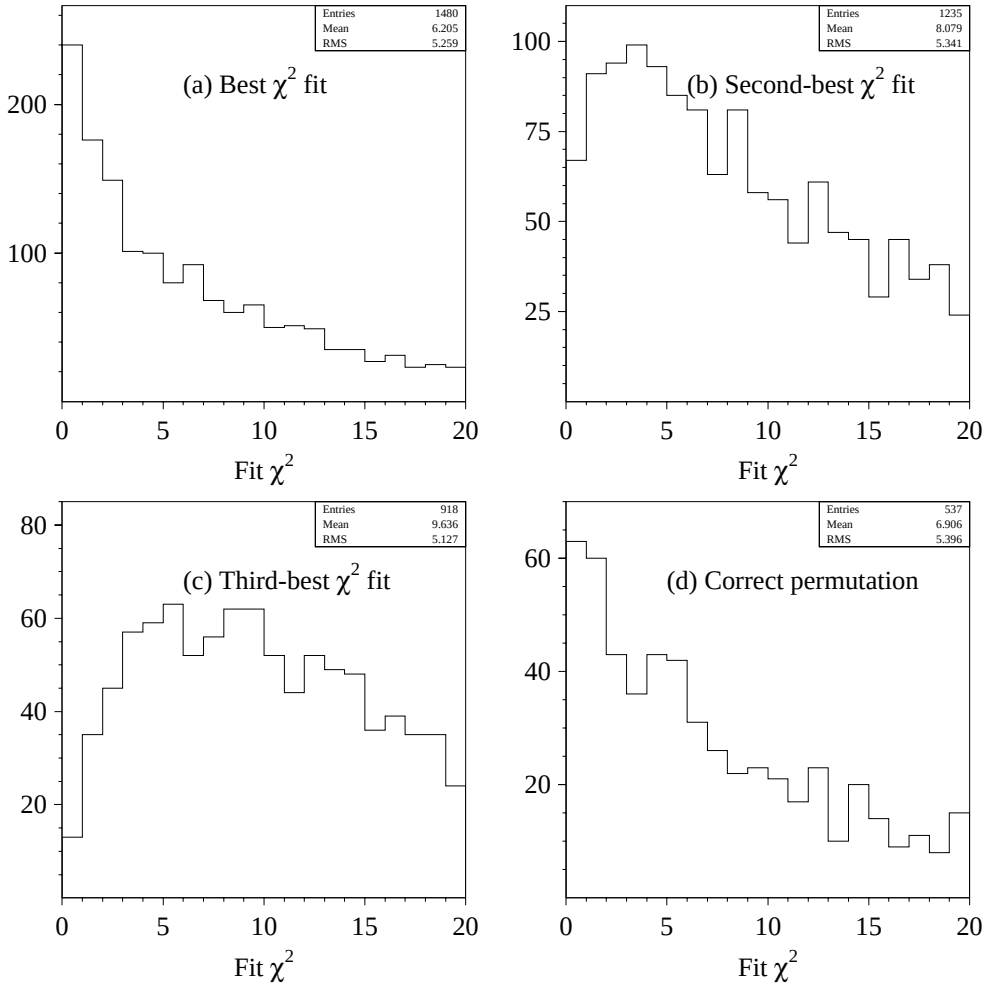

Figure 6.5: Fit χ^2 distributions for (a) best, (b) second-best, and (c) third-best χ^2 solutions and (d) the correct jet permutation. For smeared parton-level $t\bar{t} \to e +$ jets ISAJET Monte Carlo,with $m_t = 180\,\mathrm{GeV}/c^2$.

second- and third-best fits start to peak away from zero. Note, however, that even though the resolutions used for the fit are the same as were used for smearing the data, the mean of the χ^2 distribution for the correct permutations is rather larger than one would expect (2, for a 2C fit). This is again due to the effects of radiation and the non-zero widths of the W's. For fits using the full detector simulation, these effects will get absorbed into the object resolutions.

As mentioned in Section 6.3.4, what is actually used as the result from the fit is not the best χ^2 solution, but instead a weighted average of up to the three best solutions. (The weight used is the χ^2 probability $e^{-\chi^2/2}$.) Figure 6.6 shows the mass distributions that result from using the best, second-best, and third-best χ^2 permutations, as well as the weighted average of the three. As the χ^2 values go up, the resulting mass values tend to be biased low. The weighted average, however, is not biased, and is slightly narrower than the result of using the best χ^2 solution.

Figure 6.7 shows the effect of varying the amount of jet smearing. As the amount of smearing increases, the widths of the distributions increase, as expected, although the increase is rather gradual. Also note that as the smearing is increased, the distributions remain centered at the correct value.

Figure 6.8 shows effect of scaling the energies of all the jets up or down by the 10% estimated systematic error in the jet energy scale (see Section 6.2.1). This translates into a change in the fitted mass of about 10 GeV in either direction, or $\approx 5.5\%$.

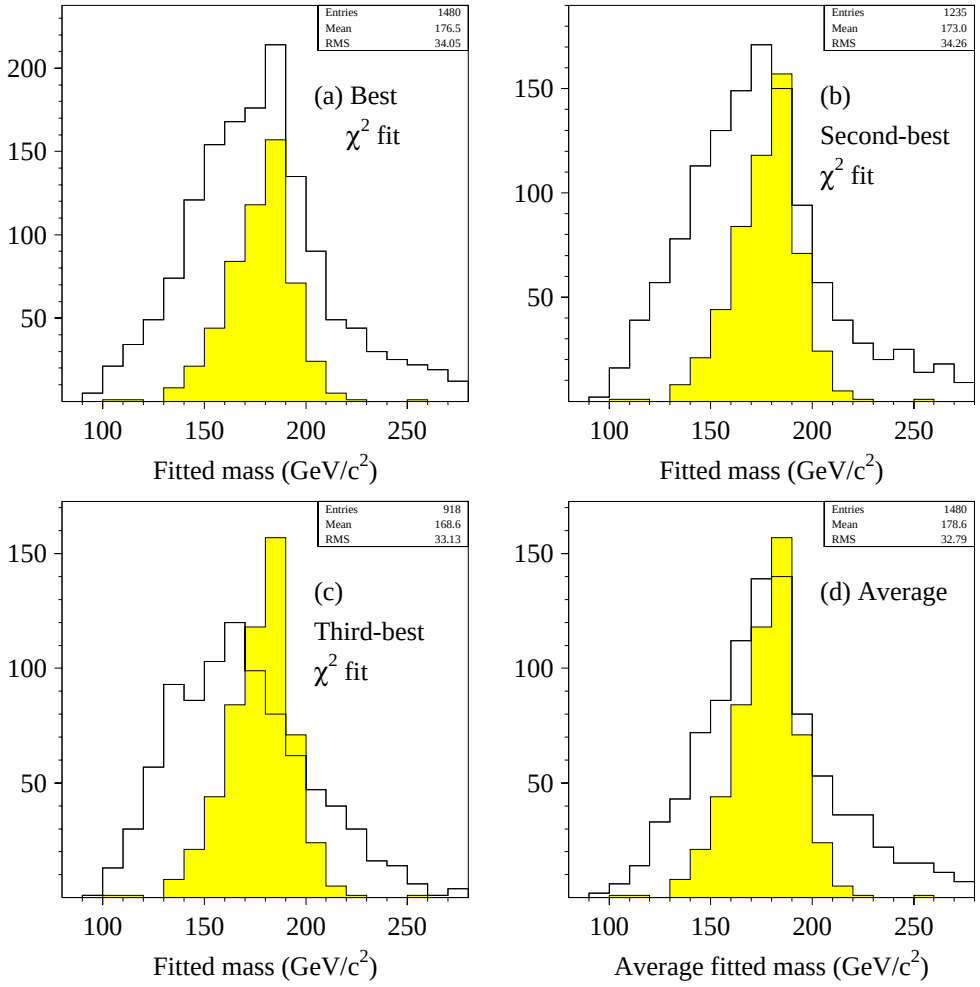

Figure 6.6: Fitted mass distributions for (a) best, (b) second-best, and (c) third-best χ^2 solutions and (d) weighted average. For smeared $t\bar{t} \to e+$jets ISAJET Monte Carlo with $m_t = 180\,\mathrm{GeV}/c^2$. Shaded plots show the result for the correct jet permutation (the same for all four plots).

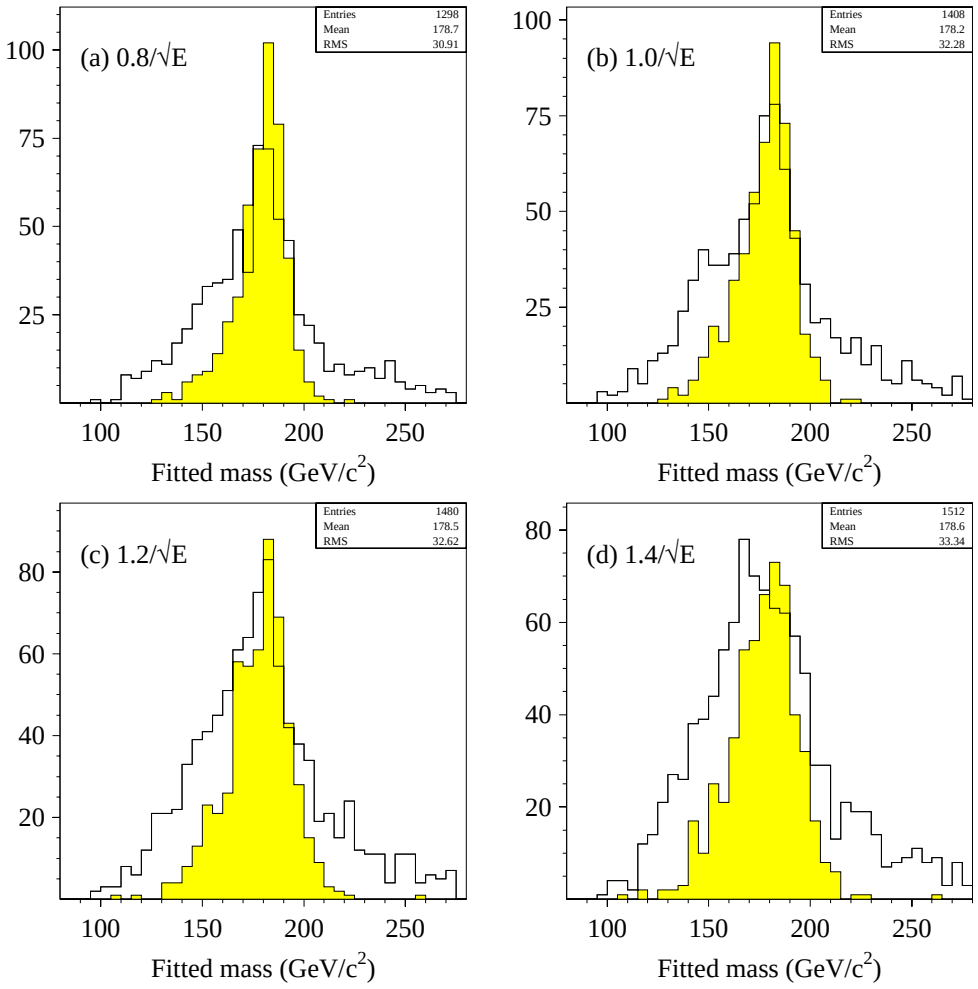

Figure 6.7: Fitted mass distributions for smeared $t\bar{t} \rightarrow e +$ jets ISAJET Monte Carlo with $m_t = 180\,\mathrm{GeV}/c^2$, for different jet smearings. (a) $0.80/\sqrt{E}$. (b) $1.00/\sqrt{E}$. (c) $1.20/\sqrt{E}$. (d) $1.40/\sqrt{E}$. Unshaded plots are the weighted average of the best three χ^2 solutions, shaded plots are the correct permutations.

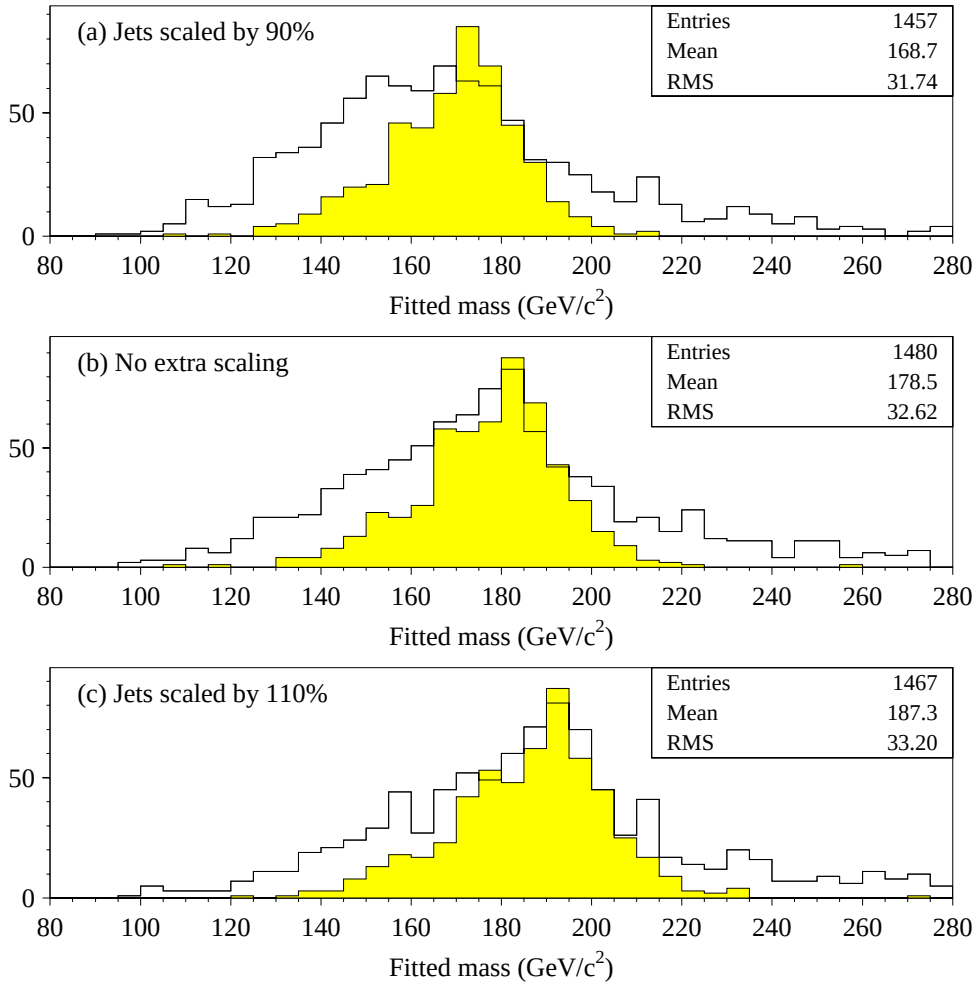

Figure 6.8: Fitted mass distributions for smeared $t\bar{t} \to e + \text{jets}$ ISAJET Monte Carlo with $m_t = 180\,\text{GeV}/c^2$, for different jet scalings. (a) Jets scaled by 90%. (b) Jets unscaled. (c) Jets scaled by 110%. Unshaded plots are the weighted average of the best three χ^2 solutions, shaded plots are the correct permutations.

6.5 Tests With Full Detector Simulation

The next step is to examine the results of the fitting algorithm for samples which have been through the full DØ detector simulation and reconstruction. All the Monte Carlo samples in this section have been subjected to the kinematic selection cuts described in Chapter 5.

The object resolutions which will be used for the fit are as follows.

$\frac{\sigma E(e)}{E(e)} = \frac{0.15}{\sqrt{E}} \oplus 0.03$	$\frac{\sigma k(\mu)}{k(\mu)} = \frac{0.01}{k} \oplus 0.2$	$\frac{\sigma E(\text{jet})}{E(\text{jet})} = \frac{0.82}{\sqrt{E}} \oplus 0.19$
$\sigma\phi(e) = 0.0060\,\text{rad}$	$\sigma\phi(\mu) = 0.0060\,\text{rad}$	$\sigma\phi(\text{jet}) = 0.035\,\text{rad}$
$\sigma\eta(e) = 0.0072$	$\sigma\eta(\mu) = 0.0072$	$\sigma\eta(\text{jet}) = 0.040$
$\sigma k_{Tx} = 12\,\text{GeV}$	$\sigma k_{Ty} = 12\,\text{GeV}$	

(6.12)

(Where E is measured in GeV and $k = 1/p$ is measured in GeV^{-1}.) Most of these values were determined by comparing the parton-level objects which were input to the detector simulation with the final objects output by the reconstruction program. The errors for the jet and electron energies are taken from [140]. The error for the muon momentum is taken from [43]. The results for the electron angular errors were reused for the muons. The large constant errors in the jet and electron energy resolutions above serve to account for the effects of radiation and the W width, which tend to spoil the kinematic constraints.

The muon momentum resolution in the Monte Carlo is known to be unrealistically good [47, p. 100]. To partially compensate for this, the momenta of high-p_T muons in the μ + jets Monte Carlo samples are smeared with the

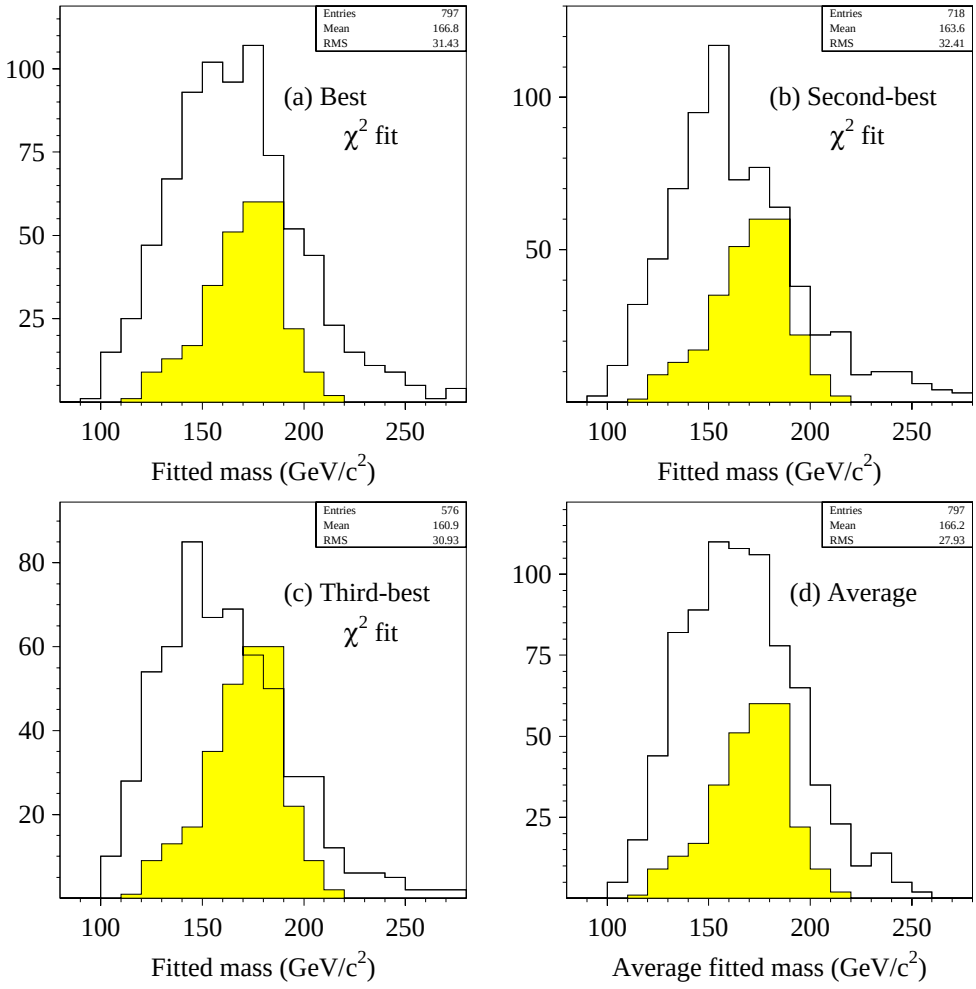

Figure 6.9: Fitted mass distributions for (a) best, (b) second-best, (c) third-best χ^2 solutions and (d) weighted average. For $t\bar{t} \to e + \text{jets}$ ISAJET Monte Carlo, with $m_t = 180\,\text{GeV}/c^2$. Shaded plots show the result for the correct jet permutation.

resolution given in (6.12).

The results of the fitter are shown for an ISAJET $t\bar{t} \to e+$jets Monte Carlo sample with $m_t = 180\,\mathrm{GeV}/c^2$ in Figure 6.9 (with full detector simulation and reconstruction). A χ^2 cut of $\chi^2 < 7$ was applied. As was seen with the parton-level tests, the averaging procedure reduces the tails of the distribution, while not introducing a significant bias. Also shown in Figure 6.10 is a comparison between the averaging procedure described above and an alternate procedure which consists of taking all solutions which satisfy the χ^2 cut (up to five per event) and plotting them with each one weighted by $e^{-\chi^2/2}$. This appears to be inferior to the averaging procedure.

Figure 6.11 shows the fit χ^2 distributions for the three lowest χ^2 jet permutations and for the correct permutation. The arrow shows the cut value. Note that the mean χ^2 for the correct jet permutation is close to 2, as expected for a 2C fit. However, this agreement was to some extent engineered by the inclusion of the large constant errors in the energy resolutions.

Figure 6.12 shows the fit results for all four lepton + jets channels. Compared to the $e+$jets channels, the $\mu+$jets channels show no additional biases, but are somewhat wider, due to the poor muon momentum resolution. The tagged channels are offset about 5–10 GeV/c^2 higher than the untagged channels; this is approximately independent of the input top mass. This is due to the additional corrections made for tagged jets. The resolution for the tagged channels is also slightly worse than for the untagged channels.

In the previous figure, the b-tag was enforced for the tagged channels, i.e., only jet permutations in which the tagged jet was used as a b were fit.

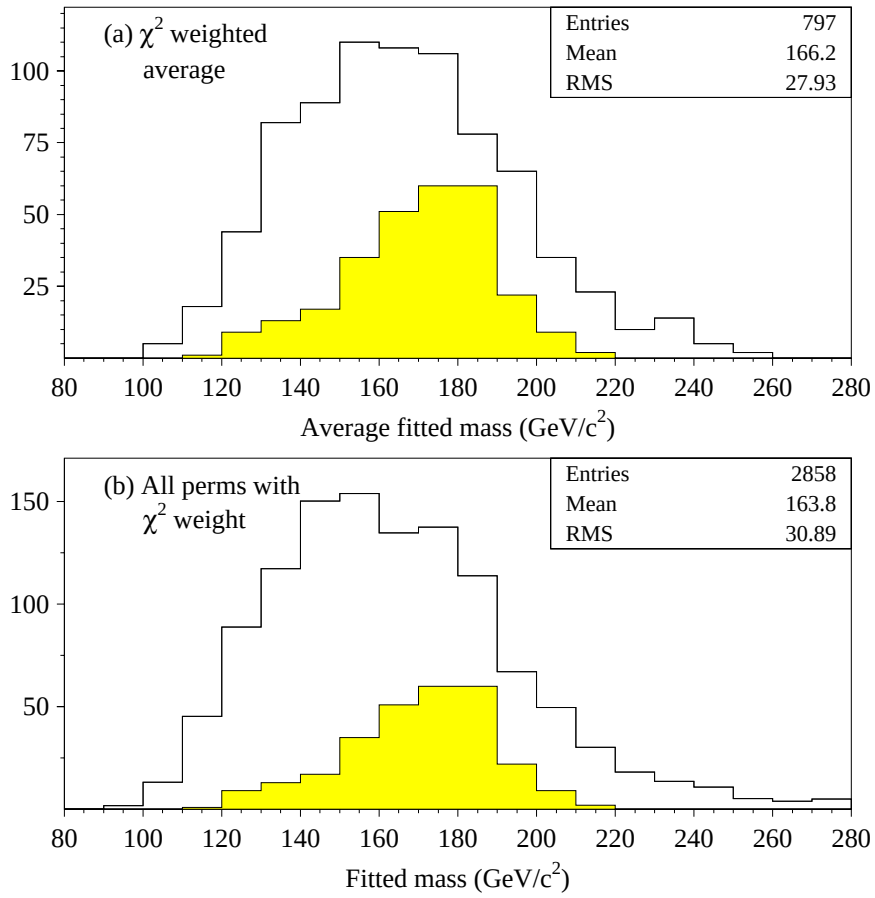

Figure 6.10: Comparison of (a) χ^2 weighted average of best three solutions, and (b) all solutions plotted, each one with a χ^2 weight. For $t\bar{t} \to e + \text{jets}$ ISAJET Monte Carlo, with $m_t = 180\,\text{GeV}/c^2$. Shaded plots show the result for the correct jet permutation.

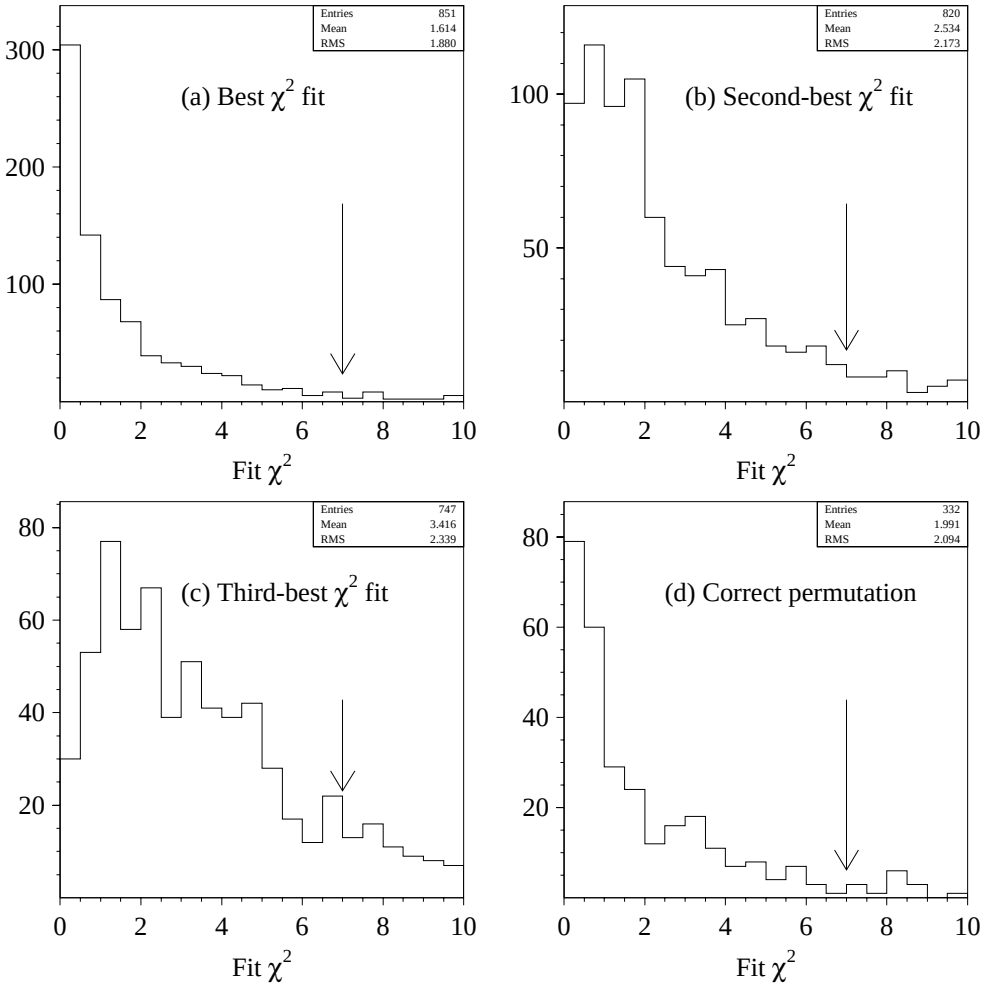

Figure 6.11: Fit χ^2 distributions for (a) best, (b) second-best, and (c) third-best χ^2 solutions and (d) the correct jet permutation. For $t\bar{t} \rightarrow e+$jets ISAJET Monte Carlo, with $m_t = 180\,\mathrm{GeV}/c^2$. The arrow shows the cut value.

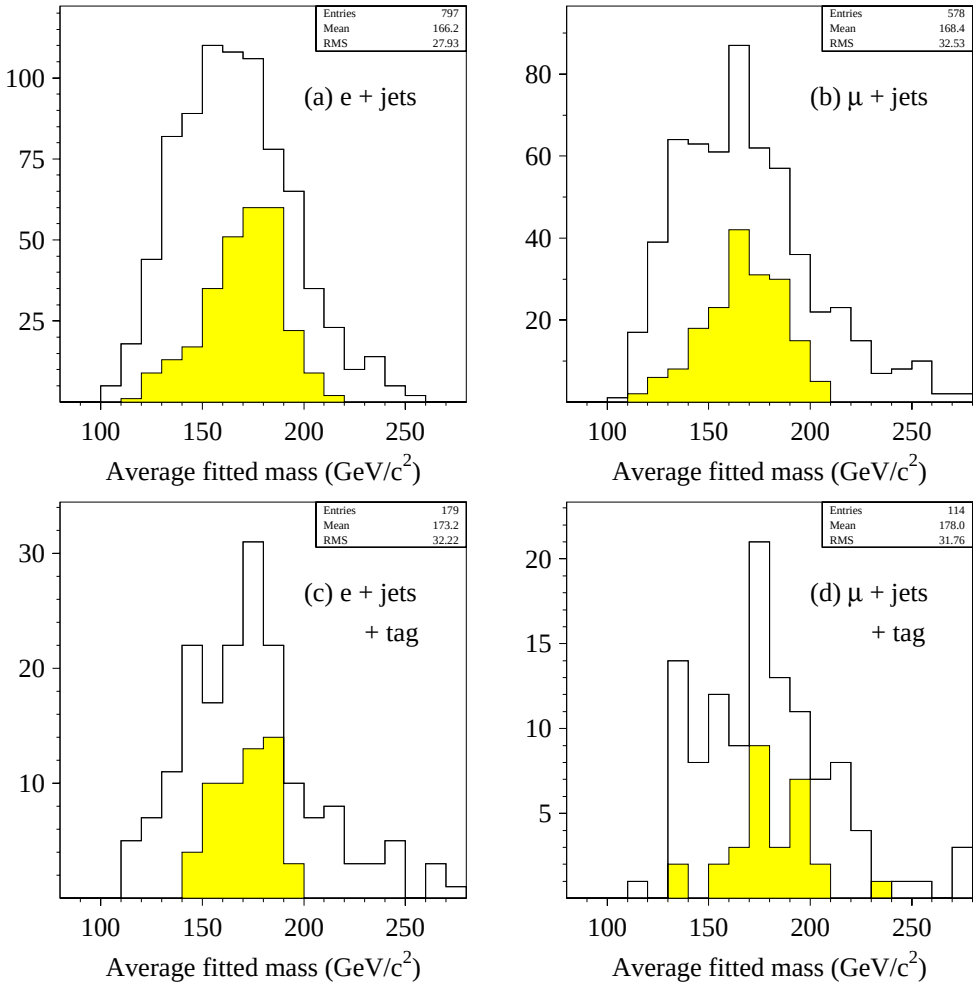

Figure 6.12: Fitted mass distributions for $t\bar{t}$ ISAJET Monte Carlo, with $m_t = 180\,\mathrm{GeV}/c^2$. For (a) e+jets, (b) μ+jets, (c) e+jets+tag, and (d) μ+jets+tag. Muon tags are enforced in the tagged channels. Shaded plots show the result for the correct jet permutation.

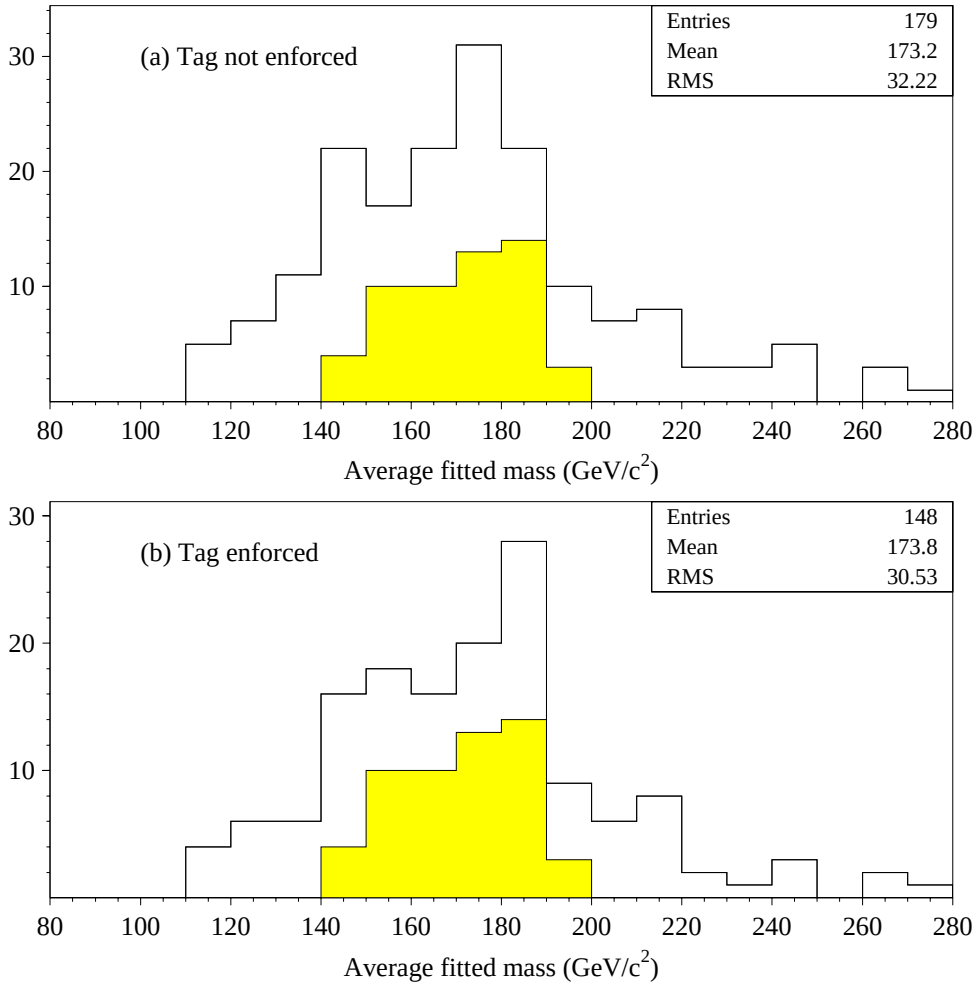

Figure 6.13: Fitted mass distributions, for the e + jets + tag channel (a) with and (b) without enforcing b-tagging. For $t\bar{t}$ ISAJET Monte Carlo, with $m_t = 180\,\mathrm{GeV}/c^2$. Shaded plots show the result for the correct jet permutation.

The effect of enforcing the tag is shown in Figure 6.13, which shows the same sample fit both with and without the tagging requirement. As can be seen, the difference is fairly small.

Figure 6.14 shows how the fitted mass distribution varies with the number of jets in the event. Looking at the RMS values, the degradation is surprisingly slow with increasing multiplicity.

In the previous figures, only the top four jets were used in the fit, no matter what the jet multiplicity in the event was. One can also consider trying to use additional jets, as discussed in Section 6.3.2. The result of doing so, without making any other alterations in the fitting procedure, is shown in Figure 6.15. Up to six jets were used, and all possible permutations were considered. However, although this procedure takes about an order of magnitude more computation time than always taking just the top four jets, the final result is not significantly better.

Another approach to try to deal with additional radiation is to increase the size of the jet cone. Figure 6.16 compares the results using the $R = 0.3$ jet cone to those obtained with a $R = 0.5$ cone. As can be seen, the results from the larger jet cone are somewhat inferior to those with the $R = 0.3$ cone. The caveat here is that the full set of out-of-cone corrections was not available for the $R = 0.5$ cone; however, the corrections one would make for that cone size are small [137].

Figure 6.17 shows the result of varying the jet scale up or down by the 10% error mentioned in Section 6.2.1. This translates into a change in the mean fit mass of about $9\,\mathrm{GeV}/c^2$ or $\approx 5\%$.

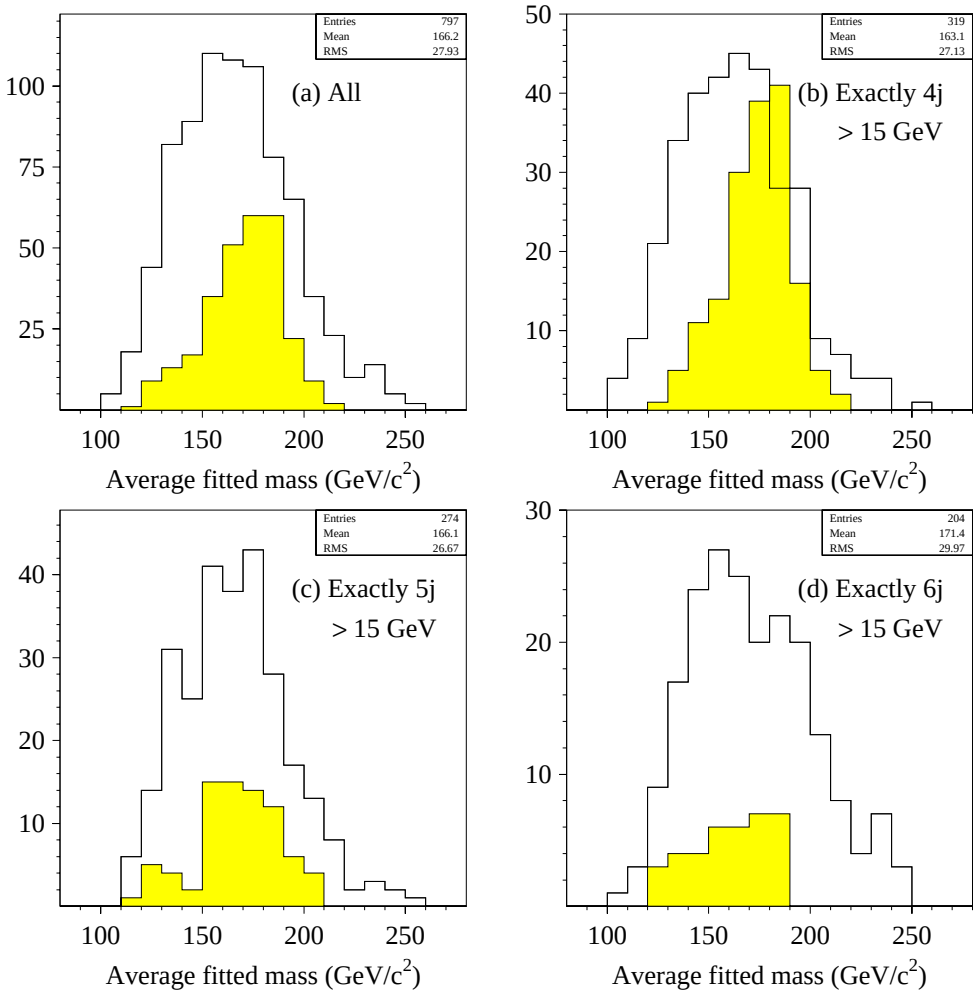

Figure 6.14: Fitted mass distributions for different jet multiplicities. (a) All multiplicities, (b) exactly four jets, (c) exactly five jets, and (d) exactly six jets. For $t\bar{t}$ ISAJET Monte Carlo, with $m_t = 180\,\mathrm{GeV}/c^2$ in the $e+$jets channel. Shaded plots show the result for the correct jet permutation.

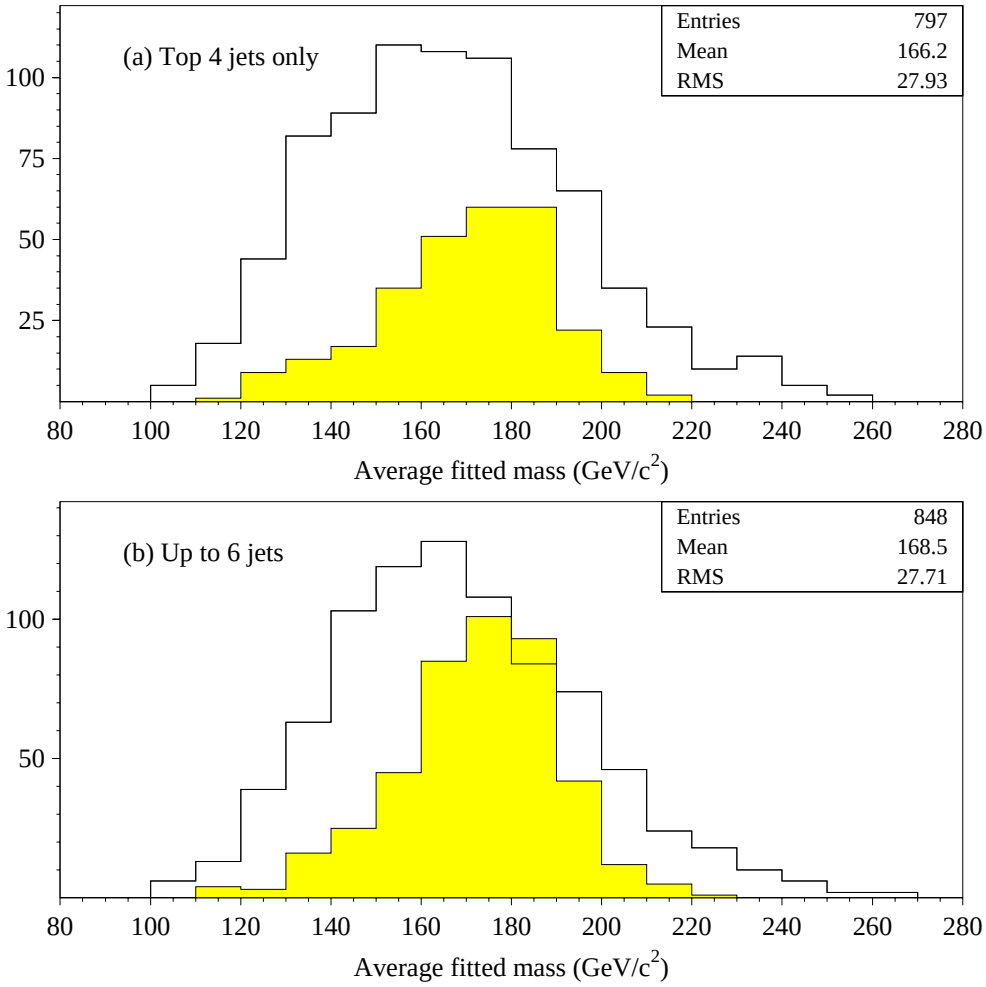

Figure 6.15: Fitted mass distributions for the $t\bar{t} \rightarrow e + \text{jets}$ channel using (a) only the top four jets and (b) up to six jets. For $t\bar{t}$ ISAJET Monte Carlo, with $m_t = 180\ \text{GeV}/c^2$. Shaded plots show the result for the correct jet permutation.

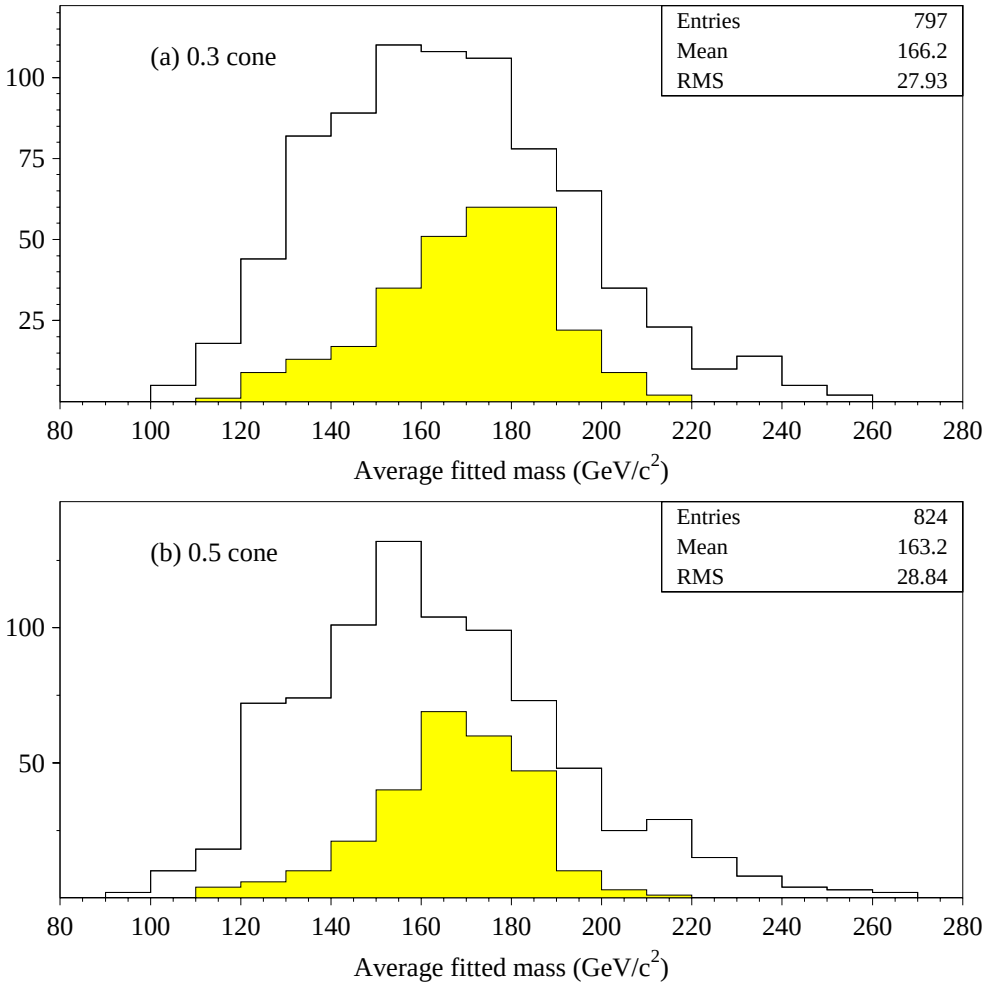

Figure 6.16: Fitted mass distributions for the $t\bar{t} \rightarrow e + \text{jets}$ channel using (a) $R = 0.3$ cone jets and (b) $R = 0.5$ cone jets. For ISAJET Monte Carlo, with $m_t = 180\ \text{GeV}/c^2$. Shaded plots show the result for the correct jet permutation.

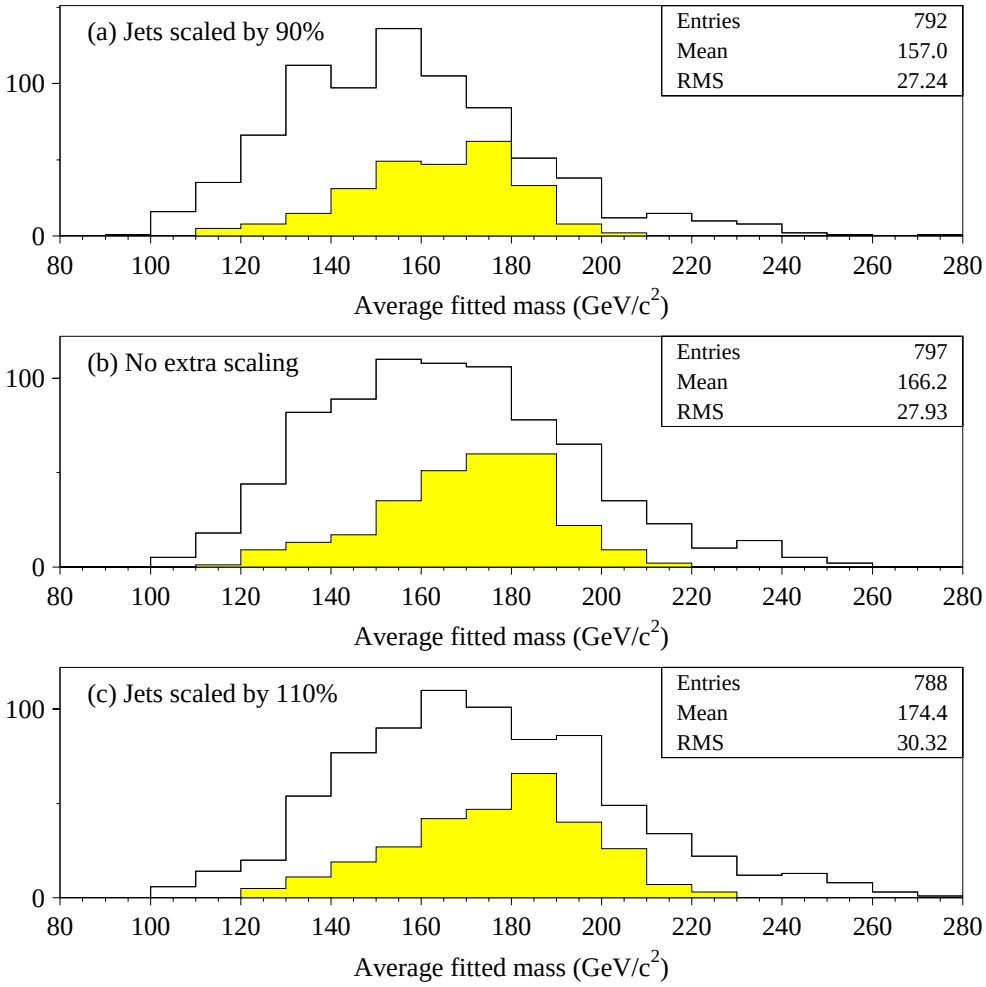

Figure 6.17: Fitted mass distributions for the $t\bar{t} \rightarrow e +$ jets channel with jets scaled by (a) 90%, (b) 100%, and (c) 110%. For $t\bar{t}$ ISAJET Monte Carlo, with $m_t = 180\ \mathrm{GeV}/c^2$. Shaded plots show the result for the correct jet permutation.

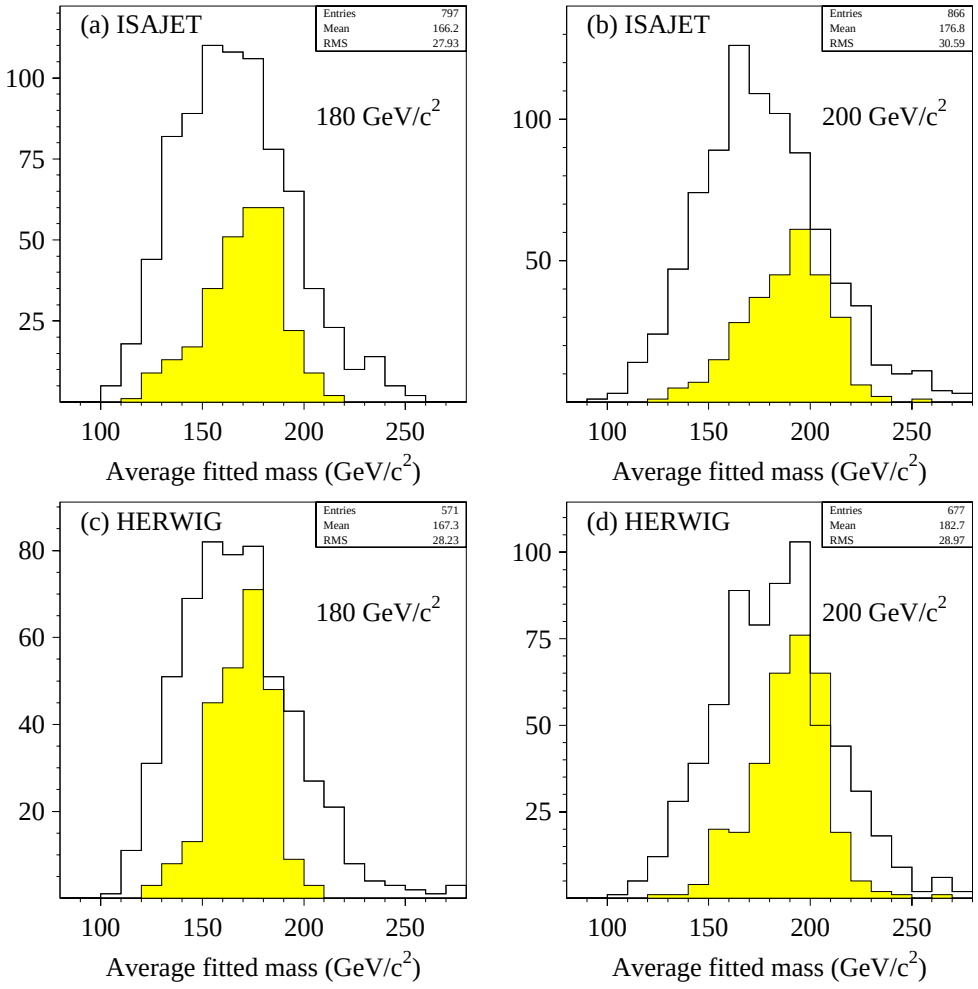

Figure 6.18: Fitted mass distributions for the $t\bar{t} \rightarrow e + \text{jets}$ channel for (a) $t\bar{t}$ ISAJET with $m_t = 180\,\text{GeV}/c^2$, (b) $t\bar{t}$ ISAJET with $m_t = 200\,\text{GeV}/c^2$, (c) $t\bar{t}$ HERWIG with $m_t = 180\,\text{GeV}/c^2$, and (d) $t\bar{t}$ HERWIG with $m_t = 200\,\text{GeV}/c^2$. Shaded plots show the result for the correct jet permutation.

Besides ISAJET, another popular Monte Carlo event generator is HERWIG. The results from the two are compared in Figure 6.18. Empirically, HERWIG is seen to produce somewhat less radiation than ISAJET does. This is reflected in the fact that it is more likely to find the correct jet permutation with HERWIG. This can be seen in the figure and also in the numbers in Table 6.1. As for the shape of the fitted mass distribution, there is not much difference between ISAJET and HERWIG for a top mass of 180 GeV/c^2. However, at 200 GeV/c^2, the difference is much more pronounced: the HERWIG distribution has somewhat less bias, and is also slightly narrower. (It is interesting to note that there is now some evidence that HERWIG underestimates the amount of radiation which is produced [138].)

Table 6.1 also shows that the efficiency for finding the correct permutation goes up slightly with increasing top mass. By contrast, these efficiencies do not change appreciably when the kinematic cuts ($\mathcal{A}$ and H_T) are changed.

So far, all the plots which have been shown have used the loose cuts. Figure 6.19 compares those results with those obtained when either the standard cuts or no $\mathcal{A}$, H_T cuts are used, for several top masses and the VECBOS W+jets background. The loose cuts, which consist of only a relaxed aplanarity cut, are seen not to change the shapes of the distributions appreciably. The standard cuts, on the other hand, add a fairly high H_T cut and bias the fitted mass distributions for low input masses significantly upwards. The background is also biased upwards. This bias is not large, however, for input top masses above about 180 GeV/c^2.

For the final set of fitted mass distributions, the results from all channels

	ISAJET		HERWIG	
Top mass (GeV/c^2)	160	200	160	200
(a) Probability of having the correct four partons as the top four jets	$39 \pm 2\%$	$42 \pm 2\%$	$47 \pm 3\%$	$52 \pm 2\%$
(b) Given (a), probability of having a good fit for the correct permutation	$85 \pm 5\%$	$78 \pm 4\%$	$89 \pm 5\%$	$85 \pm 4\%$
(c) Given (b), probability of the correct fit having the lowest χ^2	$32 \pm 3\%$	$40 \pm 3\%$	$42 \pm 4\%$	$50 \pm 3\%$
(d) Given (b), probability of the correct fit being in the best three fits	$67 \pm 4\%$	$75 \pm 4\%$	$79 \pm 5\%$	$84 \pm 4\%$
Total chance of getting the correct permutation by picking the lowest χ^2	$11 \pm 1\%$	$13 \pm 1\%$	$18 \pm 2\%$	$22 \pm 2\%$
Total chance of getting the correct permutation in the best three fits	$22 \pm 2\%$	$25 \pm 1\%$	$33 \pm 2\%$	$37 \pm 2\%$

Table 6.1: Efficiencies for finding the correct jet permution, for ISAJET and HERWIG. From $t\bar{t} \rightarrow e +$ jets Monte Carlo. Errors are statistical only.

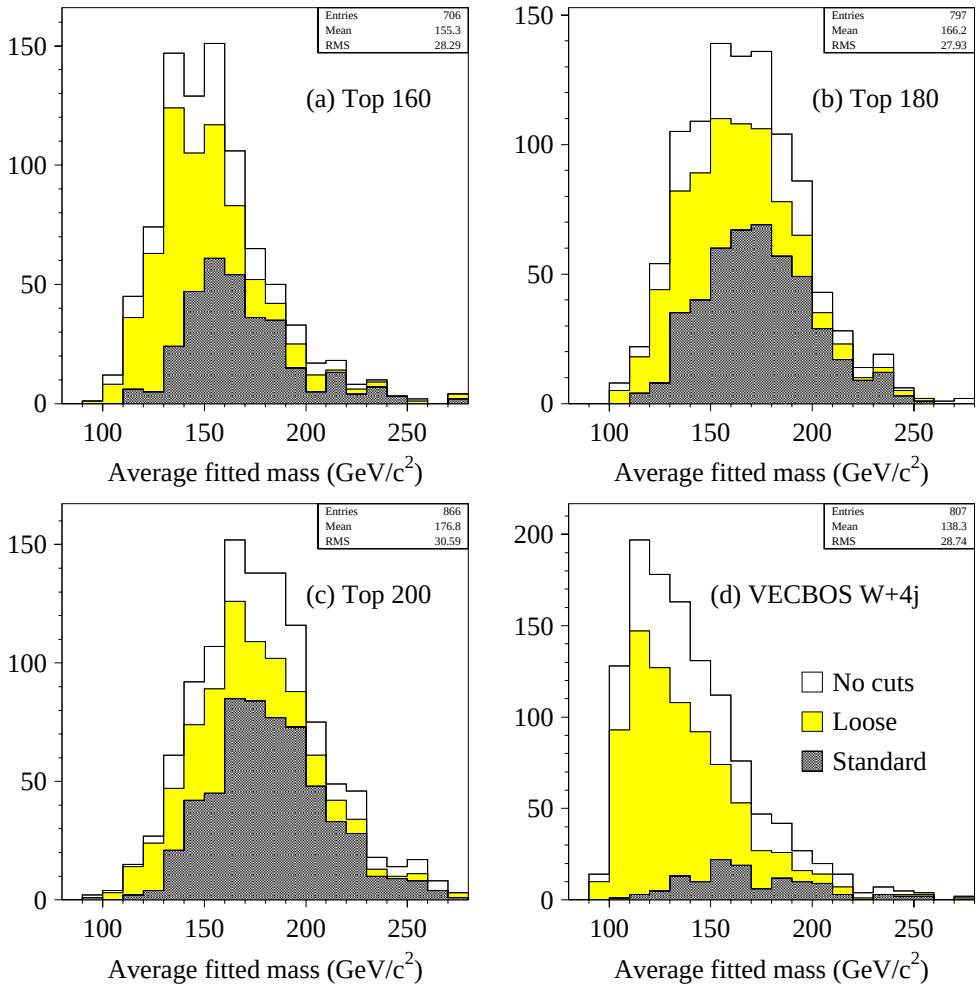

Figure 6.19: Fitted mass distributions for the e + jets channel using either no $\mathcal{A}$, H_T cuts (unshaded plot), loose cuts (lightly shaded plot), or standard cuts (dark plot) for (a) $m_t = 160\,\mathrm{GeV}/c^2$ ISAJET $t\bar{t}$ Monte Carlo, (b) $m_t = 180\,\mathrm{GeV}/c^2$ ISAJET $t\bar{t}$ Monte Carlo, (c) $m_t = 200\,\mathrm{GeV}/c^2$ ISAJET $t\bar{t}$ Monte Carlo, and (d) VECBOS W + jets Monte Carlo. Statistics reflect the loose cuts histogram.

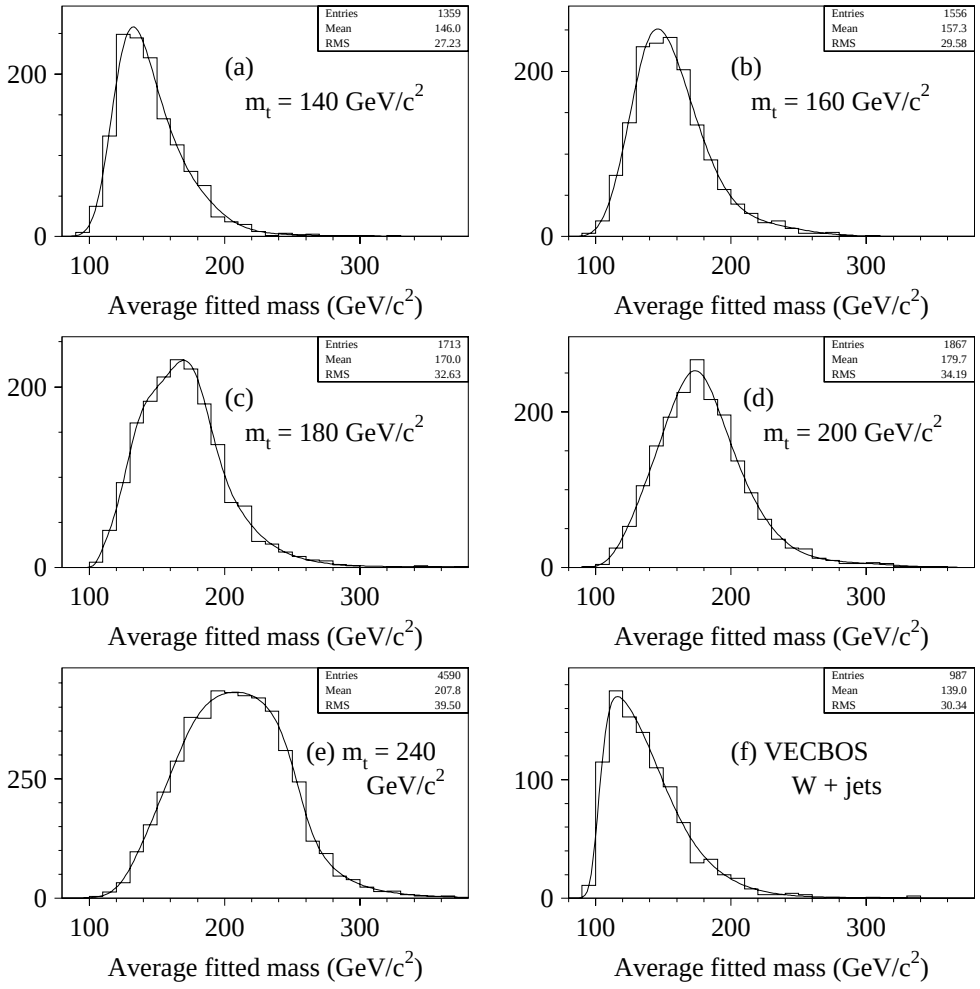

Figure 6.20: Final fitted mass distributions for all channels combined For loose cuts, using ISAJET. (a) $m_t = 140\,\mathrm{GeV}/c^2$, (b) $m_t = 160\,\mathrm{GeV}/c^2$, (c) $m_t = 180\,\mathrm{GeV}/c^2$, (d) $m_t = 200\,\mathrm{GeV}/c^2$, (e) $m_t = 240\,\mathrm{GeV}/c^2$, and (f) VECBOS W + jets background.

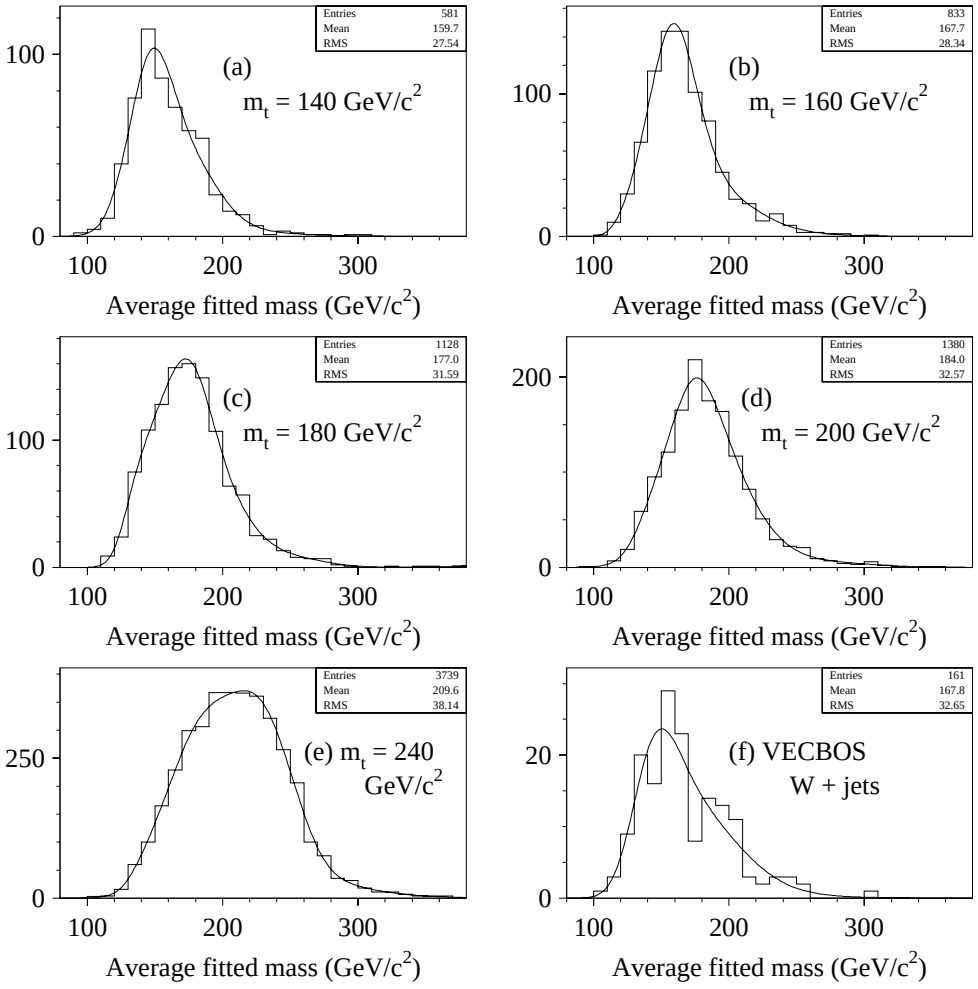

Figure 6.21: Final fitted mass distributions for all channels combined. For standard cuts, using ISAJET. (a) $m_t = 140\,\mathrm{GeV}/c^2$, (b) $m_t = 160\,\mathrm{GeV}/c^2$, (c) $m_t = 180\,\mathrm{GeV}/c^2$, (d) $m_t = 200\,\mathrm{GeV}/c^2$, (e) $m_t = 240\,\mathrm{GeV}/c^2$, and (f) VECBOS W + jets background.

are combined together, by simply adding together the appropriate Monte Carlo samples (with no additional weights; the samples were originally generated according to the appropriate branching ratios). Figures 6.20 and 6.21 show the results for all five top masses used as well as the VECBOS W + jets background, for both the loose and standard cuts. (There was no μ + jets Monte Carlo for $m_t = 240\ \mathrm{GeV}/c^2$, so the e + jets sample was reused for the μ + jets channel at that mass point.) Note that the lower mass tops have tails on the high side which bias them upwards, while the higher mass tops are biased downwards. If one plots the relation between the mean fitted mass and the input mass, one finds that it is close to linear, with a slope of about 0.6 with the loose cuts and 0.5 with the standard cuts (this is shown in Figure 6.22). The standard cuts have a smaller slope due to the larger biases present for small masses. This has the unfortunate consequence that any errors in the fitted masses will be inflated by approximately the inverse of this slope when converting to the final mass value.

For reference, the results are shown broken down by the individual channels in Figures 6.23, 6.24, 6.25, and 6.26. Note that the available Monte Carlo statistics are rather poor for the tagged channels; this is the major motivation for combining all channels together.

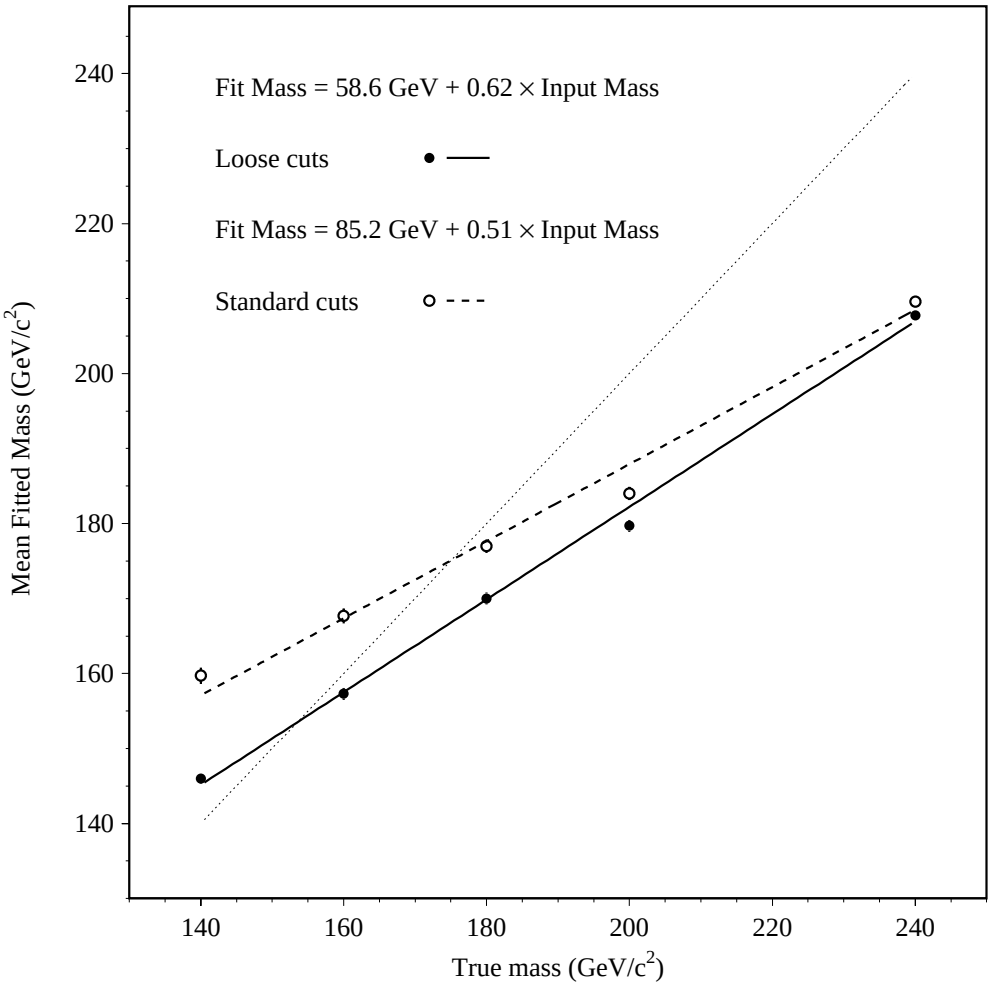

Figure 6.22: Mean fitted mass as a function of input mass for loose and tight cuts. From $t\bar{t}$ ISAJET Monte Carlo with all channels combined. The dotted line is drawn along the diagonal for comparison.

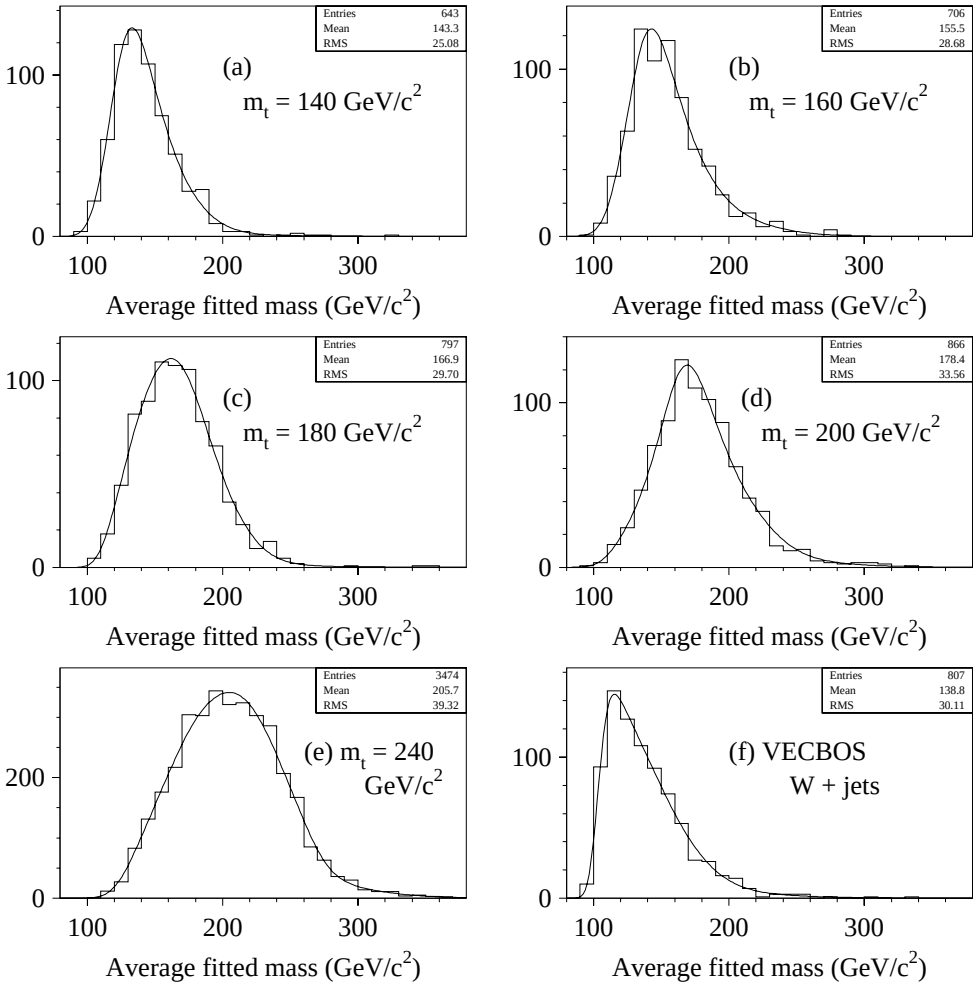

Figure 6.23: Final fitted mass distributions for e + jets channel. For loose cuts, using ISAJET. (a) $m_t = 140\,\mathrm{GeV}/c^2$, (b) $m_t = 160\,\mathrm{GeV}/c^2$, (c) $m_t = 180\,\mathrm{GeV}/c^2$, (d) $m_t = 200\,\mathrm{GeV}/c^2$, (e) $m_t = 240\,\mathrm{GeV}/c^2$, and (f) VECBOS W + jets background.

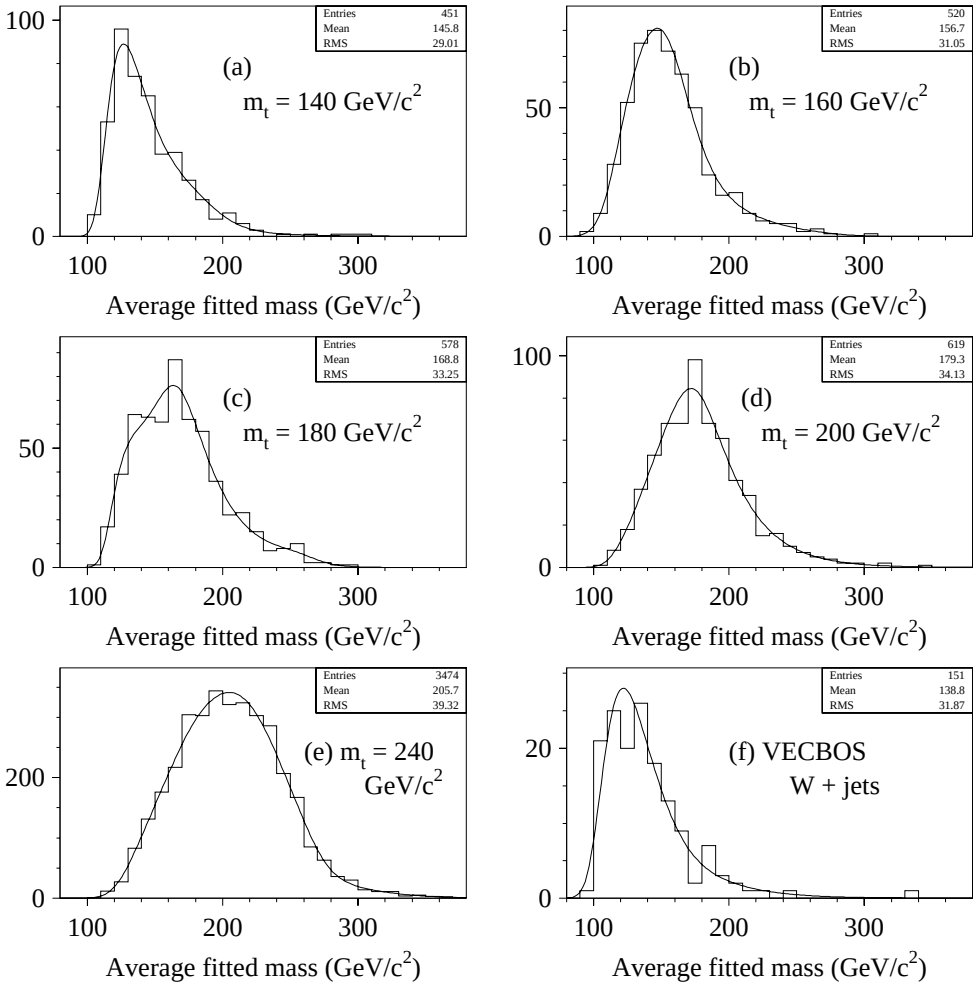

Figure 6.24: Final fitted mass distributions for μ + jets channel. For loose cuts, using ISAJET. (a) $m_t = 140\,\mathrm{GeV}/c^2$, (b) $m_t = 160\,\mathrm{GeV}/c^2$, (c) $m_t = 180\,\mathrm{GeV}/c^2$, (d) $m_t = 200\,\mathrm{GeV}/c^2$, (e) $m_t = 240\,\mathrm{GeV}/c^2$, and (f) VECBOS W + jets background.

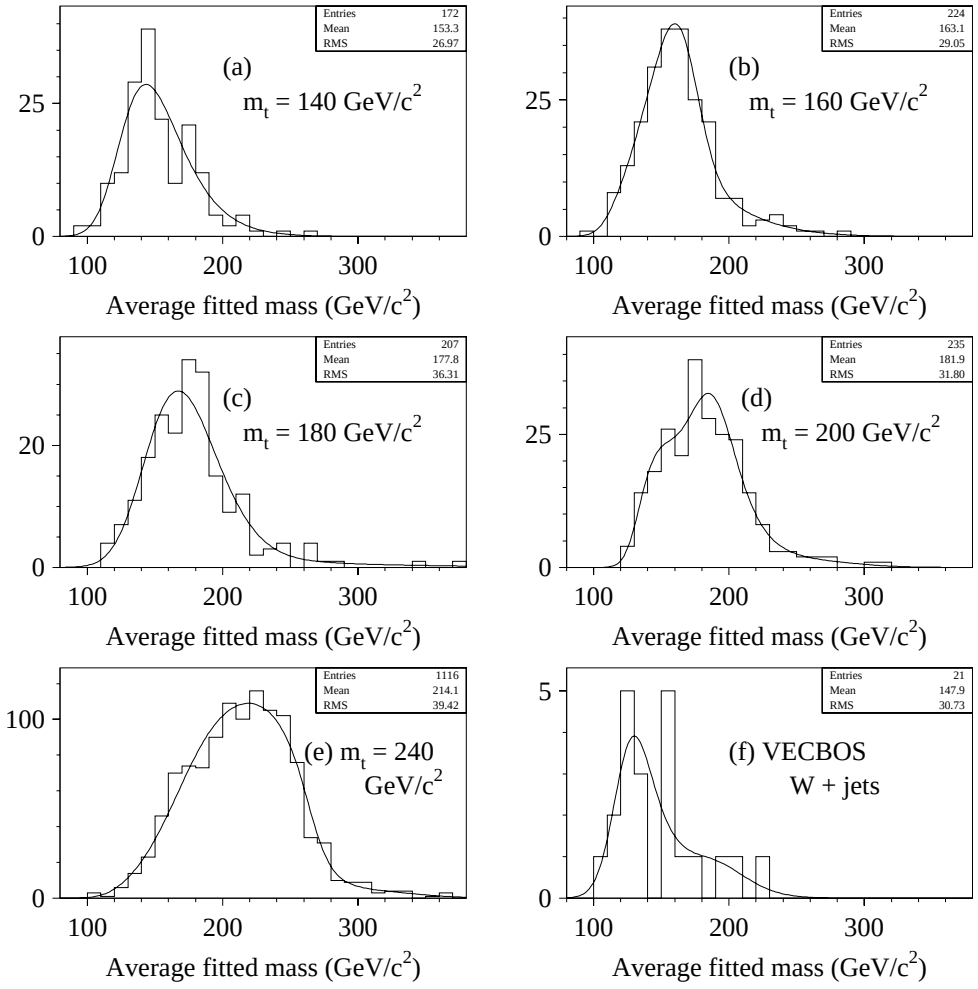

Figure 6.25: Final fitted mass distributions for e + jets/μ channel. For loose cuts, using ISAJET. (a) $m_t = 140\,\mathrm{GeV}/c^2$, (b) $m_t = 160\,\mathrm{GeV}/c^2$, (c) $m_t = 180\,\mathrm{GeV}/c^2$, (d) $m_t = 200\,\mathrm{GeV}/c^2$, (e) $m_t = 240\,\mathrm{GeV}/c^2$, and (f) VECBOS W + jets background.

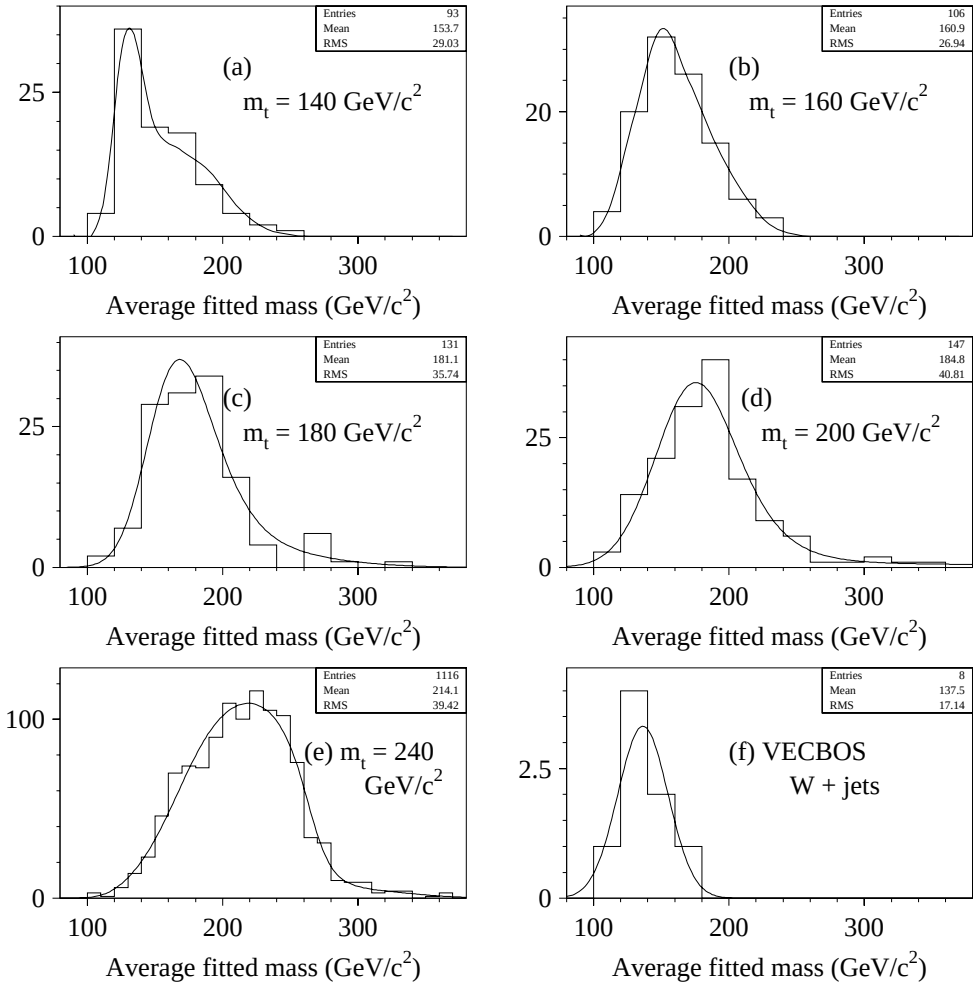

Figure 6.26: Final fitted mass distributions for μ + jets/μ channel. For loose cuts, using ISAJET. (a) $m_t = 140\,\mathrm{GeV}/c^2$, (b) $m_t = 160\,\mathrm{GeV}/c^2$, (c) $m_t = 180\,\mathrm{GeV}/c^2$, (d) $m_t = 200\,\mathrm{GeV}/c^2$, (e) $m_t = 240\,\mathrm{GeV}/c^2$, and (f) VECBOS W + jets background.

6.6 Summary

Figure 6.27 summarizes the effects which contribute to the broadening of the fitted mass distributions. Figure 6.27(a) shows the result from an unsmeared, parton-level Monte Carlo with only the correct permutation chosen. The major factors contributing to the width here are final-state radiation which is lost by choosing only the top four jets and (to a lesser extent) the W widths and initial state radiation which is merged into final-state pjets. The width this plot is already half of the width of the final distribution (Figure 6.27(d)). Figures 6.27(b) and (c) show what happens when parton smearing is turned on (using the same resolutions as in Section 6.4) and when the full detector simulation is used. Both of these increase the width of the distribution, but only by a few GeV/c^2 each. Note, however, that the result from the full detector simulation is shifted down somewhat; this is probably due to reconstruction inefficiencies for low-energy jets. Finally, Figure 6.27(d) shows the result when all jet permutations are considered, not just the correct one. This step shows a large increase in the width of the distribution. Taken together, these plots imply that the major contributions to the widths of the fitted mass distributions come from radiation and from the jet combinatorics and not from the measurement errors on the jets themselves.

Recall that the slope seen in the relation between the input mass and the mean fitted mass was about 0.62. If one makes the same calculation using only the correct jet permutations, the result is a slope of about 0.94; if one does it using only jet permutations which do not assign the correct jets to the correct

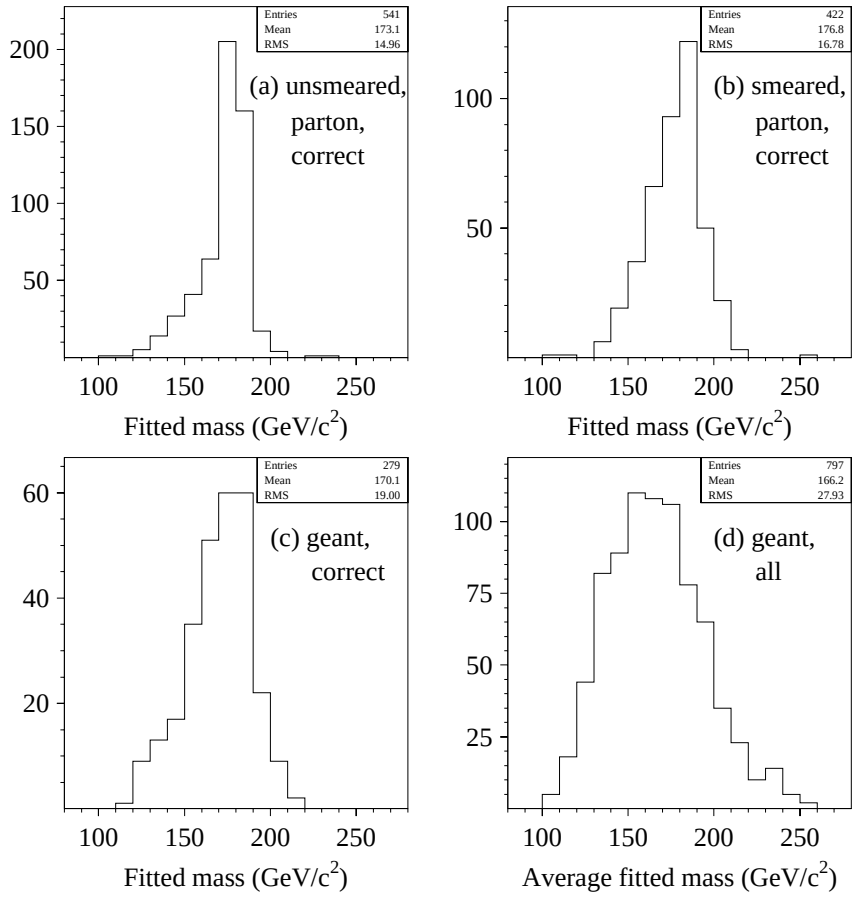

Figure 6.27: Fitted mass distributions for $t\bar{t} \to e + \text{jets}$ ISAJET Monte Carlo, with $m_t = 180\,\text{GeV}/c^2$. (a) Parton-level, unsmeared, correct permutation. (b) Parton-level, smeared, correct permutation. (c) Detector simulation, correct permutation. (d) Detector simulation, average of best three permutations.

tops, the resulting slope is about 0.56. This indicates that the shallowness of this slope is another effect of the jet combinatorics.

These results imply that a possible avenue for improving the resolutions is to try to develop methods for identifying initial and final state radiation.

Chapter 7

Likelihood Fitting

The next step in the analysis is to take a set of fitted mass measurements from a set of data, which may include background, and calculate the value of the top mass which is most consistent with the data. This is accomplished using a maximum likelihood fit.

7.1 Resolution Functions

The major inputs to the likelihood fit (besides the data itself) are the *resolution functions* for the signal and background. These are simply the probability distributions for the fitted mass, given that an event is background or top with some specific mass m_t. The background resolution function is thus a one-dimensional function of the fitted mass,

$$p(\text{measure fitted mass } m|\text{event is background}) = f_b(m), \qquad (7.1)$$

but the signal resolution function is a function of both the fitted mass and the presumed top mass:

$$p(\text{measure fitted mass } m|\text{event is top with mass } m_t) = f_s(m, m_t). \quad (7.2)$$

7.1.1 Smoothing

The resolution functions are approximated by the histograms of fitted masses for the various samples, as summarized at the end of Chapter 6, with a suitable normalization. However, the raw histograms do not provide an acceptable input to the final fit since they are full of discontinuities and can exhibit significant statistical fluctuations. To solve this problem, the histograms are smoothed to produce an analytic function which then serves as the input to the likelihood fit. For most cases, the standard HBOOK multiquadric smoothing algorithm was used [141, 142]. This smoothing algorithm takes several parameters which govern its sensitivity to fluctuations in the input and the smoothness of the final curve. Usually, the default values of these parameters worked well. However, for some of the histograms with relatively poor statistics, these parameters had to be manually adjusted to avoid oversensitivity to statistical fluctuations or other pathologies. For a few histograms with particularly small statistics, a double-gaussian fit was used instead.

7.1.2 Interpolation of Signal Resolution Functions

The signal resolution functions involve the added complication of being two-dimensional in nature. However, the Monte Carlo studies give this function only for certain discrete values of the input top mass. In order to obtain a function which is continuous with respect to both the fitted mass and the input mass, one must interpolate on the input top mass.

The general idea of the interpolation procedure is motivated by two facts. First, the resolution functions are single-peaked, and their shapes change gradually with changes in the input top mass. Second, if one plots the location of the peaks of the resolution functions versus the input top mass, the relation is approximately linear (peak $\approx Am_t + B$). Thus, the interpolation procedure first removes the dominant linear dependence of the resolution functions on the input mass, and then uses a polynomial interpolation on the input mass. In more detail:

1. The inputs to the interpolation procedure are N smoothed functions of fitted mass $l_i(m)$, each being the resolution function for a different input top mass m_{ti}. Each of these input functions should be normalized to unity.

2. For each resolution function l_i, find its maximum. Take the location of these maxima as a function of the input top mass and fit it to a linear form $Am_t + B$ with a standard least-squares fit. Call the resulting linear function $P(m_t)$.

3. Now suppose that the resolution function is to be evaluated for top mass m_t and fitted mass m.

4. First of all, if m_t is less than a cutoff (80 GeV), the result of the interpolation is taken to be zero.

5. Otherwise, for each input resolution function l_i, define the value

$$f_i = l_i\left(m - P(m_t) + P(m_{ti})\right). \tag{7.3}$$

6. Consider the f_i values as a function of the input top masses m_{ti}, and perform a (at most) fourth degree polynomial fit. (The degree is limited to two less than the number of mass values.) This fit is then evaluated at m_t; this is the result of the interpolation. If m_t is larger than some high-side cutoff (usually 10–20 GeV/c^2 above the highest mass input point), the polynomial is evaluated at the cutoff point instead. This prevents the result from blowing up too rapidly as the polynomial interpolation becomes an extrapolation, but unfortunately, it also introduces a discontinuity in the derivative of the resolution function. This can adversely affect any later fitting done in the vicinity of the cutoff.

   Note that the definition in (7.3) ensures that if m_t equals some m_{ti}, then the corresponding f_i satisfies $f_i = l_i(m)$, as one would expect.

7. The input resolution functions l_i are normalized to unity. Although it is not explicitly imposed, it is seen empirically that the result of the interpolation for any m_t within the interpolation region is within a percent or so of being normalized to unity.

The results of the interpolation procedure are illustrated for the loose and tight cuts in Figures 7.1 and 7.2.

7.2 Maximum Likelihood Fit

This section discusses the form of the likelihood function used and the actual fit procedure.

7.2.1 Derivation of the Likelihood Function

Start by making some definitions. The data consist of N events, each of which has a fitted mass:

$$D \equiv m_1, \ldots, m_N \equiv \text{Observed data.} \tag{7.4}$$

There are three parameters which could be measured:

$$m_t \equiv \text{Mass of top.} \tag{7.5}$$

$$\sigma_s \equiv \text{Cross section of top (signal).} \tag{7.6}$$

$$\sigma_b \equiv \text{Cross section of background.} \tag{7.7}$$

Instead of dealing with the cross sections directly, however, it is convenient to multiply them by the appropriate constants to obtain the expected number of events for the signal and background:

$$n_s \equiv \text{Expected number of signal events} \tag{7.8}$$

$$n_b \equiv \text{Expected number of background events.} \tag{7.9}$$

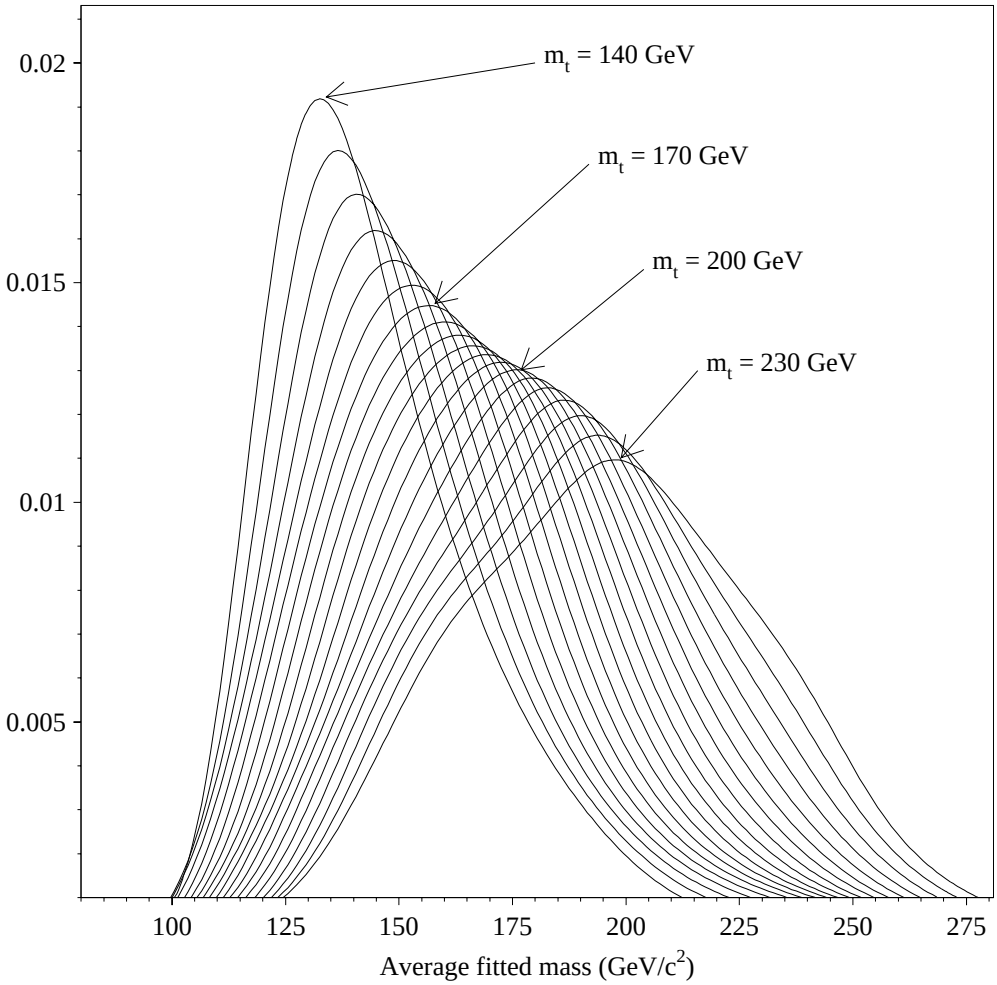

Figure 7.1: Interpolated signal resolution functions with loose cuts, plotted for different values of input top mass m_t. Made with ISAJET Monte Carlo with all channels combined.

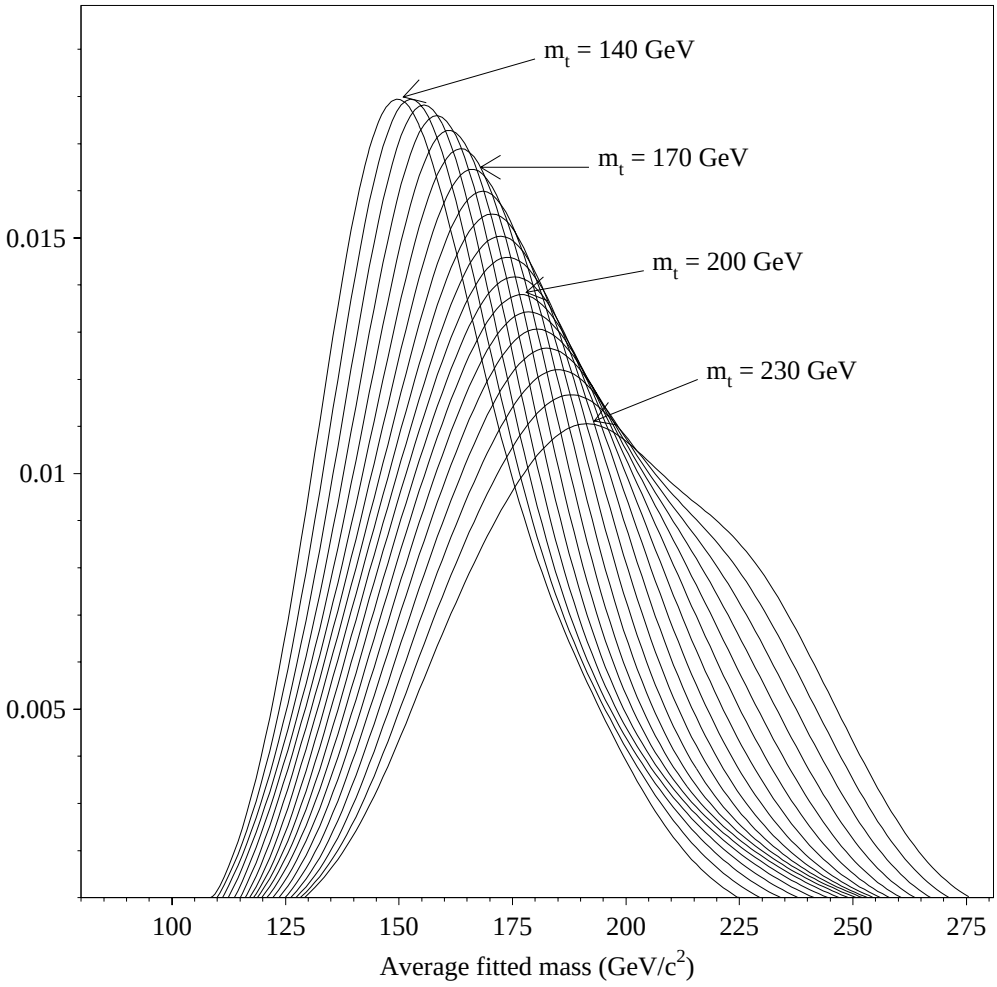

Figure 7.2: Interpolated signal resolution functions with the standard cuts, plotted for different values of input top mass m_t. Made with ISAJET Monte Carlo with all channels combined.

Define the *likelihood* of the data as the probability of seeing the observed data set D given the values of n_s, n_b, and m_t. This is notated $p(D|n_s, n_b, m_t)$. Denote the (unknown) actual number of signal events in the data sample D by N_s, and the actual number of background events by N_b:

$$N_s \equiv \text{Actual number of signal events in } D. \tag{7.10}$$

$$N_b \equiv \text{Actual number of background events in } D. \tag{7.11}$$

These values must satisfy

$$N_s + N_b = N. \tag{7.12}$$

The likelihood $p(D|n_s, n_b, m_t)$ can then be expanded by summing over all possible values of N_s and applying the product rule:

$$\begin{aligned} p(D|n_s, n_b, m_t) &= \sum_{N_s=0}^{N} p(D, N_s, N_b|n_s, n_b, m_t) \\ &= \sum_{N_s=0}^{N} p(D|N_s, N_b, n_s, n_b, m_t)\, p(N_s, N_b|n_s, n_b, m_t) \end{aligned} \tag{7.13}$$

Since the signal and background processes are independent, the second factor, $p(N_s, N_b|n_s, n_b, m_t)$, should be the product of two Poisson forms. Define the notation

$$\begin{aligned} q(N, m) &\equiv \frac{e^{-m} m^N}{N!} \\ &= \text{Probability of observing } N \text{ events in a Poisson} \\ &\quad \text{distribution with mean } m. \end{aligned} \tag{7.14}$$

Then,

$$p(N_s, N_b|n_s, n_b, m_t) = q(N_s, n_s)\, q(N_b, n_b). \tag{7.15}$$

Now to evaluate $p(D|N_s, N_b, n_s, n_b, m_t)$. This can again be done by summing over all the possible ways in which this can happen; i.e., over all the possible ways of picking N_s signal events out of the N data events in D. Define some more notation:

$$s \equiv \text{An element of the set of all } 2^N \text{ ways to divide up the } N \text{ events in } D \text{ into signal and background events.} \tag{7.16}$$

$$s_{N_s} \equiv \text{An element of the set of all } \binom{N}{N_s} \text{ ways of picking } N_s \text{ signal events from the } N \text{ events in } D. \tag{7.17}$$

$$S(s) \equiv \text{The set of signal events in combination } s. \tag{7.18}$$

$$B(s) \equiv \text{The set of background events in combination } s. \tag{7.19}$$

$$N_s(s) \equiv \text{The number of signal events in combination } s. \tag{7.20}$$

$$N_b(s) \equiv \text{The number of background events in combination } s. \tag{7.21}$$

Then,

$$\begin{aligned} p(D|N_s, N_b, n_s, n_b, m_t) &= p(D|N_s, N_b, m_t) \\ &= \sum_{s_{N_s}} p(D|s_{N_s}, N_s, N_b, m_t)\, p(s_{N_s}|N_s, N_b, m_t) \\ &= \sum_{s_{N_s}} p(D|s_{N_s}, m_t)\, p(s_{N_s}|N_s, N_b). \end{aligned} \tag{7.22}$$

To evaluate the second factor, note that each of the $\binom{N}{N_s}$ ways of picking N_s

signal events out of the N data events is equally likely, so that

$$p(s_{N_s}|N_s,N_b) = \frac{1}{\binom{N}{N_s}}. \tag{7.23}$$

Combining equations (7.13), (7.15), and (7.22),

$$p(D|n_s,n_b,m_t) = \sum_s p(D|s,m_t)\,\frac{1}{\binom{N}{N_s(s)}}\,q(N_s(s),n_s)\,q(N_b(s),n_b). \tag{7.24}$$

Finally, the factor $p(D|s,m_t)$ can be written in terms of the resolution functions:

$$p(D|s,m_t) = \left(\prod_{i\in S(s)} f_s(m_i,m_t)\right)\left(\prod_{i\in B(s)} f_b(m_i)\right). \tag{7.25}$$

So the likelihood is

$$\begin{aligned} p(D|n_s,n_b,m_t) = \sum_s &\left(\prod_{i\in S(s)} f_s(m_i,m_t)\right)\left(\prod_{i\in B(s)} f_b(m_i)\right)\times \\ &\frac{1}{\binom{N}{N_s(s)}}\,q(N_s(s),n_s)\,q(N_b(s),n_b). \end{aligned} \tag{7.26}$$

This expression can be written in a couple of other useful forms. First, note that

$$\begin{aligned} q(N_s,n_s)\,q(N_b,n_b)\frac{1}{\binom{N}{N_s(s)}} &= \frac{e^{-n_s}n_s^{N_s}}{N_s!}\frac{e^{-n_b}n_b^{N_b}}{N_b!}\frac{N_s!N_b!}{N!} \\ &= \frac{e^{-(n_s+n_b)}(n_s+n_b)^N}{N!}\frac{n_s^{N_s}n_b^{N_b}}{(n_s+n_b)^N} \\ &= q(N,n_s+n_b)\frac{n_s^{N_s}n_b^{N_b}}{(n_s+n_b)^N}. \end{aligned} \tag{7.27}$$

The likelihood can thus be rewritten

$$p(D|n_s,n_b,m_t) = \frac{q(N,n_s+n_b)}{(n_s+n_b)^N}\sum_s \left(\prod_{i\in S(s)} n_s f_s(m_i,m_t)\right)\left(\prod_{i\in B(s)} n_b f_b(m_i)\right)$$
$$= q(N,n_s+n_b)\prod_{i=1}^{N}\frac{n_s f_s(m_i,m_t)+n_b f_b(m_i)}{n_s+n_b}. \tag{7.28}$$

This is probably the form of the likelihood function which most people are used to seeing (i.e., from [33]).

Finally, instead of using the variables n_s and n_b, one can write the likelihood in terms of their relative fractions. Define

$$\epsilon \equiv \frac{n_s}{n_s+n_b}, \tag{7.29}$$

$$t \equiv n_s+n_b. \tag{7.30}$$

Then the likelihood becomes

$$p(D|\epsilon,t,m_t) = q(N,t)\prod_{i=1}^{N}\left(\epsilon f_s(m_i,m_t)+(1-\epsilon)f_b(m_i)\right). \tag{7.31}$$

7.2.2 Fit Procedure

The previous section derived the probability $p(D|n_s,n_b,m_t)$ of seeing the observed data given specific values for the parameters n_s, n_b, and m_t. However, what is actually needed is the inverse of this, the joint probability distribution for the parameters given the data D. This is readily obtained by applying Bayes' Theorem:

$$p(n_s,n_b,m_t|D) = \frac{p(D|n_s,n_b,m_t)\,p(n_s,n_b,m_t)}{p(D)}. \tag{7.32}$$

The factor in the denominator here does not depend on the parameters being estimated, so it can be ignored. The prior probability $p(n_s, n_b, m_t)$ is assumed to factor:

$$p(n_s, n_b, m_t) = p(n_s)\, p(n_b)\, p(m_t). \tag{7.33}$$

The priors for n_s and m_t are taken to be uniform, to express complete ignorance about their values[1] [143]. However, an estimate for the expected background n_b is available from the counting experiment, so the prior probability for n_b is taken to be a gaussian, with the mean and width taken from the counting experiment background calculation.

The estimate of the parameters is made using a maximum likelihood fit. Given the data D and the expected background $p(n_b)$, the parameters are varied until the function $p(n_s, n_b, m_t|D)$ is maximized. An estimate of the error on the parameters can then be made by finding how far from the maximum one must go for the natural logarithm of the likelihood function to change by 0.5. (This prescription for the error is exact if the likelihood function is gaussian. Under fairly general conditions, a likelihood function will approach a gaussian in the limit of large statistics; however, this is usually a good approximation even with limited statistics as long as the likelihood has a single maximum and its logarithm is roughly parabolic. See [144, sec. 9.7.1].)

The actual maximization is performed using the MINUIT package [145].

[1] If one wished to include the standard model cross section in the mass determination, one could define an expected number of signal events as a function of top mass $\bar{n}_s = \sigma_{t\bar{t}}(m_t) \int \mathcal{L}\, dt$, assign it an error $\sigma(\bar{n}_s)$, and then take the prior probability for n_s and m_t to be the gaussian form $p(n_s, m_t) = g(n_s; \bar{n}_s, \sigma(\bar{n}_s))$.

The MIGRAD minimizer is used, and MINOS is used for the error estimates.

7.2.3 Multiple Channels

Finally, consider the modifications which would have to be made to the fit procedure in order to handle multiple channels. This is not used for the final result (due to low Monte Carlo statistics in some of the individual channels), but it's interesting to see how it could be done.

Let there be M channels, numbered $1 \ldots M$. Each channel has its own signal and background cross sections, n_s^j and n_b^j. Let D^j be the set of N^j data events in channel j, and let $C(i)$ give the channel for data event i. Also, call the individual signal and background resolution functions for each channel f_s^j and f_b^j. Then, the combined likelihood can be written as the product of individual likelihoods for each channel:

$$p(D|n_s^1 \ldots n_s^M, n_b^1 \ldots n_b^M, m_t) = \prod_j p(D^j|n_s^j, n_b^j, m_t). \tag{7.34}$$

One could do the fit like this, fitting for each of the n_s^j and n_b^j variables. However, in order to keep the problem more manageable, an additional assumption can be made. This is that the cross sections for each channel are related by known constant factors. That is, instead of fitting for the $2M$ variables $n_s^1 \ldots n_s^M$ and $n_b^1 \ldots n_b^M$, fit instead only for the two variables n_s and n_b, where

$$\begin{aligned} n_s^j &= \epsilon_s^j n_s \\ n_b^j &= \epsilon_b^j n_b \end{aligned} \tag{7.35}$$

and

$$\sum_j \epsilon_s^j = \sum_j \epsilon_b^j = 1. \tag{7.36}$$

The constants ϵ_s^j and ϵ_b^j are determined using Monte Carlo.

With these definitions, the combined likelihood in equation (7.34) can be rewritten using equations (7.27) and (7.28) as

$$\begin{aligned} & p(D|n_s^1 \ldots n_s^M, n_b^1 \ldots n_b^M, m_t) \\ & = \left[\prod_{j=1}^{M} \frac{q(N^j, \epsilon_s^j n_s + \epsilon_b^j n_b)}{\left(\epsilon_s^j n_s + \epsilon_b^j n_b\right)^{N^j}} \right] \prod_{i=1}^{N} \left(\epsilon_s^{C(i)} n_s f_s^{C(i)}(m_i, m_t) + \epsilon_b^{C(i)} n_b f_b^{C(i)}(m_i) \right) \\ & = q(N, n_s + n_b) \binom{N}{N^1 \ldots N^M} \prod_{i=1}^{N} \frac{\epsilon_s^{C(i)} n_s f_s^{C(i)}(m_i, m_t) + \epsilon_b^{C(i)} n_b f_b^{C(i)}(m_i)}{n_s + n_b}. \end{aligned} \tag{7.37}$$

7.3 Monte Carlo Tests

The maximum likelihood fit and signal interpolation procedures can be tested using Monte Carlo. Samples were prepared consisting of signal and background (VECBOS) events satisfying the loose cuts, with all channels mixed together in their natural proportions. Here, an 'event' is represented simply as the average fitted mass for that event. The signal and background samples were mixed together in a 1:1 ratio (about what is expected for the loose cuts). These mixed samples were then grouped together to form an ensemble of fixed-size experiments. Monte Carlo statistics were extended through resampling [146], where the samples were used several times, but were shuffled before each reuse.

Each experiment was then fit using the method of the previous section.

For these tests, the background constraint was taken to be half the sample size, with an error of 50%. The results are shown in Figure 7.3 for top masses of 160 and 200 GeV/c^2, and for sample sizes of $N = 24$ and $N = 200$. The results of the fit are seen to reproduce the input mass to within a few GeV/c^2 (< 3 GeV/c^2 here; the worst cases seen in other tests for large sample sizes were $\lesssim 4$ GeV/c^2). The expected statistical error of a measurement is given by the RMS values of these distributions. For $N = 24$ (which the same size as the observed data), this is about ± 21 GeV/c^2. For $N = 200$, the RMS is about ± 7.8 GeV/c^2, indicating that the error is getting smaller with increasing N by slightly less than $1/\sqrt{N}$. Figure 7.4 shows the distribution of the the errors deduced for each experiment for the $N = 24$, $m_t = 200$ GeV/c^2 case. (Since the errors are, in general, asymmetric, what is actually plotted is the mean of the high and low side errors.) The values found here are close to, but slightly less than, the RMS width of the mass plot in Figure 7.3(b).

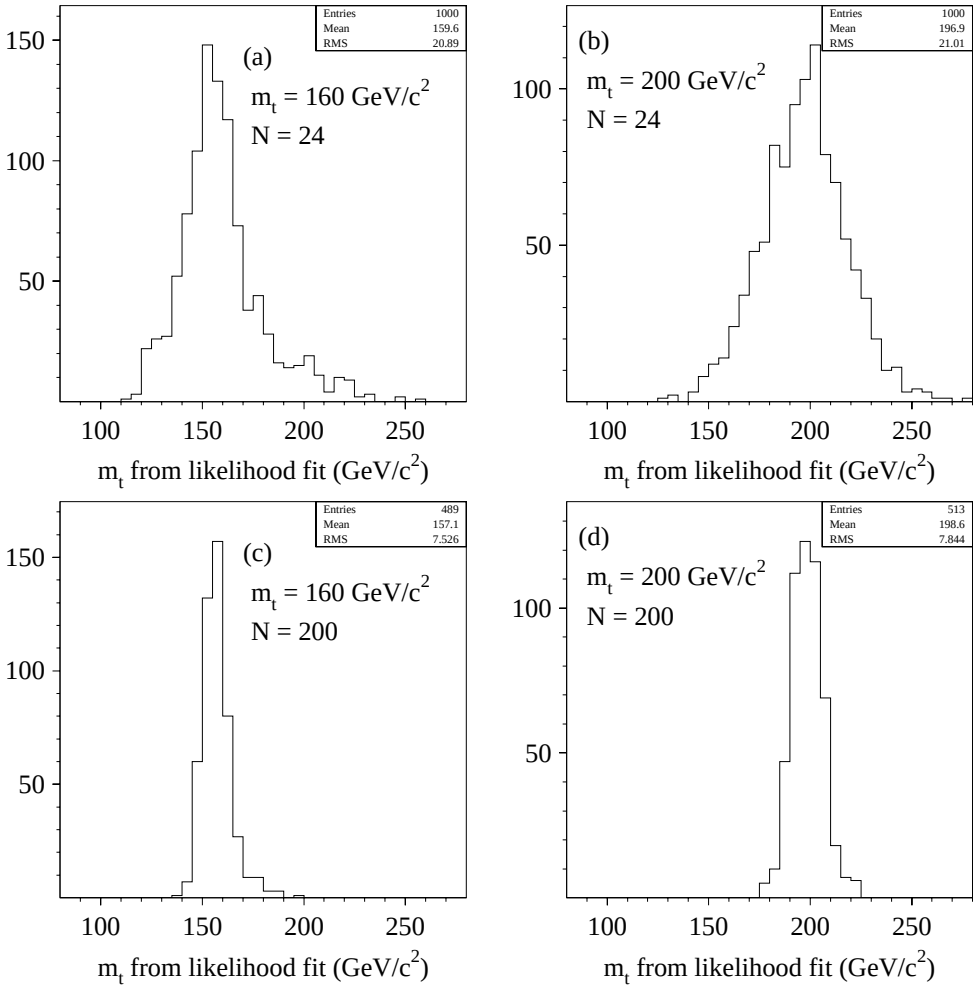

Figure 7.3: Results from Monte Carlo ensemble tests of the likelihood fit method. (a) $N = 24$ and $m_t = 160\,\mathrm{GeV}/c^2$. (b) $N = 24$ and $m_t = 200\,\mathrm{GeV}/c^2$. (c) $N = 200$ and $m_t = 160\,\mathrm{GeV}/c^2$. (d) $N = 200$ and $m_t = 200\,\mathrm{GeV}/c^2$.

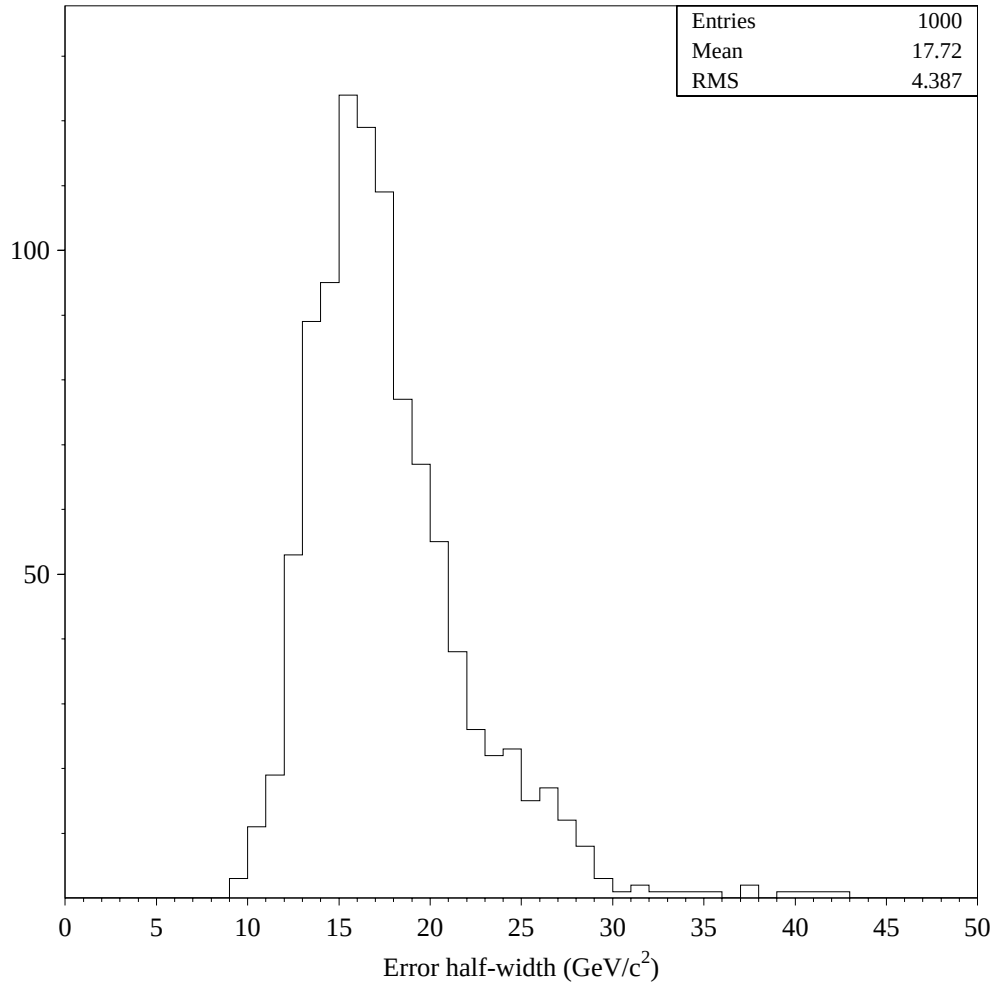

Figure 7.4: Half-width of error interval on m_t from Monte Carlo ensemble tests of the likelihood fit method, with $N = 24$ and $m_t = 200\,\mathrm{GeV}/c^2$.

Chapter 8

Data Analysis

The final step is to take the tools from the last two chapters and apply them to the real data sample.

8.1 Final Background Model

The results of the counting experiment summarized in Chapter 5 gave the number of background events expected in the standard and loose samples, broken down by channel and into W + jets and QCD contributions.

However, this is not the correct background to use for the mass fitting analysis, since additional cuts are made; namely, that there be at least four (0.3 cone) jets and at least one jet permutation with a good fit (enforcing the b-tag, if one is present). The expected background must be scaled down by the efficiencies for these cuts.

The VECBOS Monte Carlo was used to model the W+jets background. To model the QCD background, a sample was prepared consisting of events which

passed the electron triggers but failed the subsequent offline electron-ID cuts. The electromagnetic object in these events was then treated as an electron for the remainder of the analysis. All other cuts were applied, except that for the QCD sample with the standard cuts, the missing E_T cut was lowered to 10 GeV in order to get reasonable statistics.

The results of the efficiency calculation are summarized in Table 8.1. The efficiencies for the corresponding $e+$jets and $\mu+$jets channels are very similar, so they are averaged together. The errors on the efficiencies come from Monte Carlo statistics plus a systematic component which was guessed to be $\approx 5\%$. The final estimated backgrounds to the fitted samples are 11.58 ± 2.19 events for the loose cuts and 2.08 ± 0.41 events for the standard cuts. About 30% of the expected background is due to QCD, with the remainder from $W+$jets.

The shapes of the final fitted mass distributions for the background are shown in Figure 8.1 for the loose and standard cuts.

8.2 Fitting the Data

Now for the data. A total of 14 events pass the standard cuts, 11 of which have a good fit. With the loose cuts, 29 events pass the selection, 24 of which have a good fit. The fitted mass distributions for these events are shown in Figure 8.2. (The results are presented in numeric form in Appendix B.) The results of the maximum likelihood fits are also shown in the figure. The fitted values of the parameters are as follows.

Backgrounds from Chapter 5

		Loose	Standard
W + jets	untagged	12.14 ± 2.89	1.46 ± 0.41
	tagged	1.17 ± 0.19	0.63 ± 0.14
QCD	untagged	3.58 ± 1.09	0.47 ± 0.20
	tagged	0.99 ± 0.16	0.65 ± 0.10

Fitting efficiencies

		Loose	Standard
W + jets	untagged	$64.6 \pm 5.4\%$	$82.2 \pm 8.5\%$
	tagged	$30.5 \pm 7.6\%$	$38.3 \pm 10.3\%$
QCD	untagged	$82.8 \pm 5.8\%$	$71.2 \pm 11.1\%$
	tagged	$42.2 \pm 6.9\%$	$46.4 \pm 10.4\%$

Backgrounds after fitting

		Loose	Standard
W + jets	untagged	7.84 ± 1.98	1.20 ± 0.36
	tagged	0.36 ± 0.11	0.24 ± 0.09
	total	8.20 ± 1.98	1.44 ± 0.37
QCD	untagged	2.96 ± 0.92	0.34 ± 0.15
	tagged	0.42 ± 0.10	0.30 ± 0.08
	total	3.38 ± 0.93	0.64 ± 0.17

Total background

	Loose	Standard
total background	11.58 ± 2.19	2.08 ± 0.41
QCD fraction	0.29 ± 0.08	0.31 ± 0.08

Table 8.1: Summary of final background calculation.

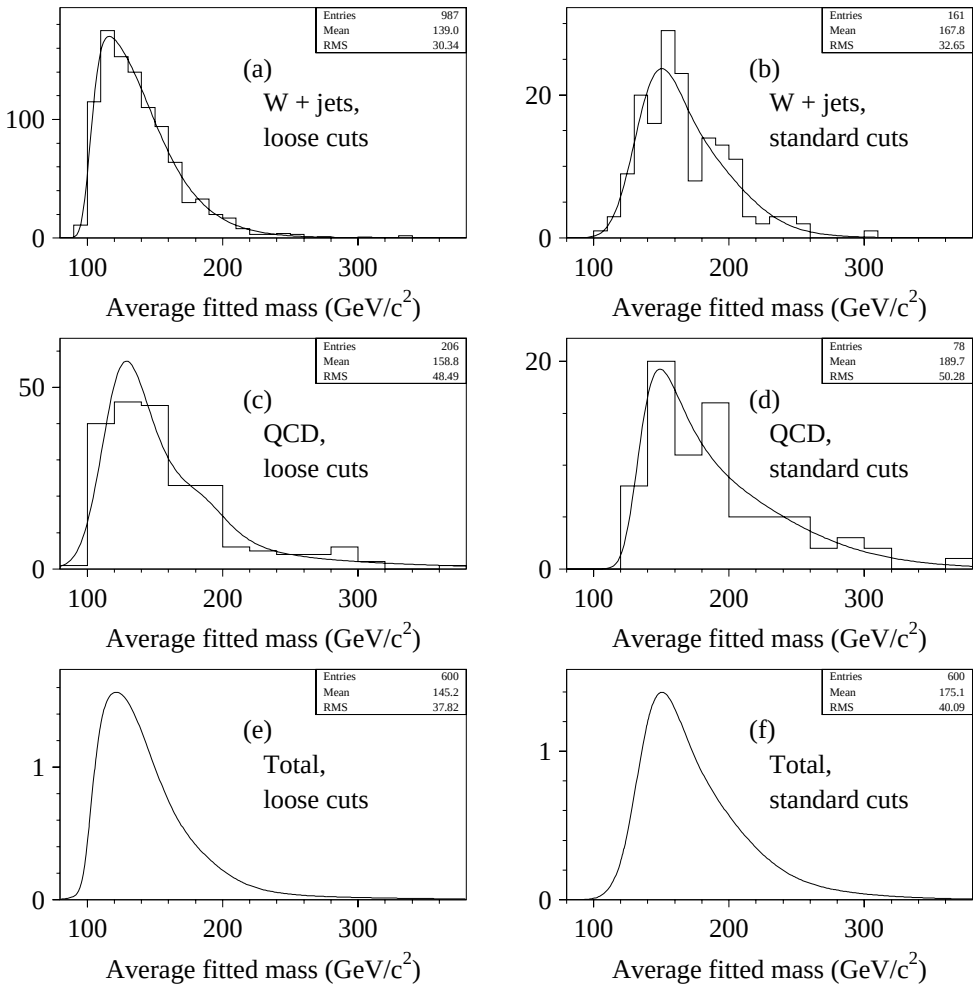

Figure 8.1: Final background shapes for mass fitting, for loose and tight cuts. (a), (b) W + jets background. (c), (d) QCD fake background. (e), (f) Total background (arbitrary normalization).

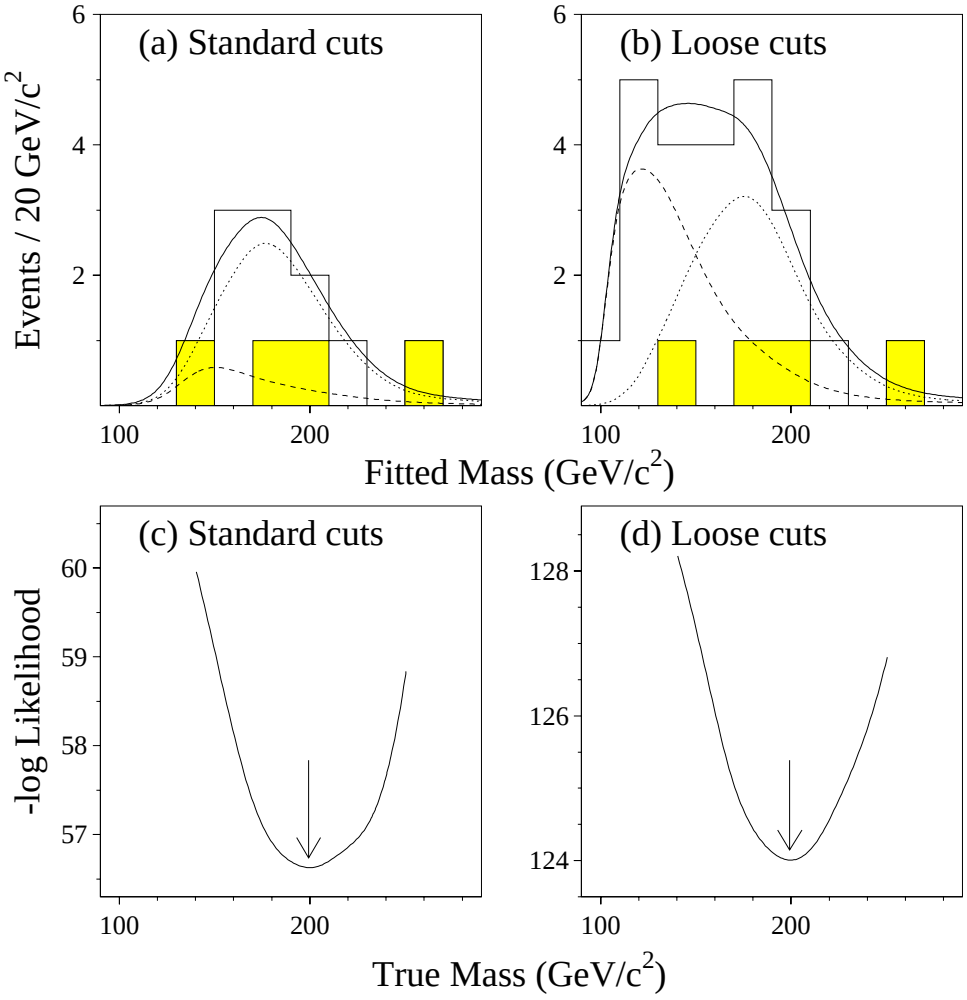

Figure 8.2: Results of fitting the candidate samples. (a), (b) Fitted mass distributions for the data samples and the results of the likelihood fit, for the standard and loose cuts. The dotted line is the fitted background, the dashed line is the signal, and the solid line is the sum of the two. The shaded histogram shows the tagged events. (c), (d) $-\ln$ Likelihood as a function of top mass, for the two sets of cuts. The arrows indicate the location of the minima.

Loose cuts	Standard cuts
$m_t = 199.3^{+19.1}_{-21.2}$	$m_t = 199.4^{+30.9}_{-25.3}$
$n_s = 12.3^{+5.0}_{-4.2}$	$n_s = 9.0^{+3.7}_{-3.0}$
$n_b = 11.6^{+2.0}_{-2.0}$	$n_b = 2.1^{+0.4}_{-0.4}$

For both sets of cuts, the observed data are fit well by a combination of a top signal with $m_t = 200\,\mathrm{GeV}/c^2$ and the predicted background. (Kolmogorov test probability [141, sec. 6.2] > 0.99 for both cases.) Another check of the background calculation is to redo the fit without the gaussian constraint on n_b (i.e., a uniform prior); the value of n_b which results from the fit is then a background estimate which is independent of that made in the previous section. When that is done for the loose sample, the result for the top mass changes by less than $1\,\mathrm{GeV}/c^2$, while the best-fit value for n_b is $11.8^{+5.8}_{-5.2}$. This is in quite good agreement with the counting experiment estimate of $n_b = 11.6 \pm 2.2$. (This cannot be done for the sample selected with the standard cuts, since in that case, there is not much difference between the shapes of the signal and background curves. When a fit is attempted with no background constraint, n_b goes to zero.)

Certain other tests have been carried out. If the analysis is repeated using signal resolution functions derived using HERWIG instead of ISAJET, the result for the top mass is $m_t = 195.7^{+15.6}_{-17.0}\,\mathrm{GeV}/c^2$ with the loose cuts and $m_t = 195.6^{+23.4}_{-22.4}\,\mathrm{GeV}/c^2$ with the standard cuts. The difference in the sizes of the error intervals between HERWIG and ISAJET is due mainly to the fact that it is easier to find the correct jet permutation with HERWIG (see Table 6.1). As a consequence, resolution functions generated using HERWIG tend to track

the input top mass better than those from ISAJET: if one plots the location of the peak of each resolution curve versus the input top mass, the slope is about 0.70 for ISAJET (loose cuts) compared to about 0.77 for HERWIG.

Varying the QCD fraction used in the fit within its errors changes the result by less than $2\,\mathrm{GeV}/c^2$.

If the event with the largest fitted mass ($266\,\mathrm{GeV}/c^2$) is removed from the sample and the fit redone (with the same background), the results are $m_t = 197.1^{+16.6}_{-22.0}\,\mathrm{GeV}/c^2$ for the loose cuts and $m_t = 197.2^{+21.9}_{-28.8}\,\mathrm{GeV}/c^2$ for the standard cuts. Note that this single event has a large effect on the high-side error, especially with the standard cuts.

8.3 Significance Tests

This section considers the question of the extent to which the mass fit information can be used to strengthen the claims for the existence of top.

8.3.1 Confidence Limits

One approach is to define a 'confidence limit' by analogy to the conventional χ^2 test. A χ^2 test is applicable to the case where one has a collection of N normally distributed variables x_i with means $\bar{x}_i$ and widths σ_i. Call this hypothesis H. One defines the function

$$\chi^2 = \sum_i \left(\frac{x_i - \bar{x}_i}{\sigma_i} \right)^2 . \tag{8.1}$$

Then, given some particular measurement $x_1 \ldots x_N$ (call it D), one can test how consistent it is with H by asking how likely it would be, assuming H, to get a measurement with a larger χ^2 than the one actually seen. That is, one evaluates the integral

$$\int_{\chi^2(D') \geq \chi^2(D)} p(D'|H)\, dD'. \tag{8.2}$$

This is in some sense the probability for the presumed model H to fluctuate to give the observed data D. (Note that it is *not* the probability that H is true; that is not well-defined unless one specifies the complete set of alternatives to H.)

In order to generalize this procedure for other types of distributions, note that the probability for observing a particular measurement D assuming H is just

$$p(D|H) = e^{-\chi^2(D)/2}. \tag{8.3}$$

This suggests that one can obtain an analogous significance for a problem with an arbitrary likelihood $p(D|H)$ by computing the integral

$$\int_{p(D'|H) \leq p(D|H)} p(D'|H)\, dD'. \tag{8.4}$$

That is, by computing the total probability of all possible data samples which have a lower probability than the one actually observed.

For the mass fitting problem, the hypothesis to test is that the data are described entirely by the background model. The appropriate likelihood is thus obtained by setting $n_s = 0$ in equation (7.28), yielding

$$p(D|n_b, H) = q(N, n_b) \prod_{i=1}^{N} f_b(m_i). \tag{8.5}$$

The remaining parameter n_b is then integrated out:

$$p(D|H) = \int p(D, n_b|H)\, dn_b = \int p(D|n_b, H)\, p(n_b)\, dn_b. \tag{8.6}$$

The prior $p(n_b)$ is again taken to be a gaussian.

The integral over the data space can then be written

$$\int dD' = \sum_{N'=0}^{\infty} \prod_{j=1}^{N'} \int dm'_j. \tag{8.7}$$

Note that if f_b is uniform, this prescription yields the same result as was used for the counting experiment (equation (5.12)). Strictly speaking, this is true only if N' is restricted to be larger than N; this will make a difference only in cases where the expected background is not small in comparison to the number of observed events. This is because the prescription developed here tests the consistency of D with H regardless of the direction of any disagreement, while (5.12) counts only upward fluctuations in the number of events. For example, consider some hypothetical experiment where H predicts that 100 events should be expected, but 1000 events are actually observed. Both methods would assign a small probability to this occurrence. However, if 100 events are expected, it is also quite unlikely to see zero events. The prescription developed here will also assign a small probability to this latter case; however, the counting experiment significance (5.12) would assign it a probability of 1.

The integral in equation (8.4) can be evaluated by Monte Carlo techniques. An outline of a procedure for doing so is as follows.

1. Before starting, evaluate the likelihood for the data point being tested, $p(D|H)$.

2. Loop over the number of events in the Monte Carlo ensemble, N'.

3. Evaluate the piece of the likelihood which depends only on the number of events:

$$R(N') = \int_0^\infty q(N', n_b)\, p(n_b)\, dn_b. \tag{8.8}$$

   Define a probability threshold τ by $\tau = p(D|H)/R(N')$.

4. Generate a large number N_0 of N' event experiments, picking each mass m'_j from the background probability distribution f_b. This forms a set of samples D'. For each of these samples, compute the remaining likelihood factor $\prod_j f_b(m'_j)$, and count the number of times that this is less than the threshold τ. Call this $N_<$.

5. The contribution to the significance is then $R(N')N_</N_0$. Return to step 2, and continue looping until the terms being summed become insignificantly small.

The results of this calculation are 1.6×10^{-3} for the loose cuts, and 1.9×10^{-5} for the standard cuts. If the calculation is repeated with f_b taken to be uniform (i.e., using only counting information), the results are 5.5×10^{-3} for the loose cuts and 3.7×10^{-5} for the standard cuts. (If the counting experiment prescription of equation (5.12) were used instead, the result is unchanged for the standard cuts, but goes down to 4.7×10^{-3} for the loose cuts.)

It is also interesting to try to construct a significance which uses only the shapes of the distributions and which does not depend on the scale of n_b. This can be done by fixing $N' = N$ and taking the likelihood to be simply $p(D|H) = \prod_i f_b(m_i)$. The results from this are 0.06 for the loose cuts and 0.30 for the standard cuts.

8.3.2 Likelihood Ratio

Another approach for evaluating significance is to directly compare the likelihoods for two competing hypotheses [143, 146]. Let the first hypothesis, H_b, be that the data are due entirely to background, with its cross section calculated from the counting experiment. Let the second hypothesis, H_t, be that there is also top in the data, with a mass and cross section given by the maximum likelihood fit. Then, by Bayes' theorem, the likelihood ratio is

$$R = \frac{p(H_b|D)}{p(H_t|D)} = \frac{p(H_b)}{p(H_t)} \times \frac{p(D|H_b)}{p(D|H_t)}. \tag{8.9}$$

Note that

$$p(H_b|D) = \frac{R}{1+R}, \tag{8.10}$$

and

$$p(H_t|D) = \frac{1}{1+R}. \tag{8.11}$$

The individual likelihoods are then

$$p(D|H_b) = \int p(D|n_b, H_b)\, p(n_b)\, dn_b \quad \text{and} \tag{8.12}$$

$$p(D|H_t) = \int p(D|n_s, n_b, m_t, H_t)\, p(n_s)\, p(n_b)\, p(m_t)\, dn_s\, dn_b\, dm_t. \tag{8.13}$$

The factor $p(D|n_s, n_b, m_t, H_t)$ is given by equation (7.28). The corresponding factor for H_b, $p(D|n_b, H_b)$, is the same thing with n_s set to zero:

$$p(D|n_b, H_b) = q(N, n_b) \prod_{i=1}^{N} f_b(m_i). \tag{8.14}$$

The priors for the parameters are taken to be gaussians. The result of the background calculation is used for $p(n_b)$, while $p(n_s)$ and $p(m_t)$ are taken from the result of the maximum likelihood estimation. For definiteness, the priors $p(H_b)$ and $p(H_t)$ are taken to be equal ($p(H_b) = p(H_t) = 1/2$).

The results of this calculation are $R = 4.4 \times 10^{-4}$ for the loose cuts, and $R = 1.1 \times 10^{-4}$ for the standard cuts. For R this small, $p(H_b|D) \approx R$, so the interpretation is that the data are very unlikely to be due entirely to background. As a check, the calculation was redone using the loose cut parameters on a sample consisting of 12 (VECBOS) background events. The result was $R = 1.03$, or $p(H_b|D) \approx 50\%$.

These results complement the confidence limit results of the previous section. Both indicate that the data are very unlikely to be entirely due to the known backgrounds. The confidence limit result, however, does not say anything about the plausibility of the top hypothesis; the likelihood ratio test shows that the top hypothesis is indeed much more plausible than the background-only hypothesis.

8.4 Systematic Error

This section discusses the systematic errors on the top mass determination.

8.4.1 Jet Scale Error

The dominant systematic error is that due to the uncertainty in the jet energy scale, including the out-of-cone corrections. As explained in Section 6.2.1, an upper limit on the jet scale error is about 10%. For the analysis in [1], the error in the final top mass due to this uncertainty was estimated as follows. In a $t\bar{t}$ ISAJET Monte Carlo sample with $m_t = 160\,\mathrm{GeV}/c^2$, the energies of all the jets were varied up and down by 10%. This resulted in a 7.7% shift in the mean fitted mass (using a slightly different fitting program than that described here) [137]. This number was then divided by 0.62, the slope of the relation between input mass and average fitted mass (see Figure 6.22). This yielded an error of 10.8%. A further uncertainty of about $\pm 5\,\mathrm{GeV}/c^2$ for the difference in event generators was then folded in, giving the quoted error of $\pm 22\,\mathrm{GeV}/c^2$ for a central value of $199\,\mathrm{GeV}/c^2$.

However, a better estimate of the jet scale systematic error can be made by redoing the ensemble likelihood fits (see Section 7.3) using both the $m_t = 200\,\mathrm{GeV}/c^2$ signal sample (which is much closer to the experimental value) and the background sample with the jet energies scaled up and down by 10%. The resulting fitted mass distributions for the signal and background are shown in Figure 8.3. Here, the 10% change in jet scale changes the mean fitted mass

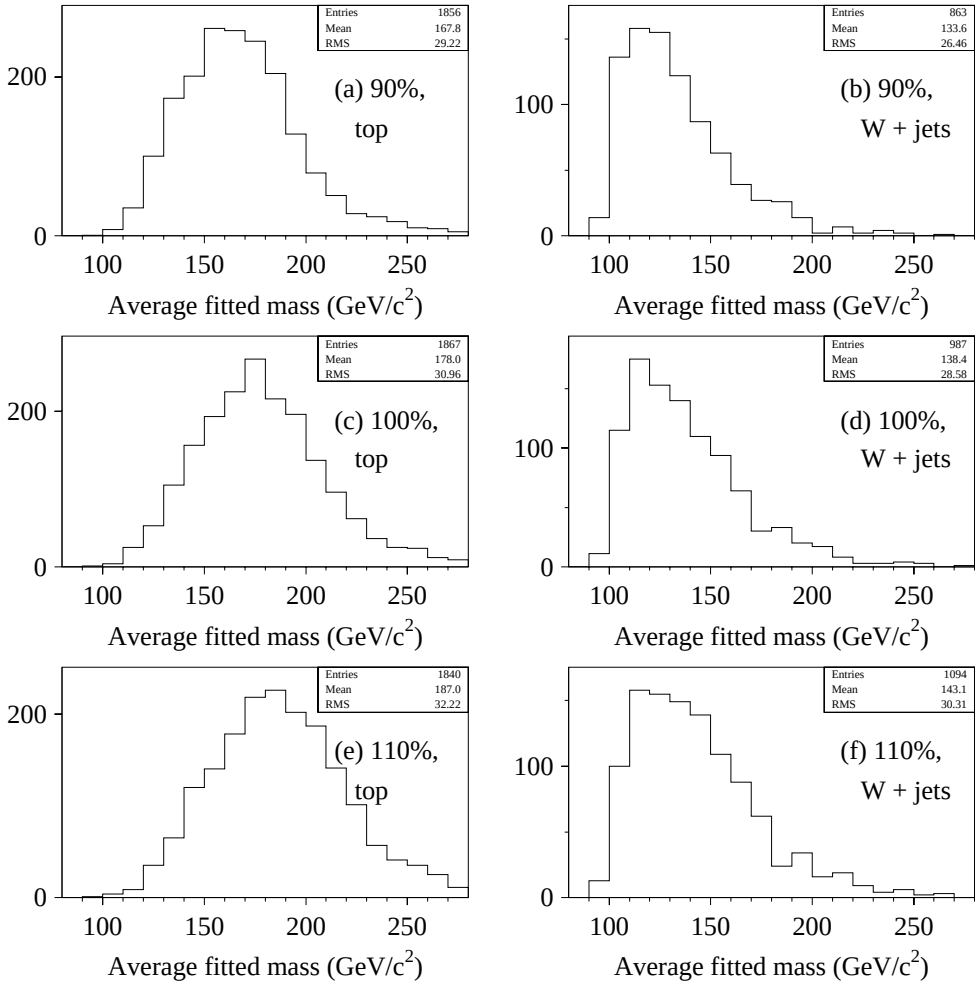

Figure 8.3: Fitted mass distributions for ISAJET $t\bar{t}$ and VECBOS W + jets Monte Carlos, with $m_t = 200\,\mathrm{GeV}/c^2$, for different jet scalings. All channels are summed together. (a), (b) Jet energies scaled by 90%. (c), (d) Jet energies unscaled. (e), (f) Jet energies scaled by 110%.

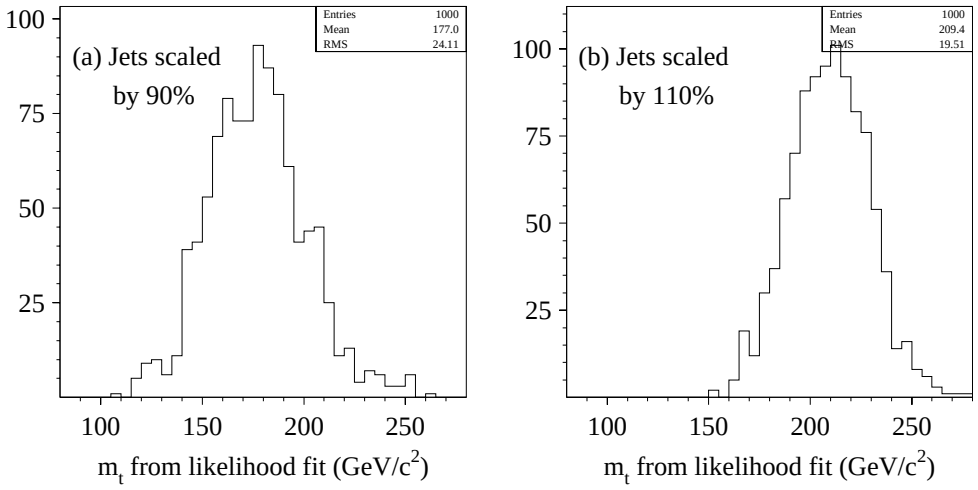

Figure 8.4: Results from Monte Carlo ensemble tests with differing jet scales. For $N = 24$ and $m_t = 200\,\mathrm{GeV}/c^2$ (ISAJET). (a) Jets scaled by 90%. (b) Jets scaled by 110%.

by about $\pm 5\%$. These samples were then gathered together in groups of 24, with a 1:1 signal/background ratio, and passed through the likelihood fit as in Section 7.3. The results are shown in Figure 8.4; they should be compared with Figure 7.3(b). The resulting error estimate is asymmetrical, and is about $^{+12}_{-20}\,\mathrm{GeV}/c^2$.

There are several factors which contribute to the asymmetry in this error. First, a careful examination of Figure 8.3 shows that the fitted mass shifts slightly more for the 90% jet scaling than for the 110% scaling (about $1\,\mathrm{GeV}/c^2$ more). Second, the relation between true mass and fitted mass produced by the interpolation procedure does not necessarily have a constant slope. In fact, for the resolution functions used, it is somewhat steeper in the range of true masses from 180 to $200\,\mathrm{GeV}/c^2$ than it is in range from 200 to $240\,\mathrm{GeV}/c^2$. For a true mass around $200\,\mathrm{GeV}/c^2$, this tends to make the errors on the low side

larger than on the high side. This is exacerbated by the fact that the mean fitted mass for the $m_t = 200\,\mathrm{GeV}/c^2$ ISAJET Monte Carlo sample lies somewhat below the trend averaged over all the Monte Carlo samples (see Figure 6.22). Note that the statistical error shows a similar asymmetry when the one event with an unusually large fitted mass is removed (see previous section). This effect should be reduced in the future with better Monte Carlo statistics.

If this calculation is redone using resolution functions derived from HERWIG instead of ISAJET, the result is a systematic error of about $^{+11}_{-17}\,\mathrm{GeV}/c^2$.

8.4.2 Other Systematic Errors

In comparison to the errors from the jet scale, other potential sources of systematic error are relatively small. Some of these sources are listed below.

- The $\approx 4\,\mathrm{GeV}/c^2$ difference seen between ISAJET and HERWIG in Section 8.2 is taken as an estimate of the systematic error due to the event generator.

- As mentioned in Section 7.3, there are small offsets after the likelihood fit, with an upper limit of about 3–$4\,\mathrm{GeV}/c^2$.

- The uncertainty due to the QCD fraction in the background is less than $2\,\mathrm{GeV}/c^2$ (Section 8.2).

- Other systematic errors, such as errors caused by folding all channels together and additional systematics in the background shape, are guessed to be $\lesssim 5\,\mathrm{GeV}/c^2$.

Summed together in quadrature, these yield about $7.8\,\mathrm{GeV}/c^2$. Combined with the jet scale systematic error, this gives a total systematic error of about $^{+14}_{-21}\,\mathrm{GeV}/c^2$. If HERWIG were used instead, the result would be $^{+13}_{-19}\,\mathrm{GeV}/c^2$.

8.5 Consistency with the Standard Model

A quick test was made to assess the consistency of the combined counting experiment and mass measurement with the top production cross section of [17]. Many simulated experiments were generated by varying the measured and theoretical parameters within errors and then picking a number of observed signal and background events from appropriate Poisson distributions. The fraction of such experiments which yielded ≥ 17 observed events is then a measure of the consistency of the measured results with the standard model cross section.

For this calculation, the standard cuts were used, and all seven channels were combined together. The luminosities from the channels were averaged to yield $47.4 \pm 5.7\,\mathrm{pb}^{-1}$ (with the standard 12% luminosity error). The backgrounds were summed over all channels; this was then divided by the $47.4\,\mathrm{pb}^{-1}$ luminosity to define an effective background cross section of $0.080 \pm 0.012\,\mathrm{pb}$. The values of efficiency times branching ratio for the signal were summed across all channels for each top mass. The results are tabulated below.

m_t (GeV/c^2)	$\varepsilon \times \mathcal{B}(\%)$
140	1.73 ± 0.16
160	2.60 ± 0.20
180	3.42 ± 0.37
200	4.43 ± 0.32

The relative errors for these four points were averaged to obtain a mean relative error of 8.7%.

The detailed procedure used is a follows. For each simulated event:

1. Choose a luminosity $\mathcal{L}$ from the gaussian distribution $47.4 \pm 5.7\,\mathrm{pb}^{-1}$.

2. Choose a background cross section from the gaussian distribution $0.080 \pm 0.012\,\mathrm{pb}$. Multiply this with the luminosity $\mathcal{L}$ to obtain the expected number of background events n_b.

3. Choose a top mass m_t from an asymmetric gaussian distribution with mean $199\,\mathrm{GeV}/c^2$. For the 50% of trials which are below the mean, take the width to be $30\,\mathrm{GeV}/c^2$; for the upper half, take the width to be $24\,\mathrm{GeV}/c^2$.

4. Convert the top mass m_t to a cross section by interpolating between the values from [17]. Smear the resulting value with a gaussian error of 30%.

5. Interpolate on the top mass in the above table of top efficiencies. For $m_t > 200\,\mathrm{GeV}/c^2$ use the $200\,\mathrm{GeV}/c^2$ point, and similarly for $m_t < 140\,\mathrm{GeV}/c^2$. Smear this with a gaussian error of 8.7%. Multiply together the top cross section, top efficiency, and luminosity $\mathcal{L}$ to obtain the expected number of signal events n_s.

6. If either n_s or n_b is nonpositive, reject this experiment and go on to the next (less than 0.1% of experiments were thus rejected).

7. Pick the number of background events from a Poisson distribution with mean n_b. Do the same for the number of signal events (using n_s).

8. Add the number of signal and background events to obtain the total number of events N and histogram the result.

To obtain the final result, sum the contents of the histogram bins for $N \geq 17$ and divide by the total number of accepted experiments. Out of a total of 10000 experiments which were generated, 9992 were accepted, and 891 had $N \geq 17$. Thus, assuming the standard model cross section and the measured mass, the probability of seeing at least 17 events is 8.9%.

A similar calculation was made in [147], except that all the channels were kept separate and allowed to vary independently. The result from that calculation for the probability of seeing at least 17 events was 6.6%.

Chapter 9

Summary and Conclusions

As summarized in Chapter 5, DØ observes a highly significant excess of events over background, with a significance of 2×10^{-6} (4.6σ). As demonstrated in Chapter 8, this excess is consistent with standard model top production with $m_t \approx 200\,\mathrm{GeV}/c^2$.

For the final mass result, the loose cuts and the ISAJET resolution functions are used. The measured top mass is then $199^{+19}_{-21}(\mathrm{stat.})^{+14}_{-21}(\mathrm{syst.})\,\mathrm{GeV}/c^2$. (The result reported in [1] was $199^{+19}_{-21} \pm 22\,\mathrm{GeV}/c^2$; the difference is the new systematic error estimate of Section 8.4.) If the HERWIG resolution functions are used instead of the ISAJET ones, the result is $196^{+16}_{-17}(\mathrm{stat.})^{+13}_{-19}(\mathrm{syst.})\,\mathrm{GeV}/c^2$.

For the event samples used for the mass determination, the significance of the result with the standard cuts is 1.9×10^{-5} if the mass information is used and 3.7×10^{-5} if it is not. The corresponding numbers for the loose cuts are 1.6×10^{-3} and 5.5×10^{-3}, respectively.

There are numerous ways in which this measurement can be improved in the future. Most obviously, the remainder of run 1B will bring in significantly

more statistics. This will help not only directly, in increasing the size of the candidate sample, but also indirectly, by decreasing the statistical error in the control samples used for jet calibration and background determination.

Changes to the fitting procedure will also improve this result. Different jet algorithms are being investigated; some (such as the nearest neighbor clustering algorithm) show some promise for improving the ultimate resolution. The methodology used for handling the jet corrections could be improved, as could the way in which extra jets are handled. Finally, increased Monte Carlo statistics will improve the accuracy of the final likelihood fit.

Appendix A

Constrained Fit Details

This appendix discusses the mathematical details of the constrained kinematic fit.

A.1 Basic Fitting Algorithm

The problem to be solved is this: given a set of measurements $\mathbf{x}^m$ and their error matrix $G_{ij}^{-1} = \langle \delta x_i^m \delta x_j^m \rangle$, what is the smallest change one can make to $\mathbf{x}$ to make them satisfy a given set of constraints $F_i(\mathbf{x}) = 0$?

That is, minimize the quantity $\chi^2 = (\mathbf{x} - \mathbf{x}^m)^T \mathbf{G} (\mathbf{x} - \mathbf{x}^m)$, subject to the constraints $\mathbf{F}(\mathbf{x}) = 0$. The algorithm, which is based on SQUAW kinematic fitting program [148, 149], is developed in this section.

A.1.1 Definitions

Divide the set of variables into 'well-measured' and 'poorly measured' sets. Let $\mathbf{x}$ and $\mathbf{y}$ be column vectors of the well-measured and poorly-measured

variables, respectively. (Whole vectors and matrices are denoted by **boldface** symbols, and their elements by subscripted plain symbols.) Denote the measured values of the variables by $\mathbf{x}^m$, $\mathbf{y}^m$, the 'true' values of the variables by $\mathbf{x}^t$, $\mathbf{y}^t$, their values after iteration ν by $\mathbf{x}^\nu$, $\mathbf{y}^\nu$, and their values at the end of the fit by $\mathbf{x}^f$, $\mathbf{y}^f$. The error matrices are given by

$$\begin{aligned} G_{ij}^{-1} &\equiv \langle \delta x_i^m \delta x_j^m \rangle \\ Y_{ij}^{-1} &\equiv \langle \delta y_i^m \delta y_j^m \rangle, \end{aligned} \tag{A.1}$$

where it is tacitly assumed that $\langle \delta x_i^m \delta y_j^m \rangle = 0$. (Here, $\delta \mathbf{x}^m = \mathbf{x}^m - \mathbf{x}^t$.)

The fitting algorithm will be designed to reference only the arrays $\mathbf{G}^{-1}$ and $\mathbf{Y}$, not their inverses. This permits one to specify that a well-measured variable is constant by setting the appropriate diagonal element in $\mathbf{G}^{-1}$ to zero. One can also specify that a poorly measured variable is completely unknown by setting the appropriate element in $\mathbf{Y}$ to zero.

It is convenient to define the displacement vectors

$$\begin{aligned} \mathbf{c} &= \mathbf{x} - \mathbf{x}^m \\ \mathbf{d} &= \mathbf{y} - \mathbf{y}^m. \end{aligned} \tag{A.2}$$

The constraints which must be satisfied are given by the row vector

$$\mathbf{F}(\mathbf{x}^\nu, \mathbf{y}^\nu) = \mathbf{F}^\nu = 0 \tag{A.3}$$

with gradients

$$\begin{aligned} B_{\mathrm{x}i\lambda}(\mathbf{x}^\nu, \mathbf{y}^\nu) &= B^\nu_{\mathrm{x}i\lambda} = \frac{\partial F^\nu_\lambda}{\partial x_i} \\ B_{\mathrm{y}i\lambda}(\mathbf{x}^\nu, \mathbf{y}^\nu) &= B^\nu_{\mathrm{y}i\lambda} = \frac{\partial F^\nu_\lambda}{\partial y_i}. \end{aligned} \tag{A.4}$$

A column vector of Lagrange multipliers $\boldsymbol{\alpha}^\nu$ will also be needed.

The quantity which should be minimized subject to the constraints $\mathbf{F} = 0$ is

$$\chi^2 = \mathbf{c}^T\mathbf{G}\mathbf{c} + \mathbf{d}^T\mathbf{Y}\mathbf{d}. \tag{A.5}$$

A.1.2 Fit

In order to minimize the χ^2 subject to the constraints $\mathbf{F} = 0$, use the method of Lagrange multipliers and minimize the quantity

$$M = \mathbf{c}^T\mathbf{G}\mathbf{c} + \mathbf{d}^T\mathbf{Y}\mathbf{d} + 2\mathbf{F}\boldsymbol{\alpha}. \tag{A.6}$$

At the minimum, it must be the case that

$$0 = \frac{\partial M}{\partial \boldsymbol{\alpha}} = \mathbf{F} \tag{A.7}$$

$$0 = \frac{\partial M}{\partial \mathbf{x}} = \mathbf{G}\mathbf{c} + \mathbf{B}_\mathrm{x}\boldsymbol{\alpha} \tag{A.8}$$

$$0 = \frac{\partial M}{\partial \mathbf{y}} = \mathbf{Y}\mathbf{d} + \mathbf{B}_\mathrm{y}\boldsymbol{\alpha}. \tag{A.9}$$

Write (A.7) in terms of $\mathbf{x}^{\nu+1}$, $\mathbf{y}^{\nu+1}$ and linearize around $\mathbf{x}^\nu$, $\mathbf{y}^\nu$:

$$\begin{aligned} 0 &= \mathbf{F}^\nu + (\mathbf{x}^{\nu+1} - \mathbf{x}^\nu)^T\mathbf{B}_\mathrm{x}^\nu + (\mathbf{y}^{\nu+1} - \mathbf{y}^\nu)^T\mathbf{B}_\mathrm{y}^\nu \\ &= \mathbf{F}^\nu + (\mathbf{c}^{\nu+1} - \mathbf{c}^\nu)^T\mathbf{B}_\mathrm{x}^\nu + (\mathbf{d}^{\nu+1} - \mathbf{d}^\nu)^T\mathbf{B}_\mathrm{y}^\nu. \end{aligned} \tag{A.10}$$

Make the definitions

$$\mathbf{E} \equiv \mathbf{G}^{-1}\mathbf{B}_\mathrm{x} \tag{A.11}$$

$$\mathbf{H} \equiv \mathbf{E}^T\mathbf{B}_\mathrm{x} = \mathbf{B}_\mathrm{x}^T\mathbf{G}^{-1}\mathbf{B}_\mathrm{x} \tag{A.12}$$

$$\mathbf{r} \equiv \mathbf{c}^T\mathbf{B}_\mathrm{x} + \mathbf{d}^T\mathbf{B}_\mathrm{y} - \mathbf{F}. \tag{A.13}$$

Note that $\mathbf{H}$ is symmetric, $\mathbf{H}^T = \mathbf{H}$.

Then (A.10) yields

$$\mathbf{r}^\nu = \mathbf{c}^{\nu+1T}\mathbf{B}_\mathrm{x}^\nu + \mathbf{d}^{\nu+1T}\mathbf{B}_\mathrm{y}^\nu. \tag{A.14}$$

From (A.8) one gets

$$\mathbf{c}^{\nu+1} = -\mathbf{E}^\nu \boldsymbol{\alpha}^{\nu+1} \tag{A.15}$$

so

$$\mathbf{r}^\nu = -\boldsymbol{\alpha}^{\nu+1T}\mathbf{H}^\nu + \mathbf{d}^{\nu+1T}\mathbf{B}_\mathrm{y}^\nu \tag{A.16}$$

or

$$\mathbf{r}^{\nu T} = -\mathbf{H}^\nu \boldsymbol{\alpha}^{\nu+1} + \mathbf{B}_\mathrm{y}^{\nu T}\mathbf{d}^{\nu+1}. \tag{A.17}$$

Combining this with (A.9) yields

$$\begin{pmatrix} -\mathbf{H}^\nu & \mathbf{B}_\mathrm{y}^{\nu T} \\ \mathbf{B}_\mathrm{y}^\nu & \mathbf{Y} \end{pmatrix} \begin{pmatrix} \boldsymbol{\alpha}^{\nu+1} \\ \mathbf{d}^{\nu+1} \end{pmatrix} = \begin{pmatrix} \mathbf{r}^{\nu T} \\ 0 \end{pmatrix}. \tag{A.18}$$

Thus, for each iteration, given $\mathbf{x}^\nu$ and $\mathbf{y}^\nu$, evaluate $\mathbf{F}^\nu$, $\mathbf{B}_\mathrm{x}^\nu$, and $\mathbf{B}_\mathrm{y}^\nu$. Then calculate $\mathbf{E}^\nu$, $\mathbf{H}^\nu$, and $\mathbf{r}^\nu$, and find $\boldsymbol{\alpha}^{\nu+1}$ and $\mathbf{d}^{\nu+1}$ by solving (A.18). Finally, $\mathbf{c}^{\nu+1}$ follows from (A.15).

A.1.3 χ^2 Calculation

Define

$$\begin{pmatrix} \mathbf{W} & \mathbf{V}^T \\ \mathbf{V} & \mathbf{U} \end{pmatrix} \equiv \begin{pmatrix} -\mathbf{H} & \mathbf{B}_\mathrm{y}^T \\ \mathbf{B}_\mathrm{y} & \mathbf{Y} \end{pmatrix}^{-1}. \tag{A.19}$$

Note that since both $\mathbf{H}$ and $\mathbf{Y}$ are symmetric, $\mathbf{W}$ and $\mathbf{U}$ are also symmetric.

Multiplying out (A.19) yields

$$\begin{aligned} -\mathbf{HW} + \mathbf{B}_\mathbf{y}^T\mathbf{V} &= 1 \qquad & \mathbf{B}_\mathbf{y}\mathbf{W} + \mathbf{YV} &= 0 \\ -\mathbf{HV}^T + \mathbf{B}_\mathbf{y}^T\mathbf{U} &= 0 \qquad & \mathbf{B}_\mathbf{y}\mathbf{V}^T + \mathbf{YU} &= 1. \end{aligned} \tag{A.20}$$

From these relations, one can derive the identities

$$\begin{aligned} \mathbf{WHW} + \mathbf{V}^T\mathbf{YV} &= (\mathbf{V}^T\mathbf{B}_\mathbf{y} - 1)\mathbf{W} + \mathbf{V}^T(-\mathbf{B}_\mathbf{y}\mathbf{W}) \\ &= -\mathbf{W}, \end{aligned} \tag{A.21}$$

$$\begin{aligned} \mathbf{VHV}^T + \mathbf{UYU} &= (\mathbf{UB}_\mathbf{y})\mathbf{V}^T + \mathbf{U}(1 - \mathbf{B}_\mathbf{y}\mathbf{V}^T) \\ &= \mathbf{U}, \end{aligned} \tag{A.22}$$

and

$$\begin{aligned} \mathbf{V}^T\mathbf{YU} + \mathbf{WHV}^T &= \mathbf{V}^T(1 - \mathbf{B}_\mathbf{y}\mathbf{V}^T) + (\mathbf{V}^T\mathbf{B}_\mathbf{y} - 1)\mathbf{V}^T \\ &= 0. \end{aligned} \tag{A.23}$$

Now, combining (A.19) with (A.18) yields

$$\begin{pmatrix} \boldsymbol{\alpha} \\ \mathbf{d} \end{pmatrix}^{\nu+1} = \begin{pmatrix} \mathbf{W} & \mathbf{V}^T \\ \mathbf{V} & \mathbf{U} \end{pmatrix}^{\nu} \begin{pmatrix} \mathbf{r}^T \\ 0 \end{pmatrix}^{\nu}, \tag{A.24}$$

or, multiplying it out,

$$\boldsymbol{\alpha}^{\nu+1} = \mathbf{W}^\nu \mathbf{r}^{\nu T} \tag{A.25}$$

and

$$\mathbf{d}^{\nu+1} = \mathbf{V}^\nu \mathbf{r}^{\nu T}. \tag{A.26}$$

Also, from (A.15),

$$\mathbf{c}^{\nu+1} = -\mathbf{E}^{\nu}\mathbf{W}^{\nu}\mathbf{r}^{\nu T}. \tag{A.27}$$

Pulling these all together and plugging them into the χ^2 definition (A.5), the result is (suppressing ν superscripts)

$$\begin{aligned}
\chi^2 &= \mathbf{c}^T\mathbf{G}\mathbf{c} + \mathbf{d}^T\mathbf{Y}\mathbf{d} \\
&= (-\mathbf{r}\mathbf{W}\mathbf{E}^T)\mathbf{G}(-\mathbf{E}\mathbf{W}\mathbf{r}^T) + (\mathbf{r}\mathbf{V}^T)\mathbf{Y}(\mathbf{V}\mathbf{r}^T) \\
&= \mathbf{r}(\mathbf{W}\mathbf{H}\mathbf{W} + \mathbf{V}^T\mathbf{Y}\mathbf{V})\mathbf{r}^T \\
&= -\mathbf{r}\mathbf{W}\mathbf{r}^T \\
&= -\mathbf{r}\boldsymbol{\alpha}.
\end{aligned}$$

Thus,

$$\left(\chi^2\right)^{\nu+1} = -\mathbf{r}^{\nu}\boldsymbol{\alpha}^{\nu+1}. \tag{A.28}$$

A.1.4 Convergence

The fit is deemed to have converged if the following two criteria are satisfied for two iterations in a row:

$$\sum_{\lambda} |F_{\lambda}^{\nu+1}| < \epsilon_2 \tag{A.29}$$

$$|(\chi^2)^{\nu+1} - (\chi^2)^{\nu}| < \epsilon_3 \tag{A.30}$$

That is, when the constraints are satisfied and the χ^2 has stopped changing.

A.1.5 Error Propagation

The final error matrices are defined by

$$\begin{aligned} Q_{ij} &\equiv \langle \delta x_i^f \delta x_j^f \rangle \\ R_{ij} &\equiv \langle \delta y_i^f \delta y_j^f \rangle \\ S_{ij} &\equiv \langle \delta x_i^f \delta y_j^f \rangle. \end{aligned} \tag{A.31}$$

Noting that $\mathbf{B}_\mathrm{x}$, $\mathbf{B}_\mathrm{y}$, and $\mathbf{E}$ are independent of $\mathbf{x}^m$ and $\mathbf{y}^m$, differentiate (A.13) with respect to $\mathbf{x}^m$ and $\mathbf{y}^m$ to get

$$\frac{\partial \mathbf{r}}{\partial \mathbf{x}^m} = -\mathbf{B}_\mathrm{x} \qquad \frac{\partial \mathbf{r}}{\partial \mathbf{y}^m} = -\mathbf{B}_\mathrm{y}. \tag{A.32}$$

Thus, using (A.27), (A.26), and (A.20),

$$\frac{\partial \mathbf{x}}{\partial \mathbf{x}^m} = \frac{\partial(\mathbf{x}^m + \mathbf{c})}{\partial \mathbf{x}^m} = 1 - \frac{\partial}{\partial \mathbf{x}^m}(\mathbf{E}\mathbf{W}\mathbf{r}^T) = 1 + \mathbf{B}_\mathrm{x}\mathbf{W}\mathbf{E}^T \tag{A.33}$$

$$\frac{\partial \mathbf{x}}{\partial \mathbf{y}^m} = \frac{\partial(\mathbf{x}^m + \mathbf{c})}{\partial \mathbf{y}^m} = -\frac{\partial}{\partial \mathbf{y}^m}(\mathbf{E}\mathbf{W}\mathbf{r}^T) = \mathbf{B}_\mathrm{y}\mathbf{W}\mathbf{E}^T = -\mathbf{Y}\mathbf{V}\mathbf{E}^T \tag{A.34}$$

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}^m} = \frac{\partial(\mathbf{y}^m + \mathbf{d})}{\partial \mathbf{x}^m} = \frac{\partial}{\partial \mathbf{x}^m}(\mathbf{V}\mathbf{r}^T) = -\mathbf{B}_\mathrm{x}\mathbf{V}^T \tag{A.35}$$

$$\frac{\partial \mathbf{y}}{\partial \mathbf{y}^m} = \frac{\partial(\mathbf{y}^m + \mathbf{d})}{\partial \mathbf{y}^m} = 1 + \frac{\partial}{\partial \mathbf{y}^m}(\mathbf{V}\mathbf{r}^T) = 1 - \mathbf{B}_\mathrm{y}\mathbf{V}^T = \mathbf{Y}\mathbf{U}. \tag{A.36}$$

(Note that since $\frac{\partial F_\lambda}{\partial x_i} = B_{i\lambda}$, the form $\mathbf{M} = \frac{\partial}{\partial \mathbf{x}}(\mathbf{A}\mathbf{F})$ can be written as $M_{ij} = \frac{\partial}{\partial x_i}(A_{jk}F_k) = A_{jk}B_{ik}$, or $\mathbf{M} = \mathbf{B}\mathbf{A}^T$.)

So, to evaluate $\mathbf{Q} = \langle \delta\mathbf{x}\delta\mathbf{x}^T \rangle$, note that

$$\delta\mathbf{x} = \left(\frac{\partial \mathbf{x}}{\partial \mathbf{x}^m}\right)^T \delta\mathbf{x}^m + \left(\frac{\partial \mathbf{x}}{\partial \mathbf{y}^m}\right)^T \delta\mathbf{y}^m, \tag{A.37}$$

and similarly for $\mathbf{y}$. So, using (A.21), one gets

$$\begin{aligned}
\mathbf{Q} &= \left(\frac{\partial \mathbf{x}}{\partial \mathbf{x}^m}\right)^T \langle \delta\mathbf{x}^m \delta\mathbf{x}^{mT} \rangle \left(\frac{\partial \mathbf{x}}{\partial \mathbf{x}^m}\right) + \left(\frac{\partial \mathbf{x}}{\partial \mathbf{y}^m}\right)^T \langle \delta\mathbf{y}^m \delta\mathbf{y}^{mT} \rangle \left(\frac{\partial \mathbf{x}}{\partial \mathbf{y}^m}\right) \\
&= (1 + \mathbf{E}\mathbf{W}\mathbf{B}_{\mathrm{x}}^T)\mathbf{G}^{-1}(1 + \mathbf{B}_{\mathrm{x}}\mathbf{W}\mathbf{E}^T) + (\mathbf{E}\mathbf{V}^T\mathbf{Y})\mathbf{Y}^{-1}(\mathbf{Y}\mathbf{V}\mathbf{E}^T) \\
&= \mathbf{G}^{-1} + \mathbf{E}(2\mathbf{W} + \mathbf{W}\mathbf{H}\mathbf{W} + \mathbf{V}^T\mathbf{Y}\mathbf{V})\mathbf{E}^T \\
&= \mathbf{G}^{-1} + \mathbf{E}\mathbf{W}\mathbf{E}^T.
\end{aligned} \tag{A.38}$$

Also, using (A.23),

$$\begin{aligned}
\mathbf{S} &= \langle \delta\mathbf{x}\delta\mathbf{y} \rangle \\
&= \left(\frac{\partial \mathbf{x}}{\partial \mathbf{x}^m}\right)^T \langle \delta\mathbf{x}^m \delta\mathbf{x}^{mT} \rangle \left(\frac{\partial \mathbf{y}}{\partial \mathbf{x}^m}\right) + \left(\frac{\partial \mathbf{x}}{\partial \mathbf{y}^m}\right)^T \langle \delta\mathbf{y}^m \delta\mathbf{y}^{mT} \rangle \left(\frac{\partial \mathbf{y}}{\partial \mathbf{y}^m}\right) \\
&= (1 + \mathbf{E}\mathbf{W}\mathbf{B}_{\mathrm{x}}^T)\mathbf{G}^{-1}(-\mathbf{B}_{\mathrm{x}}\mathbf{V}^T) + (-\mathbf{E}\mathbf{V}^T\mathbf{Y})\mathbf{Y}^{-1}(\mathbf{Y}\mathbf{U}) \\
&= -\mathbf{E}(\mathbf{V}^T + \mathbf{W}\mathbf{H}\mathbf{V}^T + \mathbf{V}^T\mathbf{Y}\mathbf{U}) \\
&= -\mathbf{E}\mathbf{V}^T.
\end{aligned} \tag{A.39}$$

Finally, using (A.22),

$$\begin{aligned}
\mathbf{R} &= \langle \delta\mathbf{y}\delta\mathbf{y} \rangle \\
&= \left(\frac{\partial \mathbf{y}}{\partial \mathbf{x}^m}\right)^T \langle \delta\mathbf{x}^m \delta\mathbf{x}^{mT} \rangle \left(\frac{\partial \mathbf{y}}{\partial \mathbf{x}^m}\right) + \left(\frac{\partial \mathbf{y}}{\partial \mathbf{y}^m}\right)^T \langle \delta\mathbf{y}^m \delta\mathbf{y}^{mT} \rangle \left(\frac{\partial \mathbf{y}}{\partial \mathbf{y}^m}\right) \\
&= (\mathbf{V}\mathbf{B}_{\mathrm{x}}^T)\mathbf{G}^{-1}(\mathbf{B}_{\mathrm{x}}\mathbf{V}^T) + (\mathbf{U}\mathbf{Y})\mathbf{Y}^{-1}(\mathbf{Y}\mathbf{U}) \\
&= \mathbf{V}\mathbf{H}\mathbf{V}^T + \mathbf{U}\mathbf{Y}\mathbf{U} \\
&= \mathbf{U}.
\end{aligned} \tag{A.40}$$

A.1.6 Pull Functions

Define the pull functions

$$\xi_i = \frac{c_i}{\sqrt{\langle c_i^2 \rangle}}, \qquad \phi_i = \frac{d_i}{\sqrt{\langle d_i^2 \rangle}}. \tag{A.41}$$

These quantities should be normally distributed with means of zero and widths of one.

For the moment, ignore the y's. Then, since $\delta\mathbf{x} = \mathbf{x} - \mathbf{x}^t$ and $\delta\mathbf{x}^m = \mathbf{x}^m - \mathbf{x}^t$, it follows that

$$\delta\mathbf{x} = \delta\mathbf{x}^m + \mathbf{c}. \tag{A.42}$$

Thus,

$$\langle \delta\mathbf{x}^m \delta\mathbf{x}^{mT} \rangle = \langle \delta\mathbf{x} \delta\mathbf{x}^T \rangle + \langle \mathbf{c}\mathbf{c}^T \rangle - 2\langle \mathbf{c}\delta\mathbf{x}^T \rangle, \tag{A.43}$$

or

$$\langle \mathbf{c}\mathbf{c}^T \rangle = \mathbf{G}^{-1} - \mathbf{Q} + 2\langle \mathbf{c}\delta\mathbf{x}^T \rangle. \tag{A.44}$$

But,

$$\langle \mathbf{c}\delta\mathbf{x}^T \rangle = \langle \delta\mathbf{x}\delta\mathbf{x}^T \rangle - \langle \delta\mathbf{x}^m \delta\mathbf{x}^T \rangle = \mathbf{Q} - \mathbf{G}^{-1} \left(\frac{\partial \mathbf{x}}{\partial \mathbf{x}^m} \right). \tag{A.45}$$

Distinguishing again between well-measured and poorly measured variables,

$$\begin{aligned} \langle \mathbf{c}\delta\mathbf{x}^T \rangle &\rightarrow \begin{pmatrix} \langle \mathbf{c}\delta\mathbf{x}^T \rangle & \langle \mathbf{d}\delta\mathbf{x}^T \rangle \\ \langle \mathbf{c}\delta\mathbf{y}^T \rangle & \langle \mathbf{d}\delta\mathbf{y}^T \rangle \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{Q} - \mathbf{G}^{-1} \left(\frac{\partial \mathbf{x}}{\partial \mathbf{x}^m} \right) & \mathbf{S}^T - \mathbf{Y}^{-1} \left(\frac{\partial \mathbf{x}}{\partial \mathbf{y}^m} \right) \\ \mathbf{S} - \mathbf{G}^{-1} \left(\frac{\partial \mathbf{y}}{\partial \mathbf{x}^m} \right) & \mathbf{R} - \mathbf{Y}^{-1} \left(\frac{\partial \mathbf{y}}{\partial \mathbf{y}^m} \right) \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{Q} - (\mathbf{G}^{-1} + \mathbf{E}\mathbf{W}\mathbf{E}^T) & \mathbf{S}^T + \mathbf{V}\mathbf{E}^T \\ \mathbf{S} + \mathbf{E}\mathbf{V}^T & \mathbf{R} - \mathbf{U} \end{pmatrix} \\ &= 0. \end{aligned} \tag{A.46}$$

Thus,

$$\langle \mathbf{c}\mathbf{c}^T \rangle = \mathbf{G}^{-1} - \mathbf{Q}, \tag{A.47}$$

so the pull is

$$\xi_i = \frac{c_i}{\sqrt{(G^{-1})_{ii} - Q_{ii}}}. \tag{A.48}$$

Similarly, for poorly measured variables (if the off-diagonal elements are small),

$$\frac{1}{\langle \mathbf{d}\mathbf{d}^T \rangle_{ii}} = \frac{1}{(Y^{-1})_{ii} - R_{ii}} \approx \frac{Y_{ii}}{1 - Y_{ii}R_{ii}}, \tag{A.49}$$

so

$$\phi_i = \frac{d_i}{\sqrt{\langle \mathbf{d}\mathbf{d}^T \rangle_{ii}}} = d_i \sqrt{\frac{Y_{ii}}{1 - Y_{ii}R_{ii}}}. \tag{A.50}$$

A.1.7 Cut Steps

The constraint sum $\sum_\lambda |F_\lambda|$ should get smaller with each step. If it does not, then see if the χ^2 is changing. If it is not, then the fit is near a solution, and a directed step should be tried (see Section A.1.8). Otherwise, the step is 'cut' by a fraction ϵ and retried. That is, instead of stepping to $\mathbf{x}^{\nu+1}$, step to $\mathbf{x}'$, defined by

$$\mathbf{x}' = \epsilon(\mathbf{x}^{\nu+1} - \mathbf{x}^\nu) + \mathbf{x}^\nu, \tag{A.51}$$

or

$$\mathbf{c}' = \epsilon \mathbf{c}^{\nu+1} + (1 - \epsilon)\mathbf{c}^\nu. \tag{A.52}$$

Also,

$$\mathbf{d}' = \epsilon \mathbf{d}^{\nu+1} + (1 - \epsilon)\mathbf{d}^\nu. \tag{A.53}$$

Now to find the χ^2. At $\mathbf{x}'$, $\mathbf{y}'$, the χ^2 is

$$\begin{aligned}\left(\chi^2\right)' &= \mathbf{c}'^T\mathbf{G}\mathbf{c}' + \mathbf{d}'^T\mathbf{Y}\mathbf{d}' \\ &= (\epsilon\mathbf{c}^{\nu+1T} + (1-\epsilon)\mathbf{c}^{\nu T})\mathbf{G}(\epsilon\mathbf{c}^{\nu+1} + (1-\epsilon)\mathbf{c}^{\nu}) \\ &\quad + (\epsilon\mathbf{d}^{\nu+1T} + (1-\epsilon)\mathbf{d}^{\nu T})\mathbf{Y}(\epsilon\mathbf{d}^{\nu+1} + (1-\epsilon)\mathbf{d}^{\nu}) \\ &= \epsilon^2\mathbf{c}^{\nu+1T}\mathbf{G}\mathbf{c}^{\nu+1} + (1-\epsilon)^2\mathbf{c}^{\nu T}\mathbf{G}\mathbf{c}^{\nu} + 2\epsilon(1-\epsilon)\mathbf{c}^{\nu+1T}\mathbf{G}\mathbf{c}^{\nu} \\ &\quad + \epsilon^2\mathbf{d}^{\nu+1T}\mathbf{Y}\mathbf{d}^{\nu+1} + (1-\epsilon)^2\mathbf{d}^{\nu T}\mathbf{Y}\mathbf{d}^{\nu} + 2\epsilon(1-\epsilon)\mathbf{d}^{\nu+1T}\mathbf{Y}\mathbf{d}^{\nu}.\end{aligned} \tag{A.54}$$

Define

$$\psi \equiv \mathbf{c}^{\nu+1T}\mathbf{G}\mathbf{c}^{\nu} + \mathbf{d}^{\nu+1T}\mathbf{Y}\mathbf{d}^{\nu}. \tag{A.55}$$

Then,

$$\left(\chi^2\right)' = \epsilon^2\left(\chi^2\right)^{\nu+1} + (1-\epsilon)^2\left(\chi^2\right)^{\nu} + 2\epsilon(1-\epsilon)\psi. \tag{A.56}$$

Now, using (A.20), (A.27), (A.26), and (A.25),

$$\begin{aligned}\psi &= \mathbf{c}^{\nu+1T}\mathbf{G}\mathbf{c}^{\nu} + \mathbf{d}^{\nu+1T}\mathbf{Y}\mathbf{d}^{\nu} \\ &= (-\mathbf{r}\mathbf{W}\mathbf{E}^T)^{\nu}\mathbf{G}\mathbf{c}^{\nu} + (\mathbf{r}\mathbf{V}^T)^{\nu}\mathbf{Y}\mathbf{d}^{\nu} \\ &= (-\mathbf{r}\mathbf{W}\mathbf{B}_{\mathbf{x}}^T)^{\nu}\mathbf{c}^{\nu} + (-\mathbf{r}\mathbf{W}\mathbf{B}_{\mathbf{y}}^T)^{\nu}\mathbf{d}^{\nu} \\ &= -\mathbf{r}^{\nu}\mathbf{W}^{\nu}(\mathbf{B}_{\mathbf{x}}^T\mathbf{c} + \mathbf{B}_{\mathbf{y}}^T\mathbf{d})^{\nu} \\ &= -(\mathbf{r}+\mathbf{F})^{\nu}\boldsymbol{\alpha}^{\nu+1}.\end{aligned} \tag{A.57}$$

Since the step may have to be cut multiple times, note that

$$\begin{aligned}\psi' &= \mathbf{c}'^T\mathbf{G}\mathbf{c}^{\nu} + \mathbf{d}'^T\mathbf{Y}\mathbf{d}^{\nu} \\ &= (\epsilon\mathbf{c}^{\nu+1T} + (1-\epsilon)\mathbf{c}^{\nu T})\mathbf{G}\mathbf{c}^{\nu} + (\epsilon\mathbf{d}^{\nu+1T} + (1-\epsilon)\mathbf{d}^{\nu T})\mathbf{Y}\mathbf{d}^{\nu} \\ &= \epsilon\psi + (1-\epsilon)\left(\chi^2\right)^{\nu}.\end{aligned} \tag{A.58}$$

A.1.8 Directed Steps

For a directed step, try to take the smallest step which satisfies the constraints; i.e., minimize

$$N = \boldsymbol{\gamma}^T \mathbf{G} \boldsymbol{\gamma} + \boldsymbol{\delta}^T \mathbf{Y} \boldsymbol{\delta} + 2\mathbf{F}\boldsymbol{\beta} \tag{A.59}$$

by stepping to $\mathbf{x}'$, $\mathbf{y}'$, where

$$\boldsymbol{\gamma} = \mathbf{x}' - \mathbf{x}^\nu, \qquad \boldsymbol{\delta} = \mathbf{y}' - \mathbf{y}^\nu. \tag{A.60}$$

After linearizing, this gives

$$0 = \mathbf{F}^\nu + \boldsymbol{\gamma}^T \mathbf{B}_\mathrm{x}^\nu + \boldsymbol{\delta}^T \mathbf{B}_\mathrm{y}^\nu \tag{A.61}$$

$$0 = \mathbf{G}\boldsymbol{\gamma} + \mathbf{B}_\mathrm{x}^\nu \boldsymbol{\beta} \quad \longrightarrow \quad \boldsymbol{\gamma} = -\mathbf{G}^{-1}\mathbf{B}_\mathrm{x}^\nu \boldsymbol{\beta} \tag{A.62}$$

$$0 = \mathbf{Y}\boldsymbol{\delta} + \mathbf{B}_\mathrm{y}^\nu \boldsymbol{\beta} \quad \longrightarrow \quad \boldsymbol{\delta} = -\mathbf{Y}^{-1}\mathbf{B}_\mathrm{y}^\nu \boldsymbol{\beta}. \tag{A.63}$$

Putting together (A.61) and (A.62) yields

$$-\mathbf{H}^\nu \boldsymbol{\beta} + \mathbf{B}_\mathrm{y}^{\nu T} \boldsymbol{\delta} = -\mathbf{F}^{\nu T}. \tag{A.64}$$

Combining this with (A.63),

$$\begin{pmatrix} -\mathbf{H} & \mathbf{B}_\mathrm{y}^T \\ \mathbf{B}_\mathrm{y} & \mathbf{Y} \end{pmatrix}^\nu \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\delta} \end{pmatrix} = \begin{pmatrix} -\mathbf{F}^T \\ 0 \end{pmatrix}^\nu . \tag{A.65}$$

This can be solved to find $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$, and thence $\boldsymbol{\gamma}$.

Using (A.19), the result is

$$\begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\delta} \end{pmatrix} = \begin{pmatrix} \mathbf{W} & \mathbf{V}^T \\ \mathbf{V} & \mathbf{U} \end{pmatrix}^\nu \begin{pmatrix} -\mathbf{F}^T \\ 0 \end{pmatrix}^\nu , \tag{A.66}$$

or

$$\boldsymbol{\beta} = -\mathbf{W}^{\nu}\mathbf{F}^{\nu T} \qquad \boldsymbol{\delta} = -\mathbf{V}^{\nu}\mathbf{F}^{\nu T}. \tag{A.67}$$

Then calculate the resulting χ^2 (using (A.20)):

$$\begin{aligned}
\left(\chi^2\right)' &= \mathbf{c}'^T\mathbf{G}\mathbf{c}' + \mathbf{d}'^T\mathbf{Y}\mathbf{d}' \\
&= (\boldsymbol{\gamma} + \mathbf{c}^{\nu})^T\mathbf{G}(\boldsymbol{\gamma} + \mathbf{c}^{\nu}) + (\boldsymbol{\delta} + \mathbf{d}^{\nu})^T\mathbf{Y}(\boldsymbol{\delta} + \mathbf{d}^{\nu}) \\
&= \boldsymbol{\gamma}^T\mathbf{G}\boldsymbol{\gamma} + \boldsymbol{\delta}^T\mathbf{Y}\boldsymbol{\delta} + 2\boldsymbol{\gamma}^T\mathbf{G}\mathbf{c}^{\nu} + 2\boldsymbol{\delta}^T\mathbf{Y}\mathbf{d}^{\nu} + \mathbf{c}^{\nu T}\mathbf{G}\mathbf{c}^{\nu} + \mathbf{d}^{\nu T}\mathbf{Y}\mathbf{d}^{\nu} \\
&= \boldsymbol{\beta}^T\mathbf{B}_{\mathrm{x}}^{\nu T}\mathbf{G}^{-1}\mathbf{G}\mathbf{G}^{-1}\mathbf{B}_{\mathrm{x}}^{\nu}\boldsymbol{\beta} - \boldsymbol{\beta}^T\mathbf{B}_{\mathrm{y}}^{\nu T}\mathbf{Y}^{-1}\mathbf{Y}\boldsymbol{\delta} \\
&\quad - 2\boldsymbol{\beta}^T\mathbf{B}_{\mathrm{x}}^{\nu T}\mathbf{G}^{-1}\mathbf{G}\mathbf{c}^{\nu} - 2\boldsymbol{\beta}^T\mathbf{B}_{\mathrm{y}}^{\nu T}\mathbf{Y}^{-1}\mathbf{Y}\mathbf{d}^{\nu} + \left(\chi^2\right)^{\nu} \\
&= \boldsymbol{\beta}^T\mathbf{H}^{\nu}\boldsymbol{\beta} - \boldsymbol{\beta}^T\mathbf{B}_{\mathrm{y}}^{\nu T}\boldsymbol{\delta} - 2\boldsymbol{\beta}^T(\mathbf{B}_{\mathrm{x}}^{\nu T}\mathbf{c}^{\nu} + \mathbf{B}_{\mathrm{y}}^{\nu T}\mathbf{d}^{\nu}) + \left(\chi^2\right)^{\nu} \\
&= \boldsymbol{\beta}^T(-\mathbf{H}^{\nu}\mathbf{W}^{\nu}\mathbf{F}^{\nu T} + \mathbf{B}_{\mathrm{y}}^{\nu T}\mathbf{V}^{\nu}\mathbf{F}^{\nu T}) - 2\boldsymbol{\beta}^T(\mathbf{r} + \mathbf{F})^{\nu T} + \left(\chi^2\right)^{\nu} \\
&= \mathbf{F}^{\nu}\boldsymbol{\beta} - 2(\mathbf{r} + \mathbf{F})^{\nu}\boldsymbol{\beta} + \left(\chi^2\right)^{\nu} \\
&= \left(\chi^2\right)^{\nu} - 2\mathbf{r}^{\nu}\boldsymbol{\beta} - \mathbf{F}^{\nu}\boldsymbol{\beta}.
\end{aligned} \tag{A.68}$$

A.1.9 Summary of Fitting Algorithm

1. The inputs are the vectors of measured values $\mathbf{x}^m$, $\mathbf{y}^m$, the initial starting point for the fit $\mathbf{x}$, $\mathbf{y}$, and the error matrices $\mathbf{G}^{-1}$ and $\mathbf{Y}$. There's also a constraint function `constraint_fcn`, which takes a point $\mathbf{x}$, $\mathbf{y}$ and returns the values of the constraints $\mathbf{F}$ and their gradients $\mathbf{B}_{\mathrm{x}}$ and $\mathbf{B}_{\mathrm{y}}$.

2. Call `constraint_fcn` for the initial values of $\mathbf{x}$ and $\mathbf{y}$ to get the initial values for $\mathbf{F}$, $\mathbf{B}_{\mathrm{x}}$, and $\mathbf{B}_{\mathrm{y}}$. If the initial starting point is rejected by `constraint_fcn` as invalid, return an error status.

3. Initialize $\mathbf{c} \leftarrow \mathbf{x} - \mathbf{x}^m$ and $\mathbf{d} \leftarrow \mathbf{y} - \mathbf{y}^m$.

4. For each iteration:

5. Compute $\mathbf{E} \leftarrow \mathbf{G}^{-1}\mathbf{B}_\mathrm{x}$, $\mathbf{H} \leftarrow \mathbf{E}^T\mathbf{B}_\mathrm{x}$, and $\mathbf{r} \leftarrow \mathbf{c}^T\mathbf{B}_\mathrm{x} + \mathbf{d}^T\mathbf{B}_\mathrm{y} - \mathbf{F}$.

6. Solve the set of equations $\begin{pmatrix} -\mathbf{H} & \mathbf{B}_\mathrm{y}^T \\ \mathbf{B}_\mathrm{y} & \mathbf{Y} \end{pmatrix} \begin{pmatrix} \boldsymbol{\alpha} \\ \mathbf{d}_1 \end{pmatrix} = \begin{pmatrix} \mathbf{r}^T \\ 0 \end{pmatrix}$ for $\boldsymbol{\alpha}$ and $\mathbf{d}_1$. Also remember the matrices $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{W}$, defined by the equations $\begin{pmatrix} \mathbf{W} & \mathbf{V}^T \\ \mathbf{V} & \mathbf{U} \end{pmatrix} = \begin{pmatrix} -\mathbf{H} & \mathbf{B}_\mathrm{y}^T \\ \mathbf{B}_\mathrm{y} & \mathbf{Y} \end{pmatrix}^{-1}$.

7. Compute the new value of $\mathbf{c}$ and χ^2 by $\mathbf{c}_1 \leftarrow -\mathbf{E}\boldsymbol{\alpha}$ and $\chi^2 \leftarrow -\mathbf{r}\boldsymbol{\alpha}$. In case the step needs to be cut, compute the value of ψ by $\psi \leftarrow -(\mathbf{r}+\mathbf{F})\boldsymbol{\alpha}$.

8. Compute the trial destination of this step, $\mathbf{x}$, $\mathbf{y}$, by $\mathbf{x} \leftarrow \mathbf{c}_1 + \mathbf{x}^m$ and $\mathbf{y} \leftarrow \mathbf{d}_1 + \mathbf{y}^m$.

9. Initialize the total cut size for this step, `this_step_cutsize` $\leftarrow 1$.

10. Call `constraint_fcn` at $\mathbf{x}$, $\mathbf{y}$. If `constraint_fcn` rejects the point, the step must be cut, as described below. Compute the sum of the absolute values of the constraint functions, `constraint_sum`. If it has decreased since the last step, or if it is below the convergence threshold `constraint_sum_eps`, the step is accepted. The first step is accepted no matter what the constraint functions are doing. Otherwise, the constraints have increased, and the step should be cut:

    (a) If this is the first attempt to cut this step, test to see if the χ^2 is stationary. If $|$`chisq` $-$ `chisq_last`$| <$ `chisq_diff_eps`, try a directed step:

i. Solve the set of equations $\begin{pmatrix} -\mathbf{H} & \mathbf{B}_y^T \\ \mathbf{B}_y & \mathbf{Y} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\delta} \end{pmatrix} = \begin{pmatrix} -\mathbf{F}^T \\ 0 \end{pmatrix}$ for $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$, using the values of $\mathbf{B}_y$ and $\mathbf{F}$ from the beginning of the step. Also remember $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{W}$, as before.

ii. Compute $\boldsymbol{\gamma} = -\mathbf{E}\boldsymbol{\beta}$.

iii. Compute the destination of the directed step by $\mathbf{x} \leftarrow \mathbf{c} + \mathbf{x}^m + \boldsymbol{\gamma}$ and $\mathbf{y} \leftarrow \mathbf{d} + \mathbf{y}^m + \boldsymbol{\delta}$.

iv. Call `constraint_fcn`. If $\mathbf{x}$, $\mathbf{y}$ is an acceptable point and the value of `constraint_sum` is now less than the value on the previous step, then accept the directed step and make the assignments

- `chisq` $\leftarrow$ `chisq_last` $- (\mathbf{F} + 2\mathbf{r})\boldsymbol{\beta}$
- $\mathbf{c}_1 \leftarrow \mathbf{x} - \mathbf{x}^m$
- $\mathbf{d}_1 \leftarrow \mathbf{y} - \mathbf{y}^m$

Otherwise, continue with the attempt to cut the step.

(b) If too many attempts have already been made to cut this step, give up and return an error.

(c) Set up the size by which to cut this step, ϵ. Normally, this is set to a constant `cutsize`. However, it is fairly common to have a number of steps in a row, all of which get cut to about the same size. To speed up this case, remember the total cut size from the last step. If this is the first cut for this step, and the last step was also cut, set ϵ to twice the final cut size from the last step, provided that this

value is less than `cutsize`.

(d) Compute the the new total cut size for this step by using the assignment `this_step_cutsize` $\leftarrow \epsilon \cdot$ `this_step_cutsize`. If the total cut size falls below a threshold, `this_step_cutsize` $<$ `min_tot_cutsize`, give up and return an error.

(e) Cut the step, by $\mathbf{c}_1 \leftarrow \epsilon \cdot \mathbf{c}_1 + (1-\epsilon)\mathbf{c}$ and $\mathbf{d}_1 \leftarrow \epsilon \cdot \mathbf{d}_1 + (1-\epsilon)\mathbf{d}$.

(f) Calculate the new χ^2 and ψ from (A.56) and (A.58), and the new $\mathbf{x}$ and $\mathbf{y}$ from $\mathbf{x} \leftarrow \mathbf{c}_1 + \mathbf{x}^m$ and $\mathbf{y} \leftarrow \mathbf{d}_1 + \mathbf{y}^m$.

(g) Return to step 10 to test the cut step.

11. At this point, there's an acceptable step. Shuffle the variables around to prepare for the next step:

- `z2` $\leftarrow$ |`chisq` $-$ `chisq_last`|
- `chisq_last` $\leftarrow$ `chisq`
- `constraint_sum_last` $\leftarrow$ `constraint_sum`
- `last_step_cutsize` $\leftarrow$ `this_step_cutsize`
- $\mathbf{c} \leftarrow \mathbf{c}_1$
- $\mathbf{d} \leftarrow \mathbf{d}_1$

12. Test for convergence by requiring that the system satisfy

`z2` $<$ `chisq_diff_eps` and `constraint_sum` $<$ `constraint_sum_eps`

for two iterations in a row. If these conditions are satisfied, exit the fit loop and go to step 15.

13. Count the total number of steps taken. If it exceeds `maxit`, give up and return an error.

14. Return to step 4 for the next iteration.

15. This fit has succeeded. Calculate the error matrices $\mathbf{Q}$ and $\mathbf{R}$ from (A.38) and (A.40), and the vectors of pull functions $\boldsymbol{\xi}$ and $\boldsymbol{\phi}$ from (A.48) and (A.50).

A.2 Evaluation of Constraint Functions

Now turn to the problem of evaluating the kinematic constraints and their gradients. In order to do this, the variables used to describe an event must be specified, as well as the forms of the constraints themselves.

An event consists of a collection of N final state particles with four-momenta $\mathbf{p}_i$. Each particle is described by three variables: the absolute value of its three-momentum P_i and its direction in $\eta - \phi$ space η_i and ϕ_i. The masses of the particles are taken to be constants, and denoted m_i.

There can be at most one neutrino in the final state, with four-momentum $\mathbf{p}_\nu$. The neutrino is also described by three variables: the z-component of its

momentum p_ν^z, and ${k_T}^x$ and ${k_T}^y$, which are defined as

$$\begin{aligned} {k_T}^x &= p_\nu^x + \sum_i p_i^x \\ {k_T}^y &= p_\nu^y + \sum_i p_i^y. \end{aligned} \tag{A.69}$$

The neutrino, of course, is considered massless.

Now for the constraints. Two general forms of mass constraint are implemented. The first asserts that some collection of particles has a fixed invariant mass (such as in the W mass constraints). The second asserts that two collections of particles have the same invariant mass (as in the top mass equality constraint). The first of these can be written

$$F = \left(\sum_{i \in C} \mathbf{p}_i \right)^2 - M^2, \tag{A.70}$$

and the second

$$F = \left(\sum_{i \in C_1} \mathbf{p}_i \right)^2 - \left(\sum_{i \in C_2} \mathbf{p}_i \right)^2. \tag{A.71}$$

The sums in parentheses can be expanded out as

$$\left(\sum_{i \in C} \mathbf{p}_i \right)^2 = \sum_{i \in C} m_i^2 + 2 \sum_{\substack{i,j \in C \\ i<j}} \mathbf{p}_i \cdot \mathbf{p}_j. \tag{A.72}$$

Note that since the masses m_i are treated as constant, their gradients are zero. Thus, the problem of calculating the gradients of the constraints reduces to calculating the gradients of the dot product of two four-vectors.

If there is no final-state neutrino in the event, then there are the two

additional constraints

$$0 = {k_T}^x = \sum_i p_i^x$$
$$0 = {k_T}^y = \sum_i p_i^y. \tag{A.73}$$

These constraints will be dealt with separately below.

Unfortunately, the variables used to describe an event are not very convenient for calculating the gradients. So the strategy used requires several steps: first, the gradients are calculated with respect to a more convenient set of variables. Then, the gradients are converted to the variables actually desired by a series of Jacobian transformations. For this initial set of variables, the polar angle θ will be used instead of the pseudorapidity η. In addition, the neutrino will be treated just like any other final-state particle, and parameterized using P, θ, and η.

So, the problem is to evaluate the gradients of $\mathbf{p}_1 \cdot \mathbf{p}_2$ with respect to the variables P_1, P_2, θ_1, θ_2, ϕ_1, and ϕ_2. Since, by definition,

$$\begin{aligned} E &= \sqrt{P^2 + m^2} \\ p^x &= P \sin\theta \cos\phi \\ p^y &= P \sin\theta \sin\phi \\ p^z &= P \cos\theta, \end{aligned} \tag{A.74}$$

the dot product

$$\mathbf{p}_1 \cdot \mathbf{p}_2 = E_1 E_2 - p_1^x p_2^x - p_1^y p_2^y - p_1^z p_2^z \tag{A.75}$$

can be written as

$$\begin{aligned}\mathbf{p}_1 \cdot \mathbf{p}_2 =& \sqrt{P_1^2 + m_1^2}\sqrt{P_2^2 + m_2^2} - \\ & P_1 P_2 \left(\sin\theta_1 \sin\theta_2 \cos(\phi_1 - \phi_2) + \cos\theta_1 \cos\theta_2\right).\end{aligned} \tag{A.76}$$

The partial derivative with respect to P_1 is thus

$$\begin{aligned}\frac{\partial\, \mathbf{p}_1 \cdot \mathbf{p}_2}{\partial P_1} &= \frac{P_1}{\sqrt{P_1^2 + m_1^2}}\sqrt{P_2^2 + m_2^2} - P_2(\ldots) \\ &= \frac{\mathbf{p}_1 \cdot \mathbf{p}_2}{P_1} - \frac{E_1 E_2}{P_1} + \frac{P_1}{E_1}E_2 \\ &= \frac{\mathbf{p}_1 \cdot \mathbf{p}_2}{P_1} + \frac{P_1^2 - E_1^2}{P_1 E_1}E_2 \\ &= \frac{1}{P_1}\left(\mathbf{p}_1 \cdot \mathbf{p}_2 - m_1^2\frac{E_2}{E_1}\right).\end{aligned} \tag{A.77}$$

Similarly for P_2,

$$\frac{\partial\, \mathbf{p}_1 \cdot \mathbf{p}_2}{\partial P_2} = \frac{1}{P_2}\left(\mathbf{p}_1 \cdot \mathbf{p}_2 - m_2^2\frac{E_1}{E_2}\right). \tag{A.78}$$

For the angular gradients, the result is

$$\frac{\partial\, \mathbf{p}_1 \cdot \mathbf{p}_2}{\partial \theta_1} = -P_1 P_2 \left(\cos\theta_1 \sin\theta_2 \cos(\phi_1 - \phi_2) - \sin\theta_1 \cos\theta_2\right). \tag{A.79}$$

Noting that

$$\begin{aligned}\sin\phi &= p^y/p_T \\ \cos\phi &= p^x/p_T \\ \sin\theta &= p_T/P \\ \cos\theta &= p^z/P\end{aligned} \tag{A.80}$$

(where $p_T = \sqrt{p^x + p^y}$), (A.79) can be rewritten as

$$\begin{aligned}\frac{\partial\, \mathbf{p}_1 \cdot \mathbf{p}_2}{\partial \theta_1} &= -P_1 P_2 \left[\frac{p_1^z}{P_1}\frac{p_{T2}}{P_2}\frac{1}{p_{T1}}\frac{1}{p_{T2}}(p_1^x p_2^x + p_1^y p_2^y) - \frac{p_{T1}}{P_1}\frac{p_2^z}{P_2}\right] \\ &= p_{T1} p_2^z - \frac{p_1^z}{p_{T1}}(p_1^x p_2^x + p_1^y p_2^y).\end{aligned} \tag{A.81}$$

Similarly,

$$\frac{\partial \mathbf{p}_1 \cdot \mathbf{p}_2}{\partial \theta_2} = p_{T2} p_1^z - \frac{p_2^z}{p_{T2}}(p_1^x p_2^x + p_1^y p_2^y). \tag{A.82}$$

Finally,

$$\begin{aligned} \frac{\partial \mathbf{p}_1 \cdot \mathbf{p}_2}{\partial \phi_1} &= P_1 P_2 \sin\theta_1 \sin\theta_2 \sin(\phi_1 - \phi_2) \\ &= P_1 P_2 \frac{p_{T1}}{P_1}\frac{p_{T2}}{P_2}\frac{1}{p_{T1}}\frac{1}{p_{T2}}(p_1^y p_2^x - p_1^x p_2^y) \\ &= p_1^y p_2^x - p_1^x p_2^y \end{aligned} \tag{A.83}$$

and

$$\frac{\partial \mathbf{p}_1 \cdot \mathbf{p}_2}{\partial \phi_2} = p_1^x p_2^y - p_1^y p_2^x. \tag{A.84}$$

If there is no neutrino in the final state, then the gradients of the k_T constraints (A.73) also need to be calculated. This reduces to evaluating the derivatives of the x and y components of a four-momentum p^x, p^y with respect to P, θ, and ϕ. From (A.74), these gradients are

$$\begin{aligned} \frac{\partial p^x}{\partial P} &= p^x/P \\ \frac{\partial p^y}{\partial P} &= p^y/P. \end{aligned} \tag{A.85}$$

Also,

$$\begin{aligned} \frac{\partial p^x}{\partial \phi} &= -p^y \\ \frac{\partial p^y}{\partial \phi} &= p^x. \end{aligned} \tag{A.86}$$

And finally,

$$\begin{aligned} \frac{\partial p^x}{\partial \theta} &= P\cos\theta\cos\phi \\ &= p^x \cot\theta \\ \frac{\partial p^y}{\partial \theta} &= p^y \cot\theta. \end{aligned} \tag{A.87}$$

Now the variables need to be transformed from the set in which the gradients were evaluated to the final set used for describing the event. The first step will be to change the variables for the neutrino from spherical (P_ν, θ_ν, ϕ_ν) to rectangular ($p_\nu^x = x$, $p_\nu^y = y$, $p_\nu^z = z$) coordinates. This transformation is independent of all other event variables. The equations of transformation can be written

$$\begin{aligned} P &= \sqrt{x^2 + y^2 + z^2} \\ \theta &= \tan^{-1} \frac{p_T}{z} = \tan^{-1} \frac{\sqrt{x^2 + y^2}}{z} \\ \phi &= \tan^{-1} \frac{y}{x}. \end{aligned} \tag{A.88}$$

Thus, by the chain rule,

$$\begin{aligned} \frac{\partial}{\partial x} &= \frac{\partial P}{\partial x}\frac{\partial}{\partial P} + \frac{\partial \phi}{\partial x}\frac{\partial}{\partial \phi} + \frac{\partial \theta}{\partial x}\frac{\partial}{\partial \theta} \\ &= \left(\frac{x}{P}\right)\frac{\partial}{\partial P} - \left(\frac{y}{{p_T}^2}\right)\frac{\partial}{\partial \phi} + \left(\frac{xz}{P^2 p_T}\right)\frac{\partial}{\partial \theta} \\ \frac{\partial}{\partial y} &= \frac{\partial P}{\partial y}\frac{\partial}{\partial P} + \frac{\partial \phi}{\partial y}\frac{\partial}{\partial \phi} + \frac{\partial \theta}{\partial y}\frac{\partial}{\partial \theta} \\ &= \left(\frac{y}{P}\right)\frac{\partial}{\partial P} + \left(\frac{x}{{p_T}^2}\right)\frac{\partial}{\partial \phi} + \left(\frac{yz}{P^2 p_T}\right)\frac{\partial}{\partial \theta} \\ \frac{\partial}{\partial z} &= \frac{\partial P}{\partial z}\frac{\partial}{\partial P} + \frac{\partial \phi}{\partial z}\frac{\partial}{\partial \phi} + \frac{\partial \theta}{\partial z}\frac{\partial}{\partial \theta} \\ &= \left(\frac{z}{P}\right)\frac{\partial}{\partial P} - \left(\frac{p_T}{P^2}\right)\frac{\partial}{\partial \theta}. \end{aligned} \tag{A.89}$$

The next step is to change from the x and y components of the neutrino momentum p_ν^x and p_ν^y to ${k_T}^x$ and ${k_T}^y$, as defined by (A.69). That is, the transformation is

$$x_\nu, y_\nu, z'_\nu, P'_i, \phi'_i, \theta'_i \rightarrow {k_T}^x, {k_T}^y, z_\nu, P_i, \phi_i, \theta_i \tag{A.90}$$

and the equations of transformation are

$$\begin{aligned} x_\nu &= {k_T}^x - \sum_i P_i \sin\theta_i \cos\phi_i \\ y_\nu &= {k_T}^y - \sum_i P_i \sin\theta_i \sin\phi_i \\ z'_\nu &= z_\nu \\ P'_i &= P_i \\ \phi'_i &= \phi_i \\ \theta'_i &= \theta_i. \end{aligned} \tag{A.91}$$

Thus, a Jacobian transform gives

$$\begin{aligned} \frac{\partial}{\partial {k_T}^x} &= \frac{\partial}{\partial x_\nu} \\ \frac{\partial}{\partial {k_T}^y} &= \frac{\partial}{\partial y_\nu} \\ \frac{\partial}{\partial z_\nu} &= \frac{\partial}{\partial z'_\nu} \\ \frac{\partial}{\partial P_i} &= \frac{\partial P'_i}{\partial P_i}\frac{\partial}{\partial P'_i} + \frac{\partial x_\nu}{\partial P_i}\frac{\partial}{\partial x_\nu} + \frac{\partial y_\nu}{\partial P_i}\frac{\partial}{\partial y_\nu} \\ &= \frac{\partial}{\partial P'_i} - \frac{1}{P_i}\left(x_i\frac{\partial}{\partial x_\nu} + y_i\frac{\partial}{\partial y_\nu}\right) \\ \frac{\partial}{\partial \phi_i} &= \frac{\partial \phi'_i}{\partial \phi_i}\frac{\partial}{\partial \phi'_i} + \frac{\partial x_\nu}{\partial \phi_i}\frac{\partial}{\partial x_\nu} + \frac{\partial y_\nu}{\partial \phi_i}\frac{\partial}{\partial y_\nu} \\ &= \frac{\partial}{\partial \phi'_i} + y_i\frac{\partial}{\partial x_\nu} - x_i\frac{\partial}{\partial y_\nu} \\ \frac{\partial}{\partial \theta_i} &= \frac{\partial \theta'_i}{\partial \phi_i}\frac{\partial}{\partial \theta'_i} + \frac{\partial x_\nu}{\partial \phi_i}\frac{\partial}{\partial x_\nu} + \frac{\partial y_\nu}{\partial \phi_i}\frac{\partial}{\partial y_\nu} \\ &= \frac{\partial}{\partial \theta'_i} - \left(x_i\frac{\partial}{\partial x_\nu} + y_i\frac{\partial}{\partial y_\nu}\right)\cot\theta_i. \end{aligned} \tag{A.92}$$

Finally, the polar angles need to be converted from θ to η. This is independent for each object. The transformation is

$$\theta = 2\tan^{-1} e^{-\eta}. \tag{A.93}$$

Thus,

$$\frac{\partial\theta}{\partial\eta} = \frac{2}{1+e^{-2\eta}}\left(-e^{-\eta}\right) = -\frac{2}{e^{\eta}+e^{-\eta}} = -\frac{1}{\cosh\eta} \tag{A.94}$$

and the associated Jacobian transform is

$$\frac{\partial}{\partial\eta} = \frac{\partial\theta}{\partial\eta}\frac{\partial}{\partial\theta} = -\frac{1}{\cosh\eta}\frac{\partial}{\partial\theta}. \tag{A.95}$$

There is one more transformation which might need to be made. If the lepton in the event is a muon, then the variable used to describe its momentum is not actually P but instead its inverse

$$K = 1/P. \tag{A.96}$$

The Jacobian transformation for this case is

$$\frac{\partial}{\partial K} = \frac{\partial P}{\partial K}\frac{\partial}{\partial P} = -\frac{1}{K^2}\frac{\partial}{\partial P} = -P^2\frac{\partial}{\partial P}. \tag{A.97}$$

Appendix B

Candidate Details

This appendix lists the kinematic parameters of all lepton + jets candidate events which pass the loose cuts, plus the results of the mass fit. Each event is identified by a run and event number. The H_T and $\mathcal{A}$ values for the event are also listed. If the event also passes the standard cuts, the event number is followed by an asterisk (*). Next, the lepton, missing E_T, jets, and the tag muon (if present) are listed. All objects are dumped as a four vector, plus the E_T ($= \sqrt{p_x^2 + p_y^2}$), ϕ, and η values. Jets are $R = 0.3$ cone, with the out-of-cone corrections applied. Tagged jets are listed twice, both with and without the correction for the tagging muon (the latter is indicated by a dagger (†)). All jets with $E_T > 15\,\mathrm{GeV}$ and $|\eta| < 2.5$ are listed, although only the first four are used for mass fitting.

Following the dump of the object four-vectors is another table giving the result of the mass fits for that event. Up to three fits are listed, giving the fitted mass (m_t), χ^2, and the jet permutation used (the letters correspond to the first four jets; $b_l \equiv b$ associated with leptonic top, $b_h \equiv b$ associated with

hadronic top, $w \equiv$ hadronic W). The χ^2-weighted average is also listed at the bottom. Fits with $\chi^2 > 7$ are not listed. For tagged candidates, only the fits which properly assign the tagged jet as a b are listed.

Due to an implementation error, a cut which was intended to reject before fitting jet permutations in which there was a large difference between the two top masses was actually rejecting permutations for which the smallest solution for the z-component of the neutrino momentum was large compared to the $\not\!\!E_T$. This is noted in the few cases where this affects the results for a candidate event. Note, however, that since the procedure was applied consistently to both data and Monte Carlo, this does not affect the reliability of the final result.

B.1 Electron + Jets Candidates

Run 62431, Event 788*				$H_T = 205.3$	$\mathcal{A} = 0.089$		
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Electron	−13.7	−50.0	59.1	78.6	51.8	4.45	0.98
$\not\!\!E_T$	20.8	14.6	0.0	25.4	25.4	0.61	0.00
Jet 1	−81.3	10.1	3.6	82.5	81.9	3.02	0.04
Jet 2	45.5	56.8	−170.3	185.5	72.8	0.90	−1.59
Jet 3	13.2	−18.7	−64.1	68.3	22.9	5.33	−1.75
Jet 4	6.2	−15.7	−15.5	23.0	16.9	5.09	−0.82

	m_t	χ^2	Permutation
Fit 1	230.0	0.62	wb_hb_lw
Fit 2	238.5	1.62	b_hwb_lw
Fit 3	119.1	6.19	b_lwwb_h
	Average fitted mass: 229.0		

Run 63066, Event 13373*				$H_T = 224.0$	$\mathcal{A} = 0.124$		
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Electron	−49.8	−12.5	9.3	52.1	51.3	3.39	0.18
$\not\!\!E_T$	−42.7	−31.8	0.0	53.2	53.2	3.78	0.00
Jet 1	87.5	17.8	−2.6	89.7	89.3	0.20	−0.03
Jet 2	13.4	52.7	−123.4	135.0	54.4	1.32	−1.56
Jet 3	−23.7	30.7	7.0	39.8	38.8	2.23	0.18
Jet 4	21.4	−26.9	−50.9	61.7	34.4	5.38	−1.18
Jet 5	−7.8	−28.5	34.1	45.4	29.6	4.45	0.99
Jet 6	19.4	11.8	−9.8	25.2	22.8	0.55	−0.42

	m_t	χ^2	Permutation
Fit 1	159.5	2.37	wb_lb_hw
Fit 2	177.8	3.01	wb_hwb_l
Fit 3	167.2	3.27	b_lwb_hw
	Average fitted mass: 167.2		

Run 65358, Event 225				$H_T = 138.7$	$\mathcal{A} = 0.033$		
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Electron	10.4	24.0	−83.9	87.9	26.1	1.16	−1.88
$\not\!\!E_T$	53.7	14.8	0.0	55.7	55.7	0.27	0.00
Jet 1	−59.9	−38.0	−86.0	111.8	71.0	3.71	−1.02
Jet 2	−1.1	29.3	38.8	48.7	29.3	1.61	1.09
Jet 3	−8.9	−16.9	−18.1	26.5	19.2	4.23	−0.84
Jet 4	12.0	−9.0	−36.6	39.6	15.0	5.64	−1.62

	m_t	χ^2	Permutation
Fit 1	168.0	0.13	b_hwb_lw
Fit 2	133.8	5.70	b_lwwb_h
Fit 3	—	—	—
	Average fitted mass: 166.0		

Run 79667, Event 2422					$H_T = 186.5$	$\mathcal{A} = 0.066$	
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Electron	19.1	32.0	−56.8	68.0	37.3	1.03	−1.21
$\not\!\!E_T$	−3.8	41.5	0.0	41.6	41.6	1.66	0.00
Jet 1	23.6	−69.9	20.3	77.3	73.8	5.04	0.27
Jet 2	−25.3	45.8	−99.6	113.3	52.3	2.08	−1.40
Jet 3	−1.7	−26.8	−27.2	38.5	26.9	4.65	−0.89
Jet 4	−13.3	−18.1	−3.3	23.0	22.4	4.08	−0.15
Jet 5	20.1	−9.6	1.0	22.5	22.3	5.84	0.05

	m_t	χ^2	Permutation
Fit 1	185.4	1.95	$b_h w w b_l$
Fit 2	167.9	5.11	$b_h w b_l w$
Fit 3	—	—	—
	Average fitted mass: 182.4		

Run 81543, Event 21845					$H_T = 97.9$	$\mathcal{A} = 0.086$	
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Electron	19.4	−55.9	137.8	150.0	59.1	5.05	1.58
$\not\!\!E_T$	−31.2	24.0	0.0	39.4	39.4	2.49	0.00
Jet 1	33.6	−3.9	−13.4	36.6	33.9	6.17	−0.38
Jet 2	−31.7	−0.7	0.3	31.9	31.7	3.16	0.01
Jet 3	−14.2	25.4	1.8	29.5	29.1	2.08	0.06
Jet 4	20.2	12.3	−42.4	48.7	23.6	0.55	−1.35

	m_t	χ^2	Permutation
Fit 1	123.9	6.54	$w w b_l b_h$
Fit 2	—	—	—
Fit 3	—	—	—
	Average fitted mass: 123.9		

Run 81949, Event 12380					$H_T = 174.6$	$\mathcal{A} = 0.038$	
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Electron	15.8	−25.2	−35.0	45.9	29.7	5.27	−1.00
$\not\!\!E_T$	−26.3	19.1	0.0	32.5	32.5	2.51	0.00
Jet 1	−39.2	−74.2	−14.0	85.6	83.9	4.23	−0.17
Jet 2	35.4	38.5	−20.9	56.5	52.3	0.83	−0.39
Jet 3	−11.4	43.1	51.2	68.4	44.5	1.83	0.98
Jet 4	19.4	−19.6	−1.0	27.8	27.6	5.49	−0.04

	m_t	χ^2	Permutation
Fit 1	132.5	0.07	$b_l b_h w w$
Fit 2	167.7	0.12	$b_h b_l w w$
Fit 3	125.8	0.69	$b_l w w b_h$
	Average fitted mass: 143.4		

Run 85837, Event 5417					$H_T = 134.4$	$\mathcal{A} = 0.198$	
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Electron	19.7	21.5	84.8	89.6	29.1	0.83	1.79
$\not\!\!E_T$	−34.5	33.2	0.0	47.9	47.9	2.37	0.00
Jet 1	45.8	−35.2	−19.6	61.9	57.7	5.63	−0.33
Jet 2	−39.0	−40.3	−51.1	76.1	56.1	3.94	−0.82
Jet 3	−9.0	22.0	−24.5	34.3	23.8	1.96	−0.90
Jet 4	9.5	−16.3	30.1	35.6	18.8	5.24	1.25

	m_t	χ^2	Permutation
Fit 1	135.5	0.51	$w b_h w b_l$
Fit 2	131.4	0.54	$w w b_l b_h$
Fit 3	131.7	0.57	$b_h w b_l w$
	Average fitted mass: 132.9		

Run 86518, Event 11716*				$H_T = 281.6$	$\mathcal{A} = 0.079$		
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Electron	−36.5	5.2	7.3	37.6	36.9	3.00	0.20
$\not\!\!E_T$	−57.8	−72.9	0.0	93.0	93.0	4.04	0.00
Jet 1	134.0	103.7	−176.1	244.9	169.4	0.66	−0.91
Jet 2	−60.5	16.5	−87.0	107.5	62.7	2.87	−1.13
Jet 3	3.7	−50.1	−75.4	90.8	50.3	4.79	−1.19
Jet 4	0.5	19.5	−16.5	25.7	19.5	1.54	−0.77
Jet 5	4.8	−17.5	2.5	18.7	18.2	4.98	0.14
Jet 6	15.8	7.5	−45.5	48.8	17.5	0.45	−1.68

	m_t	χ^2	Permutation
Fit 1	207.1	6.84	$b_h b_l w w$
Fit 2	—	—	—
Fit 3	—	—	—
Average fitted mass: 207.1			

Run 86601, Event 33128				$H_T = 227.9$	$\mathcal{A} = 0.047$		
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Electron	−29.7	13.7	−26.0	41.8	32.7	2.71	−0.73
$\not\!\!E_T$	13.7	29.4	0.0	32.5	32.5	1.14	0.00
Jet 1	60.6	78.5	100.7	141.8	99.2	0.91	0.89
Jet 2	−10.5	−86.6	−76.0	116.0	87.3	4.59	−0.79
Jet 3	−47.3	−19.4	136.1	145.5	51.1	3.53	1.71
Jet 4	7.3	−18.3	28.5	34.8	19.7	5.09	1.17

	m_t	χ^2	Permutation
Fit 1	181.4	0.12	$w b_l b_h w$
Fit 2	250.8	3.44	$b_l w b_h w$
Fit 3	226.7	3.69	$b_h w b_l w$
Average fitted mass: 196.7			

Run 86922, Event 9328*					$H_T = 217.9$	$\mathcal{A} = 0.118$	
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Electron	−30.2	48.4	35.8	67.3	57.0	2.13	0.59
$\not\!\!E_T$	−9.3	24.0	0.0	25.8	25.8	1.94	0.00
Jet 1	101.6	−3.9	7.7	102.6	101.7	6.24	0.08
Jet 2	−51.2	−22.2	29.6	63.5	55.8	3.55	0.51
Jet 3	−43.1	−33.0	−53.2	76.4	54.3	3.80	−0.87
Jet 4	35.8	−24.8	55.7	70.9	43.5	5.68	1.07

	m_t	χ^2	Permutation
Fit 1	181.4	0.37	$b_l w w b_h$
Fit 2	173.8	1.11	$b_l w b_h w$
Fit 3	168.9	1.33	$b_h w b_l w$
	Average fitted mass: 175.8		

Run 87089, Event 7533*					$H_T = 219.7$	$\mathcal{A} = 0.055$	
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Electron	−19.6	81.6	−4.2	84.1	84.0	1.81	−0.05
$\not\!\!E_T$	−2.6	−27.0	0.0	27.2	27.2	4.62	0.00
Jet 1	69.1	47.8	23.0	88.1	84.1	0.61	0.27
Jet 2	−67.5	−26.7	52.5	90.9	72.6	3.52	0.67
Jet 3	32.9	−24.9	−3.1	41.4	41.3	5.63	−0.07
Jet 4	4.3	−35.6	−66.9	76.1	35.8	4.83	−1.38

	m_t	χ^2	Permutation
Fit 1	174.6	1.13	$w b_l w b_h$
Fit 2	199.1	3.30	$w b_h w b_l$
Fit 3	166.5	3.42	$b_h w w b_l$
	Average fitted mass: 178.1		

Run 87433, Event 2573				$H_T = 115.2$	$\mathcal{A} = 0.100$		
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Electron	14.7	57.3	−36.9	69.8	59.2	1.32	−0.59
$\not\!\!E_T$	29.4	2.5	0.0	29.6	29.6	0.08	0.00
Jet 1	−36.2	−28.0	−44.1	63.9	45.8	3.80	−0.86
Jet 2	12.9	−27.4	50.4	58.9	30.3	5.15	1.28
Jet 3	20.3	15.0	−51.9	57.9	25.3	0.64	−1.47
Jet 4	−10.2	−23.0	−122.2	124.8	25.2	4.29	−2.28

	m_t	χ^2	Permutation
Fit 1	183.7	0.76	wb_hwb_l
Fit 2	164.0	3.38	b_lwwb_h
Fit 3	138.9	3.78	b_hwwb_l
	Average fitted mass: 173.5		

Run 87446, Event 14294				$H_T = 100.9$	$\mathcal{A} = 0.077$		
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Electron	−20.9	22.1	7.6	31.4	30.5	2.33	0.25
$\not\!\!E_T$	25.9	−45.9	0.0	52.7	52.7	5.23	0.00
Jet 1	−37.7	7.3	−6.3	39.3	38.4	2.95	−0.16
Jet 2	25.1	9.4	11.0	29.3	26.8	0.36	0.40
Jet 3	1.0	−24.6	82.9	86.6	24.6	4.75	1.93
Jet 4	9.1	16.1	44.7	48.4	18.5	1.06	1.61

	m_t	χ^2	Permutation
Fit 1	110.3	0.57	wb_lwb_h
Fit 2	115.7	0.57	wb_hwb_l
Fit 3	115.2	1.80	wb_hb_lw
	Average fitted mass: 113.4		

B.2 Muon + Jets Candidates

Run 61275, Event 9188					$H_T = 173.6$	$\mathcal{A} = 0.177$	
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Muon	11.3	−16.5	−36.1	41.3	20.0	5.31	−1.35
$\not\!\!E_T$	−53.3	−10.6	0.0	54.3	54.3	3.34	0.00
Jet 1	36.7	−29.2	−12.8	49.2	46.9	5.61	−0.27
Jet 2	−17.2	36.7	52.0	66.3	40.5	2.01	1.07
Jet 3	1.3	34.7	−14.1	37.7	34.7	1.53	−0.39
Jet 4	−6.3	−31.7	−0.4	32.6	32.3	4.52	−0.01
Jet 5	21.5	−2.5	−14.3	26.2	21.7	6.17	−0.62
Jet 6	−12.3	10.1	6.0	17.3	15.9	2.46	0.37
Jet 7	15.3	−2.3	−2.8	16.0	15.5	6.13	−0.18

	m_t	χ^2	Permutation
Fit 1	124.6	0.19	$b_l w b_h w$
Fit 2	141.0	0.67	$b_h w b_l w$
Fit 3	152.7	0.94	$w b_h w b_l$
	Average fitted mass: 137.6		

Run 61514, Event 4537					$H_T = 114.4$	$\mathcal{A} = 0.076$	
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Muon	−32.7	44.4	13.7	56.8	55.1	2.20	0.25
$\not\!\!E_T$	−2.8	21.3	0.0	21.4	21.4	1.70	0.00
Jet 1	32.4	−29.1	−17.1	47.1	43.5	5.55	−0.38
Jet 2	−0.1	−37.0	−14.1	39.9	37.0	4.71	−0.37
Jet 3	18.6	−24.0	11.3	32.7	30.4	5.37	0.36
Jet 4	2.9	21.0	−61.5	65.1	21.2	1.43	−1.79

	m_t	χ^2	Permutation
Fit 1	126.2	1.97	wb_lb_hw
Fit 2	119.2	2.57	wb_hb_lw
Fit 3	119.0	2.57	b_hwb_lw
	Average fitted mass: 122.0		

Run 62431, Event 215					$H_T = 85.6$	$\mathcal{A} = 0.039$	
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Muon	9.0	15.8	4.5	18.7	18.2	1.05	0.25
$\not\!\!E_T$	−43.7	−11.7	0.0	45.3	45.3	3.40	0.00
Jet 1	37.9	7.6	12.3	41.1	38.7	0.20	0.31
Jet 2	2.2	−17.6	23.9	30.0	17.7	4.84	1.11

	m_t	χ^2	Permutation
Fit 1	—	—	—
Fit 2	—	—	—
Fit 3	—	—	—
	Average fitted mass: —		

(Two of the 0.5 cone jets were close to the 15 GeV threshold, and were not found with the 0.3 cone.)

Run 63740, Event 14197*				$H_T = 274.0$	$\mathcal{A} = 0.084$		
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Muon	18.2	9.5	−9.1	22.5	20.6	0.48	−0.43
$\not\!\!E_T$	−45.3	−31.4	0.0	55.1	55.1	3.75	0.00
Jet 1	54.8	61.3	48.6	96.1	82.2	0.84	0.56
Jet 2	−67.1	−44.3	−30.7	86.6	80.4	3.73	−0.37
Jet 3	57.7	−41.8	31.6	78.5	71.2	5.66	0.43
Jet 4	−41.1	23.2	38.8	61.2	47.2	2.63	0.75
Jet 5	34.2	29.5	54.7	71.4	45.2	0.71	1.02

	m_t	χ^2	Permutation
Fit 1	184.4	1.37	$b_l w b_h w$
Fit 2	162.2	1.49	$w b_l b_h w$
Fit 3	157.8	2.29	$w b_l w b_h$
	Average fitted mass: 169.8		

(The permutation $b_h w b_l w$ ($\chi^2 = 1.39$, $m_t = 201.8\,\mathrm{GeV}/c^2$) was discarded due to a large initial p_z^ν.)

Run 81372, Event 4899					$H_T = 94.3$	$\mathcal{A} = 0.291$	
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Muon	7.4	14.9	−7.7	18.3	16.6	1.11	−0.45
$\not\!\!E_T$	−35.5	22.8	0.0	42.2	42.2	2.57	0.00
Jet 1	38.2	11.8	11.7	41.9	40.0	0.30	0.29
Jet 2	−18.0	−18.8	−29.4	39.4	26.0	3.95	−0.97
Jet 3	5.3	−23.9	31.4	40.0	24.5	4.93	1.07
Jet 4	13.9	−17.4	−21.1	30.8	22.3	5.39	−0.84

	m_t	χ^2	Permutation
Fit 1	119.7	0.43	wwb_hb_l
Fit 2	129.2	2.05	b_hwwb_l
Fit 3	103.3	2.60	wwb_lb_h
	Average fitted mass: 118.9		

Run 82694, Event 25595*				$H_T = 246.9$	$\mathcal{A} = 0.128$		
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Muon	38.1	−43.2	−24.4	62.5	57.6	5.44	−0.41
$\not{E}_T$	29.8	30.7	0.0	42.8	42.8	0.80	0.00
Jet 1	−116.0	45.6	−14.8	126.3	124.7	2.77	−0.12
Jet 2	−24.6	−48.4	−2.7	54.5	54.3	4.24	−0.05
Jet 3	50.1	4.0	−104.0	115.6	50.2	0.08	−1.47
Jet 4	−30.0	6.7	−32.2	45.2	30.7	2.92	−0.91
Jet 5	6.7	23.7	31.4	40.2	24.7	1.30	1.06

	m_t	χ^2	Permutation
Fit 1	166.3	2.27	$b_l b_h ww$
Fit 2	159.8	3.57	$b_h w b_l w$
Fit 3	107.8	4.13	$ww b_l b_h$
	Average fitted mass: 152.4		

(The permutation $b_h b_l ww$ ($\chi^2 = 3.84$, $m_t = 195.3\,\mathrm{GeV}/c^2$) was discarded due to a large initial p_z^ν.)

Run 83074, Event 22558				$H_T = 155.1$	$\mathcal{A} = 0.046$		
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Muon	−21.5	8.2	−8.6	24.6	23.0	2.78	−0.37
$\not\!\!E_T$	−18.2	27.8	0.0	33.2	33.2	2.15	0.00
Jet 1	18.7	−33.1	26.8	46.9	38.0	5.23	0.66
Jet 2	17.2	30.9	−37.2	51.7	35.4	1.06	−0.92
Jet 3	25.1	−20.4	56.2	65.1	32.4	5.60	1.32
Jet 4	3.8	−19.5	48.9	52.8	19.8	4.91	1.63
Jet 5	−15.9	6.5	−19.3	25.9	17.1	2.75	−0.97

	m_t	χ^2	Permutation
Fit 1	127.7	0.53	wwb_lb_h
Fit 2	128.6	1.97	b_lwb_hw
Fit 3	121.1	2.37	b_hwb_lw
	Average fitted mass: 126.5		

Run 84534, Event 15306					$H_T = 188.8$	$\mathcal{A} = 0.171$	
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Muon	−17.6	20.8	−10.8	29.4	27.3	2.27	−0.39
$\not\!\!E_T$	−12.5	27.1	0.0	29.9	29.9	2.00	0.00
Jet 1	−8.2	−41.0	−47.9	63.8	41.8	4.51	−0.98
Jet 2	27.0	−7.6	9.8	29.9	28.0	6.01	0.34
Jet 3	16.6	−13.1	9.2	23.4	21.2	5.62	0.42
Jet 4	−3.0	−18.7	0.9	19.2	19.0	4.55	0.05
Jet 5	14.3	8.0	22.7	28.1	16.4	0.51	1.13
Jet 6	−8.2	13.2	34.5	38.0	15.6	2.13	1.54

	m_t	χ^2	Permutation
Fit 1	104.0	3.52	wwb_hb_l
Fit 2	—	—	—
Fit 3	—	—	—
Average fitted mass: 104.0			

Run 86214, Event 13721*				$H_T = 365.8$	$\mathcal{A} = 0.187$		
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Muon	17.1	−4.4	−9.1	19.9	17.7	6.03	−0.49
$\not{E}_T$	4.3	−21.6	0.0	22.0	22.0	4.91	0.00
Jet 1	79.0	105.8	−34.0	136.9	132.0	0.93	−0.25
Jet 2	47.0	−74.7	−47.8	101.1	88.3	5.27	−0.52
Jet 3	−66.6	−49.6	−25.1	87.0	83.0	3.78	−0.30
Jet 4	−36.3	59.3	93.3	116.7	69.6	2.12	1.10
Jet 5	−20.9	26.4	−17.4	38.1	33.7	2.24	−0.49
Jet 6	25.6	−10.4	−15.5	31.9	27.6	5.90	−0.53
Jet 7	−15.2	2.0	−51.7	53.9	15.3	3.01	−1.93

	m_t	χ^2	Permutation
Fit 1	—	—	—
Fit 2	—	—	—
Fit 3	—	—	—
	Average fitted mass: —		

(The permutation $b_l w w b_h$ ($\chi^2 = 4.97$, $m_t = 238.7\,\mathrm{GeV}/c^2$) was discarded due to a large initial p_z^ν. All other permutations have $\chi^2 > 7$.)

Run 86282, Event 13774				$H_T = 179.8$	$\mathcal{A} = 0.047$		
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Muon	−18.1	−6.4	1.9	19.3	19.2	3.48	0.10
$\not\!\!E_T$	−48.6	−3.2	0.0	48.7	48.7	3.21	0.00
Jet 1	85.6	−49.6	−94.7	137.6	99.0	5.76	−0.85
Jet 2	−56.7	10.6	−10.8	59.3	57.7	2.96	−0.19
Jet 3	8.3	16.9	−9.9	21.5	18.8	1.12	−0.50
Jet 4	16.2	1.4	9.3	18.9	16.3	0.09	0.54

	m_t	χ^2	Permutation
Fit 1	—	—	—
Fit 2	—	—	—
Fit 3	—	—	—
	Average fitted mass: —		

(The permutations wb_lb_hw ($\chi^2 = 4.09$, $m_t = 116.0\,\mathrm{GeV}/c^2$),

wb_lwb_h ($\chi^2 = 4.16$, $m_t = 117.0\,\mathrm{GeV}/c^2$), and

wb_hb_lw ($\chi^2 = 4.14$, $m_t = 179.0\,\mathrm{GeV}/c^2$) were discarded due to large initial p_z^ν values.)

B.3 Electron + Jets + Tag Candidates

Run 57144, Event 15138*				$H_T = 217.2$	$\mathcal{A} = 0.017$		
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Electron	27.6	−41.9	−95.9	108.2	50.2	5.29	−1.40
$\not\!\!E_T$	43.4	−79.2	0.0	90.4	90.4	5.21	0.00
Jet 1†	45.9	85.3	117.9	152.8	96.9	1.08	1.03
Jet 1	54.6	99.8	137.8	179.0	113.8	1.07	1.02
Jet 2	−104.4	36.4	115.2	160.0	110.6	2.81	0.91
Jet 3	−32.9	28.7	15.8	46.7	43.7	2.42	0.35
Jet 4	12.2	11.5	9.5	19.5	16.8	0.76	0.54
Jet 5	−8.7	−13.2	76.1	77.8	15.8	4.13	2.28
Tag μ	4.3	7.3	10.0	13.1	8.5	1.03	1.00

	m_t	χ^2	Permutation
Fit 1	197.8	0.26	$b_h w b_l w$
Fit 2	—	—	—
Fit 3	—	—	—
	Average fitted mass: 197.8		

Run 62199, Event 13305*				$H_T = 228.4$	$\mathcal{A} = 0.037$		
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Electron	61.5	−21.9	91.4	112.3	65.3	5.94	1.14
$\not\!\!E_T$	−8.9	−25.4	0.0	26.9	26.9	4.37	0.00
Jet 1†	−48.5	62.4	5.1	79.7	79.0	2.23	0.06
Jet 1	−67.5	85.6	5.8	109.7	109.0	2.24	0.05
Jet 2	83.1	−8.1	−40.1	93.0	83.5	6.19	−0.46
Jet 3	−48.7	−0.7	17.9	52.1	48.7	3.16	0.36
Jet 4	−39.0	−6.9	58.8	71.6	39.6	3.32	1.19
Jet 5	25.1	−20.0	−23.7	40.2	32.1	5.61	−0.68
Tag μ	−9.5	11.6	0.3	15.0	15.0	2.26	0.02

	m_t	χ^2	Permutation
Fit 1	—	—	—
Fit 2	—	—	—
Fit 3	—	—	—
	Average fitted mass: —		

(None of the tagged permutations gave a good fit. The best tagged permutation is $b_h w w b_l$ with $\chi^2 = 8.46$, $m_t = 161.2\,\mathrm{GeV}/c^2$.)

Run 85129, Event 19079*					$H_T = 162.2$	$\mathcal{A} = 0.093$	
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Electron	20.1	−38.0	27.3	51.0	43.0	5.20	0.60
$\not\!\!E_T$	15.5	35.0	0.0	38.3	38.3	1.15	0.00
Jet 1†	−49.5	−23.6	2.2	55.1	54.8	3.59	0.04
Jet 1	−60.9	−27.3	2.1	67.2	66.8	3.56	0.03
Jet 2	−58.8	−5.4	−58.0	83.2	59.0	3.23	−0.87
Jet 3	21.2	29.3	8.2	37.3	36.1	0.94	0.22
Jet 4	25.6	−11.7	−12.9	31.2	28.2	5.86	−0.44
Jet 5	14.2	16.0	−49.5	54.0	21.4	0.84	−1.57
Tag μ	−5.7	−1.8	0.0	6.0	6.0	3.45	−0.01

	m_t	χ^2	Permutation
Fit 1	138.8	0.32	$b_l w b_h w$
Fit 2	126.4	0.47	$b_h w b_l w$
Fit 3	126.3	1.89	$b_h w w b_l$
	Average fitted mass: 131.6		

B.4 Muon + Jets + Tag Candidates

Run 58192, Event 137*				$H_T = 246.7$	$\mathcal{A} = 0.049$		
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Muon	−42.9	−62.3	−6.6	75.9	75.7	4.11	−0.09
$\not\!\!E_T$	−8.1	−77.2	0.0	77.6	77.6	4.61	0.00
Jet 1	49.0	135.8	90.8	171.9	144.3	1.22	0.59
Jet 2†	38.2	26.3	53.8	71.1	46.3	0.60	0.99
Jet 2	58.9	39.8	81.8	108.5	71.0	0.59	0.98
Jet 3	−10.7	−41.0	−90.1	99.7	42.4	4.46	−1.50
Jet 4	−33.5	22.9	−9.7	42.1	40.6	2.54	−0.24
Tag μ	10.3	6.8	14.0	18.7	12.4	0.58	0.97

	m_t	χ^2	Permutation
Fit 1	277.0	0.38	$b_h b_l w w$
Fit 2	247.0	2.30	$b_l b_h w w$
Fit 3	247.1	3.98	$w b_l b_h w$
	Average fitted mass: 266.4		

Run 58203, Event 4980*					$H_T = 200.2$	$\mathcal{A} = 0.056$	
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Muon	−20.7	76.7	44.1	90.9	79.5	1.83	0.53
$\not\!\!E_T$	−38.8	29.8	0.0	48.9	48.9	2.49	0.00
Jet 1	26.9	−129.7	16.8	134.5	132.4	4.92	0.13
Jet 2†	40.4	−16.1	12.9	45.7	43.5	5.90	0.29
Jet 2	65.6	−30.4	18.4	75.2	72.3	5.85	0.25
Jet 3	13.7	38.5	65.8	77.8	40.9	1.23	1.25
Jet 4	−17.3	−8.0	−14.9	24.2	19.1	3.57	−0.72
Tag μ	12.6	−7.2	2.8	14.8	14.5	5.77	0.19

	m_t	χ^2	Permutation
Fit 1	199.3	0.09	$b_h b_l w w$
Fit 2	144.8	0.10	$w b_h b_l w$
Fit 3	206.6	0.10	$w b_l b_h w$
	Average fitted mass: 183.6		

Run 84695, Event 29699*				$H_T = 182.2$	$\mathcal{A} = 0.071$		
Object	p_x	p_y	p_z	E	E_T	ϕ	η
Muon	53.6	−23.7	4.9	58.8	58.6	5.87	0.08
$\not\!\!E_T$	50.2	−46.7	0.0	68.6	68.6	5.53	0.00
Jet 1	−107.8	−57.9	−82.3	148.0	122.4	3.63	−0.63
Jet 2†	27.4	40.1	−12.8	50.5	48.6	0.97	−0.26
Jet 2	68.3	100.0	−32.0	125.5	121.1	0.97	−0.26
Jet 3	0.5	35.0	49.1	60.5	35.0	1.56	1.14
Tag μ	20.4	29.9	−9.6	37.5	36.2	0.97	−0.26

	m_t	χ^2	Permutation
Fit 1	—	—	—
Fit 2	—	—	—
Fit 3	—	—	—
	Average fitted mass: —		

(Only three jets.)

Bibliography

[1] S. Abachi *et al.* (DØ). *Phys. Rev. Lett.*, **74**:2632, April 1995.
(URL: `http://d0wop.fnal.gov/d0pubs/journals/PRL_top`)

[2] F. Abe *et al.* (CDF). *Phys. Rev. Lett.*, **74**:2626, April 1995.
(URL: `http://www-cdf.fnal.gov/top_status/prl_cdf.ps`)

[3] Particle Data Group. *Phys. Rev.*, **D50**(3):1173–1876, August 1994.
(URL: `http://pdg.lbl.gov/`)

[4] The program for this calculation was written by Gerrit Burgers (CERN/TH).

[5] Gordon L. Kane. *Modern Elementary Particle Physics.* Addison-Wesley, 1993. (Updated edition).

[6] Chris Quigg. *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions.* Addison-Wesley, 1983.

[7] Lewis H. Ryder. *Quantum Field Theory.* Cambridge, 1985.

[8] Vernon D. Barger and Roger J. N. Phillips. *Collider Physics.* Addison-Wesley, 1987.

[9] Gordon L. Kane. Top quark topics. In J. Hawthorne, editor, *Gauge Bosons and Heavy Quarks: Proceedings of the Eighteenth SLAC Summer Institute on Particle Physics*, pages 123–142. SLAC-REPORT-378, July 1990.

[10] Paula J. Franzini. *Phys. Rep.*, **173**:1–62, 1989.

[11] E. Elsen *et al.* (JADE). *Z. Physik*, **C46**:349–359, 1990.

[12] H. J. Behrend *et al.* (CELLO). *Z. Physik*, **C47**:333–342, 1990.

[13] A. Shimonaka *et al.* (TOPAZ). *Phys. Lett.*, **B268**:457–464, 1991.

[14] Gordon L. Kane and Michael E. Peskin. *Nucl. Phys.*, **B195**:29–38, 1982.

[15] A. Bean *et al.* (CLEO). *Phys. Rev.*, **D35**:3533–3536, 1987.

[16] C. Albajar *et al.* (UA1). *Phys. Lett.*, **B262**:163–170, 1991.

[17] E. Laenen, J. Smith, and W. L. van Neerven. *Phys. Lett.*, **B321**:254–258, 1994.
(URL: `http://xxx.lanl.gov/abs/hep-ph/9310233`)

[18] P. Nason, S. Dawson, and R. K. Ellis. *Nucl. Phys.*, **B303**:607–633, 1988.

[19] P. Nason, S. Dawson, and R. K. Ellis. *Nucl. Phys.*, **B327**:49–92, 1989.

[20] P. Nason, S. Dawson, and R. K. Ellis. *Nucl. Phys.*, **B335**:260, 1990.

[21] W. Beenakker, H. Kuijf, W. L. van Neerven, and J. Smith. *Phys. Rev.*, **D40**:54–82, 1989.

[22] W. Beenakker, W. L. van Neerven, R. Meng, G. A. Schuler, and J. Smith. *Nucl. Phys.*, **B351**:507–560, 1991.

[23] E. Laenen. A note on the lower limit of the top quark production rate at the Tevatron. Fermi-Pub 93/155-T, FNAL, June 1993.

[24] S. Abachi *et al.* (DØ). *Phys. Rev. Lett.*, **72**:2138, 1994.
(URL: `http://d0wop.fnal.gov/d0pubs/journals/PRL72.2138/PRL72.2138.html`)

[25] S. Abachi *et al.* (DØ). Search for the top quark in DØ. (*Phys. Rev. D* paper in preparation.).

[26] F. Abe *et al.* (CDF). *Phys. Rev. Lett.*, **73**:220, 1994.

[27] Darien Wood. Electroweak physics results from hadronic interactions. In *Proceedings of the Physics in Collision Conference*, Tallahasse, Florida, June 1994. DØ note 2296.

[28] F. Abe *et al.* (CDF). *Phys. Rev.*, **D43**:664–686, 1991.

[29] F. Abe *et al.* (CDF). *Phys. Rev.*, **D45**:3921–3948, 1992.

[30] D. Decamp *et al.* (ALEPH). *Phys. Lett.*, **B236**:511–522, 1990.

[31] G. S. Abrams *et al.* (MARK II). *Phys. Rev. Lett.*, **63**:2447–2451, 1989.

[32] D. Schaile. Precision tests of the electroweak interaction. CERN PPE/94-162-2, CERN, December 1994.
(URL: http://preprints.cern.ch/archive/electronic/cern/9410/ppe94-162-2.ps.Z)

[33] F. Abe *et al.* (CDF). *Phys. Rev.*, **D50**:2966, 1994.
(URL: http://www-cdf.fnal.gov/physics/pub94/cdf2561_top_prd.ps)

[34] F. Abe *et al.* (CDF). *Phys. Rev. Lett.*, **73**:225, 1994.
(URL: http://www-cdf.fnal.gov/physics/pub94/cdf2595_top_prl.ps)

[35] S. Abachi *et al.* (DØ). *Phys. Rev. Lett.*, **74**:2422, March 1995.
(URL: http://d0wop.fnal.gov/d0pubs/journals/PRLTOP/PRLTOP.html)

[36] S. Abachi *et al. Nucl. Inst. Meth.*, **A338**:185–253, 1994.
(URL: http://d0wop.fnal.gov/d0pubs/journals/NIMA338.185/NIMA338.185.html)

[37] N. Amos, J. Linnemann, R. Partridge, and L. Paterno. Luminosity calculations for DØ. DØ note 2031, 1994.

[38] N. Amos *et al.* Change to the DØ luminosity monitor constant. DØ note 2186, 1994.

[39] Leon M. Lederman. *Scientific American*, **264**(3):48–55, March 1991.

[40] Helen T. Edwards. *Ann. Rev. Nucl. Part. Sci.*, **35**:605–660, 1985.

[41] FNAL. Design report Tevatron 1 project. FNAL internal note, September 1984.

[42] FNAL. A report of the design of the Fermi National Accelerator Laboratory superconducing accelerator. FNAL internal note, May 1979.

[43] William Joseph Thompson. *Search for the Top Quark in the Muon + Jets channel at DØ.* PhD thesis, State University of New York at Stony Brook, Stony Brook, New York, February 1994.

[44] M. P. Church and J. P. Marriner. *Ann. Rev. Nucl. Part. Sci.*, **43**:253–295, 1993.

[45] D. Möhl, G. Petrucci, L. Thorndahl, and S. van der Meer. *Phys. Rep.*, **C58**:73–119, 1980.

[46] Srinivasan Rajagopalan. *The dE/dx Capabilities of the DØ Tracking System.* PhD thesis, Northwestern University, Evanston, Illinois, June 1992.

[47] James Herbert Cochran, Jr. *Search for the Truth in the eμ Channel at DØ.* PhD thesis, State University of New York at Stony Brook, Stony Brook, New York, December 1993.

[48] Ties Behnke. *The Central Drift Chamber for the DØ Detector: Design, Construction, and Test.* PhD thesis, State University of New York at Stony Brook, Stony Brook, New York, 1989.

[49] Richard C. Fernow. *Introduction to Experimental Particle Physics*, chapter 10, pages 234–258. Cambridge University Press, 1986.

[50] Konrad Kleinknecht. *Detectors for Particle Radiation.* Cambridge University Press, 1987.

[51] F. Sauli. Principles of operation of multiwire proportional and drift chambers. In Thomas Ferbel, editor, *Experimental Techniques in High Energy Physics*, pages 79–188. Addison-Wesley, 1987.

[52] A. R. Clark *et al. Nucl. Inst. Meth.*, **A261**:420–426, 1987.

[53] A. R. Clark *et al. Nucl. Inst. Meth.*, **A279**:243–248, 1989.

[54] A. R. Clark *et al. Nucl. Inst. Meth.*, **A315**:193–196, 1992.

[55] D. Buchholz *et al. Nucl. Inst. Meth.*, **A257**:556–566, 1987.

[56] Domenico Pizzuto. *D0 Central Tracking Chamber Performance Studies.* PhD thesis, State University of New York at Stony Brook, Stony Brook, New York, December 1991.

[57] Jeffrey W. Bantly. *The D0 Detector Forward Drift Chamber Performance and Physics Capability in the 1990 FNAL Testbeam Run.* PhD thesis, Northwestern University, Evanston, Illinois, June 1992.

[58] Robert E. Avery *et al. IEEE Trans. Nucl. Sci.*, **40**(4):573–577, 1993.

[59] J. F. Detœuf *et al. Nucl. Inst. Meth.*, **A265**:157–166, 1988.

[60] J. F. Detœuf *et al. Nucl. Inst. Meth.*, **A279**:310–316, 1989.

[61] Y. Ducros *et al. Nucl. Inst. Meth.*, **A277**:401–406, 1989.

[62] R. G. Angstadt *et al. IEEE Trans. Nucl. Sci.*, **39**(5):1297–1301, 1992.

[63] C. Fabjan. Calorimetry in high-energy physics. In Thomas Ferbel, editor, *Experimental Techniques in High Energy Physics*, pages 257–324. Addison-Wesley, 1987.

[64] Richard C. Fernow. *Introduction to Experimental Particle Physics*, chapter 11, pages 259–284. Cambridge University Press, 1986.

[65] Richard Wigmans. *Nucl. Inst. Meth.*, **A259**:389–429, 1987.

[66] Jaehoon Yu. *Determination of the Strong Coupling Constant (α_s) and a Test of Perturbative QCD Using W + jets Processes in the DØ detector.* PhD thesis, State University of New York at Stony Brook, Stony Brook, New York, August 1993.

[67] Anthony Spadafora, July 1992.

[68] Stephen J. Wimpenny *et al. Nucl. Inst. Meth.*, **A279**:107–113, 1989.

[69] Anthony L. Spadafora *et al. Nucl. Inst. Meth.*, **A315**:279–284, 1992.

[70] Paolo Franzini *et al. Nucl. Inst. Meth.*, **A289**:438–445, 1990.

[71] Hiroaki Aihara *et al. IEEE Trans. Nucl. Sci.*, **38**(2):398–402, 1991.

[72] C. Brown *et al. Nucl. Inst. Meth.*, **A279**:331–338, 1989.

[73] J. M. Butler *et al. Nucl. Inst. Meth.*, **A290**:122–130, 1990.

[74] Yu. M. Antipov *et al. Nucl. Inst. Meth.*, **A297**:121–125, 1990.

[75] Maris Abolins *et al.* The level one framework: D0 note 328 revised. DØ note 705, 1988.

[76] Maris Abolins *et al. IEEE Trans. Nucl. Sci.*, **36**(1):384–389, 1989.

[77] Maris Abolins, Daniel Edmunds, Philippe Laurens, and Bo Pi. *Nucl. Inst. Meth.*, **A289**:543–560, 1990.

[78] G. S. Gao and R. Partridge. *IEEE Trans. Nucl. Sci.*, **38**(2):286–289, 1991.

[79] J. Bantly *et al.* The level 0 trigger for the DØ detector. DØ note 1996, 1993. Submitted to IEEE Transactions on Nuclear Science.

[80] John M. Butler. Main ring deadtime. DØ note 1682, 1993.

[81] Daniel Edmunds. Changes to the first level trigger timing. DØ note 827, 1989.

[82] M. Fortner *et al. IEEE Trans. Nucl. Sci.*, **38**(2):480–485, 1991.

[83] D. Cutts *et al. IEEE Trans. Nucl. Sci.*, **36**(1):738–739, 1989.

[84] David Cullen-Vidal *et al.* DØ level-2/data acquisition; the new generation. In *Proceedings of the Conference on Computing in High-Energy Physics*, Tsukuba, Japan, March 1991.

[85] Dave Cutts *et al.* Operation of the D0 data acquistion system. In *Proceedings of the Conference on Computing in High-Energy Physics*, pages 262–264, Annecy, France, 1992. CERN 92-01.

[86] CERN. *Zebra*, 1992. CERN Program Library Number Q100.
(URL: `http://www.hep.net/cernlib/docs/www.dir/zebraold/ZEBRAMAIN.html`)
(URL: `http://wwwcn.cern.ch/asdoc/zebra/ZEBRAMAIN.html`)

[87] James T. Linnemann *et al.* The DØ software trigger. In *Proceedings of the Conference on Computing in High-Energy Physics*, pages 199–201, Annecy, France, 1992. CERN 92-01.

[88] John Featherly. *ITC User's Guide*, 1991. (unpublished).
(URL: `http://d0hs19.fnal.gov/rnodocs/itc.html`)

[89] Bruce Gibbard *et al.* Run control and resource management in the D0 run time system. In *Proceedings of the Conference on Computing in High-Energy Physics*, pages 265–268, Annecy, France, 1992. CERN 92-01.

[90] Bruce Gibbard and Scott Snyder. *"Taker" A Generic Human Interface for D-Zero Data Taking*, 1993. (unpublished).
(URL: `http://d0hs19.fnal.gov/rnodocs/taker.html`)

[91] Bruce Gibbard. *Working with COOR's Configuration Files*, 1993. (unpublished).
(URL: `http://d0hs19.fnal.gov/rnodocs/coor_configurations.html`)

[92] Scott Snyder. *Trigparse*, 1994. (unpublished; available at `d0$configs$coor_sim:trigparse.doc`).

[93] Scott Snyder. *The DØ Data Logger*, 1993. (unpublished).
(URL: `http://d0hs19.fnal.gov/daqdocs/html/logger_toc.html`)

[94] David Quarrie. Data acquisition system programmers reference manual. Computing Division manual PN251.04, FNAL, 1990.

[95] S. Banerjee, F. Bruyant, R. Mount, and E. Nagy. *A Software Utility Package for the L3 Data Bases*, 1992. (unpublished).
(URL: `http://d0wop.fnal.gov/offline/d0library/dbl3/docs/dbl3.doc.html`)

[96] Adam Para *et al.* The D0 offline production and analysis. In *Proceedings of the Conference on Computing in High-Energy Physics*, pages 593–596, Annecy, France, 1992. CERN 92-01.

[97] Jonathan Kotcher. *Response of the D0 Calorimeter to Cosmic Ray Muons.* PhD thesis, New York University, New York, New York, October 1992.

[98] Myungyun Pang. *Search for top in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.8 TeV by constrained kinematic fitting.* PhD thesis, Iowa State University, Ames, Iowa, 1994.

[99] Saul Youssef. *Comp. Phys. Comm.*, **45**:423–426, 1987.

[100] Dhiman Chakraborty. *A search for $t\bar{t} \rightarrow electron + \not\!\!E_T + jets$ signature in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.8 TeV with the DØ detector.* PhD thesis, State University of New York at Stony Brook, Stony Brook, New York, September 1994.

[101] T. C. Awes *et al. Nucl. Inst. Meth.*, **A311**:130–138, 1991.

[102] R. Engelmann *et al. Nucl. Inst. Meth.*, **A216**:45–56, 1983.

[103] Meena Narain. Electron identification in the DØ detector. In Carl H. Albright *et al.*, editors, *The Fermilab Meeting: DPF 92*, volume 2, pages 1678–1680. World Scientific, November 1992.

[104] Meena Narain. Electron identification in the DØ detector. DØ note 1548, 1992.

[105] H. T. Diehl and D. Hedin. D0 muon level 2 filter. DØ note 1325, January 1992.
(URL: `http://d0wop.fnal.gov/offline/d0library/muon_util/muon_l2.mem.html`)
(URL: `http://d0wop.fnal.gov/offline/d0library/muon_util/muon_analysis.mem.html`)

[106] C. Gerber *et al.* Muon momentum determination. DØ note 2140, October 1994.

[107] Jim Cochran, private communication.

[108] H. T. Diehl. Cosmic ray muon rejection in the level 2 filter at D0. DØ note 1517, 1992.

[109] G. Arnison *et al.* (UA1). *Phys. Lett.*, **B123**:115–122, 1983.

[110] G. Arnison *et al.* (UA1). *Phys. Lett.*, **B132**:214–222, 1983.

[111] Marc Paterno. *A search for Squarks and Gluinos in $p\bar{p}$ Collisions at $\sqrt{s}$ = 1.8 TeV with the DØ Detector.* PhD thesis, State University of New York at Stony Brook, Stony Brook, New York, May 1994.

[112] Andrew James Milder. *Dijet Angular Distributions at $\sqrt{s}$ = 1800 GeV Using the DØ Detector.* PhD thesis, University of Arizona, 1993.

[113] N. J. Hadley. Cone algorithm for jet finding. DØ note 904, November 1989.

[114] F. Abe *et al.* (CDF). *Phys. Rev. Lett.*, **69**:2896–2900, 1992.

[115] Victor Daniel Elvira. *Measurement of the Inclusive Jet Cross Sections at $\sqrt{s}$ = 1.8 TeV with the DØ Detector.* PhD thesis, Universidad de Buenos Aries, Argentina, 1994.

[116] Bob Kehoe. Resolution bias in jet response measurement. DØ note 2052, 1994.

[117] R. Astur. Correction of MET. DØ note 2088, 1994.

[118] Balamurali V. *Search for the Top Quark in the Dielectron Channel in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.8 TeV.* PhD thesis, University of Notre Dame, Notre Dame, Indiana, November 1994.

[119] Raymond Edward Hall. *Search for the Top Quark in Dimuon Events at DØ.* PhD thesis, University of California Riverside, Riverside, California, June 1994.

[120] Frank Paige and Serban Protopopescu. ISAJET. BNL Report 38034, Brookhaven, 1986. Release v6.49. (Unpublished).

[121] CERN. GEANT *Detector Description and Simulation Tool*, 1993. CERN Program Library Number Q123.
(URL: `http://www.hep.net/cernlib/docs/www.dir/geant/GEANTMAIN.html`)
(URL: `http://wwwcn.cern.ch/asdoc/geant/GEANTMAIN.html`)

[122] R. Raja, Q. Li-Demarteau, and S. Kunori. DØGEANT. DØ note 1007, July 1990.

[123] W. G. D. Dharmaratna, R. Raja, and C. Stewart. The DØ shower library — version 2.0. DØ note 1730, May 1993.

[124] Norman Graf. The DØ shower library. In D. F. Anderson *et al.*, editors, *Proceedings of the International Conference on Calorimetry in High Energy Physics*, page 539. World Scientific, October 1990.

[125] Level 1 simulator user's guide. (unpublished; available at `d0$level1:l1sim.doc`).
(URL: `http://d0wop.fnal.gov/offline/d0library/level1/l1sim.doc.html`)

[126] Dan Claes. L2SIM. (unpublished; available at `d0$level2:l2sim.doc`).
(URL: `http://d0wop.fnal.gov/offline/d0library/level2/l2sim.doc.html`)

[127] G. Marchesini *et al. Comp. Phys. Comm.*, **67**:465–508, 1992.

[128] W. Giele, E. Glover, and D. Kosower. *Nucl. Phys.*, **B403**:633–667, 1993.
(URL: `http://xxx.lanl.gov/abs/hep-ph/9302225`)

[129] G. Hanson *et al. Phys. Rev.*, **D26**:991–1012, 1982.

[130] Paul D. Grannis (DØ). Observation of the top quark, March 1995. (Talk given at Fermilab seminar).

[131] F. A. Berends, H. Kuijf, B. Tausk, and W. T. Giele. *Nucl. Phys.*, **B357**: 32–64, 1991.

[132] Sailesh Chopra, private communication.

[133] Tom Rockwell, private communication.

[134] Ssu-Min Chang, Peter Tamburello, and Bill Cobau, private communication.

[135] John Hobbs, private communication.

[136] Tom Ferbel *et al.* Kinematic fitting for the mass of the top quark. DØ note 2408, March 1995.

[137] Myungyun Pang, private communication.

[138] Lynne H. Orr, T. Stelzer, and W. J. Stirling. Gluon radiation in $t\bar{t}$ production at the Tevatron $p\bar{p}$ collider. Preprint UR-1397, University of Rochester, November 1994.
(URL: http://xxx.lanl.gov/abs/hep-ph/9412294)

[139] Boaz Klima and Chip Stewart. PJET — parton jets in the Monte Carlo. DØ note 905, December 1989.

[140] Sailesh Chopra and Rajendran Raja. Estimation of the QCD background to $W \rightarrow e\nu +$ jets. DØ note 2098, April 1994.

[141] CERN. HBOOK *Reference Manual*, 1993. CERN Program Library Number Y250.
(URL: http://www.hep.net/cernlib/docs/www.dir/hbook/HBOOKMAIN.html)
(URL: http://wwwcn.cern.ch/asdoc/hbook/HBOOKMAIN.html)

[142] John Allison. *Comp. Phys. Comm.*, **77**:377–395, 1993.

[143] E. T. Jaynes. *Probability Theory: The Logic of Science.* June 1994. (Unpublished).
(URL: http://www.math.albany.edu:8008/JaynesBook.html)

[144] A. G. Frodesen, O. Skjeggestad, and H. Tøfte. *Probability and Statistics in Particle Physics.* Universitetsforlaget, 1979.

[145] F. James. MINUIT *Minimization Package Reference Manual*, 1994. CERN Program Library Number D506.
(URL: http://wwwcn.cern.ch/asdoc/WWW/minuit/minmain/minmain.html)

[146] Harrison Prosper, private communication.

[147] Rajendran Raja. Consistency of the DØ top discovery numbers with the standard model of top production. DØ note 2488, February 1995.
(URL: http://d0wop.fnal.gov/physics_analysis/d0notes/source/note2488.ps)

[148] J. Peter Berge, Frank T. Solmitz, and Horace D. Taft. *Review of Scientific Instruments*, **32**(5):538–548, May 1961.

[149] O. I. Dahl, T. B. Day, F. T. Solmitz, and N. L. Gould. SQUAW kinematic fitting program. Technical Report P-126, Lawrence Radiation Laboratory, Berkeley, California, July 1968.