Search for a *W′* boson in the top-quark decay channel

Anne Fabricant

Advisor: Meenakshi Narain

Brown University

May 8, 2009

Abstract

Many theoretical extensions of the Standard Model of particle physics predict the existence of undiscovered charged massive bosons (*W*′). This analysis looks at 2.3 fb^{-1} of data collected by the DØ detector at the Tevatron and reconstructed using the p17/p20 reconstruction software. In particular, Boosted Decision Trees (BDTs) are used to search for a *W*′ produced in proton-antiproton collisions which decays to a top quark and a bottom quark. The BDTs are trained using simulated *W*′ and Standard-Model processes, then applied to the collected data in order to separate signal and background events. My search focuses on events whose final state consists of a muon, missing transverse energy and two jets, one of which is tagged as a *b*-quark jet. We see no evidence of *W*′ production; limits are set on the cross section for *W*′ production and subsequent decay to *tb*. By comparing these limits to those obtained via previous methods, it is possible to evaluate the effectiveness of the BDT analysis.

Contents

Acknowledgements

I would like to thank Meenakshi Narain for advising me, and graduate student Monica Pangilinan for her frequent and patient guidance. In addition, graduate student Yunhe Xie provided much-appreciated tips.

I. Introduction

Many extensions of the Standard Model (SM) of particle physics predict the existence of undiscovered charged massive bosons. These hypothetical mediators of the weak interaction are called W', in analogy with the SM W bosons. To search for evidence of W' boson production, 2.3 fb^{-1} data was collected at the DØ detector between August 2002 and August 2007. Data analysis published in 2008 [1] examined three different models of production in which W' decays to a top quark and a bottom quark ($W' \rightarrow tb$). The approach was to look for significant excess in the invariant-mass distribution of the final state, which consists of a lepton, jets and missing transverse energy. While no such excess was found, upper limits were set on the cross section for W' production and subsequent decay to tb.

The 2008 results are summarized in Fig. 1 below. Fig. 1a shows the expected limits and 95% confidence-level limits for a left-handed W' boson (W'_L) as a function of its mass. Also included are the theoretical cross sections; these are based on the assumption that a W' with left-handed SM couplings will interfere with the W boson in the s-channel single-top-quark process. All theoretical cross sections above the curve determined by the 95% confidence-level limits are excluded by the analysis, establishing a lower mass limit of 731 GeV.[*] Fig. 1b displays analogous information for a right-handed W' boson (W'_R). In this case the limits depend on the W' mass as compared to that of the hypothetical right-handed neutrino ν_R (see chapter II). The W' lower mass limits are 738 GeV if $M(\nu_R) < M(W'_R)$ and 768 GeV if $M(\nu_R) > M(W'_R)$.[†]

The present analysis attempts to improve upon the previously published results by employing the method of boosted decision trees (BDTs). Our search focuses on s-channel events arising from a W' with approximate mass 750 GeV, near the current lower mass limit. In particular, we examine final states consisting of a muon, missing transverse energy and two jets, one of which is tagged as a b-quark jet. Rather than exclusively using the final-state invariant-mass distribution to separate signal and background events, the BDTs incorporate 31 discriminating variables. In the absence of evidence for W' production and subsequent decay to tb, we set new cross-section limits based on the events we have selected. Comparing these limits to those obtained using the invariant-mass distribution alone, we find significant improvement.

[*] Particle physics uses the natural system of units, for which $\hbar = c = 1$. Hence energy, momentum and mass are all expressed in eV. Using other unit systems, momentum is expressed in eV/c and mass is expressed in eV/c^2.
[†] A full explanation of cross-section limits will be given in chapter VI.

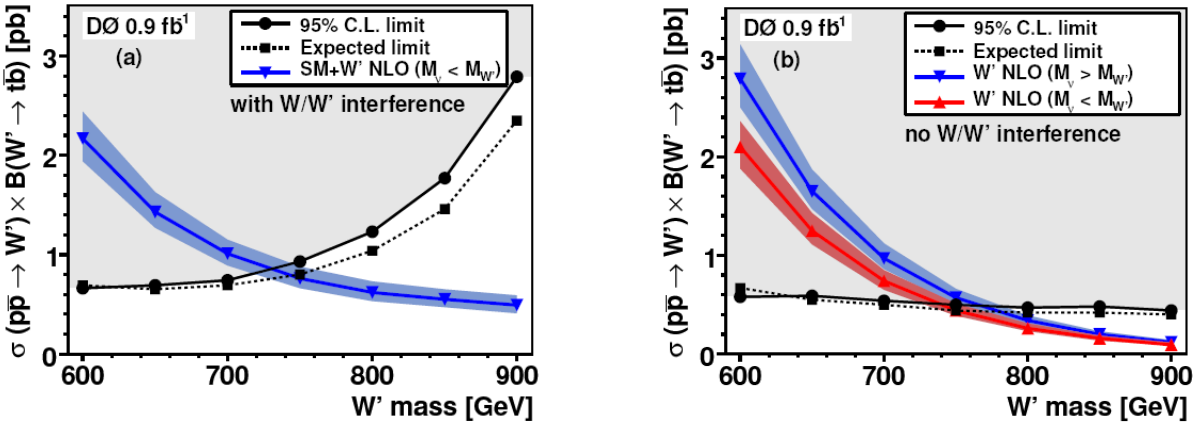

Figure 1. Results of the 2008 *W'* analysis [1]. Left to right: **(a)** 95% confidence-level upper limits for a left-handed *W'*, including interference effects with the SM *W* boson; the shaded regions are excluded by the analysis. Also shown are the expected limits and theoretical cross sections. **(b)** Analogous information for the case of a right-handed *W'*, for which no interference effects are included.

II. Theory

In order to motivate our search for the hypothetical *W*′ boson, we first need to understand the Standard Model of particle physics. This will prepare us for a brief overview of Standard-Model extensions that predict the existence of undiscovered charged massive bosons. Next we will examine the specific processes by which a *W*′ boson could be produced and subsequently decay to *tb*. These processes vary depending on the boson's mass and couplings. As we saw above, recent experiments have already constrained the possible masses of a *W*′ boson with exclusively left-handed or right-handed couplings.

The Standard Model[‡]

All the constituents of the Standard Model were discovered over the course of a century, beginning with the electron in 1897 and ending (for now) with the tau neutrino in 2000. The Standard Model describes three types of particle interactions[§] and their associated charges: electromagnetic (electric charge), strong nuclear (color charge) and weak nuclear (weak isospin charge). The electromagnetic and weak nuclear processes have actually been unified—both are manifestations of the electroweak force. Current SM physics deals with energy scales up to about 1 TeV (10^{12} eV) and length scales down to about 1 fm (10^{-15} m). (For comparison, human senses evolved to interpret physical processes with energies on the order of 1 eV). When the Large Hadron Collider (LHC) at CERN becomes operational, we expect to expand the energy frontier beyond 7 TeV.

We may divide the fundamental SM particles into two categories based on their spin, then further subdivide them based on other measurable properties. *Fermions* have half-integer spin while *bosons* have integer spin. The fermions comprise twelve matter particles and twelve antimatter particles; each matter particle has an antimatter counterpart of opposite electric charge. There are two groups of fermions: six *quarks* and six *leptons*. Quarks participate in all interactions, which is to say that they carry all types of charge. There are three quark varieties ("flavors") with electric charge +⅔ (up, charmed, top)

[‡] Most of the material in this section may be found in any standard text on particle physics. See, for example, [2].
[§] A note on terminology: Although the fundamental building blocks of the SM are referred to as *particles*, we should not be tempted to adopt a colloquial mechanistic picture of them. Note also that we are using the terms *interaction* and *force* interchangeably in this context.

and three with electric charge -⅓ (down, strange, bottom). Unlike quarks, leptons do not carry color charge. Three of them, the neutrinos, are also electrically neutral. Each neutrino corresponds to one of the three negatively charged leptons (electron, muon, tau).

The fermions—quarks and leptons—are arranged in three generations of increasing mass (Table 1). Mixing between quark flavors is greatest for quarks of neighboring generations. Both quarks and leptons come in pairs; the upper member of each pair has weak isospin +½ while the upper member has weak isospin -½. We should note that quarks have not been found in isolation—fractional electric charges and free color charges are apparently not physically manifested. Instead we observe baryons, color-neutral combinations of three quarks bound together by the strong force. First-generation quarks compose the most familiar baryons, the proton (*uud*) and the neutron (*udd*). Quarks also combine via the strong force into mesons, which are color-neutral quark/antiquark bound states.

Table 1. The three generations of fermions (spin ½). Approximate masses are shown in parentheses. For each fermion there exists an antiparticle of opposite electric charge. Charges are given in units of *e*.

	Electric charge	Generation 1	Generation 2	Generation 3
Quarks	+⅔	$\begin{pmatrix} u\ (1.5 - 3.3\ \text{MeV}) \\ d\ (3.5 - 6.0\ \text{MeV}) \end{pmatrix}$	$\begin{pmatrix} c\ (1.27\ \text{GeV}) \\ s\ (104\ \text{MeV}) \end{pmatrix}$	$\begin{pmatrix} t\ (171\ \text{GeV}) \\ b\ (4.20\ \text{GeV}) \end{pmatrix}$
	-⅓			
Leptons	0	$\begin{pmatrix} \nu_e\ (\leq 3\ \text{eV}) \\ e^-\ (0.511\ \text{MeV}) \end{pmatrix}$	$\begin{pmatrix} \nu_\mu\ (\leq 0.2\ \text{MeV}) \\ \mu^-\ (106\ \text{MeV}) \end{pmatrix}$	$\begin{pmatrix} \nu_\tau\ (\leq 18\ \text{MeV}) \\ \tau^-\ (1.78\ \text{GeV}) \end{pmatrix}$
	-1			

Bosons are the quanta of the force fields, as described by quantum field theory; they mediate each type of fermion interaction. As shown in Table 2, there are five known types of bosons. The photon mediates the electromagnetic interaction but carries no charge (other than mass-energy). The weak-nuclear mediators are the *W* and *Z* bosons—both carry weak isospin, and the *W* bosons carry electric charge (±1). Finally we have the gluon, strong-force mediator. There are actually eight gluons, all of which carry color charge.

Table 2. The known bosons (integer spin). Not listed are the hypothetical graviton (massless, spin 2) and Higgs boson (mass > 100 GeV, spin 0). Spin is given in units of $\hbar/2\pi$.

Interaction	Boson(s)	Mass-energy (GeV)	Spin
Electromagnetic	Photon (γ)	0	1
Weak nuclear	W^+ W^- Z^0	80.40 80.40 91.19	1 1 1
Strong nuclear	Gluon (g) (8 total)	0	1

Two important concepts, *coupling* and *helicity*, will be important in our discussion. Coupling describes the relationship between a boson and the particles which interact via that boson. Helicity refers to the direction of a particle's spin relative to its momentum—either right-handed or left-handed. We say, for example, that a boson has right-handed couplings if it couples to right-handed fermions. Experiments show that neutrinos are left-handed and antineutrinos are right-handed, although the existence of a right-handed neutrino is not prohibited in theory. In fact, if the neutrino mass is found to be nonzero, then a right-handed neutrino *must* exist.

At this point, a bit more detail on the mathematical underpinnings (Lie algebra) of the Standard Model is desirable. Quantum chromodynamics (QCD), our model of the strong force, is based on SU(3) gauge symmetry. This corresponds to rotations in "color space" of three complex dimensions. Color space is an internal symmetry (non-physical) space. The three dimensions correspond to the three quark colors: red, blue and green. Combining all three color charges yields a net color charge of zero. The eight generators of SU(3) correspond to eight field quanta, the gluons.

Our model of the electroweak interaction is based on a combined U(1) and SU(2) gauge symmetry. The three generators of SU(2) give rise to three field quanta, two electrically charged (W^+ and W^-) and one electrically neutral (W^0). U(1) has a single generator which contributes an additional neutral field quantum B^0. The photon and the Z^0 boson are quantum-mechanical mixtures of the B^0 and W^0 quanta—only these mixtures are physically real.

The precise masses of SM particles are determined experimentally. Although the photon and gluon are massless, the *W* and *Z* bosons are massive. A crisis arises when we introduce these nonzero masses into

the mathematical theory: gauge symmetry is destroyed. We require a hypothetical scalar field, called the Higgs, in order to salvage the Standard Model. The idea is that the *W* and *Z* bosons, as well as the fermions, acquire mass by coupling to the Higgs field. The Higgs boson, which is predicted to be a massive spinless particle, is the quantum of the Higgs field.

While the Higgs boson is needed to make the Standard Model consistent, we still have not accounted for gravity. To incorporate gravity into the SM as a fourth type of particle interaction, we would need to discover yet another boson. This hypothetical "graviton," if it exists, is expected to be a massless particle of spin 2.

The incomplete nature of the Standard Model is strong motivation to perform experimental tests of theoretical SM extensions, such as the *W*′ search under consideration here. In addition, the weak nuclear interaction is arguably the most fascinating of the three SM interactions. The so-called weak force is not so much weak as short-ranged—its range of influence is much less than the radius of a proton. This short range is due to the fact that the field quanta, the *W* and *Z* bosons, are massive. Weak-nuclear interactions are also probably responsible for matter/antimatter symmetry in our universe. Yet another interesting property is parity violation. *W* bosons couple exclusively to left-handed particles; the *Z* boson shows a preference for left-handed particles, but right-handed couplings are observed. For all of these reasons, the discovery of a new weak-nuclear gauge boson and the observation of its properties are exciting propositions.

The *W*′ boson

As mentioned earlier, a *W*′ boson arises independently in various theoretical extensions of the Standard Model. These include the following:

- *Composite Higgs*. In this class of models, the Higgs field is constructed from the bound states of fermions and hence is a "composite" field. A composite Higgs boson emerges, as well as a new massive boson from the U(1) gauge group [3].
- *Noncommuting extended technicolor (ETC)*. Technicolor models fall under the category of composite Higgs, explaining the masses of the weak-interaction bosons without recourse to an elementary scalar Higgs field [4].

- *Little Higgs*. The mass of the Higgs boson is unstable under radiative SM corrections. This problem can be solved by Little Higgs models, in which global symmetry-breaking at high (TeV) energy scales gives rise to a light Higgs boson and new massive bosons like *W*′ and *Z*′ [5].
- *Composite gauge bosons*. These models abandon the Higgs mechanism altogether; a series of new excited weak-interaction bosons (including *W*′ and *Z*′) is constructed from composite states of quarks and leptons [6].
- *Grand Unification*. In models which unify quantum chromodynamics and the electroweak theory using gauge groups larger than SU(5), right-handed *W*′ bosons and various *Z*′ bosons are expected [7].
- *Superstring*/M theory attempts to unify all forces including gravity, thereby providing a quantum theory of gravity. Unfortunately, it is divorced from reality at this time. In addition to new gauge bosons, notable byproducts include 10-11 spacetime dimensions [8].
- *Universal extra dimensions* refers to a large class of models, including string theory, which deal with fields in higher dimensions [9].
- *Top flavor*. In these models, third-generation fermions take part in a different SU(2) weak interaction than do members of the first two generations. Depending on the gauge coupling parameters, the *W*′ couples more strongly to either the third generation or the first two [10].
- *Left-right symmetric*. Multiple variants have been studied; perhaps the most promising of these minimally extends SM gauge symmetry in four dimensions [11].

Additional references for the above can be found in Ref. [12]. While we may say in general that the hypothetical *W*′ emerges via gauge-group extensions, theoretical constraints on its properties vary considerably from model to model. Due to this extreme model dependence, experimental limits on the mass of *W*′ are desirable. Indirect limits have been extracted from leptonic and semileptonic decays, as well as from cosmological and astrophysical data. However, the indirect approach gives a wide range of upper mass limits, from 549 GeV to 23 TeV [13].[**] In the direct approach, we assume that the *W*′ has purely left-handed or purely right-handed SM couplings (justification for this is given in chapter IV).

A promising experimental technique in the *W*′ search is to look for decays to third-generation quarks (*tb*) via the top-quark decay channel. Benefits of this approach include smaller QCD backgrounds than for decays to light quarks, as well as reduced model dependence. The hypothetical s-channel process for *W*′

[**] The present author apologizes for this brief foray into experiment in an allegedly theoretical thesis chapter.

is depicted in Figure 2; the top quark may subsequently decay to a lepton plus neutrino (by way of *W*) and a bottom quark. In the s-channel process, a left-handed *W*′ interferes with the SM *W* boson, leading to destructive interference [14]. The mass of a right-handed *W*′ depends on the mass of the hypothetical right-handed neutrino ν_R. If $M(W'_R) < M(\nu_R)$, the decay to ν_R plus lepton is not allowed and W'_R decays only to quark-antiquark pairs. In this analysis we focus on the s-channel, but a t-channel process ($W' \rightarrow tqb$) is also possible.

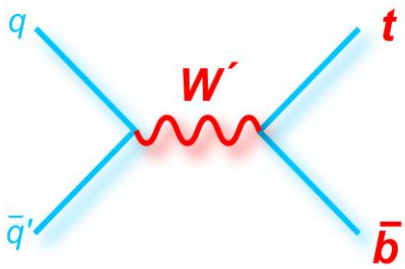

Figure 2. Hypothetical s-channel process for *W*′. Unless we are dealing with $M(W'_R) < M(\nu_R)$, the top quark can subsequently decay to a lepton (electron or muon), neutrino and bottom quark.

III. The DØ Detector

The Tevatron particle accelerator at Fermilab is a four-mile-long underground proton-antiproton collider. A chain of accelerators produce and accelerate particles, which are then injected into the Tevatron ring. Proton and antiproton beams with energy up to almost 1 TeV circle the ring in opposite directions and collide at the center of two 5000-ton detectors. Data for this experiment was collected by the DØ detector during Tevatron Run II. In preparation for Run II, which began in 2001, DØ received substantial upgrades. Full details of the enhancements are given in Ref. [15]. Each second, the Tevatron produces over two million proton-antiproton collisions and DØ records about 100 events.

As shown in Fig. 3, DØ is a multipurpose detector comprising three main systems: central tracking detectors, calorimeter and muon spectrometer. Let us work our way through the detector from the inside out.

- DØ's completely upgraded *central tracking* system now consists of a silicon microstrip tracker and a scintillating-fiber tracker within a 2-T solenoidal magnet. The silicon microstrip tracker is used for *b*-quark tagging. The magnetic field serves multiple purposes, including lepton identification and calorimeter calibration.
- *Preshower detectors* located between the magnet and central calorimeter aid in electron *identification.*
- Uranium/liquid-argon *calorimeters* primarily measure the energies of electrons, photons, and jets[††], as well as the transverse energy of events.
- The *muon tracking* system is made up of drift tubes, toroidal magnets, tracking detectors and scintillation trigger counters, all of which work together for effective muon identification and measurement.
- A *luminosity monitor* determines the Tevatron's luminosity in the DØ region by detecting inelastic proton-antiproton collisions.
- The *trigger* system looks for interesting physics and selects events to be recorded by the *data acquisition* system.
- DØ operations are coordinated by the *detector controls and monitoring* system.
- Extensive computer *software* is needed for data collection, hardware operation, and event simulation and reconstruction.

[††] A particle jet is a stream of hadrons produced by a quark or gluon.

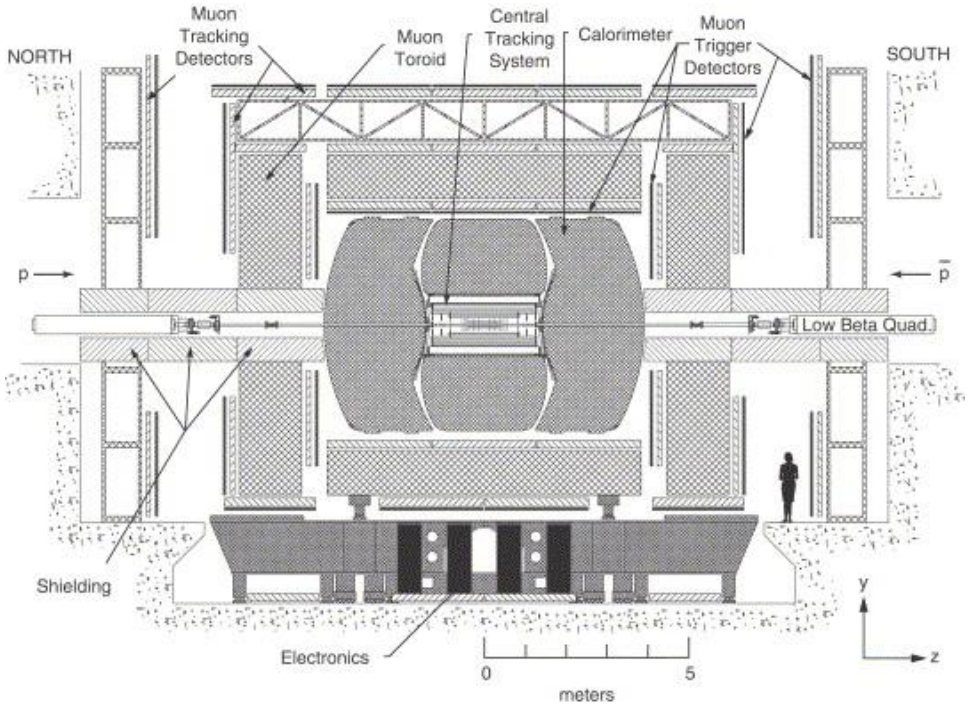

Figure 3. Schematic of the DØ detector (from [15]).

DØ uses a right-handed coordinate system, with the positive *z*-axis defined by the direction of the proton beam, the *y*-axis upward, and the *x*-axis toward the center of the ring. The detector is centered at the origin (0,0,0). The *r* coordinate of a particle is its perpendicular distance from the *z*-axis. To denote the particle's angular position we use the azimuthal angle φ and the polar angle ϑ:

$$\emptyset = \tan^{-1} y/x, \qquad \theta = \tan^{-1} r/z$$

We are also interested in measuring transverse energy and momentum, i.e. the components of a particle's energy and momentum perpendicular to the *z*-axis (beam):

$$p_T = p \sin\theta, \qquad E_T = E \sin\theta$$

In general, we replace the polar coordinate with the pseudorapidity

$$\eta = -\ln\tan\theta/2$$

This approximates the true rapidity, which is Lorentz-invariant.

IV. Event samples

As noted above, the Tevatron produces over two millions collisions per second. Out of this large pool, the DØ detector must select and record events of interest. The Standard Model predicts that a single top quark is produced for every 330 millions collisions. Only a fraction of single-top quark events is expected to arise from *W'* decay, if it even occurs at Tevatron energies. Unfortunately, many events can look like a signal event (in this case, a *W'* event) but actually are other processes. There are always far more background than signal events, so we require a powerful analysis tool to classify our data. In this analysis we use the technique of Boosted Decision Trees. First, BDTs are trained using simulated signal and background samples. We then apply the trained BDTs to data collected from the detector, in order to classify events as either signal or background. In preparation for a more detailed presentation of the BDT analysis, this chapter briefly discusses how the collected and simulated data samples are obtained.

Collected

Data for the *W'* search was recorded from August 2002 through August 2007, in two sets (Aug. 2002 – Dec. 2005 and Jan. 2006 – Aug. 2007). The first and second data sets were reconstructed using the p17 and p20 reconstruction software, respectively. For the combined data samples, the integrated luminosity[‡‡] is about 2.3 fb^{-1}. To select events, DØ looks for an electron or muon, missing transverse energy, and 2 to 3 jets—at least one of which is tagged (classified) as a *b*-quark jet. The invariant mass $\sqrt{\hat{s}}$ constructed from the leading two jets, lepton and neutrino must be greater than 400 GeV.

This analysis studies events with a muon plus two jets, one of which is *b*-tagged. Muons are identified as tracks in the muon spectrometer and central tracker. The missing transverse momentum p_T is computed by vectorial addition of transverse energies in all cells of the hadronic calorimeters; we require 15 GeV < $\not\!\!E_T$ < 200 GeV. Jets are reconstructed (using the p17/p20 reconstruction software) from energy deposits in the calorimeter. To be selected, the leading jet must have p_T > 25 GeV and $|\eta|$ < 2.5, and the second

[‡‡] Integrated luminosity essentially measures the probability of proton-antiproton collisions. The units are inverse femtobarns (femto = 10^{-15}, barn = 10^{-24} cm^2).

leading jet must have p_T > 20 GeV. Exactly one of these jets must be tagged as a *b*-jet by the neural-network tagger.

Simulated

Signal and background for s-channel events are modeled using the SingleTop generator. SingleTop simulates *W*′, the top quark (mass 175 GeV), and the *W*-boson decay into a lepton plus neutrino. For this analysis, we are concerned with the s-channel signal files for which the mass of W' is 750 GeV. The purely left-handed W'_L sample includes *W*-*W*′ interference while the purely right-handed W'_R sample does not; no other distinction is made between the different couplings. Due to the limits set in the 2008 analysis (Fig. 1), for W'_R at 750 GeV we assume M(ν_R) < M(W'_R). Since the simulated invariant-mass distribution for *W*′ bosons with "mixed" couplings is the same as for a mixture of "pure" right- and left-handed *W*′ bosons, we need only consider the latter.

Simulated background files include the following:

- Single-top quark events (s-channel or t-channel) not involving *W*′
- Production of a *W* or *Z* boson plus jets
- Final states consisting of a top-antitop quark pair
- Multiple jets with one misidentified as a lepton
- Diboson production (*WW* or *WZ*)

Background processes are estimated using Monte Carlo samples as well as data from the detector.

V. Boosted Decision Trees

As applied to particle-physics analysis, a decision tree classifies events in a data sample as signal or background by performing a sequence of binary splits. Yes/no (signal/background) decisions are made using one variable at a time to cut the sample. We continue cutting, using the most discriminating variable each time, until a specified stop criterion is reached. Ultimately, the data sample is split into regions (in the phase space of all discriminating variables) classified as either signal or background. This chapter introduces some theoretical aspects of boosted decision trees, then discusses their use in the current *W*' analysis. "Boosting" here refers to the generation of multiple decision trees which are then compiled into a single tree, in order to attain better classification.

Overview

Tree classification methods can be traced back to the early 1970s, when it emerged as an effective computational tool for social scientists [16]. One of the first applications involved ship classification based on radar range profiles [17]. The idea was to recognize a ship as belonging to one of six structural classes, using only radar data collected by an airplane circling around the ship. Each profile plotted the intensity of radar returns versus the distance of the corresponding part of the ship from the airplane. Although the shape of the profiles varied depending on the angle between the airplane and the central axis of the ship, peak positions were fairly constant. Hence it was desirable to assign to each profile a vector consisting of local-maxima locations. The inconsistent dimensionality (number of peaks) in this problem was best handled using a tree-structured approach. Thirty years later, the same basic concepts once used to classify ships are being used to classify interactions among fundamental particles.

Binary tree-structured classification begins with a complete data set, known as the *root node*. The initial binary split divides the set into two subsets, which are subsequently split into additional subsets. If a subset is not split further, it becomes a *terminal* or *leaf node* and is assigned to a class. Subsets which do not terminate are termed, unsurprisingly, *nonterminal nodes*. Thus we see that decision-tree construction consists of three steps:

1. Use the data set to select appropriate splits.

2. Decide whether to continue splitting a node or to declare it terminal.
3. Assign each terminal node to a particular class.

Obviously, we want the split at a nonterminal node to be such that the data in each daughter subset is "purer" (i.e. has more elements belonging to a single class) that the data in the parent subset. Fig. 4 gives a schematic of a decision tree in which every terminal node has been classified as either signal (S) or background (B), based on the majority of its data content.

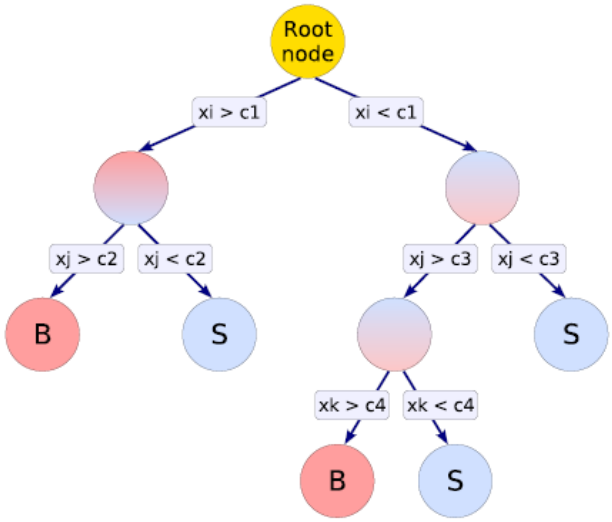

Figure 4. Schematic of a decision tree using the discriminating variables x_i (from [17]).

For this analysis we employ the ROOT-integrated Toolkit for Multivariate Data Analysis (TMVA) [18].[§§] TMVA contains a number of multivariate classification algorithms for high-energy physics analyses, including boosted decision trees. We first train a BDT using the simulated signal and background samples described in the previous chapter. Then we apply the trained BDT to a data sample from the detector to classify actual events as signal or background.

Various parameters are specified during the training phase:

- *Minimum leaf size* refers to the minimum number of events allowed in a terminal (leaf) node, i.e. a node at which further splitting is stopped. Nodes containing this minimum number of events will remain unsplit regardless of data purity.

[§§] ROOT (http://root.cern.ch) is an interactive data-analysis system for high-energy physics.

- *Boosting* is a way of stabilizing decision trees to offset fluctuations in the training sample. If we set the number of boosts to *n*, we obtain *n* individually trained decision trees by successively reweighting events. Correlations between variables are accounted for.
- The *forest* is the collection of trained decision trees, obtained through boosting, which are combined to form a single classifier. This is accomplished by taking a majority vote of the individual trees.

The usual boosting algorithm is the adaptive boost, which gives a higher event weight to events which were misclassified during the training of the preceding tree. All the weights of the previously misclassified event are multiplied by a common boost weight α, derived from the misclassification rate (denoted err):

$$\alpha = \frac{1 - \text{err}}{\text{err}}$$

After multiplication by α, the data sample is renormalized to maintain a constant sum of event weights.

The split at each nonterminal node is determined by the variable, at a particular cut value, which gives the best separation between signal and background. Hence it is necessary to define a selection criterion, in this case the Gini index $p \cdot (1-p)$. The purity p is the ratio of signal events to all events in the node. A pure background node has purity zero, while a fully mixed node has purity 0.5. Note that the separation criterion is symmetric with respect to signal and background, reaching a maximum at purity 0.5 and falling off to zero at purity 0 or 1. To find the optimal variable cut value, the training algorithm scans over 20 possible cut values—finer stepping has not been shown to improve performance.

Like any classifier, BDT analysis has its own set of advantages and limitations. Main advantages include the simplicity of the method (one-dimensional cut optimization), as well as the ability to deal with poorly discriminating variables (which can be ignored at a given node). Although in theory we could continue splitting until each leaf node is completely pure (all background or all signal), in practice that strategy results in an "overtrained" decision tree which would then need to be "pruned." We can avoid pruning by setting a larger leaf size. This reduces unwanted amplification of fluctuations in the training sample, but it also worsens leaf-node purity.

Analysis

In this analysis we consider four data sets: the p17 and p20 muon/2-jet/1-*b*-tag events, each with the left- and right-handed 750-GeV *W*′ signal files. For a given set, we first create signal and background training files in ROOT from the corresponding simulated sample. Decision trees are then trained using a list of 31 relevant input variables. Minimum leaf size is set to 100 events; 20 trees are trained (number of boosts = 20) and compiled into a single BDT (forest). The next step is to check the output variable rating table to see which variables offer the best signal/background separation. Variable ranking is determined by a weighted count of how often each variable has been used to split tree nodes. Rankings are provided for the 20 individual trees as well as for the entire forest. Using the overall top-5 variables (see next section), we train a new BDT in the same manner. Both output BDTs will be applied to the collected data for purposes of comparison.

The second stage of analysis deals with ROOT files created from the yield sample (detector data). We apply our trained BDTs to the yield files, thereby classifying events in the yield sample as signal or background. The results for each input file (signal, background and collected data) are summarized in an output histogram. We also create invariant-mass histograms for the three input files, since invariant mass was the sole discriminating variable used in the 2008 *W*′ analysis. At this point, for each of the four data sets we have three sets of histograms:

- All-variable BDT result
- 5-variable BDT result
- Invariant-mass distribution

Finally we are ready to run limit-setting codes. Cross-section limits (expected and measured) will be discussed in the next chapter, following a review of the BDT training results.

Variable rankings

Table 3 displays the five top-ranked variables from each of the trained 31-variable BDTs. To appreciate the significance of the rankings, a brief description of variables is in order.

- *Shat*: Invariant mass of jets, lepton (muon) and neutrino
- *Jet1Pt*: Transverse energy of the first leading jet
- *Jet2Pt*: Transverse energy of the second leading jet

- *HT_AllJets*: Sum of all jets' transverse energies
- *LeptonPt*: Transverse energy of the lepton (muon)
- *Cos_LeptonQZ_BestTop*: Cosine of the angle between the muon and *z*-axis, in the rest frame of the top quark reconstructed with the best jet
- *HT_AllJetsLeptonMET*: Sum of all jets' transverse energies, plus the missing transverse energy of the lepton (muon)
- *HT_AllJetsMinusBestJet*: Sum of all jets' transverse energies, minus that of the best jet

Refer to Figs. 5-6 for normalized histograms of *Shat* and *HT_AllJets* for the p17 training and yield samples.

Table 3. The top-5-ranked variables from each trained 31-variable BDT. *Percent contribution* refers to the fraction of total node-splitting for which a particular variable was used.

Training sample (Reco version / helicity of W' couplings)	Top-ranked BDT variables	Percent contribution
p17 / left-handed signal	*Shat*	31.0
	HT_AllJets	20.0
	Jet2Pt	12.7
	Jet1Pt	8.1
	LeptonPt	5.1
p20 / left-handed signal	*Shat*	33.1
	HT_AllJets	14.6
	Jet2Pt	12.8
	Jet1Pt	8.4
	LeptonPt	4.7
p17 / right-handed signal	*Shat*	64.7
	HT_AllJets	9.7
	Cos_LeptonQZ_BestTop	4.2
	Jet2Pt	2.1
	HT_AllJetsLeptonMET	2.0
p20 / right-handed signal	*Shat*	64.8
	HT_AllJets	5.2
	Cos_LeptonQZ_BestTop	4.2
	HT_AllJets_MinusBestJet	3.3
	HT_AllJetsLeptonMET	2.9

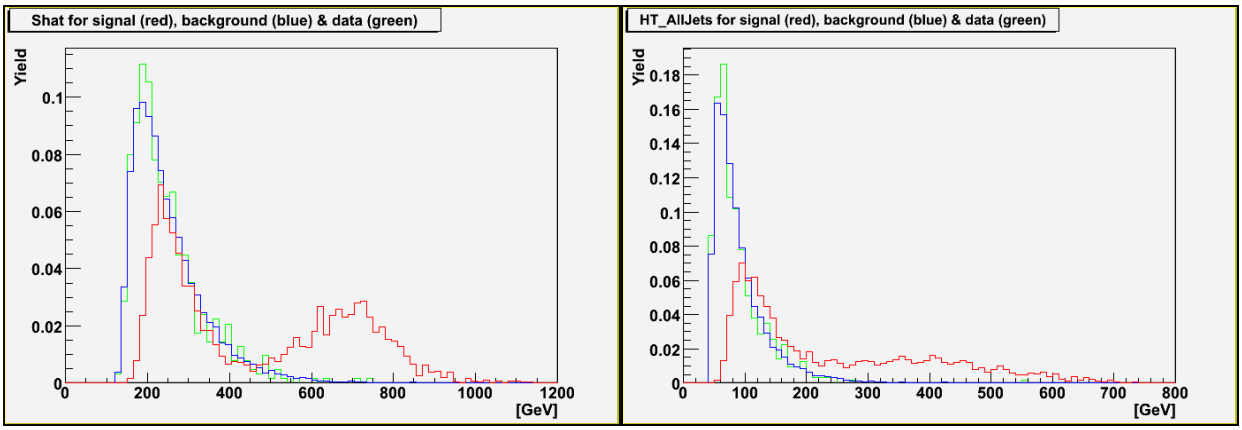

Figure 5. Normalized histograms of *Shat* and *HT_AllJets* for the p17 data/background and left-handed *W'* signal (mass 750 GeV). Note the destructive interference in the invariant-mass distribution for signal events.

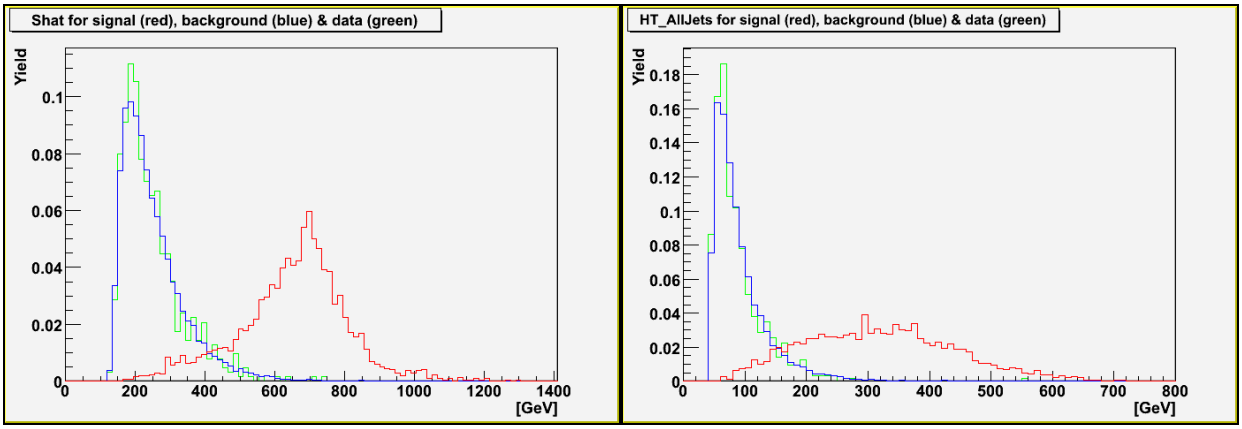

Figure 6. Normalized histograms of *Shat* and *HT_AllJets* for the p17 data/background and right-handed *W'* signal (mass 750 GeV).

VI. Cross-section limits

The final stage of analysis entails using limit-setting codes to search for evidence of W' production in the collected data. In the absence of such evidence, we set limits on the cross section $\sigma(p\bar{p} \to W') \times B(W' \to tb)$ for *W'* production and subsequent decay to a top quark and a bottom quark.

Two types of cross sections are calculated using the TopStatistics package: *expected* and *measured*. The expected cross section, in this case 1 picobarn (pb)[***], corresponds roughly to the number of signal events we expect to see based on theory. To obtain the expected number from the cross section, we multiply by the luminosity (2.3 fb^{-1}). So hypothetically we should see about 2300 *W'* events in the collected data. It turns out, however, that the measured limits are centered at 0 pb.

Although we have not discovered new physics, we are interested in learning whether multivariate BDT analysis yields more precise cross sections than those obtained using only the invariant-mass distribution to discriminate signal and background events. Each cross section corresponds to the peak position of a Bayesian posterior probability distribution generated by TopStatistics. We would like the distribution to be as narrow as possible. One way to measure this is to calculate the width in which 95% of the distribution is contained—the 95% confidence-level limit. These limits have been calculated for both the expected and measured cross sections (Tables 4-5), using the combined p17/p20 data and simulated background samples. Each expected limit is calculated from the simulated background and either the left-handed or right-handed simulated signal, while the corresponding measured limit derives from the collected data as well. Since our purpose here is strictly comparative, only statistical uncertainties are incorporated in the results. Rigorous limit-setting would require consideration of global (systematic) uncertainties.

[***] pico = 10^{-12}

Table 4. Combined p17/p20 95.4% C.L. cross-section limits for events with muon, 2 jets and 1 *b*-tag, using the *left-handed* simulated signal sample for which M(W'_L) = 750 GeV

Limit-setting input	Expected limit (pb)	Measured limit (pb)
BDT result: 31 variables	2.96592	1.37904
BDT result: top 5 variables	3.22416	1.64640
Invariant mass (*Shat*)	3.56952	2.25048

Table 5. Combined p17/p20 95.4% C.L. cross-section limits for events with muon, 2 jets and 1 *b*-tag, using the *right-handed* simulated signal sample for which M(W'_R) = 750 GeV

Limit-setting input	Expected limit (pb)	Measured limit (pb)
BDT result: 31 variables	2.08320	0.80616
BDT result: top 5 variables	2.15952	0.88560
Invariant mass (*Shat*)	2.39208	1.20456

We see that use of BDT analysis offers a significant improvement over use of the invariant-mass distribution alone, in both the expected and measured limits. The differences in the measured limits are especially dramatic, with the 31-variable BDT result providing improvements of 38.7% for W'_L and 33.1% for W'_R. Use of the top-5-variable BDT analysis is not as effective but still yields better results than use of invariant mass alone—in the measured limits we see improvements of 26.8% for W'_L and 26.5% for W'_R.

VII. Conclusion

Boosted-decision-tree analysis is a promising method which offers considerable improvements (> 30%) over single-variable analysis in computing cross-section limits for the *W*′ experiment. The more variables we include during BDT training, the better signal/background separation we see upon applying the BDTs to the data samples. This is true despite the fact that most of the variables are relatively non-discriminating compared to the final-state invariant-mass distribution (the single discriminating variable used in the 2008 *W*′ analysis [1]). Boosted decision trees are a fairly recent addition to the high-energy-physics data-classification toolkit and have enjoyed only limited application so far. It will be interesting to see what role they will play in future HEP discoveries.

[1] V. M. Abazov *et al*. (DØ Collaboration), Phys. Rev. Lett. **100**, 211803 (2008).
[2] D. Griffiths, *Introduction to elementary particles*, 2nd ed. (Wiley, Germany, 2008).
[3] M. J. Dugan, H. Georgi and D. B. Kaplan, Nucl. Phys. **B254**, 299 (1985).
[4] R. S. Chivukula, E. H. Simmons and J. Terning, Phys. Rev. D **53**, 5258 (1996).
[5] M. Schmaltz and D. Tucker-Smith, Ann. Rev. Nucl. Part. Sci. **55**, 229 (2005).
[6] U. Baur *et al.*, Phys. Rev. **D35**, 297 (1987).
[7] P. Langacker, R. W. Robinett, and J. L. Rosner, Phys. Rev. **D30**, 1470 (1984).
[8] J. C. Pati, arXiv:hep-ph/0606089v2.
[9] R. Sundrum, arXiv:hep-ph/0508134v2.
[10] E. Malkawi, T. Tait and C. P. Yuan, Phys. Lett. B **385**, 304 (1996).
[11] Y. Mimura and S. Nandi, Phys. Lett. B **538**, 406 (2002).
[12] T. Bose and M. Narain, DØ Note 5602 (2008).
[13] W.-M. Yao *et al.*, J. Phys. G **33**, 1 (2006).
[14] E. Boos, V. Bunichev, L. Dudko and M. Perfilov, Phys. Lett. B **655**, 245 (2007).
[15] V. M. Abazov *et al*. (DØ Collaboration), Nucl. Instrum. Methods Phys. Res. A **565**, 463 (2006).
[16] L. Breiman, J. H. Friedman, R. A. Olshen and C. J. Stone, *Classification and regression trees* (Chapman & Hall, Boca Raton, 1998).
[17] R. Hooper and A. Lucero, *Radar profile classification: a feasibility study*. (Technology Service Corp., Santa Monica, 1976).
[18] A. Höcker *et al*., TMVA User's Guide, http:// tmva.sf.net (2008).