Version: 0.9

Model Independent Search for New Physics at DØ in Final States Containing Leptons

The DØ Collaboration
(Dated: March 15, 2009)

This note describes the methods used to search for physics beyond the standard model in a model independent way and presents the results of the search based on 1.07 fb^{-1} of data collected with the DØ detector.

Preliminary Results for Winter 2009 Conferences

I. INTRODUCTION

Particle physics is at a stage where there is no unique way forward. The standard model of particle physics has been remarkably successful: all fundamental particles predicted by this model have been discovered, with the notable exception of the Higgs boson. Despite its success, there are strong motivations from the theory to expect new physics at energies at or just above the electroweak scale. For example, the Higgs boson receives quantum corrections to its mass through loop diagrams. The scalar nature of the Higgs boson leads to quadratic divergence, where the upper limit of the integral is set by the highest scale, i.e. the Planck mass (10^{19} GeV). To avoid correcting the Higgs mass to above the electroweak scale it is necessary to fine tune a parameter in the theory to within $M_W/M_{Planck} \sim 10^{-16}$.

To overcome this problem, there are few options. If the Higgs boson does not exist, then there must be new physics at the electroweak scale. If the Higgs boson exists but the theory is not fine tuned, new physics is also expected to be present at the electroweak scale. However, if the Higgs boson exists and the theory is fine tuned, one can avoid the need for new physics at the electroweak scale.

Assuming that new physics does exist, we do not know what this new physics is and thus precisely how to search for it. There are many theories and models which predict differences between reality and the standard model. But generally these theories and models do not give precise energy and phase space regions to search for new physics. Motivated by this, we performed a scan over many channels to look for significant deviations from the standard model. After this, we focused on events containing objects of high transverse momentum in a quasi-model independent search. Our background model is currently most developed for the case of final states containing leptons, which is the focus of the search reported in the note. Similar approaches to search for new physics have been applied to data from DØ Run I [1–3], H1 at HERA [4] and CDF Run II [5, 6].

II. DETECTOR

The DØ detector is described in detail elsewhere [7]. The central tracking, the calorimetry and muon systems are the components most important to this analysis. The central tracking system consists of a silicon microstrip tracker (SMT) and a central fiber tracker (CFT), both located within a 2T superconducting solenoidal magnet, providing charged particle tracking for pseudorapidities $|\eta| < 3$, where $\eta = -\ln[\tan(\theta/2)]$ and θ is the polar angle. The DØ cylindrical coordinate system is right-handed with the z-axis aligned along the proton direction, $\theta = 0$ along the positive z-axis, and $\phi = \pi/2$ straight up along the positive y-axis.

The three components of the liquid-argon/uranium calorimeter are housed in separate cryostats. A central section, lying outside the tracking system, covers up to $|\eta| = 1.1$. Two end calorimeters extend the coverage to $|\eta| \approx 4$.

Outside of the calorimeter, the outer muon system consists of a layer of tracking detectors and scintillation trigger counters in front of 1.8T iron toroids, followed by two similar layers after the toroids and has pseudorapidity coverage $|\eta| < 2.0$ [8].

No explicit trigger requirement was made, although most of the sample was collected with single muon and single electron triggers. The nature of our analysis is such that we do not need to apply absolute trigger efficiencies, which are accounted for in our scale factors described later in this note. The redundancy of the D0 trigger system, in conjunction with the relatively high transverse momenta required for the final state objects, leads to per event trigger efficiencies with little or no kinematic dependence, such that a global correction factor is a sufficiently good approximation.

The data for this analysis were collected during the data-taking period that ran from 2002 to 2006, referred to throughout this note as RunIIa.

III. STRATEGY

The search technique we have developed trades sensitivity for breadth of search: we do not make data selections specific to a particular model and intentionally avoid the consideration of systematic uncertainties. These conditions allow us to incorporate many channels without the large devotion of time and resources necessary to understand modeling inconsistencies specific to each individual channel. This approach limits claims that can be made about any individual discrepancy, but it helps point to any gross differences that we might see relative to the standard model expectation. If a particular final state or distribution that has been found to be discrepant withstands a full application of possible systematic uncertainties, then we may be able to make a new physics claim. The benefit of this approach is that we are able to look systematically at many channels in a coordinated fashion, applying knowledge of the SM and a consistent model of detector effects in a quick and relatively simple manner. We have also chosen to focus on high-p_T phenomena. The search for new physics starts with selecting only events with objects that have large p_T values collected by the DØ experiment. To enter our data set, high thresholds are set on objects and events

are saved in a reduced format. The net effect of the format and the high-p_T selection is to keep a factor of 10^4 less storage volume than the standard full DØ data set. The motivation behind this compact data set is that a model independent search requires frequent cycling over very large data and Monte Carlo samples and running over the full DØ data set would significantly slow down the analysis.

In order to define our final states, we first define objects such as isolated electrons, muons, taus, photons and jets, according to the DØ standard object criteria [7].. Electrons are characterized by an isolated shower in the EM calorimeter and an isolated track in the central tracker. Isolated showers without isolated tracks are identified as photons. The muons are identified in the muon chambers outside the calorimeter and then matched to tracks in the central tracker. An isolated muon is characterized by little activity in the calorimeter around the muon track with tracker isolation further enhancing the isolated muon purity. We use jets built out of calorimeter cells with significant energy deposition in the hadronic calorimeter, while taus are track-matched with a narrow hadronic calorimeter signature. Additionally, neutrinos and potentially other objects arising from undiscovered physics processes escape detection. We account for these missing objects by determining the transverse energy imbalance($\not{E}_T$) using all calorimeter cells. To account for the fact that our calorimeter is not perfectly hermetic, we further correct the $\not{E}_T$ by scaling the energies of the EM objects and jets to account for known inefficiencies. Also, because muons deposit relatively little of their energy into the calorimeter, we also correct the $\not{E}_T$ by the tracks of reconstructed muons rather than the calorimeter cell energy. Our Monte Carlo does not properly model background that arises from multijet processes, such as jets misidentified as electrons or b-jets that decay semileptonically. To account for this, we reverse some of our object selection cuts in data to produce samples with electron, muon, or tau objects which are mostly misidentified jets from multijet backgrounds.

We then divide the whole data set into seven nonoverlapping final states. In each of these, we apply previously determined standard weights using well-understood areas of phase space that are dominated by particular standard model processes to account for necessary corrections to the Monte Carlo and detector modeling. An example of this is the tracking efficiency scale factor applied to muons shown in Fig. 1. The seven states are inclusive in jets and additional objects as specified in Table I where each state is listed with the objects that define it and the associated object cuts. The additional objects (X in the table) require cuts as seen in Table II.
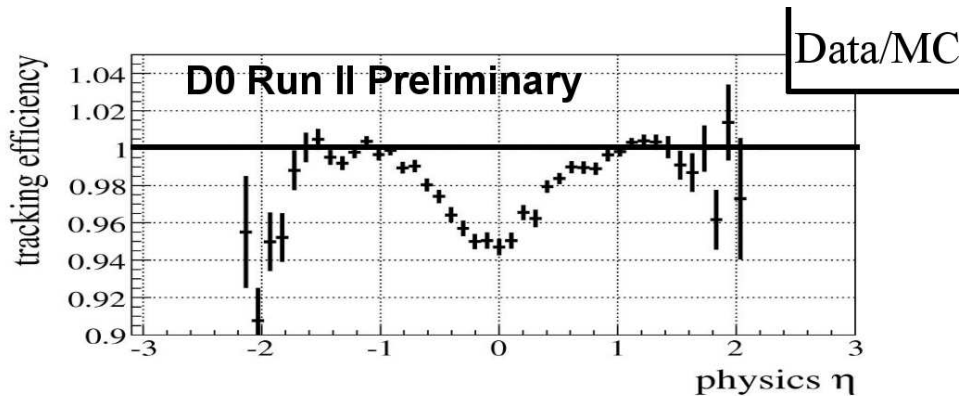

FIG. 1: The scale factor weight applied to events containing a muon to account for the differences in tracking efficiency between the DØ detector and the detector simulation.The physics η is the pseudorapidity defined with respect to the point of interaction as opposed to detector η which is defined with respect to the center of the detector.

Our simplified modeling implementation does not directly account for certain normalization factors due to such things as trigger efficiencies and some K-factors. In order to avoid gross errors in normalization, we perform a fit, described below, for each of these states to obtain the scale factors which reproduce the distributions of the selected data with the background from Monte Carlo and multijet background determined from data. These seven states were picked to fit with particular standard model processes and share overall normalization factors. Since they are nonoverlapping, they can be combined as an input to the Vista algorithm without fear of double-counting.

All of the Drell-Yan, W and $t\bar{t}$ Monte Carlo samples that we currently use are produced using ALPGEN [9] matched to PYTHIA [10] for partonic showering and hadronization. We use the MLM matching scheme to avoid double counting in areas of phase space where PYTHIA and ALPGEN may overlap [11]. PYTHIA is used exclusively for the production of our diboson(WW,WZ,ZZ) events.

Our Monte Carlo generation method using ALPGEN matched with PYTHIA is shown to have inconsistencies with data in the Z and W boson p_T spectra at low values of boson p_T. Because of this a Z p_T reweighting is performed

TABLE I: Table of Final State Object Cuts: The seven inclusive final states that are being considered with .

MIS Final State	Object	Min p_T (GeV)	Max $|\eta|$
e + jets + X[a]	e	35	1.1
	jet	20	2.5
	MET	20	NA
μ + jets + X[b]	μ^h	25	1.7
	jet	20	2.5
	MET	20	NA
ee + X[c]	e	15	1.1
$\mu\mu$ + X[d]	μ^h	15	2.0
$\mu\tau$ + X[e]	μ^h	15	2.0
	τ	15	2.5
eτ + X[f]	e	15	2.5
	τ	15	2.5
μe + X[g]	μ^h	15	2.0
	e	15	2.5

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

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

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

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

to carefully match the behavior seen in the measured Z p_T distribution from the $Z \rightarrow ee$ process [12]. This is then carried over to the W p_T distribution by utilizing the theoretical ratio of W/Z p_T spectra [13].

The fits for normalization use several histograms of basic object quantities to determine a scale factor altering the overall normalization of each input process, so that the χ^2 probability is minimized for the combined fit. In order to avoid fitting to the high-p_T tails that will eventually be searched for new physics, we check each object in the event to see if the object p_T is outside the bulk of the distribution. Basic histograms like $\not{E}_T$, p_T, η, $cos(\phi_{obj} - \phi_{\not{E}_T})$ for the leptons and jets are used to fit while we reserve more complex variables to check the fit quality. These histograms include variables like the mass or transverse mass of two or more objects, jet multiplicities, $\Delta\phi$ between two objects, inclusive jet p_T, W and Z p_T, etc. If an event contains any object in the tails, then none of the objects in the event will be used in the fit. A full list of the processes which are normalized based on these inclusive fits along with the final states that are used to determine their values are shown in Table III. A slightly simplified example using the electron + jets + X final state (X is not an e, μ or τ) would work as follows. We know this state to be dominated by the W processes with a significant amount of multijet background and Drell-Yan. We use a constant normalization factor for the Drell-Yan process determined by a separate fit to the ee + X final state (X $\neq$ μ or τ). This parameter will be held fixed in the e + jets fit along with other rare processes which have contributions that would be too small to fit. Then, the W and multijet contributions will find the best agreement to fit the given histograms and two scale factors will be used to give an overall weight to the $W \rightarrow e\nu$ and $multijet \rightarrow e + jets$ contributions. Once the fit values are found, the histograms are plotted again taking into account the values obtained from this fit. An example of a fitting histogram and a checking histogram can be seen in Figs. 2(a) and 2(b).

After determining the normalization scale factors, the seven inclusive subsets are merged to create an input file for the algorithm called Vista [5]. Each Monte Carlo and background event is given a weight calculated from the scale factors and known corrections. The Vista algorithm was developed by the CDF collaboration as a tool that

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

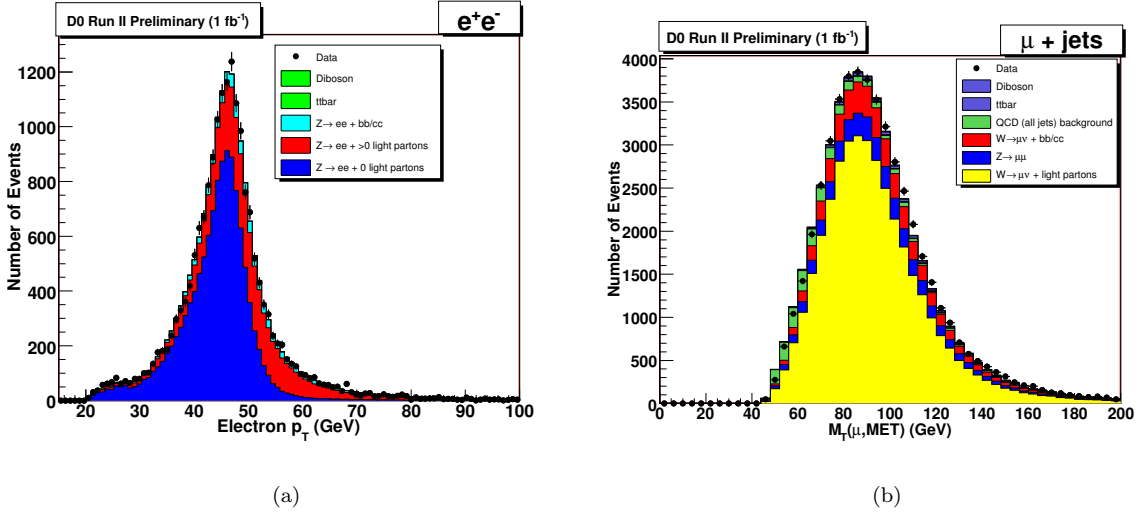| Input Process	Final States
$W \to e\nu$ + light partons	e + jets
e multijet background (e + jets)	e + jets
$W \to \mu\nu$ + light partons	μ + jets
μ multijet background (μ + jets)	μ + jets
$Z/\gamma \to ee$ + light partons	ee
$Z/\gamma \to \mu\mu$ + light partons	$\mu\mu$
Heavy flavor/light flavor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

FIG. 2: The leading electron p_T in a dielectron final state used in the MIS fit, and the transverse mass of the muon and missing energy in a muon + jets final state used to check the MIS fit.

could be used by collider experiments to perform a broad check of the agreement between data and the standard model background. We found it necessary to modify this algorithm according to the needs of the DØ detector and our analysis strategy. The VISTA@DØ algorithm is an attempt to broadly understand DØ high-p_T data. Using this framework, we try to see whether the DØ high-p_T data can be adequately accommodated by the standard model or if significant discrepancies with the Standard Model exist. Vista mainly concentrates on discrepancies that affect the bulk of distributions rather than narrow regions of phase space because it looks at the raw number of events and Monte Carlo/data agreement across full distributions. Since Vista looks at many final states and histograms, the sensitivity to any individual discrepancy is reduced. This makes it sensitive only to relatively significant discrepancies.

The original Vista algorithm as used by CDF also had the responsibility of determining many of the correction factors on its own. We chose instead to use a standard set of corrections developed by the collaboration and scale factors determined from our normalization fits described above which all come from well-understood regions of phase space at DØ.

The use of standard object identification criteria as defined above allows a great simplification: the data can be partitioned into exclusive final states. Each event is identified by the objects contained in the event. This results in

180 exclusive final states each of which contains a homogeneous subset of events. The following are examples of these exclusive final states: $\mu^{\pm}\tau^{\mp}$ + 2 jets + $\not{E}_T$, $e^{\pm}\mu^{\mp}$ + 2 jets + $\not{E}_T$, $e^{\pm}e^{\mp}$ + 3 jets and μ + 4 jets + $\not{E}_T$.

Based on event weights supplied by the correction factors, Vista performs two sorts of checks. First it does a normalization-only check on the number of events in each exclusive state; the goodness of fit can be calculated for the normalizations by Poisson probabilities. Second, it performs a shape-only analysis of histograms within a state by calculating a Kolmogorov-Smirnov statistic (and resulting fit probability) for the consistency of the shape with the predicted SM backgrounds. Both of these numbers require additional interpretation, because of the number of trials involved. When observing many final states, some disagreement is expected due to statistical fluctuations in the data. Thus the Poisson probability used in determining event count agreement is corrected to reflect this multiple testing. A similar effect occurs when comparing histograms, and again the probabilities are first converted to equivalent numbers of standard deviations and then corrected for the number of histograms examined. Two examples of 1-D histograms created by the Vista program are shown in Figs. 3(a) and 3(b).
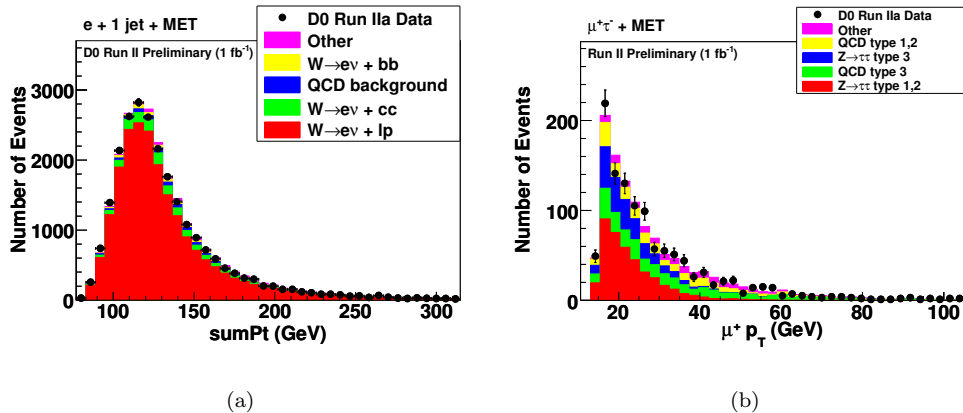

(a)

(b)

FIG. 3: Examples of Vista 1-D histograms. In 3(a), the sum of the electron p_T, jet p_T, and $\not{E}_T$ in the electron + 1 jet + $\not{E}_T$ final state. In 3(b), the muon p_T in an opposite sign $\mu\tau\not{E}_T$ final state.

The other algorithm we use to search for new physics is called Sleuth [2], developed at the DØ experiment during Run I of the Tevatron. Sleuth is an attempt to systematically search for new physics as an excess in the tails of Σp_T distributions. This variable adds the absolute values of the p_T of each object in the event to the $\not{E}_T$. The unclustered energy is not included. The Sleuth algorithm is often described as being quasi-model independent, where "quasi" refers to the assumption that the first new physics will appear as an excess of events with high-p_T objects. Thus, Sleuth would be expected to be most sensitive to high-mass objects decaying into relatively few final state particles.

In Sleuth, first Vista exclusive channels are combined by charge conjugation (so e^+X and e^-X are combined), and light-lepton universality is assumed (simply put, eX is combined with μX). By making these basic theoretical assumptions about new physics signatures, the number of states considered in Sleuth is greatly reduced from that used in Vista. This will improve the statistical sensitivity by diminishing the chances of a large fluctuation. Next, the distribution of Σp_T in each channel is searched for a lower cut which maximizes the significance of the data excess over the SM backgrounds producing Σp_T > cut. Finally, this probability needs to be corrected for the number of trials in both the number of possible positions of the cut in the histogram, and at a higher level, the number of final states Sleuth examines. This final corrected probability corresponds to the probability that an individual final state would produce one or more probabilities as small as observed. We have followed CDF by defining a significant output from Sleuth as one with corrected probability < 0.001 (that is around 3 Gaussian standard deviations).

In an attempt to limit potential bias, we did not look at all the data in the beginning. Instead we started our analysis with just 10% of the RunIIa data. This way we were able to get a broad sense of our large scale problems before focusing on individual discrepancies. The sample was chosen by randomly sampling files of the final sample. This results in a data set which samples any time or luminosity-dependent systematic uncertainties in order to closely match those of the final sample. Intrinsic to our procedure is checking inclusive states first, then systematically examining Vista plots, before proceeding to the Sleuth plots. We attempted to understand as many data/Monte Carlo discrepancies as possible using this sample before looking at the full sample. During this process we occasionally ran into cases in which there was no prospect of determining a scale factor because the background is dominated by multijet processes in such a way that we were unable to distinguish these from the other background processes by histogram shape. This made it impossible to determine the normalization factors through the fit described above. This caused us to carefully

choose how we incorporate multijet background from data and forced us to drop our consideration of $\gamma\gamma + X$ and $\gamma + jet + X$ final states. Given our method of fixing normalization factors with inclusive fits, our overall correction model is rather coarse, and we spend our effort fine-tuning it in the areas where it demands the most attention. Our current version of the model does not predict the fake rate of jets to photons in a way that is sophisticated enough to identify discrepancies in lepton + photon channels. This is a case where our naive modeling assumption is insufficient and a more complex method will have to be applied.

Once we looked at the 10% sample of RunIIa data and corrected or understood most of the discrepancies, we moved on to analyze the 100% of RunIIa data sample of 1.07 fb^{-1} of integrated luminosity.

IV. TEST OF THE METHOD

In order to test the ability of our method to discover new physics, we performed a $t\bar{t}$ sensitivity test. Re-discovering $t\bar{t}$ pairs in our data sample is a crucial test of the validity and robustness of our analysis strategy. We ran the Sleuth algorithm on our data after removing the $t\bar{t}$ Monte Carlo from the standard model background. In our analysis, we use a simplified b-tagging method that separates out jets that would have a strong probability of originating from a b-quark. For the purpose of brevity, these "b-enriched" jet objects will be identified as b-jets. Sleuth finds discoveries in $b\bar{b}\,\ell + 2$ jets $+ \not{E}_T$, $b\bar{b}\,\ell + \not{E}_T$ and $b\bar{b}\,e^{\pm}\,\mu^{\mp} + \not{E}_T$. One of these plots is shown in Fig. 4(a) with and Fig.4(b) without $t\bar{t}$ Monte Carlo included. The probability $\mathcal{P}$ of the distribution without $t\bar{t}$ arising from a statistical fluctuation would be $< 1.6 \times 10^{-7}$. The obvious discrepancies show that the Sleuth method could successfully "rediscover" top quark pairs.
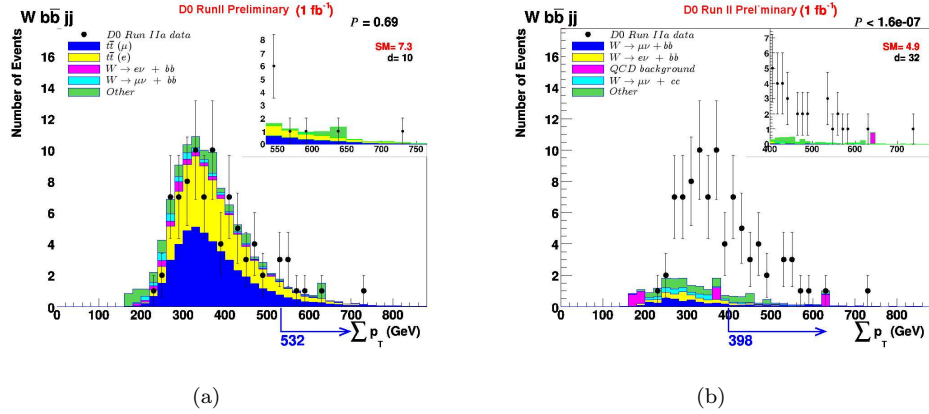

FIG. 4: Sleuth plots with and without $t\bar{t}$ Monte Carlo for $b\bar{b}\,\ell + 2$ jets $+ \not{E}_T$. The $\mathcal{P}$ value at the top right corner is the probability before final state trials factor. The probability, $\mathcal{P}$, of the distribution without $t\bar{t}$ being due to a statistical fluctuation is $< 1.6 \times 10^{-7}$.

V. RESULTS

In the 100% data sample, Vista finds a total of 180 exclusive final states. For these 180 final states, the probability is determined from $p = 1 - (1 - p_{fs})^{180}$, where p_{fs} is the probability that the number of events predicted in the standard model background would fluctuate up to or down to what is observed in data before applying the trials factor. This is then converted into units of standard deviation using $\int_{\sigma}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = p$. The final state probabilities converted into standard deviations before adding the trials factor correction are shown in Fig. 5. This distribution shows most final states near the center with some excess at the tails. Of the 180 distributions, four show significant discrepancy. These are the final states $\mu + 2$ jets $+ \not{E}_T$ with a converted probability of 9.3σ after trials correction, $\mu + \gamma + 1$ jet $+ \not{E}_T$ with 6.6σ, $\mu^+\mu^- + \not{E}_T$ with a discrepancy of 4.4σ and $\mu^+\mu^- + \gamma$ at 4.1σ.

As mentioned previously, two of these states are directly related to an oversimplified modeling of the photon misidentification rate. The $\mu + 2$ jets $+ \not{E}_T$ final state discrepancy shows an excess of events with a muon at $\eta > 1.0$ as seen in Fig. 6(a). The excess points to an oversimplification in our approach to trigger efficiencies. The proportion

of events that are brought in by single muon vs. muon plus jets triggers changes significantly as we increase jet multiplicity. These triggers introduce η-dependent efficiencies which are not properly incorporated into our simple fits. The dimuon with missing energy final state shows an excess of data compared to the standard model Monte Carlo prediction. A study into the track curvature of data and MC muons, and of the associated resolution, has shown that an additional smearing should be applied in the Monte Carlo to appropriately simulate very high p_T muons. The prime signature of these muons is an excess of $\not{E}_T$ because of the lack of compensation for the mismeasured, unbalanced track. The $\Delta\phi$ distribution of the muon and $\not{E}_T$ can be seen in Fig. 6(b).

FIG. 5: Vista final state σ distribution for 100% sample before accounting for the trials factors. The curve represents a Gaussian distribution centered at zero.
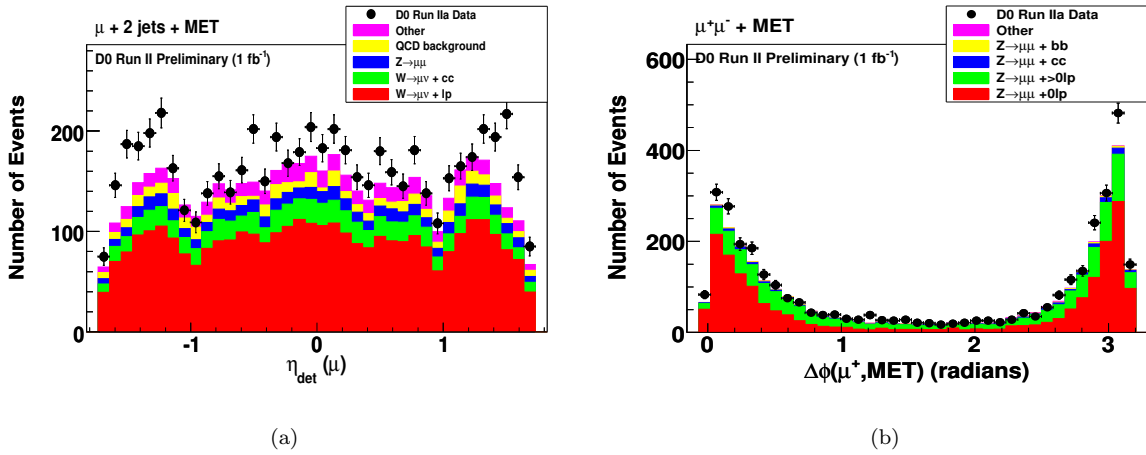

(a)

(b)

FIG. 6: Two distributions from discrepant final states. Figure 6(a) shows the excess of data in $\mu + 2$ jets $+ \not{E}_T$ to be focused on events with muons that have η values > 1.0. Figure 6(b) shows the $\Delta\phi$ distribution between a muon and the $\not{E}_T$ where the excess tends to be with events where the missing transverse energy is pointing opposite to a muon.

The 180 final states include a total of 9335 individual 1-D histograms, and a shape comparison is performed for each. The trials factor adjusted probability is determined with $p = 1 - (1 - p_{shp})^{9335}$, where p_{shp} is the KS probability to observe an individual shape discrepancy before applying the trials factor. As with the probability for a final state discrepancy, the probability for a shape discrepancy is converted into units of standard deviation and the discrepancies are shown. For the histogram shapes, any deviation $>3\sigma$ is considered discrepant. The distribution
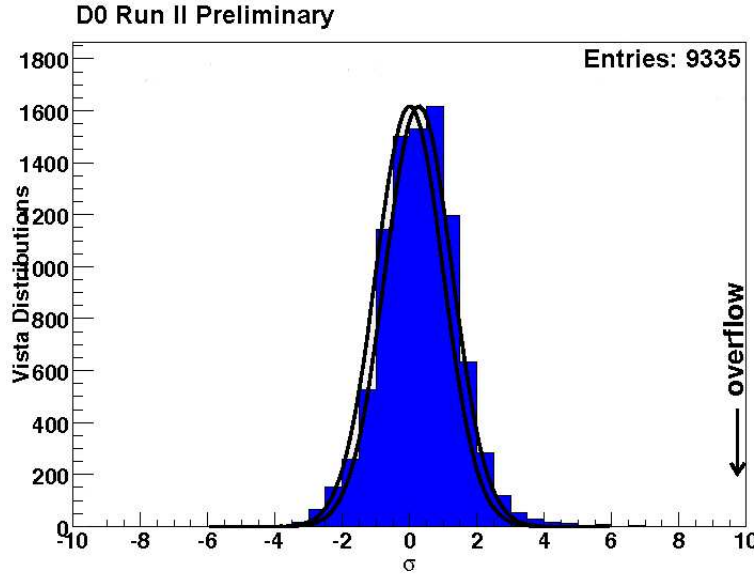
FIG. 7: Vista histogram σ distribution for 100% sample before accounting for the trials factor. Each curve is a Gaussian distribution. The curve that is shifted to lower values is centered at zero while the second curve is centered at the mean.

of standard deviations before trials correction is shown in Fig. 7. This distribution approximates a slightly shifted Gaussian of the expected width, but several distributions appear in the tails.

A total of 23 distributions are found to be discrepant at the 3σ level. The majority of these are related to spatial distributions involving jets, low $\not{E}_T$ excesses in dilepton distributions and multijet background dominated τ distributions. All of these types of discrepancies are related to known oversimplifications in our modeling assumptions and would not be expected to severely affect the Sleuth search for new physics in the high-p_T tails. Two shape discrepancies from states that agree in number of events are shown in Figs. 8(a) and 8(b).

All Vista final states are input to Sleuth and the 180 final states are folded into 44 final states after applying global charge conjugation, rebinning in number of jets and using light lepton universality as described above. The several Vista final states that show broad numerical excesses are found again with the Sleuth algorithm as would be expected.
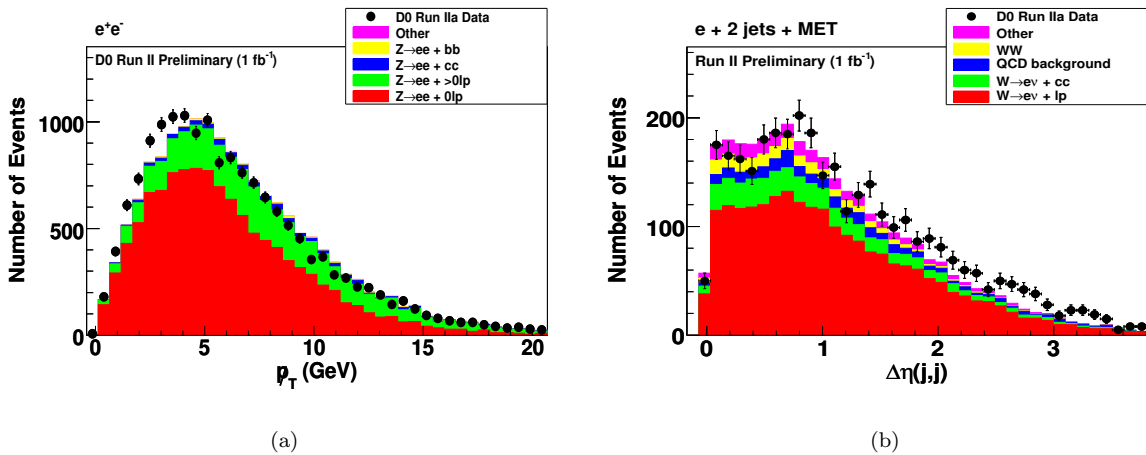

FIG. 8: Two discrepant distributions in final states that agree in number of events. The plot 8(a) shows the $\not{E}_T$ distribution in the opposite sign dielectron final state, and 8(b) shows the $\Delta\eta$ distribution between the two jets in the e + 2 jets + $\not{E}_T$ final state.

One additional distribution crosses the discovery threshold of $\tilde{\mathcal{P}} < 0.001$, where $\tilde{\mathcal{P}}$ is the probability after all trials factors, described in detail in [5] and briefly in Appendix A. The final state that crosses the discovery threshold is $\mu^{\pm} + e^{\mp} + \not{E}_T$ as can be seen in Fig. 9. Currently the evidence suggests that the muon tracking resolution is responsible for this discrepancy from the standard model, as well. A large fraction of the events in the tail of the Sleuth distribution have a muon with a very large p_T and large missing energy. With the present modeling of muon resolution, straight track events are underrepresented in the standard model background estimation. This state has 46 data events in the tail of this distribution compared to only 17 predicted by the Monte Carlo. A table of the top five Sleuth final states that contain only leptons and jets is shown in Table IV. The known Vista numerical excesses have been removed since this information is already known. All of these states are subject to the muon resolution issues discussed above.
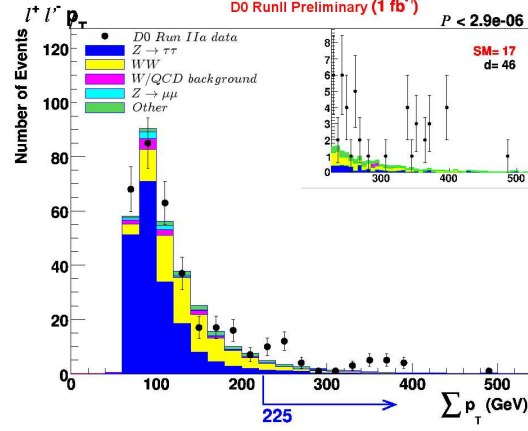

FIG. 9: Sleuth plot for OS $\ell\ell' + \not{E}_T$. The $\mathcal{P}$ value at the top right corner of the plot is the probability before final state trials factor.

TABLE IV: Top Five Sleuth States With Only Leptons and Jets. The value $\mathcal{P}$ represents the probability that the standard model background for an individual final state would have a fluctuation at any cut that would be more significant than what is seen in data. The variable $\tilde{\mathcal{P}}$ calculates the probability that one would observe a final state with $\mathcal{P}$ less than or equal to the one observed in data based on a statistical fluctuation.

Final State	$\mathcal{P}$	$\tilde{\mathcal{P}}^a$
$\ell^+\ell'^- + \text{MET}$	2.9 E-6	0.00018
$\ell + \text{MET}$	.00082	0.049
$\ell^+\ell'^-$	.0031	0.17
$\ell^+\tau^- + \text{MET}$	0.006	0.31
$\ell^+\tau^+$	0.0066	0.33

[a]The value of $\tilde{\mathcal{P}}$ is not necessarily accurate below 0.001. The important check is whether the value drops below the threshold. Further discussion can be found in Appendix A and [5].

In the Sleuth runs performed at CDF using a slightly different analysis strategy, the four most interesting observed final states were $\mu^{\pm} + e^{\pm}$, $\mu^{\pm} + e^{\pm} + 2$ jets $+ \not{E}_T$, $\mu^{\pm} + e^{\pm} + \not{E}_T$ and $\ell^{\pm} + \ell^{\mp} + \ell' + \not{E}_T$ with 2.0 fb^{-1} [6]. These states were also observed among the most discrepant for CDF at 927 pb^{-1} [5]. At DØ with 1.07 fb^{-1}, the $\mathcal{P}$ value is rather low in Figs. 10(a) and 10(c), however, none of these states are among the most discrepant. All four corresponding DØ plots can be seen in Figs. 10(a), 10(b), 10(c) and 10(d).

VI.   CONCLUSION

In conclusion, we have performed a broad search for new physics using our full RunIIa data set (1.07 fb^{-1}). A total of 180 exclusive data final states and 9,335 relevant kinematic distributions were compared to the complete standard model background predictions using Vista. Only four out of 180 exclusive final states show a statistically significant discrepancy. Given the known modeling difficulties in all four final states, we refrain from attributing the
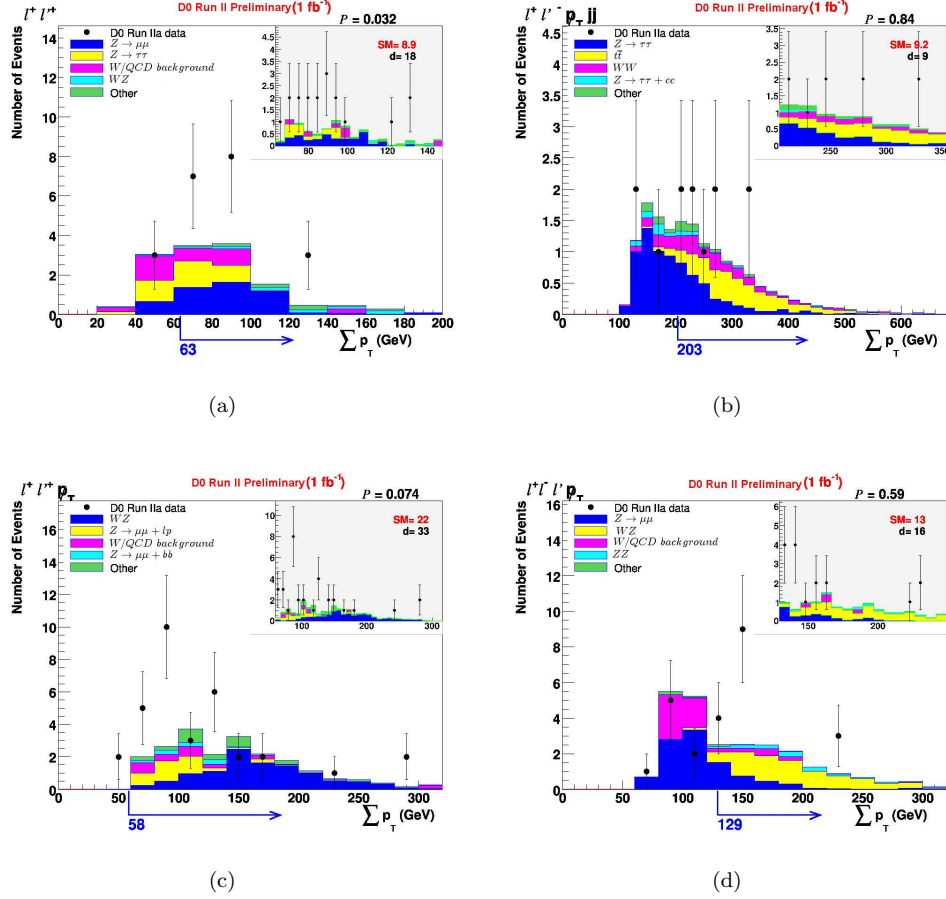
FIG. 10: Check of most discrepant CDF plots. Fig. 10(a) SS $\ell\ell'$. Fig. 10(b) SS $\ell\ell'$ + jj + $\not{E}_T$. Fig. 10(c) SS $\ell\ell'$ + $\not{E}_T$. Fig. 10(d) $\ell^\pm\ell^\mp\ell'$ + $\not{E}_T$. The $\mathcal{P}$ values at the top right corner of the plots are the probabilities before final state trials factors.

observed discrepancies to new physics. A quasi-model independent search for new physics was also performed using the algorithm Sleuth by looking at the regions of excess on the high-$\sum p_T$ tails of exclusive final states. Only $\mu^\pm + e^\mp$ + $\not{E}_T$ surpasses the discovery threshold beyond the obvious excesses noticed in Vista, and this seems to be related to difficulties in modeling the muon p_T resolution. Although we did not find any hint of new physics in the DØ RunIIa data, a factor of 5 more data has already been recorded by the DØexperiment. As we incorporate this data set into our analysis and continue implementing improvements to our correction model, we will become much more sensitive to possible new physics.

APPENDIX A: CALCULATION OF $\tilde{\mathcal{P}}$

The probability that a discrepancy seen in a given Sleuth final state is due to a statistical fluctuation in the standard model background has been defined as $\mathcal{P}$. Once the minimum value of this probability $\mathcal{P}_{min}$ over all final states is found, an additional trials factor must be determined to account for the number of states that are checked. The value $\tilde{\mathcal{P}}$ represents the probability of seeing a final state as unlikely as the value of $\mathcal{P}_{min}$ based purely on the standard model background. This is determined by the formula

$$\tilde{\mathcal{P}} = 1 - \Pi_a(1 - \hat{p}_a), \tag{A1}$$

where a represents all Sleuth final states. The variable $\hat{p}_a$ is defined as the minimum of $\mathcal{P}_{min}$ and the probability of the total number of predicted events in a final state a to fluctuate up to three data events. Three events is found to be

the minimum necessary to reasonably determine a value of $\tilde{\mathcal{P}}$ on the order of 0.001. A discussion of the determination of the minimum number of events can be found in [5].

ACKNOWLEDGMENTS

[1] B. Abbott *et al.*, Phys. Rev. Lett. **86**, 3712 (2001).
[2] B. Abbott *et al.*, Phys. Rev. D **62**, 092004 (2000).
[3] V.M. Abazov *et al.*, Phys. Rev. D **64**, 012004 (2001).
[4] A. Aktas *et al.*, Phys. Lett. B **602**, 14 (2004).
[5] T. Aaltonen *et al.*, Phys. Rev. D **78**, 012002 (2008).
[6] T. Aaltonen *et al.*, Phys. Rev. D **79**, 011101 (2009).
[7] V.M Abazov *et al.*, Nucl. Instrum. Methods Phys. Res. A **565**, 463 (2006).
[8] V.M Abazov *et al.*, Nucl. Instrum. Methods Phys. Res. A **552** 372 (2005).
[9] M. Mangano, *et al.*, Alpgen, J. High Energy Phys. **07** 001 (2003).
[10] T. Sjostrand *et al.*, Pythia 6.2, arxiv:hep-ph/0108264(2001).
[11] S. Hoech *et al*, Matching Parton Showers and Matrix Elements, arxiv:hep-ph/0602031(2008)
[12] V.M. Abazov *et al.*, Phys. Rev. Lett. **100**, 102002 (2008).
[13] K. Melnikov and F. Petriello, Phys. Rev. D **74**, 114017 (2006).