DØNote XXXX - v1.33 - CONF

Evidence for $Z \rightarrow b\bar{b}$ Decays at DØ

Amber Jenkins, Per Jonsson, Gavin Davies, and Andrew Haas

(Dated: March 9, 2006)

A search for $Z \rightarrow b\bar{b}$ production and decay has been performed at DØ using the initial 300 pb^{-1} of Run II data. Candidate events are selected by triggering on muons coming from the semileptonic decay of the b-jets. Two slightly different methods of estimating the background from data have been explored. In both approaches, a significant excess of events is observed after background subtraction. The first yields an excess of 1353 $\pm$ 151 $_{\text{(stat.)}}$ $\pm$ 306 $_{\text{(syst.)}}$ events, which corresponds to a $Z \rightarrow b\bar{b}$ signal of 4.0σ and a cross-section for $p\overline{p}$→Z of 1.XX $\pm$ XX $_{\text{(stat.)}}$ $\pm$ XX $_{\text{(syst.)}}$ nb. The second method reveals an excess of 810 $\pm$ 230 $_{\text{(stat.)}}$ events, using different event criteria. Both results are in good agreement with Monte Carlo predictions of the Standard Model.

Preliminary Results for Winter 2006 Conferences

I. MOTIVATION

The measurement of $Z \to b\bar{b}$ is an important part of the Run II physics programme at DØ. $Z \to b\bar{b}$ is an essential tool in the calibration of b-jets, which affects much of the high p_T physics studied at the Tevatron. The current uncertainty on the b-jet energy scale - of the order of 3% - dominates the uncertainty on the mass of the top quark. In addition, $Z \to b\bar{b}$ can be used to better understand the b-jet energy resolution. $Z \to b\bar{b}$ also serves as an important benchmark signal for Higgs physics, since it is the closest observable process to the dominant decay of a light Higgs to $b\bar{b}$.

The upgraded tracking system, improved muon triggering, and the possibility of online b-jet tagging make the accumulation of a significant $Z \to b\bar{b}$ signal possible during Run II. The main challenge is to sufficiently reduce the QCD backgrounds such that $b\bar{b}$ events from the Z can be observed. To this end, careful analysis techniques are critical and effective triggers must be employed. This paper describes the first search for $Z \to b\bar{b}$ at DØ in Run II of the Tevatron collider.

II. EVENT SAMPLES

The data used in the analysis come from the "BID" skim, processed and fixed with the p14 versions of the reconstruction software. (Background subtraction method 1 uses PASS2 corrections, whereas method 2 is based on the older PASS1 corrections.) The skim has the following properties:

- A "loose" offline reconstructed muon in each event, with $p_T > 4$ GeV/c, matched within ΔR <0.7 to a Run II cone jet [4] of radius 0.5. This requirement enhances the fraction of heavy-flavor events due to the decays of $b \to \mu$ and $b \to c \to \mu$; [5]
- The skim contains about 90 million events and corresponds to an integrated luminosity of about 300 pb^{-1} for pre-v13 triggers used.

The following run selection is imposed to ensure data quality:

- Events are required to be in good luminosity blocks;
- Good calorimeter and jet/Missing E_T runs are selected;
- Bad muon runs are removed;
- Runs with bad central tracking system information (SMT and CFT) are excluded.

In addition, a range of Monte Carlo (MC) samples are used, as listed in Table I. Each MC sample is generated with Pythia, overlaid with an average of 0.8 minimum bias events, passed through a full GEANT simulation of the detector, and processed with version p14 of the reconstruction software.

The data skim and MC samples are processed using custom ROOT ntuple-generating packages. Jet energy scale (JES) corrections are applied to all jets, including corrections for muons in jets, using version 5.3 (5.0) of the jet energy correction software for method 1 (2). No specific b-jet energy scale corrections are applied, as no officially approved such correction exists for this data set.

Monte Carlo events are corrected to account for the b-tagging efficiency and jet measurement effects observed in data. Jet energies are smeared to accurately reflect the jet energy resolution measured in data, and event weighting is applied in order to account for the fact that jet reconstruction/identification and b-tagging are less efficient in data than MC.

A. Triggers

$Z \to b\bar{b}$ events are characterized by the presence of two b-tagged jets. Light quark rejection is needed at the trigger level, prior to any offline event selection, in order to achieve an acceptable trigger rate at instantaneous luminosities in excess of 50 x 10^{30}cm^{-2}s^{-1}. Without this rejection the data would have to be undesirably prescaled, and candidate events would be thrown away.

Ideally one would trigger on di-jet events with displaced vertices at Level 2; such a capability is however only since recently provided by the Level 2 Silicon Track Trigger (STT). However, a large reduction in rate must already be achieved at Level 1. Candidate events can be selected at the trigger level using the semileptonic decay of the b-jets.

Monte Carlo Sample	p_T Cut Applied at Parton Level, GeV/c	Number of Events
Signal		
$Z \rightarrow b\bar{b}$	None	100–200,000
Heavy-flavor QCD Backgrounds		
$b\bar{b}$	5–20	200,000
$b\bar{b}$	20–40	125,000
$b\bar{b}$	40–80	150,000
$b\bar{b}$	80–160	100,000
$b\bar{b}$	160–320	25,000
$b\bar{b}$	> 40	200,000
Light-quark (u,d,s) QCD Backgrounds		
$q\bar{q}$	5–20	200,000
$q\bar{q}$	20–40	125,000
$q\bar{q}$	40–80	150,000
$q\bar{q}$	80–160	100,000
$q\bar{q}$	160–320	25,000

TABLE I: Summary of Monte Carlo samples used in this analysis. q refers to the light quarks u, d or s.

v12 Trigger	Trigger Description		
	Level 1	Level 2	Level 3
MU_JT25_L2M0	single μ + jet	$\geq 1\mu + \geq 1$ jet	≥ 1 jet
MU_2TRK3_L2M0	single μ + jet	$\geq 1\mu$	2 global tracks
MUW_W_L2M3_TRK10	single μ	$\geq 1\mu$	1 global track

TABLE II: Description of the three leading v12 muon triggers. Candidate signal events were required to pass MU_JT25_L2M0.

Triggering on muons from one or both of the b-jets from the Z limits the signal efficiency however, because of the small $b \rightarrow \mu$ branching fraction[6].

Trigger efficiencies for the v12 and v11 triggers are listed in Tables IV and V. The most frequently fired v12 trigger, MU_JT25_L2M0, was exposed to a luminosity of 297.5 pb^{-1}. To simplify the analysis and the comparison to the MC predictions all events in the pre-v13 data-set were selected from this single trigger. The second method presented does not apply this trigger selection, and is thus not able to measure a Z cross-section.

III. EVENT SELECTION

There are few kinematic handles with which to discriminate between the $Z \rightarrow b\bar{b}$ signal and the QCD $b\bar{b}$ background. The two most powerful variables, besides the invariant di-jet mass, are the number of jets in the event, n_{jet} and the angular separation of the two leading b-jets in the plane perpendicular to the beam, $\Delta\varphi$. Both of these variables are sensitive to the prediction that QCD background, due to the color connection between initial and final state in the QCD processes, should have more gluon radiation. With this in mind, candidate events are selected offline using the following prescriptions:

Cut 1 Trigger selection: MU_JT25_L2M0;

Cut 2 The event must contain two and only two good quality jets and at least one loose offline muon. (No offline muon is explicitly required in method 2);

Cut 3 The two jets must both have $|\eta| < 2.5$ and corrected $p_T > 15$ GeV/c. (Method 2 requires applies the p_T cut before jet energy scale corrections.);

Cut 4 For b-tagging, the two jets must both be taggable;

v11 Trigger	Trigger Description		
	Level 1	Level 2	Level 3
MU_JT20_L2M0	single μ + jet	$\geq 1\mu + \geq 1$ jet	≥ 1 jet
MU_JT25_L2M0	single μ + jet	$\geq 1\mu + \geq 1$ jet	≥ 1 jet
MUW_W_L2M3_TRK10	single μ	$\geq 1\mu$	1 global track

TABLE III: Description of the three leading v11 muon triggers. Candidate signal events were required to pass MU_JT25_L2M0.

v12 Trigger	L1L2L3 Trigger Efficiency(%)	
	Absolute	w.r.t. Offline Cuts
MU_JT25_L2M0	5.5	15.6
MU_2TRK3_L2M0	7.0	18.1
MUW_W_L2M3_TRK10	3.4	8.9
OR of the above 3 triggers	8.7	20.2

TABLE IV: v12 trigger efficiencies for MC signal events.

Cut 5 The primary vertex of the event must have more than 2 tracks attached and be located within ± 50 cm in the z-direction (± 35 cm for method 2);

Cut 6 The two jets must both be "loose" SVT-tagged (using the standard p14 definitions);

Cut 7 The two b-jets must be nearly back-to-back in the plane perpendicular to the beam-line, i.e. $\Delta\varphi > 2.75$ radians (2.9 radians in method 2).

Tables VI shows the number of events passing each analysis cut in signal MC and data for method 1. Monte Carlo event counts are weighted by cross-section and luminosity.

The b- / light-jet fraction is about 10% after a single b-tag requirement. With the additional muon and trigger selection (Cuts 1 and 2) the fraction is about 20% and after requiring a double offline b-tag in the data (Cut 6) the light-quark QCD background component is reduced to about 10% of the sample. However, a large heavy-flavor component remains which still swamps the signal; S:(S+B) is of the order of 1:30. By restricting the number of jets and the angle between the two b-jets (i.e. cuts 2 and 7) the signal significance can be improved. We have investigated the effect of these two cuts on the effective signal significance, where effective significance is defined as:

$$\textit{Effective Significance} = \frac{N_{signal}}{\sigma_{bkg}^{stat} \oplus \sigma_{bkg}^{syst} \oplus \sqrt{N_{tot}}} \tag{1}$$

Here N_{signal} is the number of MC signal events passing the cuts, weighted by cross-section and luminosity, σ_{bkg}^{stat} and σ_{bkg}^{syst} are the statistical and systematic errors on the estimated number of background events passing all cuts and N_{tot} is the total number of events in data passing the cuts. The square root of N_{tot} - N_{signal} is used as a conservative estimate of the size of the statistical error on the background.

IV. BACKGROUND SUBTRACTION

Understanding the shape of the background to the $Z \to b\bar{b}$ signal in the double b-tagged data sample is central to this analysis. The background is composed of heavy-flavor di-jet production and mis-tagged gluon/light-quark jet production, all of which cannot be accurately simulated using current techniques, particularly in the quantities needed for the statistical accuracy required. Thus the background is derived from data, using either single-tagged and/or un-tagged events.

We have developed two slightly different methods for measuring the background. The two approaches are discussed in sections V and VI respectively. Both methods use a *tag rate function (TRF)* to estimate the amount of double b-tagged background in the data. We measure the *tag ratio* of double to single b-tags as a function of jet p_T and jet η to form the TRF. The TRFs are applied, as an event weight, to the single tagged events.

v11 Trigger	L1L2L3 Trigger Efficiency(%)	
	Absolute	w.r.t. Offline Cuts
MU_JT20_L2M0	7.4	18.3
MU_JT25_L2M0	5.5	15.6
MUW_W_L2M3_TRK10	3.4	8.9
OR of the above 3 triggers	8.4	19.7

TABLE V: v11 trigger efficiencies for MC signal events.

Analysis Cut	MC $Z \to b\bar{b}$	Data Passing Cut
No selection	336475 ± 1174	∼ 90000000
1	18567 ± 276	10882864
2	10524 ± 209	5135730
3	8055 ± 183	4030914
4	6915 ± 169	2726889
5	6643 ± 165	2698544
6	1451 ± 77	33499
7	1444 ± 65	31183

TABLE VI: Cut flow table for $Z \to b\bar{b}$ events in Monte Carlo and events in data. Event counts are weighted by cross-section and luminosity, the errors are statistical only.

In the first method, the final invariant mass background template is normalized to the double tagged data with a fit to the signal free invariant mass region above 120 GeV. The second approach instead makes use of the difference in shape between single b-tagged and un-tagged invariant mass distributions.

V. BACKGROUND ESTIMATION METHOD 1

A TRF can be constructed on a jet-by-jet basis, a similar approach to that adopted in the DØ Run II $hb\bar{b}$ analysis [2]. Consider those events in which the first-leading-p_T jet is b-tagged. For these events, the second-leading-p_T jet is categorized according to its location in one of three different eta regions of the detector:
$|\eta| < 1.1$, $1.1 < |\eta| < 1.5$ and $1.5 < |\eta| < 2.5$. For each η region, the TRF is then parameterized as a function of the p_T of the second-leading-p_T jet. This generates a TRF *per jet*.

To estimate the background, each event passing cuts with the leading-p_T jet b-tagged is weighted by the TRF corresponding to the second-leading-p_T jet. The resulting background distribution is then corrected for the dependence of the TRF upon $\Delta\varphi$ by fitting it to the observed invariant mass distribution above 120 GeV. Figure 1 illustrates the jet-based TRF as a function of the second leading jet p_T, for jets in the three different η-bins.

Figure 2 shows the fit of the background prediction to the double tagged data in the high mass region. By construction this method produces a background invariant mass template with an identical number of integrated events as the observed double-tags. If the invariant mass shape of the background and observed data were identical then the normalization would be equal to 1. The fact that the fit yields 0.95 indicates a difference in the relative fraction of events above and below 120 GeV in the background template and the double-tagged data. (We have more events in the Z mass region in the double-tagged data.)

The normalized invariant mass background compared to data and signal MC is shown in Figure 3. Figure 4 shows the invariant mass after subtraction of the background, compared to Monte Carlo.

A complete mass-bin-by-mass-bin breakdown of the excess in data and predicted from Monte Carlo is presented in Table VII. After background subtraction we observe an excess of 1352.8 ± 151.33 (stat.) ± 305.70 (syst.) events in the 50–100 GeV/c^2 mass window. The excess is fit with a Gaussian and compared to the number of events expected from Monte Carlo. We expect to see 1388.92 ± 107.09 (stat.) events and therefore observe good agreement between data and Monte Carlo within the errors. Table VIII summarizes the overall excess.

The total statistical error is calculated by combining in quadrature the statistical errors on the number of observed

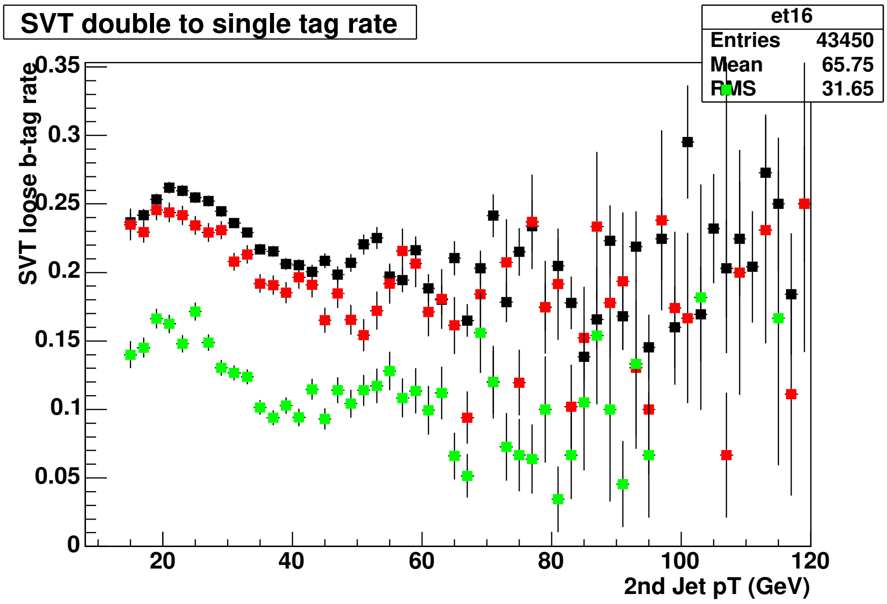

FIG. 1: TRF as a function of second-leading jet p_T for jets in the regions $|\eta| < 1.1$ (black), $1.1 < |\eta| < 1.5$ (red) and $1.5 < |\eta| < 2.5$ (green), evaluated for events in data with the leading-p_T jet b-tagged.

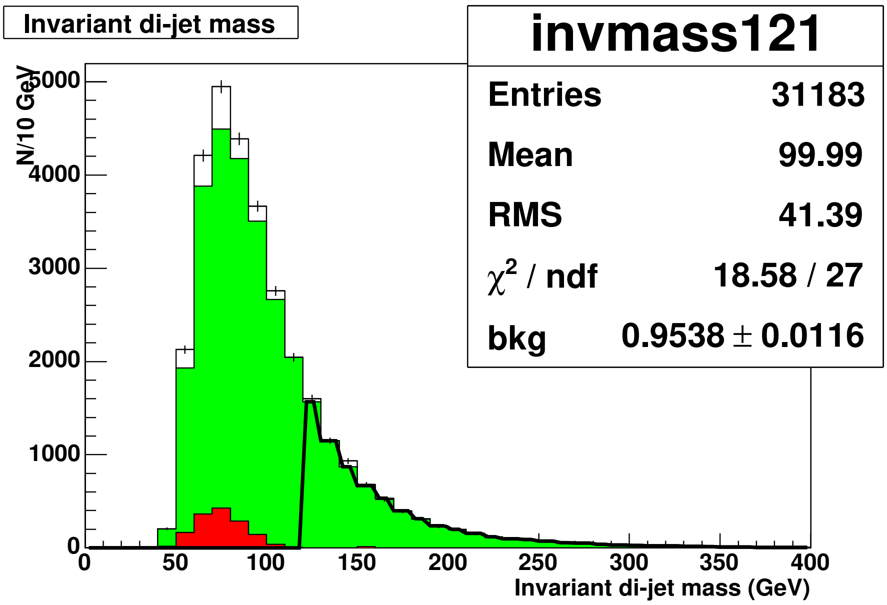

FIG. 2: Invariant di-jet mass in the pre-v13 data-set *before* background subtraction. Green (light gray) shaded histogram: expected background, estimated using a jet-based TRF from data and fit with one parameter to the observed data above 120 GeV. Black points: total di-jet mass distribution observed in data. Red (dark gray) histogram: expected $Z \rightarrow b\bar{b}$ signal from MC. The fit results in a correction to the background scale of 0.95 ± 0.01.

events and expected background events. The main source of systematic uncertainty on the observed excess comes from the normalization fit of the background and is taken, from the error on the fitted parameter, to be 1.3% of the number of background events. The contribution from the invariant mass shape of the TRF based background not being modeled perfectly, is estimated to be small from the small χ^2/NDF of the normalization fit at masses above 120 GeV. A study of the difference of the invariant mass shape of an un-tagged (light quark dominated) and single-tagged (b contaminated) event sample reveals a net uncertainty on the number of background events of less than 1% in the Z

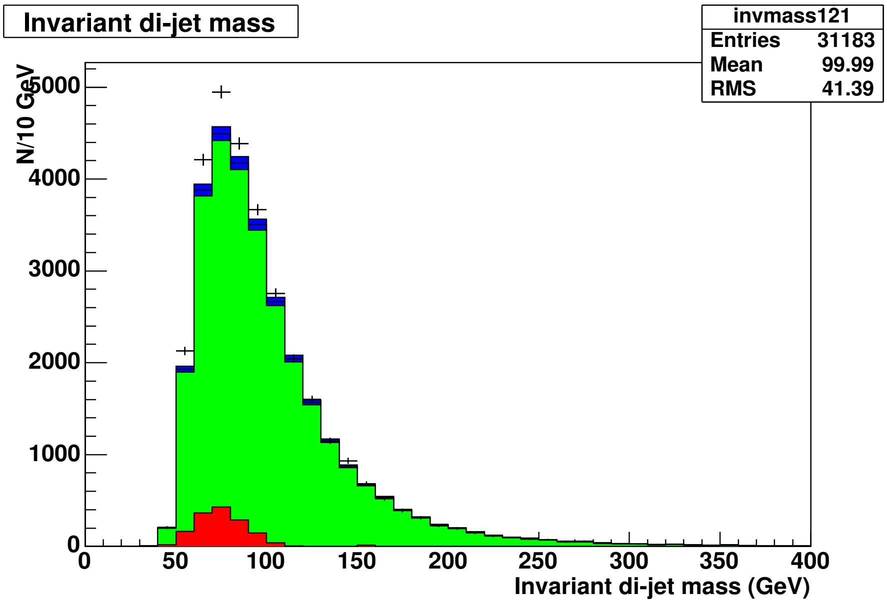

FIG. 3: Invariant di-jet mass in the data *before* background subtraction. Green (light gray) shaded histogram: expected background, estimated using a jet-based TRF from data. The dark blue band indicates the ± 1.7 % systematic error on the background. Black points: total di-jet mass distribution observed in data. Red (dark grey) histogram: expected $Z \to b\bar{b}$ signal from MC.

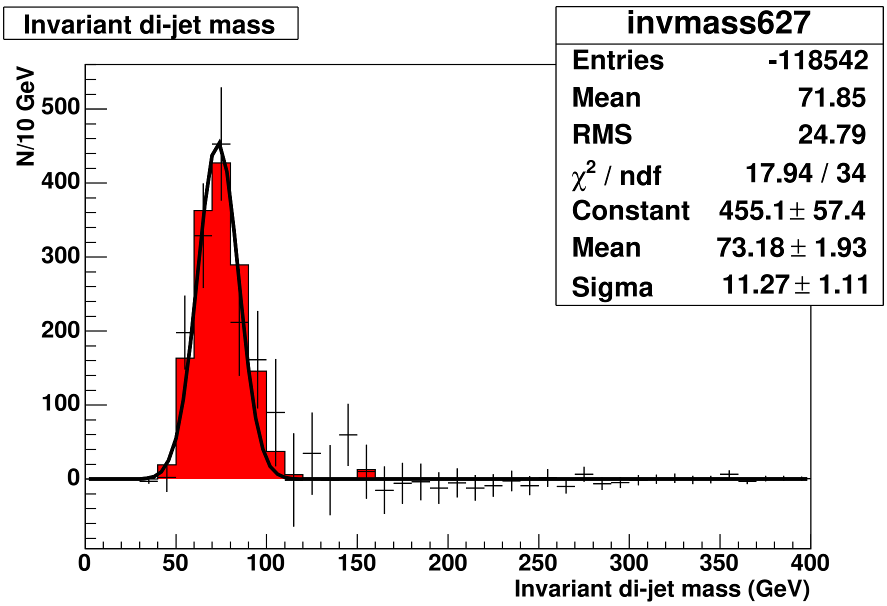

FIG. 4: The invariant di-jet mass spectrum in the data *after* background subtraction. Black points: excess observed in data, fitted with a Gaussian. Error bars show statistical errors only. Red (dark gray) histogram: expected signal from MC. The excess peak has a mean mass of 73.18 GeV/c^2 and a width of 11.27 GeV/c^2.

mass region (50-100 GeV). Hence, an additional 1% systematic uncertainty was added to the error on the background.

The systematic error on the background is added in quadrature to the total statistical error in the last column of Table VIII. Taking the systematic uncertainty of the background into account, the observed excess has a significance of 4.0 standard deviations.

The peak of the distribution lies at approximately 75 GeV/c^2, in both Monte Carlo and data, after jet energy scale corrections have been applied. At the time of this analysis no official b-jet JES correction, other than the correction for muons in jets, is available for this data-set. Applying an unofficial b-jet JES correction to the MC shifts the peak to 79 GeV. The remaining mass shift could be explained by energy loss to hard final state radiation as suggested in [3].

Invariant Mass Bin (GeV/c^{2})	**No. Events Observed**	**No. Events Expected**	**Excess**	**Expected $Z \rightarrow b\bar{b}$**
0–30	0	0	0	0
30–60	2336	2138.45 ± 20.59	197.56 ± 52.54	166.74 ± 29.21
60–90	13543	12549.70 ± 49.88	993.30 ± 126.62	1064.96 ± 73.82
90–120	8465	8215.07 ± 40.36	249.93 ± 100.47	191.41 ± 31.30
120–150	3682	3589.04 ± 26.68	92.96 ± 66.28	4.23 ± 4.65
150–180	1590	1600.98 ± 17.82	-10.98 ± 43.67	16.15 ± 9.09
180–210	731	752.03 ± 12.21	-21.03 ± 29.67	0
210–240	354	377.46 ± 8.65	-23.46 ± 20.71	0
240–270	201	218.26 ± 6.58	-17.26 ± 15.63	0
270–300	117	121.73 ± 4.91	-4.73 ± 11.88	0
300–330	67	67.51 ± 3.66	-0.51 ± 8.97	0
330–360	44	39.39 ± 2.79	4.61 ± 7.20	0
360–390	19	20.54 ± 2.02	-1.54 ± 4.80	0

TABLE VII: Signal candidate counts before and after background subtraction in the pre-v13 data and MC. The errors are statistical only.

Search Window (GeV/c^{2})	**No. Events Observed**	**No. Events Expected**	**Overall Excess**	**Total Stat. Error**	**Total Syst. Error**	**Overall Error**
50–100	19335	17982.20	1352.80	151.33	305.70	341.10
Prediction from MC: 1388.92 ± 107.09						
Significance of Result: 3.97 σ						

TABLE VIII: Excess of events observed in the 50–100 GeV/c^{2} search window after background subtraction has been performed in the pre-v13 data, along with the signal prediction from MC. Agreement is seen between the excess observed and the number of signal events predicted.

A. Closure Test

As a closure test, the sum of the invariant mass templates of the signal MC and background from the TRF were fit to the whole invariant mass range of the observed double tagged candidates. The scale of the two templates were fitted with two unbound scale parameters: "bkg" and "signal" (see Fig. 5).

The fitted background scale parameter is 0.9541 ± 0.0091, which within the error is compatible with the value obtained from Fig. 2, 0.95 ± 0.01. Also, the signal scale parameter is within errors compatible with the expected value of unity. We therefore conclude that the sum of the signal MC and the background as calculated with our TRF-based model and normalized to the high mass region are compatible with the observed double tagged invariant mass over the full region. Furthermore, adding the signal MC to the background template is necessary in order to obtain a good fit to the data.

VI. BACKGROUND ESTIMATION METHOD 2: WITH 0→1 AND Z CORRECTIONS

The data sample used for the second method of background estimation differs from that used for method 1 in the following ways:

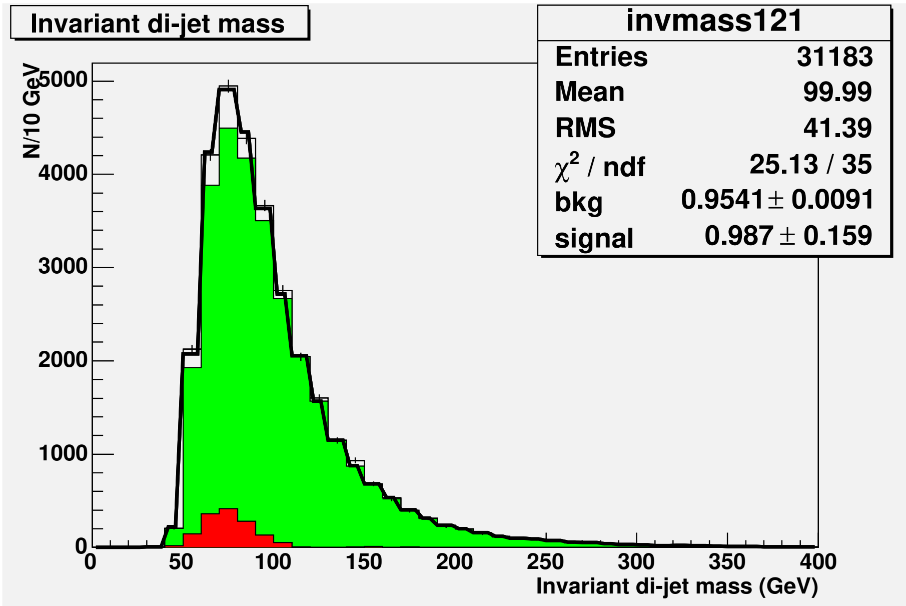

FIG. 5: Closure test of the final invariant di-jet mass spectrum observed in the data after background subtraction. Green (light gray) shaded histogram: expected background, estimated using a jet-based TRF. Black points: total di-jet mass distribution observed in data. Red (dark gray) histogram: expected signal from MC. The bold black line shows the fitted sum of the scaled background and MC templates.

- The full PASS 1 data set is used, with no trigger selection. (Pass1 data did not use the T42 correction and has a different version of the jet energy scale corrections (v5.0).)
- A slightly tighter primary vertex requirement is made, with $\geq$4 tracks attached, and $|z| < 35$ cm (the fiducial b-tagging region).
- A slightly narrower $\Delta\varphi$ cut was used, > 2.9 instead of > 2.75.

Because the TRF is applied to the single-tagged data sample, which is less rich in heavy-flavor jets than the data sample to which it is compared (the double b-tagged data), any differences in either the jet-energy-scale of heavy-flavor jets as opposed to gluon and light-quark jets or their natural invariant mass spectrum will result in a shift of the invariant mass distribution expected for the background, as compared to the data. (In method 1 this is accounted for by fitting the background to double-b-tagged data above invariant mass of 120 GeV.) In this method, this shift is observed and measured using an *un-tagged* data sample. A TRF is derived for the un-tagged data and applied to the un-tagged data, and then compared to the actual single-tagged data. The shift which is derived is called the "0→1" correction, which is then subtracted from the expected background in the double-tagged data sample. This correction relies on the fact that each successive b-tag that is required increases the fraction of heavy-flavor by the same amount. This is verified by comparing the data with MC $b\bar{b}$ events.

This method also makes an additional correction for the effects of $Z \rightarrow b\bar{b}$ events which are present in both the un-tagged and single-tagged data samples from which the 0→1 correction and the single-tag TRF are derived, respectively. A signal peak is first measured in double-tagged data, including the 0→1 correction. The number of events which would exist from this signal peak in the un-tagged and single-tagged data samples is then extrapolated using a $Z \rightarrow b\bar{b}$ MC sample with no, one, and two b-tags required. The signal peak, *measured in data*, is then scaled by these factors and subtracted from the un-tagged and single-tagged samples. Then the TRF and the 0→1 correction are re-derived, the expected background in the double-tagged data is re-calculated, and a new signal peak is observed. This correction process is then repeated, using the new signal peak to estimate the $Z \rightarrow b\bar{b}$ events in the un-tagged and single-tagged data samples, a total of three times, after which the correction and signal peak are stable. Effects of contributions from W→cs production and decay have been investigated and determined to be small.

A. The $0 \rightarrow 1$ tag Shift

The shift in the invariant mass spectrum of the di-jet system caused by applying a single b-tag is first measured. Later this will be subtracted from the estimated background of the double b-tagged data. A TRF is derived on the un-tagged sample and re-applied to the same un-tagged data sample, to predict the background to the single-tagged data sample, as shown in Figure 6. A comparison is also shown in this figure to the $b\bar{b}$ MC di-jet invariant mass spectrum (which was normalized using double-tagged data). The $b\bar{b}$ events make up about 10% of the single-tagged data sample, as opposed to about 2% in the un-tagged data sample. The small contribution of $Z \rightarrow b\bar{b}$ expected in the single-tagged data is also shown. Most of the events in the single-tagged data sample contain only gluon/light-quark jets.

The difference between the single-tagged data and the expected background from the TRF is shown in Figure 7. A comparison of the background-subtracted single-tagged data to the MC $Z \rightarrow b\bar{b}$ invariant mass distribution shows that it is not the result of a signal peak in the single-tagged data, but rather an overall shift in the invariant mass distribution. As mentioned above, this shift is understood to come from either a difference between the gluon/light-quark and the b-quark jet energy scale factors or a difference in the ratio of true cross-section of heavy-flavor jets to light-jets as a function of di-jet invariant mass.

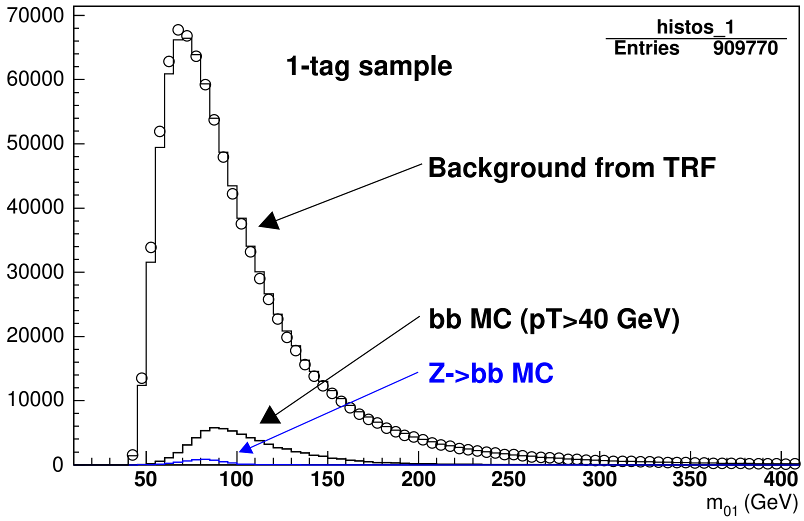

FIG. 6: Comparison between the single-tagged data and the background expected using the TRF method. Comparisons are also shown to the $b\bar{b}$ MC di-jet invariant mass spectrum (which was normalized using double-tagged data) and to the $Z \rightarrow b\bar{b}$ MC di-jet invariant mass spectrum, to give a feel for the composition of the sample. The rest of the events are thought to be gluon/light-quark jet events.

B. The 0- and 1-tag Z Peak Correction

The un-tagged and single-tagged data samples contain signal events, which will now be corrected for. The signal peak observed in double-tagged data, after subtracting the estimated background using the TRF method, is scaled by a factor of 6.5, which is the ratio of single-tagged to double-tagged events in the $Z \rightarrow b\bar{b}$ MC sample. It is important to note that the MC is only used for an overall normalization (the b-tagging efficiency), and does not affect the shape of the signal peak. This scaled signal peak is then divided (bin by bin) by the single-tag data, to determine the estimated fraction of $Z \rightarrow b\bar{b}$ events in each bin of the single-tag data.

The expected fraction of $Z \rightarrow b\bar{b}$ events in each invariant mass bin is then subtracted from the signal-tagged data (the events are weighted by 1-f, where f is the fraction of $Z \rightarrow b\bar{b}$ expected). Then the TRFs are re-derived and re-applied to this corrected, weighted, single-tagged data. The 0→1 tag correction is also re-derived, using the same principles to estimate the fraction of $Z \rightarrow b\bar{b}$ in the un-tagged data sample. The effect from $Z \rightarrow b\bar{b}$ events in the

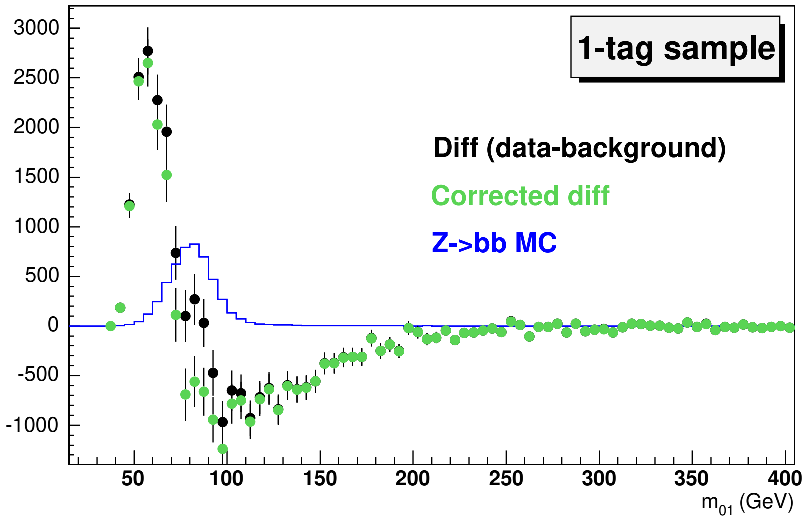

FIG. 7: Difference between the single-tagged data and the background expected using the TRF method, a measure of the 0→1 tag shift, which will be subtracted from the expected double b-tagged data (after proper normalization). The difference is also shown after correcting for the expected $Z \to b\bar{b}$ events in the un-tagged and single-tagged data samples, using the methods described below in Section VI B.

un-tagged sample is relatively small, as expected, so the modification to the 0→1 tag correction is slight, as seen in Figure 7.

Since the size and shape of the original Z peak in data changes slightly after the corrections to the un-tagged and single-tagged data, the corrections themselves can be re-derived using the new Z peak observed. This cyclic process could be continued ad-infinitum, but in practice the solution is stable after only a few iterations, which results in a final Z peak and set of corrections.

C. Applying to Data

The final $Z \to b\bar{b}$ peak derived from data, after all corrections, is considered to be the difference between the background-subtracted double-tagged data and the corrected 0→1 tag shift. This invariant mass distribution difference is shown in Figure 8, and compared to the shape of the $Z \to b\bar{b}$ distribution in MC. (The slightly different shape of the difference between the background-subtracted double-tagged data and the 0→1 tag shift without Z peak corrections to the un-tagged and single-tagged data samples is also shown, to give a feel for the size of the effect.)

The final Z peak in data (with all corrections included) is fit to a Gaussian distribution, which describes the shape of the $Z \to b\bar{b}$ peak well in MC. Both the mean (79±3 GeV) and width (9±3 GeV) of the distribution observed for data are comparable to those derived from $Z \to b\bar{b}$ MC: 81 GeV and 11 GeV, respectively.

810±230 $Z \to b\bar{b}$ events are observed, where the error is taken from the uncertainty in the height of the Gaussian fit. Recall that the data set used to test this background estimation method uses a smaller $\Delta\varphi$ range (2.9 as opposed to 2.75), makes no explicit trigger requirements, and corrects for the signal events in the data from which the TRFs were derived. (570±160 events are observed if the signal events in the data from which the TRFs were derived are not corrected for.) The total number of events expected from $Z \to b\bar{b}$ MC can not be estimated very accurately, because the data were not selected by a given trigger. However, for an assumed luminosity of 300 pb^{-}1, the size of the $Z \to b\bar{b}$ peak observed in data would correspond to an overall trigger/skimming efficiency of about 15% (note that this agrees with the efficiencies presented in Section II A).

VII. SUMMARY OF RESULTS

We have investigated two slightly different methods for extracting an observable $Z \to b\bar{b}$ signal. The first method shows the following excess in 297.5 pb^{-1} of data collected with the MU_JT25_L2M0 trigger:

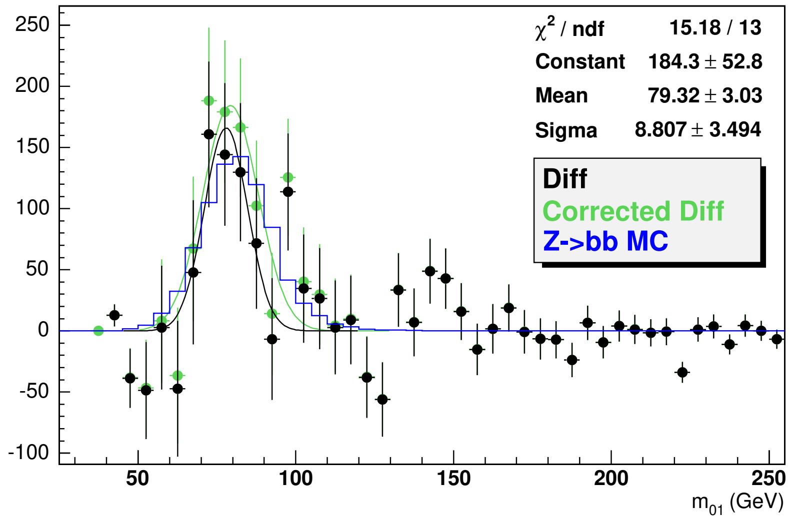

FIG. 8: The final $Z \to b\bar{b}$ peak derived from data, after all corrections (green), compared to the shape of the $Z \to b\bar{b}$ distribution in MC (blue). (The slightly different shape of the Z peak in data without the Z peak corrections to the un-tagged and single-tagged data samples (black) is also shown, to give a feel for the size of the effect.)

$$1352.80 \pm 151.33 \text{ (stat.)} \pm 305.70 \text{ (syst.)}$$

compared to MC expectations of:

$$1388.92 \pm 107.09 \text{ (stat.)}$$

The observed signal amounts to a significance of 4.0 standard deviations.

A second approach, without trigger selection, in a slightly narrower $\Delta\phi$ range, and correcting for the Z contribution to the TRFs, also confirms the observed excess:

$$810 \pm 230 \text{ (stat.)}$$

The position and width of the observed mass peaks are in agreement with MC expectations in both cases. Method 1 finds a peak with a fitted mean of 73±2 GeV and method 2 a fitted mean of 79±3 GeV. The difference between the observed mass peaks can be attributed to the different data samples and JES corrections (p14 pass2 versus pass1) and to the different event selection applied.

VIII. CONCLUSIONS

We have demonstrated that a $Z \to b\bar{b}$ di-jet mass resonance can be successfully extracted from the massive QCD background, already with the first 300 pb^{-1} of Run II data collected by the DØ experiment. The observed excess agrees well with the expectations from Monte Carlo simulations of the Standard Model.

Thanks to the new dedicated $Z \to b\bar{b}$ triggers we have already collected a several times larger data sample which is now waiting to be analyzed. The larger statistics from this sample is not only expected to provide a larger $Z \to b\bar{b}$ peak but also enable a more precise background determination, thus enabling a precision measurement.

Acknowledgments

We thank the staffs at Fermilab and collaborating institutions, and acknowledge support from the DOE and NSF (USA), CEA and CNRS/IN2P3 (France), FASI, Rosatom and RFBR (Russia), CAPES, CNPq, FAPERJ, FAPESP and FUNDUNESP (Brazil), DAE and DST (India), Colciencias (Colombia), CONACyT (Mexico), KRF (Korea), CONICET and UBACyT (Argentina), FOM (The Netherlands), PPARC (United Kingdom), MSMT (Czech Republic), CRC Program, CFI, NSERC and WestGrid Project (Canada), BMBF and DFG (Germany), SFI (Ireland), A.P. Sloan Foundation, Research Corporation, Texas Advanced Research Program, Alexander von Humboldt Foundation, and the Marie Curie Fellowships.

---

[1] T. Dorigo,*Observation of Z Decays to b Quark Pairs at the Tevatron Collider*, hep-ex/9806022.
[2] A. Haas *et.al.*,*DØ Search for Neutral Higgs Bosons at High tanβ in Multi-jet Events Using p14 Data*, DØNote 4671, hep-ex/0504018.
[3] A. Jenkins and A. Goussiou, *An Investigation of b-jet Energy Resolution in* $Z^0H^0 \rightarrow ee b\bar{b}$ *and* $Z^0 \rightarrow b\bar{b}$, DØNote 4136.
[4] G. C. Blazey *et al.*, in *Proceedings of the Workshop: "QCD and Weak Boson Physics in Run II,"* edited by U. Baur, R. K. Ellis, and D. Zeppenfeld, (Fermilab, Batavia, IL, 2000) p. 47; see Sec. 3.5 for details.
[5] The BID skim was the best data-set available at the time of this analysis. A new skim,"2JET", has been proposed for future use, which would require exactly two 0.5 cone jets with $p_T > 15$ GeV/c but not demand any muons. Initial studies indicate that the total efficiency for signal events would increase by 50–100% using this new skim.
[6] B.R.(b $\rightarrow \mu$) = 0.11. Muons from the cascade decays $b \rightarrow c \rightarrow \mu$ and $b \rightarrow \tau \rightarrow \mu$ are significant, and the total probability for a b-jet to have a muon is nearly 20%. For $Z \rightarrow b\bar{b}$ events in which *at least one* b-jet has a muon which passes muon cuts, the maximum signal efficiency hence is $1-(1-0.2)$x$(1-0.2) = 36\%$. For events with two muons, this number reduces to 4%.