DØnote 4388

Measurement of the Ratio of Inclusive Cross Sections $\sigma(p\bar{p} \to Z + b)/\sigma(p\bar{p} \to Z + j)$ at $\sqrt{s} = 1.96$ TeV

The DØ Collaboration
URL: http://www-d0.fnal.gov
(Dated: March 29, 2004)

This note describes a preliminary measurement of the ratio of $Z + b$ and $Z + j$ inclusive cross sections at DØ. The analysis is based on 152 pb^{-1} of data in the $Z \to \mu^+\mu^-$ channel, and 184 pb^{-1} in the $Z \to ee$, where there is no explicit requirement on the electric charges of the electrons. The ratio $\sigma(Z + b)/\sigma(Z + j)$ is 0.024 ± 0.007, for jets with $E_T > 20$ GeV and $|\eta| < 2.5$.

Preliminary Results for Winter 2004 Conferences

I. INTRODUCTION

Heavy flavor production in association with a Z boson is a primary source of background to several of the Higgs production channels at the Tevatron. Inclusive $Z + b$ production in particular, provides background to Standard-Model (SM) Higgs production of associated Z and H, and to Supersymmetric (SUSY) Higgs production $gb \rightarrow hb$. An interesting aspect of this channel is that it probes the parton distribution function of the b quark, because the initial state involves a b quark from the sea. There is recently renewed interest in this channel [1]. This note describes the measurement of the ratio of cross sections of inclusive $Z + b$ production to inclusive $Z + j$ production in both the dimuon and dielectron decay channels of the Z.

II. DATA SAMPLE AND EVENT SELECTION

The analysis is based on a total recorded integrated luminosity of 189 pb^{-1}, with varying luminosities for different channels that depend on the operating quality of any detector during the acquisition of data. For the dimuon and dielectron samples, analyses are based on 184 pb^{-1} and 152 pb^{-1} of integrated luminosities respectively. There is a 6.5% uncertainty on the absolute luminosity, which cancels in the measurement of ratio of cross sections.

In the subsections below, we describe event selection and yields specific to each channel, but selection of hadronic jets is common to both. Acceptable jets must :

- Be reconstructed with Run II cone algorithm of $\Delta R = 0.5$ [2].
- Have transverse energies $E_T > 20$ GeV (after jet energy-scale corrections) and have pseudorapidity $|\eta| < 2.5$.

In addition to the calorimeter jet selection, we also impose further requirements on the tracks and hits in the tracking system. Taggability of jets is a concept introduced to select a sample of stable jets with reduced dependence on the conditions of calorimeter and tracking detectors. A calorimeter jet is considered taggable if

- A jet formed from charged tracks (track jet) is within $\Delta R < 0.5$ of the calorimeter jet.
- Track jet consists of applying cone algorithm of $\Delta R = 0.5$ to charged tracks $p_T > 0.5$ GeV/c with a seed track $p_T > 1.0$ GeV/c. Each track must be within 2 cm of the interaction vertex in z and have at least one hit in the Silicon Microstrip Tracker (SMT).

A. Dielectron Sample

Events in the dielectron channel are required [3]:

- To be triggered through the presence of two electromagnetic (EM) objects.
- To have ≥ 2 isolated EM objects, each with $E_T > 15$ GeV and $|\eta| < 2.5$, to be reconstructed offline.
- That 2 EM objects must satisfy a multidimensional Gaussian estimator based on energy fractions of showers observed in the EM sections of the calorimeter, the total deposited energy, extrapolated vertex z position and transverse shower shape.
- At least one EM object must have a matching charged track.

The signal region of interest is the di-EM invariant mass of 80 GeV/c^2 $< M_{ee} <$ 100 GeV/c^2. A total of 15,613 events remain in the mass window after implementation of the above criteria. The background from multijets or "QCD", where the jets satisfy the criteria for EM objects, and Drell-Yan (DY) non-resonant dilepton production is estimated from side bands in M_{ee}. The method used is the so-called "matrix method", that relies on the fact that the efficiency for QCD background is different than for signal for any chosen selection criterion. Track matching was used as the discriminant, with the DY having the same efficiency as the Z events. The QCD and the DY contribution is derived from the data by fitting the di-EM invariant mass distribution, with a combination of Gaussian-convoluted Breit-Wigner used to model the Z peak, and an exponential, to describe the combination of QCD and DY background. We ignored any interference between Z and DY amplitudes in this analysis. The level of DY and QCD in the mass region is estimated as 4.7%. As a cross check, events in the side bands were used to extrapolate into the Z mass window, with good agreement observed between the two methods.

The level of background in inclusive $Z + jets$ candidates is different from that in the inclusive Z distribution. Out of 2,219 events with one or more jets, approximately 7.3% or 161.9 events are estimated to be background. Applying the taggability to 2,661 jets in these events, 1,910 jets or $(72 \pm 1)\%$ are found to be taggable. In a sample with tighter requirements on electrons, and therefore less background, the taggability is observed to be $(75 \pm 1)\%$. We use this value of taggabilty, since it is more applicable to the sample after background subtraction. We take the difference as a source of systematic uncertainty, but since we are taking ratios, the effect of thd difference in taggability is small.

B. Dimuon Sample

Events in the dimuon channel are required [4]:

- To have at least 2 muons matched from the muon detectors to central tracks of $p_T > 15$ GeV/c and $|\eta| < 2.0$.
- To have hits in the the muon detectors in time with beam crossings, and matching tracks to be within 2.5 mm in the transverse plane.
- To have both muons isolated and of opposite electric charge.

The mass window for Z is 65 GeV/$c^2 < M_{\mu\mu} <$ 115 GeV/c^2, and 9,856 Z candidate events remain in the mass window after imposition of these criteria.

Out of 1,239 events with one or more jets, 1,008 events remain after the requirement that there be at least one taggable jet. The taggability measured in the data is $(79.7 \pm 1.1)\%$. The background is estimated using side band in a manner similar to that used in dielectron channel, but using isolation as a way to differentiate background from signal. The main background in this channel is $b\bar{b}$ production, where both b-jets have muons that satisfy the isolation criteria. The side band method yields an expected background of 16.8 events. Distribution in transverse energy and pseudorapidity of taggable jets are compared to $Z + j$ Monte Carlo and background in Figs 1 and 2 in the dimuon channel.

III. ANALYSIS

A. b-jet Tagging

A secondary vertex b-jet tagging algorithm is used to identify heavy flavored jets in the analysis. A jet is considered tagged when it is taggable, and it has:

- A secondary reconstructed vertex with a decay length significance > 7
- $\Delta R_{jet,SV} < 0.5$ in the opening angle between the direction of the calorimeter-based jet axis and the momentum vector of the components of the reconstructed secondary vertex.

The b-tagging efficiency (ϵ_b) and light flavored tagging rate (ϵ_L), also known as the tag rate function, of the b-tagging algorithm in DØ are parametrized as function of jet E_T and η. The parametrization of ϵ_b is derived from data using events with muons embedded in jets, where there can be contributions from light, charm and b-jets. The b-tagging efficiency is extracted from the heavy-flavor component in this $\mu + jets$ through various methods, one of which involves fitting the distribution of relative transverse momentum (p_{Trel}) of the muon with respect to the jet. Mistag rate (ϵ_L) is also derived from data, and compensates for the effects of long-lived particles. Vertices identified as originating from K_s, Λ and photon conversions are removed during secondary vertex finding. However, if one of the tracks is misreconstructed or not reconstructed, then this will enhance the tagging rate in these events. Checks of systematics using different types of samples is used to assign systematic uncertainties to ϵ_L and ϵ_b.

An estimate of the expected number of tags from the light-flavored jets can be made from the data, based on the jet E_T and η distributions of jets, most of which are not heavy-flavored. Comparison of inclusive $Z + j$ events generated with the ALPGEN LO matrix element and passed to the PYTHIA Monte Carlo for showering, with inclusive $Z + b$ PYTHIA events, shows good agreement in terms of the kinematic distributions, we therefore expect that we can use the shapes derived from the data to estimate expected b-tagging efficiency and mistag rate [3, 5, 6]. Given the kinematic distributions of the jets, the average b-tagging efficiency and mistag rate per jet are expected to be $(33.4 \pm 4.7)\%$ and $(0.25 \pm 0.02)\%$, respectively, in the dielectron channel, and $(33.6 \pm 4.7)\%$ and $(0.29 \pm 0.02)\%$ in the dimuon channel, where the errors are systematic uncertainties. To obtain the event mistag rate, we take into consideration the jet multiplicity and the average event mistag rate is $(0.28 \pm 0.02)\%$ in case of the dielectron channel.

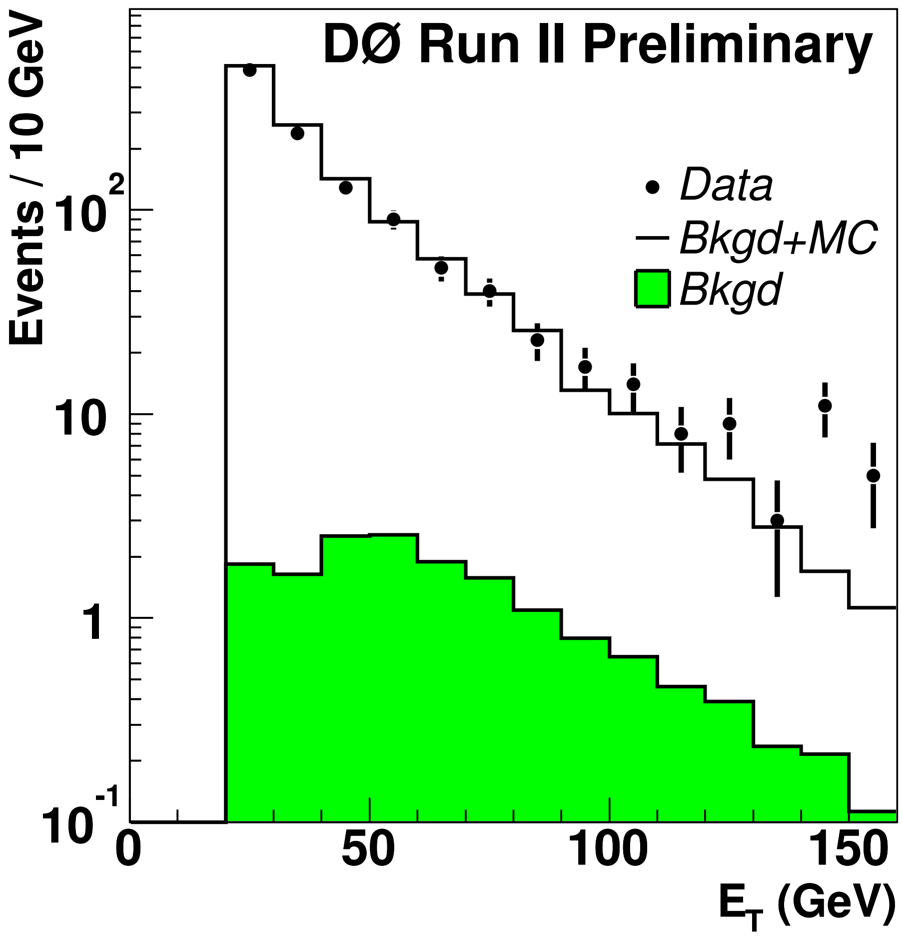

FIG. 1: E_T distribution of taggable jets in the dimuon channel compared to $Z + j$ ALPGEN with PYTHIA showering and full detector simulation (open histogram), and QCD/DY background derived from data (also shown separately as the shaded histogram). The errors on the data points are statistical. Prediction is normalized to data. Mean transverse energy of jets in data is 42 GeV.

After applying b-tagging, 27 $ee + jets$ events are left, with an expected background from the continuum and QCD of 3.4 events, as estimated using the side-band method. Assuming all the jets are light-flavored would yield 4.6 events, based on mistag rates.

In the dimuon channel, 15 events are observed, with 4.9 events from the continuum and QCD, as estimated using the side-band method. One of these events has two jets that have b-tags. We would expect about 2.8 events due to mistags, assuming that all jets were light jets.

B. Extraction of Ratio

After subtracting contributions from QCD and the Drell-Yan continuum, two equations, one before b-tagging is applied and one after the application of b-tagging, determine the contributions from different flavors in the remaining events:

$$N_{beforeb-tag} = t_b N_b + t_c N_c + t_L N_L \quad (1)$$

$$N_{b-tagged} = \epsilon_b t_b N_b + \epsilon_c t_c N_c + \epsilon_L t_L N_L, \quad (2)$$

where N_b, N_c and N_L are number of events with b, c and light jets, t's are taggabilities per event for different jet types, and the ϵ_i are the corresponding event tagging efficiencies. We assume that per jet tagging efficiencies ϵ_b and

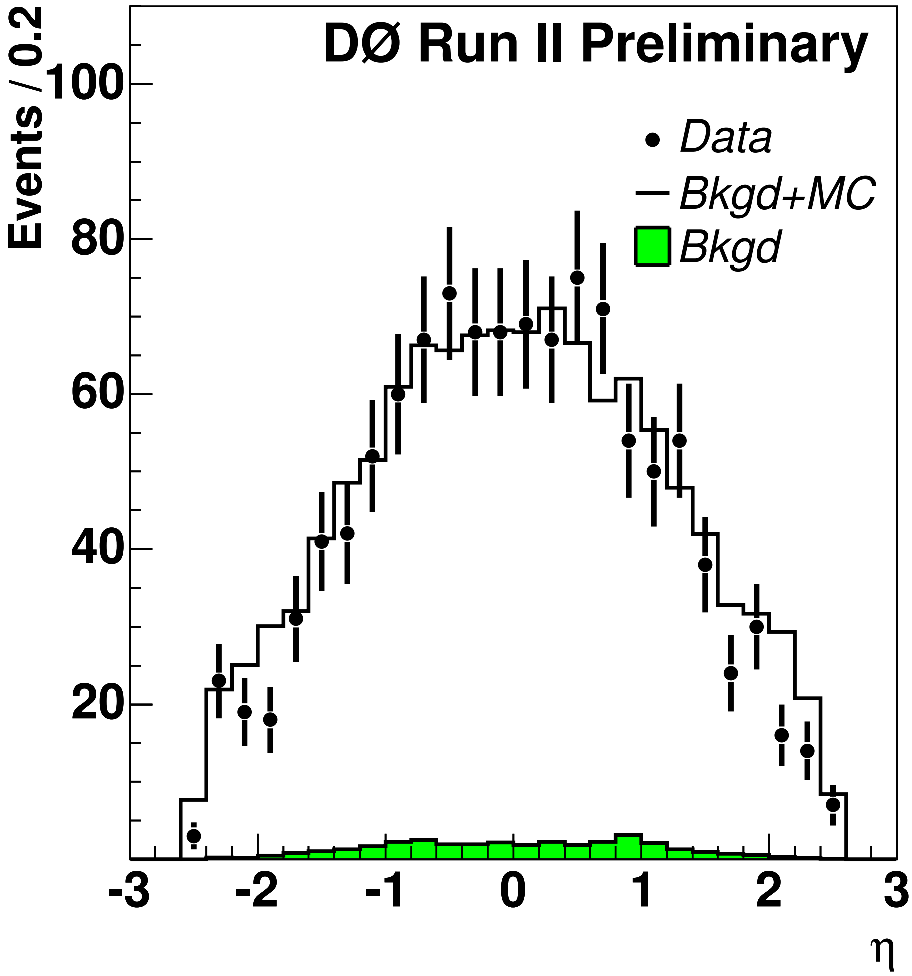

FIG. 2: η distribution of taggable jets in the dimuon channel compared to $Z + j$ ALPGEN with PYTHIA showering and full detector simulation (open histogram), and QCD/DY background derived from data (also shown separately as the shaded histogram). The errors on the data points are statistical. Prediction is normalization is to data.

ϵ_c are the same as per event tagging efficiencies. Tagging efficiency for events with no heavy flavor is calculated $\epsilon_{Levent} = 1 - \prod^{N_{jets}}(1 - \epsilon_{Ljet})$, whereas $\epsilon_{bevent} = \epsilon_{bjet}$ and $\epsilon_{cevent} = \epsilon_{cjet}$.

To measure the $Z + b$ cross section, requires an unfolding of N_b for the jet reconstruction efficiency and the effects of resolution. However, we are interested in a ratio of cross sections $N_b/(N_b + N_c + N_L)$, in which such effects tend to cancel out, and the measurement is sensitive only to the difference in jet reconstruction efficiencies and resolutions between species. The cancelation would be exact if the shape of the distributions were the same, and a systematic uncertainty is assigned to account for any differences.

The taggability t_L is measured using the data, as alluded to in earlier sections, and t_b is measured from the Monte Carlo and scaled such that $(t_b)_{data} = (t_L)_{data} * \left(\frac{t_b}{t_L}\right)_{MC}$. Expected values of t_b in data are $(79.2 \pm 1.3)\%$ in the dielectron channel and $(81.7 \pm 1.1)\%$ the dimuon channel. We assume $t_c = t_b$.

Since the ϵ_b is derived from events with μ embedded inside a jet, whereas most of our b-tagged jets do not contain such μ, the difference in b-tagging efficiencies for "hadronic" b-jets and "muonic" b-jets is derived from Monte Carlo and the ratio is used to correct the original ϵ_b. Tagging efficiency of the hadronic b-jets is 86% that of muonic b-jets in MC. We cannot at this point derive ϵ_c from data and we rely on MC in a comparison of the samples $Z \to b\bar{b}$ and $Z \to c\bar{c}$ samples. We assume that $(\epsilon_c/\epsilon_b)_{Data} = (\epsilon_c/\epsilon_b)_{MC} = 0.266 \pm 0.003$.

Equations 1 and 2 have 3 unknowns, and we therefore assume a calculation of $N_c = 1.69N_b$ [1] to solve for the unknowns. With this assumption, the ratios of $\sigma(p\bar{p} \to Z + b)/\sigma(p\bar{p} \to Z + j)$ are 0.031 ± 0.008 for the dielectron

	Dielectron Channel	Dimuon Channel
N_b	61.0	18.7
N_c	102.8	31.6
N_L	1797.4	1273.4

TABLE I: Central values of N_i for each channel.

channel and 0.015 ± 0.005 for the dimuon channel, where the errors are statistical. The central values of N_i are listed in Table I. The combination using N_b, N_c and N_L, where the individual N_i correspond to a sum of both channels, yield 0.024 ± 0.005. The shape of the b-tagged jet E_T spectrum is compared to the expected background plus $Z + b$ Monte Carlo in Fig. 3. A distribution of the secondary vertex decay length significance (Fig. 4) shows evidence for heavy flavor component in the b-tagged candidate events. A display of a b-tagged candidate event is shown in Fig. 5.

C. Systematic Uncertainty

The sources of systematic errors are as follows:

- Tagging efficiency for b and c jets – Uncertainty in the b tagging efficiency is the largest source of systematic uncertainty in the measurement. The b and c tagging efficiency is fluctuated $\pm\sigma$ assuming 100% correlation and the ratios of cross sections are extracted. For c jets, there is additional uncertainty from ϵ_c/ϵ_b obtained from MC.

- Tag-rate function – The tag-rate function, which parametrizes the expected rate of tagging of light-flavored jets, is derived from the data. However, there remains a dependence on sample and this is a source of the systematic uncertainty. Using a generic hadronic events triggered using the calorimeter, light-jet tagging efficiency is measured to be 0.23%, and for a sample of events with enhanced EM fraction and low missing E_T this number is 0.26%. 0.25% per jet tagging (or 0.28 per event tagging) efficiency is obtained for the combined data. ϵ_L is fluctuated to estimate its effects.

- Jet energy scale – The jet energies are fluctuated by the ± 1 standard deviation uncertainty in energy scale, and the analyses are repeated to gauge the impact on the ratio of cross sections. The effect is largely mitigated by the fact that we are measuring a ratio.

- Uncertainty of background estimation based on different methods.

- There is an uncertainty from theory in $\sigma(Z + c)/\sigma(Z + b)$ [1].

- A small difference observed in t_b/t_L using different samples $Z + b/Z$+light jet and $Z \to b\bar{b}/Z \to q\bar{q}$ is regarded as an uncertainty.

- Muonic jet vs hadronic jets – The tagging efficiency for hadronic jets and muonic jets is different in MC, whereas the b-tagging efficiency is measured in the data using muonic jets. The tagging efficiency in hadronic jets is estimated to be 86% that of muonic jets, derived from $Z \to b\bar{b}$ MC. The same ratio measured in $Z + bb$ MC is measured to be 84% and the difference of 2% is taken to be the systematic uncertainty.

- Different p_T dependence in the jet reconstruction efficiency between light, b and c is accounted for as a systematic uncertainty in the method.

The effects of systematic uncertainties on the combined measurement are listed in Table II. Folding in these systematic uncertainties, yield the result $0.024 \pm 0.005(stat)^{+0.005}_{-0.004}(syst)$ for the ratio of $\sigma(Z + b)/\sigma(Z + j)$.

IV. CONCLUSION

We observe $Z + b$ production at DØ, and measure a preliminary ratio of inclusive cross sections:

$$\frac{\sigma(Z+b)}{\sigma(Z+j)} = 0.024 \pm 0.007, \tag{3}$$

This corresponds to jets of $E_T > 20$ GeV and $|\eta| < 2.5$. This is compatible with the theoretical expectation of Ref. [1].

Source	Upward Fluctuation (%)	Downward Fluctuation (%)
b/c tag efficiency	15.4	11.7
background estimation	6.3	6.0
mistag rate	3.4	3.1
jet energy scale	9.9	3.4
taggability	2.8	2.7
correction for hadronic jet	1.7	1.9
jet reconstruction efficiency	1.8	2.0
$\sigma(Z+c)/\sigma(Z+b)$	2.5	2.4
Total	20.3	15.5

TABLE II: Systematic uncertainties in the combined ratio of cross sections

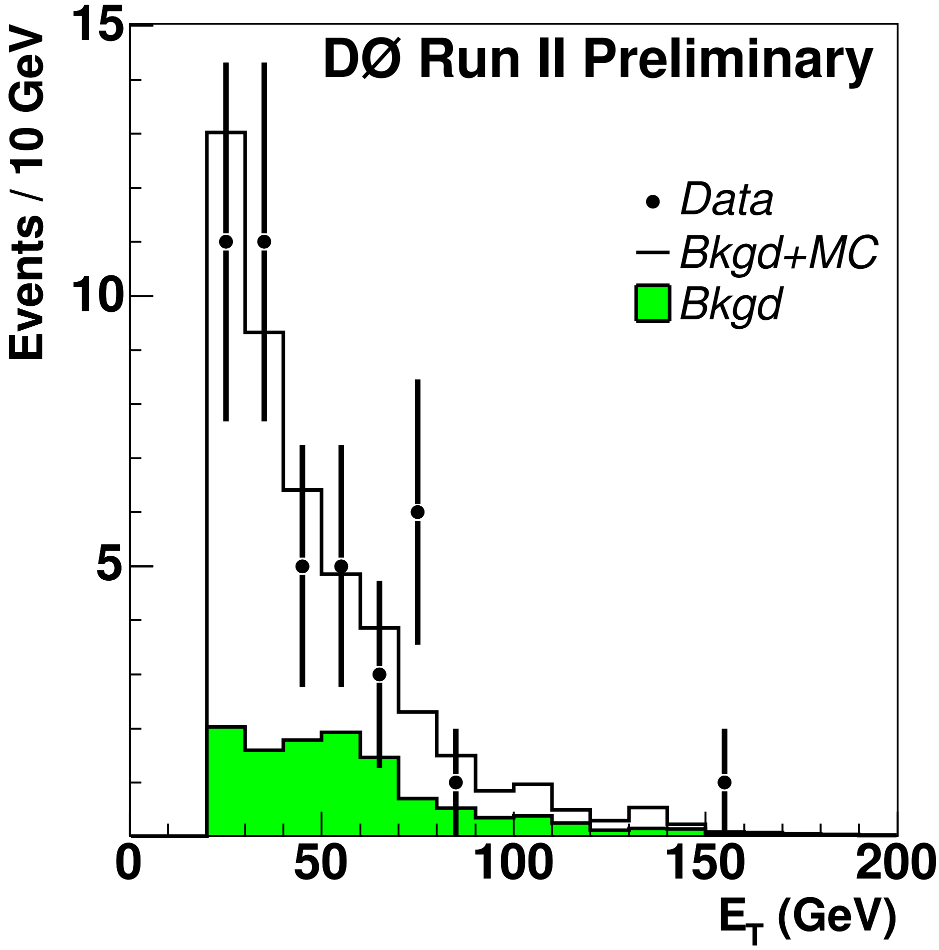

FIG. 3: Transverse energy spectrum of b-tagged jets. The shaded histogram is the contribution from QCD and DY background sources and mistag from light jets, as estimated from data. The solid histogram represents the sum of the background and simulation of $Z+b$ via PYTHIA. The contribution from $Z+b$ is scaled such that the sum of $Z+b$ and background is normalized to the data. The error bars on the data points are statistical.

Acknowledgments

We thank the staffs at Fermilab and collaborating institutions, and acknowledge support from the Department of Energy and National Science Foundation (USA), Commissariat à L'Energie Atomique and CNRS/Institut National de Physique Nucléaire et de Physique des Particules (France), Ministry of Education and Science, Agency for Atomic Energy and RF President Grants Program (Russia), CAPES, CNPq, FAPERJ, FAPESP and FUNDUNESP (Brazil),

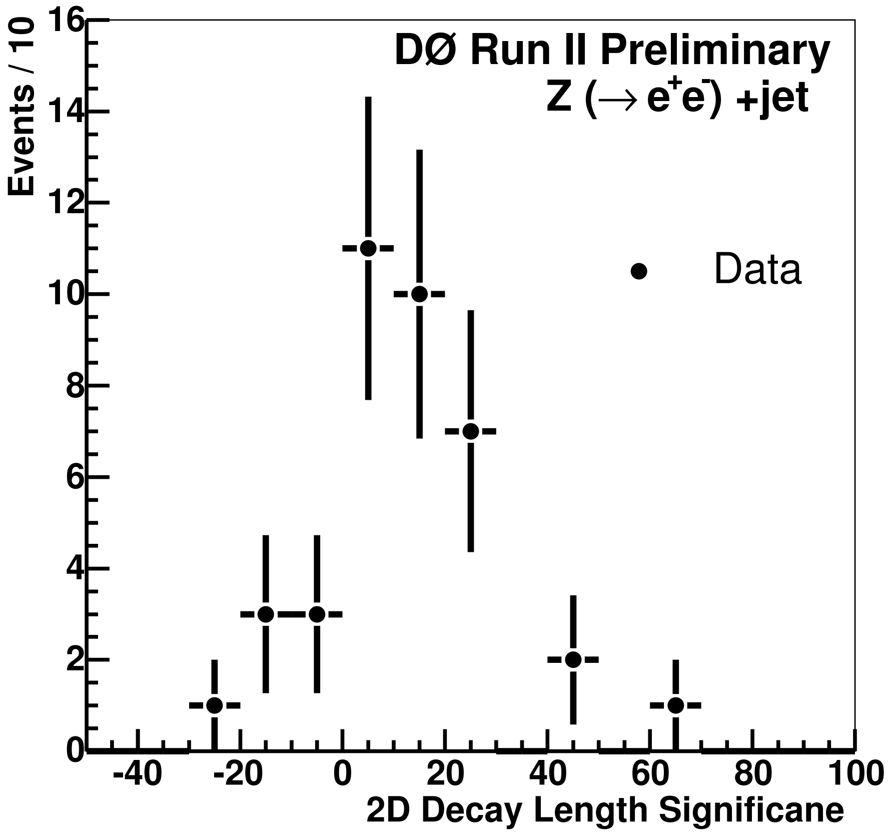

FIG. 4: Decay length significance distribution of secondary vertices in the transverse plane in $Z \to ee$ b-tagged candidate events. The errors bars on the data points are statistical.

Departments of Atomic Energy and Science and Education (India), Colciencias (Colombia), CONACyT (Mexico), Ministry of Education and KOSEF (Korea), CONICET and UBACyT (Argentina), The Foundation for Fundamental Research on Matter (The Netherlands), PPARC (United Kingdom), Ministry of Education (Czech Republic), Natural Sciences and Engineering Research Council and WestGrid Project (Canada), BMBF (Germany), A.P. Sloan Foundation, Civilian Research and Development Foundation, Research Corporation, Texas Advanced Research Program, and the Alexander von Humboldt Foundation.

[1] J.M. Campbell, R.K. Ellis, F. Maltoni and S. Willenbrock, hep-ph/0312024.
[2] G.C. Blazey *et.al.,* Run II Jet Physics, hep-ex/0005012.
[3] K. Hanagaki, DØ internal note 4353.
[4] Y. Mutaf, DØ internal note 4354.
[5] M.L. Mangano, M. Moretti, F. Piccinini, R. Pittau and A. Polosa, hep-ph/0206293.
[6] T. Sjöstrand, P. Eden, C. Friberg, L. Lönnblad, G. Miu, S. Mrenna and E. Norrbin, Computer Phys. Commun. 135 (2001) 238.

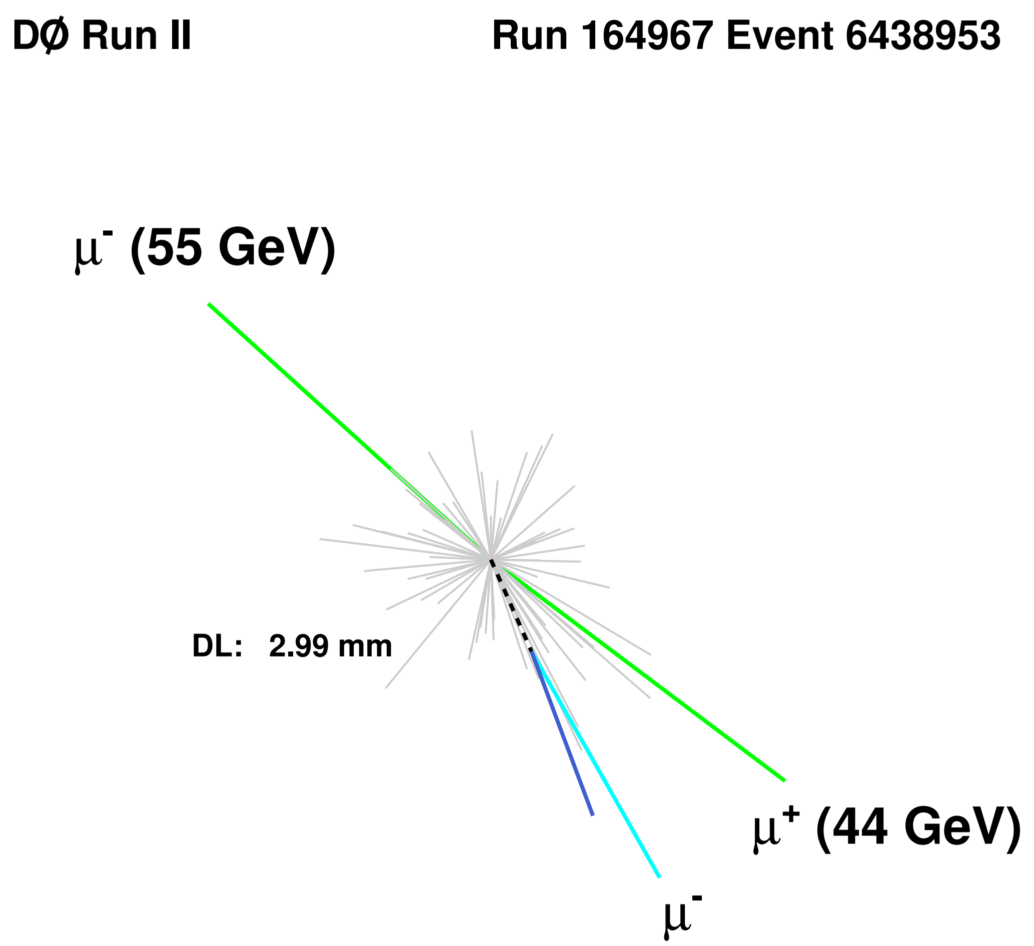

FIG. 5: An event display of a $Z \to \mu^+\mu^-$ b-tagged candidate event. The tagged jet has a secondary vertex about $3mm$ away from the main interaction point. A third muon (in lighter shading) emerges from the secondary vertex.