CDF Note 9506

Search for Gluino-Mediated Sbottom Production in the MET+b-jet Sample

The CDF Collaboration
URL http://www-cdf.fnal.gov
(Dated: September 12, 2008)

Using data collected with the CDF detector in the Run II of the Tevatron, we present the results of a search for the supersymmetric partner of the bottom quark produced from gluino decays. We analyzed events containing two or more jets and missing transverse energy searching for the presence of physics beyond the Standard Model. Two categories were made by requiring only one of the jets or at least two jets to be tagged as originating from a heavy-flavour quark in order to enhance the sensitivity to final states containing b-quarks. Results were obtained with 2.5 fb^{-1} of data and constrain the cross section production to be less than 40 fb at 95 % C.L.

Preliminary Results for the Summer Conferences 2008

INTRODUCTION

Supersymmetry (SUSY) is one of the primary theories for physics beyond the Standard Model (SM). SUSY proposes a symmetry between fermions and bosons and predicts a superpartner for all standard model particles. The superpartner for a standard model particle differs from it only by a half-unit of spin. With the addiction of this symmetry, SUSY offers solutions to the fine-tuning problem and a possible mechanism for electroweak symmetry breaking (EWSB). It also makes possible for a unification of the gauge couplings at about the Planck scale (10^{19} GeV).

In the Minimal Supersymetric extension of the Standard Model (MSSM), to each fermionic SM particle a bosonic superpartner is assigned and vice versa. Therefore, the SM quark helicity states q_L and q_R have scalar MSSM superpartners, which are also the mass eigenstates (in good approximation) for the first two generations. However, for the third generation, strong mixing can appear depending on the theoretical parameters: tanβ (the ratio of the VEV's of the two Higgs fields) and A_b (the Higgs-sbottom trilinear coupling).

If tanβ is large, then there can be a large mass splitting in the scalar bottom (sbottom) sector, yielding a mass to the lightest sbottom state ($\tilde{b}$) in the reach of the Tevatron center of mass energy. Assuming R-parity conservation, SUSY particles are pair produced and the Lightest Supersymetric Particle (LSP) must be stable. If in addition it is colorless and neutral, then it will escape from the detector undetected yielding large missing transverse energy ($\not\!\!E_T$).

At Tevatron, $\tilde{b}$ production through gluino ($\tilde{g}$) decays has a very distinctive signature. Under the assumption that m($\tilde{g}$) is smaller than m($\tilde{q}$), but larger than m($\tilde{b}$), the gluino pair production, $p\bar{p} \to \tilde{g}\tilde{g}$ is one of the dominant SUSY processes. After production the gluino decays to $\tilde{g} \to b\tilde{b}$ with the subsequent sbottom decay to a b-quark and the lightest neutralino ($\tilde{\chi}^0$), $\tilde{b} \to b\tilde{\chi}^0$. In this approach and assuming R-parity conservation, the neutralino is taken to be the LSP.

The gluino pair production cross section is large ($\sigma(g\tilde{g}) \sim 10 \times \sigma(b\tilde{b})$) compared to direct sbottom pair production of similar mass. At the Tevatron energies, gluinos are produced mainly through quark-antiquark annihilation and gluon fusion (Figure 1), the former dominating for high x. If the sbottom is light enough, then the two body decay $\tilde{g} \to b\bar{b}$ would be kinematically allowed.

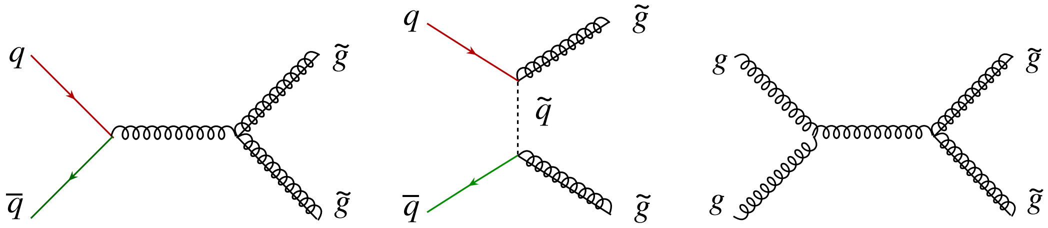

FIG. 1: Leading Order gluino pair production mechanisms at the Tevatron center of mass energies.

In our region of interest ($m_t, m_{\tilde{\chi}^+} > m_{\tilde{b}} > m_{\tilde{\chi}^0}$) the dominant decay channel is sbottom into bottom quark and neutralino $\tilde{b} \to b\tilde{\chi}^0$, with no other available decays channels since we require $m_{\tilde{b}} < m_t, m^{+}_{\tilde{\chi}}$. Hence, we assume a Branching Ratio of 100% for the $\tilde{b} \to b\tilde{\chi}^0$ decay. In summary, for this particular signature we expect 4 b-jets and large $\not\!\!E_T$ due to the stable neutralinos.

EVENT SELECTION

The present analysis is based on an integrated luminosity of 2.5 fb^{-1} collected with the CDF II detector, using a three-level logic inclusive $\not\!\!E_T$ trigger. A sequence of cuts on the $\not\!\!E_T$ is required at each level. At Level 1 it requires $\not\!\!E_T$ above 25 GeV, at Level 2 it requires $\not\!\!E_T$ above 35 GeV and at Level 3 it requires $\not\!\!E_T$ above 45 GeV.

Events computed in the present analisis are required to have a reconstructed vertex with Z-position within 60 cm of the nominal interaction point, $\not\!\!E_T \geq 70$ GeV and tracking activity consistent with the energy measured in the calorimeter to reject cosmics and beam-halo background. Two or more jets are required to accept the event. Jets are defined using a cone-based algorithm [1] with radius 0.4 and required to have a transverse energy (E_T) above 25 GeV and a pseudorapidity [2] $|\eta| \leq 2.4$. At least one of the jets is required to be central ($|\eta| \leq 0.9$) and the jet with the highest transverse energy must satisfy $E_T \geq 35$ GeV.

Since it is expected a 4 b-jets final state, two categories were made by requiring only one of the jets (one exclusive tag) or at least two jets to be tagged (two inclusive tags) as originating from a heavy-flavour quark. In order to identify jets originating from a b-quark, a Secondary Vertex tagging algorithm [3] is used. The double-tag category provides the most sensitivity in the analysis and it is used to extract the limits. The single-tag category is used as an additional control region.

SIGNAL AND BACKGROUND MODELING

In order to get the predictions for our signal, we used the program PROSPINO [4, 5] to compute the total production cross section $p\bar{p} \rightarrow \tilde{g}\tilde{g}$ and PYTHIA [6] to estimate the event acceptance in the detector and in the application of our selection cuts.

Several signal Monte Carlo samples were generated with PYTHIA and passed through the detector simulation in order to cover the phase space under study as a function of the sbottom mass and gluino mass. These samples were generated using the Tune AW [7] and setting explicitly the SUSY parameters of the model, which only affects the masses of the involved particles since the production process is via the strong interaction and all the decay branching ratio are set to 100%. The gluino mass was varied between 240 GeV/c^2 to 400 GeV/c^2 and the sbottom masses from 150 GeV/c^2 to 350 GeV/c^2. The neutralino mass was fixed to 60 GeV/c^2, while the squark mass was fixed to 500 GeV/c^2.

Part of the SM background in this analysis is predicted with Monte Carlo simulation, being the $t\bar{t}$ production the dominant one. In addition, there are contributions from Z and W production in association with jets, single top and diboson production. All this background samples were generated with PYTHIA using the same tuning parameters set described above, and processed in a similar way as the signal events.

Background contribution from QCD multijet production and light flavour jets falsely tagged as a b-quark ("mis-tags"), are estimated from data. The QCD multijet production has a very large cross section in comparison with the expected signal, however these processes usually do not produce $\not\!\!E_T$ in the final state. Events from QCD multijet production pass our selection if a mismeasurement of the jets or a semi-leptonic decay from a meson produces $\not\!\!E_T$. In both cases the $\not\!\!E_T$ tends to be aligned with the first or second most energetic jet.

The "mis-tags" are estimated from the jets in data, the "mis-tags" rate of which is parametrized by their E_T, $|\eta|$ and Secondary Vertex track multiplicity. Events are also parametrized in terms of a given number of vertexes, primary vertex z-position and the scalar summed E_T of all jets in the event.

To estimate the multijet background from data we have developed the MUlti TAg Rate Estimator (MUTARE) method. The method is based on a tag-rate matrix built over each jet in a event parametrized by their E_T, $|\eta|$ and the scalar summed E_T of all jets in the event. The final prediction is obtained after subtracting the "mis-tags" and non-QCD SM heavy-flavour contributions.

CONTROL REGIONS

The SM processes predicted with Monte Carlo or data-driven methods were tested in control regions defined such that they result in background-dominated samples in which signal contribution is negligible. Two regions were defined by reversing the selection requirements introduced to suppress specific background processes and a third one was defined in order to check the analysis tool in a signal-like environment, but avoiding the application of cuts that would enhance the signal contribution to a measurable level.

- Multijet QCD control region: second jet required to be aligned with the $\not\!\!E_T$ azimuthal direction within an angle of 0.4 radians
- Lepton control region: lepton required and second jet required to be not aligned with the $\not\!\!E_T$ azimuthal direction, with and angular distance greater than 0.7 radians
- Pre-Optimization control region: first and second jets required to be not aligned with the $\not\!\!E_T$ azimuthal direction with and angular distance greater than 0.7 radians and $E_{T,j} \geq 50$ GeV

The first jet transverse energy and the $\not\!\!E_T$ distributions for the three control regions, Multijet QCD control region (top), Lepton control region (middle) and Pre-Optimization (bottom), are shown in Figure 2 for the one exclusive tag analysis and in Figure 3 for the two inclusive tags analysis. Good agreement between data and SM predictions was obtained in all control regions (Table I).

CDF Run II Preliminary 2.5 fb^{-1}

One Exclusive Tag	QCD Region	Lepton Region	Preoptimization Region
W/Z + jets production	84 ± 36	132 ± 57	376 ± 162
Diboson production	4 ± 1	20 ± 5	41 ± 10
Top pair production	60 ± 15	371 ± 93	478 ± 119
Single top production	5 ± 1	34 ± 8	45 ± 11
HF QCD Multijets	9741 ± 4870	195 ± 97	1660 ± 830
Light-flavour contamination	5430 ± 2226	190 ± 78	919 ± 377
Total expected	15325 ± 5355	943 ± 166	3520 ± 934
Observed	15390	890	3525

CDF Run II Preliminary 2.5 fb^{-1}

Two Inclusive Tags	QCD Region	Lepton Region	Preoptimization Region
W/Z + jets production	10 ± 7	19 ± 14	29 ± 22
Diboson production	0.4 ± 0.1	2 ± 0.6	4 ± 1
Top pair production	18 ± 6	107 ± 34	140 ± 45
Single top production	1 ± 0.2	4 ± 1	6 ± 1
HF QCD Multijets	864 ± 432	23 ± 11	273 ± 136
Light-flavour contamination	238 ± 48	8 ± 2	57 ± 11
Total expected	1132 ± 435	164 ± 38	510 ± 145
Observed	1104	156	455

TABLE I: Table of predicted and observed events in the control regions with one exclusive tag (left) and two inclusive tags (right).

SIGNAL OPTIMIZATION

An optimization process via two Neural Networks (NN) was made in order to reduce the background contribution. We chose two SUSY signal points to perform a MC optimization cut study for each point. The two selected points were chosen to have large and small Δm between the gluino mass and the sbottom mass due to the strong dependence of the kinematic variables on this quantity. The large Δm optimization was made in the SUSY point M($\tilde{g}$) = 335 GeV/c^2, M($\tilde{b}$) = 260 GeV/c^2 and small Δm optimization was made in the SUSY point M($\tilde{g}$) = 335 GeV/c^2, M($\tilde{b}$) = 315 GeV/c^2

The optimization process takes as a starting benchmark the Pre-optimization selection. In addiction, for the large Δm optimization a cut on the number of jets greater than two was applied. For the small Δm optimization this cut was not applied because of the small amount of momentum available in the gluino decay which implies a softer jet multiplicity in the final estate. Over this selection, a first Neural Network was applied to distinguish between gluino signal and QCD Multijets background. This Neural Network was trained with signal MC versus taggable jets (QCD-like) in the Pre-optimization region with one exclusive tag in order to have enough statistics. A second Neural Network was applied to remove the remaining backgrounds, mainly top-pair production, and was trained with signal MC versus top pair MC also over the Pre-optimization region with one exclusive tag after applying the cut on the previous NN. The previous process was made for the two different SUSY points as well as for one exclusive and two inclusive tag selection.

After choosing the set of variables used as input for the Neural Network a training and test evaluations were made with the TMVA package [8] in both Neural Networks. The variables used in all NN are $\not\!\!E_T$, E_T of the jets, $\Delta\phi(\not\!\!E_T, jets)$ (this two variables are applied to three or two jets depending on the optimization) and the vectorial summed E_T of all jets in the event.

By applying to the Pre-optimization region the QCD Multijets Neural Networks we obtained the outputs showed in Figure 4 for the large Δm optimization and for the small Δm optimization (one exclusive tag and two inclusive tags).

For all the cases showed in Figure 4 we found an optimal selection cut on 0.8 in the NN output. This cut optimizes the sensitivity keeping a reasonable amount of signal.

At this stage of the optimization we apply the second NN, based on top pair discrimination, to the events obtained after the cut on 0.8 on the output of the QCD Multijets Neural Network. The result of applying the top pair NN to this events is shown in Figure 5 for the large Δm optimization as well as for the small Δm optimization (one exclusive tag and two inclusive tags).

We found 0.6 an optimal selection cut in the large Δm optimization NN output and 0.8 selection cut in the small Δm optimization NN output.

By applying the optimizations we obtain four final regions depending on the tagging and the signal point used in the optimization.

SYSTEMATICS UNCERTAINTIES

We address systematic uncertainties from different sources:

- Jet Energy Scale [9]: A systematic error in the calorimeter energy scale affect the total transverse energy on the jets. The effect in the final regions varies in a range between 5% and 25% depending on the optimization.

- b-tagging Scale Factor: The difference between data and MC in b-tagging efficiency (5%) is taken as systematic uncertainty. The resulting uncertainty in the final regions varies between 1.5% and 5% depending on the optimization.
- "Mis-tag" estimation: The systematic error assigned to the tag rate matrix is 4.8%.
- Luminosity: The systematic uncertainty in the luminosity is taken to be 6%, affecting to the normalization of all the MC estimated backgrounds.
- ISR/FSR: The uncertainty associated with the initial and final state radiation was evaluated by generating sample with more/less ISR/FSR. The effect in the final regions varies in a range between 2% and 5% depending on the optimization.
- PDF: The PDF uncertainty has been determined to be 2% on the acceptance.
- QCD Multijet Background: We assign a conservative 50% uncertainty in the prediction based on the variation observed when matrix definition is changed.
- Top-Pair Production cross section: We quote the uncertainty in the CDF measured value (11%) of the top-pair production cross section.
- Single Top Production cross section: We quote the theoretical uncertainty in the single-top cross section (13%) [10–13].
- Diboson Production cross section: We quote the theoretical uncertainty being 10% in the WW and WZ cross sections and 20% for the ZZ process.
- Single EWK Boson Production cross section: Although the cross section for Z and W production are known to a high precision, we are using the inclusive processes in PYTHIA to perform estimations of Z/W+multijet contributions since PYTHIA parton showering does not properly reproduce the multijet spectrum, we estimate a 40% uncertainty in the predictions.
- Top quark mass: In the current analysis, the $t\bar{t}$ production background is estimated using MC with a top quark mass of 171.5 GeV/c^2. Since our signal optimization is based on a Neural Network trained with $t\bar{t}$ processes we include a systematic error due to the top pair NN output dependence on the top quark mass. We compute this error measuring the number of top-pair events in the final selection by using a top quark mass of 174.5 GeV/c^2. The effect in the final regions varies in a range between 0.3% and 17% depending on the optimization.

RESULTS

Four different signal regions were optimized for the present analysis. In this four regions the number of observed events are in good agreement with the expected events from the SM processes as shown in Table II and Table III.

CDF Run II Preliminary 2.5 fb^{-1} (One Exclusive Tag)

Region	Large Δm Optimization	Small Δm Optimization
W/Z + jets production	1.8 ± 0.8	7.1 ± 3.1
Diboson production	0.6 ± 0.1	1.3 ± 0.3
Top pair production	7.1 ± 2.7	4.4 ± 1.1
Single top production	0.08 ± 0.03	0.4 ± 0.1
HF QCD Multijets	0.9 ± 0.5	9.2 ± 4.6
Light-flavour contamination	5.6 ± 1.2	19.6 ± 8.1
Total expected	16.2 ± 3.7	42.1 ± 10.1
Observed	18	29
Signal M($\tilde{g}$)=335, M($\tilde{b}$)=260	13.7 ± 4.5	-
Signal M($\tilde{g}$)=335, M($\tilde{b}$)=315	-	26.3 ± 8.2

TABLE II: Table of predicted and observed events in the large Δm and small Δm optimized regions with one exclusive tag.

CDF Run II Preliminary 2.5 fb^{-1} (Two Inclusive Tags)

Region	Large Δm Optimization	Small Δm Optimization
W/Z + jets production	0.1 ± 0.05	0.4 ± 0.3
Diboson production	0.07 ± 0.02	0.1 ± 0.03
Top pair production	1.9 ± 1.0	0.6 ± 0.4
Single top production	0.03 ± 0.01	0.04 ± 0.01
HF QCD Multijets	1.5 ± 0.7	0.6 ± 0.3
Light-flavour contamination	0.9 ± 0.3	0.6 ± 0.1
Total expected	4.5 ± 1.4	2.3 ± 0.8
Observed	5	2
Signal M($\tilde{g}$)=335, M($\tilde{b}$)=260	14.9 ± 5.0	-
Signal M($\tilde{g}$)=335, M($\tilde{b}$)=315	-	8.5 ± 2.8

TABLE III: Table of predicted and observed events in the large Δm and small Δm optimized regions with two inclusive tags.

Since no significant deviation from the SM prediction was observed, the results were used to extract an exclusion limit for the cross section of the described process. Since the two-tag sample is much more sensitive to the signal, only those results were used.

Using a Bayesian approach, we found that the observed results sets a 95% C.L. limit in the production cross section such that no more than 8.1 (5.4) events are observed in the large (small) Δm signal Region. This translates in the limit shown in Fig. 6, as a function of the gluino-pair production cross section for given values of the sbottom mass.

For the assumed model, the limit is nearly independent of the sbottom mass and the cross section limit is around 40 fb. In addition, using the assumed model, a 95% C.L. limit was obtained in the mass parameter plane of the model, for a fixed mass of the neutralino. Fig. 7 shows the excluded region in the gluino-sbottom mass plane of the analysis, compared with results from previous analysis [14–17].

Currently the limit obtained with the present analysis clearly improves the results of previous searches using similar topology and also, under the assumptions discussed above, sets more stringent limits to the sbottom and gluino production than dedicated and more inclusive sbottom searches.

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 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 Science and Technology Facilities Council and the Royal Society, UK; the Institut National de Physique Nucleaire et Physique des Particules/CNRS; the Russian Foundation for Basic Research; the Ministerio de Ciencia e Innovación, Spain; the Slovak R&D Agency; and the Academy of Finland.

[1] CDF Collaboration, F. Abe *et al.*, Phys. Rev. **D45**, 1448 (1992).

[2] We use a cylindrical coordinate system with the z coordinate along the proton beam direction and the origin at the center of the detector, the azimuthal angle ϕ, and the polar angle θ usually expressed through the pseudorapidity $\eta = -lntan(\theta/2)$. The rapidity y is defined as $y = 1/2ln((E+p_z)/(E-p_z))$ where E denotes the energy and p_z is the momentum component along z. The transverse energy and transverse momentum are given by $E_T = Esin(\theta)$ and $p_T = psin(\theta)$ where p is the magnitude of the momentum vector.

[3] CDF Collaboration, D. Acosta *et al.*, Phys. Rev. **D71**, 052003 (2005).

[4] W. Beenakker, M. Hopker, M. Spira, and P. Zerwas, Nucl. Phys **B**, 51 (1997).

[5] W. Beenakker, M. Hopker, and M. Spira, Prospino: A program for the production of supersymmetric particles in next-to-leading order qcd., hep-ph/9611232, 1996.

[6] T. Sjöstrand *et al.*, Comput. Phys. Commun. , 238 (2001).

[7] R.D. Field (CDF Collaboration), contribution to the APS/DPF/DPB Summer Study on the Future of Particle Physics (Snowmass 2001), Snowmass, Colorado, 30 June 21 July 2001 [hep-ph/0201192]; CDF Note 6403; R.D. Field (CDF Collaboration), presentations at the the 'Matrix Element and Monte Carlo Tuning Workshop', Fermilab, 4 October 2002 and 2930 April 2003, talks available from webpage http://cepa.fnal.gov/psm/MCTuning/; R. Field and R. C. Group (CDF Collaboration), hep-ph/0510198; R.D. Field (CDF Collaboration), further recent talks available from webpage http://www.phys.ufl.edu/rfield/cdf/.
[8] A. Hocker *et al.*, arXiv:physics/0703039 .
[9] CDF Collaboration, A.Bhatti *et al.*, Nucl. Instrum. Methods **A566**, 2 (2006).
[10] B. Harris *et al.*, Phys. Rev. **D66**, 054024 (2004).
[11] Z. Sullivan *et al.*, Phys. Rev. **D70**, 114012 (2004).
[12] J. Campbell, K. Ellis, and F. Tramontano, Phys. Rev. **D70**, 094012 (2004).
[13] N. Kidonakis, Phys. Rev. **D74**, 114012 (2006).
[14] CDF Collaboration, A. Abulencia *et al.*, Phys. Rev. Lett. **96**, 171802.
[15] DØ Collaboration, V. Abazov *et al.*, Phys. Rev. Lett **97**, 171806 (2006).
[16] CDF Collaboration, T. Affolder *et al.*, Phys. Rev. Lett. **84**, 5273 (2000).
[17] CDF Collaboration, T. Affolder *et al.*, Phys. Rev. Lett. **88**, 041801 (2002).

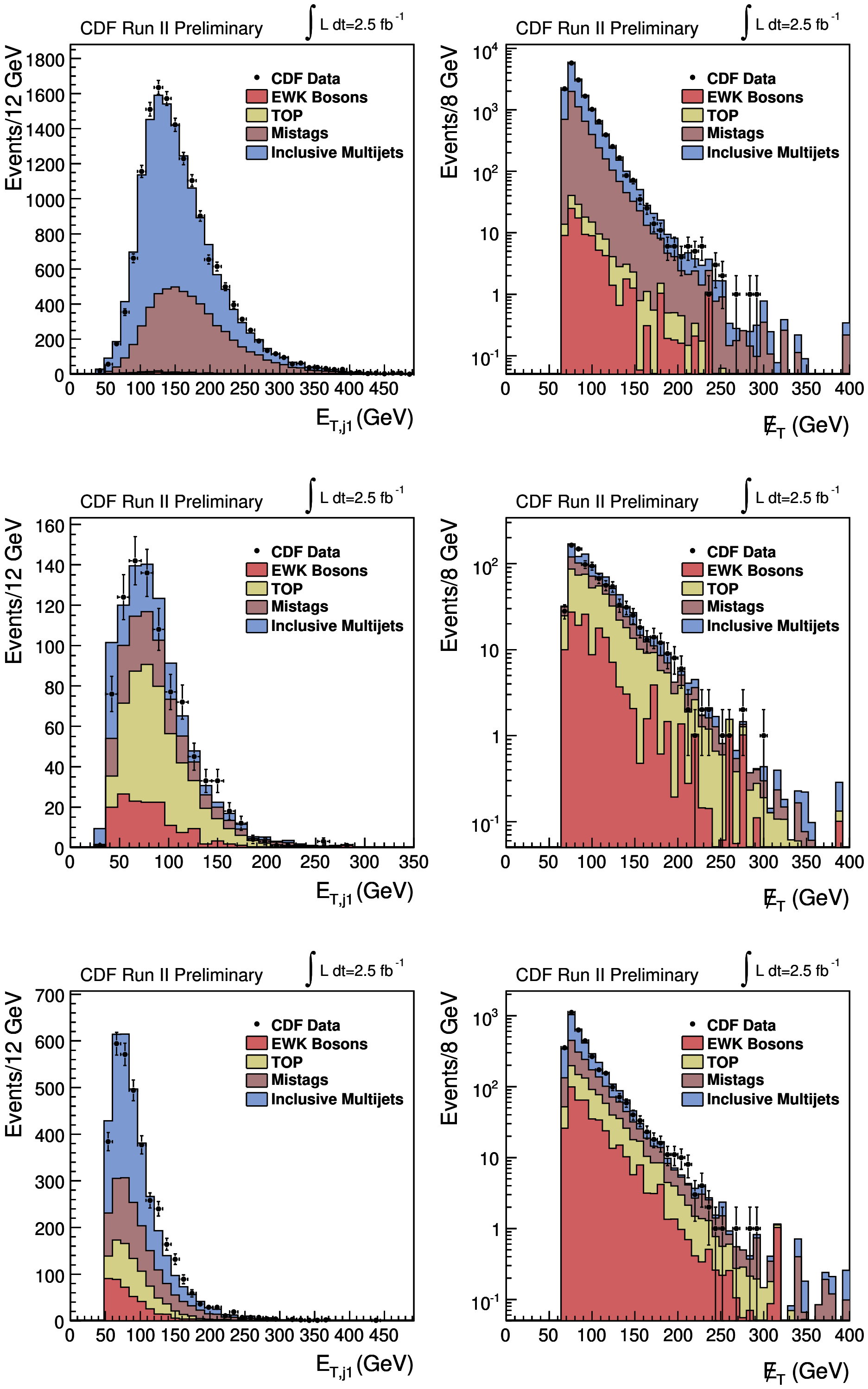

FIG. 2: Leading jet E_T, and $\not\!\!E_T$ in the QCD (top), Lepton (middle) and Pre-Optimization (bottom) control regions for the one exclusive tag selection.

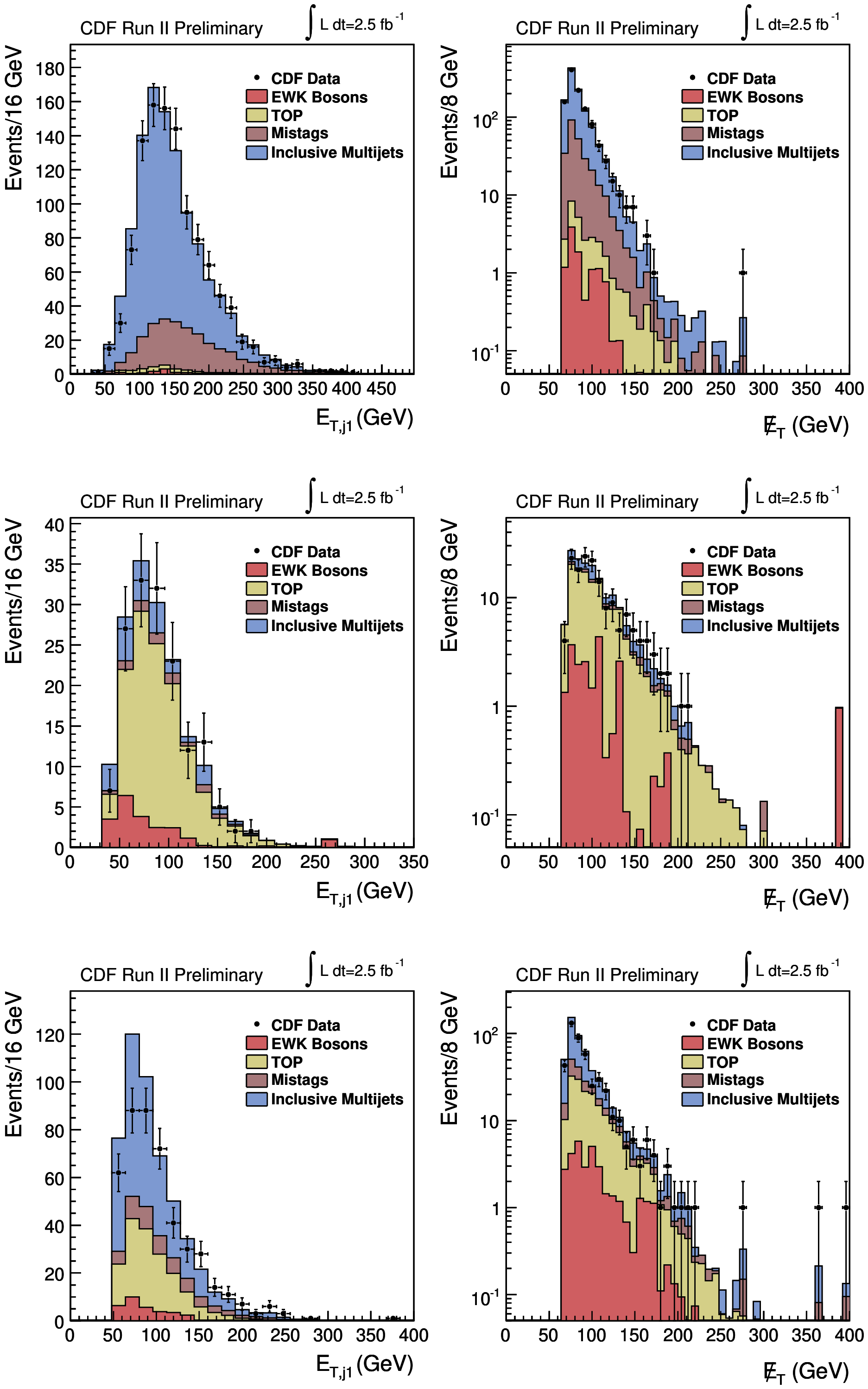

FIG. 3: Leading jet E_T and $\not{E}_T$ in the QCD (top), Lepton (middle) and Pre-Optimization (bottom) control regions for the two inclusive tags selection.

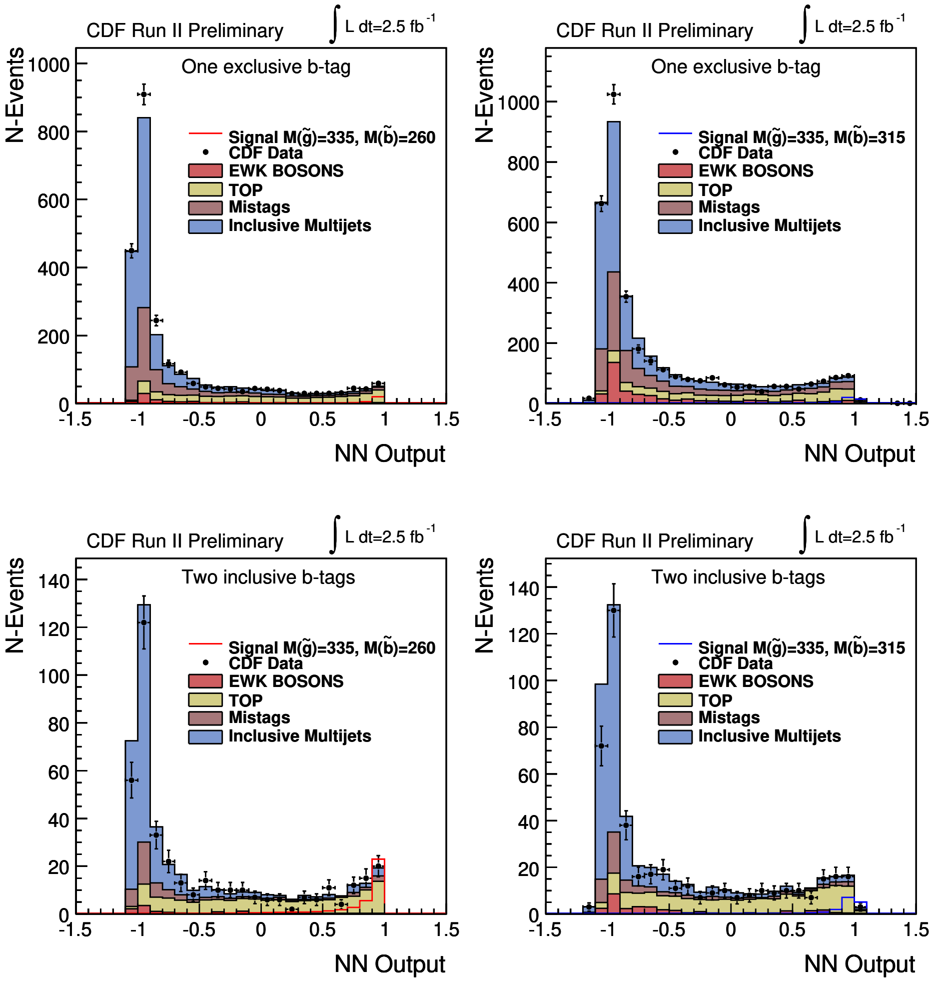

FIG. 4: QCD NN output plots for the large Δm optimized point (left) and small Δm optimized point (right) with one exclusive tag (top) and two inclusive tags (bottom)

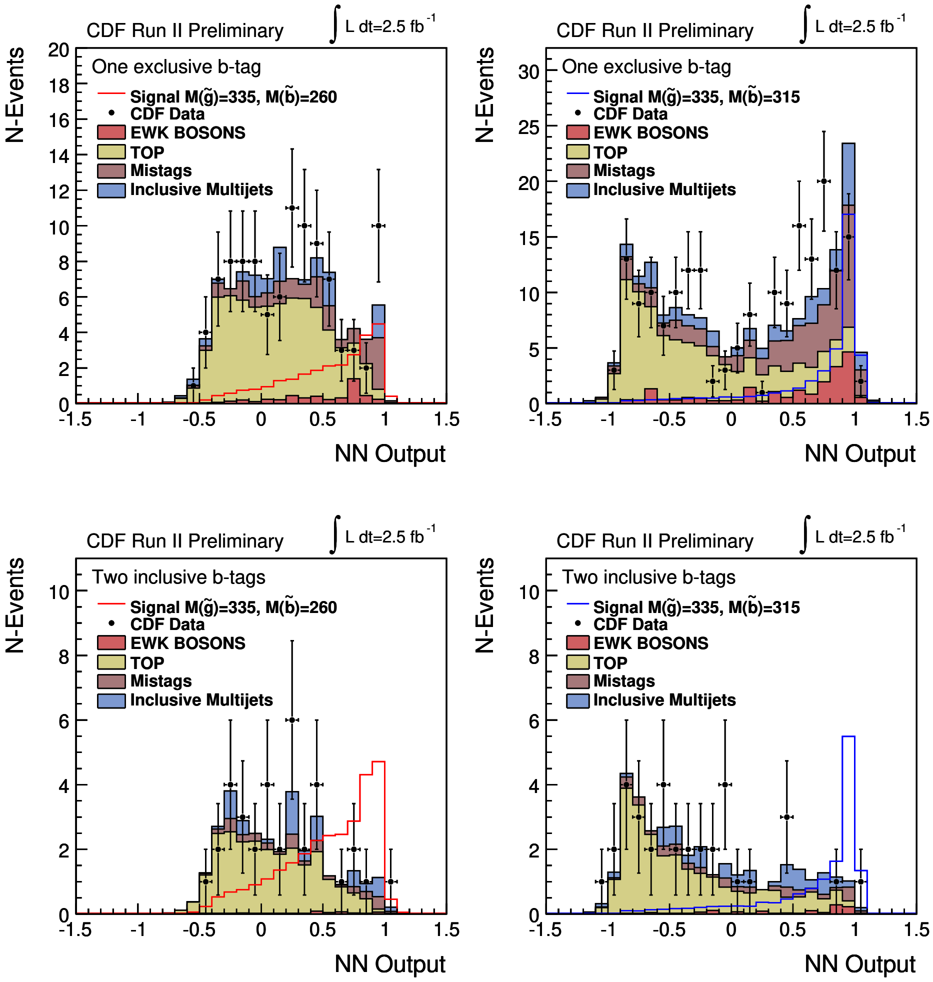

FIG. 5: Top NN output plots for the large Δm optimized point (left) and small Δm optimized point (right) with one exclusive tag (top) and two inclusive tags (bottom)

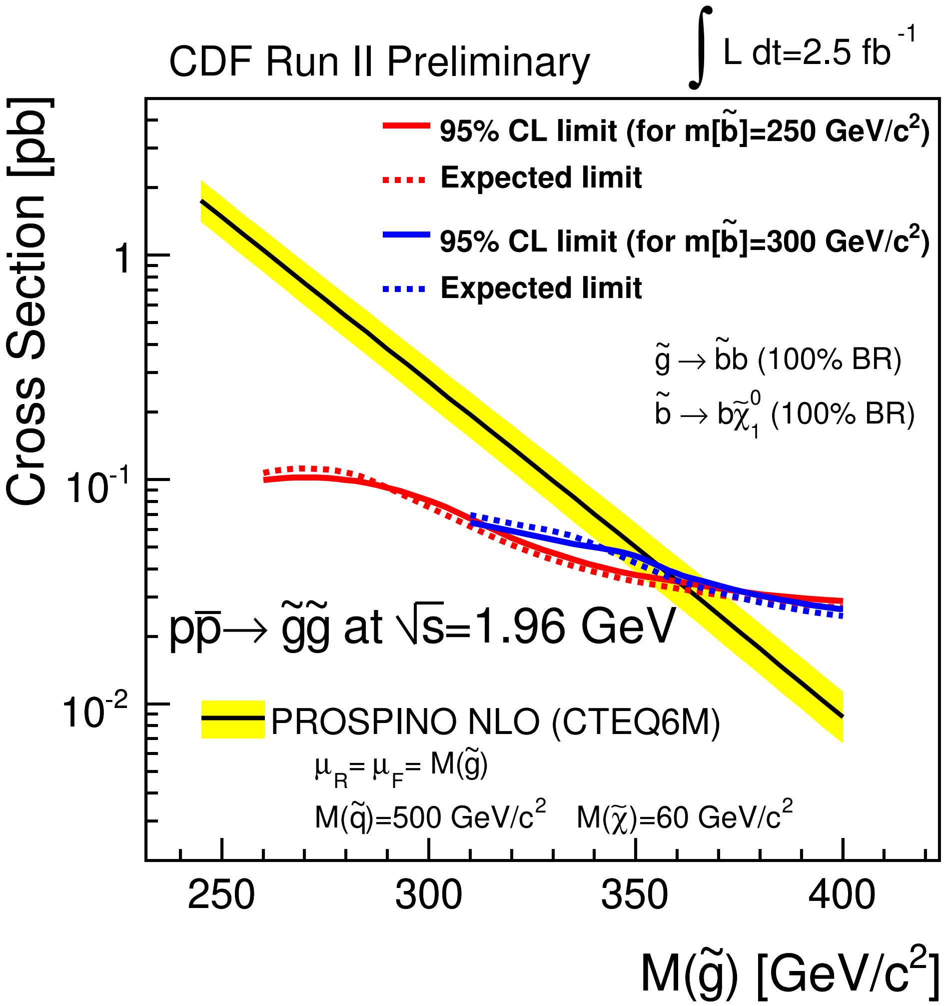

FIG. 6: Observed and expected cross section limit in terms of cross section as a function of the gluino mass for two different sbottom masses, $M(\tilde{b}) = 250$ GeV/c^2 (red) and $M(\tilde{b}) = 300$ GeV/c^2 (blue).

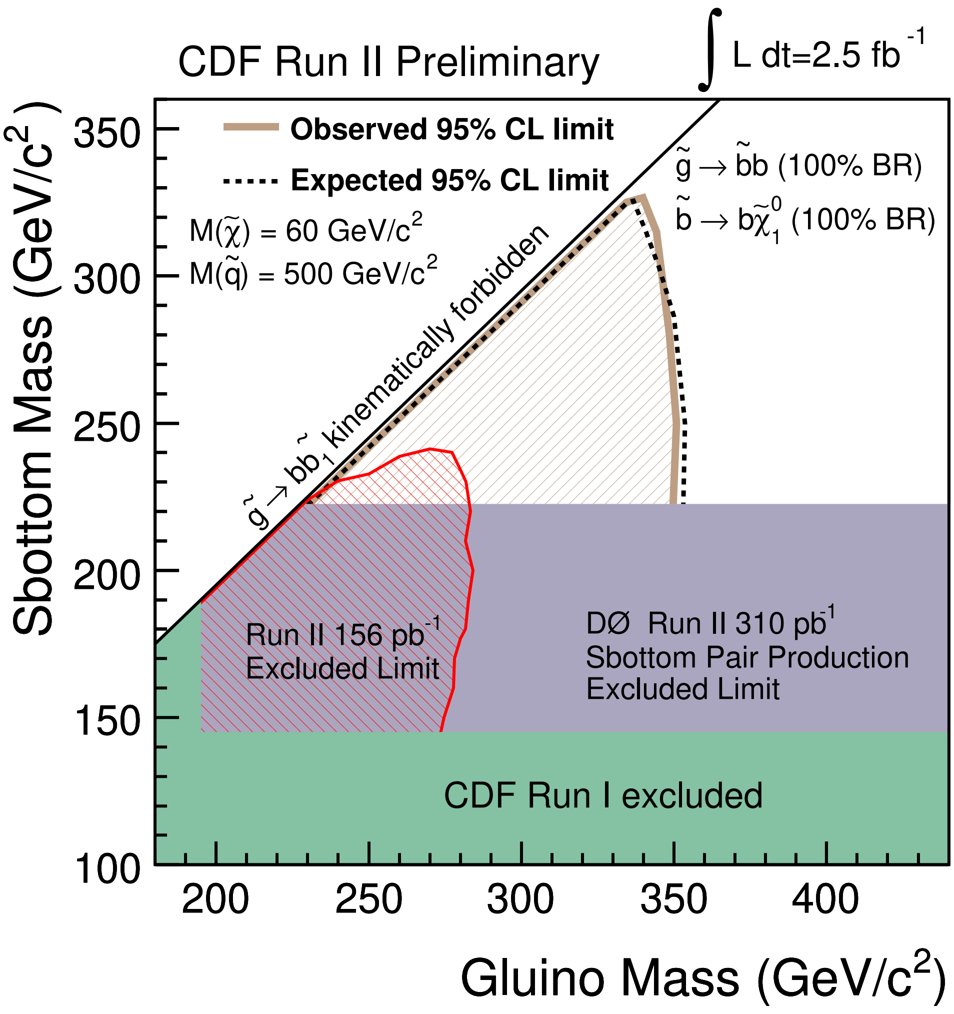

FIG. 7: M($\tilde{b}$)-M($\tilde{g}$) exclusion plot showing observed (brown) and expected (black dashed) limits at 95% C.L. assuming R-parity conservation. The present analysis made with 156 pb^{-1} by CDF [14] Collaboration (red), the excluded region by the sbottom pair production analysis made by DØ [15] Collaboration and the region excluded by CDF Run I related analysis [16, 17] are shown.