CDF note 8590

Search for New Particles Decaying to Z^0+jets

The CDF Collaboration
URL http://www-cdf.fnal.gov
(Dated: November 6, 2006)

We present the results of a search for new physics that couples to Z^0 bosons in conjunction with jets. We describe a method that uses data alone to predict the background from Standard Model Z^0+jet events, which is the dominant background and is virtually indeterminable with Monte Carlo. This method can be similarly applied to other analyses requiring background predictions in multi-jet environments, as we show when cross-checking the method to predict the background from W+jets in $t\bar{t}$ production. We see no significant excess in the data above the background prediction, and set a limit using a 4th-generation quark model to quantify the acceptance.

Preliminary Results for Winter 2006/2007 Conferences

I. INTRODUCTION

This note presents a search for new particles decaying to Z^0 gauge bosons created in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV with the CDFII detector [1] at the Fermilab Tevatron. Searches for new physics that couples to Z^0's have large background from Standard Model Z^0 production, as the cross section for new models of interest are generally much lower than the Standard Model cross section. Therefore, understanding how to measure and reject this large background constitutes the bulk of the effort in analyses searching for new physics coupling to Z^0 bosons. This analysis searches for new particles that decay to Z^0's in conjunction with jets, extending and complimenting other work with Z^0's in the final state [2, 3, 4, 5, 6, 7, 8].

There are a variety of new models predicting new particles decaying to Z^0's. We strive to retain model independence, but for optimization and specific acceptance studies use the 4th-generation model [9]. The 4th-generation down-type quark (called the b') may have a large branching ratio to bZ^0 via the loop diagram in figure 1, if kinematically allowed.

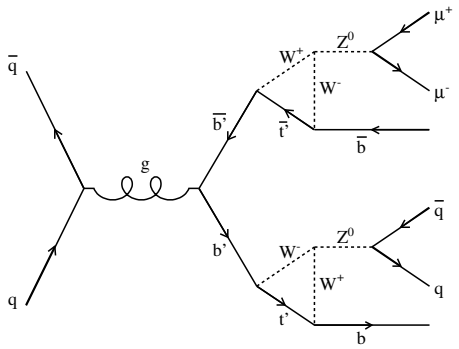


FIG. 1: Feynman diagram for b' production.

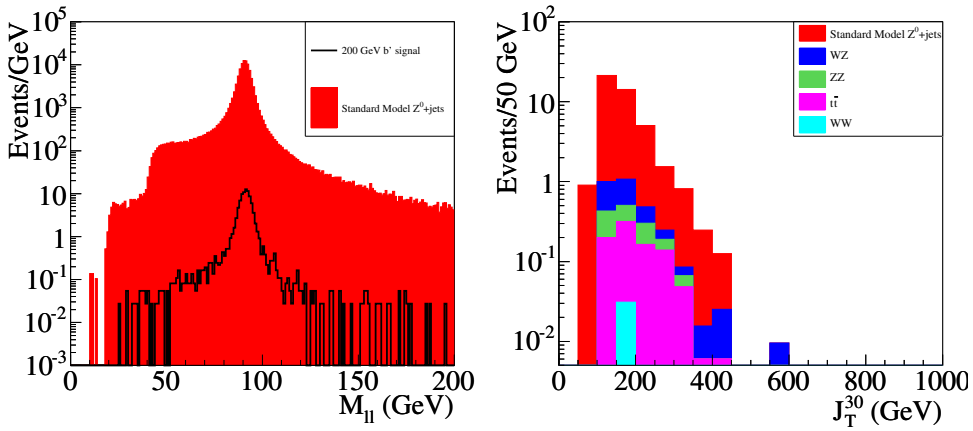
II. DATA SAMPLE & EVENT SELECTION

This search is performed using 1.055 fb^{-1} of data collected with electron and muon triggers. The electron trigger requires at least one electromagnetic energy cluster with $E_T > 18$ GeV and a matching track with $p_T > 9$ GeV. The muon trigger requires at least one track with $p_T > 18$ GeV with matching hits in the muon drift chambers. We select Z^0 candidate events offline by requiring at least one pair of electrons or muons with $p_T > 20$ GeV and invariant mass in the range $81 < M_{\ell\ell} < 101$ GeV. The search was performed using a "blind" analysis technique, in which the selection was chosen and backgrounds were predicted before looking in the signal region.

The cross section for new models of interest are many orders of magnitude smaller than the cross section for Standard Model Z^0 production. For illustration, we plot the expected invariant mass distribution from Standard Model $Z^0 \to \ell\ell$ events compared to an example b' signal with $m_{b'} = 200$ GeV (both generated using `PYTHIA` [10]) in figure 2a. It is apparent that in order to observe new signals, the Standard Model background needs to be rejected by several orders of magnitude while the signal is kept with high efficiency. To reject this background, this analysis requires the presence of high-E_T jets. The variables we use are:

$$N_{\rm jet}^{30} = \text{Number of jets in the event with } E_T > 30 \text{ GeV and } |\eta| < 2$$
$$J_T^{30} = \text{Scalar sum of } E_T\text{'s of all jets in the event with } E_T > 30 \text{ GeV and } |\eta| < 2$$

In order to be sensitive to a range of new particle masses, we design a selection that takes into account that, as a function of mass, the cross sections decrease but the jet energies increase. That is, for higher masses, we cut harder on the jet energies to remove more of the Standard Model background, becoming more sensitive to lower cross sections, while keeping the efficiency as high as at lower mass. We have found that the selection $N_{\rm jet}^{30} \geq 3$ and $J_T^{30} > m_{b'}$ is nearly maximally sensitive for b' masses of interest, in the range $150 < m_{b'} < 350$ GeV. That is, we perform the search by first requiring $N_{\rm jet}^{30} \geq 3$ and then requiring $J_T^{30} > X$, where X is scanned through in 50 GeV steps.



(a) (b)

FIG. 2: (a) Invariant mass distribution of Standard Model single $Z^0 \to \ell\ell$, compared to a b' signal. (b) J_T^{30} distribution after the $N_{\rm jet}^{30} \geq 3$ cut of various backgrounds in Monte Carlo. Each contribution is stacked on top of the one below it.

III. BACKGROUNDS

In this signal region, there are potential backgrounds from the following sources:

- Standard Model single-Z^0 production with associated jets (Z^0+jets)
- Standard Model WZ+jets, where the W decays to jets
- Standard Model ZZ+jets, where one of the Z's decays to jets
- Standard Model $t\bar{t}$+jets, where both W's decay to leptons
- QCD multijet events, where two of the jets fake leptons
- Multijet events occurring in conjunction with a cosmic ray

As a first step to understanding the relative size of each background, in figure 2b we plot the J_T^{30} distribution of backgrounds using Monte Carlo (all but the QCD and cosmic backgrounds) after selecting events with $N_{\rm jet}^{30} \geq 3$. It is clear that, according to the Monte Carlo, the Z^0+jets background is dominant. Additionally, using the sidebands of the $M_{\ell\ell}$ distribution, we find that the QCD background is an order of magnitude smaller than the Z^0+jets background. Using timing information of the lepton tracks, we find the cosmic background is completely negligible.

Using `PYTHIA` to estimate the dominant Z^0+jets background is problematic, as this Monte Carlo does not contain higher order hard-scattering diagrams. Other higher-order Monte Carlos attempt to include scattering terms beyond leading order, although doing so is not a trivial theoretical problem, so these calculations need careful validation using data. Rather than using the data indirectly, as merely a tool to validate the higher order Monte Carlos, we chose to develop a procedure to estimate the Z^0+jets background directly and solely from data.

For the background prediction, two quantities are needed: the total number of events after requiring $N_{\rm jet}^{30} \geq 3$, and the shape of the J_T^{30} distribution after the $N_{\rm jet}^{30}$ requirement. We first describe the method for finding the former, then describe the method for finding the latter. In both cases, the method is validated with data from control samples.

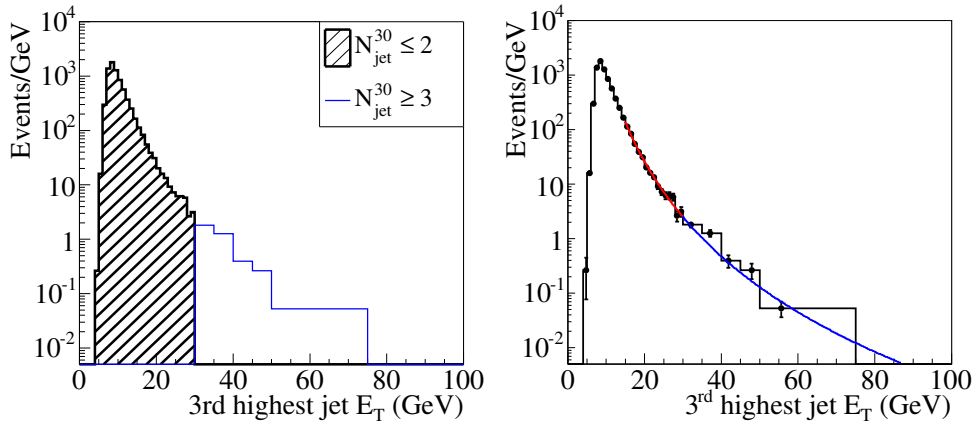
A. The Z^0+jets Background in the $N_{\rm jet}^{30} \geq 3$ Bins

We predict the total background from Standard Model Z^0+jets in the $N_{\rm jet}^{30} \geq 3$ bins using data. In order to make this prediction, we use the intuition that, since jets are counted above an E_T threshold, the $N_{\rm jet}$ distribution is completely determined from the jet E_T distributions. So, we use jet E_T distributions in the $N_{\rm jet}^{30} \leq 2$ bins from the

Z^0+jets data itself to predict the number of background events expected with $N_{\rm jet}^{30} \geq 3$. The approach is illustrated in figure 3a, where we plot the E_T distribution of the 3rd-highest E_T jet in the event after requiring $N_{\rm jet}^{30} \leq 2$ (the hatched histogram) using Standard Model $Z^0 \to \mu\mu$ Monte Carlo. The distribution cuts off at $E_T = 30$ GeV, since we have selected events with $N_{\rm jet}^{30} \leq 2$. In the same figure, we overlay the same jet E_T distribution, but for events with $N_{\rm jet}^{30} \geq 3$. These events complete the remaining portion of the distribution above 30 GeV. In other words, the $N_{\rm jet}^{30} \leq 2$ and $N_{\rm jet}^{30} \geq 3$ cuts separate the E_T distribution of the 3rd highest E_T jet at $E_T = 30$ GeV. We can therefore fit a parameterization to this distribution in the region $E_T < 30$ GeV, and extrapolate it to the region $E_T > 30$ GeV to get the expected background in the $N_{\rm jet}^{30} \geq 3$ region. The parameterization we use is:

$$f(E_T) = p_0 \frac{e^{-E_T/p_1}}{(E_T)^{p_2}} \tag{1}$$

This parameterization was motivated by the theoretically expected distribution of event mass, which is a convolution of the generic $1/s$ behavior of cross sections and the parton distribution functions. Additionally, the parameterization matches the dependence of jet E_T distributions seen in Monte Carlo and observed in control regions of data.



(a) (b)

FIG. 3: E_T distribution of 3rd-highest E_T jet in Standard Model $Z^0 \to \mu\mu$ Monte Carlo,
(a) for two different ranges of $N_{\rm jet}^{30}$.
(b) fit to the parameterization in equation (1) in the region $15 < E_T < 30$ GeV.

We show the jet E_T distribution from Monte Carlo $Z^0 \to \mu\mu$ events (the same distribution shown in figure 3a) fit to our parameterization in the range $15 < E_T < 30$ GeV in figure 3b. The extrapolation matches the actual distribution in Monte Carlo well.

We have cross-checked that this extrapolation procedure functions properly in data in various control samples. One such control sample is jet triggers, dominated by QCD dijet data. In this sample, a dijet pair is selected, and additional jets above the 30 GeV threshold are counted. We have verified the extrapolation procedure accurately predicts the number of events with $N_{\rm jet}^{30} \geq 3$ in this control sample.

Another such control sample is W+jet data (with the W decaying to $\ell\nu$), selected by requiring a single muon with $p_T > 20$ GeV in addition to large missing tranverse energy, specifically $\not\!\!E_T > 25$ GeV. In this control sample, however, a real signal is present from the top quark, via $t\bar{t} \to WWb\bar{b}$. We show the fit to the 3rd-highest jet E_T distribution in figure 4. In this case, the fit predicts $439\ ^{+20}_{-19}$ (stat.) events (after propagating the fit parameter uncertainties); we observe 762 events. Using this excess (323 ± 34 events) to measure the $t\bar{t}$ cross section, we take an acceptance from Monte Carlo ($3.41 \pm 0.02\%$), with our luminosity (1.036 fb^{-1} for the muon triggers), and obtain a cross section of 9 ± 1 pb (statistical error only). The proximity to the theoretical cross section of ~ 7 pb indicates the method measures the W+jet background adequately.

We now apply this method to the $Z^0 \to \ell\ell$ data. The fit and the extrapolation is shown in figure 5 along with the data, excluding events in the blinded $N_{\rm jet}^{30} \geq 3$ region. Overall, we predict $72\ ^{+10}_{-9}$ (stat.) events with $N_{\rm jet}^{30} \geq 3$.

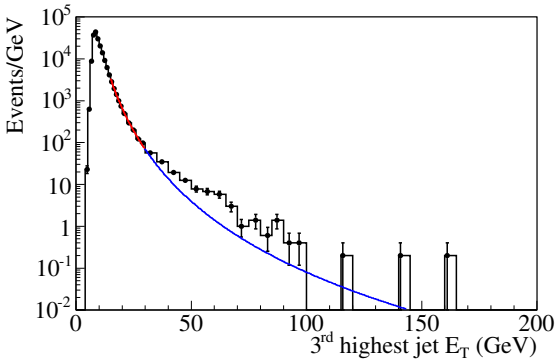


FIG. 4: E_T distribution of 3rd-highest E_T jet in W+jet data, fit to the parameterization in equation (1) in the region $15 < E_T < 30$ GeV. The excess above the extrapolated prediction is consistent with the expectation from $t\bar{t}$.

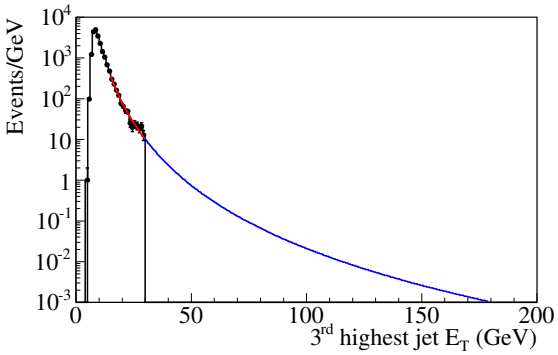


FIG. 5: E_T distribution of 3rd-highest E_T jet in $Z^0 \to \ell\ell$ data, fit to the parameterization in equation (1) in the region $15 < E_T < 30$ GeV.

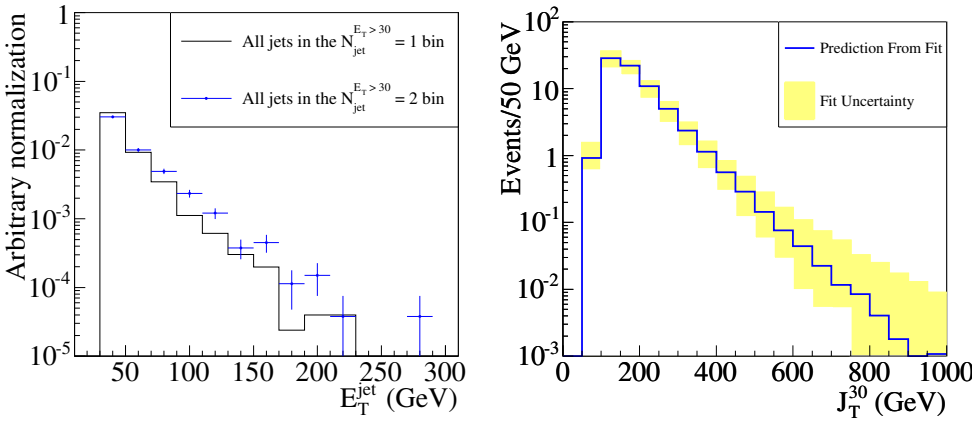
B. The J_T^{30} Shape of the Z^0+jets Background in the $N_{\rm jet}^{30} \geq 3$ Bins

Given we have a method for predicting the total number of background events in the $N_{\rm jet}^{30} \geq 3$ region, we now must predict the shape of the J_T^{30} distribution in this region. Since J_T^{30} is merely the sum of the jet E_T's (above 30 GeV), all that is needed is the shape of the jet E_T distributions in the $N_{\rm jet}^{30} \geq 3$ bins. To predict this distribution, we take the jet E_T distributions of jets in the $N_{\rm jet}^{30} = 1$ and 2 bins, and use them to predict the jet E_T shape in the $N_{\rm jet}^{30} \geq 3$ bins.

Figure 6a shows the jet E_T distributions of jets in the $N_{\rm jet}^{30} = 1$ and 2 bins, in $Z^0 \to \ell\ell$ data. The general behavior is broadly similar, indicating that it is nearly sufficient to take either one of these distributions as an estimate of the jet E_T shape in the $N_{\rm jet}^{30} \geq 3$ bins. However, the $N_{\rm jet}^{30} = 1$ bin has a slightly softer E_T spectrum than the $N_{\rm jet}^{30} = 2$ bin. In order to take into account this dependence, we fit the E_T distribution in the $N_{\rm jet}^{30} = 1$ bin to the parameterization in equation (1), fit the $N_{\rm jet}^{30} = 2$ bin to the same parameterization, and linearly extrapolate the exponential parameter (p_1) into the $N_{\rm jet}^{30} \geq 3$ bins. To avoid simultaneously extrapolating two correlated fit parameters, we fix the power-law parameter (p_2) to the fit in the $N_{\rm jet}^{30} = 1$ bin. Using this extrapolation, we have a prediction for the E_T distribution in the $N_{\rm jet}^{30} \geq 3$ bins. We obtain an estimate for the relative fractions of events in the $N_{\rm jet}^{30} = 3, 4, 5...$ bins using an exponential fit to the data in the $N_{\rm jet}^{30} \leq 2$ bins. Once these distributions are known, it is then a simple matter to obtain the J_T^{30} distribution by a random sampling of the $N_{\rm jet}^{30}$ shape and the extrapolated jet E_T shapes.

We show the J_T^{30} distribution prediction using this method on the $Z^0 \to \ell\ell$ data in figure 6b. The uncertainties are evaluated by varying each fit parameter independently and calculating the change in J_T^{30} distribution. To estimate the level of systematic uncertainty from a poor fit parameterization, we vary the ranges of the jet E_T distribution fits

and calculate the change in the J_T^{30} distribution. These individual uncertainties are then each added in quadrature to obtain the full uncertainty. The distribution is normalized to the total number of events with $N_{\rm jet}^{30} \geq 3$, predicted using the method in section III A above.



(a) (b)

FIG. 6: (a) Jet E_T distribution of jets in the $N_{\rm jet}^{30} = 1$ and 2 bins.
(b) J_T^{30} prediction in the $N_{\rm jet}^{30} \geq 3$ bins, using extrapolated jet E_T distributions in $Z^0 \rightarrow \ell\ell$ data. The distribution is normalized to the total predicted number of events with $N_{\rm jet}^{30} \geq 3$, found using the method described in section III A.

To verify that this procedure adequately predicts the J_T^{30} distribution, we again validate it using control samples in data. Using the jet triggers, we observe good agreement with the predicted and observated distribution. In the W+jet data, the observed J_T^{30} distribution agrees well with the background+$t\bar{t}$ hypothesis (where the background distribution is taken from the fit prediction, and the $t\bar{t}$ distribution is taken from Monte Carlo). We show the predicted and observed distributions in W+jet data in figure 7.

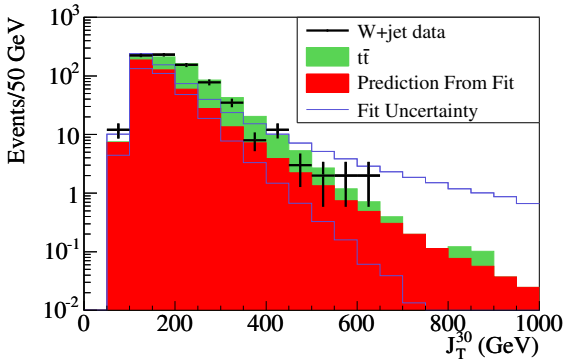


FIG. 7: J_T^{30} background prediction in W+jet data, added with the shape from $t\bar{t}$ Monte Carlo in the $N_{\rm jet}^{30} \geq 3$ bins, compared to the observation.

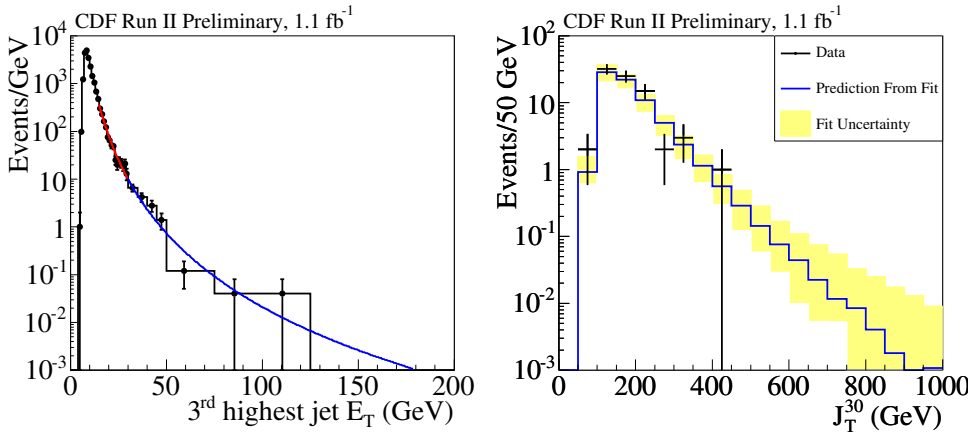
IV. RESULTS

After validating the procedure described above in control regions in data, we then compare the prediction to the observation in the Z^0+jets data, in the signal region $N_{\rm jet}^{30} \geq 3$. We show the 3rd-highest E_T jet extrapolation, and

the comparison to the data, in figure 8a. We show the J_T^{30} distribution prediction and the data in figure 8b. The data agree with the prediction quite well. We list the predicted and observed number of events integrated above various J_T^{30} thresholds in table I.

Minimum J_T^{30}	Total Bkg. (stat.+syst. errors)	Data
50	$72.1\ ^{+17.7}_{-22.6}$	80
100	$71.2\ ^{+17}_{-22.3}$	78
150	$42.7\ ^{+9.48}_{-14}$	46
200	$20.5\ ^{+5.64}_{-7.77}$	21
250	$9.67\ ^{+3.48}_{-4.04}$	6
300	$4.67\ ^{+2.17}_{-2.13}$	4
350	$2.31\ ^{+1.4}_{-1.16}$	1
400	$1.17\ ^{+0.925}_{-0.642}$	1
450	$0.605\ ^{+0.655}_{-0.378}$	0

TABLE I: The data compared to the Z^0+jets background fit prediction vs. J_T^{30}.



(a) (b)

FIG. 8: (a) 3rd-highest E_T jet extrapolation in $Z^0 \to \ell\ell$ data.
(b) J_T^{30} distribution prediction compared to $Z^0 \to \ell\ell$ data in the $N_{\rm jet}^{30} \geq 3$ bins.

Given there is no significant excess present in the data, we set a cross section limit using the b' model. At each b' mass, the counting experiment is evaluated with the requirement $J_T^{30} > m_{b'}$. The limit is set at a 95% confidence level by integrating a likelihood obtained using a Bayesian technique that smears the Poission-distributed background with Gaussian acceptance and mean background uncertainties [11]. The background and its uncertainty is taken from the fit prediction (listed in table I); the acceptance×efficiency is taken from Monte Carlo, with correction factors applied to match the observed efficiency of leptons in $Z^0 \to \ell\ell$ data. The uncertainty on the acceptance×efficiency is 10%, with the dominant source from a jet energy scale uncertainty of 6.7%.

The cross section limit as a function of mass is shown in figure 9; we do not set a cross section limit below masses of 150 GeV, as studies in Monte Carlo indicate that the large cross sections present below this mass cause signal contamination. The b' cross section is calculated at leading order using `PYTHIA`, with the assumption that $BR(b' \to bZ^0) = 100\%$. With this assumption, the mass limit observed is $m_{b'} > 270$ GeV.

Acknowledgments

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the

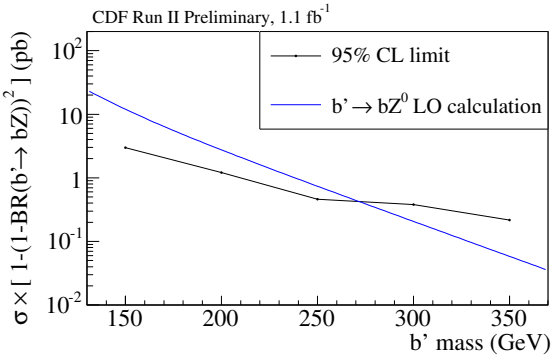


FIG. 9: Cross section limit vs. b' mass.

Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium für Bildung und Forschung, Germany; the Korean Science and Engineering Foundation and the Korean Research Foundation; the Particle Physics and Astronomy Research Council and the Royal Society, UK; the Institut National de Physique Nucleaire et Physique des Particules/CNRS; the Russian Foundation for Basic Research; the Comisión Interministerial de Ciencia y Tecnología, Spain; the European Community's Human Potential Programme under contract HPRN-CT-2002-00292; and the Academy of Finland.

[1] F. Abe, *et al.*, *Nucl. Instrum. Methods Phys. Res. A* **271**, 387 (1988); D. Amidei, *et al.*, *Nucl. Instum. Methods Phys. Res. A* **350**, 73 (1994); F. Abe, *et al.*, *Phys. Rev. D* **52**, 4784 (1995); P. Azzi, *et al.*, *Nucl. Instrum. Methods Phys. Res. A* **360**, 137 (1995); The CDFII Detector Technical Design Report, Fermilab-Pub-96/390-E
[2] D. Acosta, *et al.*, *Phys. Rev. Lett.* **94** 091803 (2005)
[3] T. Affolder, *et al.*, *Phys. Rev. Lett.* **84**, 835 (2000)
[4] V. M. Abazov, *et al.*, *Phys. Rev. D* **74**, 011104 (2006)
[5] A.L. Scott, for the CDF collaboration, *Int. J. Mod. Phys. A*, **20**, 3263 (2005)
[6] Public CDF note 8452
[7] Public CDF note 8419
[8] Public CDF note 8422
[9] P.H. Frampton, P.Q. Hung, M. Sher, *Phys.Rept.* **330**, 263 (2000)
[10] T. Sjostrand, *et al.*, *Comput. Phys. Commun.* **135**, 238 (2001)
[11] Section 4 of J. Conway, Proceedings of Workshop on Confidence Limits, Geneva, Switzerland, 17-18 Jan 2000, FERMILAB-CONF-00-048-E